La fonction de densité au carrefour entre probabilités et analyse en terminale S. Etude de la conception et de la mise en œuvre de tâches d'introduction articulant lois à densité et calcul intégral.

Charlotte Derouet

▶ **To cite this version:**

Charlotte Derouet. La fonction de densité au carrefour entre probabilités et analyse en terminale S. Etude de la conception et de la mise en œuvre de tâches d'introduction articulant lois à densité et calcul intégral. . Histoire et perspectives sur les mathématiques [math.HO]. Université Paris Diderot (Paris 7) Sorbonne Paris Cité, 2016. Français. NNT : . tel-01431913

HAL Id: tel-01431913
https://theses.hal.science/tel-01431913

Submitted on 11 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Sorbonne Paris Cité
Université Paris Diderot (Paris 7)

École Doctorale : Savoirs scientifiques
Épistémologie, histoire des sciences et didactique des disciplines
Laboratoire de Didactique André Revuz

Thèse de doctorat

Didactique des disciplines – Mathématiques

présentée par

Charlotte Derouet

La fonction de densité au carrefour entre probabilités et analyse en terminale S

Étude de la conception et de la mise en œuvre de tâches d'introduction articulant lois à densité et calcul intégral

Sous la direction de Alain Kuzniak et Fabrice Vandebrouck

Soutenue publiquement à Paris le 25 novembre 2016 devant le jury composé de :

M. Jean-Claude Régnier	Université Lumière Lyon 2	Président, rapporteur
M^me^ Maria Alessandra Mariotti	Université de Sienne	Rapporteure
M. Giambattista Giacomin	Université Paris Diderot	Examinateur
M^me^ Corinne Hahn	ESCP Europe	Examinatrice
M. Alain Kuzniak	Université Paris Diderot	Directeur
M. Fabrice Vandebrouck	Université Paris Diderot	Co-directeur

Laboratoire de Didactique André Revuz (LDAR)
Université Paris Diderot
Bâtiment Sophie Germain
Case 7018
8 place Aurélie Nemours
75 205 Paris Cedex 13

ED 400
Université Paris Diderot
CAPE
Case 7078
5 rue Thomas Mann
75 205 Paris Cedex 13

Il est impossible que l'improbable n'arrive jamais.
Gumbel (1958)

Remerciements

Avant tout, je voudrais remercier mes deux directeurs de thèse. Merci à eux pour leur soutien et leur confiance tout au long de ces trois années de thèse. Merci d'avoir passé l'été avec les différents chapitres de ma thèse sur leur chevet. Merci à Alain de m'avoir intégrée à plusieurs projets, notamment au projet Ecos. J'en garderai un très bon souvenir. Merci à Fabrice de m'avoir acceptée comme doctorante pour son premier encadrement de thèse.

Je souhaite aussi remercier les membres de mon jury. Merci à Maria Alessandra Mariotti et Jean-Claude Régnier d'avoir accepté de rapporter ma thèse. Merci pour leurs relectures attentives et les remarques et commentaires constructifs qu'ils ont su faire sur mon manuscrit. Je voudrais tout particulièrement remercier Maria Alessandra Mariotti pour cette discussion que nous avons eu à la Summer School YESS 7 à Kassel, il y a deux ans, sans laquelle je n'aurais peut-être jamais eu cette idée d'ingénierie. Je suis très heureuse qu'elle ait pu découvrir l'aboutissement de mon travail de thèse en tant que rapporteure. Merci à Corinne Hahn et Giambattista Giacomin d'avoir accepté de faire partie de mon jury et d'avoir porté de l'intérêt à mon travail.

Je voudrais ensuite remercier tout particulièrement trois chercheurs qui m'ont énormément soutenue dans mon travail de thèse. Merci à Aline Robert. Merci d'avoir passé des après-midis à m'aider à trouver des façons d'analyser mes séances. Merci d'avoir été à l'écoute et de m'avoir donné énormément d'idées et de conseils. Ces moments particuliers ont été très importants pour moi. Merci aussi à Bernard Parzysz d'avoir accepté d'écrire un article (puis une communication) avec moi. J'ai appris énormément dans cette collaboration. Ce fût un véritable plaisir de travailler avec lui. Enfin, merci à Michel Henry, qui a accepté de me rencontrer en septembre 2015 pour discuter de l'Histoire des probabilités et qui, cet été, a passé énormément de temps à relire ma thèse et à me faire des retours et des commentaires, toujours constructifs. Le chapitre 3 lui doit beaucoup.

Je voudrais aussi remercier ceux sans qui ce travail n'existerait pas (en tout cas, pas comme tel) : les enseignants qui ont accepté d'échanger, de travailler avec moi. Je commence par celle à qui je dois le plus : merci à Sylvie, qui m'a fait confiance. La séquence conçue a vraiment été un travail collaboratif qui a été au-delà de mes espérances. J'ai adoré nos échanges, nos discussions. Merci à elle de s'être empressée de discuter avec son collègue Vincent après notre première « réunion » sur le projet et de le lui avoir tellement bien « vendu » qu'il s'en est aussitôt emparé. Merci à

Vincent de s'être laissé embarquer (malgré lui) dans l'aventure. Merci d'avoir testé la séquence et d'avoir ensuite accepté de tout me raconter. Merci à tous les deux pour leur enthousiasme, leur investissement et leur bonne humeur.

Merci à mes anciens collègues du lycée Touchard et du lycée Darius Milhaud et à mes amis profs qui ont accepté d'être interviewés et de faire passer les questionnaires dans leur classe. Merci à Valérie, à Ali, à Olivier, à Virginie, à Robert, à François, à Nicolas. Merci aussi à Laurent, à Pierre.

Je voudrais remercier les chercheurs et jeunes chercheurs que j'ai pu côtoyer tout au long de ma thèse. Tout d'abord, je voudrais remercier les chercheurs du laboratoire, toujours ouverts aux doctorants pour échanger. Le LDAR est un laboratoire de recherche où il fait bon vivre et où les doctorants ont leur place à part entière. Merci d'abord à Cécile de Hosson qui contribue en tant que directrice à ce climat. Merci d'être à l'écoute des doctorants et toujours de bon conseil. J'en profite pour la remercier d'avoir en-chanté mes vendredis midis (et ceux des autres choristes) depuis 3 ans, Ô huitième.

Merci à Christophe, à Julie, à Julia, à Joris, à Maha... et tous ceux avec qui j'ai pu partager des déjeuners sympathiques à la conviv', autour de discussions scientifiques mais pas que. Merci à Michèle Artigue et Laurent Vivier de m'avoir suivie pendant mon mémoire de M2. C'est finalement là le commencement du chemin vers le doctorat.

Je voudrais aussi remercier les jeunes chercheurs du LDAR (qui pour certains sont devenus « vieux »). Merci à tous les jeunes chercheurs avec qui j'ai pu discuter, échanger et rigoler. Ces moments partagés ont été essentiels pendant ces trois années. Merci aux jeunes chercheurs qui sont devenus « vieux » : Zoé, Lynn, Valentin, Robin, Edith, Cécile... Merci aux doctorants qui ont passé ces trois années avec moi : Assia (ma co-bureau pendant ces trois années, merci pour les discussions, les fous rires et les chambres d'hôtels partagées), Dominique, Robin, Sophie, Sophie, Soraya. Merci à ceux qui sont arrivés en cours de route : Alice, Blandine, Ines, Jorge, Léonard, Stéphane, Zakaria...

Merci à Monica d'être venue passer trois mois de son doctorat en France tout au début de ma thèse. Ce fût ma première co-auteure et elle est en même temps devenue une amie.

Merci à l'équipe chilienne de l'IMA qui m'a accueillie à bras ouverts lors de mon séjour d'un mois à Valparaíso dans le cadre du projet ECOS. Je pense en particulier à Elizabeth Montoya qui m'a totalement intégrée à l'équipe et qui m'a aussi fait visiter des superbes endroits. Merci à Jaime, Soledad... Merci aussi aux doctorantes chiliennes qui ont rendu mon séjour inoubliable. Merci à Paula, aux deux Carolina, à Isabel... Merci à Monica, Carolina et Romina, nous avons formé une bonne équipe internationale de doctorantes pour nos études comparatives. Muchas gracias a mi familia chilena Andrea, Marcelo y Pia.

Merci aussi aux chercheurs que j'ai pu rencontrer au sein de l'ARDM et notamment à ceux qui m'ont fait confiance dans le comité scientifique et d'organisation de l'École d'été à Brest, en me donnant des responsabilités. Merci à Ghislaine Gueudet,

à Yves Matheron, à Valentina Celi...
Merci aux jeunes chercheurs de l'ARDM qu'il est toujours plaisant de retrouver lors des WEJCH annuels : Anne, Céline, Laetitia, Maud, Valérie...

Merci aussi à Philippe Richard, Denis Tanguay et au regretté Jean-Philippe Drouhard avec qui j'ai pu échanger à plusieurs reprises pendant ces trois années.

Merci à Karine Chemla pour la direction de l'ED et la reconnaissance qu'elle a du travail des délégués représentants des doctorants de l'ED. Merci aux délégués de SPHERE avec qui j'ai apprécié organiser les journées des doctorants.

Merci à tous ceux qui rendent la vie du doctorant plus facile. Merci à Sandrine d'avoir toujours une solution en cas de problème administratif. Merci à Jérôme d'être capable de trouver n'importe quel article les yeux fermés (presque !). Merci à Évelyne d'avoir toujours le sourire. Merci à Laetitia de m'avoir aidée à régler tous mes problèmes informatiques. Merci à Martine et à Nadine.

Merci aux doctorants du 6ème qui ont partagé mon quotidien de ces trois années. Merci à Baptiste d'être mon plus fidèle co-bureau, à Marco pour toutes nos discussions (et pour son superbe accent anglais), à Victoria pour les belles flèches dans mes tableaux, à Kévin d'avoir été là pour m'entendre crier de joie après l'envoi de ma thèse, à Élie pour sa pyramide de boîtes de bonbons, à Julie d'avoir apporté un peu de Québec à l'étage... Merci à tous ceux avec qui j'ai pu prendre des goûters durant ces trois années.

Merci à mes nouveaux collègues de l'ESPE de Paris qui m'ont très bien accueillie au sein de l'équipe. Merci en particulier à Anne qui m'a proposé et a relu un chapitre de ma thèse. Merci au groupe IREM Analyse, nos rencontres ont ponctué ces trois années.

Je voudrais aussi remercier les enseignants, que j'ai eus lors de mon parcours scolaire et universitaire, qui m'ont inconsciemment conduite vers la didactique des maths. Je pense tout d'abord à Alain Lagrais, mon prof de maths du lycée, qui m'a (encore plus !) donné envie d'enseigner les mathématiques. Puis, merci à Jean Julo de m'avoir passionnée dans l'UE Didactique de la licence de maths de l'Université de Rennes 1. Enfin, merci à Mihaï Gradinaru de m'avoir encouragée, pendant la préparation à l'agrégation, à m'inscrire au M2 de Didactique des Mathématiques de Paris 7 pour l'année suivante. Ces rencontres, ainsi que le M2 à Paris 7, ont fait que trois ans après l'agreg, je démarrais une thèse (au grand désarroi de ma maman ! ;)).

Je voudrais enfin remercier mes amies de toujours : Cycy, Lily, Lulu et Mélan. Merci à mes copains de LSDM et les « changéens » pour nos soirées, nos week-ends annuels à Saint-Maixent... Merci à la loi des 5. Merci à Aurelia, de me rappeler de mettre la musique à fond et de danser en cas de stress, à Marie, de m'avoir inspiré le prénom le plus présent dans ma thèse, à Maud, d'avoir été ma coach de thèse, à Virginie, de m'avoir donné le meilleur conseil de ma thèse « se donner de *petits*

objectifs atteignables chaque jour ». Merci aux visannais, pour les bons moments qu'on passe encore et toujours. Merci à Julie, Rodolphe, Anaël et Ludo. Merci à Myrtille et Béatrice d'avoir illuminé mes mercredis après-midis pendant ces trois années.

Merci à ma maman pour ses relectures consciencieuses. Désolée de l'avoir rendue malade avec mes 350 pages (mais elle l'a fait !). Promis, maintenant c'est fini, je ne compte pas inventer une nouvelle thèse. Merci à mon papa de toujours être fier de ses filles. Merci à ma sœur d'avoir aussi relu des bouts de ma thèse et surtout de m'avoir assistée dans les dernières corrections, notamment la veille d'envoyer la thèse aux rapporteurs.

Merci à Thibault d'avoir toujours trouvé des solutions à mes problèmes de LaTeX, merci de m'avoir supportée pendant mes gros moments de stress, merci pour les vacances en Grèce, merci pour les petits plats... La liste serait trop longue, tout simplement merci.

Pour conclure, ces trois années de thèse ont été une expérience scientifique mais avant tout une aventure humaine que je n'oublierai pas. Merci à tous.

Résumé

Cette thèse porte sur les articulations entre les probabilités et l'analyse en classe de terminale scientifique. Nous avons exploré comment se créent et sont exploités les liens entre les sous-domaines mathématiques des probabilités des lois à densité et du calcul intégral, à travers une recherche centrée sur la notion de fonction de densité. En adoptant le modèle des Espaces de Travail Mathématique et des éléments de la théorie de l'activité, nous nous sommes demandé quelles tâches permettent d'introduire cette notion et de construire la relation sémiotique $P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x$. Pour aborder cette question, nous avons commencé par faire une étude épistémologique et historique de la naissance de la notion de lois à densité, qui nous a notamment permis de dégager la place importante de la statistique dans cette genèse. Puis, nous avons effectué une analyse des documents institutionnels et des manuels. Cette analyse a montré que l'articulation entre probabilités à densité et calcul intégral est imposée aux élèves et peu exploitée dans les différentes tâches qui leur sont proposées. Enfin, nous avons étudié la conception et la mise en place de tâches d'introduction originales grâce à une méthodologie de recherche que nous qualifions d'ingénierie didactique collaborative. Ces tâches ont pour objectif de faire construire, par le « collectif » classe, la notion de fonction de densité et d'amener le besoin du calcul d'aire sous une courbe. Nous avons mis en évidence les activités de ce collectif classe, dans la construction de cette notion, en analysant les circulations entre trois sous-domaines : les probabilités à densité, la statistique descriptive et le calcul intégral.

Mots-clefs

Didactique des mathématiques, Probabilités, Statistique, Calcul intégral, Fonction de densité, Espace de Travail Mathématique, Théorie de l'activité.

The density function at the crossroads between probability and calculus in the scientific track of the Grade 12.

Studying of the design and implementation of introductory tasks articulating continuous probability distribution and integral calculus.

Abstract

This thesis focuses on the connections between probability and analysis (calculus) in the scientific track of Grade 12 (French baccalaureate program). We explored the ways in which links between the mathematics subfields of continuous probability and integral calculus are created and explored, through a research focused on the concept of density function. Using the Mathematical Working Space model and some

elements of Activity Theory, we sought to identify tasks that would allow introducing this concept and building the semiotic relationship $P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x$. In order to address this issue, we began with an epistemological and historical study of the birth of the concept of density function, which enabled us to identify the important role of statistics in this genesis. Then, an analysis of institutional documents and textbooks showed that the link between continuous probability and integral calculus is imposed on students and rarely exploited in the different tasks given to them. Finally, we studied the design and implementation of original introductory tasks through a research methodology that we call "collaborative didactic engineering". The goal of these tasks is to get the class "collective" to construct the concept of density function and trigger the need for calculating areas under a curve. We highlighted the activities of the class "collective" in the construction of this notion by analyzing articulations between the three subfields: continuous probability, descriptive statistics and integral calculus.

Keywords

Mathematics education, Probability, Statistics, Integral calculus, Density function, Mathematical Working Space, Activity Theory.

Table des matières

Introduction générale **17**

1 Préliminaires **21**
1.1 Évolution de l'enseignement des probabilités et de la statistique . . . 21
1.2 Place actuelle des probabilités et de la statistique 23
1.3 Rappels mathématiques sur les probabilités 24
1.3.1 Généralités 24
1.3.2 Probabilités discrètes 28
1.3.3 Probabilités à densité 29
1.3.4 Loi faible des grands nombres 34
1.3.5 Théorème limite central 34

2 Problématique, cadre théorique et méthodologie **37**
2.1 Problématique 37
2.1.1 Des polémiques 37
2.1.2 Les liens entre lois de probabilités à densité et calcul intégral . 38
2.1.3 Vers nos questions de recherche 41
2.2 Les travaux en didactique portant sur les lois à densité 41
2.2.1 Travaux sur les lois à densité 42
2.2.2 Des brochures pour la formation des enseignants 45
2.2.3 Quelques difficultés observées chez les élèves en probabilités . 46
2.2.4 Le manque de formation des enseignants en probabilités et statistique 46
2.2.5 Le lien entre statistique et probabilités ou entre réalité et modèle 47
2.3 Principaux choix théoriques 47
2.3.1 Le modèle des Espaces de Travail Mathématique : cadre qui guide notre étude 48
2.3.2 Quelques éléments de théories de l'activité qui complètent notre cadre théorique 54
2.4 Questions de recherche et méthodologie générale 57
2.4.1 Nos questions de recherche 58
2.4.2 Objectifs et méthodologie générale de la recherche 59

Partie I Approche épistémologique et historique **63**

3 Étude historico-épistémologique **67**

3.1 Méthodologie 67
3.2 Émergence historique du calcul intégral : absence des probabilités . . 68
3.3 La naissance des lois de probabilité à densité 70
3.3.1 Les débuts des probabilités : les jeux de hasard 70
3.3.2 Le problème de Bernoulli : apparition d'intégrales 70
3.3.3 Buffon et le Franc-carreau : première rencontre avec la loi uniforme continue 72
3.3.4 Les travaux de Bayes : présence d'une courbe de densité . . . 73
3.3.5 La théorie des erreurs : naissance des lois à densité 73
3.3.6 La théorie moderne des probabilités 82
3.4 Conclusion de l'étude historico-épistémoloqique 83

Partie II Un regard sur les articulations entre probabilités à densité et calcul intégral dans l'enseignement en terminale S 87

4 Étude des documents institutionnels 91
4.1 Méthodologie de l'étude des documents officiels 91
4.1.1 Le choix des documents 91
4.1.2 Méthode d'analyse des documents 92
4.2 Analyse du programme 93
4.2.1 Présentation du programme 93
4.2.2 Des articulations apparentes ou possibles 98
4.2.3 Un programme avec des « incohérences » entre probabilités et analyse 99
4.2.4 Conclusion sur le programme 100
4.3 Analyse du document *Ressources* 101
4.3.1 Présentation du document 102
4.3.2 Premiers éclaircissements et apparition de nouvelles ambiguïtés 103
4.3.3 L'introduction des différentes lois 104
4.3.4 Autres articulations 109
4.3.5 Des tâches qui n'exploitent pas l'articulation entre probabilités à densité et calcul intégral 111
4.3.6 Conclusion sur le document *Ressources* 113
4.4 Analyse des sujets de baccalauréat 114
4.4.1 Présentation des sujets de baccalauréat 114
4.4.2 Répartition des exercices de baccalauréat portant sur les probabilités à densité et le calcul intégral 115
4.4.3 Différences repérées entre les deux sous-domaines 117
4.4.4 Des types de tâches caractérisant les deux sous-domaines . . . 118
4.4.5 Des tâches qui exploitent les liens entre les deux sous-domaines ou pouvant en créer 120
4.4.6 Conclusion sur les exercices du baccalauréat 124
4.5 Conclusion sur les dynamiques entre les ETM de référence 124

5 Étude des manuels de terminale S 127
5.1 Méthodologie de l'étude des manuels . . . 128
5.1.1 Justification du choix de l'étude des manuels . . . 128
5.1.2 Présentation générale des manuels . . . 129
5.1.3 Outils méthodologiques pour l'analyse *a priori* des *activités* d'introduction de la notion de fonction de densité . . . 131
5.1.4 Méthodologie pour l'observation des tâches exploitant les connexions entre les deux sous-domaines . . . 132
5.2 Analyse des activités d'introduction . . . 133
5.2.1 Avant de commencer . . . 133
5.2.2 Introduction de la notion de fonction de densité : les différentes approches rencontrées dans les manuels . . . 136
5.2.3 Analyse *a priori* des *activités* introductives utilisant la notion d'histogramme . . . 139
5.2.4 Résultats . . . 166
5.2.5 Compléments sur la notion d'histogramme . . . 169
5.2.6 Conclusion . . . 174
5.3 Des connexions au niveau du calcul d'aires . . . 174
5.3.1 Mise en évidence de différents niveaux de calculs d'aires . . . 175
5.3.2 Le niveau A1 . . . 176
5.3.3 Le niveau A2 . . . 178
5.3.4 Le niveau A3 . . . 180
5.3.5 Récapitulatif . . . 183
5.4 Un point de vue sur les choix des enseignants . . . 185
5.4.1 Méthodologie relative aux entretiens . . . 185
5.4.2 Élements des entretiens . . . 186
5.5 Conclusion sur les ETM idoines . . . 188

Partie III Élaboration d'une ingénierie didactique collaborative. Des problèmes de modélisation articulant l'introduction des probabilités à densité et du calcul intégral 195

6 Présentation de l'ingénierie 199
6.1 Méthodologie de type ingénierie didactique collaborative . . . 199
6.1.1 La méthodologie de type ingénierie didactique . . . 200
6.1.2 Une recherche collaborative . . . 200
6.1.3 Vers une méthodologie de type ingénierie didactique collaborative . . . 203
6.1.4 La mise en œuvre de la méthodologie . . . 204
6.2 Description de la phase de conception . . . 208
6.2.1 Des contraintes à respecter . . . 208
6.2.2 Prise en compte des pratiques habituelles de Marie . . . 209
6.2.3 Une première proposition : une réorganisation et quelques modifications . . . 211

6.2.4 Principaux choix retenus pour la conception des tâches et justification de ces choix . . . 213
6.3 Présentation du scénario de la séquence . . . 215
6.3.1 Avant la séquence . . . 215
6.3.2 Introduction de la notion de fonction de densité (1) : le problème de la rencontre . . . 217
6.3.3 Introduction de la notion de fonction de densité (2) et du calcul intégral : le problème du volcan Aso . . . 218
6.3.4 Un cours et des exercices articulant calcul intégral et probabilités à densité . . . 219
6.3.5 Un autre devoir maison . . . 219
6.4 Précisions sur la question de recherche . . . 220
6.4.1 Éléments théoriques complémentaires . . . 220
6.4.2 Nos nouvelles questions de recherche . . . 221

7 Analyses des scénarios et des déroulements 223
7.1 Méthodologie des analyses . . . 224
7.1.1 Précisions sur l'expérimentation . . . 224
7.1.2 Base de notre méthodologie . . . 226
7.1.3 Prise en compte de la démarche de modélisation en probabilités 227
7.1.4 Adaptation de la méthodologie aux tâches de modélisation . . . 233
7.1.5 Grilles d'analyses globales des deux problèmes . . . 234
7.1.6 Grilles d'analyses locales . . . 237
7.1.7 Dernier niveau d'analyse : l'institutionnalisation . . . 240
7.2 Le problème de la rencontre . . . 241
7.2.1 Éléments de contexte . . . 241
7.2.2 Analyse *a priori* du problème de la rencontre . . . 242
7.2.3 Analyse du déroulement . . . 254
7.2.4 Analyse de la première institutionnalisation . . . 274
7.2.5 Conclusion sur le problème de la rencontre . . . 283
7.3 Le problème du volcan Aso . . . 285
7.3.1 Analyse *a priori* du problème du volcan Aso . . . 286
7.3.2 Analyse du déroulement . . . 296
7.3.3 Analyse de la deuxième institutionnalisation . . . 318
7.3.4 Conclusion sur le problème du volcan Aso . . . 323
7.3.5 Éléments sur le cours associé à la séquence . . . 326
7.4 Validation interne . . . 327
7.5 Discussion sur les choix méthodologiques . . . 329
7.5.1 Prise en compte de la démarche de modélisation . . . 329
7.5.2 Alliance entre les méthodologies des théories de l'activité et des ETM . . . 330
7.5.3 Prise en compte des changements de domaine dans les ETM et retour sur l'articulation entre probabilités à densité et calcul intégral . . . 332
7.5.4 Analyse du collectif classe . . . 332

8 Sur la portée de nos résultats 335
8.1 Validation de l'enseignante . . . 335
8.1.1 Respect des contraintes . . . 336
8.1.2 Son ressenti d'enseignante . . . 336
8.2 Limites de l'analyse du collectif . . . 337
8.3 Validation complémentaire : questionnaire de fin d'année . . . 339
8.3.1 Méthodologie relative au questionnaire . . . 340
8.3.2 Résultats de ce questionnaire . . . 341
8.3.3 Conclusion sur le questionnaire . . . 348
8.4 La question de la reproductibilité . . . 349
8.4.1 Le cas de Laurent . . . 349
8.4.2 Des conditions à la reproductibilité . . . 351
8.5 Impacts et portée sur nos résultats . . . 352

Conclusion générale 353
9.1 Synthèse des résultats . . . 353
9.2 Limites de cette recherche . . . 356
9.3 Apports de ce travail . . . 357
9.4 Prolongements et perspectives . . . 358
9.4.1 Autour de l'ingénierie . . . 358
9.4.2 Sur les connaissances des élèves . . . 359
9.4.3 Vers de nouvelles notions à étudier . . . 359
9.4.4 Un questionnement plus théorique . . . 359

Bibliographie 361

Liste des manuels de terminale S 373

Liste des manuels de seconde 375

Annexes 379

A Documents historiques 383

B Le programme de terminale S 389

C Le document *Ressources* 393

D Les sujets de baccalauréat 396

E Les manuels 428

F Les entretiens 444

G La séquence conçue 495

H Planning de la séquence 524

I Les données relatives à la séquence **526**

J Compléments **651**

Introduction générale

Depuis quelques dizaines d'années, les probabilités et la statistique ont fait leur apparition dans l'enseignement secondaire de nombreux pays. Les raisons de l'introduction de ces deux domaines mathématiques dans l'enseignement ont été mises en évidence dans de nombreuses recherches (Holmes, 1980 ; Hawkins, Jolliffe & Glickman, 1991 ; Vere-Jones, 1995 ; Wild & Pfannkuch, 1999 ; Gal, 2002). Batanero, Godino et Roa (2004) en reprennent les principales, qui sont l'utilité des probabilités et de la statistique dans la vie quotidienne, leur rôle d'outil dans d'autres disciplines (physique, économie, biologie...), leur rôle dans le développement du raisonnement critique et la nécessité de connaissances statistiques dans de nombreuses professions. Nous pouvons remarquer un choix de l'enseignement de se rapprocher des pratiques de la vie professionnelle et sociale. De nos jours, en plus d'être considérées comme utiles à la formation du citoyen (comme on peut le retrouver explicitement dans les programmes français), les probabilités et la statistique offrent de nouvelles possibilités d'exploitation des outils technologiques à la disposition des élèves (logiciels, calculatrice...).

En France, au fur et à mesure des réformes, une place grandissante est laissée à l'enseignement des probabilités et de la statistique. En probabilités, les premières notions sont abordées en classe de troisième, puis elles sont enrichies tout au long du lycée, dans les différentes filières. Actuellement, les différentes lois étudiées ne se limitent plus aux lois de probabilité discrètes. Depuis la rentrée 2012, certaines lois de probabilité à densité sont apparues dans les programmes de la classe de terminale de nombreuses filières (et pas seulement des filières générales) ; les premières lois à densité (uniforme et exponentielle) sont apparues dès la réforme des années 2000 en terminale scientifique.

D'un point de vue épistémologique, le calcul de probabilités de lois à densité marque un véritable changement par rapport au calcul de probabilités de lois discrètes. En effet, dans le premier cas, il ne s'agit plus de définir une loi en décrivant la probabilité associée à chaque issue possible, mais de la définir par la probabilité de tout intervalle. Pour une variable aléatoire X suivant une loi à densité, on va avoir :

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x, \qquad (*)$$

où f est la fonction de densité de probabilité associée à la loi de X. On voit ici apparaître une écriture sémiotique qui a la particularité d'établir une relation entre deux domaines mathématiques, les probabilités et l'analyse. Cette écriture sémiotique est un macro-signe qui représente une égalité entre deux signes, l'un associé aux probabilités et l'autre à l'analyse, plus précisément au calcul intégral. Plus qu'une simple

écriture sémiotique, nous appellerons ce signe une relation sémiotique, du fait de la relation entre domaines qu'elle induit.

La question est alors de savoir quelle véritable relation ce macro-signe établit entre les deux domaines. Nous pouvons d'abord nous poser la question de la relation entre probabilités et analyse d'un point de vue historique et épistémologique. Quelle est l'origine de la relation qui se traduit par ce signe ? Ensuite, d'un point de vue didactique, nous pouvons chercher à comprendre quel sens est donné à cette relation sémiotique dans les classes et quel sens lui donnent les élèves. Cette relation sémiotique s'arrête-t-elle simplement à ce macro-signe ou est-ce que cela va au-delà ? Une véritable relation entre les deux domaines existe-elle ?

Les probabilités à l'université

En licence de mathématiques à l'université, la théorie des probabilités de Kolmogorov est introduite en troisième année, sans prendre en compte ce qui a pu être vu antérieurement par les étudiants sur ce thème. D'une université à l'autre, le déroulement est sensiblement le même. Les étudiants rencontrent d'abord la théorie de la mesure puis la théorie de l'intégration de Lebesgue. Ensuite, la théorie des probabilités arrive en s'appuyant essentiellement sur les deux précédentes. L'enseignement des probabilités, à ce niveau, s'appuie sur l'axiomatique de Kolmogorov. Les mesures de probabilités sont un cas particulier de mesure et il en découle que le calcul de probabilités est une application de l'intégration de Lebesgue. Au sein de cette théorie, les mesures de probabilités discrètes et continues sont étudiées dans le même cadre. À l'université, avec cet enseignement des probabilités, intégration et probabilités sont donc liées par construction même des objets mathématiques.

À ce niveau universitaire, des difficultés peuvent tout de même émerger. Garet et Kurtzmann (2011), dans leur livre de cours de probabilités à destination des étudiants de L3-M1, disent que « *l'expérience montre que les difficultés* [des étudiants] *sont bien plus souvent liées à une défaillance du lien entre le cours d'intégration et ses applications en probabilités, qu'à un défaut de compréhension des probabilités elles-mêmes* » (p. i). Ces enseignants mettent le doigt sur les difficultés des étudiants à faire des liens entre intégration et probabilités, malgré cette construction des probabilités.

Le choix de la classe

Au lycée, la problématique est différente. Les théories de la mesure et de l'intégration de Lebesgue ne sont évidemment pas au programme. Cela n'empêche pas pour autant l'introduction de certaines lois à densité. Nous allons nous intéresser à ce niveau d'enseignement car il s'agit de la première rencontre avec ces lois de probabilités, c'est donc un moyen pour nous de comprendre la genèse de la relation sémiotique $(*)$ dans l'enseignement et de voir le sens qui lui est donné.

Dans beaucoup de filières du lycée, l'enseignement des lois à densité se limite à l'étude des lois normales. Dans ces classes, le calcul intégral n'est pas abordé. La relation sémiotique $(*)$ n'apparaît donc pas explicitement (bien qu'elle soit présente implicitement). Le travail sur les lois normales se restreint souvent à une utilisation

de la calculatrice. Pour notre recherche, nous avons donc voulu choisir une classe dans laquelle le calcul intégral est au programme et l'enseignement des lois à densité plus développé que simplement l'étude des lois normales. Pour cette raison nous avons choisi la classe de terminale scientifique (terminale S). C'est dans cette section du lycée que ces deux parties du programme sont les plus conséquentes.

Nous faisons l'hypothèse que peu de sens est donné à cette relation sémiotique dans les classes.

Objectif de la thèse

Nous avons commencé à introduire le contexte général de notre questionnement. La question principale qui va guider notre travail de recherche est la suivante :

Est-il possible de concevoir et de mettre en œuvre des problèmes d'introduction aux lois à densité qui permettent aux élèves de terminale scientifique de construire conceptuellement la relation sémiotique $(*)$ **?**

Les objectifs de cette thèse vont être d'ordre épistémologique et didactique. Il s'agira de mieux comprendre la relation induite par le signe $(*)$, c'est-à-dire sa genèse historique et épistémologique, son rôle dans l'enseignement en terminale S, pour ensuite se poser la question de son introduction dans la classe de terminale et de la possibilité de lui faire prendre tout son sens.

Plan de l'étude

Avant de préciser notre problématique de recherche, le **chapitre 1** présentera des éléments préliminaires. Nous commencerons par donner des précisions sur le contexte de l'enseignement des probabilités et de la statistique en France. Nous ferons aussi des rappels mathématiques sur les probabilités et nous commenterons l'adaptation qui en est faite au lycée. Dans le **chapitre 2**, nous détaillerons notre problématique, ainsi que notre cadre théorique et notre méthodologie générale. Notre étude sera ensuite structurée en trois parties.

La **Partie I** est constituée d'un seul chapitre (**chapitre 3**). Nous proposerons, dans cette partie, une analyse historique et épistémologique de la genèse des lois à densité et de la notion de fonction de densité, afin de mieux comprendre la construction historique de la relation entre les probabilités et l'analyse.

La **Partie II** est dédiée à un état des lieux de l'enseignement des lois à densité et sa mise en relation avec le calcul intégral dans la classe de terminale S. Cette partie est composée de deux chapitres (**chapitres 4 et 5**). Dans le **chapitre 4**, nous présenterons une analyse des documents institutionnels, à savoir le programme, le document *Ressources* ainsi que les sujets du baccalauréat. Dans le **chapitre 5**, il s'agira d'étudier l'introduction de la fonction de densité dans les manuels et dans les classes, à partir d'entretiens d'enseignants.

La **Partie III** (**chapitres 6, 7 et 8**) est consacrée à l'ingénierie didactique collaborative que nous avons mise en place. Dans le **chapitre 6**, nous décrirons la méthodologie de type ingénierie didactique, prenant en compte une dimension collaborative, qui a guidé notre travail. En prenant appui sur les résultats des deux parties

précédentes, nous exposerons les choix que nous avons retenus dans la conception des problèmes d'introduction. Dans le **chapitre 7**, nous analyserons dans le détail les mises en œuvre des problèmes d'introduction dans le cadre d'une expérimentation dans une classe de terminale S et nous proposerons, dans le **chapitre 8**, une ouverture sur nos résultats pour essayer d'élargir la portée de nos résultats.

Nous terminerons par une **conclusion générale** dans laquelle nous reviendrons sur nos questions de recherche en essayant de dégager ce que ce travail nous a appris sur l'introduction des lois à densité en classe de terminale S. Nous présenterons les apports de cette thèse et nous proposerons des prolongements possibles à notre travail.

Chapitre 1

Préliminaires : l'enseignement des probabilités et de la statistique en France et quelques rappels mathématiques

Avant de présenter notre problématique et les questions de recherche de notre travail, nous allons resituer l'enseignement des probabilités et de la statistique dans l'enseignement secondaire en France. Ce chapitre a pour but de « contextualiser » cet enseignement, moins étudié que celui de l'algèbre ou de la géométrie, par exemple.

Nous présenterons tout d'abord, dans la section 1.1, l'évolution des programmes en probabilités et en statistique au cours des différentes réformes (à partir de 1954). Dans la section 1.2, nous insisterons plus particulièrement sur la place actuelle des probabilités et de la statistique dans l'enseignement secondaire. Nous parlerons des programmes avant la réforme des cycles 2, 3 et 4, mise en vigueur à la rentrée 2016. Cependant, nous tenons à préciser que la réforme n'influence que peu les contenus en probabilités et statistique. Nous ferons ensuite, dans la section 1.3, quelques rappels sur les notions probabilistes du niveau de deuxième année d'université (L2) pour prendre un peu de hauteur sur les mathématiques en jeu dans la thèse, sans pour autant entrer dans la théorie de Kolmogorov. En parallèle, nous ajouterons des remarques pour préciser si ces notions sont abordées dans l'enseignement secondaire et nous préciserons les éventuelles transpositions didactiques qui en sont faites.

1.1 Évolution de la place des probabilités et de la statistique dans l'enseignement secondaire

Les probabilités et la statistique font leur apparition tardivement dans l'enseignement secondaire. Ceci est à mettre en lien avec le développement assez récent des domaines mathématiques des probabilités et de la statistique. Nous allons présenter ici quelques éléments d'une étude de Parzysz (2003) sur l'évolution des programmes en probabilités et en statistique entre 1965 et 2002.

Il faut attendre 1954 pour voir apparaître pour la première fois l'enseignement de

la statistique dans l'enseignement secondaire et seulement dans une seule section, la section économique (B). La statistique entre véritablement dans les programmes en 1965, en classe de première des sections A (littéraire), B (économique) et D (scientifique option sciences naturelles). Les probabilités sont quant à elles introduites l'année suivante, en classe de terminale des mêmes sections. Nous pouvons signaler qu'aucun enseignement des probabilités et statistique n'est donné dans la section C (scientifique option mathématiques et physique). Dans les sections A et B, plusieurs lois de probabilités sont au programme, entre autres, la loi binomiale, la loi de Poisson ainsi que la loi normale, comme loi limite de la loi binomiale.

En 1970, le programme de la réforme des « mathématiques modernes » est mis en place. C'est seulement à ce moment-là que les probabilités et la statistique arrivent dans toutes les filières scientifiques (C, D et E). La présentation des probabilités est alors axiomatique. Nous trouvons les notions d'espace probabilisable et d'espace probabilisé [1] (fini) en classes de première et de terminale. Certaines notions, comme la loi normale, sont enlevées. Seules la loi binomiale et quelques autres lois discrètes finies restent étudiées. L'étude des probabilités illustre l'intérêt des structures ensemblistes qui sont alors au cœur des programmes de la réforme des « mathématiques modernes ». En effet, les probabilités et la statistique permettent aux élèves de réinvestir leurs connaissances sur les ensembles (réunion, intersection,...). À cette époque, les probabilités et la statistique ne sont à aucun moment connectées.

Cette séparation entre les deux domaines est renforcée pendant la période de 1981 à 1986. L'enseignement des probabilités est alors moins ambitieux : l'aspect axiomatique a disparu et les probabilités sont présentées comme un « sous-produit » du dénombrement. Seule la loi binomiale reste étudiée. À partir de 1986, l'enseignement de la statistique descriptive est avancé à la classe de sixième et se poursuit jusqu'à la classe de première. Les calculs de probabilités et la loi binomiale sont introduits seulement en terminale.

La réforme des années 1990 maintient l'enseignement de la statistique (descriptive) à partir du début du collège. En revanche, un grand changement marque l'enseignement des probabilités : l'approche recommandée est alors une approche « fréquentiste » (Henry, 2010). La probabilité est définie par l'observation de la stabilisation de la fréquence dans la répétition d'une expérience aléatoire. Il s'agit d'une approche expérimentale. La liaison probabilités-statistique est présente, avec cependant un risque didactique que les élèves assimilent fréquence observée et probabilité (Girard, 2004).

La réforme suivante (dans les années 2000-2002) est celle que Girard (*ibid.*) appelle la « révolution statistique ». Un volume horaire plus conséquent est consacré à la statistique en classe de seconde. De plus, le programme introduit de nouveaux aspects comme les fluctuations d'échantillonnage et la simulation. En classe de première, la probabilité est ensuite définie comme une distribution de fréquences théoriques intervenant dans la modélisation d'une situation aléatoire. Enfin, les lois de probabilités continues (loi uniforme sur $[0; 1]$ et lois exponentielles) font leur apparition dans le programme de terminale S. Il en est de même pour les tests statistiques avec la vérification de l'adéquation d'une distribution observée à une loi équirépartie grâce à un test. Dans ce programme, il y a une volonté de lier statistique et pro-

1. Nous préciserons ces notions dans la section 1.3.

babilités, notamment à travers la loi des grands nombres (énoncée sous une forme « vulgarisée »).

Au fur et à mesure des réformes, l'enseignement de la statistique et des probabilités s'est étoffé et a pris une place de plus en plus importante dans les programmes. Des approches différentes se sont succédées, avec une tendance récente à relier les domaines de la statistique et des probabilités. En ce qui concerne les probabilités à densité, nous remarquons que la loi normale est au programme entre 1966 et 1970, mais elle est ensuite retirée ce qui limite l'enseignement des probabilités à l'étude de lois discrètes. Il faut attendre 2002 pour que réapparaissent des lois à densité. Il s'agit alors de la loi uniforme sur $[0;1]$ et des lois exponentielles. Intéressons-nous maintenant aux programmes actuels.

1.2 Place actuelle des probabilités et de la statistique dans l'enseignement secondaire

Dans le cadre de la réforme des collèges de 2008 et celle des lycées, les années suivantes, nous retrouvons la statistique descriptive à partir de la classe de sixième jusqu'en terminale. La statistique inférentielle, avec les intervalles de fluctuation et de confiance, est maintenant au programme de la seconde à la terminale. Le début de l'enseignement des probabilités est quant à lui avancé à la classe de troisième.

Au collège, « *seule la statistique exploratoire est abordée et l'aspect descriptif constitue l'essentiel de l'apprentissage* » (MEN, 2008, p. 5) autour de « *quelques notions fondamentales de statistique descriptive* » (*ibid.*, p. 10). Ensuite, en troisième, les élèves doivent « *se familiariser avec les notions de chance et de probabilité* » (*ibid.*, p. 10). Il s'agit d'initier les élèves à la probabilité sur des exemples simples. En seconde, ce sont les situations d'équiprobabilité qui sont privilégiées, suivies en première par des exemples de variables aléatoires suivant des lois binomiales (dans la filière scientifique, en particulier). Jusqu'en classe de terminale, seules les probabilités discrètes sur un ensemble fini sont au programme. En terminale scientifique (mais pas seulement), depuis la rentrée 2012, la partie sur les lois de probabilité à densité est une partie non négligeable du programme. Trois exemples sont abordés : les lois uniformes sur un intervalle $[a;b]$, les lois exponentielles et les lois normales. Dans les différents niveaux, il est précisé que statistique et probabilités sont en relation étroite et qu'il est souhaité de faire des allers-retours entre ces deux domaines.

Nous pouvons repérer dans les programmes actuels, une forte augmentation de la place des probabilités et c'est notamment le cas pour les lois à densité. En 2002, cette partie représentait trois lignes du programme et maintenant trois pages. L'enseignement des probabilités qui pouvait être « négligé » par certains enseignants auparavant ne peut actuellement plus l'être, vu son importance en ce qui concerne la proportion dans le programme (un huitième du programme de terminale S) et sa présence constante au baccalauréat.

1.3 Rappels mathématiques sur les probabilités et commentaires

Dans cette section, nous allons présenter des rappels théoriques sur les probabilités. Pour rester dans le cadre qui nous intéresse dans la thèse, nous n'allons pas présenter les notions dans le cadre de la théorie de Kolmogorov. Nous allons nous limiter à un niveau L2 (deuxième année d'université). Il ne s'agit pas ici de présenter un cours de probabilités, mais bien de s'appuyer sur les définitions et propriétés rigoureuses, dans un langage probabiliste, pour mettre en évidence les adaptations qui en sont faites en terminale S ou dans les classes antérieures. Nous ajouterons des compléments en langage naturel quand cela nous paraîtra nécessaire. Nous ne ferons pas de rappels en statistique bien que les lois normales, par exemple, puissent être utilisées dans ce domaine dans le programme de terminale. Nous rappelons que les intégrales sont étudiées pour la première fois par les élèves en classe de terminale.

Les rappels théoriques sont inspirés de plusieurs livres sur la théorie des probabilités (Jacob & Protter, 2003, Escoffier, 2010) et des notes de cours d'Alain Troesch (2007, 2012) disponibles sur internet.

1.3.1 Généralités

La théorie des probabilités est une branche des mathématiques qui fournit des modèles mathématiques pour décrire le « hasard » ou plus spécifiquement les phénomènes aléatoires.

a) Expérience aléatoire et univers

En probabilités, une *expérience aléatoire* est une expérience dont le résultat n'est pas connu d'avance et peut varier si l'on répète cette expérience. Le résultat de l'expérience dépend du « hasard ».

Définition. On appelle *univers* un ensemble représentatif des résultats possibles de l'expérience aléatoire. Ses éléments ω sont aussi appelés *issues*. L'univers est souvent noté Ω.

Exemples.
1. Jeter une pièce de monnaie : $\Omega = \{P; F\}$.
2. Lancer de dés : $\Omega = \{1; 2; 3; 4; 5; 6\}$.
3. Durée de vie d'une ampoule : $\Omega = \mathbb{R}^+$.
4. Mesure d'une longueur L avec une erreur de mesure. $w \in \Omega$ désigne le résultat de la mesure : $\Omega = \mathbb{R}^+$. $w - L \in \Omega'$ est l'erreur de la mesure : $\Omega' = \mathbb{R}$.

Adaptation dans l'enseignement secondaire
Les premiers exemples simples d'expériences aléatoires sont étudiés en classe de troisième. Ils peuvent notamment être abordés en procédant à des expérimentations. La notion d'univers est explicitée en seconde. Les univers rencontrés sont des ensembles discrets finis jusqu'en première. C'est en terminale qu'apparaissent des univers sous forme d'intervalles. Cependant, ceci n'est pas nécessairement explicité avec ce vocabulaire.

b) Événements

Définition. On dit que $\mathcal{A}$ est une tribu sur l'ensemble Ω si $\mathcal{A}$ est une famille de parties de Ω telle que :
- $\varnothing \in \mathcal{A}$,
- $\forall A \in \mathcal{A}, \overline{A} \in \mathcal{A}$,
- $\forall i \in I \subset \mathbb{N}, A_i \in \mathcal{A} \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{A}$.

Définition. On appelle *espace probabilisable* un couple $(\Omega, \mathcal{A})$, où $\mathcal{A}$ est une tribu de Ω. Les éléments de $\mathcal{A}$ sont appelés *événements*.

Exemples.
1. Si Ω est fini ou dénombrable, on considère le plus souvent $\mathcal{A} = \mathcal{P}(\Omega)$.
2. Si $\Omega = \mathbb{R}$, on ne considère jamais $\mathcal{A} = \mathcal{P}(\Omega)$. Généralement, on considère la tribu des boréliens $\mathcal{B}(\mathbb{R})$ (il s'agit de la tribu engendrée par l'ensemble des intervalles ouverts $]a; b[$, avec a et $b \in \mathbb{R}$).

En pratique, on appelle aussi événement une propriété dont on peut dire si elle est vraie ou fausse, en fonction du résultat de l'expérience une fois réalisée. Un événement (propriété) A est représenté, dans le modèle probabiliste, par l'évènement (partie de Ω) suivant :

$$A = \{w \in \Omega | A \text{ est vraie lorsque } w \text{ est l'issue de l'expérience}\}.$$

Définition. On représente :
- un *événement élémentaire* par un singleton $\{w\} \subset \Omega$
- l'*événement certain* par l'ensemble Ω,
- l'*événement impossible* par la partie $\varnothing$,
- l'*événement contraire* de l'événement A (non A) par la partie $\overline{A} = \Omega \setminus A$,
- la *disjonction* des événements A et B (A ou B) par la partie $A \cup B$,
- la *conjonction* des événements A et B (A et B) par la partie $A \cap B$.

On dit que les événements A et B sont *incompatibles* si $A \cap B = \varnothing$.

Adaptation dans l'enseignement secondaire
La notion d'événement est rencontrée à partir de la classe de seconde. On trouve en particulier les définitions d'événement certain, contraire... Cependant, la notion d'événement est traitée au lycée sous forme d'exemples sans être rigoureusement définie. Cela est la même chose sur la plupart des notions qui vont suivre. En effet, la définition d'événement que nous avons donnée nécessiterait un niveau d'abstraction trop élevé pour les élèves de lycée.

c) Probabilité

Définition. Étant donné un espace probabilisable $(\Omega, \mathcal{A})$, on appelle *probabilité* sur $(\Omega, \mathcal{A})$ toute application $P : \mathcal{A} \to \mathbb{R}$ telle que :
- $\forall A \in \mathcal{A}, P(A) \geqslant 0$,
- $P(\Omega) = 1$,
- pour toute famille dénombrable $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux incompatibles, on a : $P(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \geqslant 0} P(A_n)$ (propriété de σ-additivité).

En théorie de la mesure, on dit que P est une mesure positive finie, de masse totale égale à 1.

Définition. On appelle le triplet $(\Omega, \mathcal{A}, P)$ un *espace probabilisé*.

Adaptation dans l'enseignement secondaire
La probabilité n'est pas définie comme une application dans l'enseignement secondaire. On parle plutôt de probabilité pour désigner l'image d'un événement par cette application, sans que cela soit explicité en ces termes. En revanche, les différentes propriétés sont utilisées par les élèves dès la classe de troisième (sur des ensembles finis), avec une restriction à la propriété d'additivité.

Définition. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et soit A et B deux éléments appartenant à $\mathcal{A}$.
On dit que les événements A et B sont *indépendants* si $P(A \cap B) = P(A)P(B)$.

Définition. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et soit A et B deux éléments appartenant à $\mathcal{A}$. Si $P(B) \neq 0$, on définit la *probabilité conditionnelle* de A sachant B, notée $P_B(A)$, par la relation :

$$P_B(A) = \frac{P(A \cap B)}{P(B)}$$

.

Adaptation dans l'enseignement secondaire
Les notions d'indépendance et de probabilités conditionnelles sont étudiées en classe de terminale S, avant le cours sur les lois de probabilité à densité.

d) Variable aléatoire

Définition. On appelle *variable aléatoire réelle* sur $(\Omega, \mathcal{A})$ toute application

$$X \ : \ \Omega \ \to \ \mathbb{R}$$

telle que

$$\forall B \in \mathcal{B}(\mathbb{R}), X^{-1}(B) \in \mathcal{A}.$$

Exemple. Jet de deux dés. $\Omega = [\![1; 6]\!]^2$. On peut considérer la variable aléatoire représentant la somme des deux dés :

$$X : w = (a; b) \mapsto a + b.$$

Définition. On appelle *loi de probabilité* (ou *distribution de probabilité*) d'une variable aléatoire réelle X, définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, l'application :

$$\begin{array}{rccc} \mu & : & \mathcal{B}(\mathbb{R}) & \to & [0; 1] \\ & & B & \mapsto & P(X^{-1}(B)). \end{array}$$

Remarque. $X^{-1}(B) = \{w \in \Omega | X(w) \in B\}$. On note souvent cet événement « $X \in B$ » (par abus d'écriture).

Proposition. *L'application μ ainsi définie est une probabilité sur l'espace probabilisable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Elle est appelée probabilité image de P par X. On la note souvent P_X.*

Adaptation dans l'enseignement secondaire
Dès la classe de troisième, les élèves rencontrent des exemples de variables aléatoires. Elles restent cependant implicites. Le vocabulaire apparaît en seconde mais les programmes précisent qu'aucune théorie ne doit être faite sur les variables aléatoires. Des variables aléatoires sont donc étudiées mais jamais explicitement définies. Il s'agit d'une application

$$\begin{array}{rccc} X & : & \Omega & \to & \mathbb{R} \\ & & w & \mapsto & X(w). \end{array}$$

Mais cette « définition » non plus n'est pas donnée. Il est clair que définir rigoureusement cette notion nécessite des concepts hors de portée des élèves (tribu, application réciproque...). Cela ne les empêche pas pour autant de s'intéresser aux lois que peuvent suivre les différentes variables aléatoires qu'ils rencontrent. Nous verrons plus loin comment définir les lois de probabilités discrètes et continues.

e) Fonction de répartition

Définition. Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. On appelle *fonction de répartition* de X l'application :

$$\begin{array}{rccc} F & : & \mathbb{R} & \to & [0;1] \\ & & x & \mapsto & P(X \leqslant x). \end{array}$$

Proposition. *Toute fonction de répartition possède les propriétés suivantes :*

1. *F est croissante.*
2. $\lim\limits_{x\to-\infty} F(x) = 0$ *et* $\lim\limits_{x\to+\infty} F(x) = 1$.
3. *F est continue à droite en tout point x de $\mathbb{R}$.*
4. *F a une limite à gauche en tout point x de $\mathbb{R}$ et*

$$\forall x \in \mathbb{R}, F(x) - \lim_{t\to x^-} F(t) = P(X = x).$$

Théorème. *Toute fonction $F : \mathbb{R} \to [0;1]$ telle que :*
- *F est croissante,*
- $\lim\limits_{x\to-\infty} F(x) = 0$ *et* $\lim\limits_{x\to+\infty} F(x) = 1$,
- *F est continue à droite en tout point x de $\mathbb{R}$,*

est la fonction de répartition d'une variable aléatoire réelle X.

Proposition. *La donnée de F caractérise la loi de probabilité μ.*

Adaptation dans l'enseignement secondaire
La fonction de répartition n'est pas une notion abordée au lycée. Elle n'est effectivement pas indispensable pour définir les différentes lois rencontrées au programme. Cette fonction apparaît tout de même implicitement en terminale S, notamment dans différentes démonstrations de cours.

1.3.2 Probabilités discrètes

a) Généralités

Définition. Une variable aléatoire réelle X est dite *discrète* si $X(\Omega)$ est finie ou dénombrable.

Définition. Pour une variable aléatoire discrète X, donner la loi de X c'est donner $X(\Omega)$ et la valeur $P(X = x_i)$, pour tout $x_i \in X(\Omega)$.

Soit $I \subset \mathbb{N}$. Soit X une variable aléatoire réelle discrète prenant les valeurs distinctes x_i $(i \in I)$.

Définition. On définit l'espérance $E(X)$ par : $E(X) = \sum_{i \in I} x_i P(X = x_i)$.

L'espérance est une estimation du comportement moyen observé du caractère représenté par X.

Définition. On définit la variance $V(X)$ par : $V(X) = E((X - E(X))^2)$.
On a donc : $V(X) = \sum_{i \in I} (x_i - E(X))^2 P(X = x_i)$.

La variance décrit la dispersion du caractère représenté par X.

Adaptation dans l'enseignement secondaire
Les définitions et les propriétés de l'espérance et de la variance de lois de probabilités discrètes sont travaillées en classe de première S.

b) Lois discrètes usuelles

Loi uniforme discrète

Définition. Soit $n \in \mathbb{N}$. Une variable aléatoire réelle X suit une *loi uniforme* sur $[\![1; n]\!]$ si : $X(\Omega) = [\![1; n]\!]$ et

$$P(X = k) = \frac{1}{n}, k \in [\![1; n]\!].$$

Exemple. Le résultat d'un dé équilibré suit une loi uniforme sur $[\![1; 6]\!]$.

Proposition. *Soit X une variable aléatoire suivant une loi uniforme sur $[\![1; n]\!]$. Alors $E(X) = \frac{n+1}{2}$ et $V(X) = \frac{n^2-1}{12}$.*

Loi de Bernoulli

Définition. Soit X une variable aléatoire réelle et $p \in]0; 1[$. X suit une *loi de Bernoulli* de paramètre p si : $X(\Omega) = \{0; 1\}$ et $P(X = 1) = p$ et $P(X = 0) = 1 - p$.

Remarque. La loi de Bernoulli décrit une expérience aléatoire à deux issues : succès (de probabilité p) et échec (de probabilité $1 - p$).

Exemple. Le résultat du lancer d'une pièce (Pile ou Face) équilibrée suit une loi de Bernoulli de paramètre $1/2$.

Proposition. *Soit X une variable aléatoire suivant une loi de Bernoulli de paramètre p, notée $\mathcal{B}(1, p)$. Alors $E(X) = p$ et $V(X) = p(1 - p)$.*

Loi binomiale

Soit $n \in \mathbb{N}^*$. On répète n fois une expérience de Bernoulli et on note X le nombre de succès obtenus.

Définition. La variable aléatoire X suit une *loi binomiale* de paramètres n et p, notée $\mathcal{B}(n,p)$, c'est-à-dire que $X(\Omega) = [\![0;n]\!]$ et

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, k \in [\![0;n]\!].$$

Proposition. *Soit X une variable aléatoire suivant une loi de binomiale $\mathcal{B}(n,p)$. Alors $E(X) = np$ et $V(X) = np(1-p)$.*

Adaptation dans l'enseignement secondaire
Ces trois lois sont rencontrées au lycée. Les lois de Bernoulli et binomiales sont travaillées à partir de la classe de première et reprises en terminale.

1.3.3 Probabilités à densité

a) Généralités

Définition. Une variable aléatoire réelle X est dite *continue* si $X(\Omega)$ est un intervalle de $\mathbb{R}$.

Définition. X est une variable aléatoire. On dit que X admet une fonction de densité f si sa fonction de répartition est continue et peut s'écrire sous la forme :

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t, x \in \mathbb{R},$$

avec f une fonction positive, qui possède un nombre fini de points de discontinuité et telle que $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$.

Toutes les variables aléatoires continues n'admettent pas de fonction de densité. Une variable aléatoire continue peut posséder une fonction de répartition non dérivable presque partout. Les variables aléatoires discrètes possèdent une fonction de répartition qui n'est pas continue sur $\mathbb{R}$, donc ne peuvent pas admettre de densité.

On suppose maintenant que la variable X admet une fonction de densité f.

Proposition. *En tout point x_0 où f est continue, sa fonction de répartition F est dérivable et $F'(x_0) = f(x_0)$.*

Proposition.
- *Sa fonction de répartition F est la primitive de f qui tend vers 0 en $-\infty$.*
- *Si f est continue, la dérivée de la fonction de répartition F est la fonction f elle-même.*

Dans le cas d'une variable aléatoire à densité, F est continue mais f ne l'est pas nécessairement : elle peut posséder des points de discontinuité. Pour cette raison, f n'est pas unique.

Adaptation dans l'enseignement secondaire
En terminale S, les variables aléatoires continues étudiées admettent toutes une fonction de densité. Les différentes lois de probabilité considérées sont définies par leur fonction de densité et non par leur fonction de répartition. Cela s'explique notamment par le fait que les intégrales impropres ne sont pas au programme en calcul intégral. Les fonctions de densité rencontrées sont continues sur $X(\Omega)$ (qui n'est pas toujours $\mathbb{R}$).

Proposition. *Soit X une variable aléatoire de fonction de répartition F, admettant f pour densité. Alors pour tout a et b réels tels que $a \leqslant b$,*

$$P(a \leqslant X \leqslant b) = \int_a^b f(t)\,\mathrm{d}t.$$

Pour tout x réel, $P(X = a) = 0$.

Remarque. Il s'agit de la relation sémiotique $(*)$ qui nous intéresse tout particulièrement.

Définition. Soit X une variable aléatoire réelle à densité. Sous réserve d'existence, on définit l'espérance $E(X)$ par :

$$E(X) = \int_{-\infty}^{+\infty} tf(t)\,\mathrm{d}t.$$

Adaptation dans l'enseignement secondaire
Cette définition ne peut être donnée comme telle en terminale S du fait que les intégrales impropres ne sont pas au programme. L'espérance est définie au cas par cas pour les différentes lois étudiées. Par exemple, dans le cas d'une variable aléatoire à valeurs sur $\mathbb{R}$, comme pour la loi normale centrée réduite par exemple (cf. plus loin), on ne considère pas $E(X) = \displaystyle\int_{-\infty}^{+\infty} tf(t)\,\mathrm{d}t$ mais

$$E(X) = \lim_{x\to-\infty} \int_x^0 tf(t)\,\mathrm{d}t + \lim_{y\to+\infty} \int_0^y tf(t)\,\mathrm{d}t.$$

C'est le moyen proposé dans le programme pour contourner le problème de l'existence de ce type d'intégrale.

Définition. Soit X une variable aléatoire réelle à densité. Sous réserve d'existence, on définit la variance $V(X)$ par :

$$V(X) = E((X - E(X))^2) = \int_{-\infty}^{+\infty} (t - E(X))^2 f(t)\,\mathrm{d}t.$$

Adaptation dans l'enseignement secondaire
Le calcul de la variance des variables aléatoires à densité considérées en terminale S n'est pas un attendu du programme. Il est seulement mentionné que l'« *on admet que la variance* [d'une loi normale centrée réduite]*, définie par* $E\left((X-E(X))^2\right)$*, vaut 1* ». On peut remarquer que $(X-E(X))^2$ est bien une variable aléatoire, cependant il n'y a pas de raison qu'elle suive une loi uniforme, exponentielle ou normale, qui sont les trois seuls cas pour lesquels est définie l'espérance dans le programme.

b) Lois à densité usuelles

Loi uniforme sur $[a;b]$

Proposition. *Soit* $a<b$ *et* f *la fonction définie sur* $\mathbb{R}$ *par :*

$$f(x)=\begin{cases}\dfrac{1}{b-a} & \text{si } x\in[a;b],\\ 0 & \text{sinon.}\end{cases}$$

Alors f *est une densité de probabilité.*

Définition. Si X est une variable aléatoire dont f est une densité, on dit que X suit une loi uniforme sur $[a;b]$, notée $\mathcal{U}_{[a;b]}$.

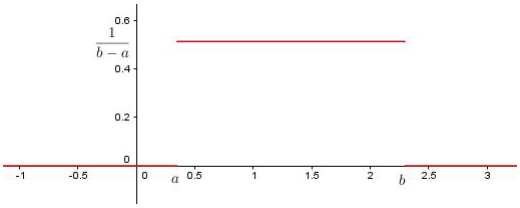


Figure 1.1 – Fonction de densité associée à la loi uniforme sur $[a;b]$

Exemple. Choisir un nombre réel au hasard entre 0 et 1 peut être modélisé par une variable aléatoire X qui suit une loi uniforme sur $[0;1]$.

Proposition. *Soit* X *une variable aléatoire qui suit une loi uniforme sur* $[a;b]$*. Alors pour tous réels* c *et* $d\in[a;b]$ *tels que* $c<d$*,* $P(c\leqslant X\leqslant d)=\dfrac{d-c}{b-a}$.

Proposition. *Soit* X *une variable aléatoire qui suit une loi uniforme sur* $[a;b]$*. Alors* X *admet une espérance :* $E(X)=\dfrac{a+b}{2}$.

Proposition. X *admet une variance :* $V(X)=\dfrac{(b-a)^2}{12}$.

Adaptation dans l'enseignement secondaire
En terminale S, la fonction de densité de la loi uniforme est définie sur $[a;b]$ (dans ce cas, la fonction f est continue sur son ensemble de définition). Le calcul de l'espérance fait partie des démonstrations exigibles au baccalauréat. La variance n'est quant à elle pas au programme.

Loi exponentielle

Proposition. *Soit $\lambda > 0$ et f la fonction définie sur $\mathbb{R}$ par :*

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases}$$

Alors f est une densité de probabilité.

Définition. Si X est une variable aléatoire dont f est une densité, on dit que X suit une *loi exponentielle* de paramètre λ, notée $\mathcal{E}(\lambda)$.

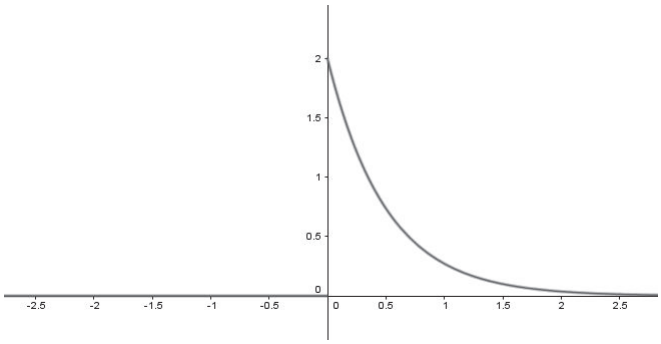


FIGURE 1.2 – Fonction de densité associée à la loi exponentielle de paramètre 2

Proposition. *Soit X une variable aléatoire qui suit une loi exponentielle $\mathcal{E}(\lambda)$, avec $\lambda > 0$. Alors pour tout réel a,*

$$P(X \leqslant a) = 1 - e^{-\lambda a}.$$

Remarque. On en déduit que pour tous réels a et $b \in \mathbb{R}^+$ tels que $a < b$,

$$P(a \leqslant X \leqslant b) = e^{-\lambda a} - e^{-\lambda b}.$$

Proposition. *Soit X une variable aléatoire qui suit une loi exponentielle $\mathcal{E}(\lambda)$, avec $\lambda > 0$. Alors X admet une espérance : $E(X) = \dfrac{1}{\lambda}$.*

Proposition. *(Propriété de loi sans vieillissement). Soit X une variable aléatoire qui suit une loi exponentielle $\mathcal{E}(\lambda)$, avec $\lambda > 0$. Alors*

$$P_{X \geqslant a}(X \geqslant a + h) = P(X \geqslant h), \text{ pour tous réels } a \text{ et } h > 0.$$

Adaptation dans l'enseignement secondaire
De la même manière que pour la loi uniforme, la fonction de densité de la loi exponentielle est définie sur $\mathbb{R}^+$, en terminale S. Toutes ces propriétés sont au programme. Elles sont démontrées en cours et sont exigibles à l'épreuve du baccalauréat. Les intégrales impropres en jeu sont contournées en s'intéressant à la limite de l'intégrale quand la borne supérieure tend vers $+\infty$. La démonstration pour l'espérance nécessite une intégration par parties qui n'est pas un attendu de la classe de terminale S. Des détours sont proposés pour appliquer cette méthode sur cet exemple particulier. Les probabilités conditionnelles (présentes dans la propriété de loi sans vieillissement) sont aussi au programme de terminale S.

Loi normale

Proposition. *Soit $\mu \in \mathbb{R}$ et $\sigma > 0$. Soit f la fonction définie sur $\mathbb{R}$ par :*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

Alors f est une densité de probabilité.

Définition. Si X est une variable aléatoire admettant f pour densité, on dit que X suit une *loi normale* de paramètres μ et σ, notée $\mathcal{N}(\mu, \sigma^2)$.

Proposition. *Soit X une loi normale $\mathcal{N}(\mu, \sigma^2)$, alors X admet une espérance et une variance : $E(X) = \mu$ et $V(X) = \sigma^2$.*

Définition. La loi normale $\mathcal{N}(0, 1)$ est appelée la *loi normale centrée réduite.*

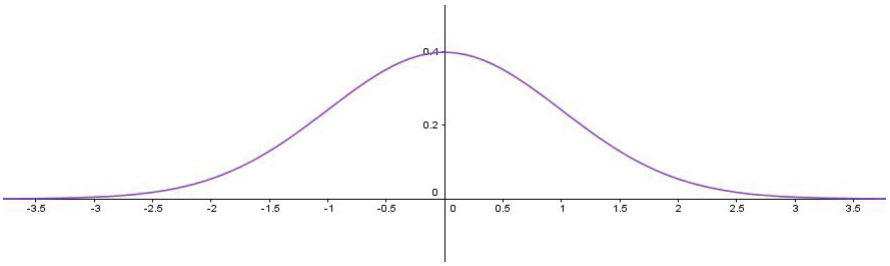


Figure 1.3 – Fonction de densité associée à la loi normale centrée réduite

Adaptation dans l'enseignement secondaire
Dans le programme de terminale S, il est demandé d'introduire, tout d'abord, la loi normale centrée réduite à l'aide du théorème de Moivre Laplace (cf. section 1.3.5), puis de définir les autres lois normales à l'aide de la loi normale centrée réduite comme ceci :

Définition. On dit que la variable aléatoire X suit une loi normale de paramètres μ et σ, si la variable aléatoire $\dfrac{X-\mu}{\sigma}$ suit une loi normale centrée réduite.

Aucun calcul de probabilités n'est attendu des élèves de terminale S pour la loi normale. Ces calculs nécessitent en effet des techniques non disponibles à ce niveau (changements de variables, intégrales doubles,...). Les élèves doivent savoir utiliser la calculatrice pour déterminer les probabilités. Le programme demande que l'influence des paramètres soit connu. Plusieurs propriétés sont à connaître, notamment les probabilités associées aux intervalles centrés en μ à plus ou moins 1, 2 ou 3 σ.[2] Le théorème suivant est aussi à connaître et sa démonstration est exigible :

Théorème. *Soit X une variable aléatoire qui suit une loi normale $\mathcal{N}(0, 1)$. Alors : pour $\alpha \in]0; 1[$, il existe un unique u_α tel que $P(-u_\alpha \leqslant X \leqslant u_\alpha) = 1 - \alpha$.*

La démonstration de ce théorème utilise (implicitement en terminale S) la fonction de répartition.

2. Nous appellerons ces propriétés « 1, 2, 3 sigmas », conformément au vocabulaire utilisé en terminale.

1.3.4 Loi faible des grands nombres

Théorème. *(Loi faible des grands nombres) Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles deux à deux indépendantes et identiquement distribuées, ayant pour espérance commune $E(X)$. Soit $(\overline{X}_n)_{n\in\mathbb{N}^*}$ définie par :*

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \text{ pour tout } n \in \mathbb{N}^*.$$

Alors $(\overline{X}_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers $E(X)$. Autrement dit,

$$\forall \varepsilon > 0, \lim_{n\to+\infty} P\left(\left|\overline{X}_n - E(X)\right| \geqslant \varepsilon\right) = 0.$$

En langage naturel, on peut dire que quel que soit un majorant de l'écart entre la fréquence empirique et la probabilité, la probabilité de cet écart est aussi proche de 1 que l'on veut, pourvu que n soit assez grand. On peut déduire de la loi des grands nombres que la fréquence empirique d'apparition d'un caractère « converge » vers la probabilité associée.

Adaptation dans l'enseignement secondaire
En pratique, la loi des grands nombres est utilisée à partir de la classe de troisième. Elle permet notamment de valider l'approche « fréquentiste » des probabilités. Cependant, ce théorème n'est évidemment jamais explicité.

1.3.5 Théorème limite central

Théorème. *(Théorème limite central). Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles deux à deux indépendantes et identiquement distribuées, ayant pour espérance commune $E(X)$ et pour écart-type $\sigma(X)$. Alors*

$$\frac{\sum_{i=1}^{n} X_i - nE(X)}{\sigma(X)\sqrt{n}} \xrightarrow[n\to+\infty]{\mathcal{L}} Y, \text{ où } Y \sim \mathcal{N}(0,1).$$

Autrement dit,

$$\forall a < b, \lim_{n\to+\infty} P\left(a \leqslant \frac{\sum_{i=1}^{n} X_i - nE(X)}{\sigma(X)\sqrt{n}} \leqslant b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

Théorème. *(Théorème de Moivre Laplace). Soit $p \in]0;1[$, soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles qui suivent une loi $\mathcal{B}(n,p)$. Alors*

$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow[n\to+\infty]{\mathcal{L}} X, \text{ où } X \sim \mathcal{N}(0,1).$$

Autrement dit, pour tous réels a et b tels que $a < b$,

$$\lim_{n\to+\infty} P\left(a \leqslant \frac{X_n - np}{\sqrt{np(1-p)}} \leqslant b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}\,\mathrm{d}t.$$

Le théorème de Moivre Laplace est une conséquence du théorème limite central en prenant une suite de variables aléatoires de loi $\mathcal{B}(1, p)$. Cependant, il peut être prouvé directement.

Remarque. En principe, on considère que l'on peut approcher une loi binomiale $\mathcal{B}(n, p)$ par une loi normale $\mathcal{N}(np, npq)$ si $n > 30$, $np > 5$ et $n(1 - p) > 5$.

Adaptation dans l'enseignement secondaire
Dans le programme de terminale S, il est demandé d'introduire la loi normale centrée réduite par l'observation du théorème de Moivre Laplace. Bien entendu, ce théorème est admis. Il doit seulement être « visualisé ». Il sert ensuite dans la partie *Intervalle de fluctuation* du programme pour déterminer des intervalles de fluctuation.

Chapitre 2

Problématique, cadre théorique et méthodologie générale

Notre questionnement de départ se place autour du macro-signe

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x, \qquad (*)$$

qui met en relation les probabilités avec l'analyse, en terminale S. Dans le chapitre précédent, nous avons présenté la place grandissante des probabilités dans l'enseignement secondaire actuel en France, notamment au travers l'évolution des programmes depuis les années 1970. Ensuite, les rappels mathématiques accompagnés de commentaires sur l'adaptation des notions au lycée ont eu pour but d'entrevoir superficiellement la transposition didactique existant dans les classes. Dans ce chapitre, nous commencerons par préciser notre problématique et nos questions de recherche (section 2.1). Dans la section 2.2, nous proposerons un état des lieux de la recherche en didactique sur les probabilités continues. Ensuite, dans la section 2.3, nous présenterons le cadre théorique dans lequel nous allons nous situer. Enfin, nous terminerons par la présentation de la méthodologie générale que nous avons mise en place pour répondre à nos questions de recherche (section 2.4).

2.1 Problématique

2.1.1 Des polémiques

Nous avons vu, à travers les évolutions des programmes au cours des dernières décennies, une augmentation importante de la place des probabilités et de la statistique dans l'enseignement secondaire, et en particulier l'apparition des probabilités continues à partir de 2002, qui s'est intensifiée depuis 2012. Ces évolutions de programme ont engendré des polémiques au sein des mathématiciens et des enseignants. Colmez (2012), directeur de recherche au CNRS, a répondu « *Oui* » à la question « *Faut-il arrêter d'enseigner les statistiques au lycée ?* ». Il considère que l'introduction des statistiques mais aussi des probabilités) a « *largement contribué à l'appauvrissement de l'enseignement des mathématiques au lycée* ». Tout comme Perrin (2015), il ne nie pas l'importance du rôle de la statistique et des probabilités dans des domaines

variés, mais reproche le fait que « *la totalité des énoncés portant sur les statistiques et les probabilités sont admis, ce qui transforme le cours de mathématiques en un cours de mauvaise magie* ». Raoult et Arnoux (2013) répondent à cet article en mettant en avant l'apport que peut retirer l'ensemble de l'enseignement des mathématiques de l'enseignement des probabilités et de la statistique. Notamment, ils mettent en avant l'attractivité des ces deux domaines mathématiques pour stimuler les élèves. Ils regrettent l'absence des probabilités continues et de la statistique en classe préparatoire scientifique, où ils disent qu'il pourrait y « *avoir profit à les associer directement aux chapitres de mathématiques qu'elles peuvent mettre en valeur* ». Notamment, ils donnent l'exemple du calcul intégral qui pourrait « *prendre vie avec les exercices sur les lois à densité* », en classe préparatoire. Perrin (2015) est d'avis de « *renvoyer à l'université (ou aux classes préparatoires) l'étude de la loi normale, de se concentrer au lycée sur les probabilités discrètes [...] et d'utiliser ces connaissances de probabilités pour faire aussi de la statistique* » (p. 63).

Il n'est pas question ici de rentrer dans cette polémique et de remettre en question le programme actuel. Ce qui est sûr, c'est que les liens entre les probabilités et l'analyse ressortent et semblent mériter d'être travaillés. Nous considérons que ces liens peuvent être exploités dès la classe de terminale, bien que certaines « incohérences » soient repérables dans le programme. Nous pouvons mentionner l'absence des intégrales impropres dans le programme de terminale S, indispensables en théorie pour l'étude de la loi normale et de la loi exponentielle. Le problème est toutefois contourné. En outre, que ce soit au lycée ou plus tard, la question de la mise en relation des probabilités et de l'analyse subsiste.

Dans cette thèse, nous nous limiterons à l'enseignement au niveau de la terminale S pour les raisons que nous avons évoquées dans l'introduction générale, dans le cadre du programme actuel.

2.1.2 Les liens entre lois de probabilités à densité et calcul intégral

Comme nous l'avons vu, le domaine mathématique actuel des probabilités ne peut se passer du domaine de l'analyse. Dans l'enseignement secondaire, l'étude de probabilités discrètes est faite sans recours à l'analyse, mais quand arrivent les lois à densité, en terminale, le lien avec l'intégrale est nécessaire.

Dans le programme de 2001 (MEN, 2001), il est précisément mentionné que l'étude des lois à densité est « *une application de ce qui aura été fait en début d'année sur l'exponentielle et le calcul intégral* ». Donc l'existence d'un lien avec l'analyse est mentionné : les probabilités sont une application du calcul intégral. Dans le programme de 2011 (MEN, 2011), aucun commentaire n'est fait sur les liens entre calcul intégral et lois à densité. Cependant, dans les trois pages qui sont consacrées aux lois à densité, le signe $\int$ apparait plusieurs fois notamment dans les définitions de l'espérance des différentes lois. Quoi qu'il en soit les liens entre les deux domaines restent.

Si nous nous penchons sur les définitions données dans le programme de terminale scientifique (MEN, 2011), le macro-signe ($*$) n'est pas présent. Cependant, la relation sémiotique, donc ce que porte ce signe au-delà du signe lui-même, est présentée sous

une autre forme. D'une part, en calcul intégral, on définit « *l'intégrale d'une fonction continue et positive sur* $[a;b]$ *comme aire sous la courbe* » (p. 7). D'autre part, au niveau des probabilités continues,

> On définit [...] une variable aléatoire X, fonction de Ω dans $\mathbb{R}$, qui associe à chaque issue un nombre réel d'un intervalle I de $\mathbb{R}$.[3] On admet que X satisfait aux conditions qui permettent de définir la probabilité de l'événement $X \in J$ comme aire du domaine : $\{M(x;y); x \in J \text{ et } 0 \leqslant y \leqslant f(x)\}$ où f désigne la fonction de densité de la loi et J un intervalle inclus dans I. (p. 12).

On voit donc apparaître une nouvelle notion, l'aire sous la courbe. Nous avons donc trois objets mathématique en jeu : l'intégrale, l'aire sous une courbe et la probabilité (cf. figure 2.1).

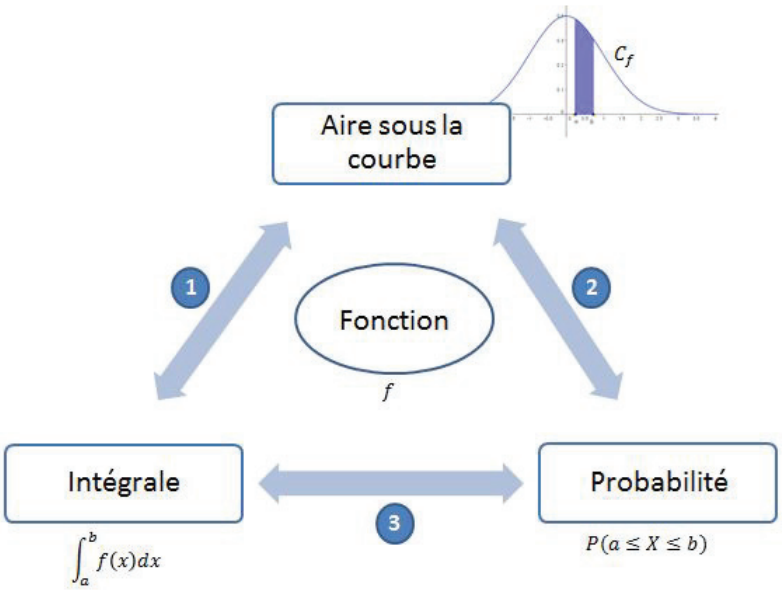


Figure 2.1 – Relations entre intégrale, aire et probabilité

Cependant, les relations entre les trois, deux à deux, ne sont pas de même nature. Entre intégrale et aire sous la courbe (flèche 1), le lien est direct : l'intégrale est définie comme aire sous la courbe dans le programme. En revanche, entre probabilité et aire sous la courbe (flèche 2), le lien est indirect et se fait par l'intermédiaire d'une fonction, la fonction de densité de probabilité. Enfin, entre probabilité et intégrale (flèche 3), le lien se fait de la même manière qu'entre probabilité et aire sous la courbe. On voit donc apparaître un objet central dans les différentes relations, qui est la fonction. Cependant, la fonction f fait partie intégrante de l'intégrale (intégrale de la fonction f) et de l'aire sous la courbe (courbe représentative de la fonction f). Ce n'est pas le cas pour la probabilité. Comment passe-t-on d'une variable aléatoire X à une fonction f ? On voit donc le rôle central de la fonction de densité dans la relation entre probabilités et analyse. Effectivement, vu la définition de l'intégrale, il est assez naturel de rassembler intégrale et aire sous la courbe et de ne considérer plus que la relation entre ce bloc et la probabilité (cf. figure 2.2).

3. Cela sous-entend $X(\Omega) = I$.

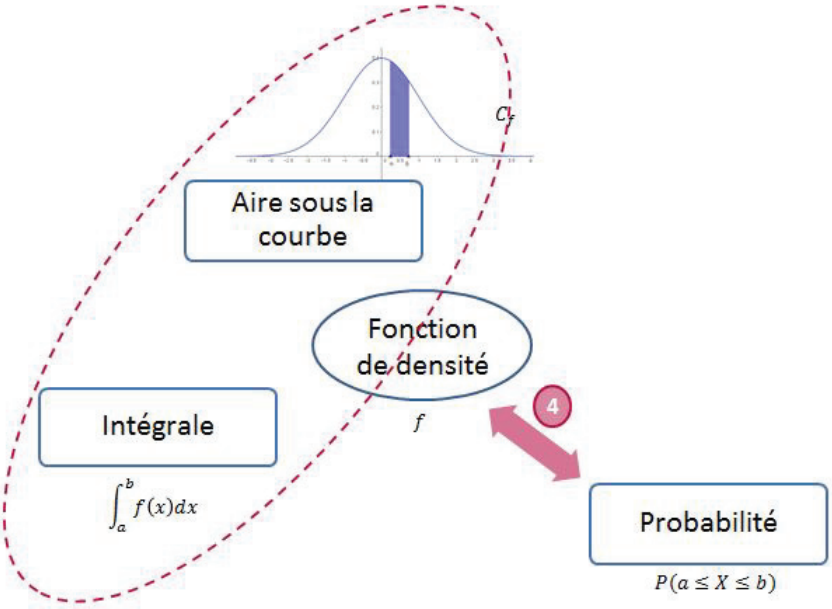


FIGURE 2.2 – Relation entre calcul intégral et probabilités

On peut donc voir la fonction de densité comme l'objet intermédiaire pour faire la mise en relation entre les probabilités et l'analyse (flèche 4). Cette fonction a un rôle central dans la relation $P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x$.

Ces réflexions nous permettent d'orienter notre réflexion en nous focalisant sur l'objet mathématique fonction de densité. Comprendre cette notion, son origine, son introduction et sa place dans les classes de terminale S, nous permettra de comprendre la relation sémiotique $(*)$ installée.

Certaines réflexions de groupes IREM laissent penser que la place accordée dans les classes à la fonction de densité n'est pas satisfaisante. Nous citerons ce paragraphe :

> Il ne viendrait à personne l'idée d'émettre la moindre réserve quant à la notion de densité. Chacun connaît sa longue carrière en Probabilités, même sous le point de vue le plus élémentaire : calculs de moments (espérance, variance,...) par exemple, ou encore convolution de Haar pour obtenir la densité de la somme de deux variables aléatoires indépendantes, etc...
> Il n'en reste pas moins une évidence difficile à réfuter : en dehors de la définition d'une loi continue « imposée » par le programme et de quelques exercices volontairement ciblés, la notion de densité n'est guère prégnante en Terminale S, sauf à considérer, par exemple, l'approximation continue d'un phénomène discret, la simulation discrétisée d'un phénomène continu... (Barbazo *et al.*, 2003, pp. 101-102)

Peut-être n'est-il pas si évident de faire que la fonction de densité soit « prégnante » en terminale ? Ces réflexions nourrissent nos questions de recherche que nous allons maintenant présenter.

2.1.3 Vers nos questions de recherche

La question de départ, présentée dans l'introduction, est la suivante :

Est-il possible de concevoir et de mettre en œuvre des problèmes d'introduction aux lois à densité qui permettent aux élèves de terminale scientifique de construire conceptuellement la relation sémiotique $(*)$?

Nos premières réflexions recentrent la problématique sur la notion de fonction de densité, qui joue un rôle central dans la relation sémiotique $(*)$. Des premières questions naïves surgissent. D'où vient cette fonction ? A quoi correspond-t-elle ? D'où vient le mot densité ? Ces questions semblent importantes avant même d'aborder la question de l'introduction de la notion dans les classes.

Voici les trois questions de recherche qui vont guider notre travail :

1. Comment historiquement et épistémologiquement sont apparus les liens entre les probabilités et le calcul intégral ? Comment a émergé le concept de lois à densité et tout particulièrement la notion de fonction de densité ?
2. En terminale S, quels sont les enjeux didactiques derrière la relation sémiotique $(*)$? Comment est introduite la notion de fonction de densité ? Puis comment est exploitée la relation $(*)$?
3. Quels éléments didactiques peut-on prendre en compte pour concevoir et mettre en œuvre des problèmes permettant de construire la notion de fonction de densité ?

Ces questions seront reformulées dans les termes de notre cadre théorique dans la section 2.4.

2.2 Les travaux en didactique portant sur les lois à densité

Nos questions de recherche posées nous amènent à se demander si des travaux ont déjà été menés en didactique des probabilités sur la notion de fonction de densité et sur les liens entre probabilités et analyse.

Malgré une recherche bibliographique importante, très peu de travaux ont été trouvés sur l'enseignement et l'apprentissage des lois à densité. La didactique des probabilités commence à trouver sa place, cependant pour l'instant, les problématiques étudiées portent, la plupart du temps, sur les probabilités discrètes (sûrement est-ce dû au fait que l'enseignement des lois à densité était jusque là très peu présent dans le secondaire en France et à l'étranger).

Nous allons présenter les différents travaux trouvés en élargissant au fur et à mesure sur les thèmes abordés. Nous resterons toujours dans l'optique de l'enseignement et l'apprentissage des lois de probabilités à densité.

Avant tout, nous voulons mentionner que de nombreux travaux en didactique des mathématiques existent sur le calcul intégral. Nous ne citons ici que trois thèses de doctorat en français :

- celle de Schneider (1988) qui s'intéresse notamment au lien entre aire sous la courbe et primitive,
- celle de Tran Luong (2006) sur une comparaison de l'enseignement de la notion d'intégrale au lycée en France et au Vietnam,
- celle d'Haddad (2012) sur cet enseignement en Tunisie.

Dans ces différentes recherches et dans d'autres, nous n'avons trouvé aucune mention aux probabilités. Ceci explique que nous allons nous centrer exclusivement sur les travaux en probabilités.

2.2.1 Travaux sur les lois à densité

Nous allons présenter ici les travaux de recherche en didactique des mathématiques que nous avons trouvés spécifiquement centrés sur les lois à densité.

Nous commençons par les recherches en France. Tout d'abord, nous citons l'article d'Henry (2003), *Des lois de probabilité continues en terminale S, pourquoi et pour quoi faire ?*, dans lequel il s'interroge sur la pertinence de l'enseignement des lois à densité en terminale S, un an après la réforme de 2002. Il s'intéresse donc à la loi uniforme sur $[0; 1]$, aux lois exponentielles, mais aussi (bien que pas encore au programme) à la loi normale. Il se pose la question de la légitimité de prolonger les lois discrètes à des lois sur un ensemble continu. Il s'interroge aussi sur la notion de densité :

> La probabilité d'un événement a été définie comme la somme des probabilités des événements élémentaires le constituant. Transposées au cas continu, ces probabilités sont nécessairement toutes nulles [...]. Comment donner un sens probabiliste à la notion de densité ? (p. 6)

Cette question nous intéresse tout particulièrement.

Au sujet de la réforme, il écrit :

> On voit que le pari d'introduire des lois continues en terminale est pour le moins audacieux. Le point de vue de la modélisation est ici incontournable. Les lois continues sont alors des outils mathématiques permettant, dans les modèles considérés, de calculer des probabilités de certains événements. (p. 6)

Il met en avant l'importance de la modélisation. Il insiste fortement dans l'article sur le statut de modèle des lois de probabilité, qui lui semble bien mis en valeur dans le programme. Il distingue la statistique, qui décrit la réalité, des probabilités, qui sont une modélisation de la réalité.

Il propose plusieurs exemples d'introduction de lois continues en passant du discret au continu, à partir de modélisations. Pour la loi uniforme continue, il précise que l'intégrale n'est pas nécessaire et du coup que la fonction de densité non plus. Pour la loi exponentielle, il explique que la discrétisation et le passage par la loi géométrique, comme il le propose dans l'article, permettent de comprendre pourquoi la densité de probabilité en jeu est de type exponentiel, mais que cette approche est hors de portée pour des élèves de terminale. Il propose donc de présenter directement aux élèves le résultat : $P(a < X < b) = \int_a^b \alpha e^{-\alpha x}\,\mathrm{d}x$. Cela revient donc à imposer aux élèves la relation sémiotique qui lie probabilités et analyse. Il précise aussi que certains exemples, comme celui de la radioactivité, permettent une bonne illustration de cette loi, en mettant en œuvre les connaissances acquises en analyse.

Il conclut en disant qu'après la terminale S, les exemples de la loi uniforme ou de la loi exponentielle permettront de dégager la notion « *fondamentale* » de densité de probabilité, qui donnera lieu à une utilisation intensive du calcul intégral et plus généralement des outils de l'analyse.

De plus, il pense que l'introduction des lois continues en terminale est source de problèmes intéressants (pour lui, « *plus intéressants que les problèmes traditionnels de combinatoire à propos de jeux de hasard* ») et « *a pour importance essentielle de rendre incontournable la notion de modèle* ».

Dantal et Henry (2003) s'intéressent à l'utilisation de l'analyse en probabilités, notamment pour illustrer le commentaire du programme de l'époque (MEN, 2001) indiquant que les lois à densité sont une « *application de ce qui aura été fait en début d'année sur l'exponentielle et le calcul intégral* ». Ils pensent que « *l'étude de certaines lois de probabilité et de leurs caractéristiques numériques permet de mettre en œuvre des connaissances essentielles des cours d'analyse et de probabilités, contribuant ainsi par leur utilisation à leur donner encore plus de sens* » (p. 132). Ils proposent quatre exemples qu'ils précisent « *en marge* » du programme, dont un sur la loi uniforme continue et un sur la loi exponentielle. Ils font notamment le choix d'introduire les lois par leur fonction de répartition, ce qui permet ensuite de décliner des questions exploitant le lien entre probabilités et calcul intégral (lien entre fonction de répartition et primitive, lien entre calcul d'espérance et de variance et calcul intégral...). Il est plutôt question ici d'exploiter la relation sémiotique qui lie les deux domaines mathématiques. Cet article met bien en avant l'intérêt que nous avons de nous pencher sur la notion de fonction de densité et sur les connexions entre probabilités et analyse.

Ce sont les seuls travaux de recherche à notre connaissance portant sur ces notions, en France. Nous pouvons ensuite citer quelques travaux de groupes IREM travaillant sur les probabilités, qui proposent des brochures notamment pour former les enseignants suite à la réforme de 2002 (Barbazo *et al.*, 2003 ; Dutarte *et al.*, 2003 ; IREM de Besançon, Groupe Probabilités et Statistique, 2007). Nous y reviendrons dans la section suivante.

Dans les travaux internationaux, les recherches sont centrées sur la loi normale. Des chercheurs ont travaillé sur la compréhension du théorème limite central par les étudiants et repéré des idées fausses au niveau de la normalité des distributions d'échantillonnage (cf. Vallecillos, 1999 ; delMas, Garfiels & Chance, 1999). Wilensly (1995, 1997) a fait une étude sur le comportement des étudiants résolvant des problèmes se rapportant à la loi normale. Il observe que les étudiants sont tout à fait capables de résoudre les différents problèmes mais sont dans l'incapacité d'expliquer pourquoi ils choisissent la loi normale plutôt qu'une autre. Wilensky parle alors d'anxiété épistémologique, qu'il définit comme étant un sentiment de confusion et d'indécision chez les étudiants lorsqu'ils sont face à des chemins différents pour résoudre un problème.

Nous avons repéré d'autres recherches sur l'enseignement et l'apprentissage de la loi normale. Nous pouvons citer les travaux de Batanero, Tauber et Meyer (1999), de

Batanero, Tauber et Sanchez (2004), de Batanero, Godino et Roa (2004), de Pfannkuch et Reading (2006) ainsi que de Bansilal (2012, 2014) qui se sont intéressés à la compréhension de la loi normale par les étudiants à l'université ou chez les futurs enseignants.

Batanero, Tauber *et al.* (2004) expliquent l'importance de travailler sur les représentations graphiques pour aider les étudiants à la compréhension de la loi normale. Ils insistent : « *it is important that students understand basic concepts such as probability, density curve* [...] *and histograms before they start the study of normal distribution* » (op. cit., p. 275). Les principales difficultés qu'ils ont repérées chez les étudiants sont les suivantes :

- difficultés de percevoir l'intérêt d'un modèle pour décrire des données empiriques ;
- difficultés d'interpréter des aires d'histogrammes de fréquences et de calculer des aires dans les cas où un changement dans les extrémités des intervalles est nécessaire ;
- difficultés d'interpréter les probabilités sous la courbe normale. Les auteurs pensent que la représentation graphique des aires sous la courbe de la loi normale est le principal outil didactique pour les étudiants pour qu'ils comprennent le calcul de probabilités et en même temps pour résoudre différents problèmes mettant en jeu la loi normale ;
- difficultés de faire un choix entre des données empiriques et des modèles mathématiques, d'interpréter des résumés statistiques et des graphiques ;
- difficultés de savoir quand une variable discrète peut ou ne peut pas être modélisée par une loi normale.

On retrouve ici des éléments d'autres recherches. Ces travaux mettent aussi en avant la difficulté de la modélisation car les étudiants doivent à la fois traiter des données empiriques et un modèle mathématique. Ils concluent en évoquant le fait que la loi normale est une idée complexe qui requiert l'intégration de beaucoup d'idées et de concepts statistiques différents. Mais ils pensent possible de concevoir des activités, notamment des activités liant données réelles et modèles, qui facilitent l'apprentissage des notions basiques sur la loi normale. Pour eux, travailler avec l'outil informatique semble améliorer la compréhension graphique et donc la compréhension de la loi normale. Ces articles ne font pas de propositions de situations, mais ils donnent des éléments qui leur semblent importants pour travailler la loi normale (et sûrement d'autres lois), à savoir de travailler sur des données réelles et d'utiliser des logiciels.

Bansilal (2014) montre que les futurs professeurs de mathématiques du secondaire, en formation initiale, en Afrique du Sud, rencontrent des difficultés dans l'utilisation des propriétés de la loi normale centrée réduite mais aussi pour traduire l'expression de la probabilité en tant qu'aire sous la courbe représentant la distribution. Elle propose que les enseignants aient plus de temps pour travailler sur des phénomènes concrets plutôt que de ne travailler que le formalisme. Nous retrouvons la même idée que précédemment, travailler sur des données réelles.

Dans les travaux cités, nous pouvons repérer à l'étranger des travaux centrés sur

la loi normale, dans lesquels sont présentées des difficultés dans la compréhension non pas du formalisme mais de ce que décrit le modèle en question. Aucuns travaux n'ont été trouvés sur d'autres lois même si nous pouvons voir chez Batanero, Tauber *et al.* (2004) des résultats qui peuvent s'ouvrir à d'autres lois. La modélisation est centrale dans les articles.

2.2.2 Des brochures pour la formation des enseignants

Dans les travaux de recherche sur l'enseignement et l'apprentissage des lois à densité que nous avons consultés, nous n'avons trouvé aucune proposition d'enseignement, type ingénierie, sur l'introduction de la notion de densité ou encore sur la construction de la relation entre les probabilités et l'analyse. En étendant nos recherches aux documents tels que les brochures IREM, nous avons pu repérer quelques idées, bien que non testées du point de vue de la recherche. Nous souhaitons en citer quelques unes ici.

Dans une brochure du groupe Probabilités et Statistique de l'IREM de Besançon (2007), il est proposé une approche expérimentale du modèle mathématique de la loi uniforme continue par la simulation et la représentation sous forme d'histogramme pour ensuite introduire la densité de probabilité de la manière suivante :

> Si nous augmentons le nombre d'observations d'une variable statistique continue et si on réduit l'amplitude des classes de l'histogramme, les polygones des fréquences qu'on construit peuvent être approchés par une courbe continue située au-dessus de l'axe des abscisses, l'aire sous la courbe étant égale à 1. (p. 17)

L'introduction de la fonction de densité à partir de la loi uniforme ne semble pas pertinente, comme le précise les auteurs de *L'esprit des lois continues* :

> C'est en vain que l'on chercherait à cet endroit [dans l'introduction des lois uniformes continues] un quelconque intérêt à introduire la notion de densité, tant du point de vue scientifique que pédagogique, sauf peut-être pour suivre le libellé du Programme. (Barbazo *et al.*, 2003, p. 41)

ce qui nous rappelle Henry (2003) :

> Il n'est donc nul besoin pour le modèle uniforme de recourir à des calculs d'intégrales, et la notion de densité ne semble pas s'imposer. (p. 24)

La non nécessité de la fonction de densité pour résoudre des problèmes faisant intervenir des lois uniformes continues ne permet pas de montrer la pertinence de cet objet mathématique. Dans cette brochure, il est proposé une démarche qui conduit à « *explorer la notion de densité à travers quelques exemples de phénomènes continus* ». Il est alors présenté un problème de somme de deux nombres pris au hasard dans $[0; 1]$.

Nous avons ensuite étendu notre bibliographie en prenant en compte des articles sur les probabilités, en général. Nous n'avons pris en compte que des éléments pouvant nous servir dans notre recherche.

2.2.3 Quelques difficultés observées chez les élèves en probabilités

Dans différents travaux ressortent des difficultés des élèves en probabilités. Nous ne citerons que les difficultés pouvant se retrouver dans l'apprentissage des lois à densité. En France, Girard (2003) met en évidence que l'introduction de la notion de probabilité est un obstacle épistémologique, car il faut voir son caractère « approximatif » : la probabilité peut être vue comme une mesure (pas au sens de la théorie de la mesure ici) avec une certaine approximation, comme lorsque l'on mesure une longueur avec une règle par exemple. Il explique aussi que les notions mathématiques en jeu, telles que les pourcentages ou le vocabulaire ensembliste, et plus largement le langage utilisé (par exemple « au moins un », « moins de un »,...) peuvent être sources de difficultés chez les élèves. Il expose aussi la difficulté des élèves à admettre qu'un événement de probabilité nulle puisse tout de même se produire. Dans les recherches internationales, Burrill et Biehler (2011) mentionnent que les étudiants qui travaillent les probabilités avec une approche formelle, apprennent les formalismes sans comprendre les phénomènes qui sont décrits par ces notions mathématiques. Nous retrouvons ici un point soulevé par Batanero, Tauber *et al.* (2004).

2.2.4 Le manque de formation des enseignants en probabilités et statistique

Gattuso et Ottaviani (2011) expliquent que la plupart des enseignants ne sont pas formés à l'enseignement de la statistique et des probabilités. Ils disposent de peu de connaissances statistiques et n'ont pas du tout de formation sur l'enseignement de ces notions. En France, nous retrouvons cette même idée (Régnier, 2002).

Wessels et Nieuwoudt (2013) précisent que les enseignants peuvent suivre des cours de statistique pendant leur formation, mais que ces derniers sont procéduraux plutôt que conceptuels. Aucune application pratique des connaissances statistiques n'est attendue des étudiants. En France, on peut retrouver le même type de résultat, avec des enseignants qui ont étudié les probabilités et la statistique seulement à l'université, dans le cadre de la théorie de la mesure (Girard, 2003). Girard (*ibid.*) précise aussi que les enseignants, notamment ceux ayant reçu leur formation initiale exclusivement dans le contexte de la théorie de la mesure, peuvent avoir une vision négative des probabilités dans le secondaire, ne les considérant pas comme un domaine des mathématiques à part entière, car non rigoureux du fait que le hasard est quelque chose de mal défini.

Ces différents travaux mettent en avant le besoin de formation des enseignants (cf. Bansilal, 2014). Pour aller dans ce sens, en France, suite à la réforme en vigueur en 2002 introduisant les lois à densité dans le programme de terminale S, quelques groupes IREM se sont attelés à l'écriture de brochures citées plus haut (Barbazo *et al.*, 2003 ; IREM de Besançon, Groupe Probabilités et Statistique, 2007) pour essayer de former les enseignants aux lois continues. Enfin, suite à la réforme de 2012, le groupe IREM de l'Université Paris Diderot par exemple a ouvert une formation pour les enseignants, car cette évolution des programmes a nécessité chez beaucoup d'enseignants une indispensable mise à jour des connaissances dans ce domaine (qui

n'a pas toujours été facile et très bien vécue).

2.2.5 Le lien entre statistique et probabilités ou entre réalité et modèle

Dans les différents travaux français et internationaux ressort énormément le lien entre statistique et probabilités ou entre réalité et modèle. Nous en avons déjà un peu parlé dans les travaux sur les lois à densité.

En France, un livre est entièrement consacré à la modélisation en probabilités (Henry, 2001a). Dans cet ouvrage sont abordées la notion de modèle et la question de la modélisation dans l'enseignement secondaire.

Reading et Canada (2011) pensent que l'étude des distributions de données, bien que complexe, est fondamentale pour la compréhension du raisonnement statistique et probabiliste. Pour eux, une distribution peut être définie comme « *the arrangement of values of a variable along the scale of measurement resultingin a représentation of the observed frequencies or the theoritical probability of a range of variables of the variable*[4] » (Reading & Canada, 2011, p. 224).

Des chercheurs (Cohen & Chechile, 1997 ; Reading & Canada, 2011) pointent l'importance de bien distinguer la différence entre la distribution statistique et la distribution probabiliste. La première est empirique, visible directement à partir des données, par le biais des fréquences ; tandis que la seconde (la loi de probabilité) est un modèle théorique formel utilisé pour décrire la probabilité qu'une variable prenne une valeur isolée ou dans un intervalle.

Un lien fort est repéré entre probabilités et statistique à travers la modélisation et il semble incontournable d'identifier les ressemblances mais aussi les contrastes entre les deux pour enseigner ces domaines des mathématiques et développer l'idée de l'aléatoire (Cohen & Chechile, 1997).

Le numéro spécial de novembre 2006 du *Statistics Education Research Journal* s'est centré sur le raisonnement sur les distributions. Dans son article, Wild (2006) s'efforce de bien distinguer distributions empirique et théorique. Il donne l'exemple de l'expression « *les poids sont normalement distribués* » qui n'a de sens que si on l'interprète par le fait que l'on peut raisonnablement approcher la distribution empirique du poids par une courbe gaussienne. Il insiste en expliquant que les distributions sont conditionnelles : si les conditions changent, les variations vont elles aussi changer. Il fait aussi un parallèle entre la distribution d'un échantillon, qui se ramène à une distribution empirique, et la distribution d'une population, dont on n'a pas accès directement et qui se ramène donc à une distribution théorique.

2.3 Principaux choix théoriques

Nos questions de recherche sont orientées du côté du savoir, mais plus particulièrement sur l'enseignement et l'apprentissage d'un certain contenu mathématique,

4. Traduction par nos soins : « *la répartition des valeurs d'une variable dans une échelle de mesure résultant d'une représentation des fréquences observées ou de la probabilité théorique dans un domaine de variations de la variable.* »

dans le contexte de l'enseignement en France et en terminale S. Notre questionnement étant d'ordre didactique, nous abordons le savoir en relation étroite avec les acteurs en interaction avec ce savoir, c'est-à-dire l'enseignant et les élèves. Comme dans les théories de l'activité, nous choisissons d'aborder les apprentissages des élèves et l'enseignement correspondant par l'intermédiaire des tâches qui leur sont prescrites et des activités mathématiques qui en découlent.

Pour étudier notre problématique, nous allons nous appuyer sur le modèle des Espaces de Travail Mathématique, introduit par Kuzniak (2011) que nous compléterons avec des éléments de la théorie de l'activité adaptés dans le cadre de la didactique des mathématiques par Robert et Rogalski (dans Vandebrouck, 2008). Nous verrons la complémentarité de ces deux cadres. Dans cette section, nous allons présenter les différents éléments théoriques qui vont nous être utiles pour préciser nos questions.

2.3.1 Le modèle des Espaces de Travail Mathématique : cadre qui guide notre étude

Comme le précise Kuzniak (2011), la mise en activité de l'élève est de plus en plus importante dans l'enseignement des mathématiques. En réaction à la réforme des mathématiques modernes et après une remise en question de l'enseignement « classique », à partir des années 70, l'enseignement des mathématiques a pour ambition de modifier l'activité attendue des élèves. Un désir de changer le rôle de l'élève dans son apprentissage semble apparaître. L'élève n'est plus là pour « absorber » les connaissances que lui apporte son enseignant, mais il est invité à les construire et les organiser. On peut alors envisager un parallèle entre l'élève en classe de mathématique et le chercheur en mathématiques : l'élève peut être vu comme un apprenti mathématicien. D'après Kuzniak (2011),

> l'insistance sur l'activité de l'élève [dans l'enseignement des mathématiques] a mis en relief deux aspects à la fois complémentaires et très différents du travail de l'élève. D'une part, il s'agit de considérer son travail d'apprenant dans le cadre scolaire avec des professeurs, des devoirs, des évaluations en contexte social. D'autre part, il importe de s'assurer qu'un travail de nature mathématique est effectivement produit par l'élève. (p. 9)

C'est le deuxième point qu'il considère dans son modèle des Espaces de Travail Mathématique. Comme nous venons de le préciser, c'est le travail mathématique qui est au cœur du modèle des Espaces de Travail Mathématique. Nous allons donc d'abord définir ce qu'est le travail mathématique tel que l'entend Kuzniak (2011), puis nous présenterons les éléments caractéristiques d'un Espace de Travail Mathématique.

a) Le travail mathématique

Le travail mathématique se réfère à l'activité des mathématiciens, « *ces êtres humains qui font avancer la compréhension humaine des mathématiques* » (Thurston, 1995, p. 29). Giaquinto (2005) distingue plusieurs phases dans l'activité globale du mathématicien : la découverte, l'explication, la justification et les applications. Kuzniak (2011) explique que dans la conception classique de l'enseignement, l'élève est essentiellement concerné par le travail d'appropriation (applications), ce qui est

restrictif de l'activité du mathématicien, mais il précise que l'on peut agrandir les perspectives si l'on souhaite rendre l'élève plus actif dans la construction de son savoir. Une place essentielle dans le travail des mathématiciens, et aussi dans l'enseignement des mathématiques, est donnée à la résolution de problèmes. Par les problèmes, les élèves vont mettre en œuvre des savoirs et des techniques. Ils vont déclencher le travail mathématique. Il semble donc essentiel d'observer les tâches demandées qui permettent à la fois de décrire le travail d'un point de vue mathématique mais aussi d'un point de vue didactique. Kuzniak envisage le terme « tâche » ici suivant l'approche de la théorie de l'activité (Rogalski, 2008) ou bien celle des praxéologies (Bosch & Chevallard, 1999).

Les mathématiques enseignées ne se limitent pas à un corpus de propriétés et d'objets réduits à des signifiants qui se manipulent à l'aide de systèmes formels. Le travail mathématique est avant tout une activité humaine. Nous voyons ici le rapprochement qui peut alors être fait avec les théories de l'activité. Les questions qui peuvent ensuite intéresser le chercheur dans ce cadre sont de comprendre comment les élèves utilisent et s'approprient les connaissances mathématiques dans leur pratique de la discipline, en restant dans un cadre didactique.

b) Le modèle des Espaces de Travail Mathématique

La notion d'Espace de Travail Mathématique, que nous noterons ETM dans la suite, est une extension de la notion d'Espace de Travail Géométrique (Houdement & Kuzniak, 2006). Nous pouvons penser un ETM comme un environnement organisé pour permettre le travail des individus résolvant des problèmes (pour nous, il s'agira d'élèves). Nous pouvons faire un parallèle avec les architectes, qui définissent les espaces de travail comme des lieux à construire pour que l'utilisateur puisse y exercer au mieux son travail (Lautier, 1999). Les problèmes (ou exercices) ne font pas directement partie de l'ETM, cependant ils en sont la raison d'être et aussi l'activateur. L'ETM peut aussi être vu comme un espace en constante évolution et la résolution de problèmes contribue à cette évolution.

Dans le modèle des ETM, la conceptualisation est conçue comme le fruit d'une interaction entre un individu et des problèmes (dans un sens large). Le travail mathématique est alors organisé dans un modèle (figure 2.3) articulant deux plans : le plan épistémologique, en rapport étroit avec les contenus mathématiques du domaine étudié, et le plan cognitif, axé sur la pensée du sujet résolvant des tâches mathématiques.

Le plan épistémologique est constitué de trois composantes :
- un ensemble de signes ou *representamen*,
- un ensemble d'artefacts, tels que les logiciels et les instruments de dessin,
- un système théorique de référence (ou référentiel théorique), basé sur des définitions et des propriétés ou théorèmes relatifs au domaine.

Les composantes épistémologiques ne peuvent suffire pour considérer l'activité mathématique. Il est aussi essentiel de comprendre comment les élèves vont donner un sens à tous ces signes, artefacts et éléments théoriques. Cela implique un deuxième niveau de l'ETM centré sur le sujet vu comme un sujet cognitif.

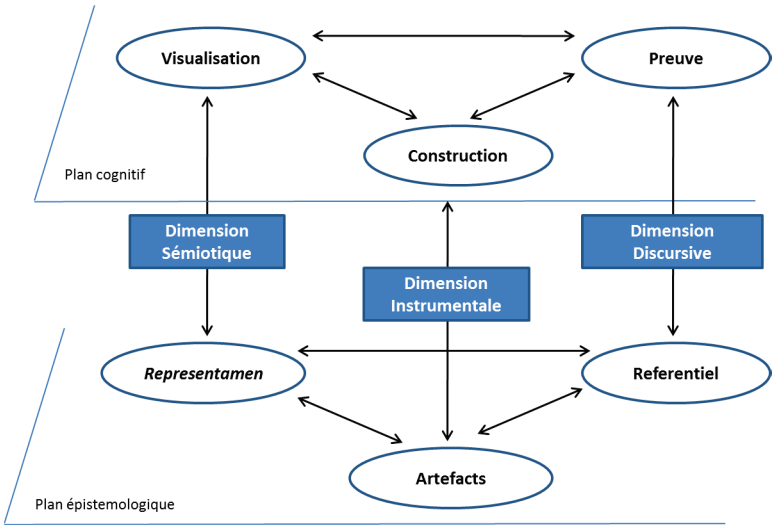


Figure 2.3 – Le modèle des ETM

Le plan cognitif est lui aussi composé de trois processus étroitement associés aux composantes du plan épistémologique :

- la visualisation,
- la construction,
- la preuve.

Pour décrire l'articulation entre ces deux plans, Kuzniak parle de genèses (2011). Étant donné que nous n'étudierons pas le processus qui relie les deux plans, nous préférons utiliser le terme « dimension », plus large. Nous distinguons donc trois dimensions :

- la dimension sémiotique, basée sur les registres de représentation sémiotique (Duval, 1993) qui confère aux objets tangibles de l'ETM un statut d'objet mathématique opérationnel ;
- la dimension instrumentale, qui permet de rendre opérationnel les artefacts dans le processus de construction, et
- la dimension discursive (de la preuve) qui donne un sens aux propriétés pour le mettre au service du raisonnement.

L'idée de *representamen* est à relier avec la notion de signe, au sens de Peirce (1931), mais aussi de représentation sémiotique (Duval, 1993). Les mathématiques sont fondées sur des représentations sémiotiques ; il est donc pertinent de s'intéresser à ces signes ou *representamen* dans l'étude du travail mathématique. Le signe ou le *representamen* est une « chose » qui en représente une autre que ce soit son objet ou peut-être aussi lui-même (Peirce,1931). Les objets mathématiques ne sont abordables que par les signes qui les représentent. Ces propos confirment l'intérêt de notre question principale qui est de s'intéresser au macro-signe $(*)$ et donc aux objets mathématiques qui en dépendent.

Les *representamen* sont à mettre en relation avec le processus de visualisation. La visualisation est à prendre dans un sens plus large, dans lequel il faut distinguer ce

processus de visualisation, « étendue », de la simple vision ou perception des objets. Kuzniak et Richard (2014) précisent qu'il « *peut être envisagé comme le processus de structuration des informations apportées par les diagrammes et les signes. Il nourrit l'intuition des propriétés et il contribue parfois à fonder cognitivement la validité de ces propriétés. Sous certaines conditions, il peut s'apparenter à un raisonnement de type discursivo-graphique (Richard, 2004) et pourra s'exprimer à l'intérieur de registres de représentation sémiotique déterminés* ». Cela constitue la dimension sémiotique.

L'ensemble des signes peut être organisé suivant différents registres de représentation (Duval, 1993). Les registres que nous allons pouvoir rencontrer dans notre recherche sont les suivants :

- le registre symbolique probabiliste : nous entendons les expressions du type $P(a \leqslant X \leqslant b)$ ou encore $P(X \in [a; +\infty[)$;
- le registre symbolique de l'intégrale : par exemple $\displaystyle\int_a^b f(x)\,\mathrm{d}x$;
- le registre graphique fonctionnel, par exemple la représentation graphique de l'aire sous la courbe ;
- le registre de la langue naturelle (probabiliste ou analytique).

Ces registres peuvent être clairement identifiés à un domaine mathématique (les deux premiers) ou être plus transversaux (le registre de la langue naturelle). Duval (1996) s'intéresse à la connaissance en tant qu'activité d'un être individuel. Son approche est cognitive. Dans l'ETM, nous nous plaçons dans le plan cognitif, plus particulièrement au niveau de la composante « visualisation ». Il considère l'activité cognitive impliquée dans les démarches mathématiques différente par rapport aux autres disciplines. Il l'explique par le fait qu'en mathématiques, on a recours à plusieurs registres de représentations sémiotiques. De plus les objets mathématiques ne sont jamais des objets accessibles par la perception. Il identifie des représentations sémiotiques discursives (telles que la langue naturelle, la langue formelle...) et d'autres non discursives (telles les figures, les graphiques, les schémas...). Ces représentations ont plusieurs fonctions cognitives fondamentales : de communication, d'objectivation et de traitement. Les représentations sémiotiques peuvent être regroupées en registres de représentation. On peut ensuite parler de changement de registre : « *il s'agit d'une conversion de la représentation de quelque chose en une représentation de cette même chose dans un autre système sémiotique* » (Duval, 1996, p. 357). Un changement de registre peut être confronté à des difficultés liées à des non-congruences entre registres. Cette non-congruence peut ne poser problème que dans un sens de conversion. Il appelle *traitement* une transformation de représentation où l'on reste dans le même registre et *conversion* lorsque l'on change de registre. Duval (*ibid.*) fait l'hypothèse que « *c'est en prenant simultanément en compte deux registres de représentations, et non chacun isolément, que l'on peut analyser le fonctionnement cognitif des activités mathématiques* » (p. 373).

Les artefacts dans le plan épistémologique sont à relier à la conception de Rabardel (1995), pour qui un artefact peut être matériel (du monde physique) ou symbolique. Par artefact, il entend plus qu'une « *chose ayant subi une transformation d'origine humaine* », mais surtout une « *chose susceptible d'un usage, élaborée*

pour s'inscrire dans des activités finalisées » (*ibid.*, p. 49). Dans le travail mathématique, en probabilités, nous pensons par exemple aux logiciels, à la calculatrice, aux tables de la loi normale, mais aussi aux méthodes données par l'enseignant. Rabardel désigne par « instrument » : « *l'artefact en situation, inscrit dans un usage, dans un rapport instrumental à l'action du sujet, en tant que moyen de celui-ci* » (*ibid.*, p. 49). Dans le modèle des ETM, le processus de construction est déterminé justement par les instruments utilisés. L'articulation entre artefact et construction se produit avec la genèse instrumentale (Artigue, 2002 ; Guin & Trouche, 2002).

Enfin, le référentiel théorique est constitué des définitions, des propriétés et des théorèmes à disposition des élèves. Cet ensemble doit être clairement identifié et donc faire l'objet d'une institutionnalisation. Il sert ensuite à une démarche de raisonnement et est utilisé dans un discours pour démontrer, nous parlerons de preuve. Cette dimension est appelé discursive, en gardant à l'esprit qu'il s'agit d'un discours de la preuve.

c) Les circulations entre les différentes composantes de l'ETM

Il ne faut pas voir ces trois dimensions (sémiotique, instrumentale et discursive), et *a fortiori* les composantes, comme isolées les unes des autres. Il existe des dépendances entre elles. Dans cette perspective, Kuzniak et Richard (2014) ont défini trois plans verticaux liant à chaque fois deux des dimensions : le plan sémiotico-instrumental, le plan sémiotico-discursif et le plan instrumentalo-discursif. Chacun mêlant deux dimensions simultanément.

À partir du modèle des ETM, Kuzniak, Nechache et Drouhard (2016) ont développé l'idée de travail mathématique complet. Ils expliquent qu'un travail mathématique est complet lorsqu'il vérifie les conditions suivantes :

1. une véritable relation entre les plans épistémologiques et cognitifs. Cela signifie que les élèves sont capables de choisir les outils adéquates pour traiter un problème et ensuite de les utiliser en tant qu'instruments pour résoudre la tâche donnée.
2. une articulation riche entre les différentes genèses et les plans verticaux. Cela signifie que des composantes variées du travail mathématique sont prises en compte. Cela reflète une diversité de contexte de travail mathématique dans les activités des élèves (Kuzniak *et al.*, 2016).

d) Les différents niveaux des ETM

Le travail mathématique, dans le contexte scolaire, peut être décrit à plusieurs niveaux (Kuzniak, 2011, p. 15-16). Le travail mathématique visé par l'institution est décrit dans l'ETM de référence. Nous pouvons étudier cet ETM à travers les programmes, ainsi que les documents officiels tels que les documents ressources ou documents d'accompagnement mis en ligne sur Eduscol. De plus, nous considérons que les sujets d'examens nationaux, comme le baccalauréat, sont aussi à considérer dans cet ETM, dans la mesure où ils constituent un ensemble de tâches proposés par l'institution et que donc l'institution considère comme des tâches réalisables

(tout du moins comme étant à savoir réaliser) par les élèves. Cet ETM ne peut être décrit que partiellement car les programmes notamment laissent une grande part d'autonomie à l'enseignant.

Cet ETM de référence doit ensuite être aménagé en ETM idoine pour permettre une mise en place effective dans les classes, où chaque élève travaille dans son ETM personnel. L'adjectif « idoine » qui, selon le dictionnaire Larousse en ligne, signifie « *qui convient exactement à la situation* », est à prendre ici de façon plutôt subjective, dans le sens où l'enseignant qui propose cet ETM le pense idoine. Cela ne signifie pas qu'objectivement il soit parfaitement adéquat pour l'enseignement de telle notion. Par cet adjectif, nous faisons l'hypothèse, qui semble raisonnable, que l'enseignant qui prépare son cours le veut idoine dans sa classe. Nous considérons deux types d'ETM idoine : les ETM idoines potentiels (Nechache, 2015) et les ETM idoines choisis. Les manuels, qui servent de ressources pour beaucoup d'enseignants, proposent des ETM idoines potentiels, inspirés par l'ETM de référence. Ensuite l'enseignant adapte ces ETM idoines potentiels, voire d'autres (diverses ressources), en ETM idoine choisi, choisi *par* l'enseignant *pour* sa classe particulière, en prenant en compte son propre ETM personnel, ainsi que les ETM personnels de ses élèves. Le mot « potentiel » est à prendre dans le sens où les manuels offrent des potentialités, des possibles : tout ce que propose le manuel ne peut être repris dans une classe donc l'enseignant doit à l'intérieur faire des choix, de plus, la plupart du temps, il module les situations choisies.

e) ETM et domaine mathématique

Les mathématiques sont structurées suivant différents domaines tels l'analyse, l'algèbre, les probabilités, la statistique,... Un domaine des mathématiques est une partie des mathématiques définie par des objets, des représentations de ces objets et un référentiel théorique portant sur ces objets. Ce domaine doit avoir une structure cohérente du point de vue mathématique et épistémologique. Le modèle considère souvent un ETM associé à un domaine mathématique ; ce qui est le cas des ETG (ETM de la géométrie). On parle alors d'ETM de l'analyse (que l'on notera $\text{ETM}_{\text{analyse}}$), d'ETM des probabilités ($\text{ETM}_{\text{probabilités}}$),... Dans notre travail, nous nous intéresserons aux domaines de l'analyse et des probabilités, mais aussi au domaine de la statistique. Comme le fait remarquer Kuzniak (2014) : « *les mathématiques ne sont pas la réunion disjointe des divers domaines qui les composent car ils entretiennent entre eux des relations complexes* » (p. 388). C'est pour cette raison qu'il peut être pertinent de s'intéresser à ces changements de domaines. Kuzniak (2014, p. 388) fait un parallèle avec la notion de changement de cadre, utilisée par Douady (1986). Pour Douady, un cadre « *est constitué des objets d'une branche des mathématiques, des relations entre les objets, de leurs formulations éventuellement diverses et des images mentales associées à ces objets et ces relations* » (p. 11). Kuzniak (2014), complété par Montoya Delgadillo et Vivier (2014), montre que la notion de cadre est plus large que celle de domaine :

> Un cadre contient les images mentales des élèves, leurs connaissances, ce qui n'est pas le cas d'un domaine qui est essentiellement mathématique. Si le cognitif est directement pris en charge dans un cadre, en revanche il n'est pas présent dans la notion de domaine. Dans le modèle des ETM que nous

> utilisons, le cognitif est considéré dans le plan cognitif et les images mentales d'un sujet peuvent être considérées à travers la notion d'ETM personnel. (Montoya Delgadillo & Vivier, 2014, p. 76).

Dans notre étude didactique, par exemple, le domaine mathématique peut être parfois considéré comme trop conséquent. Dans ce cas, on peut se restreindre à un sous-domaine. Dans le domaine des probabilités, nous pouvons distinguer le sous-domaine des probabilités discrètes et celui des probabilités continues. Cependant, dans notre recherche, nous voulons nous limiter aux lois de probabilité à densité. Or nous avons mis en évidence dans l'étude mathématique (section 1.3) qu'il existe des lois de probabilité continues qui ne sont pas à densité. Nous avons donc décidé d'appeler probabilités à densité, le sous-domaine qui nous intéresse ici. Nous avons bien conscience qu'il s'agit d'un abus de langage pour signifier les probabilités de lois à densité. Nous aurons donc affaire au sous-domaine des probabilités à densité (que nous noterons PaD), à celui du calcul intégral (CI) et aussi à celui de la statistique descriptive (SD). Nous précisons que nous considérerons les fonctions comme faisant partie de l'ETM relatif au calcul intégral si elles interviennent pour ensuite calculer une aire sous la courbe ou une intégrale.

On peut donc relever plusieurs types d'ETM : des ETM associés à un domaine, des ETM associés à un sous-domaine et des ETM globaux. Selon Montoya Delgadillo et Vivier (*ibid.*), « *un ETM global regroupe l'ensemble des composantes cognitives et épistémologiques dans le travail mathématique de plusieurs domaines* ». Ils mettent en garde en précisant que la notion d'ETM global est à réserver dans le cas où les registres, domaines, notions, etc. en jeu sont suffisamment bien intégrés, d'une manière cohérente. Ils ajoutent :

> On peut d'ailleurs interpréter comme un enjeu majeur de l'enseignement des mathématiques le fait d'aboutir à un ETM global coordonnant les ETM associés à des domaines mathématiques. On peut en particulier penser que l'ETM personnel d'un enseignant de mathématiques est global ce qui lui permet d'élaborer un ETM idoine qui prend en charge plusieurs domaines, même si l'on peut interroger la coordination des ETM associés aux différents domaines au sein de cet ETM idoine (p. 78).

2.3.2 Quelques éléments de théories de l'activité qui complètent notre cadre théorique

L'ETM est un espace qui se construit et qui évolue au cours de l'apprentissage des élèves. Ce sont les tâches données aux élèves qui déclenchent un travail mathématique et ainsi activent des dynamiques entre les différentes dimensions de l'ETM. Le rôle des tâches mathématiques nous semble central, c'est pourquoi il semble pertinent d'étudier les tâches qui déclenchent l'activité mathématique. C'est pourquoi nous pensons qu'il est intéressant de compléter le modèle des ETM, avec des éléments des théories de l'activité (Vygotsky, 1997 ; Leontiev, 1984). Ces travaux ne concernent pas explicitement les apprentissages scolaires, ni les mathématiques, mais ont ensuite été adaptés à la didactique des mathématiques.

La théorie de l'activité utilisée en didactique des mathématiques en France prend appui sur ces considérations. L'action (et ensuite l'activité) est considérée comme centrale dans l'apprentissage des élèves et en particulier dans la conceptualisation

des notions mathématiques (Piaget, 1971, 1985 ; Vergnaud, 1991). Comme l'explique Rogalski (2008, p. 23) :

> l'objet de [la] théorie [de l'activité] est une action finalisée et motivée : le sujet vise l'atteinte de buts d'action, et ce sont les mobiles de son activité qui sont le moteur de ses actions.

Elle précise en note de bas de page que la théorie de l'activité articule trois niveaux de finalités :

- le motif ou le mobile de l'activité du sujet (son intention quand il s'engage dans l'action) ;
- le but de l'action qui est un certain état de l'objet sur lequel s'exerce cette action ;
- les opérations qu'il faut effectuer pour réaliser l'action.

Elle poursuit en présentant l'objectif de cette théorie et ses notions clés :

> La théorie [de l'activité] vise l'analyse des processus en jeu chez le sujet agissant, et les processus par lesquels son activité évolue et par lesquels il se développe. Elle s'appuie sur deux notions clés : celle de sujet et celle de situation. Elle différencie par ailleurs tâche et activité, qui sont respectivement « du côté de la situation » et « du côté du sujet ». (p. 23)

Nous allons donc donner ici la définition de ces termes.

a) Tâche et activité

Dans les théories de l'activité, tâche et activité sont deux termes liés mais à bien distinguer. En reprenant les mots de Rogalski (2008) :

> La tâche est ce qui est à faire : le « *but qu'il s'agit d'atteindre sous certaines conditions* » (définition proposée par Leontiev en 1984). L'activité est ce que développe un sujet lors de la réalisation de la tâche : non seulement ses actes extériorisés, mais aussi les inférences, les hypothèses qu'il fait, les décisions qu'il prend, dans ce qu'il fait et ce qu'il se retient de faire ; l'activité comprend aussi la manière dont le sujet gère son temps, et également son état personnel, ainsi que ses interactions avec autrui dans la situation de travail. (p. 24)

Lorsque la tâche est complexe, le but à atteindre est composé de différents sous-buts. On parle dans ce cas de sous-tâches. Le sujet répond à des tâches qui lui sont proposées (voire imposées). Ces tâches peuvent être définies dans les termes des résultats attendus par le prescripteur : on est du côté de la tâche prescrite. L'activité n'est quant à elle pas la réponse directe à la tâche prescrite : la tâche est redéfinie par le sujet, qui pour la réaliser s'en fait une représentation, s'autorise ou s'interdit des possibles, lève des contraintes ou s'en pose... On est alors du côté de la tâche effective à laquelle l'activité du sujet répond. Il existe donc un décalage entre tâche prescrite et tâche effective, lié à l'existence des deux points de vue : celui du prescripteur et celui du réalisateur. La tâche effective, celle que le sujet réalise, peut s'éloigner de la tâche prescrite pour plusieurs raisons : le sujet manque de mobile pour s'engager dans l'action attendue, il ne dispose pas des compétences nécessaires pour la réaliser ou encore il s'en est construit une représentation erronée. En situation institutionnelle, l'activité est très orientée par la réalisation de la tâche. Les observables de l'activité les plus accessibles sont souvent les opérations sur les objets sur lesquels porte l'action. Cela explique que l'analyse de l'activité, en didactique des mathéma-

tiques par exemple, s'appuie dans un premier temps sur une analyse préalable de la tâche.

b) Le sujet et la situation

Le sujet est considéré comme individualisé, avec des intentions (son activité est finalisée), des compétences et des responsabilités. Ses compétences constituent un potentiel de ressources permettant d'engendrer des actions pour réaliser ses buts, marqué par une insertion sociale mais non identifié à celle-ci.

En ce qui concerne la situation, le sujet n'est pas seul maître de ses buts, ni de ses moyens. Il agit dans une situation de travail qui comporte son propre système de contraintes et de ressources ; il a une tâche à accomplir qui le lie à un prescripteur par un contrat partiellement implicite.

c) Contextualisation de ces notions au sujet « élève » en classe de mathématiques

Nous faisons le choix, dans ce travail, d'étudier les activités des élèves (en particulier toutes celles qui peuvent se passer en classe) comme intermédiaires entre les apprentissages des élèves et l'enseignement correspondant (Vandebrouck, 2008).

Si l'on reprend les termes définis précédemment, dans le cas ici du sujet « élève » en situation scolaire, dans la classe de mathématiques, quelques adaptations sont à prendre en compte. Notamment, il existe un décalage entre tâche (mathématique) prescrite et tâche (mathématique) effective. Pour l'enseignant (prescripteur), le but de production de l'activité de l'élève est l'apprentissage de l'élève (et non, par exemple, la valeur de l'inconnue dans la résolution d'une équation du premier degré donnée en exercice). Cependant, il se peut que ce ne soit pas la même chose pour l'élève, qui peut interpréter la tâche sur « *le registre de l'accomplissement de la tâche indépendamment des apprentissages attendus* » (Bautier, 2006, p. 113). L'élève qui agit dans une situation de travail (par exemple dans la résolution d'un problème) avec ses contraintes et ses ressources a une tâche à accomplir qui le lie à l'enseignant par le contrat didactique (Brousseau, 1988). Pour Robert (2008c), « *une tâche mathématique est caractérisée par les mises en fonctionnement des connaissances anciennes et nouvelles des élèves, déterminées à partir de ce qui est à leur disposition dans leur cours* » (p. 40).

Les activités des élèves sont vues comme pouvant engendrer des connaissances mathématiques. Comme le précise Robert (2008c), elles sont constituées de « *tout ce que font les élèves, mais aussi ce qu'ils ne font pas, disent, ne disent pas, pensent...* » (p. ! !). Elles sont conditionnées par beaucoup de facteurs, qui dépendent plus ou moins de l'enseignant. Nous ne retenons, comme contribuant aux transformations espérées en matière de connaissances, que les cours et les exercices proposés aux élèves ainsi que les conditions de travail organisées en classe. Les autres déterminants (facteurs affectifs, facteurs socio-culturels, éléments liés à la conjoncture,...) sont considérés comme des paramètres. De plus, bien entendu, les activités d'un sujet dépendent du sujet lui-même. Nous venons de dire que les tâches données aux élèves entraînent des activités effectives des élèves. Cependant, ces tâches effectives ne sont accessibles que partiellement, nous ne pouvons donc faire que des hypothèses sur ces

activités. Nous parlons donc d'activités possibles (Robert, 2008c), qui sont à mettre en relation avec les activités attendues (*a priori*).

Décrire la circulation dans l'ETM correspond en fait à une relecture globale des activités possibles des élèves. De plus, l'analyse des activités est locale, au niveau de la tâche (même des sous-tâches), tandis que l'analyse du travail mathématique, dans le cadre des ETM, est plus globale, il faut prendre en compte plusieurs tâches. Penser en terme d'ETM oblige à dépasser l'analyse de tâches pour penser aux effets à plus long terme.

d) Les pratiques des enseignants

Nous pourrons être amenée à considérer le rôle de l'enseignant notamment pour mieux comprendre ce qu'il organise en classe comme activité pour les élèves. Dans ce cas, nous nous placerons dans le cadre de la double approche didactique et ergonomique (Robert & Rogalski, 2002). Nous parlerons de pratiques pour qualifier tout ce qui se rapporte à ce que l'enseignant pense, dit ou ne dit pas, fait ou ne fait pas, sur un temps long, que ce soit avant, pendant ou après les séances de classe (Robert, 2008a).

Robert (*ibid.*) identifie cinq composantes qui jouent sur le métier d'enseignant de mathématiques, qui correspondent à des choix ou des contraintes :

- la composante cognitive : le choix sur les contenus, les tâches, leur organisation, leur quantité...
- la composante médiative : le choix dans les déroulements, le discours... Cela correspond au cheminement organisé pour les différents élèves ;
- la composante personnelle : les choix liés au temps long (représentations de l'enseignant, les risques...) ;
- la composante institutionnelle : liée au respect des contraintes dans l'exercice d'un métier (programme, horaires, manuels, inspection...) ;
- la composante sociale : liée au fait que l'enseignant n'est pas seul. L'enseignant fait partie de différents groupes sociaux (la classe, le personnel de l'établissement...) qui sont soumis à des exigences, des attentes...

Certains de ces déterminants sont externes à la classe (institutionnels, sociaux, personnels).

Nous pourrons prendre en compte ces composantes pour comprendre ce qui peut peser sur les choix des enseignants.

2.4 Questions de recherche et méthodologie générale

Maintenant que nous avons présenté notre cadrage théorique, nous allons reformuler nos questions de recherche, présentées dans la section 2.1.3, dans ce cadre. Nous détaillerons ensuite les différents objectifs de cette recherche, ainsi que la méthodologie que nous avons décidé d'adopter pour répondre à nos trois questions.

2.4.1 Nos questions de recherche

Il s'agit ici de reformuler nos questions de recherche, présentées page 41, dans le cadre théorique choisi.

La question de recherche principale peut se reformuler comme ceci :

Est-il possible de concevoir et de mettre en œuvre des tâches dont les activités, attendues et possibles, des élèves leur permettent de construire conceptuellement la notion de fonction de densité et donc la relation sémiotique $(*)$ **?**

Notre hypothèse de recherche est que de telles tâches existent. Pour la vérifier, nous posons les trois questions de recherche suivantes.

Question de recherche n° 1 :

> **QR1** : Comment historiquement et épistémologiquement sont apparus les liens entre probabilités à densité et calcul intégral ?
> Comment a émergé le concept de lois à densité et tout particulièrement la notion de fonction de densité ?

Cette première question de recherche est d'ordre épistémologique. Elle nous semble importante pour mieux comprendre la notion de fonction de densité et sa naissance historique, en s'intéressant à l'émergence des lois à densité.

Ensuite pour repérer les enjeux didactiques derrière la relation sémiotique liant probabilités à densité et calcul intégral, nous reformulerons notre question comme suit.

Question de recherche n° 2 :

> **QR2** : En terminale S, quels sont les enjeux didactiques derrière la relation sémiotique $P(a \leqslant X \leqslant b) = \displaystyle\int_a^b f(x)\,\mathrm{d}x$?
> Quelles sont les tâches proposées et les activités attendues des élèves pour l'introduction de la notion de fonction de densité ?
> Quelles sont les tâches qui permettent d'exploiter les articulations entre l'$\mathrm{ETM}_{\mathrm{PaD}}$ et l'$\mathrm{ETM}_{\mathrm{CI}}$ [5] ?

Dans cette question, nous nous demandons quelles tâches sont proposées et quelles activités sont attendues des élèves dans les classes de terminale S pour la construction de l'$\mathrm{ETM}_{\mathrm{PaD}}$, et notamment le référentiel théorique lié à la fonction de

5. Nous rappelons que $\mathrm{ETM}_{\mathrm{PaD}}$ signifie espace de travail mathématique associé aux probabilités à densité et que $\mathrm{ETM}_{\mathrm{CI}}$ signifie espace de travail mathématique associé au calcul intégral.

densité, et ensuite pour l'exploitation de l'articulation entre l'ETM_{PaD} et l'ETM_{CI}. Nous nous demanderons si l'on retrouve l'approche historique.

Question de recherche n° 3 :

> **QR3** : Quels éléments didactiques peut-on prendre en compte pour concevoir et mettre en œuvre effectivement des tâches qui permettent aux élèves de construire la notion de fonction de densité ?

Cette question porte sur les contenus mathématiques à prendre en compte et l'approche à favoriser pour la conception de tâches qui permettent la construction par les élèves du référentiel théorique lié à la notion de fonction de densité, mais elle va plus loin. Il s'agit aussi de prendre en compte les contraintes de l'enseignant et de se questionner sur la gestion du travail mathématique par l'enseignant pour arriver à l'objectif d'apprentissage visé.

2.4.2 Objectifs et méthodologie générale de la recherche

Les questions de recherche viennent d'être reformulées. Nous allons maintenant exposer nos objectifs et notre méthodologie associés à chaque question de recherche, de façon assez générale. Nous donnerons plus de précisions sur la méthodologie au fur et à mesure des chapitres.

L'objectif de la thèse est de comprendre quels sont les enjeux didactiques dans la relation sémiotique $(*)$ qui met en relation les probabilités à densité et le calcul intégral, pour ensuite concevoir et expérimenter en classe des tâches permettant une construction de la notion de densité plus porteuse de sens afin de mieux justifier les liens entre probabilités et analyse et ainsi mieux les exploiter.

a) Question de recherche QR1

La première question de recherche sera traitée dans la **Partie I** (**chapitre 3**). Elle porte sur l'émergence de la notion de lois à densité et du lien entre les probabilités à densité et le calcul intégral. Il est important, dans un premier temps, de se focaliser sur le contenu mathématique.

L'objectif relatif à cette question de recherche est de décrire historiquement et épistémologiquement la naissance des lois à densité et la genèse du lien entre probabilité et calcul intégral. Le but est ensuite d'en dégager des éléments qui paraissent importants voire incontournables et révélateurs de sens dans la construction de ces connaissances mathématiques.

Pour répondre à cette première question de recherche, le travail sera bibliographique. Il ne s'agit pas de lire l'ensemble des ouvrages des mathématiciens de l'époque ayant travaillé les probabilités ou l'analyse. Nous nous contenterons de la lecture d'articles d'épistémologie à visée didactique ou d'articles de didactique des mathématiques avec un point de vue historique et épistémologique autour des probabilités à densité. Nous compléterons cette étude par deux ouvrages d'histoire des

sciences assez révélateurs, ainsi que quelques textes originaux. Ce travail bibliographique nous permettra d'avoir un point de vue assez général sur l'émergence des lois à densité et de la notion de densité.

b) Question de recherche QR2

La deuxième question de recherche porte sur les tâches proposées et les activités attendues des élèves en classe de terminale S pour l'introduction de la notion de densité, mais aussi au niveau de l'exploitation des liens entre les deux sous-domaines (PaD et CI). Elle sera traitée dans la **Partie II** (**chapitres 4** et **5**).

L'objectif général ici est d'analyser ce qui se fait dans les classes de terminale S. Nous voulons déterminer comment se construit l'ETM_{PaD} et notamment le lien entre l'ETM_{PaD} et l'ETM_{CI}. Étant donné que la finalité de cette thèse est de proposer et d'expérimenter en classe des tâches favorisant la construction de la notion de fonction de densité, le temps limité de la thèse nous a poussé à faire des choix moins coûteux en temps. C'est la raison pour laquelle, pour répondre à cette question de recherche, nous avons préféré nous limiter en grande partie à l'analyse de ressources plutôt qu'une analyse effective dans les classes. Nous allons donc nous focaliser essentiellement sur l'ETM de référence et les ETM idoines potentiels proposés par les manuels. L'analyse de l'ETM de référence doit se faire en amont de celle des ETM idoines car c'est l'ETM de référence qui est censé guider les ETM idoines. Dans chaque cas, nous n'allons pas analyser l'ensemble de l'ETM des probabilités à densité et celui du calcul intégral mais nous allons regarder les deux d'un point de vue global avec un angle de vue particulier, celui des connexions et des articulations effectives ou possibles entre l'ETM_{PaD} et l'ETM_{CI}.

Tout d'abord, comme nous nous plaçons dans le contexte institutionnel de la terminale S, et donc du programme correspondant, il nous a semblé judicieux de commencer par l'étude des ETM de référence (institutionnelle) pour bien comprendre les choix visibles à ce niveau. L'objectif est donc dans un premier temps de décrire et analyser l'ETM_{PaD} de référence en lien avec l'ETM_{CI}, à travers des tâches proposées ou ne serait-ce que des éléments mis en avant. Cette étude constitue le **chapitre 4**. Pour étudier les ETM de référence, nous avons bien entendu décidé de prendre en considération le programme de terminale S (MEN, 2011), ainsi que les anciens programmes afin d'observer les évolutions. Nous avons aussi fait le choix d'analyser le document *Ressources* en probabilités et statistique proposé par le Ministère de l'Éducation nationale qui accompagne le programme. Enfin, nous avons pris en compte les sujets de baccalauréat de mathématiques de la filière scientifique des différents centres d'examen (France métropolitaine et centres étrangers) des années 2013, 2014 et 2015, plus exactement les exercices se rapportant aux probabilités à densité et au calcul intégral. Pour nous, il s'agit de l'ensemble des documents permettant de décrire l'ETM de référence, dont la description reste de toute manière partielle.

Après une étude de l'ETM de référence, pour analyser ce qui peut se faire dans les classes, l'étude des manuels de terminale S nous a semblé être un bon compromis pour entrevoir un panel de possibles. Nous avons bien conscience que cela ne permet pas d'avoir réellement connaissance de ce qui se passe effectivement dans les classes

mais cela nous permet d'analyser des ETM idoines potentiels qui nous permettent d'approcher quelque peu des ETM idoines effectifs.

Le deuxième objectif de la question de recherche QR2 est donc d'analyser les ETM idoines potentiels des manuels de terminale S. Nous étudierons les chapitres concernant le calcul intégral et les lois de probabilité à densité des huit manuels de mathématiques disponibles sur le marché français. Ce point sera abordé dans le **chapitre 5**.

Pour apporter un point de vue d'enseignants, nous avons conduit des entretiens individuels avec quatre enseignants, travaillant dans des contextes différents, qui nous permettent de valider les éléments mis en avant dans les ETM idoines potentiels des manuels.

L'objectif n'est pas ici de décrire et d'analyser l'ETM choisi des quatre enseignants car un entretien ne permet pas d'avoir accès à tout cela, mais l'objectif est de renforcer les considérations faites sur les ETM idoines potentiels.

Nous pouvons résumer la méthodologie associée à QR2 par le schéma ci-dessous (2.4).

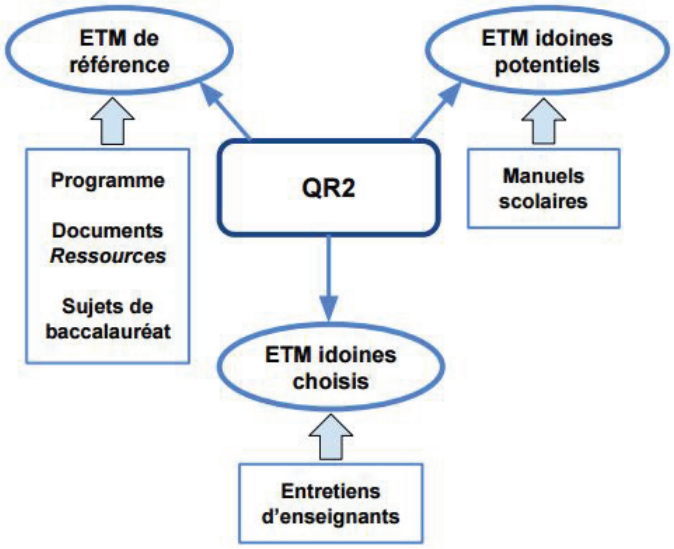


Figure 2.4 – Schéma de la méthodologie associée à QR2

c) Question de recherche QR3

La troisième question de recherche QR3 constitue la **Partie III** de la thèse (**chapitres 6**, **7** et **8**). Cette question interroge sur la possibilité de concevoir et mettre effectivement en place en classe de terminale S des tâches qui permettent d'introduire la notion de fonction de densité. L'objectif de cette partie est de concevoir, expérimenter et analyser un ETM idoine dont l'objectif est de favoriser et enrichir la compréhension de la relation entre probabilités à densité et calcul intégral en donnant du sens à la fonction de densité. Le but est de proposer des tâches dont les activités attendues des élèves leur permettent réellement de construire le référentiel théorique lié à la notion de fonction de densité.

La méthodologie que nous avons décidé de mettre en place est celle de l'ingénierie didactique (Artigue, 1988). Cependant, désirant que ce projet d'enseignement s'intègre parfaitement dans un contexte de classe habituel, avec notamment toutes les contraintes que cela implique (les contraintes du programme, mais aussi les contraintes institutionnelles), nous avons fait le choix de mettre en place une ingénierie didactique prenant en compte une dimension collaborative de type recherche collaborative (Desgagné, 1997). Nous avons donc conçu, conjointement avec une enseignante, les tâches (et même la séquence) et nous les avons expérimentées dans sa classe de terminale S (**chapitre 6**).

Dans cette question de recherche, nous nous limiterons à l'analyse de la construction de l'ETM_{PaD} idoine effectif dans la classe (par rapport à l'ETM idoine prévu), en interaction avec l'ETM_{CI} mais aussi avec l'ETM de la statistique descriptive, pendant la résolution des tâches données en introduction de la séquence. Nous analyserons le travail mathématique en jeu et nous identifierons les conditions notamment au niveau de la gestion du travail mathématique de l'enseignante qui permettent à la classe d'arriver à l'objectif d'apprentissage visé (**chapitre 7**).

En plus de la comparaison entre l'analyse *a priori* et *a posteriori*, pour valider les choix faits, nous ajouterons des éléments pour élargir la portée des résultats (**chapitre 8**).

Nous pouvons schématiser notre méthodologie générale, qui présente le cheminement global de la recherche, par la figure 2.5.

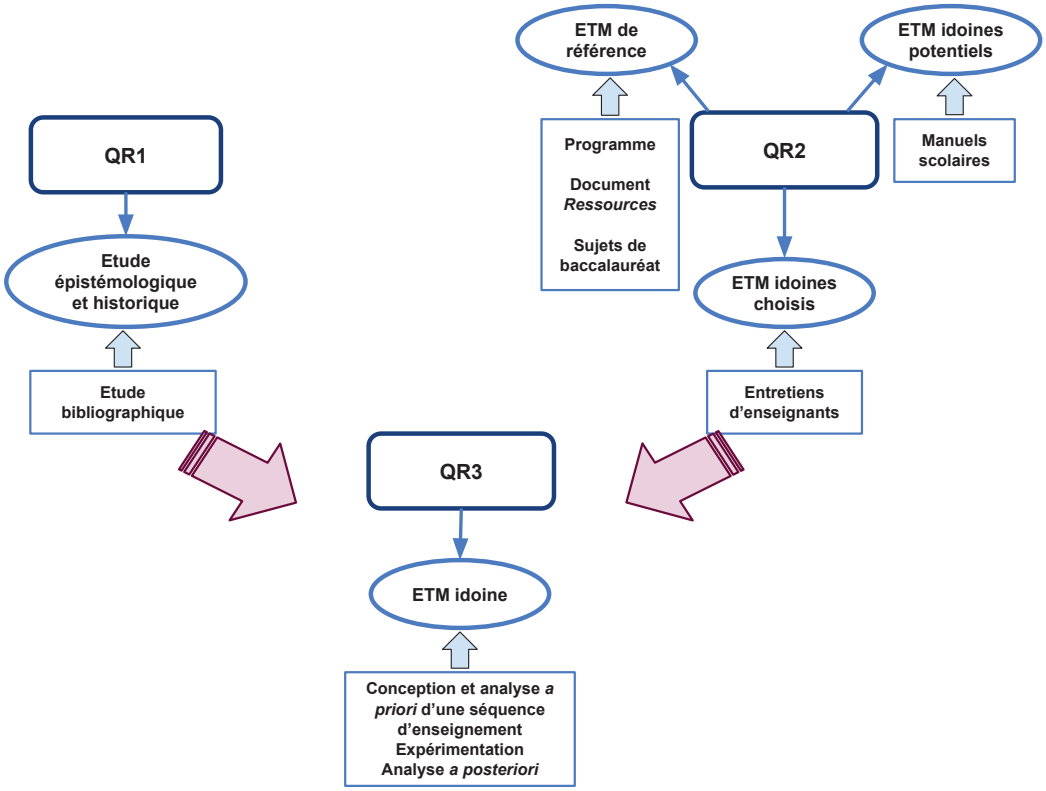


FIGURE 2.5 – Schéma de la méthodologie générale

Première partie

Approche épistémologique et historique

Cette première partie, constituée d'un seul chapitre (**chapitre 3**), a pour objectif de répondre à notre première question de recherche que nous rappelons ici :

> **QR1** : Comment historiquement et épistémologiquement sont apparus les liens entre probabilités à densité et calcul intégral ?
> Comment a émergé le concept de lois à densité et tout particulièrement la notion de fonction de densité ?

Cette étude historique et épistémologique est un premier contact nécessaire avec les concepts de fonction de densité et de lois à densité pour nous permettre d'aborder nos questions didactiques. Comme l'écrit Dorier :

> [L'] étude [historique et épistémologique] est une phase fondamentale pour que le chercheur puisse prendre ses distances par rapport aux enjeux didactiques. Le sens des concepts, les problèmes qui s'y attachent, la position relative d'un élément de savoir dans un champ de savoir plus large qui l'englobe, sont autant de questions qui aident à mieux comprendre le fonctionnement d'un système didactique. (Dorier, 2000, p. 9)

Dorier nous incite à nous intéresser aux « *problèmes qui s'y attachent* ». Dans sa thèse de doctorat, Robinet (1984) insiste sur ce point. Elle explique :

> Les concepts doivent être introduits par leur fonctionnement et par conséquent « au plus près de leur sens » comme outil, donc dans les problèmes où ils sont une réponse. Historiquement les problèmes, pour lesquels les concepts ont été inventés, remplissent ces conditions, mais d'une part ils ne sont pas toujours adaptables à l'enseignement, d'autre part le fonctionnement du concept le fait apparaître parfois « après coup » comme plus nécessaire dans d'autres types de problèmes. (p. 18)

Dans tous les cas, ressort l'importance de comprendre à quels problèmes historiques la notion répond. Ceci nous a conduit à faire une étude historico-épistémologique de la notion concernée, pour, notamment, nous éclairer sur l'état de la science au moment de l'apparition de cette notion, ainsi que sur les problèmes qui ont servi à créer le nouveau concept.

Nous dégagerons ensuite parmi ces éléments ceux qui pourraient être favorables à l'introduction de la notion de fonction de densité et à la mise en relation entre probabilités à densité et calcul intégral. Nous repérerons aussi des obstacles épistémologiques (Brousseau, 1989) qui pourraient être une difficulté pour cette introduction.

Chapitre 3

Étude historico-épistémologique

Dans ce chapitre, nous donnons un aperçu historique sur la genèse des lois de probabilités à densité afin de comprendre comment sont apparues les notions de lois à densité et de fonction de densité. Les questions qui guident notre travail sont :

- Quels problèmes ont conduit à l'introduction des probabilités à densité ?
- Comment a émergé la notion de fonction de densité ?
- Comment a été établi le lien épistémologique entre le calcul intégral et les probabilités à densité ? Ce lien provient-il de l'intégrale ou des probabilités ?

Il ne s'agit pas ici d'une recherche d'historien des sciences ; le point de vue adopté est celui d'une didacticienne des mathématiques qui cherche à comprendre et mettre en lumière quelques idées, sans prétendre à l'exhaustivité.

Nous commencerons par présenter notre méthodologie (section 3.1), puis, dans la section 3.2, nous donnerons quelques éléments sur l'émergence du calcul intégral, sans trop nous attarder car nous verrons qu'aucun élément des probabilités n'est visible dans ce domaine. Nous poursuivrons, dans la section 3.3 plus conséquente, par des éléments sur la naissance des lois de probabilité à densité, en essayant de dégager différentes étapes dans la construction de ces nouvelles notions. Nous terminerons ce chapitre (section 3.4) en répondant à la question de recherche QR1 et en dégageant les apports de cette étude pour la suite de notre travail.

3.1 Méthodologie

Pour mener cette étude historico-épistémologique, nous avons pris appui, dans un premier temps, sur la thèse de Haddad (2012) consacrée à l'enseignement du calcul intégral pour présenter brièvement l'émergence du calcul intégral.

Pour aborder la genèse des lois à densité, nous nous sommes tout d'abord référée à des articles de Parzysz (2013), Bru (2006) et Pichard (2005). Ces différentes sources portent sur ces questions historiques avec un regard didactique. Bien que ces articles soient à première vue exclusivement centrés sur la loi normale, il est tout de même possible d'en dégager des informations sous un angle plus large et de généraliser aux lois continues.

Nous avons complété notre étude avec deux livres que nous pouvons qualifier d'histoire des mathématiques, ou plus particulièrement d'histoire de la statistique.

Il s'agit de *The history of statistics. The Measurement of Uncertainty before 1900* de Stigler (1986) et de *Histoire de la statistique* de Droesbeke et Tassi (1990).

Enfin, nous avons aussi étudié quelques textes originaux dont certains ont été soigneusement choisis par M.-F. Jozeau lors du colloque inter-IREM « Histoire des Probabilités et des Statistiques » (du 31 mai au 1er juin 2002) pour son atelier portant sur la naissance de la théorie des erreurs.

Notre étude est en grande partie chronologique. Nous nous sommes principalement intéressée au début de l'apparition des lois à densité, lorsque la théorie de la mesure et de l'intégration et *a fortiori* l'axiomatisation de Kolmogorov n'existaient pas encore, et donc que le lien entre probabilités et calcul intégral se faisait par le biais de l'intégrale de Riemann et non celle de Lebesgue. Ceci nous permet de pouvoir faire un parallèle avec l'enseignement actuel de la classe de terminale.

3.2 Émergence historique du calcul intégral : absence des probabilités

Comme nous l'avons expliqué dans l'introduction du chapitre, cette section a pour objectif de donner un rapide aperçu de l'histoire du calcul intégral. Il ne se veut ni exhaustif, ni détaillé [6].

Les premières idées qui touchent au calcul intégral remontent à l'époque de la Grèce antique et sont attribuées à Archimède (287-212 avant J.-C.), qui utilisait notamment la méthode d'exhaustion [7] pour vérifier et valider les solutions de problèmes de rectification et de quadrature. Les véritables fondements du calcul intégral commencent cependant plus tard, simultanément avec la création du calcul infinitésimal par Newton (1643-1727) et Leibniz (1646-1716), dans la deuxième moitié du XVII^{e} siècle et le début du XVIII^{e}. Les deux mathématiciens privilégient des démarches différentes. Newton, dans son traité de 1687 traitant de la quadrature des courbes, considère l'intégration comme étant l'opération réciproque de la dérivation. Il définit donc l'intégrale avec la notion de primitive. Leibniz, lui, développe la notion d'intégrale définie, à partir de la sommation d'infiniment petits. On peut faire remarquer cependant que cette idée de sommation « d'infiniment petits » est déjà présente chez Archimède dans ses calculs d'aires et de volumes, notamment ceux de la boule et du tronc de paraboloïde. Il n'y a pas de développement théorique mais on peut déjà y voir des travaux en gestation.

C'est au XIX^{e} siècle que le calcul intégral se fonde sur des bases plus rigoureuses avec l'introduction de la notion de limite. En 1823, Cauchy (1789-1857), dans son *Résumé des leçons sur le calcul infinitésimal*, est le premier à donner une définition précise de l'intégrale. Il introduit une subdivision x_0,..., x_n de l'intervalle d'intégra-

6. Pour plus de précisions, se reporter au livre de Hawkins (2001), ainsi qu'à la thèse de Haddad (2012).

7. D'après la définition de l'encyclopédie libre Wikipédia : Dans le cas du calcul de l'aire A d'une figure plane, la méthode d'exhaustion consiste en un double raisonnement par l'absurde : on suppose que son aire est strictement supérieure à A, puis on aboutit à une contradiction ; on suppose ensuite que son aire est strictement inférieure à A, puis on aboutit à une autre contradiction. On parvient ainsi à montrer que l'aire de la figure est A.

tion $[a;b]$, puis considère la somme $S = \sum_{i=1}^{n}(x_i - x_{i-1})f(x_{i-1})$. Il montre ensuite que, quelque soit la subdivision choisie, plus les pas de la subdivision sont petits, plus cette somme se rapproche d'une « limite » [8] qui ne dépend que de la fonction f et des bornes de l'intervalle d'intégration $[a;b]$. Cette intégrale est définie pour toute fonction continue par morceaux sur l'intervalle $[a;b]$ et s'étend aussi à des intervalles non bornés.

Ensuite, en 1854, Riemann (1826-1866) expose sa théorie de l'intégration, qui sera considérée comme étant la mieux fondée pour l'époque. Il apparaît ici un saut épistémologique. Sa méthode consiste à approcher la fonction f par une fonction en escalier ϕ constante sur chaque intervalle de la subdivision. Si la fonction est assez régulière, la « limite » de la somme de Riemann, quand les pas de la subdivision sont de plus en plus petits, existe et est appelée l'intégrale de la fonction f sur $[a;b]$. Darboux (1842-1917) propose ensuite une autre approche de l'intégrale, mais nous ne la détaillerons pas ici.

Borel (1871-1956), en 1897, va introduire la théorie des ensembles mesurables ; ce qui permet par la suite à Lebesgue (1875-1941) d'introduire sa propre théorie de l'intégration, plus générale que celle de Riemann, qui va permettre de dépasser certaines « situations gênantes » que l'intégrale de Riemann ne permet pas de surmonter (Droesbeke & Tassi, 1990, p. 38). Par exemple, l'intégrale de Lebesgue va permettre d'intégrer la fonction de Dirichlet (la fonction indicatrice sur $\mathbb{Q}$), non intégrable au sens de Riemann. L'achèvement de la construction des théories de la mesure et de l'intégration se situe dans les années 1930. L'intégrale de Lebesgue, contrairement à l'intégrale de Riemann qui s'applique seulement à des fonctions à variables réelles, découle d'une mesure pouvant éventuellement porter sur des ensembles abstraits. Nous pouvons constater un saut épistémologique entre ces deux intégrales, avec l'intégrale de Lebesgue qui atteint un niveau d'abstraction beaucoup plus élevé que l'intégrale de Riemann.

Par ce rapide aperçu sur l'histoire du calcul intégral, nous pouvons remarquer que les problèmes de calculs d'aires à l'origine du calcul intégral remontent à l'antiquité. Le concept d'intégrale a ensuite mis du temps à se formaliser rigoureusement, il faut attendre le XIX^{e} siècle pour qu'émerge une théorie rigoureuse de l'intégrale, par Riemann. Sa définition de l'intégrale nécessite la formalisation des notions de fonction et de limite ; ce qui justifie, en partie, l'arrivée tardive de l'intégrale.

À aucun moment, nous ne voyons apparaître une quelconque relation avec les probabilités. La mise en relation entre intégrales et probabilités ne vient donc pas de l'analyse.

8. Cette notion de limite appellerait un approfondissement.

3.3 La naissance des lois de probabilité à densité

3.3.1 Les débuts des probabilités : les jeux de hasard

Contrairement au calcul intégral, dont les problématiques datent de l'antiquité, il faut attendre le XVII^{e} siècle pour qu'apparaissent les premiers problèmes de probabilités. On attribue la naissance des probabilités à Pascal (1623-1662) et Fermat (1601-1665), à travers leur correspondance sur le problème « des partis » (1654) proposé à Pascal par le chevalier de Méré, non résolu jusqu'alors.
Le problème est le suivant (Henry, 2004) :
Comment répartir les mises (faire le parti) entre les joueurs d'un jeu de hasard en plusieurs manches, interrompu avant que le sort attribue l'enjeu à l'un des joueurs ?
La correspondance de Pascal et Fermat débouche sur deux démarches différentes pour résoudre ce problème.

Huygens (1629-1695) dans *De ratiociniis in ludo alea* (1657) postule l'existence de la notion de probabilité : « *Quoique dans les jeux de hasard pur les résultats soient incertains, la chance qu'un joueur a de gagner ou de perdre a cependant une valeur déterminée* » (Lanier, 1992). Chez Huygens, la notion de probabilité (qui ne porte pas encore ce nom) découle de celle d'espérance.

Après cette époque fondatrice du calcul de probabilités où se sont côtoyés Pascal, Fermat et Huygens entre autres, au milieu du XVII^{e} siècle, les mathématiciens cherchent à mieux cerner la notion de probabilité.

Ainsi, les débuts des probabilités sont à mettre en relation forte avec les jeux de hasard ; les probabilités étudiées sont donc discrètes. Il nous faut aller chercher plus tard dans l'histoire pour voir apparaître les lois de probabilité à densité.

3.3.2 Le problème de Bernoulli : apparition d'intégrales

Jacques Bernoulli (1654-1705) se pose la question de savoir comment déterminer la valeur de la probabilité [9], de façon objective, dans le cas d'une urne de Bernoulli. Dans son ouvrage posthume *Ars conjectandi* (1713), il propose, dans le cas où l'on ne peut pas déterminer *a priori* par un moyen quelconque (par exemple par un argument de symétrie), d'évaluer la valeur de la chance que survienne l'une des issues (la probabilité de « succès ») à l'aide de la répétition de l'expérience, et cela d'autant mieux que le nombre de répétitions est grand. Bernoulli prouve cette convergence (théorème de Bernoulli), que nous appellerons dans un sens plus général, à partir de 1835, la loi des grands nombres. Il cherche ensuite à déterminer un majorant de l'erreur commise lorsque l'on évalue une probabilité inconnue par la fréquence observée après un grand nombre d'épreuves. Cependant, avec ses majorations trop grossières, il ne réussit pas à trouver une approximation de cette probabilité, pour n (assez grand) donné, qui lui convienne. Il arrive au résultat qu'il faut faire plus de 25 000 expériences pour que la fréquence soit suffisamment proche de la probabilité [10], ce qui n'est pas satisfaisant pour Bernoulli car, pour l'époque, il n'est pas

9. Le philosophe Antoine Arnauld (1612-1694) est le premier à introduire le terme de probabilité en 1662 dans *La logique ou l'art de penser*, terme qui sera repris par Jacques Bernoulli.

10. Il faut exactement 25 550 expériences pour approcher la probabilité à 2% près.

raisonnable de réaliser autant d'expériences. L'énoncé du théorème de Bernoulli, traduit en français contemporain dans Henry (2004), est le suivant :

> Théorème : Soit E une issue possible d'une épreuve de Bernoulli[11]. On suppose que la probabilité de E est $p = \frac{r}{t}$. Cette épreuve est répétée un nombre nt de fois. Soit F_{nt} la fréquence de réalisations de E au cours de ces nt épreuves. Alors, pour tout nombre c aussi grand que l'on veut, pour n assez grand, la probabilité que F_{nt} soit comprise entre $\dfrac{r-1}{t}$ et $\dfrac{r+1}{t}$ est plus grande que $\dfrac{c}{c+1}$. (p. 128)

En prenant r et t suffisamment grands tels que $\frac{r}{t} = p$, on obtient pour n assez grand, l'inégalité suivante : $P\left(p - \frac{1}{t} < F_{nt} < p + \frac{1}{t}\right) > \dfrac{c}{c+1}$.
En d'autres termes, F_{nt} converge en loi vers p.
En posant $\varepsilon = \frac{1}{t}$ et $\alpha = \dfrac{1}{c+1}$, l'inégalité devient :

$$P(F_{nt} - \varepsilon < p < F_{nt} + \varepsilon) > 1 - \alpha.$$

Ce qui revient à dire que $]F_{nt} - \varepsilon; F_{nt} + \varepsilon[$ est un intervalle de confiance de p au niveau $1 - \alpha$. Il s'agit du premier intervalle de confiance de l'histoire pour estimer une probabilité. La démonstration de ce théorème est étudiée dans Henry (2004).

Le problème de Bernoulli sera résolu par De Moivre (1667-1754), dans la 3e édition de sa *Doctrine of chances* (1756), dans laquelle il utilise des équivalents au lieu de majorations. De plus, grâce à la formule de Stirling[12], découverte par De Moivre et Stirling, la situation se débloque et l'on voit apparaître une fonction du type $x \longmapsto e^{-kx^2}$, qui nous fait penser bien évidemment aujourd'hui à la fonction de densité de la loi normale. En 1770, s'attaquant au même problème, Daniel Bernoulli (1700-1782) arrive, lui aussi, à une fonction de ce type et en donne une table succincte.

Laplace (1749-1827), dans sa *Théorie analytique des probabilités* (1812), complète les travaux de Bernoulli et De Moivre, en utilisant, outre les formules de Stirling et de Taylor, une formule qui ramène la somme d'une série à une intégrale. Nous n'avons plus seulement que la fréquence observable se situe dans un intervalle donné centré sur la probabilité (intervalle de fluctuation), mais aussi que la probabilité de succès se situe dans un intervalle donné centré sur la fréquence observable (intervalle de confiance) par « retournement de la formule ». Apparait dans la formule une intégrale du type $\displaystyle\int_0^T e^{-t^2}\,\mathrm{d}t$, dont on ne connait pas de primitive exprimable à l'aide des fonctions élémentaires, mais dont la valeur peut être approchée à l'aide de la table de Daniel Bernoulli. L'intégrale en question peut être vue actuellement, à constantes près, comme une probabilité associée à une variable aléatoire suivant une loi normale.

Nous pouvons voir, dans ces travaux, l'émergence d'une relation entre intégrale et probabilité. Cependant, bien que l'on voit apparaître la fonction de densité de la

11. L'épreuve de Bernoulli consiste en un tirage au hasard d'une boule dans une urne à deux couleurs, contenant t boules dont r blanches (*fertiles* selon Bernoulli) et s noires (*stériles*).

12. Rappelons la formule de Stirling : $n! \underset{n\to+\infty}{\sim} \sqrt{2\pi n}(\frac{n}{e})^n$.

loi normale (avec notre point de vue actuel), celle-ci n'est alors qu'une « simple » fonction, un outil de l'analyse, qui apparaît dans les calculs et n'est pas reliée à une loi de densité. La fonction associée à la courbe de Gauss est donc « apparue » chez De Moivre, puis chez Laplace, sans toutefois être tracée et sans qu'ils se rendent compte qu'elle définissait une sorte de distribution de probabilité idéale. Ces travaux ne débouchent pas sur la découverte de la notion générale de densité. Ce n'est que plus tard que le théorème de Moivre-Laplace sera reformulé en ces termes. Ce n'est donc pas de ces travaux qu'est née la notion de loi de probabilité à densité.

3.3.3 Buffon et le Franc-carreau : première rencontre avec la loi uniforme continue

Buffon (1707-1788), né sous le nom de Georges-Louis Leclerc, propose dans son *Mémoire sur le Jeu de Franc-Carreau* (1733), pour rentrer à l'Académie Royale des Sciences, le premier essai de calcul de probabilités dans un contexte géométrique. La règle du jeu du Franc-carreau, comme la présente Buffon, est la suivante :

> Voici un problème qui m'a occupé ces jours passés, et qui sera peut-être du goût de Mr de Moivre. Vous ne savez peut-être pas ce que nous appelons en français le jeu du franc-carreau. Dans une chambre pavée de carreaux, on jette en l'air un écu. S'il retombe sur un seul carreau, on dit qu'il tombe franc, et celui qui l'a jeté gagne. S'il tombe sur deux ou plusieurs carreaux, celui qui l'a jeté perd. C'est un problème à résoudre et qui n'a point de difficulté : trouver la probabilité de gagner ou de perdre, les carreaux et l'écu étant donnés. (Nowacki & Milliard, 2013)

Buffon est le premier à avoir recours bien qu'implicitement aux probabilités continues et plus particulièrement à la loi uniforme dans une partie bornée du plan. Il résout le problème avec des rapports d'aires pour calculer les probabilités. Buffon propose ensuite de s'intéresser au problème dit de l'aiguille. Il s'agit cette fois de jeter une baguette sur un sol parqueté, et le joueur parie qu'elle ne rencontrera pas les joints du parquet (Métin, 2000). Buffon met en œuvre un cas simple d'espace de probabilités bidimensionnel et continu, où rentre en considération, la longueur de l'aiguille et la largeur de latte fixées, l'angle géométrique formé entre l'aiguille et la rainure ainsi que la distance du centre de l'aiguille à la rainure la plus proche. On peut considérer que ces deux variables (aléatoires) suivent respectivement des lois uniformes continues sur $[0; \frac{\pi}{2}]$ et $[0; \frac{l}{2}]$ avec l la largeur d'une latte. Pour autant, ces hypothèses restent implicites dans les écrits de Buffon. Malgré des erreurs présentes dans son mémoire, il s'agit du premier exemple connu de probabilités à densité, et plus spécifiquement des lois uniformes continues. Comme le précise Métin (2000), « *Buffon y emploie les ressources du tout jeune calcul intégral, mais ne les maîtrisant probablement pas il a recours en final à une astuce faisant appel à l'aire d'une arche de cycloïde.* » (*ibid.*, p. 176).

Nous voyons dans ces travaux apparaître un lien entre le calcul de probabilités et le calcul intégral, par le biais de calcul d'aires. Nous rencontrons ici la loi uniforme continue. Cependant, la notion de densité n'apparaît pas encore.

3.3.4 Les travaux de Bayes : présence d'une courbe de densité

Dans son mémoire posthume intitulé *Essay*, publié en 1763 (traduit et commenté par Cléro (1988)), Bayes (1702-1761) se propose « *d'évaluer, au moyen de son dispositif ingénieux de boules lancées sur une table carrée et dont la position d'arrêt est au hasard sur la table, la probabilité que la première boule lancée se soit arrêtée dans une bande donnée (c'est-à-dire une probabilité inférieure et une probabilité supérieure), connaissant la position d'arrêt de boules lancées après* » (Pichard, 2005, p. 225). Sur le graphique suivant (figure 3.1), nous pouvons repérer le dessin de la table carrée $ABCD$ et en-dessous une courbe en forme de cloche. Si on considère que la boule initiale a une position d'arrêt d'abscisse x, connaissant les abscisses des positions d'arrêt des boules lancées après (p boules ayant une abscisse plus petite que x et q boules ayant une abscisse plus grande), alors l'abscisse de la première boule suit une loi représentée par la courbe d'équation $y = Kx^p(1-x)^q$, où K est un coefficient adéquat.

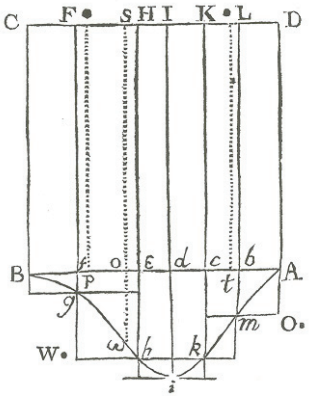


Figure 3.1 – Graphique de la loi de l'abscisse de la position d'arrêt de la première boule lancée (Bayes, 1763, dans Pichard, 2005, p. 226).

Il s'agit de la loi que nous appelons actuellement la loi Bêta de paramètre p et q. Euler (1707-1783) a étudié l'intégrale, appelée fonction Bêta, associée à cette distribution $B(p,q) = \displaystyle\int_0^1 x^{p-1}(1-x)^{q-1}\,\mathrm{d}x$ pour p et q strictement positifs.

Dans ces travaux, Bayes fait apparaître ce que nous pourrions appeler maintenant une courbe de densité, ou tout du moins une courbe qui s'en rapproche. Cependant, nous ne disposons pas de détails sur la construction de celle-ci. Nous repérons ici une première approche de la notion de densité, avec une place accordée au graphique.

3.3.5 La théorie des erreurs : naissance des lois à densité

Les premières distributions à densité sont apparues à la naissance de la théorie des erreurs. Au cours du siècle des Lumières, l'astronomie et la géodésie[13] sont en

13. La géodésie est la science destinée, à l'origine, à la mesure de la Terre et au tracé des cartes.

plein essor. Les instruments de mesure sont de plus en plus précis mais se pose la question des écarts entre les différentes mesures observées d'une même quantité. Armatte (2004) détaille les enjeux de ces erreurs. Les mathématiciens cherchent alors à déterminer la « vraie » ou sinon la « meilleure » valeur de la quantité mesurée et à trouver une « loi des erreurs ». Va alors naître la théorie des erreurs : il s'agit de la théorie mathématique des erreurs d'observation, qui permet, par exemple à l'astronomie de précision de pouvoir concilier les résultats des observations avec la mécanique céleste qui en rend compte.

Lorsque l'on fait une mesure d'une quantité à l'aide d'un instrument (optique, par exemple), même si on enlève les biais systématiques de l'instrument, il reste les erreurs inévitables, défauts d'attention, variations anarchiques de la réfraction atmosphérique, vibrations diverses, etc. Toutes ces causes accidentelles d'erreurs affectent la précision de la mesure. Si on mesure dix fois, cent fois la même quantité, nous trouverons à chaque fois des valeurs différentes. C'est même pire quand les observations sont indirectes (par exemple la masse de Jupiter) car elles ne donnent le résultat cherché que par l'intermédiaire d'équations mettant en jeu plusieurs mesures de natures différentes (Bru, 2006).

La théorie des erreurs cherche à modéliser la dispersion des erreurs de mesure d'une valeur qui doit être centrale et qui devrait être la valeur théorique. Dans son principe, elle consiste à décrire la loi de probabilité des erreurs (écarts entre la « vraie » valeur d'une quantité et les mesures fournissant des valeurs observées). Il ne s'agit plus ici d'une série de données de pile ou face mais d'une série de nombres tous à peu près égaux.

Un des précurseurs dans ce domaine est Galilée (1564-1643) qui, en 1632, s'intéresse à la détermination de la distance entre la Terre et une nouvelle étoile. Il dispose de 78 mesures d'élévation de cette étoile, fournies par 13 observateurs répartis à plusieurs endroits. Dans son écrit intitulé *Dialogo sopra i due massimi sistemi del mondo : Tolemaico e Copernicano* (1632), Galilée écrit :

> Étant donné que ces observateurs sont capables, qu'ils ont tous fait des erreurs, et que leurs erreurs doivent être corrigées pour nous permettre d'avoir la meilleure information possible à partir de ces observations, nous nous consacrerons à appliquer les modifications minimales et les corrections les plus petites possibles, juste assez pour ôter les observations de l'impossible et les remettre dans le possible... (dans Droesbeke & Tassi, 1990, p. 11)

a) Une approche purement théorique

En 1757, Simpson (1710-1761) explique que ce ne sont pas les observations qui importent mais les écarts entre ces observations et la « vraie » valeur, qui est inconnue. Il fait une analyse probabiliste de ces écarts, qu'il appelle erreurs (d'où le nom de « théorie des erreurs »), et suppose que les erreurs positives et négatives également distantes de la valeur théorique ont la même probabilité de survenir (Parzysz, 2013). Il suppose donc la symétrie de la distribution. La publication de Simpson est considérée comme la première portant sur les lois des erreurs. Il introduit d'abord la loi uniforme discrète, puis la loi triangulaire discrète des erreurs. À l'aide de la fonction génératrice (utilisée par De Moivre et théorisée par Laplace), Simpson obtient ensuite la loi de probabilité de la somme de n erreurs indépendantes. Par passage à

la limite, il introduit la distribution de la moyenne arithmétique d'erreurs, suivant chacune une loi triangulaire continue définie par la fonction de facilité (que nous appelons maintenant fonction de densité) : $f(x) = \dfrac{a - |x|}{a^2}$, avec $-a \leqslant x \leqslant a$. Le fait d'introduire la moyenne arithmétique est très courant pour réduire les imprécisions des mesures.

À propos de la figure 3.2, Simpson écrit :

> Let, then, the line AB represent the whole extent of the given interval, within which all the observations are supposed to fall ; and conceive the same to be divided into an exceeding great number of very small, equal particles, by perpendiculars terminating in the sides AD, BD of an isosceles triangle ABD formed upon the base AB : and let the probability or chance whereby the result of any observation tends to fall within any of these very small intervals Nn, be proportional to the corresponding area $NMmn$, or to the perpendicular NM ; then, since these chances (or areas) reckoning from the extremes A and B, increase according to the terms of the arithmetical progression 1, 2, 3, 4, &c. it is evident that the case is here the same with that in the latter part of Prop. II [where the discrete triangular distribution was studied] ; only, as the number v (expressing the particles in AC or BC) is indefinitely great, all (finite) quantities joined to v, or its multiples, with the signs of addition or substraction, will here vanish, as being nothing in comparison to v. (Simpson, 1757, dans Stigler, 1986, p. 96)

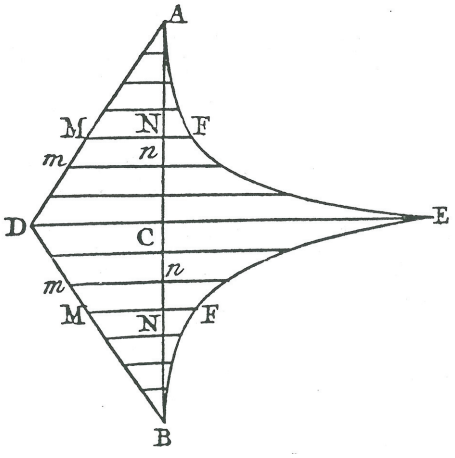


Figure 3.2 – Graphique de la loi d'erreurs triangulaire ADB et de la loi de la moyenne des erreurs AFEFB (Simpson, 1757, dans Stigler, 1986, p. 96).

Nous voyons apparaître ici une courbe de densité, bien qu'elle ne soit pas sous la forme actuelle (la théorie des fonctions est encore en évolution à cette époque), et la citation mentionne bien le lien entre probabilité et aire sous la courbe : « *la probabilité (ou la chance) que le résultat des observations tende à tomber dans l'un des petits intervalles Nn est proportionnelle à l'aire correspondante $NMmn$* » [14].

14. Traduit par nos soins.

En parallèle, en 1757, Boscovic (1711-1787) écrit que la meilleure estimation est celle pour laquelle :

– la somme des écarts positifs est égale à celle des écarts négatifs ;
– la somme des valeurs absolues des écarts est aussi petite que possible.

Dans ce cas, la « vraie » valeur n'est pas la moyenne arithmétique mais la médiane de la série de mesures.

En 1760, Lambert (1728-1777) consacre une partie de son traité *Photometria sive de mensura* (1760) à l'étude de courbes de facilité continues, symétriques et unimodales, dans laquelle il introduit la méthode du maximum de vraisemblance (qui ne porte alors pas encore ce nom) en recherchant la position du centre d'une distribution continue symétrique. On voit apparaître dans ce traité une courbe (figure 3.3). Lambert explique :

> Soit AC la valeur vraie à déterminer au moyen d'expériences, soit CB et CD les valeurs maximales de part et d'autre. Que les ordonnées QM, PN, RL, SK de la courbe $BMLD$ représente les retours des erreurs CQ, CP, CR, CS, à savoir leurs fréquences. Nous appellerons ici celles-ci fréquences absolues ou vraies pour les distinguer des fréquences observées. Il est clair que celles-ci coïncident avec celles-là si l'expérience est infiniment répétée. Mais puisque cela n'arrive jamais, il faut découvrir ce qui demeure pour un nombre fini d'expériences (Lambert, 1760, dans Lefort & Boyé, 1996, p. 244)

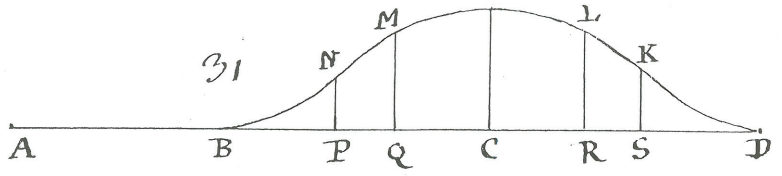


FIGURE 3.3 – Graphique d'une loi de facilité des erreurs (Lambert, 1760, dans Lefort & Boyé, 1996).

Il ne s'agit pas exactement d'une courbe de densité. Lambert précise que l'ordonnée correspond à la fréquence. Il s'agit finalement d'un diagramme en bâtons des fréquences « *vraies* » dont les sommets sont ensuite reliés par une courbe continue. Lambert propose ensuite, en 1765, une nouvelle loi continue semi-circulaire.

En 1768, Lagrange (1736-1813) rédige un *Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations ; dans lequel on examine les avantages de cette méthode par le calcul des probabilités ; et où l'on résout différents problèmes relatifs à cette matière*. Dans ce mémoire, on retrouve les résultats de Simpson, ainsi que de nouvelles lois continues : la loi uniforme sur un segment, la distribution parabolique et la distribution sinusoïdale. Jusque là, les mathématiciens considèrent que les erreurs doivent être bornées : une valeur trop éloignée des autres sera éliminée.

Laplace est le premier à considérer une loi de facilité sur un domaine infini. En 1774, il indique quelques conditions que doit vérifier la courbe de facilité des erreurs

(Parzysz, 2013) :

> La loi suivant laquelle cette vraisemblance [probabilité d'apparition] diminue à mesure que l'observation s'éloigne de la vérité nous est inconnue. Supposons donc (fig. 2) [il s'agit ici de la figure 3.4] que le point V soit le véritable instant du phénomène [il s'agit d'une mesure de temps] [...] en nommant x l'abscisse VP, et y l'ordonnée correspondante PM, nous représenterons l'équation par celle-ci : $y = \phi(x)$.

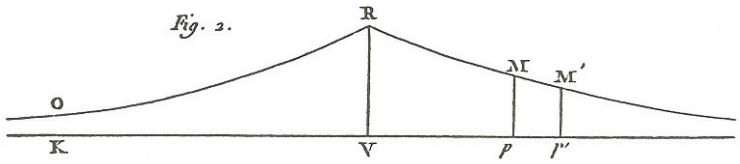


Figure 3.4 – Courbe de la première loi de Laplace (Laplace, 1774, dans Stigler, 1986, p. 107)

> Or voici les propriétés de cette courbe [figure 3.4] :
>
> 1. elle doit être partagée en deux parties entièrement semblables par la droite VR, car il est tout aussi probable que l'observation s'écartera de la vérité à droite comme à gauche ;
> 2. elle doit avoir pour asymptote la ligne KP, parce que la probabilité que l'observation s'éloigne de la vérité à une distance infinie est nulle ;
> 3. l'aire entière de cette courbe doit être égale à l'unité, puisqu'il est certain que l'observateur tombera sur un des points de la droite KP.

Laplace fait à nouveau apparaître, après Simpson, le lien entre la probabilité et l'aire sous la courbe. On voit ici émerger des contraintes sur la fonction de densité : la symétrie de la fonction par rapport à la droite verticale de la valeur « réelle », la limite de la fonction en $+\infty$ est nulle et l'aire sous la courbe vaut 1. On retrouve, dans les deux derniers points, les propriétés qu'une fonction de densité en général doit vérifier. La propriété de la positivité de la fonction reste implicite.

Laplace est le premier à étendre l'intervalle des possibles à $\mathbb{R}$, en introduisant la première loi de Laplace, aussi appelée loi double exponentielle définie par la fonction $f(x) = \frac{k}{2}e^{-k|x|}$.

Dans ses travaux ultérieurs, il rencontre fréquemment l'intégrale $\int e^{-t^2}\,\mathrm{d}t$, qui lui semble importante et explique qu'il serait utile d'en faire une table sur $]-\infty;0]$ (Parzysz, 2013). Mais, il faut attendre Gauss (1777-1855), en 1809, pour introduire la loi de Gauss, aujourd'hui appelée loi normale. Elle est introduite en recourant à la méthode des moindres carrés. Gauss reprend, à ce moment-là, le problème « d'ajustement ». Son approche est probabiliste, ce qui est novateur pour l'époque. La démarche qu'il met en place est détaillée dans Droesbeke et Tassi (1990, p. 32-33) ou dans Parzysz (2013).

À la même époque, en 1810, Laplace introduit la « seconde loi des erreurs », qui n'est autre que la loi de Gauss : il montre que la distribution de la moyenne arithmétique de n erreurs d'une même quantité, indépendantes, peut être approchée

par une distribution d'une loi normale, avec une erreur d'approximation de plus en plus faible lorsque n tend vers l'infini.

En 1824, Poisson considère la loi définie sur $\mathbb{R}$ par la fonction $f(x) = \dfrac{1}{\pi(1+x^2)}$, que nous appelons aujourd'hui loi de Cauchy.

De la théorie des erreurs émergent les premières lois à densité (toutes symétriques) et la notion de fonction de densité (qui ne porte pas encore ce nom) et notamment ses propriétés. Nous voyons apparaître une relation entre les probabilités et la statistique. La courbe de facilité représente alors la fréquence « théorique ».

Nous avons fait ici un aperçu (non exhaustif) des mathématiciens de l'époque ayant travaillé sur la théorie des erreurs. Cependant, nous pouvons aussi rencontrer des lois à densité du côté des astronomes et géodésiens de terrain.

b) La vérification empirique

Dans les années 1830, astronomes et géodésiens ont commencé à dresser des tableaux statistiques pour rassembler toutes leurs données, ce qui a permis de confronter les données expérimentales aux modèles théoriques. Selon Bru (2006), un des premiers à l'avoir fait serait Bessel (1784-1846) (cf. Annexe A.1). Bru (*ibid.*) précise que le tableau de fréquences associé ressemble fortement aux tables de la courbe de Gauss.

Dans le cours de *Calcul des probabilités et Théorie des erreurs* (1852), Liagre présente des données des travaux de Bessel et compare les valeurs d'erreurs observées et les valeurs théoriques (cf. figure 3.5), à l'aide de la table de l'intégrale $\dfrac{2}{\sqrt{\pi}}\displaystyle\int_0^t e^{-t^2}\,\mathrm{d}t$ (cf. Annexe A.2). Les valeurs sont très proches, ce qui valide le choix de la loi de Gauss. Il en est ainsi car les phénomènes étudiés sont de type gaussien (il s'agit de la somme de multitudes d'interventions du hasard indépendantes).

101. — Comme exemple de l'accord qui règne entre la théorie et l'expérience, nous choisirons une série de 470 observations faites par Bradley, le modèle des observateurs suivant l'expression de Laplace, pour déterminer directement la différence d'ascension droite entre le soleil et l'une des deux étoiles Ataïr et Procyon. Bessel, qui a discuté ce sujet, a trouvé que l'erreur probable $r = 0'',2637$; d'où $0'',1 = 0{,}3792\, r = 0{,}3792$, en prenant r pour unité. Cherchant ensuite le nombre d'erreurs qui, suivant la théorie, doivent tomber entre $0''.0$ et $0'',1$; $0'',2$; $0'',3$..., il a calculé, au moyen de la table mentionnée au § 99 (table n° 3), la valeur de l'intégrale P_z pour ces diverses limites *. On trouve ainsi :

Pour	0″,1 . . .	0,3792	le nombre 0,20186,
»	0 ,2 . . .	0,7584	» 0,39402,
»	0 ,3 . . .	1,1376	» 0,55705,
»	0 ,4 . . .	1,5168	» 0,69372,
»	0 ,5 . . .	1,8960	» 0,79904.
»	0 .6 . . .	2,2752	» 0,87511.
»	0 ,7 . . .	2,6544	» 0,92661.
»	0 ,8 . . .	3,0336	» 0,95926,
»	0 ,9 . . .	3,4128	» 0,97866,
»	1 ,0 . . .	3,7920	» 0,98983.
	. . .	∞	» 1,00000.

Si l'on retranche chaque nombre du suivant, et qu'on multiplie le reste par le nombre des observations, ou par 470, on trouve

ENTRE	NOMBRE DES ERREURS.	D'APRÈS la théorie.	D'APRÈS l'observation.
0″,0 et 0″,1	0,20186	95	94
0 ,1 0 ,2	0,18916	89	88
0 ,2 0 ,3	0,16603	78	78
0 ,3 0 ,4	0,13667	64	58
0 ,4 0 ,5	0,10532	50	51
0 ,5 0 .6	0,07607	36	36
0 .6 0 ,7	0.05150	24	26
0 ,7 0 ,8	0,03265	15	14
0 ,8 0 ,9	0,01940	9	10
0 ,9 1 ,0	0,01117	5	7
au delà de 1 ,0	0,01017	5	8

Figure 3.5 – Distribution des erreurs d'observation de la différence d'ascension droite entre le soleil et l'une des deux étoiles Ataïr et Procyon, données empiriques et données théoriques (Liagre, 1852)

À la suite, Liagre justifie le fait de considérer une courbe de facilité définie sur $\mathbb{R}$ et non sur un intervalle borné :

> Il arrive fréquemment, comme ici, que l'observation présente *plus* de grandes erreurs que la théorie n'en prévoit. Cela prouve la légitimité de l'hypothèse que nous avons faite [...] en intégrant le dénominateur entre les limites $-\infty$ et $+\infty$, au lieu de l'intégrer entre des limites finies ; car *le contraire* aurait lieu si l'inexactitude de notre hypothèse devait avoir quelque influence dans la pratique. (Liagre, 1852)

Bien que les tableaux de fréquences soient de plus en plus courants, il faut attendre encore un peu pour voir apparaître des courbes de fréquences, maintenant appelées histogrammes [15]. Elles deviendront systématiques seulement à la fin du XIX^{e} siècle. Quetelet (1796-1855) montre que la théorie des erreurs peut s'appliquer

15. Le mot anglais « histogram » date de 1895 et sa traduction française normalisée, de 1948.

à d'autres domaines que les erreurs de mesures, et notamment il repère que les tours de poitrine des soldats écossais sont régies par la loi de Gauss. Pour lui, c'est aussi le cas pour différentes mesures de caractéristiques humaines (taille, tour de poitrine, taille de crâne,...) et pour les phénomènes en science morale (ou physique sociale). Pour des données d'observation, Quetelet utilisait l'ajustement à une courbe des erreurs comme un test sur l'homogénéité de la population. Si l'histogramme construit sur les observations n'a pas la forme d'une courbe en cloche, cela montre une hétérogénéité, à déterminer. Bertillon (1821-1883) offre un exemple frappant sur les tailles des conscrits du département du Doubs en 1863 avec une courbe à double bosse (figure 3.6). L'hétérogénéité est expliquée ici par un mélange de types de personnes différents.

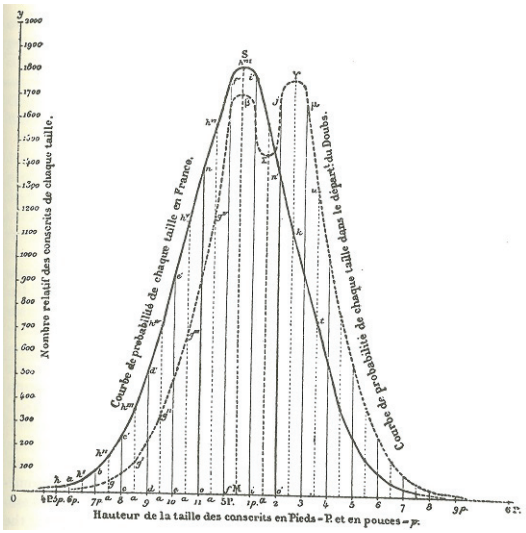


FIGURE 3.6 – Courbe des tailles des conscrits du département du Doubs en 1863 (Bertillon, 1876, dans Stigler, 1986, p. 217)

Dans son essai pour donner un fondement expérimental à la théorie de l'évolution de Darwin, Galton (1822-1911) utilise la méthodologie de Quetelet, en particulier celle d'ajustement de données à une courbe des erreurs (Pichard, 2005). Pour Quetelet, une persistance des causes se traduit par une tendance pour que la moyenne soit stable et inversement une modification de la moyenne indique une variation des causes, une évolution. Dans le cas de nombreuses petites causes indépendantes et de même influence, la résultante suit une loi des erreurs, comme par exemple pour la planche de Galton. Ceci sera confirmé par l'énoncé du théorème limite central. Les diverses expérimentations menées par Galton en anthropométrie, par le biologiste Weldon (1860-1906) sur les crevettes, crabes... et par d'autres ont permis de montrer que les distributions de fréquences des mesures effectuées dans les différents phénomènes ne suivent pas toujours une courbe des erreurs.

Pearson (1857-1936) cherche à caractériser la forme de différentes distributions empiriques. Il propose ensuite, dans les années 1893-1895, un système de distributions théoriques :

> c'est-à-dire de fonctions qui sont des densités de probabilité et vérifient l'équa-

> tion différentielle $\frac{dy}{dx} = \frac{y(x+a)}{b_0 + b_1 x + b_2 x^2}$ sur un certain domaine D de valeurs de x. Les valeurs des paramètres a, b_0, b_1, $b-2$ sont déterminées par les moments d'ordre 1 à 4 de la loi ayant pour densité la fonction y. Suivant les valeurs des paramètres a, b_0, b_1, $b-2$, Pearson obtient différents types de distribution ; par exemple :
>
> – *si* $b_0 \neq 0$, $b_1 = b_2 = 0$, *les solutions de (*) donnent les lois normales* $N(\mu, \sigma^2)$ *avec* $b_0 = -\sigma^2$ *et* $a = -\mu$ *;*
> – *si* $b_1 \neq 0$, $b_2 = 0$, *les lois exponentielles et Gamma ;*
> – *si* $b_0 \neq 0$, $b_2 \neq 0$ *tels que le trinôme* $b_0 + b_1 x + b_2 x^2$ *est irréductible sur* $\mathbb{R}$*, les fonctions puissance de ce trinôme, dont les lois de Cauchy et de Student ; si ce trinôme est réductible sur* $\mathbb{R}$*, les loi Bêta ; etc.*
>
> (Pichard, 2005, p. 242)

Il fait donc apparaître de nouvelles lois d'un point de vue théorique, mais que l'on retrouve ensuite dans la pratique : la loi de Pareto fait son apparition pour les distributions de revenus, la loi exponentielle pour les durées de vie... D'autres lois discrètes font aussi leur apparition telles que la loi de Poisson pour modéliser les évènements rares, la loi géométrique...

De nos jours, ces différentes lois sont toujours utilisées pour modéliser divers phénomènes. Elles permettent notamment de faire des estimations.

Dans ces travaux, nous voyons l'importance de la prise en compte de la distribution empirique pour vérifier la pertinence de la distribution théorique choisie. Effectivement, la loi des erreurs ou les autres lois ensuite sont là pour répondre à un problème concret d'où l'importance de s'assurer de la correspondance entre données empiriques et données théoriques. Nous voyons apparaître des courbes de fréquences (histogrammes) pour permettre ces vérifications. Nous remarquons qu'émergent aussi de nouvelles courbes non symétriques mais qui conservent les autres propriétés, notamment d'aire sous la courbe égale à 1. Ces nouvelles courbes permettent de résoudre d'autres problèmes que ceux liés aux erreurs de mesures.

c) La détermination empirique de lois : usage des histogrammes

Une fois les résultats théoriques vérifiés empiriquement, les résultats mathématiques sont diffusés à la communauté des astronomes et des géodésiens afin qu'ils puissent les appliquer à leurs données. Dans son *Cours d'astronomie et de géodésie* de 1881, Faye décide de présenter les choses différemment de celles exposées dans les cours de probabilités. Il explique son point de vue :

> Nous allons appliquer ici la méthode expérimentale, c'est-à-dire chercher par expérience, s'il existe une relation entre la probabilité d'une erreur et la grandeur de cette erreur, puis opérant par des mesures réellement faites, sur des écarts dûment constatés, nous tâcherons de découvrir graphiquement cette loi (Faye, 1928, p. 271)

Il prend le même exemple que Liagre (figure 3.5). La démarche n'est pas de comparer les valeurs empiriques et théoriques, mais de partir des valeurs observées pour déterminer empiriquement la loi que suivent les erreurs. Il s'agit de problèmes d'ajustement, qui restent très importants et d'actualité. Il propose donc de représenter graphiquement les données sous forme d'histogramme (il explique comment le tra-

cer) puis de tracer une courbe qui laisse autant de triangles au-dessus qu'en-dessous de la courbe (cf. figure 3.7).

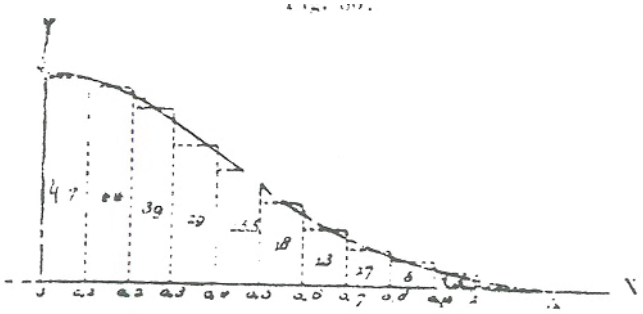


FIGURE 3.7 – Détermination empirique d'une loi, à l'aide d'une construction graphique (Faye, 1928)

L'expression de la courbe est du type $y = ae^{-h^2x^2}$, où a et h sont des paramètres à déterminer. Il justifie aussi le fait de considérer une fonction strictement positive sur $\mathbb{R}$, bien qu'elle devrait être nulle en dehors des valeurs extrêmes, en expliquant qu'au-delà de ces abscisses, l'aire sous la courbe est négligeable donc que cela ne pose pas de problème. Il donne ensuite un second exemple, non issu de la théorie des erreurs, tiré du *Traité des probabilités appliquées au tir* de Didion de 1858, où l'on retrouve à nouveau une courbe de la même famille. Tout cet extrait se trouve en annexe A.3.

On voit ici apparaître une autre démarche, proposant de partir des données statistiques pour aboutir aux probabilités. Il s'agit d'une approche didactique que Faye choisit dans son cours à des non mathématiciens. Il y a ici clairement des histogrammes construits avec les données expérimentales qui permettent la construction de la courbe de densité. Il ne s'agit plus d'une comparaison entre les deux mais d'un ajustement. La place du graphique est très importante. Mais les résultats s'appuient ensuite sur des calculs, prenant en compte les propriétés de la fonction de densité.

3.3.6 La théorie moderne des probabilités

Pour présenter brièvement la théorie moderne des probabilités, les fondements sont dûs à Kolmogorov (1933) qui développe une théorie axiomatique s'inspirant du développant de la théorie de la mesure de Borel et de la théorie de l'intégration de Lebesgue. Les concepts de mesure et d'intégration ne se limitent plus à $\mathbb{R}$ et $\mathbb{R}^n$ mais laissent la place à la notion de mesure abstraite définie sur un ensemble Ω quelconque muni d'une tribu. C'est justement cette notion qui permet le développement de la théorie mathématique des probabilités, la probabilité étant considérée comme une mesure de masse totale 1.

On définit notamment une mesure de probabilité et la loi de probabilité d'une variable aléatoire comme suit :

Définition. Soit $(\Omega, \mathcal{A})$ un espace mesurable[16]. $\mathbb{P}$ est une *mesure de*

16. Pour plus de détails théoriques, se reporter à Ouvrard (2009).

probabilité si :

1. $\mathbb{P}$ est une application de $\mathcal{A}$ dans $[0;1]$;
2. $\mathbb{P}(\Omega) = 1$;
3. $\mathbb{P}$ est σ-additive, c'est-à-dire pour toute famille finie ou dénombrable d'éléments disjoints $(A_i, i \in I)$ de $\mathcal{A}$:

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} \mathbb{P}(A_i)$$

.

Définition. Soit X une variable aléatoire réelle sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, c'est-à-dire une fonction mesurable $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. La loi de probabilité de la variable aléatoire X est la mesure de probabilité, notée $\mathbb{P}_X$, définie sur l'espace mesurable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ par :

$$\mathbb{P}_X(B) = \mathbb{P}\Big(X^{-1}(B)\Big) = \mathbb{P}(X \in B),$$

pour tout borélien réel $B \in \mathcal{B}(\mathbb{R})$. Autrement dit, $\mathbb{P}_X$ est la mesure image de $\mathbb{P}$ par X.

Cette axiomatique a permis à Kolmogorov de prouver de nombreux résultats relatifs aux probabilités, jusqu'ici non démontrés. La théorie des probabilités de Kolmogorov marque l'aboutissement de l'histoire des probabilités au XX$^{\text{e}}$ siècle. Cette théorie mettra en avant la fonction de répartition plutôt que la densité, comme outil indispensable dans les démonstrations théoriques où apparaîtront des fonctions de répartition singulières, non dérivables presque partout. La théorie de Kolmogorov permet de présenter avec rigueur la théorie des probabilités, avec un niveau d'abstraction très élevé, loin des problèmes de la réalité qui ont guidé les premiers travaux dans ce domaine.

À ce niveau supérieur, la mise en relation entre probabilités et analyse ne se fait pas au niveau de la fonction de densité, mais au niveau de la mesure. Ce niveau d'abstraction n'est bien entendu pas à la portée des élèves de terminale.

3.4 Conclusion de l'étude historico-épistémoloqique

Rappelons la question de recherche QR1, qui a guidé cette étude :

> **QR1** : Comment historiquement et épistémologiquement sont apparus les liens entre probabilités à densité et calcul intégral ?
> Comment a émergé le concept de lois à densité et tout particulièrement la notion de fonction de densité ?

Dans ce chapitre, nous avons essayé de mettre en évidence des éléments historiques importants permettant de comprendre l'émergence épistémologique de ces

notions. Nous allons présenter les éléments qui en ressortent, puis nous dégagerons les apports de cette étude pour la suite de notre travail.

L'introduction des lois à densité est liée aux progrès de l'intégration et plus généralement de la théorie des fonctions. Cependant, elle n'est pas venue de l'analyse mais de problèmes concrets traités avec une approche probabiliste. La relation entre probabilités et analyse vient donc des probabilités. Les lois continues émergent en réponse à des problèmes voulant décrire les erreurs de mesures et non à des problèmes liés à des jeux de hasard. C'est donc bien pour répondre à des problèmes concrets et pour interpréter des données statistiques sur des mesures, qu'elles font leur apparition. Cependant, dans un premier temps, l'approche est purement théorique. Les mathématiciens font des hypothèses en amont sur les distributions statistiques dont ils disposent qui leur permettent de proposer des lois, dont la courbe de facilité vérifie certaines propriétés : notamment que l'aire sous la courbe doit valoir 1. La courbe la plus rencontrée à l'époque est la courbe de Gauss qui ensuite sera jugée adéquate pour la loi des erreurs mais aussi pour modéliser de nombreux autres phénomènes. Ce choix de courbe est confirmé en confrontant la théorie avec les données empiriques.

Pour l'enseignement des probabilités à des astronomes et géodésiens, Faye expose une approche différente dans son cours. Il propose une démarche inverse : le but est alors de partir des données, de passer par la représentation sous forme d'histogramme des fréquences associées aux données, pour ensuite trouver une courbe qui « épouse » bien l'histogramme. Une place importante est laissée aux graphiques, dans ce cas. Cette approche par les histogrammes fait apparaître le lien entre probabilité et calcul d'aire, et donc entre probabilités et calcul intégral. Nous pouvons repérer un rôle clé de la statistique dans la découverte des lois à densité : dans un premier temps, ce sont les données statistiques qui sont à l'origine de la recherche de modèles, et ensuite ces données ont permis de valider les modèles. Nous voyons donc apparaître une connexion forte entre statistique et probabilités, qui permet qui plus est la justification du lien avec l'aire et donc le calcul intégral, avec notamment le passage par les histogrammes de fréquences.

Le théorème de Moivre-Laplace a une place importante dans l'histoire des probabilités mais il n'est pas à l'origine des lois à densité, pas même de la loi normale. L'expression de la fonction de densité de la loi normale est apparue dans la résolution du « problème » de Bernoulli par De Moivre et Laplace, mais la loi normale n'existe pas encore en tant que telle à l'époque. La notion de lois à densité doit sa naissance à la théorie des erreurs, bien que cela ne soit pas explicité comme tel.

Cette étude historico-épistémologique apporte bien plus à notre recherche que des détails historiques intéressants. Nous pouvons retenir les points suivants :

- La fonction de densité est apparue bien avant la fonction de répartition. L'approche du programme par la fonction de densité est plus conforme à l'approche historique. L'approche par la fonction de répartition, que l'on peut retrouver dans l'enseignement supérieur, est une autre approche, apparue plus tardivement.
- La statistique descriptive a joué un rôle important dans l'émergence de la

notion de fonction de densité. La statistique (via les données empiriques) est à l'origine de l'introduction des lois à densité. Elle a ensuite permis de valider des modèles en permettant de confronter réalité et modèle.
– Toujours en lien avec la statistique, Faye a proposé dans son cours des problèmes d'ajustement, par le biais d'histogrammes de fréquence, pour amener à la construction de fonction de densité. Ce passage par le graphique histogramme/courbe nous semble intéressant notamment pour faire émerger le rôle de l'aire sous la courbe. Ce passage par l'histogramme nous semble donc une idée intéressante à retenir pour la suite.

La statistique descriptive semble donc avoir été un bon levier pour permettre la connexion entre les probabilités à densité et le calcul intégral et notamment à travers la représentation graphique.

Nous retrouvons à travers cette étude épistémologique et historique des points qui ont été jugés importants pour la compréhension notamment de la loi normale par des chercheurs en didactique, comme l'importance du lien entre distribution empirique et distribution théorique ou encore le rôle de l'histogramme (cf. section 2.2.1). Il nous semble donc intéressant dans la suite de se demander si dans l'ETM de référence, mais aussi dans les ETM idoines, est prise en compte cette approche historico-épistémologique, ou tout du moins si la statistique est présente dans l'introduction de la fonction de densité. Est-ce que la statistique descriptive joue un rôle dans la construction de la notion de densité dans les classes ?

Deuxième partie

Un regard sur les articulations entre probabilités à densité et calcul intégral dans l'enseignement en terminale S

Nous venons, dans la Partie I, d'exposer un point de vue historique et épistémologique sur la naissance des lois à densité. Cette étude nous a éclairée sur notre première question de recherche et nous a aussi donné des pistes à explorer, notamment du côté de la place accordée à la statistique dans l'introduction de la notion de fonction de densité dans la classe de terminale S.

Dans cette deuxième partie, nous allons chercher à répondre à la deuxième question de recherche. Il s'agit de la question suivante :

QR2 : En terminale S, quels sont les enjeux didactiques derrière la relation sémiotique

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x \ ? \qquad (*)$$

Quelles sont les tâches proposées et les activités attendues des élèves pour l'introduction de la notion de fonction de densité ?
Quelles sont les tâches qui permettent d'exploiter les articulations entre l'$\mathrm{ETM}_{\mathrm{PaD}}$ et l'$\mathrm{ETM}_{\mathrm{CI}}$?

L'étude historico-épistémologique (Partie I) ayant mis en évidence l'importance de la statistique descriptive dans la genèse des lois à densité et de la notion de fonction de densité, nous allons aussi nous demander :

Est-ce que la statistique descriptive joue un rôle dans la construction de la notion de densité dans les classes ?

L'objectif principal de la thèse étant de concevoir et mettre en place des tâches permettant de construire la notion de fonction de densité, il est, avant tout, nécessaire de mettre en évidence les enjeux didactiques qui existent autour de cette notion et de la relation sémiotique $(*)$ qui en découle. Ce travail nous permettra aussi de voir si les tâches d'introduction proposées dans l'enseignement construisent réellement la notion de fonction de densité et si non, de dégager des pistes qui pourraient aider à la conception de telles tâches. Étendre l'étude aux autres tâches qui exploitent la relation entre probabilités à densité et calcul intégral va nous permettre d'avoir un regard plus complet sur la notion, qui permettra de dégager d'autres points importants à prendre en compte.

Nous allons, tout d'abord, étudier ces questions au niveau de l'ETM de référence. Nous ferons une étude des documents institutionnels : le programme actuel, le document *Ressources* qui accompagne le programme en probabilités et statistique ainsi que les sujets de baccalauréat. Cette première étude fera l'objet du **chapitre 4**.

Ensuite, dans le **chapitre 5**, nous nous poserons la question de ce qui est fait de cet ETM de référence dans les classes. Pour cela, nous analyserons les manuels de mathématiques de terminale S. Nous complèterons cette analyse par des entretiens avec quatre enseignants, qui montreront que leurs choix de tâches sont similaires à ceux des manuels.

Chapitre 4

Étude des documents institutionnels

Pour aborder notre deuxième question de recherche, comme nous venons de le présenter dans l'introduction de cette Partie II, nous allons, dans ce chapitre, limiter notre étude au niveau de l'ETM de référence (institutionnelle). Il s'agit de décrire le paysage dans lequel les enseignants s'inscrivent pour construire leur propre ETM idoine. Le but est donc de caractériser l'ETM de référence relatif aux probabilités à densité (ETM_{PaD}) et celui relatif au calcul intégral (ETM_{CI}), et surtout d'en dégager les interactions établies entre ces deux ETM. Comme nous l'avons précisé dans le chapitre 2, la description d'ETM de référence ne peut être que partielle car les documents officiels qui permettent de le décrire ne donnent que des informations générales ou ponctuelles. Nous pouvons toucher du doigt ce qui est ou peut être attendu de l'institution mais sans pouvoir toujours entrer dans les détails précis.

Les objectifs de ce chapitre sont donc, à travers les documents institutionnels, de :

- Identifier, au niveau de l'ETM de référence, les articulations existantes entre les deux sous-domaines essentiellement au niveau du plan épistémologique.
- Repérer les tâches proposées pour l'introduction de la notion de densité.
- Repérer les tâches proposées exploitant le lien entre les probabilités à densité et le calcul intégral.

Dans la section 4.1, nous présenterons la méthodologie que nous avons choisie pour cette étude, ensuite successivement nous analyserons le programme (section 4.2), puis le document *Ressources* de la classe de terminale S en probabilités et statistique (section 4.3) et enfin, les sujets du baccalauréat (section 4.4).

4.1 Méthodologie de l'étude des documents officiels

4.1.1 Le choix des documents

Les documents qui nous intéressent pour décrire les ETM de référence sont ceux mis à disposition directement par le ministère de l'Éducation nationale, en lien avec

le programme et ce qui en est attendu. Les documents officiels que nous allons étudier sont les suivants :

- le programme de la classe de terminale S de 2011 (MEN, 2011), appliqué dès la rentrée 2012, qui est le programme actuellement en vigueur ;
- le document *Ressources pour la classe terminale générale et technologique* destiné aux probabilités et à la statistique (MENJVA & DGESCO, 2012) qui accompagne ce programme ;
- les exercices portant sur les lois à densité ou le calcul intégral des sujets de baccalauréat des années 2013, 2014 et 2015 des différents centres d'examen.

Les documents *Ressources* accompagnent le programme, le précisent, en donnent des compléments et font aussi des propositions. Dans les *Ressources* proposées sur le site Eduscol, figurent aussi des exercices pour la classe de terminale générale et technologique des filières S, ES, STI2D, STMG (MENESR & DGESCO, 2014). Nous ne les étudierons pas car ce sont des exercices travaillant la différentiation, mais ne proposant pas des exercices « nouveaux » sur le contenu mathématique, par rapport aux sujets du baccalauréat, par exemple.

Aucun document *Ressources* n'est dédié à l'analyse en terminale, *a fortiori* nous ne trouvons aucun élément sur le chapitre calcul intégral (cela est sans doute lié au fait que l'enseignement de l'analyse en terminale S, et donc du calcul intégral, en vigueur depuis plusieurs réformes, est assez stabilisé, malgré quelques suppressions et quelques modifications dans l'ordre d'apparition des notions, par exemple). De ce fait, de nombreuses ressources sont déjà à disposition des enseignants, ne serait-ce que les manuels scolaires.

La partie sur les probabilités à densité étant composée de notions nouvelles, le document *Ressources* permet de donner une ligne directrice, qui sert aux enseignants mais aussi aux concepteurs de manuels.

Trois types de documents se dessinent donc pour décrire ces ETM de référence : le programme qui est une directive ministérielle à respecter, les documents *Ressources* qui sont là dans le but de préciser le programme et de donner des idées, et enfin les sujets de baccalauréat, qui regroupent les types de tâches que les élèves peuvent rencontrer en classe. Nous faisons le choix de mettre les sujets du baccalauréat dans les documents institutionnels car ce sont des tâches validées par l'institution.

Nous compléterons parfois nos analyses avec le programme de 2001 (MEN, 2001) et le document d'accompagnement de ce programme (MJENR, 2002).

Dans ces différents documents, nous nous occuperons bien entendu en particulier des parties concernant le calcul intégral et les probabilités à densité.

4.1.2 Méthode d'analyse des documents

L'analyse de ces documents ne se fait pas linéairement et nécessite des axes clairs. Un ETM de référence existe pour chacun des deux sous-domaines. Il ne s'agit pas de les décrire individuellement mais dans un souci de cohérence permettant de dégager les liens existants ou possibles entre les deux.

On repérera les endroits où apparaissent des liens entre deux sous-domaines : des liens explicites ou implicites ou simplement potentiels. Le cadre des ETM, avec

la structure de modèle qui invite à regarder suivant trois dimensions (sémiotique, instrumentale et discursive) va guider notre analyse.

Les trois types de documents ne vont pas répondre aux mêmes questions et ne vont pas nécessiter le même type d'analyse. Le programme et le document *Ressources* sur les probabilités et la statistique vont plutôt nous informer sur le point de vue qui doit être adopté sur l'enseignement des probabilités à densité, et donc de la façon de mettre en évidence ses liens avec le calcul intégral, au niveau du plan épistémologique. Tandis que les sujets de baccalauréat – mais aussi le document *Ressources* – vont plutôt nous éclairer sur les tâches proposées dans ces sous-domaines. Nous ajouterons quelques précisions sur l'analyse des différents documents quand cela sera nécessaire.

4.2 Analyse du programme

Avant d'aborder les tâches proposées dans le contexte de la terminale S, il paraît incontournable d'analyser le programme correspondant. Bien que nous ayons déjà fait quelques commentaires sur les contenus mathématiques du programme dans les chapitres 1 et 2, nous n'avons pas encore étudié le programme en tant que tel. Dans cette section, il s'agit de le lire avec un regard fixé sur les liens perceptibles entre les probabilités à densité et le calcul intégral. Le programme, mis à part quelques fois où il explicite des types de tâches, n'a pas pour vocation de proposer des tâches à donner aux élèves. Cependant, il nous permet d'obtenir des éléments du plan épistémologique de l'ETM de référence, en précisant les définitions, propriétés et théorèmes à connaître (référentiel théorique), en donnant des représentations sémiotiques (qui se limitent souvent à du symbolique ou du langage naturel) et précisant les artefacts technologiques (TICE) qui peuvent être utilisés. Nous avons donc bien conscience que l'étude du programme ne nous permettra pas de lister des tâches mais de décrire partiellement le plan épistémologique de l'ETM qui sert de référence aux enseignants pour ensuite proposer des tâches.

Nous allons, tout d'abord, faire une présentation des deux parties qui nous intéressent dans le programme, à savoir celle concernant le calcul intégral et celle concernant les lois à densité. Puis, nous dégagerons quelques premiers éléments sur les liens présents entre les deux sous-domaines. Les citations seront toutes extraites du programme (MEN, 2011), nous ne le préciserons pas à chaque fois.

4.2.1 Présentation du programme

Le programme de terminale S est composé de trois grandes parties : l'analyse, la géométrie ainsi que les probabilités et statistique. Chaque partie est partagée en sous-parties qui sont détaillées sous forme d'un tableau à trois colonnes (Contenus/Capacités attendues/Commentaires). Les deux sous-domaines qui nous intéressent dans notre recherche se retrouvent dans les sous-parties nommées *Intégration* et *Notion de loi à densité à partir d'exemples* (les extraits correspondants sont dans l'Annexe B).

À titre indicatif, il est précisé qu'« *on pourrait consacrer la moitié du temps à l'analyse, l'autre moitié se répartissant équitablement entre géométrie et probabilités-*

statistique ». Le paragraphe du programme sur l'enseignement des lois à densité correspond à peu près (au niveau de la place textuelle y étant consacrée) à la moitié du programme sur l'enseignement des probabilités, on peut donc considérer que cet enseignement représente près d'un huitième du temps d'enseignement de la classe de terminale S. L'enseignement de l'intégration représente quant à lui à peu près un quart de l'enseignement de l'analyse. On peut donc constater qu'un temps à peu près équivalent est à consacrer aux deux thèmes, temps non négligeable (un huitième de l'année pour chacun d'eux).

Rien n'est précisé sur l'ordre à adopter entre les sous-parties, cependant la présence de signes $\int$ dans la sous-partie sur les lois à densité laisse penser que le calcul intégral doit être rencontré en amont.

a) Présentation générale de la sous-partie *Intégration*

La partie *Analyse* est principalement constituée de révisions et d'approfondissements, au niveau de l'étude de fonctions et de suites, et de l'introduction de nouvelles fonctions de référence, comme la fonction exponentielle. De plus, un nouveau concept, qualifié de « *fondamental* » en analyse, est étudié. Il est cependant « *modestement abordé et développé* ». Il s'agit du calcul intégral de fonctions continues de signe quelconque sur un intervalle borné (sans l'intégration par parties et les changements de variables). Les notions abordées dans cette sous-partie sont entièrement nouvelles pour les élèves.

Cette sous-partie n'est pas une nouveauté dans le programme, l'intégration est enseignée en terminale S depuis plusieurs réformes. Seuls l'intégration par parties a disparu par rapport à la réforme précédente (programme de 2001) et l'ordre d'apparition des notions a été modifié. Désormais, nous trouvons les notions dans cet ordre :

- L'intégrale d'une fonction continue positive sur un intervalle $[a; b]$ est définie comme aire sous la courbe.
- On introduit la fonction F, primitive de la fonction f dans le cas continu positif.
- La notion de primitive d'une fonction continue est étendue et amène à l'intégrale d'une fonction continue de signe quelconque, définie par la formule : $\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$.
- Les propriétés de linéarité et de positivité, la relation de Chasles, puis la définition de la valeur moyenne sont introduites.

Il est demandé de s'appuyer sur la notion intuitive d'aire que les élèves ont rencontrée au collège et sur ses propriétés d'additivité et d'invariance par translation et symétrie. Aucune connaissance plus théorique sur l'aire n'est exigée.

Dans les modalités de mise en œuvre, dans le programme de 2001, il est précisé : « *On indiquera que l'aire sous la courbe peut être approchée en l'encadrant par deux suites adjacentes construites en quadrillant le plan de plus en plus finement* ». La notion de suites adjacentes n'étant plus au programme, cette introduction ne peut plus être faite à l'identique, bien que l'on puisse toujours « *mener un calcul approché d'aire (parabole, hyperbole, etc.) pour illustrer cette définition* ». Le théorème « *Toute*

fonction continue sur un intervalle admet des primitives » est explicitement au programme, avec une démonstration exigible seulement dans le cas d'un intervalle fermé borné, en admettant que la fonction possède un minimum. Ceci permet de faire observer aux élèves que « *certaines fonctions comme* $x \mapsto \exp(-x^2)$ *n'ont pas de primitives « explicites »* ». Dans l'ancien programme, un paragraphe est dédié à la modélisation, dans lequel il est précisé que l'« *on illustrera l'intérêt de l'intégrale par diverses situations, entre autres :*

- *expression intégrale de la distance parcourue sur une droite par un point mobile dont on connait la vitesse instantanée* ;
- *expression intégrale du volume d'un solide dont on connait les aires des sections avec les plans d'équation* $z =$*constante* ;
- *calculs de probabilités d'intervalles pour des lois de probabilités à densité* ».

Dans le programme actuel, aucune référence n'est faite à la modélisation ou à des applications particulières du calcul intégral mis à part pour la valeur moyenne où il est proposé de l'illustrer par des exemples issus d'autres disciplines comme la science physique (mouvement uniformément accéléré) et les sciences de l'ingénieur (valeur moyenne et valeur efficace dans un transfert énergétique). Nous trouvons aussi la proposition d'étudier le calcul du volume d'un solide en AP (accompagnement personnalisé).

La seule référence aux TICE se trouve dans le cadre de l'algorithmique : « *Pour une fonction monotone positive, mettre en œuvre un algorithme pour déterminer un encadrement d'une intégrale* » à implémenter certainement sur calculatrice. Il n'est pas mentionné l'utilisation de logiciels ou de la calculatrice pour calculer des intégrales ou faire des encadrements. Seulement dans l'introduction générale de la partie *Analyse*, il est proposé d'utiliser des logiciels de calcul formel ou scientifique.

b) Présentation générale de la sous-partie *Notion de loi à densité à partir d'exemples*

Pour la partie *Probabilités et statistique*, il est mentionné qu'il s'agit d'un approfondissement « *du travail mené les années précédentes en probabilités et statistique* ». Il est précisément écrit que les lois de probabilités à densité sont introduites « *afin de traiter les champs de problèmes associés aux données continues* » (MEN, 2011, p. 12). Comme nous l'avons mentionné dans le chapitre 1, le programme de 2001 ne consacre que quatre lignes aux lois à densité et seules la loi uniforme sur $[0;1]$ et les lois exponentielles sont abordées. L'enseignement, dans ce nouveau programme, des lois normales, ainsi que l'approfondissement des autres lois (avec notamment l'étude de l'espérance) ont enrichi cette partie de manière conséquente.

Les trois types de lois étudiées sont :

- les lois uniformes sur un intervalle $[a;b]$,
- les lois exponentielles,
- les lois normales.

Dans le programme, il est précisé que pour définir la probabilité d'une variable aléatoire continue X,

> on admet que X [une variable aléatoire, fonction de Ω dans $\mathbb{R}$, qui associe à chaque issue un nombre réel d'un intervalle I de $\mathbb{R}$] satisfait aux conditions qui permettent de définir la probabilité de l'évènement $\{X \in J\}$ comme aire

> du domaine : $\{M(x;y); x \in J \text{ et } 0 \leqslant y \leqslant f(x)\}$ où f désigne la fonction de densité de la loi et J un intervalle inclus dans I. (MEN, 2011, p. 12)

Dans le cas de la loi uniforme et de la loi exponentielle, la fonction de densité ainsi que les calculs de probabilités et d'espérances associés sont à connaître. Il n'est pas mentionné la variance dans ces deux lois. La loi normale centrée réduite doit être introduite en s'appuyant « *sur l'observation des représentations graphiques de la loi de la variable aléatoire* $Z_n = \dfrac{X_n - np}{\sqrt{np(1-p)}}$ *où* X_n *suit la loi binomiale* $\mathcal{B}(n,p)$*, et cela pour de grandes valeurs de* n *et une valeur de* p *fixée entre* 0 *et* 1 ». Ensuite, le théorème de Moivre Laplace assure que « *pour tous réels* a *et* b*,* $P(Z_n \in [a;b])$ *tend vers* $\displaystyle\int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,\mathrm{d}x$ *lorsque* n *tend vers* $+\infty$ ». Ce théorème est admis. La visualisation de ce théorème qui permet de le « valider ».

Dans les capacités attendues, les élèves doivent connaître la fonction de densité de la loi normale $\mathcal{N}(0;1)$, sa représentation graphique et certaines valeurs approchées remarquables. L'espérance de la loi normale centrée réduite peut être établie, tandis que la variance est admise. Enfin, la loi normale centrée réduite est généralisée à la loi normale de paramètres μ et σ.

La loi normale et tout particulièrement le théorème de Moivre Laplace sont ensuite exploités dans la sous-partie du programme qui suit, *Intervalle de fluctuation.* Le théorème de Moivre Laplace permet en effet de déduire un nouvel intervalle de fluctuation dans le cas où la variable aléatoire considérée suit une loi binomiale qui vérifie les conditions $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$.

Enfin, dans la partie *Notions de loi à densité à partir d'exemples*, des interactions avec la science physique et les sciences de l'ingénieur sont possibles, notamment autour des mesures physiques sur un système réel en essai. Au niveau des intervalles « 1, 2 ou 3 sigmas », il est indiqué que ces nouvelles notions doivent être illustrées par des « *exemples issus des autres disciplines* ».

Au niveau des TICE, « *utiliser une calculatrice ou un tableur pour calculer une probabilité dans le cadre d'une loi normale* $\mathcal{N}(\mu;\sigma^2)$ » fait partie des capacités attendues.

c) Un programme avec des ambiguïtés

Contrairement à la sous-partie *Intégration* qui, mis à part quelques changements minimes entre chaque réforme, est assez stable, nous pouvons déceler des ambiguïtés dans la sous-partie sur les lois à densité qui, elle, est nouvelle (tout du moins dans ces proportions). Nous notons deux ambiguïtés dans ce texte institutionnel.

Tout d'abord, dans l'intitulé même de la sous-partie *Notion de loi à densité à partir d'exemples*, à quoi se rapporte le mot « *exemples* » ? Dans la colonne *Commentaires*, il est écrit que « *les exemples étudiés s'appuient sur une expérience aléatoire et un univers associé* Ω*, muni d'une probabilité* ». Cela laisse entendre que le mot « *exemple* » n'est pas associé à expérience aléatoire. S'agit-il alors d'exemples de lois à densité ? Et cela sous-entend-t-il qu'il faut traiter seulement les trois lois au programme, c'est-à-dire les lois uniformes, exponentielles et normales ? En d'autres termes, qu'il s'agit de trois études de cas ? Ou s'agit-il de rencontrer différentes lois à

densité (même en dehors des trois au programme) mais sans tendre vers une « généralisation » formalisée et seulement à travers des exemples d'expériences aléatoires ? Ou, finalement, les différentes lois à densité sont-elles un point de départ pour introduire la notion de lois à densité, de façon décontextualisée, ou encore des exemples d'applications des lois à densité ?

Dans le premier cas, on peut comprendre qu'il n'est recommandé de n'étudier que les trois lois au programme, sans même en rencontrer d'autres, ni essayer de tirer des généralités sur les lois à densité et en particulier sur les fonctions de densité par exemple. Dans le deuxième cas, on peut envisager cette « généralisation » mais seulement en contexte, avec peut-être une importance accordée à la modélisation de phénomènes aléatoires. Enfin, dans le troisième cas, une « généralisation » est clairement à étudier.

Malgré les *exemples* demandés, qui pourraient être interprétés comme des exemples de situations, nous pouvons remarquer que le mot « modélisation » n'apparaît à aucun moment dans la partie *Probabilités et statistique* du programme. Cependant, certaines phrases peuvent nous laisser penser à des problèmes de modélisation comme :

- « *traiter les champs de problèmes associés aux données continues* » ;
- « *cette partie du programme se prête particulièrement à l'étude de situations concrètes, par exemple sur la radioactivité ou la durée de fonctionnement d'un système non soumis à un phénomène d'usure* » ;
- « *on illustre ces nouvelles notions par des exemples issus des autres disciplines* », en particulier avec des propositions d'interdisciplinarité avec les sciences de l'ingénieur et les sciences physique-chimie sur des « *mesures physiques sur un système réel en essai* ».

On peut quand même remarquer que le mot « *exemples* » est lui aussi présent dans le titre du paragraphe concernant les probabilités à densité dans le précédent programme : « *Exemples de lois continues* ». À la différence qu'ici il est clairement associé aux lois, qui sont les deux lois à densité au programme. Cette première question est soulevée. A première vue, le questionnement que nous avons ici peut paraître éloigné de nos questions sur les liens entre probabilités et analyse, mais en fait il n'en est rien car l'approche de ces notions est différentes selon les cas.

Ensuite, nous pouvons repérer une seconde ambiguïté, pouvant renforcer la première, dans la phrase : « *toute théorie générale des lois à densité* [...] *est exclue* ».

S'agit-il de ne pas « généraliser » les idées sur les lois à densité ? Nous prenons le soin de mettre le mot « généraliser » entre guillemets car nous avons bien conscience du fait que cette généralisation peut atteindre différents niveaux et nous voulons ici parler seulement d'une généralisation partielle, par exemple sur les fonctions de densité dans le contexte de la terminale S. Cela signifie donc de ne traiter que des trois exemples, de façon « cloisonnée » ? Cela reviendrait à trois études de cas distincts, sans essais de les unifier. Ou doit-on prendre « *théorie générale* » dans un sens de rigueur et de généralisation très forte, ce qui signifie qu'il n'est pas demandé de parler de théorie de la mesure, de tribus... ? Ce qui bien entendu n'a pas les mêmes conséquences sur l'enseignement. Dans le deuxième cas, cela paraît bien entendu légitime de ne pas traiter de notions qui ne sont pas à la portée des élèves. Dans le premier cas, il y a matière à discussion. Ce qui est sûr, c'est que dans l'extrait

du programme considéré, il est mis en avant le fait que certains éléments semblent ne pas devoir être explicités, mais pour autant cela n'est pas écrit. Nous pouvons remarquer que, dans le programme, aucune propriété sur la fonction de densité et sur les probabilités associées n'est mentionnée, mis à part dans la définition de la probabilité associée à une loi à densité. Ce dernier point peut quand même laisser penser qu'une place est possible pour les généralités.

4.2.2 Des articulations apparentes ou possibles

Dans l'ancien programme apparaît le commentaire « *ce paragraphe est une application de ce qui aura été fait en début d'année sur l'exponentielle et le calcul intégral* » (MEN, 2001, p. 70), qui mentionne de façon claire l'objectif de l'étude des lois à densité et de son lien avec l'analyse. Dans le programme actuel, aucun discours précisant ce lien n'est présent, ni dans la sous-partie *Intégration*, ni dans la sous-partie *Fonction exponentielle*, ni même dans celle relative aux lois à densité. Pour la fonction exponentielle, il est seulement écrit, dans l'introduction de la partie *Analyse*, qu'elle intervient dans différents champs du programme et donc qu'« *il est souhaitable de l'introduire assez tôt dans l'année* ». Mais aucune mention plus explicite quant à son intervention en probabilités n'est citée ici.

Malgré cette absence de discours explicite, le lien entre probabilités à densité et calcul intégral apparaît dès le départ dans le programme, au niveau des définitions. Comme nous l'avons déjà mentionné dans le chapitre 2, dans un cas, on définit « *l'intégrale d'une fonction continue et positive sur* $[a;b]$ *comme aire sous la courbe* », et dans l'autre, on considère une variable aléatoire X, fonction de Ω dans $\mathbb{R}$, qui associe à chaque issue un nombre réel d'un intervalle I de $\mathbb{R}$ et :

> on admet que X satisfait aux conditions qui permettent de définir la probabilité de l'évènement $\{X \in J\}$ comme aire du domaine :
>
> $$\{M(x;y);\ x \in J \text{ et } 0 \leqslant y \leqslant f(x)\}$$
>
> où f désigne la fonction de densité de la loi et J un intervalle inclus dans I.

Nous avons donc déjà identifié la fonction de densité comme l'objet mathématique crucial pour permettre de faire le lien entre les deux sous-domaines.

Finalement, mis à part la phrase citée ci-dessus, aucune indication n'est donnée sur la notion de fonction de densité et sur son introduction d'un point de vue général. Dans le cas précis de la loi normale centrée réduite en revanche, il est précisé que l'on doit s'appuyer sur l'observation du théorème de Moivre Laplace, qui « *assure que pour tous réels a et b,* $P(Z_n \in [a;b])$ *tend vers* $\displaystyle\int_a^b \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}\,\mathrm{d}x$ *lorsque n tend vers* $+\infty$ ». Cette approche est donc censée justifier un lien entre loi à densité (de la loi normale centrée réduite) et calcul intégral. Ce lien se fait en passant d'une loi discrète (la loi binomiale) à une loi continue, la loi normale centrée réduite. Mais cette introduction est exclusivement réservée à la loi normale.

Les trois familles de fonctions de densité au programme sont les fonctions du type $x \mapsto \dfrac{1}{b-a}$, $x \mapsto \lambda e^{-\lambda x}$ et $\dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Nous appellerons ces fonctions, à constante multiplicative près, les fonctions de référence probabilistes. Ces fonctions font à la fois partie du domaine de l'analyse et de celui des probabilités et sont justement des

intermédiaires entre les deux. On peut remarquer que l'exponentielle est une fonction de référence introduite en terminale S, donc cette fonction est beaucoup étudiée à ce niveau. Dans la sous-partie dédiée au calcul intégral, on peut aussi repérer la présence de la fonction $x \longmapsto exp(-x^2)$. Elle est introduite en commentaire comme exemple de fonction qui n'admet pas de primitive « *explicite* » (dans le sens, exprimable avec des fonctions de référence à ce niveau). On peut voir ici, bien qu'implicite, une référence à la fonction de densité de la loi normale. Ceci justifie (implicitement toujours) que l'on ne peut pas déterminer par le calcul (par « primitivation ») les valeurs des probabilités des variables aléatoires suivant une loi normale.

Comme autre lien apparent dans le programme, nous pouvons constater la présence du signe $\int$ dans la définition de l'espérance mathématique des trois lois à densité au programme. La mise en relation entre probabilités et calcul intégral se fait ici par le signe. Cette présence est justifiée dans le cas de la loi uniforme par le commentaire suivant : « *On note que cette définition constitue un prolongement dans le cadre continu de l'espérance d'une variable aléatoire discrète* ». La justification ici de la présence d'une intégrale est une justification du type passage du discret au continu, avec le passage de la somme finie $\sum$ à la somme intégrale $\int$.

Enfin, nous pouvons aussi voir la proposition d'étudier la méthode de Monte-Carlo en AP (accompagnement personnalisé) comme un autre moyen de connecter lois à densité et calcul intégral. Dans ce cas, il s'agit de déterminer la valeur approchée d'une intégrale à l'aide de lois uniformes, donc d'utiliser ici les probabilités comme un outil pour déterminer une intégrale. Cette méthode étant proposée pour des séances d'AP, nous pouvons penser qu'elle n'est pas abordée par l'ensemble des classes de terminale S.

4.2.3 Un programme avec des « incohérences » entre probabilités et analyse

Nous allons maintenant mettre en évidence plusieurs « incohérences » repérables entre les deux sous-parties qui font l'objet de notre étude. Nous mettons des guillemets au terme « incohérence » car il ne s'agit pas de faire de jugement de valeur (notamment en discréditant les choix faits par les concepteurs du programme) mais simplement de soulever des éléments qui, tout du moins à première vue, semblent pouvoir poser problème.

La première « incohérence » est le fait que seules les intégrales de fonctions continues sur un intervalle borné sont traitées en calcul intégral, alors que des intégrales impropres apparaissent en probabilités à densité, notamment dans la définition de l'espérance. Ces intégrales ne sont certes jamais traitées comme telles, on parle seulement de limite d'intégrales sans se poser plus de question. Une contradiction apparaît car, pour autant, il est bien précisé dans cette sous-partie que « *toute théorie générale* [...] *des intégrales sur un intervalle non borné est exclue* ». Cela laisse entendre qu'il y a effectivement un problème (nous allons calculer des intégrales impropres) mais que nous n'allons pas nous arrêter dessus. Aucune solution n'est apportée pour la gestion de ce conflit. Peut-être cela doit-il se gérer seulement dans le cas particulier des intégrales rencontrées ?

L'espérance, par exemple dans le cas de la loi exponentielle, est définie « *comme la limite quand x tend vers $+\infty$ de $\int_0^x tf(t)\,\mathrm{d}t$, où f est la fonction de densité de la loi exponentielle considérée* ». Aucun commentaire n'est fait sur cette limite : Existe-t-elle ? Est-elle finie ? Pourquoi ? Des discussions pourraient être engagées sur ce thème. Pour la loi exponentielle, comme il s'agit d'une fonction positive, la limite peut être soit finie, soit égale à $+\infty$. Mais cela n'est pas une proposition du programme.

Une seconde « incohérence », bien que non explicite dans le programme, existe. Il est précisé dans la sous-partie *Intégration* que « *l'intégration par parties n'est pas un attendu du programme* », et pour autant il est demandé ensuite dans la sous-partie *Notion de loi à densité à partir d'exemples* de « *démontrer que l'espérance d'une variable aléatoire suivant une loi exponentielle de paramètre λ est $\frac{1}{\lambda}$* ». Cette démonstration nécessite en effet une intégration par parties. Bien entendu, celle-ci peut-être contournée en introduisant directement la fonction intermédiaire qui permet le calcul.

4.2.4 Conclusion sur le programme

Les lois à densité tout comme le calcul intégral (et la fonction exponentielle) ont la particularité d'être des parties où les notions sont totalement nouvelles pour des élèves de terminale S. Il s'agit de notions rencontrées pour la première fois.

Le programme semble imposer un ordre. L'exponentielle et le calcul intégral sont à étudier en amont des probabilités à densité. De même au sein des probabilités, un ordre semble s'imposer : la loi uniforme, suivie de la loi exponentielle et enfin, la loi normale.

Ce programme laisse apparaître quelques articulations entre les probabilités à densité et le calcul intégral. Cependant la compréhension ou même la justification de ces liens ne sont pas abordées. Pour l'expliquer, nous pouvons dire que les mises en relation entre les deux sous-domaines se font par les signes dans les définitions (nous pensons, par exemple, à la définition de la probabilité comme aire sous la courbe de densité, ou encore aux différentes définitions de l'espérance) et dans le théorème de Moivre Laplace (avec une relation entre une probabilité et une intégrale). La relation sémiotique $(*)$ est donc contenue dans le référentiel théorique, cependant aucun commentaire n'est fait dans le programme pour justifier cet élément du référentiel théorique. Notamment, rien n'est proposé pour l'introduction de la notion de fonction de densité. L'intégrale semble avoir un statut outil pour le calcul de probabilités, mais sans que cela ne soit jamais justifié.

Pour le cas de la loi normale, nous avons vu que le théorème de Moivre Laplace, faute de pouvoir être démontré doit être visualisé. La visualisation (graphique *a priori*) semble donc avoir sa place dans la construction des articulations entre les deux sous-domaines. Là aussi, rien n'est précisé sur cette visualisation.

Au niveau des artefacts, on peut repérer que le programme invite à l'utilisation des logiciels dans les classes. En calcul intégral, les TICE sont absents (en considérant la partie *Analyse*, on en déduit que les logiciels de calcul formel ou scientifique

peuvent être utilisés). Pour les probabilités à densité, l'utilisation de la calculatrice et du tableur est fortement souhaitée, notamment dans le cas du calcul de probabilités dans le cadre d'une loi normale. Il ne semble pas question ici de faire du calcul approché d'intégrale, qui aurait pourtant tout à fait lieu d'être. Nous pouvons nous demander si ces artefacts différents ne peuvent pas justement être source de rupture entre les deux domaines.

Cette analyse nous permet de dire que des relations sémiotiques sont présentes, mais le programme ne nous permet pas de dire comment ces relations sont justifiées et la question de la compréhension de ces liens n'est pas visible. L'articulation entre les sous-domaines est gérée par le biais de la dimension sémiotique, qui semble imposer ce lien.

Nous avons bien conscience que le programme se limite à du contenu disciplinaire et quelques commentaires. Pour cette raison, l'analyse du document *Ressources* qui accompagne le programme semble importante pour peut-être nous éclairer davantage.

Nous avons de plus soulevé des ambiguïtés et « incohérences » dans ce programme, qui laisse entretenir un certain flou, que nous espérons lever avec l'analyse des documents officiels qui complètent le programme.

4.3 Analyse du document *Ressources*

Après l'analyse du programme, nous allons poursuivre notre description et analyse des ETM de référence à l'aide du document *Ressources* dédié aux probabilités et à la statistique en terminale (MENJVA & DGESCO, 2012) mis à disposition par l'Éducation nationale pour accompagner le programme. Nous compléterons nos analyses par endroit avec le document d'accompagnement du programme antérieur (MJENR, 2002). Cette analyse va nous permettre d'avoir des précisions par rapport au programme et de découvrir des premières tâches qui peuvent être proposées aux élèves.

Nous allons donc aborder l'analyse de ce document en nous posant plusieurs questions :

- Est-ce que les ambiguïtés du programme sont levées ? Entre autres, est-ce que les trois lois au programme constituent trois études de cas ou est-ce qu'elles doivent être abordées comme des exemples de lois à densité qui font suite à une présentation plus « générale » ? Quel degré de généralité est attendu ?
- Comment sont prises en charge les « incohérences » repérées sur la gestion des intégrales impropres ou encore le calcul de l'espérance de la loi exponentielle ?
- Quelles approches sont proposées pour justifier les relations sémiotiques mises en avant dans le programme ? Quelle introduction de la notion de densité est attendue ?
- Existe-t-il d'autres liens entre probabilités à densité et calcul intégral mis en évidence dans le document *Ressources* ?
- Existe-t-il des tâches exploitant les articulations entre les deux sous-domaines ?

Nous allons dans un premier temps décrire le document *Ressources* et les parties qui nous intéressent en essayant ensuite d'en dégager des compléments sur le plan épistémologique de l'ETM de référence et sur les articulations entre probabilités à densité et calcul intégral. Enfin, nous nous intéresserons particulièrement aux tâches proposées.

Les citations de cette section (sauf mention contraire) seront extraites du document Ressources (MENJVA & DGESCO, 2012).

4.3.1 Présentation du document

Ce document *Ressources* est un support pour les enseignants, assez conséquent (environ 70 pages). Il est écrit dans l'introduction qu'il a pour objectif de donner « *des éléments détaillés permettant aux professeurs de construire leur propre cours* ». Cependant, il est précisé qu'il « *ne s'agit pas d'un modèle reproductible tel quel mais d'un support théorique sur les notions introduites pour la première fois dans les programmes du secondaire* ». On peut peut-être voir ce document comme une ressource proposant des éléments pour enrichir l'ETM personnel des enseignants et non seulement pour compléter l'ETM de référence. Cette ressource va permettre à l'enseignant d'élaborer un ETM idoine à l'aide des éléments proposés. Cependant, certains éléments vont tout de même permettre d'éclairer l'ETM de référence attendu par l'institution. C'est ce qui va nous intéresser ici.

Ce document est majoritairement centré sur la loi normale et les notions de statistique inférentielle au programme (intervalle de fluctuation asymptotique et intervalle de confiance). Il est composé de six parties. Les trois premières portent sur les lois normales, les deux suivantes sur des notions de statistique et enfin, la dernière sur un complément sur les lois uniformes et exponentielles. Les six parties sont complétées par six annexes, dont trois sont susceptibles de nous intéresser dans cette analyse : l'annexe 2 apporte des compléments sur les lois normales, l'annexe 4 est dédiée à l'utilisation des TICE et enfin, l'annexe 5 porte sur la méthode de Monte-Carlo. Les annexes 3 et 4 traitent de statistique et l'annexe 1 donne des éléments historiques sur le théorème de Moivre Laplace. En nombre de pages, hormis les annexes, 14 pages sont consacrées à la loi normale, 22 à la statistique inférentielle et seulement 3 aux lois uniformes et exponentielles. Nous allons bien entendu nous centrer seulement sur les paragraphes portant sur les lois à densité (et non sur la statistique).

La majeure partie du document consacrée aux lois à densité traite de la loi normale. Nous pouvons penser que ce choix est justifié par le fait que cette loi est nouvelle dans le programme (les lois uniformes et exponentielles étaient déjà présentes dans le programme de 2001) mais nous pouvons aussi nous demander si cette loi ne doit pas être favorisée dans son étude par rapport aux autres. Une remarque (p. 10) laisse entendre que les lois uniformes et exponentielles doivent être introduites avant les lois normales, ce qui renforce l'ordre donné dans le programme.

4.3.2 Premiers éclaircissements et apparition de nouvelles ambiguïtés

a) Premières réponses

Nous pouvons repérer que seules les trois lois au programme apparaissent dans ce document. Cela laisse donc penser que le terme « *exemples* » se rapporte exclusivement aux trois exemples de lois précisés dans le programme, ce qui répond à une des questions que nous nous posions à la lecture de ce dernier. L'étude des lois à densité semble se limiter à ces trois lois. La seule « généralité » présente dans le document se trouve en remarque (p. 11) dans ce que l'on pourrait qualifier de définition d'une loi à densité : « *il s'agit d'une loi à densité c'est-à-dire qu'il existe une fonction g définie sur $\mathbb{R}$ telle que, pour tous réels a et b vérifiant $a \leqslant b$, on a* $P(a \leqslant X \leqslant b) = \int_a^b g(t)\,\mathrm{d}t$ ». Nous pouvons voir ici un pas vers la généralisation de ce qu'est une loi à densité et le lien avec la fonction de densité, mais c'est l'unique endroit du document où apparaît ce type de « généralités ».

Nous pouvons cependant faire remarquer que le terme « *exemples* » est aussi présent dans l'ancien programme et que pour autant une partie de généralités est présente dans le document d'accompagnement de 2002 (MJENR, 2002). On y trouve notamment un tableau récapitulatif sur les propriétés liées au calcul de probabilités dans le cas de « *lois de probabilités définies sur un intervalle I borné ou borné à gauche (i.e. $I = [a; b]$, ou $I = [a; +\infty[$ et dites à densité continue* » (il s'agit donc des cas où la fonction de densité est continue sur l'intervalle I). Toutes les propriétés sont données sous forme générale et non uniquement pour la loi uniforme sur $[0; 1]$ et la loi exponentielle (bien que ce soient les deux seules lois explicitement au programme). On voit aussi apparaître dans ce document d'accompagnement des exercices qui proposent de déterminer des paramètres pour que des fonctions (autres que la densité de la loi exponentielle) soient des fonctions de densité.

b) Apparition de nouvelles ambiguïtés dans l'introduction du document

Selon l'introduction du document *Ressources*, l'approche des probabilités et statistique du programme de terminale S se veut différent de celle de l'enseignement supérieur : ces notions nouvelles « *sont enseignées dans différents cursus de l'enseignement supérieur mais le point de vue adopté dans le programme de la classe terminale est assez différent* » (p. 1). Il n'est pas précisé en quoi il est différent. Ceci peut être dû au fait que les probabilités dans le secondaire ne sont pas abordées selon l'axiomatique de Kolmogorov (pour autant cela est aussi le cas dans beaucoup de filières du supérieur, comme en BTS ou en classe préparatoire économique par exemple) ou encore au fait que les intégrales impropres n'étant pas au programme, cela ne permet pas de présenter les probabilités avec la fonction de répartition ou encore pour une autre raison, d'où de nouvelles ambiguïtés.

La deuxième ambiguïté de cette introduction concerne les lois uniformes et exponentielles pour lesquelles il est mentionné qu'« *un complément [...] est proposé, leur approche ayant été modifiée* » (p. 1). Là encore, la modification apportée n'est pas explicitée ni dans l'introduction, ni dans le complément. Cela sous-entend-t-il seulement que la loi uniforme est définie non plus seulement sur $[0; 1]$ mais sur $[a; b]$

quelconque ? Cela signifie-t-il que les propositions faites dans le document d'accompagnement du programme antérieur ne sont plus conformes ?

Dans le document d'accompagnement des programmes de 2002 (MJENR, 2002), comme dans le programme associé, il est explicité le lien direct avec le calcul intégral et plus généralement l'analyse. Il est écrit que l'introduction des lois de probabilité à densité « *fait naturellement suite au cours sur l'intégration et l'enrichit d'applications importantes, telles la modélisation de la durée de vie d'un noyau d'une substance radioactive* » (*ibid.*, p. 142). Voilà une première hypothèse qui semble incontestable quant à la différence d'approche. Dans le programme de 2001, il est clairement attendu, en modalité de mise en œuvre, d'introduire la loi exponentielle comme application à la désintégration radioactive. Les équations différentielles n'étant plus au programme de 2011, cette approche détaillée dans le document d'accompagnement des programmes de 2002 (MJENR, 2002, pp. 77-79) n'a plus lieu d'être aujourd'hui.

4.3.3 L'introduction des différentes lois

Aucune proposition n'est faite dans le document pour introduire la notion de fonction de densité (dans sa généralité). Nous pouvons préciser ce qui est présenté pour chaque loi.

a) L'introduction de la loi uniforme

Pour introduire la loi uniforme sur $[a; b]$, il est tout d'abord proposé de l'introduire sur $[0; 1]$, par exemple en étudiant l'expérience aléatoire du choix au hasard d'un nombre réel entre 0 et 1, puis ensuite de définir la loi uniforme sur $[a; b]$ en « *remarquant que pour que l'aire sous la courbe soit égale à 1, il faut et il suffit que la valeur de la constante soit* $\frac{1}{b-a}$ ». Aucun détail n'est fourni. Il n'est même pas précisé comment amener dans un premier temps la fonction de densité de la loi uniforme sur $[0; 1]$, pour ensuite la définir sur $[a; b]$. Plus généralement, aucun élément n'est donné sur les raisons qui font que l'on s'intéresse à une aire sous une telle fonction, ni d'où vient cette fonction... Il semble donc que cette fonction et ses propriétés doivent être imposées sans réflexion particulière sur l'objet « fonction de densité ».

b) L'introduction de la loi exponentielle

Seulement une demi-page est consacrée aux lois exponentielles. Il semble s'agir de points à écrire ou à faire dans le cours. La loi exponentielle de paramètre λ est définie par sa fonction de densité, qui est imposée. Aucun travail n'est proposé pour l'introduire.

c) L'introduction de la loi normale centrée réduite

Contrairement aux deux lois précédentes, un travail conséquent est proposé pour l'introduction de la loi normale centrée réduite, qui se fait via le théorème de Moivre Laplace. Le document commence par expliquer comment centrer et réduire une

variable aléatoire, en utilisant les propriétés sur l'espérance et la variance vues en classe de première, et justifier l'intérêt d'une telle transformation, en particulier dans le cas d'une loi binomiale. Le passage de la variable aléatoire X_n (suivant une loi binomiale de paramètres n et p) à la variable aléatoire $Z_n = \dfrac{X_n - np}{\sqrt{np(1-p)}}$ permet de travailler avec une variable aléatoire dont les paramètres ne dépendent plus de ceux de X_n et donc plus de n et p. La variable aléatoire Z_n a pour espérance 0 et pour variance 1. De plus, les informations sur Z_n permettent de préciser les fluctuations de $\dfrac{X_n}{n}$ autour de son espérance p. Les explications sont accompagnées par des graphiques de diagrammes en bâtons permettant d'illustrer le propos. Ces préliminaires permettent d'amorcer l'activité d'introduction au théorème de Moivre Laplace qui est détaillée et doit pouvoir être proposée en classe (vu que tous les fichiers sont téléchargeables sur le site Eduscol). Il est proposé de s'intéresser à la variable aléatoire centrée réduite Z_n associée à une variable aléatoire X_n suivant une loi $\mathcal{B}(n,p)$ et d'« *associer la loi (discrète) de Z_{100} à des aires de rectangles, comme on le fait pour l'histogramme d'une variable continue* ». La construction des rectangles est détaillée pour permettre une compréhension de ce graphique, qui ne représente pas un histogramme mais bien des rectangles contigus. La visualisation de ce graphique notamment sur le fichier *GeoGebra* présenté (voir la figure 4.1) mais également disponible sur le site Eduscol permet de visualiser les rectangles pour un n aussi grand que voulu et de remarquer la proximité entre l'aire des rectangles et l'aire sous la courbe dessinée. On peut remarquer que cette courbe est donnée et non questionnée (est-ce la seule qui convient ?). L'expression de la fonction en question $x \mapsto \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ (qui est la fonction de densité de la loi normale centrée réduite) étant donnée (imposée), on peut ensuite assimiler l'aire sous la courbe à l'intégrale associée (à l'aide du cours de terminale sur l'intégration). Ensuite un travail de comparaison entre les valeurs de l'aire des rectangles qui correspondent à la probabilité $P(X = k)$ (en considérant bien le changement de paramètres), et la valeur de l'intégrale, en utilisant la calculatrice, est proposé pour vérifier la proximité des valeurs. Pour l'intégrale, il est proposé ici de déterminer la valeur approchée à l'aide directement de la touche dédiée au calcul d'intégrales. Comme précisé dans le document, l'animation proposée permet donc de constater – et non de démontrer – certaines propriétés, notamment le théorème de Moivre Laplace. Cette introduction par le théorème de Moivre Laplace présente de nombreux difficultés, notamment dans la compréhension des différentes étapes pour arriver à la visualisation du théorème.

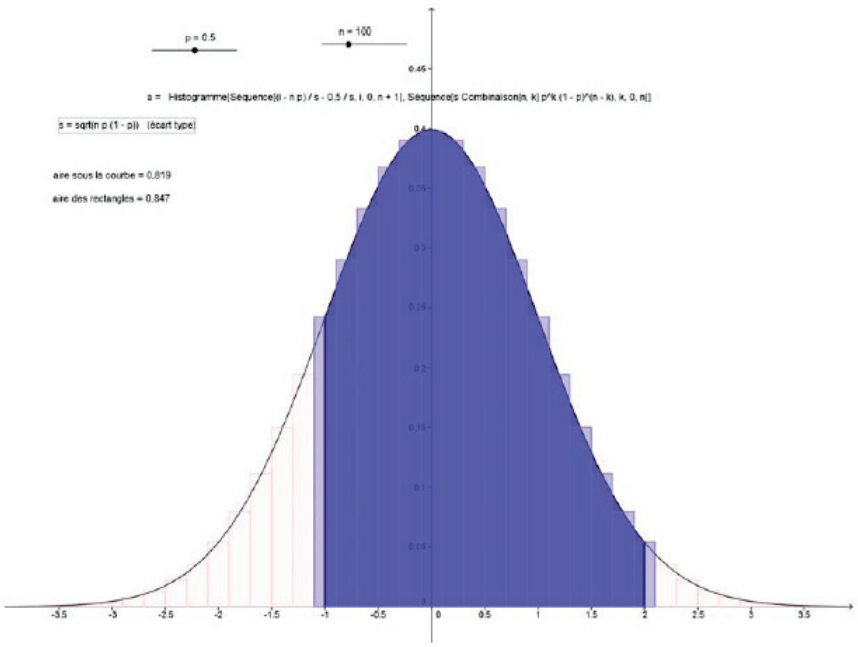


FIGURE 4.1 – Visualisation du théorème de Moivre Laplace, proposée dans le document *Ressources* (MENJVA & DGESCO, 2012, p. 6) pour introduire la loi normale centrée réduite

Juste après dans le document, nous pouvons repérer des propriétés relatives à la fonction de densité f de la loi normale centrée réduite mais aussi des probabilités associées (p. 8). Cependant, elles ne sont pas démontrées et rien ne met en évidence ce qui est particulier à cette fonction de densité f et ce qui général à toutes les fonctions de densité. Aucune remarque ou justification n'est donnée non plus sur le fait que l'aire sous la courbe est égale à 1, mis à part qu'elle représente la probabilité $P(X \in]-\infty;+\infty[)$, point déjà important. Cependant, est-ce que l'aire sous une courbe d'une fonction strictement positive sur $\mathbb{R}$ peut effectivement être finie ? Cela n'est pas questionné.

d) L'introduction de la loi normale généralisée

Pour la généralisation de la loi normale, il est proposé un échantillon de 50000 tailles d'hommes adultes, dont est donné un résumé statistique et l'histogramme disponible en figure 4.2. Une courbe est déjà tracée sur l'histogramme. Il est ensuite dit que si l'on centrait et réduisait cette variable « taille » (il n'est pas précisé comment et aucun graphique n'est donné, cela reste seulement hypothétique), on observerait que « *l'histogramme obtenu présente une analogie évidente avec la figure 3* [qui correspond à la figure 4.1] ». Cette remarque permet l'introduction de la loi normale $\mathcal{N}(\mu;\sigma^2)$.

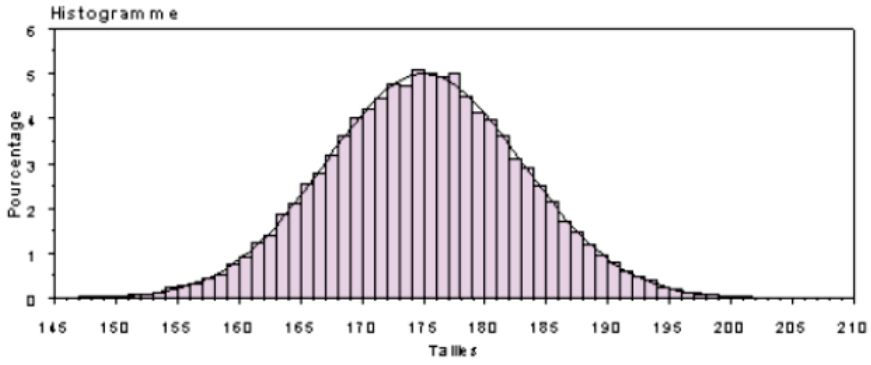


Figure 4.2 – Histogramme de 50000 valeurs de taille d'hommes adultes, proposé dans le document *Ressources* (MENJVA & DGESCO, 2012, p. 10) pour introduire la loi normale généralisée

Il est précisé en note qu'ici il s'agit d'un histogramme, ce qui n'est pas le cas pour la figure 4.1 où il s'agit d'un diagramme en bâtons contigus. Cette subtilité a son importance. Autant, pour un histogramme, c'est l'aire des rectangles qui représente la fréquence des classes ; autant, pour un diagramme en bâtons, c'est la hauteur du bâton qui représente la fréquence de la valeur isolée (l'épaisseur du bâton n'ayant aucune influence). Il n'est cependant rien précisé de plus que cette note dans le document. Cet histogramme pourrait pourtant permettre le basculement vers l'intégration. On peut faire le rapprochement ici avec l'approche que nous avons vu dans le chapitre 3 dans le cours de Faye, mais dans cette activité, le passage par le domaine des statistiques n'est pas vraiment exploité, ce qui était différent dans le document d'accompagnement de 2002.

e) Une différence d'approche avec le document d'accompagnement de 2002

Cet exemple des tailles est emprunté au document d'accompagnement publié en 2002 (MJENR, 2002), alors que la loi normale n'était pas encore au programme. Nous proposons donc de regarder comment il était traité à ce moment là car nous allons pouvoir observer une approche véritablement différente. La sous-partie du programme dédiée aux lois à densité est extrêmement réduite en 2001 et pour autant le document d'accompagnement de 2001 propose un véritable travail pour l'introduction de ce concept. L'idée première est d'étudier ce que signifie choisir au hasard un nombre dans $(0;1)$ [17]. On retrouve cette même idée pour l'introduction de la loi uniforme dans le document *Ressources* actuel, mais le document d'accompagnement donne, lui, beaucoup plus de détails. Le choix fait ici est de s'intéresser dans un premier temps à la probabilité attribuée à chacun des nombres quand on choisit au hasard un nombre dans l'ensemble E_2, ensemble des nombres de $[0;1[$ comportant au plus 2 décimales (qui vaut 10^{-2}) ; puis ensuite dans l'ensemble E_k (qui vaut 10^{-k}), pour en déduire que « *pour définir le choix au hasard d'un nombre réel dans* $[0;1[$*, on ne peut plus passer par la probabilité p de chaque élément, puisqu'on devrait alors*

17. Cette notation $(0;1)$ désigne tous types d'intervalles d'extrémités 0 et 1

avoir $p = 0$ ». Il est précisé que cette difficulté a été historiquement une difficulté épistémologique à dépasser.

Un autre argument est ensuite avancé pour justifier le fait que la probabilité d'un point est nulle. Le paragraphe suivant met en avant que justement contrairement au cas fini (discret) pour lequel on peut connaître avec exactitude le résultat d'une expérience aléatoire d'un choix au hasard, dans le cas de $[0; 1[$, c'est différent. Le résultat d'une mesure ne peut être donné que sous forme d'un intervalle,conséquence du manque de précision de l'instrument utilisé ou encore de l'arrondi choisi (le nombre de décimales étant fini, cela revient à donner le résultat à l'aide d'un intervalle).

Ces deux paragraphes du document d'accompagnement ont pour but de faire émerger l'idée que l'on ne peut plus utiliser concrètement la probabilité d'un élément pour donner la loi d'une variable continue, mais celle d'un intervalle. Nous pouvons remarquer qu'aucune proposition n'est donnée, dans le document *Ressources* actuel, pour aborder le problème de la probabilité nulle des valeurs isolées.

Ensuite, le document d'accompagnement (2001) pose la question de comment définir les probabilités d'intervalles. La proposition faite est de passer par les histogrammes via une situation « fabriquée ». On dispose d'un résumé numérique sous forme de tableau et d'un histogramme de l'échantillon des 50 000 tailles d'hommes adultes (cf. figure 4.2). Certains détails sont donnés sur le lien entre histogramme et fréquence, mais avec quelques confusions. Nous reviendrons sur l'analyse de ce type de confusion dans le chapitre 5. Ensuite, il est écrit :

> L'idée est ici de trouver une fonction f dont la courbe représentative *épouse* l'histogramme, l'aire sous cette courbe devant être égale à 1 : la probabilité d'un intervalle (a, b) sera alors l'aire sous la courbe délimitée par les droites d'équations $x = a$ et $x = b$, c'est-à-dire le nombre $\int_a^b f(x)\,\mathrm{d}x$. [...] Déterminer une telle fonction est un problème délicat, mais pour de nombreuses situations, dont celle qui est traitée ici, on cherche la fonction f parmi une famille paramétrée de fonctions : il suffit alors d'ajuster les paramètres. (MJENR, 2002, p. 143)

Une telle fonction a été tracée directement sur l'histogramme (cf. figure 4.2). La loi en jeu ici est une loi normale, bien qu'elle ne soit pas au programme de 2001. Il s'agit ici de comprendre l'intérêt d'une telle fonction et non pas d'étudier une des deux lois alors au programme. Il est ajouté :

> De même que dans un modèle défini par une loi P sur un ensemble fini E, les fréquences fluctuent autour de la loi P, de même ici, l'invariant est la fonction f et pour des grandes séries de données, les histogrammes *fluctuent autour* du graphe de f. (MJENR, 2002, p.143)

Cette remarque est illustrée par un second histogramme représentant une autre série de données.

Dans ce document d'accompagnement, la question de l'introduction de la fonction de densité est posée. On retrouve une approche similaire à celle du cours de Faye (cf. chapitre 3). Cependant, ici, la courbe est donnée, il ne s'agit pas réellement d'un ajustement. Notons qu'il serait difficile pour les élèves de la trouver seuls. Le passage par la statistique descriptive avec les histogrammes est le moyen choisi ici pour arriver à considérer une fonction particulière qui « *épouse* » l'histogramme, avec pour aire sous cette courbe 1, ce qui semble être donné comme postulat sans

être justifié. Ce travail ne se retrouve pas dans le document *Ressources* de 2012. Nous ne trouvons même que peu de passages sur la fonction de densité et ses propriétés.

Pour conclure, ce type d'approche n'est pas visible dans le document *Ressources* actuel, cependant cela signifie-t-il qu'il n'aurait pas lieu d'être en classe ? Nous gardons bien à l'esprit que le document *Ressources* n'est pas un document exhaustif pour un cours (il n'a même pas vocation à servir de cours à proprement dit). Donc, mis à part pour la loi normale centrée réduite, où la visualisation à l'aide du logiciel *GeoGebra* semble attendue, aucune tâche particulière n'est proposée pour l'introduction de la notion de densité. Les fonctions de densité des lois au programme semblent être imposées par leurs expressions algébriques. Aucun élément n'est donné pour favoriser la compréhension de cette notion et de son utilité.

4.3.4 Autres articulations

Nous allons maintenant présenter d'autres articulations entre probabilités et analyse qui peuvent être repérées dans le document *Ressources*.

a) L'espérance

Un paragraphe est dédié à l'espérance d'une variable aléatoire de loi uniforme sur $[a; b]$. Il est proposé d'observer que l'espérance par la formule $E(X) = \int_a^b xf(x)\,\mathrm{d}x$ prolonge bien celle de l'espérance d'une variable aléatoire discrète. Il semblerait qu'il y ait un besoin de justifier cette définition. Cependant, le document *Ressources* étant à destination des enseignants, nous pouvons nous demander si cette justification est à présenter aux élèves ou si elle est là seulement pour convaincre les enseignants. La justification de cette définition se fait par l'intermédiaire d'un fichier *GeoGebra* (disponible en ligne), qui permet de visualiser le passage de l'aire de rectangles avec une base de plus en plus petite à l'aire sous la courbe d'équation $y = xf(x)$ (cf. figure 4.3). Les rectangles ont chacun pour largeur (axe des abscisses) $f(x)dx$ (qui vaut $\frac{1}{n}$, dans le cas de la loi uniforme) et pour longueur (axe des ordonnées) x. Il reste tout de même le fait que l'on admet que l'aire $f(x)dx$ correspond bien à la probabilité. Ceci est à prendre comme tel, par rapport à ce qui est imposé en amont. L'ensemble de l'extrait en question se trouve en Annexe C.1.

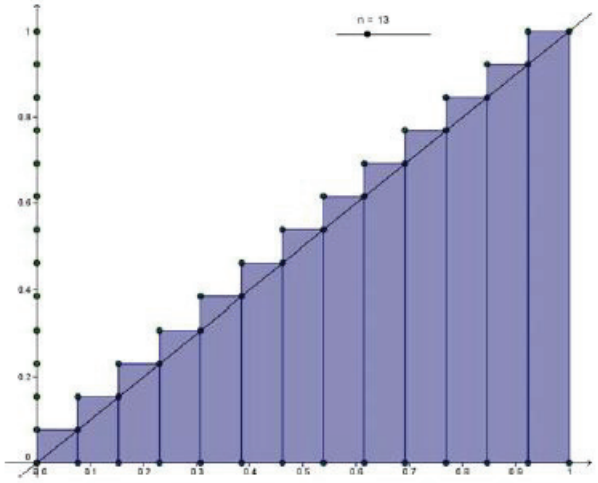


FIGURE 4.3 – Justification « visuelle » de la définition de l'espérance d'une variable aléatoire de loi uniforme dans le document *Ressources* (MENJVA & DGESCO, 2012, p. 42)

Pour la loi exponentielle, après avoir imposé la fonction de densité, la formule pour calculer les probabilités est déterminée à l'aide d'un calcul intégral. La définition de l'espérance est aussi imposée : « *L'espérance est définie comme la limite quand x tend vers $+\infty$ de $\int_0^x xf(x)\,\mathrm{d}x$ où f est la fonction de densité de la loi exponentielle considérée* ». Après avoir défini l'espérance de la loi uniforme, la définition de l'espérance de la loi exponentielle peut être vue comme une extension de l'autre. La question de la possibilité ou non de l'étendre à un intervalle non borné n'apparaît pas. L'espérance de la loi exponentielle existe seulement car la limite en question est finie (et non seulement parce qu'elle existe). Nous retrouvons ici le problème des intégrales impropres qui ne sont finalement pas abordées du tout, comme si cela ne posait aucun problème. Pourtant l'idée qu'un domaine infini ait une aire finie n'est pas simple à concevoir.

Dans le document, il est détaillé trois méthodes pour démontrer la valeur de l'espérance d'une variable aléatoire suivant une loi exponentielle. Nous rappelons que, dans l'analyse du programme, nous avions repéré une « incohérence » sur ce point, du fait que l'intégration par parties n'est pas un attendu du programme. Les deux premières méthodes proposées pour contourner le fait que l'intégration par parties ne soit pas au programme sont les suivantes :

- l'une propose de chercher une primitive de $\lambda te^{-\lambda t}$ sous la forme $(at+b)e^{-\lambda t}$, où a et b sont à déterminer,
- l'autre de vérifier que la fonction $t \mapsto -te^{-\lambda t}$ est la primitive de $\lambda te^{-\lambda t}$ qui s'annule en 0.

La solution apportée est donc de donner directement la forme d'une primitive, vu qu'elle n'est pas connue des élèves. La troisième méthode est d'utiliser la méthode d'intégration par parties directement sur cet exemple particulier. Il est bien écrit que ce n'est pas une capacité exigible du programme, mais que le principe pourra éventuellement être expliqué sur cet exemple. L'utilisation de l'intégration par par-

ties n'est donc pas « interdite » mais elle reste appliquée simplement à cet exemple et ne doit pas être expliquée de façon formelle. On trouve donc une solution au fait que l'intégration par parties ne soit pas au programme dans le calcul de $\int_a^b xe^{-\lambda x}\,\mathrm{d}x$. En ce qui concerne le droit de s'intéresser à la limite d'une telle intégrale, rien n'est précisé, ni même questionné. Cela parait aller de soi et être « normal ».

b) Des démonstrations qui exploitent ces liens

On retrouve dans ce document plusieurs autres démonstrations, qui sont exigibles, et, comme on pouvait s'y attendre, à l'intérieur desquelles sont nettement repérables des jeux entre probabilités et analyse. Nous venons de le voir pour la démonstration de l'espérance de la loi exponentielle ; il en est de même pour :

- la propriété de durée de vie sans vieillissement ;
- la démonstration du théorème sur l'existence et l'unicité du réel u_α tel que $P(-u_\alpha \leqslant X \leqslant u_\alpha) = 1 - \alpha$, si X est une variable aléatoire suivant la loi normale $\mathcal{N}(0;1)$, dans laquelle intervient la fonction H définie par $H(u) = \int_0^u f(x)\,\mathrm{d}x$;
- la démonstration de l'espérance de la loi normale faisant intervenir du calcul intégral avec, dans ce cas, une primitive explicite.

Il est même ensuite proposé de calculer la variance en utilisant une méthode analogue à celle utilisée pour le calcul de l'espérance d'une loi exponentielle. Il est donc proposé ici d'appliquer la méthode d'intégration par parties à ce deuxième exemple.

c) La méthode de Monte Carlo

Nous retrouvons, comme proposé dans le programme, la méthode de Monte-Carlo qui est détaillée dans l'annexe 5 du document *Ressources* (pp. 66-69). Il est précisé que la méthode de Monte-Carlo est « *une méthode probabiliste permettant le calcul approché d'intégrales (simples ou multiples) de fonctions quelle que soit leur régularité* ». Deux méthodes, la méthode dite du « rejet » et la méthode de l'espérance, sont présentées pour déterminer une valeur approchée de l'intégrale $\int_0^1 f(x)\,\mathrm{d}x$ pour une fonction continue sur $[0;1]$ et à valeurs dans $[0;1]$.

Nous pouvons voir, dans ce cas, le statut s'inverser entre les probabilités et l'intégrale. Ici, ce sont les probabilités qui sont un outil pour le calcul intégral, alors que dans la grande majorité des cas, c'est l'inverse.

4.3.5 Des tâches qui n'exploitent pas l'articulation entre probabilités à densité et calcul intégral

Dans ce document *Ressources*, quelques exemples d'exercices sont donnés. Cela nous permet d'identifier quelles sont les tâches qui peuvent être proposées aux élèves et qui exploitent les liens entre probabilités et analyse. Le constat est décevant car aucun exercice n'exploite réellement cette relation entre les deux sous-domaines.

Pour commencer, aucun exemple d'exercice n'est donné, ni même d'exemples d'expériences aléatoires correspondant à la loi exponentielle. Pour la loi uniforme, ils sont au nombre de deux (cf. Annexes C.2 et C.3). Dans ces deux exercices, nous pouvons remarquer que la fonction de densité n'est absolument pas nécessaire pour résoudre les différentes tâches, bien qu'elle soit demandée dans l'exercice 2 (partie A). Un raisonnement « intuitif » suffirait. Dans la partie B de l'exercice 2, la variable aléatoire en question suit en fait une loi de Xénakis (différence en valeur absolue de deux variables aléatoires uniformes sur $[0;1]$) mais cette loi n'est pas connue des élèves et la résolution donnée ne propose pas de déterminer la fonction de densité associée à cette loi mais seulement de déterminer la probabilité à l'aide d'une configuration géométrique (graphique). Cet exercice est analysé dans la thèse de Nechache (2016). Ces exercices ne mettent donc pas en avant l'intérêt de la fonction de densité et ce qu'elle représente. Ils ne suscitent pas non plus l'utilisation d'outils technologiques, comme les logiciels ou la calculatrice, et aucune démarche de modélisation n'est en jeu. Dans chaque cas, le modèle est imposé par l'énoncé et jamais questionné, alors que les lois des variables aléatoires considérées, qui sont des temps d'arrivée, pourraient être discutées : les temps d'arrivée suivent-ils toujours des lois uniformes ? Il s'agit donc ici seulement d'un habillage de l'énoncé pour calculer des probabilités associées à une loi uniforme.

Enfin, pour la loi normale, un premier exercice (p. 7) est proposé. Il s'agit d'une étude de la fonction $x \mapsto e^{-\frac{x^2}{2}}$. La fonction est issue des probabilités (c'est ce que nous avons appelé une fonction de référence probabiliste), mais le problème proposé fait partie du domaine de l'analyse : il s'agit d'une étude de fonction. On peut ici repérer un lien fait entre les deux domaines, mais cela reste très léger (et n'est pas au niveau du calcul intégral).

Des pages 12 à 17 sont proposés des exemples d'exercices sur les lois normales. On peut repérer que sur les six exercices, seul le premier, qui est théorique, utilise la fonction de densité et ses propriétés (ici la parité). Dans les autres exercices, tous contextualisés à l'aide d'énoncés « fabriqués », la fonction de densité, par son expression ou sa représentation graphique, n'apparait jamais. Comme pour la loi uniforme, à chaque fois, le modèle est donné. Il n'y a aucune réflexion à avoir sur ce modèle sauf parfois sur les paramètres, qui sont à déterminer. Dans l'exercice 3 (p. 14), ce sont des données statistiques qui permettent d'estimer les paramètres.

Nous pouvons observer la présence de tableaux pour expliquer les démarches à suivre pour les calculs liés à une probabilité associée à une loi normale, ainsi que ceux liés à une probabilité « inverse » de la loi normale, à l'aide du logiciel *R* ou d'une calculatrice, de marque TI ou CASIO [18]. Nous remarquons ici la présence indispensable de l'artefact calculatrice ou logiciel pour calculer des probabilités associées à la loi normale. Cela semble créer une rupture avec le calcul intégral, qui ne propose pas ce type de résolution dans le cas d'intégrales dont on ne connait pas de primitive. Le document *Ressources* insiste fortement sur le logiciel *R* en lui consacrant notamment toute une annexe, pour apprendre à l'utiliser. Cet artefact semble mis en avant.

Pour conclure, dans aucun des exercices proposés, l'expression de la fonction

18. Il s'agit des deux marques les plus utilisées dans les classes.

de densité n'est réellement utile : cela ne met donc pas en avant son intérêt et efface le lien entre probabilités à densité et calcul intégral. Pour la loi uniforme, une approche « intuitive » suffit et, pour la loi normale, c'est la calculatrice (ou le logiciel) qui prend tous les calculs en charge. Cela serait sûrement différent pour la loi exponentielle.

Finalement nous pouvons remarquer que seules les démonstrations exploitent les articulations entre les deux sous-domaines.

4.3.6 Conclusion sur le document *Ressources*

L'analyse du document *Ressources* donne des éclaircissements et des enrichissements par rapport au programme. Dans ce document, nous pouvons remarquer qu'il est question exclusivement des trois lois au programme. Cela laisse donc penser qu'il s'agit de faire, en classe de terminale S, trois études de cas. Il n'est fait aucune référence à d'autres lois, ni même à des propriétés générales sur la fonction de densité ou encore sur les probabilités, mis à part un commentaire. Seule la fonction de densité de la loi normale centrée réduite est introduite par une activité de visualisation du théorème de Moivre Laplace (comme demandé dans le programme) notamment avec l'utilisation du logiciel *GeoGebra*. Cependant, la courbe est déjà tracée. Il n'y a pas de réflexion sur le choix de cette courbe et donc sur ses propriétés. Nous ne trouvons aucune indication sur la façon d'introduire la fonction de densité de façon « générale », ni même dans les cas spécifiques de la loi uniforme et de la loi exponentielle. Elle semble être imposée, sans que soit posée la question de ce qu'est cet objet, d'où il vient et pourquoi il sert à calculer des probabilités.

Le domaine de la statistique descriptive (mis à part très rapidement pour généraliser la loi normale) n'est jamais utilisée pour créer du lien entre probabilités et analyse.

Concernant l'absence de l'intégration par parties dans le programme, la question est clairement prise en compte et le document propose plusieurs méthodes pour le calcul de l'espérance.

Pour ce qui est des intégrales impropres, aucun problème n'est soulevé. Le passage à la limite semble aller de soi, sans engendrer d'éventuels problèmes dans cette extension.

Sur la description du plan épistémologique de l'ETM de référence, les relations sémiotiques restent bien présentes et quelques enrichissements sont proposés pour les justifier. Au niveau du théorème de Moivre Laplace, c'est une justification par la visualisation à l'aide d'un artefact (logiciel *GeoGebra*) qui est proposée. De même, on retrouve ce type de visualisation pour l'espérance de la loi uniforme. L'artefact semble ici jouer un rôle dans l'articulation entre les deux sous-domaines. Cependant, le point de départ du lien entre les deux, qui repose sur l'introduction de la notion de fonction de densité (dans le cas général ou tout du moins pour les lois uniformes et exponentielles) n'est absolument pas justifié. Cela semble devoir être imposé aux élèves.

Nous avons aussi remarqué qu'aucune tâche n'est proposée, mis à part les démonstrations de cours, pour exploiter les liens entre probabilités et analyse. Au

niveau de la loi normale, il semblerait que la composante artefact, avec la calculatrice ou autre logiciel, soit responsable d'une rupture entre les deux sous-domaines. La fonction de densité, que ce soit pour la loi uniforme ou la loi normale, n'apparaît pas indispensable.

Les sujets du baccalauréat vont nous aider à identifier des tâches exploitant les liens entre probabilités à densité et calcul intégral.

4.4 Analyse des sujets de baccalauréat

Dans cette section, nous allons analyser les tâches concernant les deux sous-domaines qui nous intéressent, proposées au baccalauréat de la filière scientifique. Le but est d'essayer d'en dégager des tendances et peut-être des tâches typiques et de voir si des liens ou des interactions sont présents ou pourraient l'être dans ces tâches.

Bien entendu, les sujets de baccalauréat ne proposent pas l'ensemble des tâches possibles et envisageables dans la classe de terminale S, mais donnent une base de ce qui doit être travaillé *a minima*, dans un souci de réussite à l'examen. Ces tâches ne sont évidemment pas des tâches d'introduction de la notion.

4.4.1 Présentation des sujets de baccalauréat

a) Choix des sujets

Étant donné que le programme actuel est en vigueur depuis la rentrée 2012, les sujets de baccalauréat ont pris en compte ces changements depuis la session 2013. Nous avons donc à disposition au moment de l'étude seulement trois années d'examen. Pour avoir un échantillon assez large de tâches, nous avons décidé de considérer l'ensemble des sujets de l'épreuve de mathématiques du baccalauréat de tous les centres d'examen de ces trois années. Nous disposons de 13 sujets pour chaque « session » (ils ne correspondent pas forcément tous à la même année civile, car suivant les pays, les examens ne sont pas au même moment de l'année) soit de 39 sujets au total (baccalauréat 2013, 2014 et 2015). L'ensemble de ces sujets est disponible sur le site de l'APMEP [19]. Ce nombre assez conséquent permet de disposer d'un panel assez large d'exercices qui doivent nous permettre de dégager des tendances générales.

b) Organisation d'un sujet de baccalauréat

Chaque sujet de baccalauréat comprend cinq exercices : trois exercices communs à chaque élève, un exercice pour les élèves suivant l'enseignement de spécialité et un exercice pour ceux n'ayant pas suivi l'enseignement de spécialité. Ainsi chaque élève doit résoudre quatre exercices. Nous nous intéressons dans ces sujets seulement aux exercices traitant des probabilités à densité et du calcul intégral.

19. www.apmep.fr

c) Grille d'analyse des exercices

Nous allons essayer de dégager des tendances dans les tâches proposées au baccalauréat. Dans un premier temps, nous chercherons l'ensemble des exercices ou plutôt des questions se rapportant aux probabilités à densité et celles se rapportant au calcul intégral. Pour ce faire, nous passerons en revue tous les exercices de chaque sujet et repérerons les domaines en question, en fonction du contexte de l'exercice et des registres utilisés. En annexe D, se trouvent la liste des sujets et des exercices correspondants aux deux sous-domaines que nous étudions. Nous ferons ensuite une étude quantitative pour mettre en évidence la fréquence des tâches portant sur les probabilités à densité et le calcul intégral.

Dans un second temps, nous nous focaliserons plus particulièrement sur chaque exercice sélectionné précédemment. Pour étudier chaque énoncé, guidé par le modèle des ETM, nous dégagerons plusieurs points :

- la proportion de questions de l'exercice concernées par le sous-domaine des probabilités à densité ou du calcul intégral ;
- le contexte de l'énoncé (mathématique, extra-mathématique) ;
- les notions abordées et les types de tâches [20] en jeu dans les questions nous intéressant ;
- les registres (Duval, 1993) en jeu dans l'énoncé des questions, notamment avec un regard particulier sur la présence ou non du registre graphique ;
- la présence ou non d'artefacts (calculatrice, table...) ;
- la présence ou la possibilité de liens avec le calcul intégral, si nous analysons un exercice de probabilités, et inversement, dans l'autre cas.

L'analyse suivant cette grille est disponible en annexes D.2 et D.3. Les énoncés des exercices dont nous parlerons plus spécifiquement se trouvent en Annexe D.4.

La description des types de tâches en jeu plus détaillée pour les probabilités à densité, permettra de dégager une ou plusieurs tâches classiques relatives aux lois à densité.

4.4.2 Répartition des exercices de baccalauréat portant sur les probabilités à densité et le calcul intégral

Sur l'ensemble des 39 sujets de baccalauréat étudiés, une grande majorité comporte une ou plusieurs questions portant sur les probabilités à densité et le calcul intégral : 33 sujets abordent des questions relatives aux lois à densité et 30 sujets des questions relatives au calcul intégral (cf. figures 4.4 et 4.5).

20. Nous regroupons les tâches avec une catégorisation en types de tâches. Il ne faut pas voir de connotation particulière à la théorie anthropologique du didactique (Chevallard, 1999).

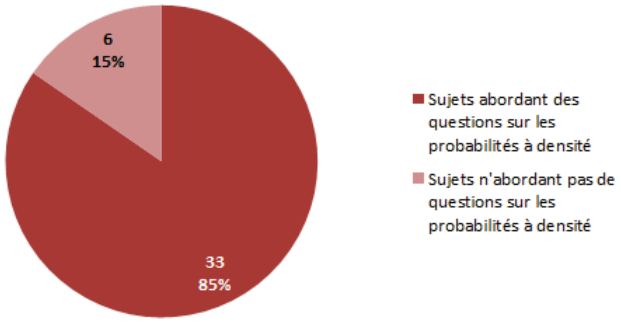


FIGURE 4.4 – Répartition des sujets de baccalauréat entre 2013 et 2015 comportant des questions relatives aux probabilités à densité

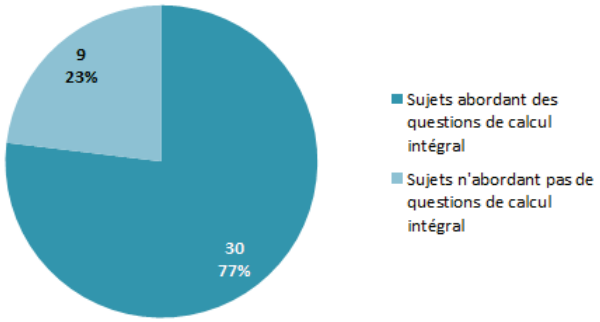


FIGURE 4.5 – Répartition des sujets de baccalauréat entre 2013 et 2015 comportant des questions relatives au calcul intégral

Seuls deux sujets comportent deux exercices différents traitant tous les deux de calcul intégral. Sinon dans les autres cas, seul un exercice est concerné pour chaque domaine.

Aucun des exercices n'est entièrement consacré à l'un ou à l'autre des sous-domaines, mis à part dans trois sujets, où les exercices concernés portent entièrement sur le calcul intégral. Le sujet de Métropole - La Réunion de septembre 2015 comprend, lui, deux exercices portant sur le calcul intégral et les deux sont entièrement dédiés à ce sous-domaine (cf. Annexe D.4.1). Nous pouvons tout de même nuancer ce propos car, dans trois exercices sur les quatre en question, il s'agit d'exercices traitant d'une étude de suite définie par une intégrale, donc le choix de considérer toutes les questions comme portant sur le calcul intégral est discutable. Quoi qu'il en soit, dans tous les autres sujets (et il s'agit d'une majorité des cas), que ce soit pour les probabilités à densité ou le calcul intégral, ces sous-domaines sont mélangés à d'autres sous-domaines, la plupart du temps rattachés au même domaine. Il n'est pas rare que le domaine des probabilités soit aussi au contact du domaine de la statistique dans le même exercice (cf. Annexe D.4.2). Cependant, les questions

relatives aux probabilités et à la statistique inférentielle sont indépendantes. L'étude de lois à densité est faite, dans la majorité des cas, en parallèle avec la loi binomiale et/ou des problèmes d'estimation.

Le calcul intégral constitue la plupart du temps une partie d'un exercice traitant une étude de fonction. Seul dans un sujet (Nouvelle-Calédonie, novembre 2015, cf. Annexe D.4.3), l'exercice traitant des probabilités à densité est le même que celui relatif au calcul intégral. Cependant, les questions relatives à l'un ou à l'autre des sous-domaines ne sont pas dans la même partie de l'exercice (les parties étant indépendantes).

Nous pouvons déjà dire que les deux sous-domaines sont clairement séparés dans les exercices, mais fortement présents dans les sujets de baccalauréat. Les pourcentages des figures 4.4 et 4.5 reflètent bien l'importance accordée à ces deux sous-domaines dans le programme actuel. Dans la section 4.2 de ce chapitre, nous avons précisé qu'environ un huitième de l'année scolaire de terminale S est consacré à l'enseignement des probabilités à densité et il en est de même pour le calcul intégral.

Voyant la séparation faite dans les exercices entre les deux sous-domaines, nous allons repérer des éléments relatifs à chacun de ces sous-domaines, les différenciant et ensuite pouvant les rassembler.

4.4.3 Différences repérées entre les deux sous-domaines

Une analyse des exercices de baccalauréat a été faite en Annexes D.2 et D.3, suivant la grille d'analyse présentée plus haut. Ces analyses nous permettent de faire quelques constats. Dans cette section, nous allons détailler certaines différences remarquables repérées entre ces deux sous-domaines.

a) Registres de représentations sémiotiques et place de la visualisation

Dans les questions relatives aux probabilités à densité, nous repérons bien les registres correspondants, tels le registre symbolique probabiliste et la langue naturelle ; de même pour le calcul intégral, avec le registre symbolique intégral et la langue naturelle. Cependant, nous pouvons remarquer une différence quant au registre graphique, qui aurait lieu d'être dans ces deux sous-domaines. Un graphique accompagne les questions de calcul intégral dans pratiquement 85% des cas, et une interprétation graphique de l'intégrale peut être demandée. Dans le cas des probabilités à densité, le graphique est nettement moins présent (on pourrait même dire absent) : seuls trois exercices sur les 33 en proposent un et proposent un travail dessus. La place de la visualisation n'est donc pas la même dans les sous-domaines, au niveau du baccalauréat. Ce n'est pourtant pas ce qui ressort du document *Ressources*.

b) Utilisation d'artefact

Dans le paragraphe précédent, nous n'avons pas parlé d'un registre, qui se trouve pourtant présent dans 7 sujets sur 9 de la session 2013 : le registre table. Dans tous ces sujets, une table de la loi normale est fournie (celle en question dans l'exercice ou celle de la loi normale centrée réduite).

La présence de ces tables est inattendue au vu des recommandations du programme et du document *Ressources*, qui précisent la diminution de l'intérêt des tables avec la calculatrice. Cependant, en observant la disparition totale de ces tables dès la session suivante du baccalauréat, nous pouvons faire l'hypothèse que les tables étaient une aide pour ne pas pénaliser des élèves qui n'auraient pas appris à utiliser la calculatrice pour déterminer des probabilités associées à une loi normale.

Dès 2014, l'utilisation de la calculatrice est indispensable dans une grande partie des cas. Parfois elle est possible mais non obligatoire, car l'utilisation des propriétés sur les intervalles « 1, 2, 3 sigmas » est applicable.

Pour la loi exponentielle, la calculatrice peut être requise mais seulement pour disposer d'une valeur approchée, ce que nous ne considérerons pas ici comme une utilisation d'artefact relative au sous-domaine des probabilités à densité. On retrouve aussi ce type d'utilisation en calcul intégral.

Nous remarquons donc un rapport totalement différent à la calculatrice dans le sous-domaine des probabilité à densité et celui du calcul intégral : l'artefact calculatrice est indispensable dans le cadre de l'étude de la loi normale.

c) Contexte des énoncés

Le contexte des exercices est différent pour les probabilités à densité et le calcul intégral. Pour le calcul intégral, dans près de 80% des cas, les exercices proposés sont dans un contexte purement mathématique ; ce qui va complètement à l'inverse des exercices de probabilités à densité, où seul un exercice dans les 39 sujets est entièrement dans ce contexte (Métropole, juin 2015, cf. Annexe D.4.4).

Les contextes sont variables en probabilités, cela va du temps de parcours entre le domicile et le lycée aux normes de conformité (par exemple, le poids d'une baguette). Bien entendu, il s'agit de contextualisation « fabriquée » pour l'exercice.

4.4.4 Des types de tâches caractérisant les deux sous-domaines

Le détail plus poussé des types de tâches dans le sous-domaine des probabilités à densité s'explique par le fait que les mêmes types de tâches reviennent souvent. Nous pourrons donc en dégager des types de tâches caractéristiques. Ce n'est pas le cas pour le calcul intégral, où beaucoup plus de choix sont proposés.

a) Le sous-domaine des probabilités à densité

En premier constat, nous pouvons remarquer que, sur les 33 sujets comportant des questions relatives aux probabilités à densité, seules les lois exponentielles et normales sont étudiées, avec une majorité pour les lois normales (cf. figure 4.6). Aucune question ne porte sur la loi uniforme, ou bien sur une autre loi. Ces choix renforcent bien l'idée que le programme favorise l'étude de cas, mais on peut pointer ici la dévalorisation de la loi uniforme.

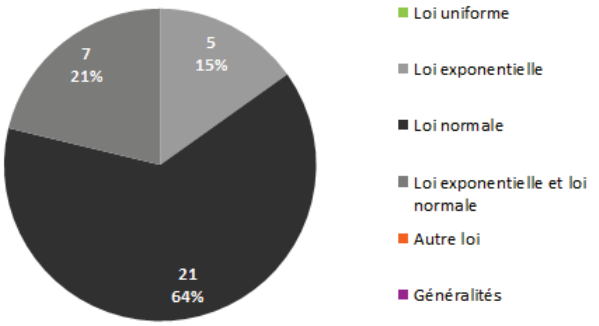


Figure 4.6 – Répartition des lois étudiées dans les sujets de baccalauréat entre 2013 et 2015

Les exercices de baccalauréat relatifs aux probabilités à densité sont extrêmement « formatés ».

Autour de la loi exponentielle, nous retrouvons toujours les mêmes types de tâches qui sont les suivantes :

- Déterminer une probabilité ;
- Déterminer une probabilité conditionnelle ;
- Déterminer le paramètre de la loi exponentielle, connaissant l'espérance ou une probabilité ;
- Déterminer l'espérance.

Nous avons bien fait attention à utiliser le terme « déterminer » plutôt que « calculer » car les activités attendues des élèves ne sont pas nécessairement un calcul. Il peut aussi y avoir des ROC (restitutions organisées des connaissances), mais cela est ponctuel.

Les types de tâches les plus présents pour les questions portant sur la loi normale sont les suivantes :

- Déterminer une probabilité ;
- Reconnaître la loi normale centrée réduite associée à une loi normale ;
- Déterminer le paramètre μ ou σ de la loi normale en jeu, connaissant une probabilité.

Dans le cas de la loi normale, comme nous l'avons identifié dans l'analyse du document *Ressources*, les activités attendues des élèves sont internes au sous-domaine des probabilités à densité avec l'utilisation de la calculatrice (ou de la table en 2013) pour déterminer des probabilités ou encore l'utilisation des propriétés « 1, 2 ou 3 sigmas ». La détermination des paramètres se fait en passant par la loi normale centrée réduite. La relation entre probabilités à densité et calcul intégral est inexistant avec la loi normale.

Au niveau de la loi exponentielle, en revanche, l'utilisation du calcul intégral pour effectuer les différents calculs (de probabilités, d'espérance, de paramètre) est possible. Cette activité possible exploite les liens entre probabilités à densité et calcul intégral. Cependant, nous ne pouvons pas ne pas ignorer une autre activité de résolution possible : l'utilisation directe des formules. Les énoncés, n'explicitant

jamais la méthode attendue, laissent cette possibilité.

b) Le sous-domaine du calcul intégral

Le même travail de précision des types de tâches n'a pas été fait pour le calcul intégral, car nous ne retrouvons pas, comme en probabilités, une standardisation aussi marquée. Nous pouvons cependant repérer des « thèmes » récurrents :

- Calcul d'aire ;
- Calcul d'intégrale ;
- Étude de suite définie par une intégrale du type $I_n = \int_a^b f_n$;
- Étude de suite définie par une intégrale du type $I_n = \int_a^n f$;
- Étude de la fonction Aire ;
- Méthode des rectangles.

4.4.5 Des tâches qui exploitent les liens entre les deux sous-domaines ou pouvant en créer

a) Des liens repérés dans les sujets de baccalauréat

Dans les tâches de probabilités à densité, la seule créant directement du lien avec le calcul intégral est le calcul de probabilité associé à la loi exponentielle, si l'on suppose que l'activité attendue est de refaire le calcul et non d'utiliser directement la formule (ce qui n'est pas explicitement précisé dans les exercices).

Les autres endroits où nous pouvons retrouver des interactions entre probabilités à densité et calcul intégral sont dans les ROC (restitutions organisées des connaissances), comme nous l'avons vu aussi à travers l'analyse du document *Ressources*. Il s'agit de la démonstration de propriétés du cours, et dans tous les cas, il est nécessaire de passer par le domaine de l'analyse pour les faire. Nous avons repéré la démonstration de la formule de l'espérance de la loi exponentielle (2 fois), la démonstration de la formule de $P(c \leqslant X \leqslant d)$, où X suit une loi exponentielle (1 fois) et enfin, la démonstration de la propriété : si X est une variable aléatoire qui suit la loi $\mathcal{N}(0;1)$, alors pour tout nombre réel $\alpha \in]0;1[$, il existe un unique nombre réel $u_\alpha > 0$ tel que $P(-u_\alpha \leqslant X \leqslant u_\alpha) = 1 - \alpha$ (1 fois).

Dans le calcul intégral, il n'y a aucune interaction explicite proposée avec les probabilités à densité. Cependant, dans certains exercices, des liens seraient envisageables. Nous allons présenter ces potentialités.

b) Des tâches propices à un lien entre les deux sous-domaines

Certains exercices traitent de l'étude de suites définies par une intégrale du type $I_n = \int_a^n f$. Prenons l'exemple de l'exercice 2 (partie A) du sujet de Métropole - La Réunion de septembre 2015 (cf. figure 4.7 et Annexe D.4.1).

Soit f la fonction définie et dérivable sur l'intervalle $[0\,;+\infty[$ telle que :

$$f(x)=\frac{x}{\mathrm{e}^x-x}$$

On admet que la fonction f est positive sur l'intervalle $[0\,;+\infty[$.
On note $\mathscr{C}$ la courbe représentative de la fonction f dans un repère orthogonal du plan.
La courbe $\mathscr{C}$ est représentée en annexe, **à rendre avec la copie**.

Partie A

Soit la suite (I_n) définie pour tout entier naturel n par $I_n=\displaystyle\int_0^n f(x)\,\mathrm{d}x$.

On ne cherchera pas à calculer la valeur exacte de I_n en fonction de n.

1. Montrer que la suite (I_n) est croissante.
2. On admet que pour tout réel x de l'intervalle $[0\,;+\infty[$, $\mathrm{e}^x-x\geqslant\dfrac{\mathrm{e}^x}{2}$.
   a. Montrer que, pour tout entier naturel n, $I_n\leqslant\displaystyle\int_0^n 2x\mathrm{e}^{-x}\,\mathrm{d}x$.
   b. Soit H la fonction définie et dérivable sur l'intervalle $[0\,;+\infty[$ telle que :

   $$H(x)=(-x-1)\mathrm{e}^{-x}$$

   Déterminer la fonction dérivée H' de la fonction H.
   c. En déduire que, pour tout entier naturel n, $I_n\leqslant 2$.
3. Montrer que la suite (I_n) est convergente. On ne demande pas la valeur de sa limite.

Figure 4.7 – Extrait de l'exercice 2 du sujet de baccalauréat de Métropole - La Réunion (septembre 2015)

Il s'agit ici de l'étude de la convergence d'une intégrale impropre, par le biais de l'introduction d'une suite de la forme $I_n=\displaystyle\int_a^n f$. On peut remarquer que ce type de travail pourrait être fait en probabilités, notamment pour justifier que l'aire sous la courbe de la fonction de densité de la loi exponentielle sur l'intervalle $[0;+\infty[$ vaut 1, ou encore pour la démonstration de la formule de l'espérance de la loi exponentielle. D'ailleurs, si l'on regarde bien cet exercice, on s'aperçoit que la fonction qui permet la comparaison est la fonction $x\mapsto xe^{-x}$, à constante multiplicative près, qui est la fonction à intégrer pour déterminer l'espérance de la loi exponentielle. Dans cet exercice, il n'est pas mentionné d'interprétation graphique, donc aucune relation à l'aire n'est faite.

À l'inverse, c'est le choix qui est fait dans l'exercice 4 du sujet d'Amérique du Nord de mai 2013 (cf. Annexe D.4.5), où il s'agit de l'étude d'une suite définie par l'aire sous une courbe entre les droites d'équation $x=1/e$ et $x=n$ (cf. figure 4.8). Il s'agit ici d'interpréter l'étude en terme d'aire finie d'un domaine infini.

4. Pour tout entier $n \geqslant 1$, on note I_n l'aire, exprimée en unités d'aires, du domaine délimité par l'axe des abscisses, la courbe $\mathscr{C}$ et les droites d'équations respectives $x = \dfrac{1}{\mathrm{e}}$ et $x = n$.

a. Démontrer que $0 \leqslant I_2 \leqslant \mathrm{e} - \dfrac{1}{2}$.

On admet que la fonction F, définie sur l'intervalle $]0\,;\,+\infty[$ par $F(x) = \dfrac{-2-\ln(x)}{x}$, est une primitive de la fonction f sur l'intervalle $]0\,;\,+\infty[$.

b. Calculer I_n en fonction de n.

c. Étudier la limite de I_n en $+\infty$. Interpréter graphiquement le résultat obtenu.

FIGURE 4.8 – Extrait de l'exercice 4 du sujet de baccalauréat d'Amérique du Nord (mai 2013)

Dans l'exercice 3 du sujet des Antilles - Guyane de juin 2013 (cf. Annexe D.4.6), il s'agit aussi d'une étude d'aire finie d'un domaine infini, mais la proposition ici est d'étudier la fonction Aire et non la suite (cf. figure 4.9). L'aire étudiée est l'aire entre deux courbes et non l'aire sous une courbe, mais c'est simplement une adaptation qui pourrait être faite.

b. Dans cette question, a désigne un réel positif, D_a la partie du plan comprise entre $\mathscr{C}_{\mathrm{e}}$, $\mathscr{C}_{-\mathrm{e}}$, l'axe (Oy) et la droite Δ_a d'équation $x = a$. On désigne par $\mathscr{A}(a)$ l'aire de cette partie du plan, exprimée en unités d'aire.

Démontrer que pour tout réel a positif : $\mathscr{A}(a) = 2\mathrm{e} - 2\mathrm{e}^{1-a}$.

En déduire la limite de $\mathscr{A}(a)$ quand a tend vers $+\infty$.

FIGURE 4.9 – Extrait de l'exercice 3 du sujet de baccalauréat des Antilles - Guyane (juin 2013)

Nous voyons donc dans ce type d'exercice, qui finalement revient à l'étude d'intégrales impropres, un potentiel pour faire un lien entre probabilités à densité et calcul intégral.

De plus, comme le premier exemple l'a montré, le choix des fonctions aussi privilégie une interaction entre les deux sous-domaines. Dans l'ensemble des sujets, on peut voir apparaître des fonctions du type $x \mapsto xe^{-x}$, $x \mapsto -(x+1)e^{-x}$, $x \mapsto xe^{-x^2}$. Nous identifions dans le choix de ces fonctions un lien direct qui pourrait s'opérer avec les probabilités. Cependant, ce n'est pas ce qui est attendu dans ces exercices. Les exercices pourraient être vus d'un double point de vue : probabiliste et analytique.

Pour résumer, plusieurs types d'exercices pourraient être propices à des interactions entre les sous-domaines, bien que ce ne soit pas fait :

- des exercices portant sur des intégrales impropres (suites, étude de la fonction Aire...) ou en d'autres termes, portant sur des aires finies de domaines infinis ;
- des exercices dans lesquels les fonctions en question sont du type fonction de référence probabiliste : $x \mapsto xe^{-x}$, $x \mapsto -(x+1)e^{-x}$, $x \mapsto xe^{-x^2}$...

c) Un exemple à forte potentialité : une occasion manquée ?

L'exercice 3 du sujet du Liban de mai 2014 (cf. Annexe D.4.7) possède les critères dégagés ci-dessus et même d'autres enrichissant encore les possibilités.

Il s'agit dans cet exercice de l'étude de la fonction f définie sur $[0;+\infty[$ par $f(x) = xe^{-x}$. On repère, comme dit plus haut, la fonction à intégrer pour déterminer l'espérance de la loi exponentielle de paramètre 1. Après une étude de cette fonction, on s'intéresse à une nouvelle fonction $\mathscr{A}$ qui est l'aire sous la courbe de f entre 0 et t. Voici la partie B de l'exercice :

Partie B

Soit $\mathscr{A}$ la fonction définie sur l'intervalle $[0\,;\,+\infty[$ de la façon suivante : pour tout réel t de l'intervalle $[0\,;\,+\infty[$, $\mathscr{A}(t)$ est l'aire, en unités d'aire, du domaine délimité par l'axe des abscisses, la courbe $\mathscr{C}$ et les droites d'équations $x = 0$ et $x = t$.

1. Déterminer le sens de variation de la fonction $\mathscr{A}$.
2. On admet que l'aire du domaine délimité par la courbe $\mathscr{C}$ et l'axe des abscisses est égale à 1 unité d'aire. Que peut-on en déduire pour la fonction $\mathscr{A}$?
3. On cherche à prouver l'existence d'un nombre réel α tel que la droite d'équation $x = \alpha$ partage le domaine compris entre l'axe des abscisses et la courbe $\mathscr{C}$, en deux parties de même aire, et à trouver une valeur approchée de ce réel.
   a. Démontrer que l'équation $\mathscr{A}(t) = \dfrac{1}{2}$ admet une unique solution sur l'intervalle $[0\,;\,+\infty[$
   b. Sur le graphique fourni en **annexe** (***à rendre avec la copie***) sont tracées la courbe $\mathscr{C}$, ainsi que la courbe Γ représentant la fonction $\mathscr{A}$.
   Sur le graphique de l'**annexe**, identifier les courbes $\mathscr{C}$ et Γ, puis tracer la droite d'équation $y = \dfrac{1}{2}$. En déduire une valeur approchée du réel α.
   Hachurer le domaine correspondant à $\mathscr{A}(\alpha)$.
4. On définit la fonction g sur l'intervalle $[0;\,+\infty[$ par
$$g(x) = (x+1)\,\mathrm{e}^{-x}.$$
   a. On note g' la fonction dérivée de la fonction g sur l'intervalle $[0\,;\,+\infty[$.
   Pour tout réel x de l'intervalle $[0\,;\,+\infty[$, calculer $g'(x)$.
   b. En déduire, pour tout réel t de l'intervalle $[0\,;\,+\infty[$, une expression de $\mathscr{A}(t)$.
   c. Calculer une valeur approchée à 10^{-2} près de $\mathscr{A}(6)$.

L'expression de $\mathscr{A}(t)$ peut être donnée par $\displaystyle\int_0^t f(x)\,\mathrm{d}x$, autrement dit $\displaystyle\int_0^t xe^{-x}\,\mathrm{d}x$.

Dans la question B. 2., il est admis que l'aire du domaine délimité par la courbe représentative de f et l'axe des abscisses est égale à 1, et il est demandé d'en déduire une information sur la fonction $\mathscr{A}$: que sa limite en $+\infty$ vaut 1. Ici, la question pourrait être vue d'un autre point de vue, celui des probabilités à densité. On pourrait très bien envisager de déterminer la limite de $\mathscr{A}$ en $+\infty$, en d'autres termes, la limite de $\displaystyle\int_0^t xe^{-x}\,\mathrm{d}x$ quand $t \to +\infty$. Il s'agit là de l'espérance de la loi exponentielle de paramètre 1, qui vaut 1 ($\frac{1}{\lambda}$ avec $\lambda = 1$).

La question B.4 pourrait aussi être une aide pour démontrer ce résultat (qui est dans le cas général une ROC exigible). De plus, on pourrait alors se poser la question de ce que l'on peut alors en déduire, non pas sur $\mathscr{A}$, mais sur f. En effet, on peut en déduire que la fonction f peut être vue aussi comme une fonction de densité sur

$[0;+\infty[$. La fonction $\mathscr{A}$ est alors la fonction de répartition associée (bien entendu cette notion n'étant pas au programme, cette dénomination ne serait pas à donner).

Ensuite, la question B.3 pourrait, elle aussi, être réinterprétée dans le domaine des probabilités. Si l'on considère X une variable aléatoire de fonction de densité f résoudre l'équation $\mathscr{A}(t) = \frac{1}{2}$ revient à résoudre $P(X \leqslant t) = \frac{1}{2}$. Cela revient à dire que l'on cherche la « médiane » de la variable aléatoire (parallèle avec la médiane d'une série statistique). Bien évidemment pour que ce type de changement de domaine soit envisageable, il ne faut pas que l'étude des lois à densité se limite à celle des trois lois au programme sans en extraire de généralités.

4.4.6 Conclusion sur les exercices du baccalauréat

Nous ne sommes pas surprise de constater que les tâches proposées dans le baccalauréat dans les deux sous-domaines sont fortement dissociées. Les tâches portant sur les probabilités à densité sont consacrées exclusivement aux lois exponentielles et normales, ce qui renforce l'idée d'études de cas.

Seules les tâches sur la loi exponentielle et les différentes ROC (restitutions organisées des connaissances) posées peuvent (mais pas obligatoirement) utiliser le calcul intégral. Nous avons mis en avant le fait que ces tâches sont extrêmement « formatées », où seul le contexte de l'exercice change mais il ne s'agit en fait que d'un habillage. La modélisation n'est pas du tout un attendu. Les tâches de calcul intégral sont, quant à elles, exclusivement bornées à l'analyse (il s'agit d'études de fonctions, de suites...). Nous avons mis en évidence des possibilités d'articulations entre les deux sous-domaines qui ne sont cependant jamais exploitées.

Cette analyse vient renforcer les analyses faites sur le document *Ressources*, qui tendaient à cloisonner l'étude des lois à densité aux seules trois lois au programme. Nous pouvons cependant nous questionner sur la loi uniforme. Est-elle trop simple pour figurer dans les sujets de baccalauréat ?

4.5 Conclusion sur les dynamiques entre les ETM de référence

Cette étude nous a permis d'identifier les différentes articulations qui existent entre probabilités à densité et calcul intégral. Les trois analyses complémentaires permettent de conclure que le programme semble cantonner l'étude des lois à densité à trois études de cas : la loi uniforme, la loi exponentielle et la loi normale. On pourrait même se demander si nous n'avons pas à faire à trois ETM disjoints pour les trois lois.

Quelle que soit la forme de généralisation, aucune n'apparaît dans les différents documents institutionnels. Notamment, il n'y a aucun élément pour l'introduction de la notion de fonction de densité de façon général. Seule la loi normale semble impérativement à introduire, avec la visualisation du théorème de Moivre Laplace.

Cependant, nous maintenons la position suivant laquelle la fonction de densité occupe une place centrale dans la compréhension des lois à densité et de la relation entre probabilités à densité et calcul intégral.

Nous avons cerné le plan épistémologique liant les deux ETM (avec quelques contours qui restent un peu mal définis) cependant la description reste floue sur le plan cognitif. Ce qui ressort de cette étude des ETM de référence est une dynamique entre les ETM des deux sous-domaines qui repose essentiellement sur des relations entre différents signes. Nous pouvons repérer des propositions d'articulations à l'aide de la visualisation par le biais d'artefact comme *GeoGebra* pour le théorème de Moivre Laplace, par exemple, mais cela reste ciblé sur la loi normale. Nous avons aussi mis en évidence que des liens notamment sémiotiques existent et pour autant ils ne sont jamais exploités dans les tâches données aux élèves, mis à part dans les démonstrations de cours et pour la loi exponentielle ; bien que, comme nous l'avons présenté, des potentialités existent.

De plus, il nous semble que les TICE, dans le cas de la loi normale, créent une rupture entre les deux sous-domaines. Cela nous a permis d'entrevoir les articulations « réelles » et, d'autres, possibles entre l'espace de travail mathématique associé aux probabilités à densité et celui relatif au calcul intégral, au niveau de l'ETM de référence, essentiellement dans le plan épistémologique. Nous avons pu mettre en lumière que ces liens entre les deux ETM sont des relations sémiotiques qui semblent la plupart du temps imposées (mis à part dans le cas de la loi normale centrée réduite pour laquelle une justification du lien entre probabilité et intégrale semble devoir se faire à l'aide de la visualisation par l'intermédiaire d'un logiciel (*GeoGebra* est préconisé dans le document *Ressources*)).

Pour conclure, nous pouvons dire qu'au niveau des ETM de référence, les relations entre probabilités à densité et calcul intégral existent mais sont presque exclusivement sémiotiques. Aucun sens, aucune justification ne semblent donnés à ces liens, qui finalement disparaissent même dans les tâches prescrites aux élèves, au baccalauréat par exemple.

Au niveau de la fonction de densité, aucune introduction n'est proposée et de ce fait, aucun lien n'est fait entre les probabilités à densité et la statistique descriptive.

Cela nous a permis d'entrevoir les articulations « réelles » et, d'autres, possibles entre l'espace de travail mathématique associé aux probabilités à densité et celui relatif au calcul intégral, au niveau de l'ETM de référence, essentiellement dans le plan épistémologique. Nous avons pu mettre en lumière que ces liens entre les deux ETM reposent sur des écritures sémiotiques qui semblent la plupart du temps imposées. Pour le signe $(*)$, il semble finalement que l'on en reste à une écriture sémiotique sans atteindre le sens de la relation qu'elle établit.

Malgré les conclusions de cette étude, nous avons pleinement conscience que le programme ainsi que le document *Ressources* ne sont pas exhaustifs, qu'ils proposent des contours flous notamment pour laisser une certaine marge de manœuvre à l'enseignant. Ce qui n'est pas écrit dans ces documents n'est pas nécessairement à exclure de l'étude. C'est la raison pour laquelle nous allons poursuivre notre travail en nous demandant maintenant ce qu'il en est au niveau des ETM idoines. Qu'est-ce qui est réellement proposé en classe ? Retrouvons-nous les mêmes conclusions ? Nous verrons que des choix différents sont faits dans les manuels.

Chapitre 5

Étude des manuels de terminale S

Dans ce chapitre, nous allons passer à une analyse au niveau des ETM idoines. Nous ne voulons plus nous cantonner aux documents officiels mais plutôt nous intéresser aux tâches qui sont réellement proposées aux élèves. Comme nous l'avons défini dans le chapitre 2, l'ETM idoine est une adaptation et un aménagement de l'ETM de référence pour permettre une mise en place effective (ou se voulant comme telle) dans les classes. Du fait du temps restreint de la thèse, nous avons préféré, pour cette étude, nous focaliser sur l'étude des manuels, plutôt que d'aller effectivement dans des classes. Il s'agit donc en fait d'une étude d'ETM idoines potentiels (Nechache, 2015) proposés par les manuels scolaires français de terminale S. Un ETM idoine potentiel est un ETM proposé par des ressources (pour les enseignants). Il ne sera peut-être jamais effectif dans une classe tel quel, mais les manuels proposent des possibles.

Pour donner tout de même du crédit à nos analyses de manuels, pour parler de ce qui pourrait se faire en classe, nous avons complété cette étude des ETM idoines potentiels par des entretiens individuels que nous avons fait auprès de quatre enseignants, pour avoir des éléments sur l'ETM idoine qu'ils choisissent effectivement dans leur classe et apprécier leur relation avec ceux proposés par les manuels.

L'étude de l'ETM de référence a laissé entendre que les trois lois doivent être étudiées comme trois études de cas. La notion de fonction de densité ne semble pas ressortir et semble être imposée dans les trois cas. Le cas particulier du théorème de Moivre Laplace essaie de justifier la présence de cette fonction, mais il est exclusivement relié à la loi normale centrée réduite et ne permet pas de faire émerger le concept général de fonction de densité.

Au vu de l'analyse des documents officiels, la recherche pourrait s'arrêter là : la notion de fonction de densité est imposée dans le cas des trois lois au programme. Cependant, contrairement à ce constat qui se dégage du programme et du document *Ressources*, il n'en est pas de même dans les manuels. Les manuels proposent un début de cours général sur les lois à densité pour seulement ensuite se focaliser sur les trois lois au programme. De plus dans beaucoup de cas, une *activité*[21] d'introduction en début de chapitre est proposée pour introduire la notion de fonction de densité.

Ceci nous permet donc de continuer nos réflexions sur notre deuxième question de recherche et de garder cet objectif de comprendre comment est introduit et tra-

21. Nous écrivons *activité* en italique pour ne pas confondre avec l'activité de l'élève.

vaillé ce concept de fonction de densité.

Les questions qui guident notre travail, dans ce chapitre, peuvent être regroupées en trois groupes de questions.

Le premier groupe de questions porte sur l'introduction de la notion de fonction de densité dans les manuels :

- Comment les manuels proposent-ils d'introduire la notion de fonction de densité ?
- Quelles sont les activités attendues des élèves dans ces tâches d'introduction ?
- Permettent-elles de donner du sens au concept de fonction de densité ?
- Retrouve-t-on des approches semblables à l'approche historique ?

Le deuxième groupe de questions porte sur les articulations qui sont faites entre les deux domaines :

- Quelles sont les tâches et les activités attendues des élèves qui exploitent les liens entre probabilités à densité et calcul intégral, et en particulier la notion de fonction de densité ?
- Retrouvons-nous les mêmes que dans l'étude précédente ?
- Y en-a-t-il de nouvelles ?
- Peut-on repérer aussi des ruptures entre les deux sous-domaines ?

Et enfin une question sur les choix des enseignants par rapport à ceux faits dans les manuels :

- Ces tâches se retrouvent-elles effectivement dans les classes ?

Chaque groupe de question fera l'objet d'une section du chapitre. Nous allons, dans la section 5.1, détailler notre méthodologie d'analyse des manuels et justifier nos choix. Dans la section 5.2, nous nous focaliserons sur l'introduction de la notion de fonction de densité dans les manuels. Nous essaierons de décrire les dynamiques en jeu pour introduire le nouvel ETM des probabilités à densité. Dans la section 5.3, nous mettrons en avant des connexions possibles entre les deux sous-domaines. Enfin, dans la section 5.4, nous présenterons des choix d'enseignants, déclarés lors d'entretiens, afin de valider les éléments mis en avant dans l'analyse de manuels.

5.1 Méthodologie de l'étude des manuels

5.1.1 Justification du choix de l'étude des manuels

Nous limiterons notre étude des ETM idoines potentiels à ceux proposés par les manuels de terminale S et non par d'autres ressources pour l'enseignant.

La première raison de ce choix est le fait que nous considérons que les manuels sont très fréquemment utilisés par les enseignants pour préparer leur enseignement (Gueudet & Trouche, 2010). Nous ne nions pas le fait que les enseignants utilisent aussi d'autres ressources (*ibid.*), comme des documents accessibles sur internet par exemple, cependant dans un souci de facilité d'accessibilité et de délimitation de notre étude, nous préférons nous limiter aux manuels. Nous faisons aussi l'hypothèse que, les probabilités à densité étant une partie nouvelle du programme (dans ces proportions) et un domaine dans lequel les enseignants sont moins formés (Régnier,

2002), les enseignants s'inspirent d'autant plus des ressources que dans d'autres parties du programme, dans un but de pallier leur manque de recul.

La deuxième raison de ce choix est que l'étude des manuels permet d'avoir accès rapidement à un panel de diversité que nous n'aurions peut-être pas pu avoir sur un petit échantillon de classes observées.

Pour justement bénéficier d'un large éventail de possibles, nous allons analyser l'intégralité des manuels de mathématiques de terminale S (conformes au programme actuel) disponibles sur le marché français. Ils sont au nombre de huit. Les manuels considérés sont les suivants (les références complètes des manuels sont disponibles dans la rubrique dédiée de la bibliographie) : *Déclic*, *Hyperbole*, *Indice*, *Math'x*, *Odyssée*, *Repères*, *Symbole* et *Transmath*. Les huit manuels ont été édités en 2012, l'année de mise en œuvre de la réforme (en terminale S).

Les élèves d'une classe disposent d'un manuel, choisi par l'établissement. Les enseignants, eux, travaillent la plupart du temps avec plusieurs manuels et s'en inspirent pour préparer leur enseignement (Gueudet & Trouche, 2010).

L'étude des manuels peut donc nous donner une ou des idée.s plus précise.s des pratiques dans les classes, notamment avec des tâches qui peuvent être données aux élèves. Nous avons bien conscience que l'ensemble d'un manuel ne constitue pas l'ensemble de ce qui a été fait dans une classe, ne serait-ce qu'au niveau de la quantité abondante des tâches proposées.

L'étude des manuels est tout de même pour nous un moyen de voir les possibles proposés et d'avoir un point de vue sur ce qui peut se faire en classe. Cela nous permet de compléter notre étude précédente qui s'arrêtait aux ETM de référence. Ici nous abordons les ETM idoines potentiels, qui peuvent éventuellement être choisis pour la classe. Pour avoir tout de même un point de vue sur des choix réels d'enseignants, nous compléterons notre analyse par des entretiens d'enseignants, dont nous détaillerons la méthodologie seulement dans la section 5.4.1.

5.1.2 Présentation générale des manuels

Dans notre étude, nous nous concentrerons sur les chapitres concernés par le calcul intégral ou les probabilités à densité. Dans chaque manuel, un chapitre est dédié au calcul intégral. Pour les probabilités à densité, cela peut varier d'un à deux chapitres, suivant si les manuels ont fait le choix de mettre à part la loi normale ou non. Parfois une partie des probabilités à densité est rattachée à une autre partie du programme.

Nous présentons les chapitres concernés par notre étude dans les différents manuels en figure 5.1.

Manuel	Chapitre sur le calcul intégral	Chapitre sur les probabilités à densité
Déclic	Chapitre 6 Calcul intégral (pp. 182-221)	Chapitre 11 Lois de probabilité continues (pp. 362-401)
Hyperbole	Chapitre 8 Calcul intégral (pp. 197-226)	Chapitre 14 Notion de loi à densité (pp. 375-396) Chapitre 15 Lois normales (pp. 397-418)
Indice	Chapitre 6 Calcul intégral (pp. 162-195)	Chapitre 11 Lois de probabilité à densité (pp. 320-353)
Math'x	Intégration (pp. 229-264	Lois à densité (pp. 397-432)
Odyssée	Chapitre 7 Intégration (pp. 236-277)	Chapitre 11 Lois à densité (pp. 390-425)
Repères	Chapitre 4 Intégration (pp. 166-227)	Chapitre 7 Probabilités conditionnelles et lois continues (pp. 326-369) Chapitre 8 Loi normale et estimation (pp. 370-419)
Symbole	Chapitre 7 Intégration et primitives d'une fonction (pp. 200-235)	Chapitre 13 Notion de loi à densité (pp. 384-415) Chapitre 14 Lois normales, intervalle de fluctuation, estimation (pp. 416-447)
Transmath	Chapitre 7 Intégration et primitives (pp. 190-229)	Chapitre 13 Lois de probabilité à densité (pp. 376-403) Chapitre 14 Lois normales (pp. 404-433)

TABLE 5.1 – Chapitres des différents manuels concernés par le calcul intégral et les lois à densité

5.1.3 Outils méthodologiques pour l'analyse *a priori* des *activités* d'introduction de la notion de fonction de densité

Un de nos objectifs est de comprendre comment est introduite la notion de fonction de densité dans les manuels de terminale S. La fonction de densité est la notion qui permet la connexion entre les probabilités à densité et le calcul intégral. La compréhension de cette connexion se fait en partie au moment de l'introduction de cette notion.

Nous faisons le choix dans cette étude de ne regarder que l'introduction de la fonction de densité en début de chapitre. Ce qui nous intéresse, c'est l'introduction de la notion de densité de façon générale. Il n'est pas question pour nous de nous atteler à l'introduction du théorème de Moivre Laplace, qui est seulement dédiée à la loi normale. Pour avoir des éléments sur le théorème de Moivre Laplace dans la classe de terminale S, il est possible de se reporter à Derouet et Parzysz (2016b).

L'introduction de la notion de densité est travaillée dans la partie *Activités* des manuels, partie précédant le cours. Nous analyserons donc les *activités* introductives des manuels ayant pour objectif cette introduction. Nous garderons la terminologie utilisée dans les manuels d'*activité* (qui sous-entend *activité* introductive), mais il ne faut pas confondre l'*activité*, c'est-à-dire l'ensemble des tâches proposées par un manuel pour introduire une notion, et l'activité attendue des élèves lors de la résolution d'une tâche. Pour le premier cas, nous mettrons le mot en italique.

Parmi ces *activités*, que nous présenterons toutes, nous nous attarderons plus longuement sur les introductions faisant intervenir l'histogramme ; nous expliquerons pourquoi les autres démarches nous paraissent moins adaptées, en temps voulu. De plus, le passage par l'histogramme n'est pas sans rappeler l'approche que nous avons trouvée dans notre étude historico-épistémologique dans le cours de Faye. Nous pourrons tout de même nous demander si nous retrouvons vraiment la même approche, à savoir de partir de l'histogramme et de déterminer la fonction de densité par un ajustement.

Dans le cas des introductions faisant intervenir l'histogramme, nous ferons une analyse critique des tâches et des activités attendues, assez poussées. Nous ferons en parallèle une analyse de la tâche prescrite et une analyse des activités (potentielles) attendues de l'élève, en s'arrêtant question par question.

Pour l'analyse de la tâche prescrite, nous nous attarderons sur le contexte, la compréhension de l'énoncé, la cohérence des données de l'énoncé... Pour l'analyse des activités attendues de l'élève, nous présenterons, tout d'abord, la ou les démarche.s mathématique.s attendue.s pour résoudre chaque tâche. Cela nous permettra parfois de faire un retour sur la tâche prescrite. Pour cette raison, nous parlons d'analyse critique de la tâche. Ensuite, nous ferons une analyse *a priori* des activités des élèves suivant plusieurs points. Nous avons répertorié ci-dessous certaines des questions que nous nous sommes posées pour faire notre analyse *a priori* et dans chaque cas, nous précisons les buts associés. Nous avons en grande partie repris les axes 3 et 4 des analyses de tâches de Robert (1998), en y ajoutant des éléments du cadre des ETM.

Nous présentons, ci-dessous, les axes que nous allons prendre en compte dans l'analyse de chaque sous-tâche de l'*activité*.

1. *(Sous)-domaines mathématiques en jeu* : Quels sont les domaines mathématiques (statistique descriptive (SD), probabilités à densité (PaD) ou calcul intégral (CI)) en jeu ?
2. *Connaissances à mettre en fonctionnement* :
   - *Ancien/Nouveau* : Les connaissances mises en jeu sont-elles anciennes (avant la terminale), récentes (de l'année de la terminale) ou nouvelles (font l'objet de l'enseignement du moment) ?
   - *Niveau de mise en fonctionnement des connaissances* : Quel est le niveau de mise en fonctionnement attendu (technique, mobilisable ou disponible) ? Le référentiel théorique à mobiliser est-il à introduire par l'élève ou est-il donné dans l'énoncé ?
3. *Adaptations* : Y a-t-il des adaptations à faire : reconnaissance de modalités d'application, introduction d'intermédiaires, introduction d'étapes, utilisation des questions précédentes, existence de choix de méthodes ?
4. *Registres* : Quels registres de représentations sémiotiques sont en jeu ?
5. *Dimensions de l'ETM en jeu* : Quelles sont les dimensions (sémiotique, instrumentale ou discursive) de l'ETM en jeu ? Quelles sont les dynamiques entre les ETM des différents domaines ? Par le biais de quelles dimensions se font les changements de domaine ?

Le niveau de mise en fonctionnement des connaissances est de l'ordre du technique, du mobilisable ou du disponible. Le niveau technique correspond à des mises en fonctionnement indiquées et isolées, mettant en jeu des applications directes du cours (théorème, définition, formule...). Le niveau mobilisable correspond à des mises en fonctionnement plus larges : encore indiquées, mais dépassant l'application simple d'une propriété à la fois. Enfin, le niveau disponible des connaissances à mettre en fonctionnement correspond au fait de savoir résoudre ce qui est proposé sans indication (Robert, 1998).

Nous allons aussi apprécier, à travers les tâches proposées, quelles sont les propriétés sur la fonction de densité qui émergent.

Les tâches sont analysées telles qu'elles sont données, sans se demander ce qu'un enseignant pourrait ajouter, adapter...

Cette étude des *activités* d'introduction, nous amènera à creuser la question de l'enseignement de la notion d'histogramme en classe de seconde, la dernière année durant laquelle l'histogramme est étudié en tant qu'objet. Nous détaillerons en tant voulu notre méthodologie pour cette étude.

5.1.4 Méthodologie pour l'observation des tâches exploitant les connexions entre les deux sous-domaines

Après avoir étudié l'introduction de la notion de fonction de densité, nous allons passer à l'identification des tâches exploitant les connexions entre les deux sous-domaines. Le but n'est pas d'être exhaustif.

Pour cette étude, nous nous limiterons à l'analyse des exercices corrigés, pour plusieurs raisons. Tout d'abord, les exercices étant corrigés, nous n'avons pas à inférer sur la méthode de résolution de la tâche et donc sur les activités des élèves.

De plus, ces exercices servent de base à de nombreux exercices des manuels, cela nous permet donc de ne pas analyser l'ensemble des exercices des manuels.

Nous essaierons, à travers ces tâches, de repérer des connexions existantes ou possibles entre probabilités à densité et calcul intégral.

Nous ne nous attarderons pas sur la méthode de Monte Carlo qui est légèrement en marge dans notre questionnement mais serait, nous en sommes sûre, une piste de recherche intéressante.

5.2 Analyse des *activités* d'introduction de la notion de fonction de densité dans les manuels

5.2.1 Avant de commencer

Dans tous les manuels, le chapitre sur le calcul intégral est toujours proposé en amont de celui sur les probabilités à densité. Cela ne veut, à première vue, rien dire car les manuels font le choix d'ordonner leurs chapitres par domaines du programme et non pas comme une progression envisageable. Cependant, dans tous les manuels qui explicitent les prérequis pour aborder le chapitre des lois à densité, il est fait référence au calcul intégral.

Nous pouvons le voir par exemple dans les manuels *Hyperbole* et *Transmath* (cf. figures 5.1 et 5.2).

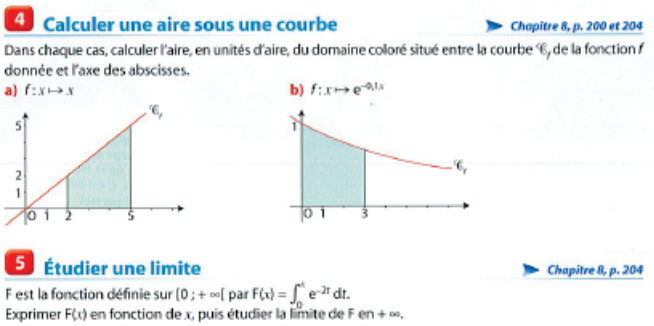


Figure 5.1 – Extrait de la partie *Vérifier les acquis* du chapitre *Notion de loi à densité* du manuel *Hyperbole* (p. 376)

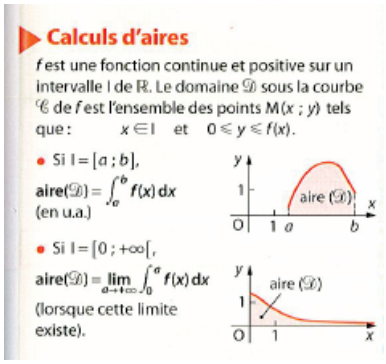


FIGURE 5.2 – Extrait de la partie *Rappels* du chapitre *Lois de probabilités à densité* du manuel *Transmath* (p. 377)

Les sous-domaines *Calcul intégral* et *Probabilités à densité* sont donc deux parties du programme bien séparées, dont le premier est clairement un prérequis pour le second.

Passons maintenant aux chapitres relatifs aux probabilités à densité. Comme nous l'avons laissé entendre dans l'introduction de ce chapitre, dans tous les manuels de terminale S, le début du cours sur les lois à densité propose une partie de généralités, non limitée aux trois lois au programme. Cette partie peut être plus ou moins longue suivant les manuels.

En annexe E.1 se trouve la partie générale des débuts de cours de chaque manuel. Ces généralités peuvent se limiter à la simple définition de la fonction de densité et parfois s'étendre aux propriétés sur les probabilités, ainsi qu'à la définition de l'espérance.

Cependant, des choix de formulation et/ou de restriction apparaissent dans les définitions choisies. Certains manuels définissent la fonction de densité seulement dans le cas borné et ensuite, seulement en temps voulu, précisent que la définition peut être étendue au cas non borné. C'est le cas du manuel *Déclic* (cf. figure 5.3).

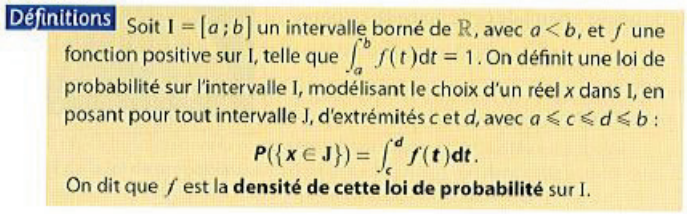


FIGURE 5.3 – Extrait du cours du manuel *Déclic* (p. 366)

Ce choix est sans doute fait dans un souci de simplicité de formulation. Certains autres manuels, dans un souci de rigueur sûrement, précisent les différents cas de figure, comme le manuel *Symbole* qui utilise la formulation du programme ou encore les manuels *Hyperbole* (cf. figure 5.4) et *Repères* qui utilisent l'intégrale mais parlent bien du cas borné et du cas borné à gauche (mais pas du cas non borné).

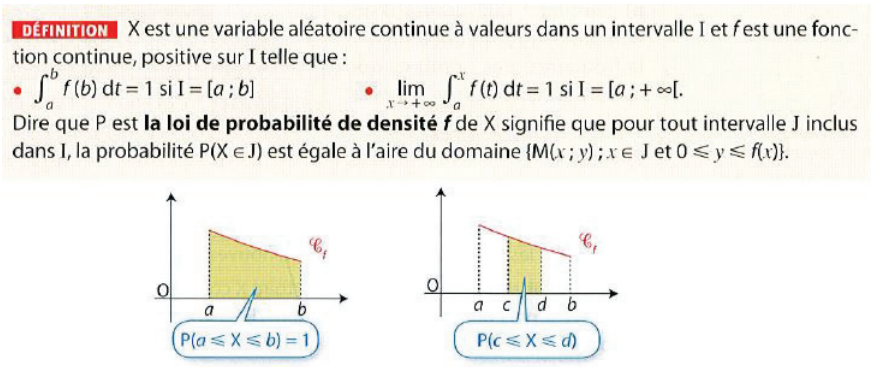


Figure 5.4 – Extrait du cours du manuel *Hyperbole* (p. 378)

Le fait de parler en termes d'aire plutôt que d'intégrale permet d'alléger les formulations. C'est souvent le choix qui est fait.

En revanche, dans aucun cours des manuels, il n'est fait référence au fait que ces intégrales « impropres » existent bien, et sont finies, et qu'il est effectivement possible que l'aire d'un domaine infini soit égale à 1. On peut alors se demander s'il y est fait référence à un autre endroit. Dans le chapitre sur les lois à densité, des calculs de limites d'intégrales sont faits, pour l'espérance notamment, sans être questionné. On peut mentionner le manuel *Symbole* qui écrit tout de même dans le cours que les définitions de l'espérance ne valent que si les limites existent. Dans le chapitre dédié au calcul intégral, on peut repérer comme dans le baccalauréat, quelques exercices sur la limite d'une suite d'intégrales de type $I_n = \int_a^n f$ (qui correspond à l'aire d'un domaine infini) mais vraiment peu (de 0 à 2 par manuel) et ils ne sont pas toujours interprétés en termes d'aires.

Les propriétés de la fonction de densité qui reviennent le plus souvent dans les manuels (6 cas sur 8), à quelques formulations près, sont les suivantes :

f est une fonction de densité associée à une variable aléatoire à valeurs sur I si :

1. f est définie et continue sur I,
2. f est positive sur I,
3. l'aire sous la courbe de f sur l'intervalle I vaut 1.

Les manuels *Symbole* et *Déclic* sont les seuls à ne pas mettre la continuité de la fonction de densité dans les propriétés. Cependant, pour le manuel *Déclic* (cf. figure 5.3), il est mentionné que l'intégrale sur $[a; b]$ est égale à 1, or une intégrale n'a été définie en calcul intégral seulement dans le cas d'une fonction continue, donc cette propriété est sous-entendue (pour les auteurs). Le manuel *Odyssée*, lui, demande la continuité sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points. Ces deux manuels proposent des exemples de fonctions de densité avec un point de non continuité.

Ici nous pouvons donc entrevoir une prise de position des manuels quant au degré de généralité qu'ils souhaitent proposer dans ce chapitre. Bien que les trois lois soient les lois majoritairement traitées, le travail n'est tout de même pas limité à

cela. Elles ne sont pas vues dans les manuels comme des études de cas mais comme des exemples particuliers de lois, ce qui est différent. On part ici du général pour nous intéresser ensuite à trois cas particuliers.

Le manuel *Odyssée* peut paraître quelque peu hors programme avec tout un paragraphe de cours dédié à la fonction de répartition. On voit dans ce manuel une volonté de présenter des résultats généraux sur les lois à densité pour ensuite les utiliser sur les trois exemples. La partie de « généralités » sur les probabilités à densité est composé de cinq pages contre entre une demi-page et une page dans les autres manuels.

5.2.2 Introduction de la notion de fonction de densité : les différentes approches rencontrées dans les manuels

Mis à part dans le manuel *Odyssée*, qui ne propose pas d'*activité* pour introduire la notion de densité et définit directement la fonction de densité dans le cours, les sept autres manuels offrent au moins une *activité* introductive. Dans ces *activité*s d'introduction, cinq (voire six) manuels présentent des courbes de densité qui ne sont pas l'une des trois au programme.

L'introduction de la fonction de densité n'est pas traitée de la même manière dans tous. Nous distinguons quatre approches différentes :

- Un manuel ne propose aucune *activité* d'introduction et impose la définition directement sans poser le problème.
- Un autre manuel utilise un exemple avec des cibles de tir de fléchettes.
- Un troisième manuel propose d'utiliser le polygone des fréquences cumulées, pour introduire dans un premier temps la fonction de répartition, et ensuite l'histogramme de fréquences, en le considérant comme la « dérivée » du polygone des fréquences cumulées, pour introduire la fonction de densité.
- Les cinq autres manuels ont une approche commune, basée sur l'histogramme de fréquences avec des classes de même amplitude approché par une courbe.

On peut retrouver ces différentes approches en fonction des manuels dans le tableau 5.2.

Type d'approche	**Sans *activité* d'introduction**	**Jeu de fléchettes**	**Via le polygone des fréquences cumulées**	**Via l'histogramme**
Manuels	*Odyssée*	*Déclic*	*Symbole*	*Hyperbole* *Indice* *Math'x* *Repères* *Transmath*
Nombre	1	1	1	5

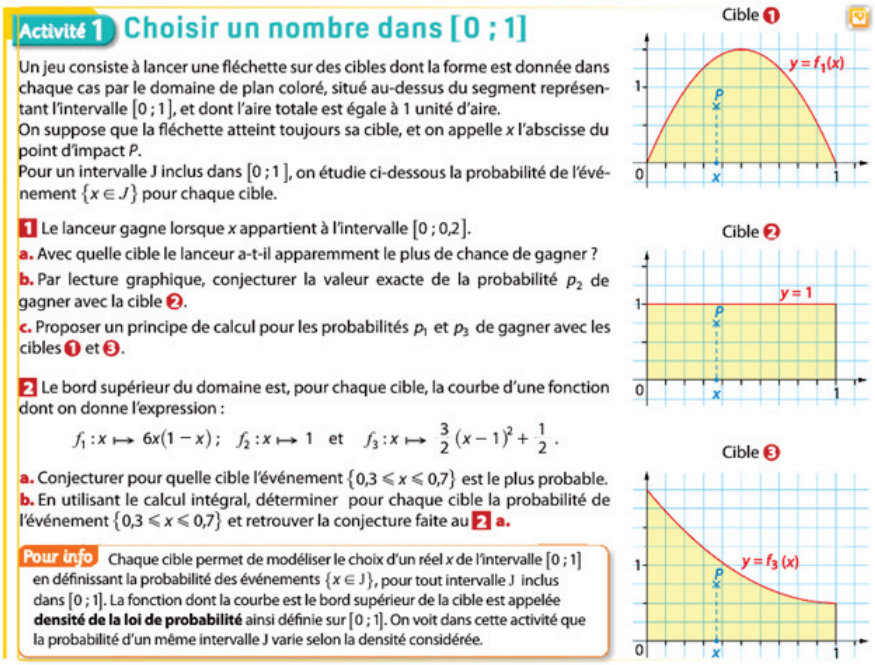
TABLE 5.2 – Les différentes approches de la notion de fonction de densité dans les manuels de terminale S

a) La fonction de densité : un objet imposé

Le manuel *Odyssée* ne propose pas d'*activité* introductive pour la notion de fonction de densité et passe directement à la partie *Cours* (cf. Annexe E.1.5). Après avoir défini une variable aléatoire continue et rappelé comment définir une loi de probabilité discrète, les auteurs écrivent directement : « *Dans le cas d'une variable aléatoire continue, on a recours à une fonction définie sur* $\mathbb{R}$ *appelée* ***densité*** » (p. 395). Cette fonction est imposée. Ceci est assumé car aucune tentative d'explication n'est donnée dans le manuel.

b) Le jeu de fléchettes

L'*activité* 1 page 364 (cf. figure 5.5), proposée par le manuel *Déclic*, utilise un exemple de jeu de fléchettes.



Activité 1 Choisir un nombre dans [0 ; 1]

Un jeu consiste à lancer une fléchette sur des cibles dont la forme est donnée dans chaque cas par le domaine de plan coloré, situé au-dessus du segment représentant l'intervalle [0 ; 1], et dont l'aire totale est égale à 1 unité d'aire.
On suppose que la fléchette atteint toujours sa cible, et on appelle x l'abscisse du point d'impact P.
Pour un intervalle J inclus dans [0 ; 1], on étudie ci-dessous la probabilité de l'événement $\{x \in J\}$ pour chaque cible.

1 Le lanceur gagne lorsque x appartient à l'intervalle [0 ; 0,2].
a. Avec quelle cible le lanceur a-t-il apparemment le plus de chance de gagner ?
b. Par lecture graphique, conjecturer la valeur exacte de la probabilité p_2 de gagner avec la cible ❷.
c. Proposer un principe de calcul pour les probabilités p_1 et p_3 de gagner avec les cibles ❶ et ❸.

2 Le bord supérieur du domaine est, pour chaque cible, la courbe d'une fonction dont on donne l'expression :

$f_1 : x \mapsto 6x(1-x)$; $f_2 : x \mapsto 1$ et $f_3 : x \mapsto \frac{3}{2}(x-1)^2 + \frac{1}{2}$.

a. Conjecturer pour quelle cible l'événement $\{0{,}3 \leqslant x \leqslant 0{,}7\}$ est le plus probable.
b. En utilisant le calcul intégral, déterminer pour chaque cible la probabilité de l'événement $\{0{,}3 \leqslant x \leqslant 0{,}7\}$ et retrouver la conjecture faite au **2 a.**

Pour info Chaque cible permet de modéliser le choix d'un réel x de l'intervalle [0 ; 1] en définissant la probabilité des événements $\{x \in J\}$, pour tout intervalle J inclus dans [0 ; 1]. La fonction dont la courbe est le bord supérieur de la cible est appelée **densité de la loi de probabilité** ainsi définie sur [0 ; 1]. On voit dans cette activité que la probabilité d'un même intervalle J varie selon la densité considérée.

Cible ❶ $y = f_1(x)$ — Cible ❷ $y = 1$ — Cible ❸ $y = f_3(x)$

Figure 5.5 – *Activité* 1 p. 364 (*Déclic*)

On considère une expérience aléatoire qui est le lancer d'une fléchette sur une cible particulière, constituée d'une partie du plan euclidien d'aire 1 comprise entre une courbe et l'axe des abscisses. On suppose que la fléchette atteint toujours la cible et, bien que cela ne soit pas mentionné, que toutes les positions sur cette cible sont équiprobables (c'est d'ailleurs cela qui va servir ensuite). On introduit alors implicitement la variable aléatoire qui correspond à l'abscisse, comprise entre 0 et 1, de la position de la fléchette lors d'un tir sur la cible. Trois cibles différentes sont successivement envisagées.

On demande ensuite la probabilité d'événements de la forme $\{x \in J\}$, avec J inclus dans $[0; 1]$. Implicitement, comme nous l'avons précisé au-dessus, la position de la fléchette suit une loi uniforme sur l'ensemble des points de la cible, ce qui implique

que la loi de son abscisse va dépendre de la forme de la cible. Cette *activité* conduit le manuel à introduire la notion de fonction de densité comme étant la fonction représentée par la courbe rouge et à conclure que « *la fonction dont la courbe est le bord supérieur de la cible est appelée densité de probabilité de la loi ainsi définie sur* $[0;1]$ » (p. 364). Nous sommes bien d'accord que, pour cette variable aléatoire, le bord de la cible « représente » bien la fonction de densité associée. Cependant, cette démarche n'est pas satisfaisante pour une variable aléatoire autre que celle-ci. Et donc, comment dans d'autres cas, définir la fonction de densité ? Cette idée de bord de cible, très contextualisée, ne permet pas de s'approprier cette notion de fonction de densité car toute variable aléatoire ne peut pas se voir comme l'abscisse d'une fléchette lors d'un lancer de fléchette sur une cible. On peut aussi se demander s'il est nécessaire que l'aire de la surface sous la courbe soit égale à 1.

c) Une introduction via le polygone des fréquences cumulées

Le manuel *Symbole*, dans l'*activité* 2 pp. 386-387 (cf. Annexe E.2.1), a fait le choix de partir de la statistique descriptive, mais d'introduire un polygone de fréquences cumulées croissantes (cf. figure 5.6).

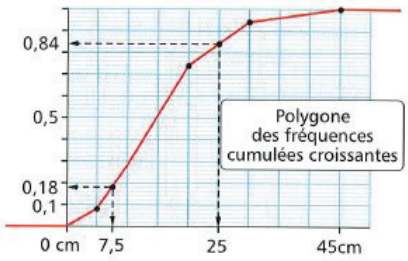


FIGURE 5.6 – Polygone des fréquences cumulées croissantes de l'*Activité* 2 p. 386-387 (*Symbole*)

À partir des fréquences cumulées croissantes F_k, il est défini la densité d_k, pour chaque classe k, par le quotient de la fréquence (cumulée croissante) de la classe par son amplitude. Il s'agit de la « dérivée » de la fonction représentée par la courbe des fréquences cumulées croissantes. Comme le précise l'objectif de cette *activité*, il s'agit ici de « *découvrir la fonction de densité et la fonction de répartition dans le cas d'une variable aléatoire non discrète* ». Le polygone des fréquences cumulées croissantes et donc la fonction de répartition sont ici les moyens d'introduire l'histogramme et par la même occasion la fonction de densité.

Cependant, la fonction de répartition n'étant pas un attendu du programme, il ne nous semble pas judicieux de passer par le polygone de fréquences pour amener l'histogramme. De plus, la notion de polygone des fréquences cumulées ajoute une difficulté pour les élèves, qui n'est pas nécessaire pour l'introduction de la fonction de densité.

Cette *activité*, en revanche, a le mérite de définir la densité d'une classe et donc d'essayer de donner du sens au terme « densité » (il en est de même dans les *activités* suivantes du manuel centrées sur la loi uniforme, puis la loi exponentielle), ce qui

n'est pas le cas dans les autres manuels.

Les trois introductions que nous venons de décrire succinctement présentent des inconvénients et ne nous semblent pas les plus adaptées pour l'introduction de la notion de densité. C'est la raison pour laquelle nous allons maintenant nous attarder plus longuement sur les cinq autres manuels, ayant tous proposés une ou deux *activités* d'introduction mettant en jeu un histogramme (sans le polygone des fréquences cumulées).

5.2.3 Analyse *a priori* des *activités* introductives utilisant la notion d'histogramme

Dans cette section, nous allons présenter nos analyses *a priori* de quatre des cinq activités d'introduction qui utilisent la notion d'histogramme pour introduire la notion de fonction de densité.

a) D'un échantillon à une population

La situation proposée

L'*activité* 2 p. 377 du manuel *Hyperbole* se trouve en figure 5.7.

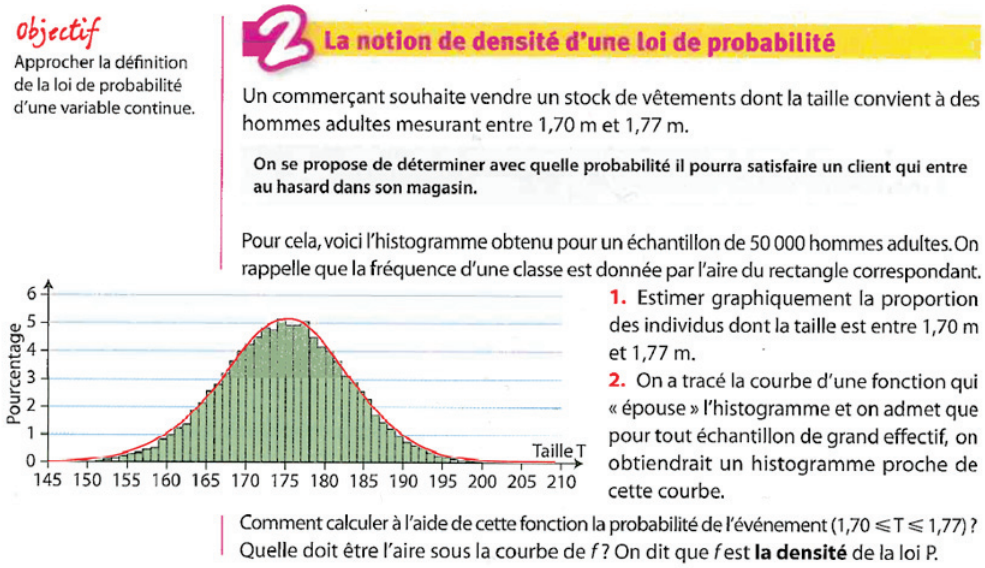


Figure 5.7 – *activité* 2 p. 377 (*Hyperbole*)

Elle a pour titre : « *La notion de densité d'une loi de probabilité* ». L'objectif de l'*activité*, écrit explicitement, est d'approcher la définition de la loi de probabilité d'une variable (certainement variable aléatoire) continue.

Le problème est contextualisé : il s'agit de la vente d'un stock de vêtements dont la taille convient à des hommes adultes mesurant entre 1,70 m et 1,77 m. L'objectif du problème (dans le cadre grisé) est de déterminer avec quelle probabilité

le commerçant pourra satisfaire un client qui entre au hasard dans son magasin. L'expérience aléatoire en question est l'entrée d'un client au hasard dans le magasin du commerçant.

On peut cependant se demander à quoi se rapporte l'expression « *au hasard* » : est-ce le choix du magasin par le client qui revient au hasard ou le choix du client qui entre dans le magasin qui est fait au hasard ? De plus, rien n'empêche le fait qu'une femme entre dans le magasin. Faut-il, dans ce cas, faire l'hypothèse qu'elle rentre pour acheter un vêtement à un homme et que l'on doit associer cette femme à l'homme pour lequel elle fait le potentiel achat ? Le fait de modéliser par une loi uniforme sur les hommes pour la loi du client entrant au hasard n'est pas si clair que cela, cependant c'est ce qui semble attendu ici.

Ensuite, à quoi correspond « satisfaire un client » ? Est-ce que cela revient vraiment au seul critère de la correspondance de la taille ? L'esthétique ne compte-t-elle pas ? De plus, si le client entre par hasard dans le magasin, il peut ne pas vouloir de vêtements.

Il pourrait y avoir beaucoup de commentaires à faire sur cette situation qui est déjà modélisée sans en expliciter les hypothèses faites. Le contexte choisi est très artificiel.

Après avoir posé le problème, l'énoncé propose un histogramme obtenu pour un échantillon de 50 000 hommes adultes. Il n'est pas précisé dans le texte qu'il s'agit de l'histogramme des tailles, mais nous pouvons trouver l'information sur l'axe des abscisses du graphique. L'histogramme fourni est sensiblement identique à celle du document *Ressources* (MENJVA & DGESCO, 2012) et du document d'accompagnement de 2002 (MJENR, 2002), qui avait pour objectif d'introduire la loi normale.

La présence du mot pourcentage sur l'axe des ordonnées peut générer des confusions, il faut bien comprendre ici qu'il ne s'agit que d'une écriture de nombre : $1\% = 0,01$. Cependant, « *pourcentage* » semble être le nom de l'axe et donc laisse penser que cet axe représente la fréquence, ce qui n'est pas rigoureusement le cas. Dans un histogramme, l'axe des ordonnées correspond à la densité de fréquence (Roditi, 2009). Ici l'amplitude des classes étant de 1, cela ne pose pas de problème d'assimiler fréquence et densité (vu qu'elles sont égales), cependant cela devient problématique dans les autres cas. Nous reviendrons plus en détail sur l'histogramme dans la section 5.2.5.

Il est, de toute façon, rappelé dans l'énoncé que la fréquence d'une classe est donnée par l'aire du rectangle de l'histogramme correspondant, ce qui permet d'enlever les doutes possibles. Cependant, pour les élèves, cela laisse quand même transparaître des ambiguïtés. L'histogramme est une notion faisant partie des connaissances anciennes, au programme des classes de la cinquième à la seconde. Cependant, le rappel montre que les auteurs des manuels ne semblent pas sûrs de pouvoir compter sur les connaissances des élèves sur cette notion.

À moins que les données n'aient été tirées d'une enquête INSEE (ce qui ne semble pas être le cas car aucune source n'est indiquée), il semblerait que l'histogramme ait été construit à partir de données simulées et non de données réelles. Il s'agit sans doute d'un problème « fabriqué » : on construit nos données (que l'on fait passer

pour des données « réelles ») pour pouvoir travailler dessus. Cela entraîne donc que le travail s'effectue avec un histogramme « idéal » : aucun trou n'est présent par exemple, ce qui est rare pour de vraies données. La modélisation est donc déjà faite implicitement. Il est probable qu'ici les données aient été calculées à partir d'une loi normale, cependant à aucun moment cela n'est indiqué.

Question 1

> 1. Estimer graphiquement la proportion des individus dont la taille est entre 1,70 m et 1,77 m.

Les valeurs de tailles considérées ici reprennent celles qui nous intéressent pour répondre à l'objectif du problème.

Démarche mathématique attendue

Estimer la proportion des individus dont la taille est entre 1,70 m et 1,77 m équivaut à estimer la fréquence de ces individus. Donc d'après le rappel de l'énoncé, il s'agit d'estimer l'aire des sept rectangles de base $[i; i+1]$, avec $i \in [\![170; 176]\!]$. Vu le manque de précision dans les graduations, les résultats peuvent être légèrement différents d'un élève à l'autre. Il faut s'attendre à une marge d'erreurs d'environ 0,01. Des produits en croix peuvent être nécessaires pour déterminer les hauteurs des rectangles. De plus, il faut bien prendre en compte le fait que les graduations des ordonnées sont en pourcentage donc 4,25 correspond en fait à 0,0425. D'où :

$$p \simeq 0,0425 + 0,045 + 0,048 + 0,0475 + 0,051 + 0,049 + 0,049$$
$$p \simeq 0,332.$$

Analyse des activités attendues de l'élève

Cette question et sa résolution se situent dans le domaine de la statistique et plus particulièrement dans le sous-domaine de la statistique descriptive (SD). Le référentiel théorique nécessaire pour la résolution est la connaissance de l'histogramme et particulièrement le traitement à opérer au sein du registre histogramme. Cependant, cette connaissance ne semble pas supposée disponible, c'est sans doute la raison pour laquelle on peut repérer la présence d'aides : avec le mot « *graphiquement* » qui appuie sur le fait que c'est l'histogramme qui va permettre de répondre à la question et ensuite comme nous l'avons précisé plus haut, le rappel donné en amont dans l'énoncé sur le lien entre fréquence et histogramme. Au niveau des activités attendues des élèves, il y a des adaptations à faire :

1. Faire une correspondance entre les mots proportion et fréquence.
2. Appliquer le rappel de l'énoncé en prenant en considération qu'il faut sommer les différentes fréquences et donc les différentes aires en question.

Les aires des rectangles (que nous nommerons aire 1) sont à estimer à l'aide du graphique donc le résultat est approché, surtout en raison du fait que les graduations sont peu précises. De plus, il est important ici de bien considérer les valeurs de

l'ordonnée comme écriture décimale. C'est la dimension sémiotique qui est à l'œuvre dans cette question. Même si la dimension discursive est mobilisée, elle n'est que très peu à la charge de l'élève, vu que les indications sont données dans l'énoncé.

Question 2

Au début de la question 2, la courbe rouge visible sur l'histogramme est introduite. Il est précisé : « *On a tracé la courbe d'une fonction qui « épouse » l'histogramme et on admet que pour tout échantillon de grand effectif, on obtiendrait un histogramme proche de cette courbe* ». On trouve ici l'idée de fluctuation d'échantillonnage à travers l'histogramme, cependant cela n'est pas plus détaillé.

Le tracé de la courbe est pris en charge par le manuel. Ce choix n'est absolument pas questionné. Il est écrit « *la courbe d'**une** fonction* », l'article indéfini laisse donc penser que cette courbe n'est pas unique mais rien ne précise pourquoi celle-ci a été choisie. On pourrait déjà se demander ce que signifie « *épouse* ». De plus, le choix de la fonction ne semble dépendre que de cette condition. Il n'y a, à ce moment-là, aucune référence au fait que l'aire sous la courbe doit être obligatoirement égale à 1.

Les auteurs du manuel font le choix d'un modèle, qu'ils imposent aux élèves.

> 2(a). Comment calculer à l'aide de cette fonction la probabilité de l'évènement $(1,70 \leqslant T \leqslant 1,77)$?
> 2(b). Quelle doit être l'aire sous la courbe de f ?

Nous avons renommé les sous-questions, 2(a) et 2(b). Avant tout, nous pouvons remarquer que la variable aléatoire T n'est pas définie. Il doit s'agir de la variable aléatoire correspondant à la taille du client entrant dans le magasin. La fonction f non plus n'est pas définie. On parle de l'aire sous la courbe de f. Il s'agit, on s'en doute, de la fonction représentée par la courbe rouge.

Démarche mathématique attendue

L'introduction de la question 2 permet de voir la courbe rouge comme une approximation de l'histogramme mais aussi de n'importe quel histogramme associé à un autre échantillon. L'élève doit déterminer la méthode à utiliser en faisant une analogie entre l'histogramme et la courbe donnée (courbe de densité) et donc entre la fréquence et la probabilité. Il y a un passage de l'aire des rectangles de l'histogramme (que nous appellerons aire 1) à l'aire sous la courbe (que nous appellerons aire 2).

La réponse attendue est : l'aire sous la courbe entre 1,70 m et 1,77 m, ou elle peut être envisagée aussi dans le registre symbolique de l'intégrale c'est-à-dire sous la forme $\int_{1,70}^{1,77} f(x)\,\mathrm{d}x$, où f est la fonction représentée par la courbe rouge. Il n'est pas attendu ici de faire des calculs car nous ne disposons pas de l'expression de la fonction.

La sous-question 2 (b) questionne sur l'aire qui doit être sous la courbe de f. L'expression « *doit être* » peut être interprétée de plusieurs façons : comme une condition nécessaire (cependant, la fonction a été choisie), comme « est certainement » au vu de ce que l'on a trouvé avant, ou « est » mais on ne sait alors pas le degré de généralité de cette réponse.

Deux démarches sont alors envisageables. Il est possible de faire à nouveau une analogie entre fréquence et probabilité et donc aire sous la courbe. Ce qui implique que l'aire doit valoir 1. Ou, avec le même raisonnement qu'à la question précédente, l'aire sous la courbe de f (ce qui sous-entend sur son ensemble de définition) correspond à la probabilité que la variable aléatoire T appartiennent à tout l'ensemble de définition et donc nécessairement elle doit valoir 1.

Analyse des activités attendues de l'élève

Dans toute la question 2, il s'agit pour l'élève d'une première rencontre avec la fonction de densité et donc avec ce type de questions. Il doit donc lui-même trouver les analogies entre fréquence et probabilité.

La sous-question 2(a) se situe dans le sous-domaine des probabilités à densité (PaD), mais il est attendu un basculement du sous-domaine de la statistique descriptive (SD) à celui des probabilités à densité (PaD), pour arriver au sous-domaine du calcul intégral (CI), avec l'aire ou l'intégrale.

Il y a un fort appui sur la visualisation, notamment pour faire l'analogie fréquence/probabilité. Le référentiel théorique de la définition de l'intégrale d'une fonction continue et positive en tant qu'aire sous la courbe doit être disponible pour que l'élève donne la réponse dans le registre symbolique de l'intégrale.

Nous pouvons remarquer que l'activité s'arrête ici. Il n'y a aucun retour sur l'objectif du problème qui était de déterminer une probabilité. Nous pourrions même dire pour être précis que le but était d'estimer une probabilité. La problématique appartient à la statistique inférentielle, mais toutes les questions que cela entraîne, comme le choix du modèle, ne sont pas examinées ici.

Conclusion sur les circulations dans cette *activité*

Les activités attendues de l'élève évoluent dans les différents ETM (SD, PaD et CI), essentiellement via la dimension sémiotique, principalement grâce à la visualisation (histogramme, courbe). Il y a une analogie qui est faite entre SD et PaD. La dimension discursive est peu présente : soit le référentiel théorique pourrait être justifié mais n'est pas nécessairement convoqué (nous pensons notamment à l'énoncé vulgarisé de la loi des grands nombres [22]), soit le référentiel théorique nécessaire est rappelé (sur l'histogramme). Les circulations en jeu sont représentées dans la figure 5.8. En rouge, sont entourées les dimensions prises en charge par les élèves (en pointillé rouge, si la dimension peut ne pas être mobilisée) et, en vert, celles prises en

22. La formulation proposée (dans le cas fini) dans le document d'accompagnement du programme de 2001 de la classe de première S (MEN, Direction de l'Enseignement scolaire, 2001), « *en langage imagé* », est la suivante : « *Si on choisit n éléments d'un ensemble fini E selon une loi de probabilité P, indépendamment les uns des autres, alors la distribution des fréquences est proche de la loi de probabilité P lorsque n est grand* ».

charge par le manuel.

Le tableau 5.3 présente plus en détail par le biais de quels objets mathématiques se font les circulations, avec comme point de départ l'histogramme. Enfin, nous avons essayé à travers la figure 5.9 de faire apparaître le cheminement et les connexions entre les trois sous-domaines.

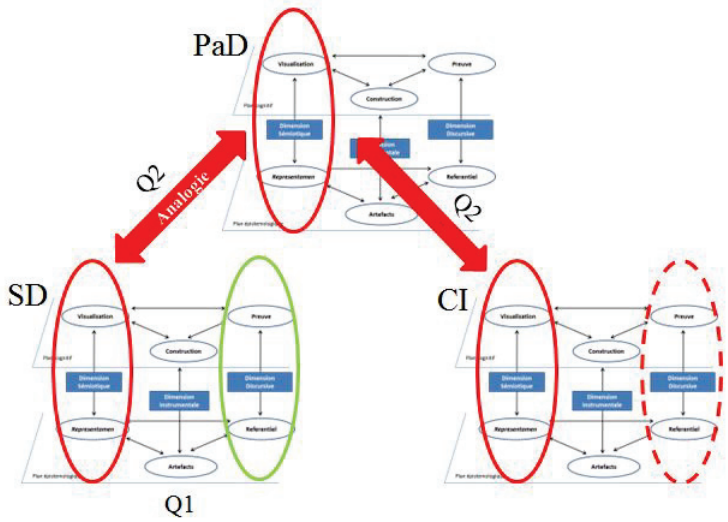


FIGURE 5.8 – Circulations entre les trois sous-domaines (*activité* 2 p. 377, *Hyperbole*)

Partie de la tâche	Sous-domaine mathématique	Objet mathématique	Intermédiaire
Enoncé introductif	SD	Histogramme (donné)	Aires des rectangles (aire 1): rappel dans l'énoncé
Question 1	SD	Proportion/fréquence	
Question 2	PaD	Courbe (donnée)	Courbe "épousant" l'histogramme
	PaD	Probabilité	Aire sous la courbe (aire 2)
	CI	Intégrale	

TABLE 5.3 – Résumé de l'analyse (*activité* 2 p. 377, *Hyperbole*)

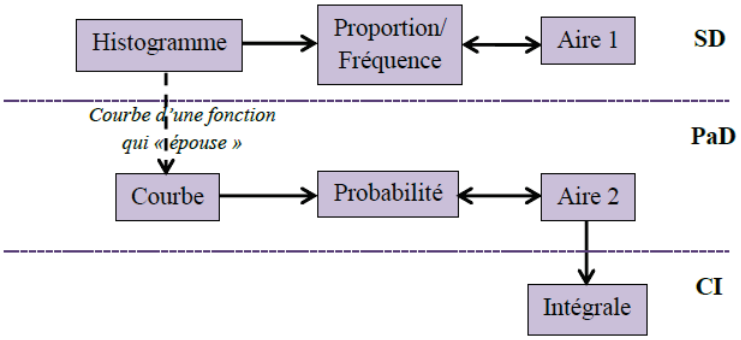


Figure 5.9 – Cheminement de l'*activité* 2 p. 377 (*Hyperbole*)

Quelles idées ressortent sur la fonction de densité ?

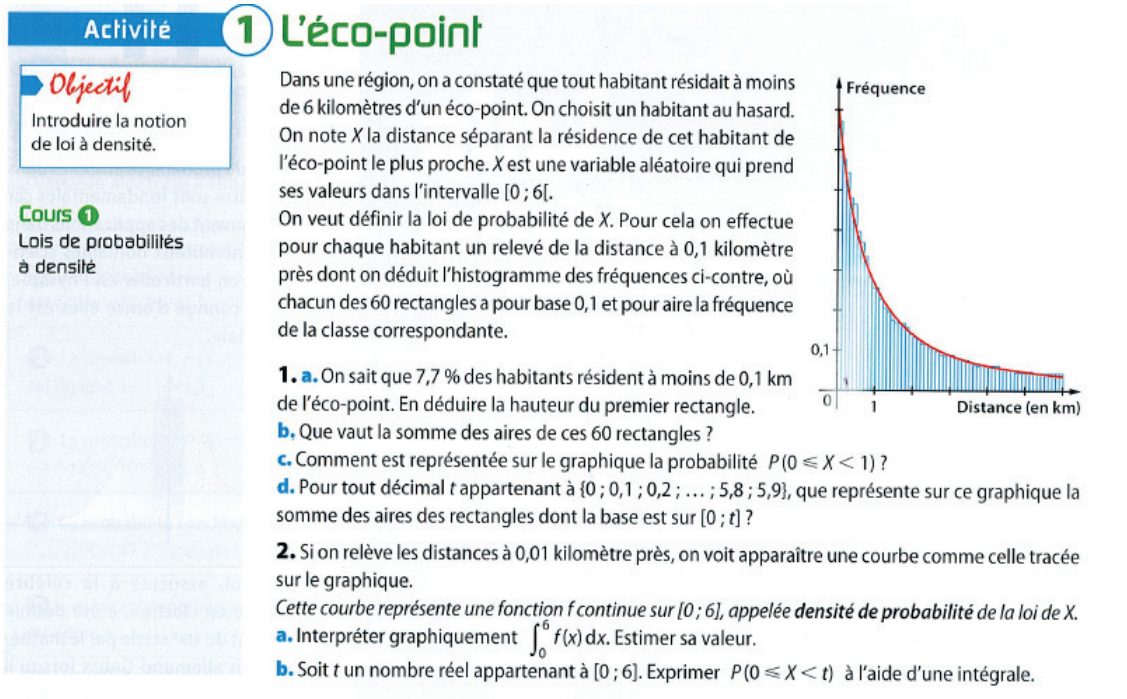
L'*activité* se conclut par : « *On dit que f est* ***la densité*** *de P* ». À nouveau, P n'est pas défini. Aucune justification de cette terminologie n'est donnée. À l'issue de cette *activité*, f est une fonction qui « *épouse* » l'histogramme. L'aire sous f doit valoir 1 (d'après la dernière question), mais nous ne savons pas bien le statut de cette propriété : Est-elle nécessaire ? Est-ce lié à ce cas particulier ? Sachant que la fonction a été donnée en amont, cette propriété s'ajoute sans avoir une place très claire. Mis à part la propriété de « proximité » avec (le haut de) l'histogramme, les autres ne ressortent pas vraiment. En revanche, il ressort qu'elle sert à calculer des probabilités.

Pour revenir sur la fonction f de l'énoncé, nous pouvons constater que l'ensemble de définition de cette fonction n'est pas explicité. Il semble borné au vu du graphique, mais c'est le cas aussi pour une loi normale (le graphique laisse croire que la fonction s'annule sur les bords ce qui n'est pas le cas). S'il s'agit d'une loi normale ici, la question 2 (b) pourrait susciter des débats, notamment sur le fait qu'un domaine infini puisse avoir une aire finie, mais aussi sur la pertinence d'une telle modélisation. Ce flou autour de l'ensemble de définition permet de contourner ces difficultés possibles. Ce qui peut paraître contradictoire ensuite, est que si l'on regarde le début du cours (p. 378), la définition d'une fonction de densité n'est donnée que dans les cas bornés ou bornés à gauche, donc cela exclut le cas de la loi normale par exemple.

b) D'une population à elle-même !

La situation proposée

Dans l'*activité* 1 p. 322 du manuel *Indice* (figure 5.10), intitulé « *L'éco-point* », l'objectif affiché est d'introduire la notion de loi à densité.



Activité 1 L'éco-point

Objectif
Introduire la notion de loi à densité.

Cours 1
Lois de probabilités à densité

Dans une région, on a constaté que tout habitant résidait à moins de 6 kilomètres d'un éco-point. On choisit un habitant au hasard. On note X la distance séparant la résidence de cet habitant de l'éco-point le plus proche. X est une variable aléatoire qui prend ses valeurs dans l'intervalle $[0\,;6[$.
On veut définir la loi de probabilité de X. Pour cela on effectue pour chaque habitant un relevé de la distance à 0,1 kilomètre près dont on déduit l'histogramme des fréquences ci-contre, où chacun des 60 rectangles a pour base 0,1 et pour aire la fréquence de la classe correspondante.

1. a. On sait que 7,7 % des habitants résident à moins de 0,1 km de l'éco-point. En déduire la hauteur du premier rectangle.
b. Que vaut la somme des aires de ces 60 rectangles ?
c. Comment est représentée sur le graphique la probabilité $P(0 \leqslant X < 1)$?
d. Pour tout décimal t appartenant à $\{0\,;0{,}1\,;0{,}2\,;\ldots\,;5{,}8\,;5{,}9\}$, que représente sur ce graphique la somme des aires des rectangles dont la base est sur $[0\,;t]$?

2. Si on relève les distances à 0,01 kilomètre près, on voit apparaître une courbe comme celle tracée sur le graphique.
Cette courbe représente une fonction f continue sur [0 ; 6], appelée ***densité de probabilité*** *de la loi de X.*
a. Interpréter graphiquement $\int_0^6 f(x)\,dx$. Estimer sa valeur.
b. Soit t un nombre réel appartenant à $[0\,;6]$. Exprimer $P(0 \leqslant X < t)$ à l'aide d'une intégrale.

FIGURE 5.10 – *Activité* 1 p. 322 du manuel (*Indice*)

Comme dans le manuel *Hyperbole*, le choix est fait d'étudier des données qui se veulent « réelles » mais, cette fois-ci, l'on dispose des données de la population entière. Ici encore, les données sont en fait « fabriquées » pour l'*activité* et même sont le fruit de la simulation d'une variable aléatoire prédéfinie (par les auteurs).

Nous disposons ici d'un histogramme rassemblant les données de toute la population étudiée, qui est l'ensemble des habitants d'une région, concernant la distance qui sépare le domicile d'un habitant de l'éco-point le plus proche.

L'expérience aléatoire en question dans l'*activité* est le choix d'un habitant au hasard. On s'intéresse ensuite à la variable aléatoire X correspondant à la distance séparant la résidence de l'habitant choisi et l'éco-point le plus proche. Ici, il ne s'agit pas d'estimer la probabilité mais de déterminer la loi de probabilité associée à la variable aléatoire X, car nous disposons de l'ensemble des données.

Un relevé de ces distances a été fait pour l'ensemble des habitants à 0,1 km près et ensuite ces données sont représentées par un histogramme avec des classes d'amplitude 0,1 km. Les données auraient donc pu être représentées par un diagramme en bâtons. Ce n'est pas ce choix qui a été fait, mais dans ce cas, une explication pourrait être donnée sur le choix de la répartition sur les classes. Par exemple, un relevé de 3,4 km a été mis dans la classe $]3,3;3,4]$ ou $[3,4;3,5[$? Nous pouvons penser qu'aucun relevé n'a été de 0 et donc que c'est le premier choix qui a été fait, mais ce choix aurait tout intérêt à être précisé. Il ne va pas de soi, car notamment si les données ont été arrondies, on pourrait penser plutôt à prendre comme classe $[3,35;3,45[$.

Au niveau de l'histogramme, on remarque tout d'abord que l'axe des ordonnées est nommé fréquence : nous avons déjà précisé que ceci est faux, d'autant plus qu'ici l'amplitude des classes est de 0,1 donc on ne peut pas faire l'amalgame entre

fréquence et densité. Cependant, l'énoncé ne dit pas la même chose que le graphique et précise bien que l'histogramme de fréquences a « *pour aire la fréquence de la classe correspondante* ». Donc nous pouvons repérer ici une contradiction entre le texte et le graphique. Comme dans le manuel *Hyperbole*, le référentiel théorique relatif à l'histogramme est donné dans l'énoncé.

Question 1

> **1.a.** On sait que $7,7\%$ des habitants résident à moins de 0,1 km de l'éco-point. En déduire la hauteur du premier rectangle ?
> **1.b.** Que vaut la somme des aires de ces 60 rectangles ?
> **1.c.** Comment est représentée sur le graphique la probabilité $P(0 \leqslant X < 1)$?
> **1.d.** Pour tout décimal t appartenant à $\{0; 0,1; 0,2; ...; 5,8; 5,9\}$, que représente sur le graphique la somme des aires des rectangles dont la base est sur $[0; t]$?

Démarche mathématique attendue

D'après le rappel donné dans l'énoncé, on sait que l'aire d'un rectangle correspond à la fréquence sur cette classe. Donc : $0,1 \times h = 0,077$, ce qui nous donne $h = 0,77$. Il est ensuite demandé la somme des aires de ces 60 rectangles, il s'agit donc de la somme des fréquences de chaque classe, c'est-à-dire la fréquence totale, qui est de 1.

La probabilité $P(0 \leqslant X < 1)$, dans cette *activité*, correspond exactement à la fréquence des habitants résidant à moins de 1 km de l'éco-point le plus proche, c'est-à-dire à l'aire des rectangles sur les 10 classes entre 0 et 1.

Enfin, pour la question 1.d., deux réponses sont envisageables : pour tout décimal t appartenant à $\{0; 0,1; 0,2; ...; 5,8; 5,9\}$, la somme des aires des rectangles dont la base est sur $[0; t]$ représente :

- la fréquence des habitants résidant à moins de t km de l'éco-point le plus proche ;
- la probabilité qu'un habitant choisi au hasard réside à moins de t km de l'éco-point le plus proche.

La première réponse est vraie dans tous les cas, dans le sens où les aires représentent les fréquences dans un histogramme, alors que la seconde est valable dans le cas particulier de l'*activité* car ici nous disposons des données pour la population entière et donc la valeur de la probabilité est égale à celle de la fréquence.

Analyse des activités attendues de l'élève

Les deux premières sous-questions sont dans le sous-domaine de la statistique descriptive (SD). Il s'agit d'un travail dans la dimension sémiotique, le référentiel théorique étant rappelé dans l'énoncé. La hauteur d'un rectangle ne correspond pas, contrairement à ce qui est écrit sur le graphique, à la fréquence mais bien à la densité de fréquence. Pour la somme des aires des 60 rectangles, il est nécessaire

de faire un changement de point de vue sur l'objet mathématique et s'intéresser à la fréquence totale. L'additivité des fréquences pour déterminer la fréquence totale doit être disponible.

Dans la question 1.c., apparaît la probabilité, nous nous situons donc dans le sous-domaine des probabilités à densité (PaD). Cependant, dans cette *activité*, probabilité et fréquence sont assimilées : la valeur de l'une est égale à celle de l'autre. Cela est dû au fait que connaissant les données pour l'ensemble de la population, la probabilité se déduit directement de la fréquence, d'où la double réponse possible à la question 1.d.

Nous voyons donc dans cette question 1, un choix d'*activité* qui a tendance à fusionner les deux sous-domaines. Cela ne pose *a priori* pas de problème, mais ne nait donc pas ici le besoin de chercher une « fonction de densité ».

Question 2

Rien n'est fait pour susciter un nouveau besoin, par exemple, en demandant « Et si on voulait connaître la probabilité de $P(0 \leqslant X < 0,98)$? ». L'énoncé propose une solution avant même d'avoir posé le problème (la question sera posée seulement ensuite). Il est dit : « *Si l'on relève les distances à 0,01 kilomètre près, on voit apparaître une courbe comme celle tracée sur le graphique* ».

Nous avons plusieurs remarques sur cette phrase. Tout d'abord, l'histogramme avec des classes d'amplitude $0,01$ n'est pas proposé dans le manuel pour confirmer cette affirmation (ce qui donnerait un ensemble de rectangles contigus de largeur presque nulle, donc des « segments verticaux accolés »). Ensuite, l'apparition de la courbe correspond en fait au haut des rectangles (pratiquement des segments) formant l'histogramme. Enfin, et surtout, à partir de l'histogramme donné, il n'est pas possible d'obtenir la courbe rouge. Par exemple, entre $0,4$ et $0,5$, le rectangle est clairement au-dessus de la courbe rouge, il est donc impossible que les 10 « petits » rectangles à intercaler dans cet intervalle aient tous une hauteur plus petite que celle du rectangle sur $[0,4; 0,5]$, car l'aire doit rester la même. Ce qui est écrit est donc faux.

Ce que l'on peut dire aussi, qui est valable dès le début, c'est que pour des données réelles, ce type de résultats est impossible, avec une telle précision, il est impossible qu'il n'y ait aucun « trou ».

> **2.a.** Interpréter graphiquement $\int_0^6 f(x)\,\mathrm{d}x$. Estimer sa valeur.
> **2.b.** Soit t un nombre réel appartenant à $[0; 6]$. Exprimer $P(0 \leqslant X < t)$ à l'aide d'une intégrale.

Démarche mathématique attendue

D'après la définition de l'intégrale d'une fonction continue et positive sur un intervalle borné, $\int_0^6 f(x)\,\mathrm{d}x$ correspond à l'aire sous la courbe de f sur $[0; 6]$. D'après

l'introduction de la question 2, on en déduit que l'aire sous la courbe est donc l'aire des rectangles pour des classes d'amplitude $0,01$ et donc que cela vaut 1. Le terme « *Estimer* » porte à confusion. Si l'on utilise ce qui est dit dans l'énoncé, ce n'est pas une estimation mais une valeur exacte. On en déduit de même que $P(0 \leqslant X < t)$ est égale à $\int_0^t f(x)\,\mathrm{d}x$.

Analyse des activités attendues de l'élève

Ces deux sous-questions sont clairement à la liaison des deux sous-domaines : celui des probabilités à densité (PaD) et celui du calcul intégral (CI). Mais étant donné que dans cette *activité*, comme nous l'avons dit plus tôt, le sous-domaine des probabilités à densité est fortement assimilé à celui de la statistique descriptive (SD), nous avons la présence des trois sous-domaines. L'explication donnée sur l'apparition de la courbe de densité rend le travail un peu bancal. Nous pouvons notamment le voir dans la question 2.a. dans laquelle il est demandé d'estimer l'intégrale, alors qu'avec les données de l'énoncé nous ne donnons pas une estimation, mais bien une valeur exacte.

Le travail est toujours avant tout dans la dimension sémiotique. La définition de l'intégrale d'une fonction continue positive en tant qu'aire est ici une connaissance récente supposée disponible.

Pour l'introduction de la fonction de densité, les auteurs auraient pu assumer le fait de vouloir trouver un modèle « régulier » qui permettrait ensuite de faire des calculs plus précis, tout en sachant que c'est un modèle et que ce ne sera plus exactement les données « réelles ». Il aurait été possible, à partir de l'histogramme, de se demander comment estimer par exemple $P(1 \leqslant X < 1,05)$ et, dans ce cas, la recherche d'un modèle aurait été inévitable. En revanche, si l'on dispose de ces données, cela n'a pas d'intérêt. Ce manuel a tendance à confondre la (pseudo) réalité et le modèle.

Conclusion sur les circulations dans cette *activité*

Dans cette *activité* aussi les trois sous-domaines communiquent mais d'une toute autre façon que dans l'*activité* précédente. La statistique et les probabilités sont ici assimilées et ensuite une connexion est faite entre ce duo PaD/SD et le sous-domaine CI. Nous avons essayé de mettre en garde contre les problèmes que peuvent entraîner cette assimilation entre probabilités et statistique.

Ici aussi les activités sont centrées sur la dimension sémiotique des différents domaines. Le référentiel théorique lié à l'histogramme est encore une fois donné dans l'énoncé. Il reste l'interprétation de l'intégrale en tant qu'aire sous la courbe qui doit être disponible pour l'élève.

Les circulations entre les ETM sont proposées en figure 5.11. Les articulations entre les objets mathématiques des différentes questions sont présentées dans le tableau 5.4. Enfin, le diagramme du cheminement de l'*activité* se trouve en figure 5.12. Nous pouvons notamment voir dans ce diagramme la flèche à double sens qui relie l'histogramme et les probabilités, et donc les fréquences et les probabilités, qui entraînent de mauvaises conceptions chez les élèves car bien entendu, ce lien n'est pas présent en général. D'ailleurs on peut remarquer que même chez les auteurs cela

pose des problèmes dans le passage de l'histogramme à la courbe.

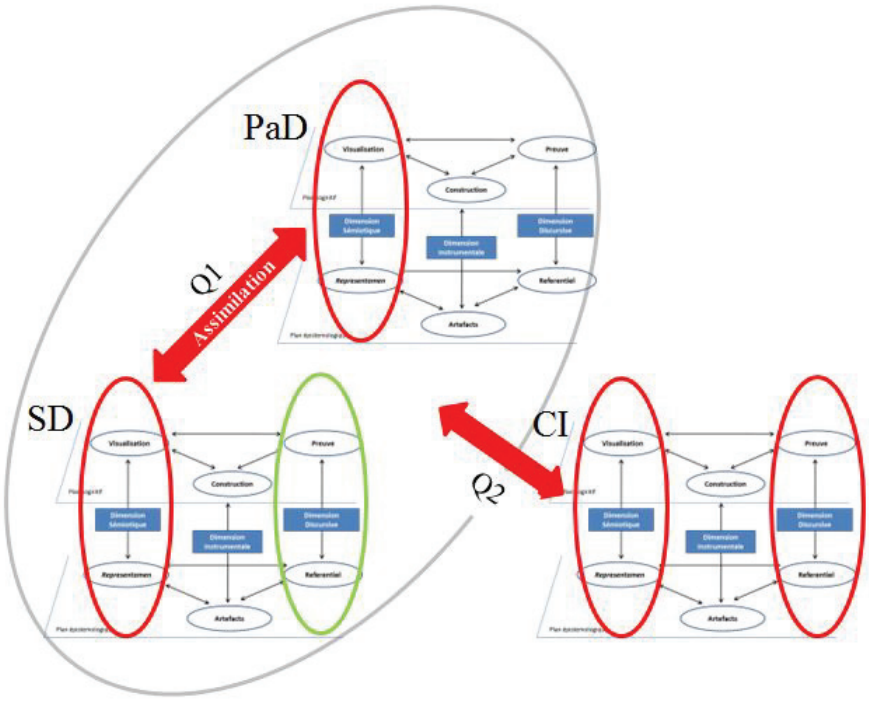


FIGURE 5.11 – Dynamiques entre les trois sous-domaines (*activité* 1 p. 322, *Indice*)

Partie de la tâche	Sous-domaine mathématique	Objet mathématique	Intermédiaire
Enoncé introductif	**PaD**	Variable aléatoire X (de loi inconnue)	Relevé de données
	SD	Histogramme (donné)	
Question 1.a.	**SD**	Hauteur d'un rectangle	Aires des rectangles (aire 1): rappel dans l'énoncé
Question 1.b.	**SD**	Aire des rectangles (aire 1)	Fréquence
Question 1.c.	**PaD**	Probabilité	
Question 1.d.	**PaD/SD**	Probabilité ou fréquence	Diminution de l'amplitude
Question 2. Enoncé	**SD**	Histogramme (non donné)	"Apparition" de la courbe
	PaD	Courbe de densité	
Question 2.a.	**CI**	Intégrale	Aire sous la courbe (aire 2)
Question 2.b.	**CI**	Probabilité/Intégrale	

TABLE 5.4 – Résumé de l'analyse (*activité* 1 p. 322, *Indice*)

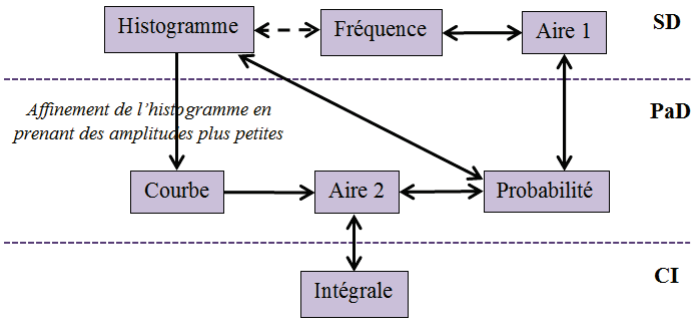


Figure 5.12 – Cheminement de l'*activité* 1 p. 322 (*Indice*)

Quelles idées ressortent sur la fonction de densité ?

Dans cette *activité*, la courbe rouge est ensuite appelée courbe de densité de probabilité de la loi de X. La première « définition » de la courbe de densité serait donc la courbe qui apparait quand on prend 0,1 comme amplitude des classes d'un histogramme (haut des rectangles). Il n'y a même pas l'idée de diminuer de plus en plus les amplitudes. La courbe de densité ne semble pas être attachée à un choix de modèle mais vraiment apparaître d'elle-même. L'intégrale de la fonction sur l'ensemble de définition vaut 1, ce qui vient du fait qu'elle représente la fréquence totale.

c) Par la simulation d'une variable aléatoire

Présentation générale des *activités* du manuel

Le manuel *Math'x* fait le choix d'introduire dans un premier temps la fonction de densité de la loi uniforme (*activité* 1 p. 398), puis celle de la loi exponentielle (*activité* 2 p. 399). La loi normale est ensuite introduite dans les *activités* 3 et 4 qui suivent. Nous remarquons que le manuel *Math'x* est désireux de faire apparaître le lien entre probabilités à densité et calcul intégral par le biais de la comparaison entre les résultats dans les deux sous-domaines indépendamment.

Dans le cas de l'*activité* 1 (cf. Annexe E.2.2), s'appuyant sur la dimension instrumentale de l'ETM, le manuel utilise une simulation de la loi uniforme sur $[0;1]$, à l'aide d'un tableur. La fonction « Histogramme » du tableur lui permet d'obtenir l'histogramme de la partie C (figure 5.13), qui est en fait un diagramme en bâtons (dont les bâtons ont été accollés). On peut le remarquer sur l'axe des abscisses où les classes sont présentées sous forme d'intervalles et non sur la graduation même de l'axe. Dans un « histogramme de tableur », chaque bâton a alors pour hauteur la fréquence et non la densité. Ici encore, cela ne pose pas de problème car les amplitudes des classes sont constantes et égales à 1. Des classes d'amplitudes inégales marqueraient vraiment la différence entre l'« histogramme de tableur », qui est en fait un diagramme en bâtons accolés, et l'histogramme de fréquences.

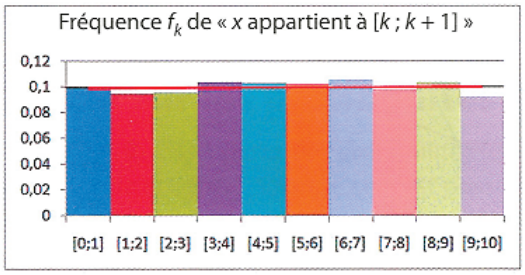


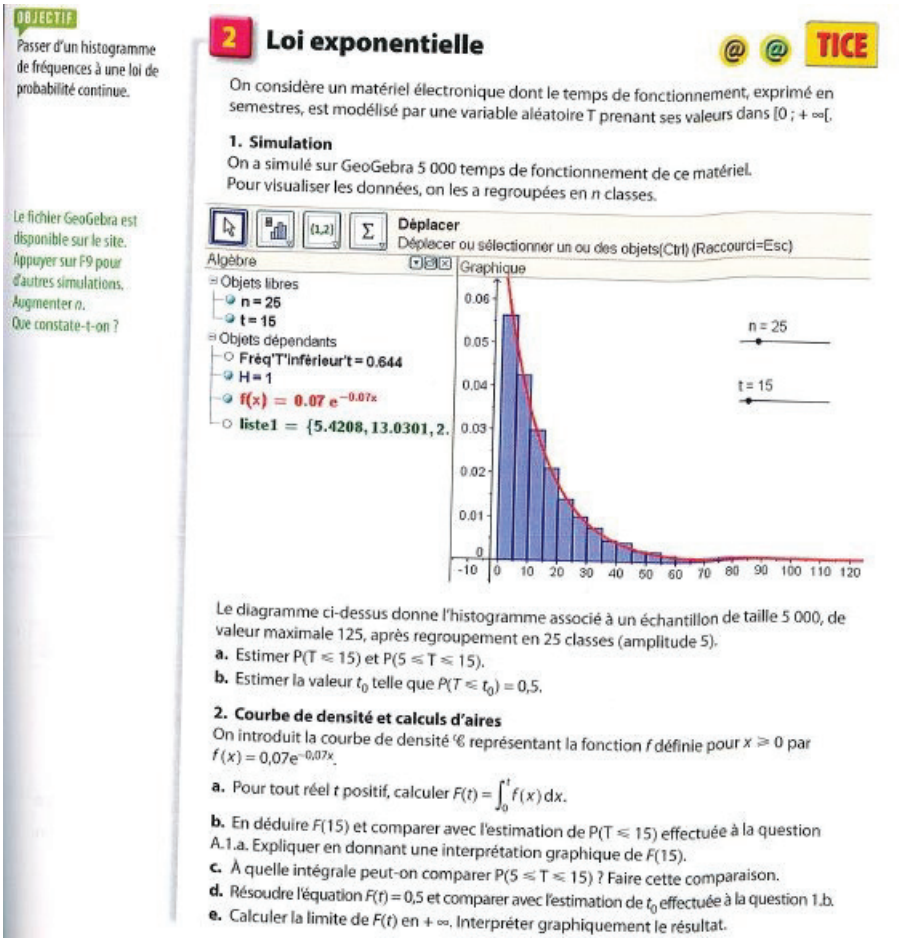
FIGURE 5.13 – Diagramme en bâtons présenté comme étant un histogramme dans l'*activité* 1 pp. 398-399 du manuel *Math'x*

L'utilisation du logiciel est pris en charge par le manuel, il n'est pas explicitement demandé à l'élève de manipuler lui-même le logiciel mais seulement d'utiliser les résultats trouvés et donnés par le manuel. Sur l'histogramme est tracée une courbe rouge, le manuel explique : « *Le tableur affiche une courbe rouge, appelée « courbe de tendance », qui « lisse » l'histogramme* ». Il faut cependant être prudent car une courbe de tendance construite par un tableur, même à partir d'un histogramme, ne vérifie pas nécessairement les conditions nécessaires à une courbe de densité.

Ensuite, l'idée est, dans un premier temps, de déterminer « intuitivement » la probabilité $P(a \leqslant X \leqslant b)$ (ce qui nous l'avons déjà dit est naturel pour la loi uniforme) et ensuite de comparer la probabilité trouvée avec l'intégrale de la fonction représentée par la courbe rouge entre a et b. Il s'agit donc de faire un travail dans le sous-domaine des probabilités à densité et un autre dans le sous-domaine du calcul intégral, tous les deux indépendamment, et ensuite de remarquer un lien entre les deux. Nous ne présenterons pas cette *activité* plus en détail. Nous allons plus longuement étudier l'*activité* suivante.

L'*activité* proposée

Dans la seconde *activité* (*activité* 2 p. 399, en figure 5.14), intitulée « *Loi exponentielle* », il n'est plus possible de déterminer intuitivement des probabilités comme avec la loi uniforme. Nous allons présenter cette *activité* question par question pour voir comment se crée le lien entre probabilités à densité et calcul intégral. L'objectif de l'*activité* est de « *passer d'un histogramme de fréquences à une loi de probabilité continue* ».



OBJECTIF
Passer d'un histogramme de fréquences à une loi de probabilité continue.

Le fichier GeoGebra est disponible sur le site. Appuyer sur F9 pour d'autres simulations. Augmenter n. Que constate-t-on ?

2 Loi exponentielle

TICE

On considère un matériel électronique dont le temps de fonctionnement, exprimé en semestres, est modélisé par une variable aléatoire T prenant ses valeurs dans $[0 ; +\infty[$.

1. Simulation

On a simulé sur GeoGebra 5 000 temps de fonctionnement de ce matériel.
Pour visualiser les données, on les a regroupées en n classes.

Déplacer
Déplacer ou sélectionner un ou des objets(Ctrl) (Raccourci=Esc)

Algèbre
Objets libres
n = 25
t = 15
Objets dépendants
Fréq'T'inférieur't = 0.644
H = 1
f(x) = 0.07 e^-0.07x
liste1 = {5.4208, 13.0301, 2.

Graphique
n = 25
t = 15

Le diagramme ci-dessus donne l'histogramme associé à un échantillon de taille 5 000, de valeur maximale 125, après regroupement en 25 classes (amplitude 5).

a. Estimer $P(T \leqslant 15)$ et $P(5 \leqslant T \leqslant 15)$.

b. Estimer la valeur t_0 telle que $P(T \leqslant t_0) = 0{,}5$.

2. Courbe de densité et calculs d'aires

On introduit la courbe de densité $\mathscr{C}$ représentant la fonction f définie pour $x \geqslant 0$ par $f(x) = 0{,}07e^{-0{,}07x}$.

a. Pour tout réel t positif, calculer $F(t) = \int_0^t f(x)\,dx$.

b. En déduire $F(15)$ et comparer avec l'estimation de $P(T \leqslant 15)$ effectuée à la question A.1.a. Expliquer en donnant une interprétation graphique de $F(15)$.

c. À quelle intégrale peut-on comparer $P(5 \leqslant T \leqslant 15)$? Faire cette comparaison.

d. Résoudre l'équation $F(t) = 0{,}5$ et comparer avec l'estimation de t_0 effectuée à la question 1.b.

e. Calculer la limite de $F(t)$ en $+\infty$. Interpréter graphiquement le résultat.

Figure 5.14 – *activité* 2 p. 399 du manuel *Math'x*

Dans cette *activité*, on s'intéresse à une variable aléatoire T prenant ses valeurs dans $[0; +\infty[$ modélisant le temps de fonctionnement, en semestres, d'un matériel électronique. Le choix de l'ensemble des valeurs prises par la variable aléatoire n'est pas discuté. Il faut comprendre que la loi de la variable aléatoire est connue par les auteurs du manuel (qui ont lancé la simulation proposée) mais pas par l'élève. C'est cette connaissance de la loi qui permet de faire la simulation.

L'*activité* est divisée en deux parties. La première porte le nom « *Simulation* » et la seconde « *Courbe de densité et calculs d'aires* ». Dans la première partie, il est introduit l'histogramme de données simulées, à l'aide de *GeoGebra*, de 5 000 temps de fonctionnement de ce matériel. Ici, il s'agit bien d'un histogramme de fréquences, même s'il n'est pas renseigné à quoi se rapporte l'ordonnée.

Un aparte sur les tâches proposées dans la marge

Dans la marge, il est précisé que le fichier *GeoGebra* est disponible sur le site du manuel et que d'appuyer sur la touche F9 permet de visualiser d'autres simulations. Les auteurs proposent aussi d'augmenter n et demandent ce que l'on constate. Cette partie met en action les dimensions instrumentale et sémiotique. Nous pouvons remarquer qu'elle est cependant à part dans l'*activité*. Elle n'en fait pas partie

intégrante : il semblerait qu'elle ne soit pas considérée comme indispensable pour ce travail.

Dans le fichier *GeoGebra*, le code entré dans les cellules du tableur est $(-ln(random()))/0,07$, qui permet de simuler une variable aléatoire suivant une loi exponentielle de paramètre 0,07, d'où le fait, plus loin dans l'*activité*, que les auteurs introduisent la fonction de densité f définie par $f(x) = 0,07e^{-0,07x}$. La fenêtre Tableur de *GeoGebra* n'est pas visible à l'ouverture du fichier, il faut demander à l'ouvrir pour voir ces simulations.

Le seul travail demandé aux élèves dans cette partie en marge est de visualiser l'histogramme pour d'autres simulations (en appuyant sur la touche F9 du clavier) et de manipuler le curseur n pour visualiser ce qui se passe quand l'amplitude des classes (d'amplitudes égales) diminue. Il est certainement attendu que l'élève remarque que les différents histogrammes semblent avoir toujours la même allure. De plus, quand n augmente, il semblerait qu'alors se dessine une courbe sur le haut des rectangles de l'histogramme. Cependant, cette partie à part n'est pas reprise dans la suite de l'*activité*, donc son objectif n'est pas clair. L'activité de l'élève est entièrement dans le plan sémiotico-instrumental.

Question 1

> 1.a. Estimer $P(T \leqslant 15)$ et $P(5 \leqslant T \leqslant 15)$.
> 1.b. Estimer la valeur t_0 telle que $P(T \leqslant t_0) = 0,5$.

Nous pouvons remarquer qu'ici le terme « Estimer » est pertinent. Ce terme est en effet à mettre en relation avec la statistique inférentielle.

Démarche mathématique attendue 1 - Cas où l'élève utilise seulement le manuel

Pour estimer la probabilité, en utilisant implicitement la loi des grands nombres, on peut l'approcher par la fréquence entre 0 et 15. Ici, il n'y a aucun rappel sur l'histogramme, il faut donc savoir que dans un histogramme la fréquence sur une classe correspond à l'aire du rectangle sur cette classe. Il faut alors déterminer l'aire des trois premiers rectangles, puis sommer les trois pour estimer $P(T \leqslant 15)$ et les deux entre 5 et 15 pour estimer $P(5 \leqslant T \leqslant 15)$. Comme pour l'*activité* du manuel *Hyperbole*, les graduations ne permettent pas une détermination exacte de la hauteur des rectangles, donc la fréquence sera seulement approximative.

On trouve

$$f_1 \simeq 0,056 \times 5 \simeq 0,28,$$

$$f_2 \simeq 0,042 \times 5 \simeq 0,21,$$

$$f_3 \simeq 0,029 \times 5 \simeq 0,145.$$

On en déduit que $P(T \leqslant 15) \simeq 0,635$ et $P(5 \leqslant T \leqslant 15) \simeq 0,355$.

Pour $P(T \leqslant 15)$, on peut aussi remarquer que dans la fenêtre Algèbre de *GeoGebra*, sur la copie d'écran dans le manuel, est donnée « Fréq'T'inférieur't=0,644 », pour t=15, d'après le curseur t de la fenêtre Graphique. Donc $P(T \leqslant 15) \simeq 0,644$, qui est plus précis.

Pour répondre à la question 1.b., d'après la question 1.a., t_0 est inférieur à 15 et d'après les valeurs approchées de f_1 et f_2, on peut remarquer que $P(T \leqslant 10) \simeq 0,49$. Donc on peut estimer $t_0 \simeq 10$. On pourrait essayer d'être plus précis en déterminant les fréquences sur des classes plus petites, en considérant la distribution uniforme sur chaque classe.

Démarche mathématique attendue 2 - Cas où l'élève utilise le fichier *GeoGebra*

Avec le fichier *GeoGebra*, après manipulation, l'échantillon n'est plus le même que celui du manuel mais en utilisant les observations faites dans les questions dans la marge, l'élève peut prendre n'importe quel histogramme et donc n'importe quel échantillon pour estimer les probabilités. Avec le logiciel, il suffit de sélectionner $t = 15$ et de regarder la valeur de « Fréq'T'inférieur't » pour estimer $P(T \leqslant 15)$, puis $t = 5$ pour estimer $P(T \leqslant 5)$.
On peut ensuite en déduire $P(5 \leqslant T \leqslant 15)$ avec $P(T \leqslant 15) - P(T \leqslant 5)$ par exemple. Pour déterminer t_0, il faut faire varier t et observer quand « Fréq'T'inférieur't » est égale à environ 0,5. Pour être plus précis qu'en papier crayon, il faut baisser le pas du curseur qui est au départ de 5.

Analyse des activités attendues de l'élève

Dans cette partie, le travail mathématique part du sous-domaine des probabilités à densité (PaD) pour ensuite aller dans le sous-domaine de la statistique descriptive (SD). La loi des grands nombres est responsable de ce basculement, cependant c'est un théorème en acte, il n'est pas connu des élèves sous ce nom. La fluctuation d'échantillonnage peut être évoquée. Comme adaptation, il faut prendre en compte la propriété d'additivité des fréquences.

Dans la démarche 1, le travail est dans le plan sémiotico-discursif dans l'ETM de la statistique descriptive avec une grande partie de visualisation et de traitement dans le registre histogramme. Les connaissances sur l'histogramme, notion ancienne (pas étudiée depuis la seconde), sont supposées disponibles.

Dans la démarche 2, c'est le plan sémiotico-instrumental qui est en jeu. Il faut comprendre ce à quoi correspondent les valeurs dans la fenêtre Algèbre « Fréq'T'inférieur't », ensuite la dimension instrumentale consiste seulement à déplacer un curseur.

Question 2

La seconde partie de l'*activité* est amenée par l'introduction d'une fonction f définie sur $[0;+\infty[$ par $f(x) = 0,07e^{-0,07}$. On peut considérer que ce terme a été introduit lors de l'*activité* précédente. Cependant, on ne sait pas ce qui est, pour le moment, caractéristique d'une telle fonction.

> 2.a. Pour tout réel t positif, calculer $F(t) = \int_0^t f(x)\,\mathrm{d}x$
> 2.b. En déduire $F(15)$ et comparer avec l'estimation de $P(T \leqslant 15)$ effectuée à la question A.1.a. Expliquer en donnant une interprétation graphique de $F(15)$.
> 2.c. À quelle intégrale peut-on comparer $P(5 \leqslant T \leqslant 15)$? Faire cette comparaison.
> 2.d. Résoudre l'équation $F(t) = 0,5$ et comparer avec l'estimation de t_0 effectuée à la question 1.b.
> 2.e. Calculer la limite de $F(t)$ en $+\infty$. Interpréter graphiquement le résultat.

Démarche mathématique attendue

Nous n'allons pas tout détailler ici. Il s'agit de déterminer l'expression de F la primitive de f qui correspond à l'aire sous la courbe entre 0 et t :

$$F(t) = \int_0^t f(x)\,\mathrm{d}x = 1 - e^{-0,07t}.$$

On en déduit une valeur approchée de $F(15)$: $F(15) = 1-e^{-0,07\times 15} \simeq 0,65$. En comparant avec l'estimation de $P(T \leqslant 15)$ de la question 1, on s'aperçoit que les deux valeurs sont très proches, ce qui s'explique par le fait que $F(15)$ correspond à l'aire sous la courbe sur l'intervalle $[0;15]$ et qu'étant donné que la courbe est proche du haut des rectangles de l'histogramme, les deux valeurs sont nécessairement proches. Le même raisonnement est à faire, avec moins d'indications, pour les questions 2.c. et 2.d. Pour la question 2.e., il s'agit d'un calcul de limite d'une fonction. On obtient

$$\lim_{t\to+\infty} F(t) = 1,$$

ce qui signifie que l'aire sous la courbe de f sur $[0;+\infty[$ est de 1.

Analyse des activités attendues de l'élève

Dans cette partie de l'*activité*, le travail est essentiellement dans le sous-domaine du calcul intégral et ce travail permet une comparaison entre le sous-domaine du calcul intégral (CI) et celui des probabilités à densité (PaD), par l'intermédiaire des résultats de la question avec ceux obtenus dans la question 1. Ce travail permet d'articuler les deux ETM avec une comparaison de la probabilité estimée via la fréquence et de l'intégrale. On en déduit aussi que l'aire sous la courbe de f vaut 1.

Il y a ici une étude de l'existence de l'intégrale impropre, même si cela n'est pas perçu comme tel pour un élève. Ensuite, l'interprétation graphique pourrait ne pas poser de problème conceptuel car bien que la fonction soit définie sur $[0;+\infty[$ dans l'énoncé, sur le graphique elle semble nulle à partir de 115 et donc le questionnement sur l'aire finie d'un domaine infini semble court-circuité. De cette propriété de la fonction (que l'aire sous la courbe vaut 1), nous ne savons pas si nous devons en tirer une propriété nécessaire pour une telle fonction ou si c'est seulement un fait lié à l'*activité*. Le statut de cette question n'est pas clair.

Nous pourrions nous demander si l'objectif du manuel qui est de « passer d'un histogramme de fréquences à une loi de probabilité continue » est bien rempli. Tout d'abord, la loi est en réalité connue (par les auteurs certes) avant d'avoir l'histogramme de fréquences car c'est elle qui permet de faire les simulations. Sinon il faudrait se demander par exemple la loi que suit la variable aléatoire $Y = -\frac{lnX}{0,07}$, ce qui n'aurait sûrement pas de sens. Dans un second temps, mis à part ce point, la courbe est ensuite directement donnée.

Conclusion sur les circulations dans cette *activité*

Le choix dans cette *activité* est, dans un premier temps, de faire une analogie entre les deux ETM des probabilités à densité et de la statistique pour estimer la probabilité, appuyé sur l'histogramme et le fait que la fréquence se rapproche de plus en plus de la probabilité (référentiel théorique), en tant que théorème en acte. Il y a ici clairement une séparation entre l'ETM_{CI} et les deux autres ETM. La connexion est ensuite faite par comparaison des deux groupes (cf. figure 5.15).

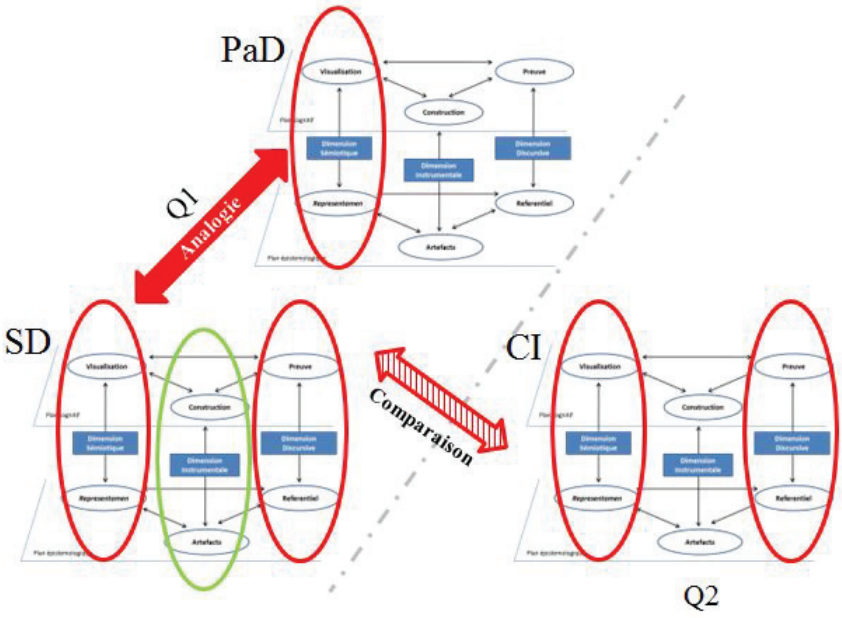


Figure 5.15 – Dynamiques entre les trois sous-domaines (*activité* 2 p. 399, *Math'x*)

Le tableau 5.5, qui propose les articulations entre les objets mathématiques des différentes questions, permet notamment de s'apercevoir que dans cette activité, le passage de l'histogramme à la courbe n'est pas du tout accompagné d'un discours de justification de choix. Nous reviendrons sur le point dans le paragraphe suivant.

Partie de la tâche	Sous-domaine mathématique	Objet mathématique	Intermédiaire
Enoncé introductif	**PaD**	Variable aléatoire T (de loi connue par les auteurs du manuel et inconnue pour les élèves)	Simulation avec GeoGebra (entièrement pris en charge
Question 1. Enoncé	**SD**	Histogramme (donné)	Aires des rectangles (aire 1)
Question 1.a.	**PaD**	Probabilité (estimation)	Pas de justification dans l'activité
Question 1.b.	**PaD**	Probabilité inverse (estimation)	
Question 2. Enoncé	**PaD**	Courbe et fonction de densité	Propriétés des primitives
Question 2.a.	**CI**	Primitive/Intégrale	
Question 2.b.	**CI**	Intégrale	Applications numériques
Question 2.c.	**CI**	Intégrale	Résolution d'équation
Question 2.d.	**CI**	Primitive	
Question 2.e.	**CI**	Limite d'une fonction/Aire	Théorème sur les limites de fonctions

Aire sous la courbe (aire 2)

TABLE 5.5 – Résumé de l'analyse (*activité* 2 p. 399, *Math'x*)

Enfin, la figure 5.16 est un moyen de décrire la cheminement de cette *activité*.

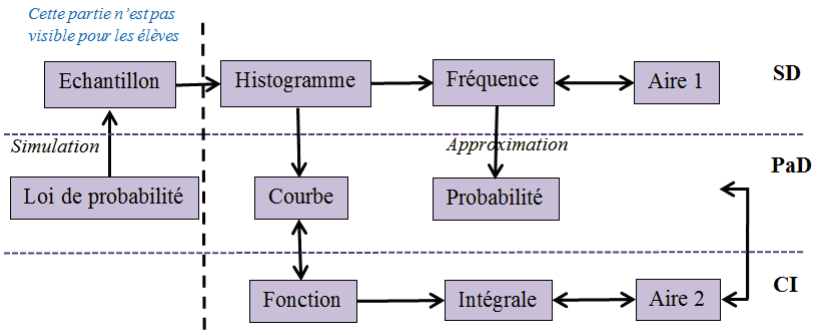


FIGURE 5.16 – Cheminement de l'*activité* 2 p. 399 (*Math'x*)

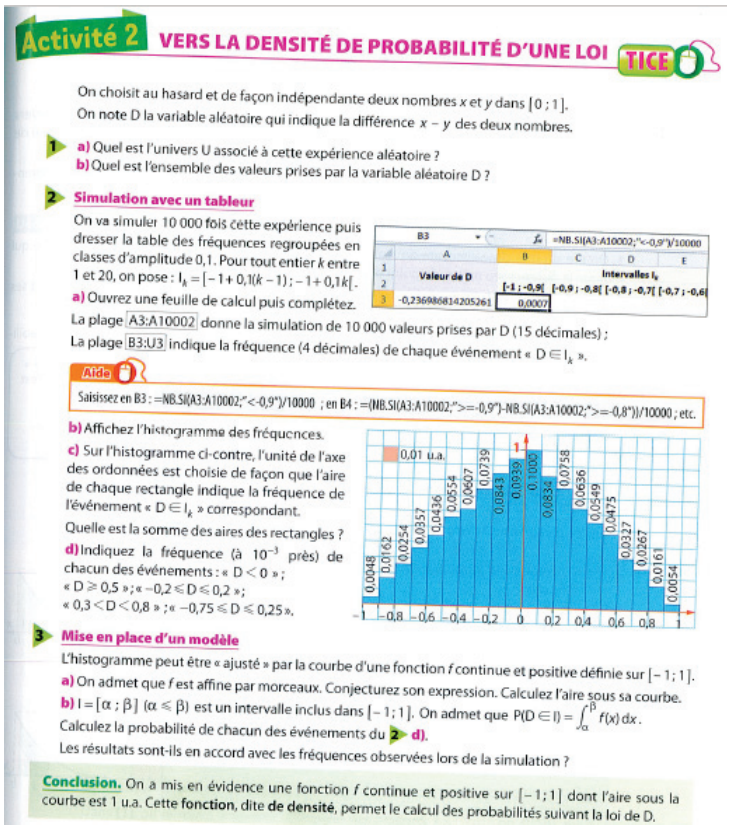
Quelles idées ressortent sur la fonction de densité ?

Dans l'*activité* qui précède celle-ci dans le manuel, pour parler de la courbe de densité, il est question de la « *courbe de tendance* » du tableur qui « *lisse* » l'histogramme. Dans cette *activité*, cette qualité de courbe qui lisse l'histogramme n'apparaît pas dans le texte, cependant elle reste visible sur le graphique. Cependant, ici, l'expression de la fonction est donnée brutalement. Les conditions pour la trouver ne sont pas explicitées. On déduit aussi des calculs que l'aire sous la courbe vaut 1, mais là encore nous ne savons pas le statut de cette propriété.

d) De lois connues vers une loi inconnue

L'*activité* proposée

Le manuel *Transmath* propose une première *activité* sur la loi uniforme, puis ensuite une *activité* intitulée « *Vers la densité de probabilité d'une loi* » (*activité* 2 p. 379, figure 5.17). Nous analysons ici seulement l'*activité* 2 p. 379 car elle a pour objectif (dans le titre) d'amener à la notion de densité de probabilité.



Activité 2 VERS LA DENSITÉ DE PROBABILITÉ D'UNE LOI TICE

On choisit au hasard et de façon indépendante deux nombres x et y dans $[0\,;1]$.
On note D la variable aléatoire qui indique la différence $x - y$ des deux nombres.

1 a) Quel est l'univers U associé à cette expérience aléatoire ?
b) Quel est l'ensemble des valeurs prises par la variable aléatoire D ?

2 **Simulation avec un tableur**

On va simuler 10 000 fois cette expérience puis dresser la table des fréquences regroupées en classes d'amplitude 0,1. Pour tout entier k entre 1 et 20, on pose : $I_k = [-1 + 0{,}1(k-1)\,;-1+0{,}1k[$.

a) Ouvrez une feuille de calcul puis complétez.

	A	B	C	D	E
1					
2	Valeur de D			Intervalles I_k	
3	-0,23698681420526	0,0007			

La plage A3:A10002 donne la simulation de 10 000 valeurs prises par D (15 décimales) ;
La plage B3:U3 indique la fréquence (4 décimales) de chaque événement « $D \in I_k$ ».

Aide
Saisissez en B3 : =NB.SI(A3:A10002;"<-0,9")/10000 ; en B4 : =(NB.SI(A3:A10002;">=-0,9")-NB.SI(A3:A10002;">=-0,8"))/10000 ; etc.

b) Affichez l'histogramme des fréquences.

c) Sur l'histogramme ci-contre, l'unité de l'axe des ordonnées est choisie de façon que l'aire de chaque rectangle indique la fréquence de l'événement « $D \in I_k$ » correspondant.
Quelle est la somme des aires des rectangles ?

d) Indiquez la fréquence (à 10^{-3} près) de chacun des événements : « $D < 0$ » ;
« $D \geqslant 0{,}5$ » ; « $-0{,}2 \leqslant D \leqslant 0{,}2$ » ;
« $0{,}3 < D < 0{,}8$ » ; « $-0{,}75 \leqslant D \leqslant 0{,}25$ ».

0,01 u.a.
0,0048 0,0162 0,0254 0,0357 0,0436 0,0554 0,0607 0,0739 0,0843 0,0939 0,1000 0,0836 0,0758 0,0656 0,0549 0,0475 0,0327 0,0267 0,0161 0,0054
−1 −0,8 −0,6 −0,4 −0,2 0 0,2 0,4 0,6 0,8 1

3 **Mise en place d'un modèle**

L'histogramme peut être « ajusté » par la courbe d'une fonction f continue et positive définie sur $[-1\,;1]$.
a) On admet que f est affine par morceaux. Conjecturez son expression. Calculez l'aire sous sa courbe.
b) $I = [\alpha\,;\beta]$ $(\alpha \leqslant \beta)$ est un intervalle inclus dans $[-1\,;1]$. On admet que $P(D \in I) = \int_\alpha^\beta f(x)\,dx$.
Calculez la probabilité de chacun des événements du 2 d).
Les résultats sont-ils en accord avec les fréquences observées lors de la simulation ?

Conclusion. On a mis en évidence une fonction f continue et positive sur $[-1\,;1]$ dont l'aire sous la courbe est 1 u.a. Cette **fonction**, dite **de densité**, permet le calcul des probabilités suivant la loi de D.

Figure 5.17 – *activité* 2 p. 379 du manuel *Transmath*

Ce manuel a fait un choix encore différent des trois premiers. Il s'agit d'une *activité* TICE, où contrairement au manuel *Math'x*, il y a réellement des questions relatives à l'utilisation de l'outil informatique, ici le tableur. Cette *activité* est similaire à l'*activité* 3 p. 329 (cf. Annexe E.2.3) du manuel *Repères*, que nous n'analyserons pas ici, pour cette raison.

L'expérience aléatoire en question dans l'*activité* est le choix au hasard de deux nombres x et y dans $[0\,;1]$. On s'intéresse alors à la variable aléatoire D qui indique la différence $x - y$, en d'autres termes la variable aléatoire D est égale à $X - Y$, avec X et Y deux variables aléatoires suivant une loi uniforme sur $[0\,;1]$. Il s'agit ensuite de faire des simulations sur un tableur pour déterminer la loi que suit la variable aléatoire D.

Question 1

> **1.a)** Quel est l'univers U associé à cette expérience aléatoire ?
> **1.b)** Quel est l'ensemble des valeurs prises par la variable aléatoire D ?

Démarche mathématique attendue

L'univers U est égal à $[0;1]^2$ et l'ensemble des valeurs prises par la variable aléatoire D est $[-1;1]$.

Analyse des activités attendues de l'élève

Ce travail est dans le domaine des probabilités. Cela repose seulement sur la loi uniforme sur $[0;1]$ qui a été vue dans l'*activité* précédente. Cependant, l'*activité* précédente n'est pas nécessaire pour répondre aux questions. Le référentiel théorique à mobiliser est la définition de l'univers, connaissance ancienne en probabilités discrètes supposée disponible.

Question 2

Ensuite, l'*activité* passe à une partie *Simulation*. Il est proposé de simuler 10 000 fois cette expérience aléatoire et ensuite de dresser un tableau de fréquences.

> **2.a)** Ouvrez une feuille de calcul puis complétez. [...]
> **2.b)** Affichez l'histogramme des fréquences.

Démarche mathématique attendue

La plage $A3 : A10002$ donne la simulation de 10 000 valeurs prises par D par la formule : $= ALEA() - ALEA()$. Il faut donc rentrer cette formule dans la cellule $A3$ puis l'étirer. La formule pour la fréquence est donnée en aide dans le manuel pour la cellule $B3$. Le manuel propose aussi une formule pour $B4$, qui est en fait une formule pour la cellule $C3$. La formule est ensuite à adapter pour chaque cellule. L'inconvénient de cette formule est qu'elle ne peut pas être étirée. Il aurait été plus judicieux, nous pensons, de rajouter une ligne (par exemple sur la plage $B1 : U1$ pour permettre une explication plus simple) avec la borne supérieure de l'intervalle de la cellule $B3 : U3$ correspondante et ensuite de proposer comme formule en $B3$:

$$= NB.SI(A3 : A10002; " < B1)/10000,$$

et en $C3$:

$$= NB.SI(A\$3\$: A\$10002\$; " < C1)/10000 - B3.$$

Cette formule serait alors étirable jusqu'en $U3$.

Après avoir sélectionné les cellules $B2 : U2$ et $B3 : U3$, il est ensuite attendu par les auteurs de sélectionner le mode « Histogramme » du tableur pour construire ce qu'ils appellent l'histogramme de fréquences associé. Il s'agit en fait d'un diagramme en bâtons des fréquences avec des rectangles contigus, comme nous l'avons déjà fait remarquer dans l'*activité* du manuel *Math'x*. Ce n'est pas simplement une hypothèse que nous formulons, cela est soutenu par le fait que dans l'*activité* précédente nous retrouvons ce même type d' « histogramme » (cf. figure 5.18). Sur cette figure, il ne s'agit effectivement pas d'un histogramme de fréquences car l'aire du rectangle sur une classe n'est pas égale à la fréquence sur cette classe, mais à $f/100$ et ne serait même pas proportionnelle à la fréquence si les classes n'étaient pas de même amplitude. Pour confirmer le fait qu'il ne s'agisse pas d'un véritable histogramme de fréquences, à partir de la question suivante, un nouvel histogramme est introduit dont « *l'unité de l'axe des ordonnées est choisie de façon à ce que l'aire de chaque rectangle indique la fréquence de l'événement « $D \in I_k$ » correspondant* », ce qui sous-entend que ce n'est pas toujours le cas. Ceci est justifiable par le fait que certaines définitions de l'histogramme exigent seulement la proportionnalité entre l'aire et la fréquence, mais cela n'enlève pas le flou autour de l'histogramme du tableur.

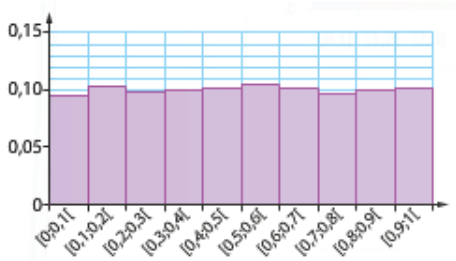


Figure 5.18 – « Histogramme », étant en fait un diagramme en bâtons, de l'*activité* 1 p. 378 du manuel (*Transmath*)

Analyse des activités attendues de l'élève

Il s'agit ici d'un travail de simulation sur tableur. Nous sommes donc dans le domaine de la statistique descriptive (SD), soutenue par le domaine des probabilités qui permet la simulation : c'est parce que nous savons que les variables aléatoires, que nous avons appelées X et Y, suivent des lois uniformes sur $[0; 1]$ que nous pouvons faire la simulation.

L'activité de l'élève se fait essentiellement dans la dimension instrumentale mais les formules à utiliser pour les fréquences sont données en aide. Cependant, nous l'avons fait remarquer, le choix fait n'est pas optimal. Seule la connaissance de la fonction $ALEA()$ est une connaissance supposée disponible ici. Ensuite, pour la construction de l'« histogramme » qui est en fait un diagramme en bâtons, il suffit de sélectionner les bonnes commandes.

Comme nous l'avons précisé juste avant, l'histogramme attendu n'est pas réellement un histogramme. Cela poserait particulièrement des problèmes si les ampli-

tudes des classes étaient inégales. Pour cette raison, les auteurs introduisent dans le manuel, un autre histogramme, qui malgré leur commentaire, est cette fois-ci l'histogramme de fréquences. Ils ont fait le choix de ne rien mettre sur l'axe des ordonnées mais de donner l'unité d'aire. Il est rappelé comment est construit cet histogramme.

> **2.c)** Quelle est la somme des aires des rectangles ?
> **2.d)** Indiquez la fréquence (à 10^{-3} près) de chacun des événements :
> « $D < 0$ » ; « $D \geqslant 0,5$ » ; « $-0,2 \leqslant D \leqslant 0,2$ » ; « $-0,3 < D < 0,8$ » ;
> « $-0,75 \leqslant D \leqslant 0,25$ ».

Démarche mathématique attendue

À l'aide du rappel donné dans l'énoncé, sur le lien entre histogramme et fréquence, on peut en déduire que la somme des aires des rectangles correspond à la somme des fréquences entre -1 et 1 et donc que cette aire est égale à 1 Ou encore, les fréquences des classes étant indiquées, il est possible de sommer toutes ces fréquences et pour obtenir 1.

Ensuite, pour déterminer la fréquence de chacun des événements, il suffit de déterminer l'aire des rectangles correspondants, sachant que la fréquence de chaque classe est donnée. Pour le dernier événement, $-0,75$ et $0,25$ n'étant pas des extrémités de classes, il faut supposer la distribution uniforme sur la classe. Par exemple, pour la fréquence sur $[-0,75; -0,7]$, il faut considérer la fréquence sur $[-0,8; -0,7]$ divisée par 2.

Nous pouvons nous demander ici quel est l'intérêt d'avoir fait construire un « histogramme » sur le tableur si c'est pour en proposer un autre ensuite, vu que ce n'est finalement pas celui qui nous intéresse. De plus, ce n'est pas précisé mais l'histogramme proposé correspond à un autre échantillon. Il aurait été possible de répondre à la question 2.d) à l'aide du tableau de fréquences fait sur le tableur, donc l'intérêt ici de ce deuxième histogramme n'est pour le moment pas clair.

Analyse des activités attendues de l'élève

Les activités attendues sont dans le sous-domaine de la statistique descriptive (SD). Il s'agit de déterminer des fréquences à partir d'un histogramme, sachant que le lien entre histogramme et fréquence a été rappelé dans l'énoncé (référentiel théorique donné). C'est essentiellement un travail dans la dimension sémiotique, pour reconnaître que les valeurs mises sur l'histogramme en haut de chaque rectangle correspond à la fois à l'aire et à la fréquence sur cette classe. Il y a notamment une place à la visualisation.

Question 3

L'intérêt de ce deuxième histogramme de fréquences (le « vrai ») arrive justement dans cette question. Il est mentionné que « *l'histogramme peut être « ajusté »*

par la courbe d'une fonction f continue et positive définie sur $[-1;1]$ ». C'est la seule *activité* d'un manuel où la courbe n'est pas directement tracée. Il est demandé à l'élève de chercher son expression.

> **3.a)** On admet que f est affine par morceaux. Conjecturez son expression. Calculez l'aire sous sa courbe.
> **3.b)** Calculez la probabilité de chacun des événements de 2.d).
> Les résultats sont-ils en accord avec les fréquences observées lors de la simulation ?

Démarche mathématique attendue

Il est précisé que la fonction f est affine par morceaux. Il semblerait qu'elle soit composée de deux morceaux. Avec une réflexion sur la symétrie de la distribution, on en vient à supposer que la fonction est paire. Le 1 en rouge sur l'axe des ordonnées laisse ensuite penser que la courbe va passer par ce point. Il faut alors trouver l'expression de la fonction affine sur [0 ;1] passant par le point $(0;1)$ et $(1;0)$. On obtient :

$$f(x) = \begin{cases} -x+1 & \text{si } x \in [0;1] \\ x+1 & \text{si } x \in [-1;0] \end{cases}$$

Si on calcule l'aire sous la courbe, on obtient 1. Il n'est pas précisé si ce calcul doit permettre ou non de valider la conjecture de l'expression de la fonction.

Ensuite, il est admis dans l'exercice que $P(D \in I) = \int_{\alpha}^{\beta} f(x)\,\mathrm{d}x$, si $I = [\alpha;\beta]$. Donc le calcul des probabilités revient à des calculs d'intégrales (ou d'aires de surfaces élémentaires). Les résultats sont proches de ceux trouvés pour les fréquences en 2.b).

Analyse des activités attendues de l'élève

La question 3.a) est dans le domaine de l'analyse pour la recherche de l'expression de la fonction, puis plus précisément dans le sous-domaine du calcul intégral (CI) pour le calcul de l'aire. Ensuite, les auteurs du manuel imposent une connexion entre le sous-domaine des probabilités à densité et celui du calcul intégral en admettant que $P(D \in I) = \int_{\alpha}^{\beta} f(x)\,\mathrm{d}x$, si $I = [\alpha;\beta]$. Pour déterminer l'expression de la fonction dans la question 3.a), la visualisation de l'histogramme, qui est censé aider l'élève, peut finalement devenir un obstacle pour lui car la symétrie n'est pas visible sur l'histogramme. Il aurait pu être intéressant justement de donner plusieurs histogrammes pour aider à appréhender le fait que la fonction doit être symétrique.

Les connaissances en jeu sont sur le calcul d'intégrales de fonctions affines. Deux démarches sont alors possibles : soit l'utilisation d'une primitive, soit l'utilisation des formules sur les aires de surfaces élémentaires. Ce sont des connaissances récentes pour la première méthode ou anciennes pour la seconde, supposées disponibles ici. Les probabilités des événements proposés sont ensuite comparées aux fréquences

trouvées ultérieurement pour voir une analogie entre la fréquence et l'intégrale, qui correspond à la probabilité.

Ce regard est différent de celui que l'on a pu trouver dans le manuel *Math'x*. Il ne s'agit pas ici de comparer d'une part la probabilité, estimée par la fréquence (d'après la loi des grands nombres), et de la comparer avec l'intégrale, mais de comparer la probabilité définie par une intégrale (définition imposée) à la fréquence. Nous retrouvons bien des valeurs proches, ceci étant assuré par la loi des grands nombres. L'objectif n'est donc pas de faire émerger le lien entre probabilités et intégrales, mais de vérifier que cela fonctionne bien.

Conclusion sur les circulations dans cette *activité*

Dans cette dernière *activité* ressort, à nouveau, une dynamique différente entre les trois ETM (cf. figure 5.19). Ici l'ETM_{SD} est travaillé d'une part, tandis que les ETM_{PaD} et ETM_{CI} sont assimilés (admis dans l'énoncé). Une comparaison est alors faite entre les deux blocs pour valider le fait de considérer la probabilité comme une intégrale. Le tableau 5.6 et la figure 5.20 nous permettent de mettre en avant aussi la rupture nette qui existe entre le travail sur tableur et le travail en papier-crayon.

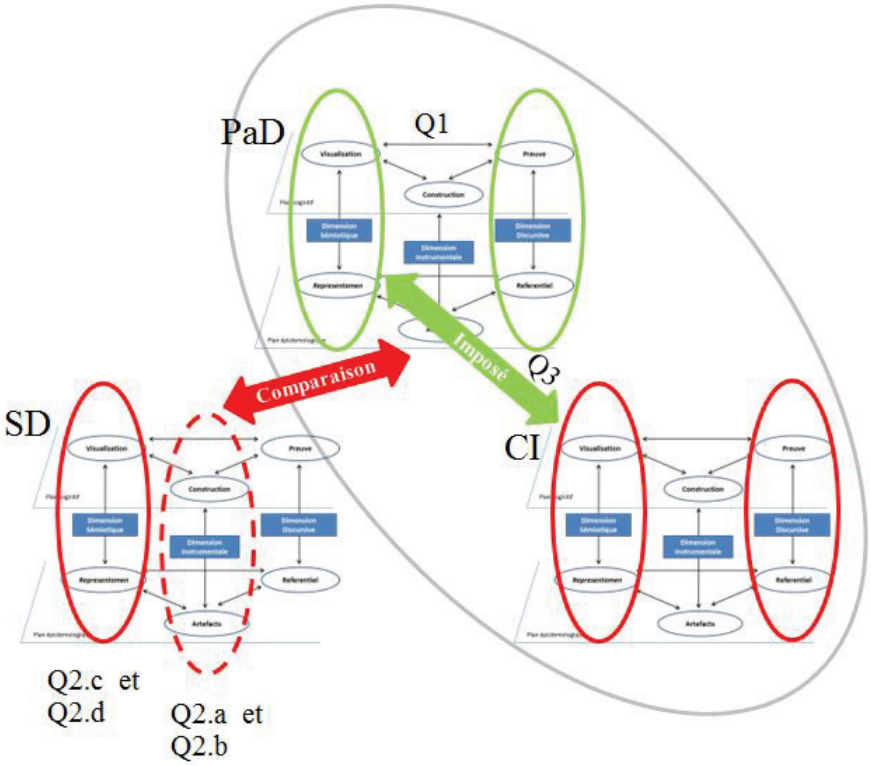


FIGURE 5.19 – Circulations entre les trois sous-domaines (*activité* 2 p. 379, *Transmath*)

Partie de la tâche	Sous-domaine mathématique	Objet mathématique	Intermédiaire
Enoncé introductif	**PaD**	Variable aléatoire D (de loi inconnue, mais étant la différence de deux lois connues)	Simulations avec le tableur Fonctionnalité du tableur Fréquence : rappel dans l'énoncé Aires des rectangles (aire 1): rappel dans l'énoncé Ajustement de l'histogramme Comparaison Intégrale : admis dans l'énoncé
Question 1.a)	**PaD**	Univers	
Question 1.b)	**PaD**	Valeurs prises par la variable aléatoire	
Question 2.a)	**SD**	Simulation, fréquence	
Question 2.b)	**SD**	" Histrogramme " de fréquences de tableur	
Question 2.c)	**SD**	Histogramme (donné)	
	SD	Aire des rectangles (aire 1)	
Question 2.d)	**SD**	Fréquence	
Question 3. Enoncé.	**PaD**	Existence d'une fonction de densité (continue et positive)	
Question 3. a)	**CI**	Fonction de densité	
	CI	Aire sous la courbe (aire 2)	
Question 3. b)	**PaD/CI**	Probabilité	

Table 5.6 – Résumé de l'analyse (*activité* 2 p. 379, *Transmath*)

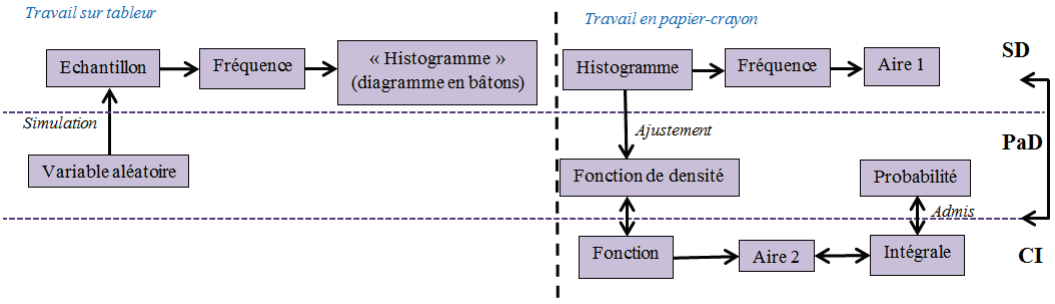


Figure 5.20 – Cheminement de l'*activité* 2 p. 379 (*Transmath*)

Quelles idées ressortent sur la fonction de densité ?

La fonction f est amenée comme une fonction continue et positive définie sur $[-1;1]$ dont la courbe « ajuste » l'histogramme. Ici encore, il est demandé plus tard de calculer l'aire sous la courbe de f, mais ce calcul vient après la conjecture de l'expression de la fonction, alors qu'au contraire elle pourrait être une condition nécessaire pour trouver l'expression. De plus, il n'est pas précisé que ce calcul d'aire pourrait être un moyen d'auto-régulation pour la détermination de l'expression.

L'*activité* 3 p. 329 du manuel *Repères* (cf. Annexe E.2.3) propose aussi de passer par la simulation, mais cette fois-ci avec le logiciel Xcas. La variable aléatoire étudiée est la somme de deux variables aléatoires suivant une loi uniforme sur $[0; 1]$.

5.2.4 Résultats

Après cette analyse des tâches proposées dans les *activités* d'introduction de la notion de fonction de densité dans les manuels, nous pouvons mettre en avant quelques résultats. Les premiers étant que la majorité des manuels (5 manuels sur 8) introduisent la fonction de densité par le biais de la notion d'histogramme.

Nous nous sommes focalisée sur quatre *activités* proposant ce type d'approche. Dans ces *activités*, le choix de faire intervenir un histogramme propose donc un passage par le sous-domaine de la statistique descriptive pour arriver aux probabilités à densité et son lien avec le calcul intégral. Les trois sous-domaines sont donc en jeu dans les activités attendues des élèves. Cela peut laisser penser à l'approche proposée par Faye (chapitre 3). Cependant, la place de la statistique descriptive n'est pas toujours la même dans les différents manuels et les interactions entre les trois ETM non plus, et nous verrons que nous ne retrouvons pas tout à fait l'approche historique.

a) Des catégories de démarches différentes

Les cinq manuels proposent deux grandes catégories de cheminement : soit l'on part de la statistique descriptive avec des données « réelles » pour aller vers les probabilités (entrée par le sous-domaine SD) ; soit l'on part des probabilités, avec une variable aléatoire, pour aller à la statistique descriptive, par la simulation et revenir aux probabilités (entrée par le sous-domaine PaD).

Dans la première catégorie, nous avons rencontré deux cas. La catégorie 1.1 est celle où nous disposons des données d'un échantillon de la population (SD) et nous cherchons à estimer des probabilités sur la population entière (PaD). Nous retrouvons cette démarche dans l'*activité* du manuel *Hyperbole*. Dans le cours de Faye présenté dans le chapitre 3, la démarche correspond à ce cas. Cependant, nous pourrons en dégager des différences par rapport à ce qui se trouve dans les manuels. La catégorie 1.2 correspond à la démarche du manuel *Indice* : nous disposons des données de l'ensemble de la population (SD) et nous voulons les probabilités associées.

Dans la deuxième catégorie, nous pouvons repérer aussi deux cas. Dans la catégorie 2.1, la loi est inconnue de l'élève, mais connue des auteurs du manuel qui proposent eux-mêmes les simulations (les formules sont déjà rentrées dans le fichier) pour ensuite permettre de retrouver la loi. C'est le cas du manuel *Math'x*. Cette démarche peut paraître artificielle. Enfin, la catégorie 2.2 est celle où la loi est inconnue mais est une combinaison de lois connues (comme la loi uniforme, elle est connue du fait que les élèves savent simuler le choix au hasard d'un nombre sur $[0; 1]$ [23]). Il est

23. En réalité, on choisit au hasard un nombre sur $[0; 1]$ avec n décimales, n fixé. La probabilité de chaque issue est de $1/10^n$, qui est quasi nulle si n est grand, d'où la convergence en loi vers la loi uniforme continue.

donc possible de simuler et après de déterminer la loi. Cette démarche se retrouve dans les manuels *Transmath* et *Repères*.

Le tableau 5.7 regroupe ces quatre sous-catégories.

Catégorie 1 Entrée : SD				**Catégorie 2** Entrée : PaD			
Catégorie 1.1		Catégorie 1.2		Catégorie 2.1		Catégorie 2.2	
SD	Données réelles Echantillon ↓	SD	Données réelles Population entière ↓	PaD	Loi inconnue ↓	PaD	Lois connues ↓
PaD	Probabilités sur la population entière	PaD	Probabilités sur la population entière	SD	Simulations (manuel) ↓	SD	Simulations ↓
				PaD	Loi	PaD	Autre loi
	Hyperbole		*Indice*		*Math'x*		*Transmath* *Repères*

Table 5.7 – Catégories de démarches dans les *activités* s'appuyant sur la notion d'histogramme

La catégorie 2.1 semble artificielle du point de vue de la problématique, les auteurs simulent des données suivant une certaine loi que les élèves doivent ensuite retrouver. La catégorie 1.2, quant à elle, a tendance à confondre (pseudo) réalité et modèle, ce qui pose un problème pour introduire la fonction de densité. Mis à part ces deux catégories, les deux autres semblent avoir des potentialités pour amener la notion de fonction de densité. Cependant, la démarche globale ne suffit pas pour que l'objectif soit pleinement atteint.

Dans les *activités* de la catégorie 1 analysées, les données « réelles » sont en fait des données « fabriquées » par simulation de la loi que l'on cherche, ce qui explique le fait que les histogrammes soient si « parfaits » et que l'histogramme soit très proche de la courbe de densité. Cela rend l'approximation par la courbe de densité bien entendu plus simple car elle est en fait sous-jacente à la simulation dont les données sont issues. Cela empêche (ou tout du moins limite) un questionnement sur le choix d'un modèle.

b) Trois domaines mais des dynamiques différentes

Du fait de la catégorie en jeu dans les manuels, mais pas seulement, nous avons pu remarquer des dynamiques et des articulations différentes entre les trois ETM pour gérer cette introduction de la notion de fonction de densité. Par exemple, dans le manuel *Transmath*, l'énoncé impose l'égalité entre probabilité et intégrale qui fait que l'ETM_{PaD} et l'ETM_{CI} forment un seul bloc. Pour vérifier cette unification, l'ETM_{SD} entre en jeu permettant de comparer les valeurs trouvées. Dans le manuel *Indice*, au contraire, ce sont l'ETM_{SD} et l'ETM_{PaD} qui sont assimilés l'un à l'autre et ensuite il y a une comparaison des valeurs entre les deux blocs PaD/SD d'un côté et CI de l'autre.

c) La dimension sémiotique très favorisée

Ce qui ressort de toutes les *activités*, c'est une très forte mobilisation de la dimension sémiotique, notamment à travers la visualisation (histogramme/courbe).

Nous remarquons que la dimension instrumentale, que l'on trouve dans l'utilisation de logiciels, n'est en réalité que peu mobilisée : les aides à l'utilisation rendent le travail purement technique. De plus, les avantages du fait de pouvoir rapidement bénéficier de nouvelles données et donc de nouveaux histogrammes ne sont pratiquement pas exploités (dans la catégorie 2).

La dimension discursive est très peu présente. Les liens entre fréquence et histogramme sont dans la majorité des cas rappelés dans l'énoncé. Ces connaissances ne sont pas supposées disponibles pour l'élève. C'est simplement la technique qui importe ici. Pour ce qui est de la loi des grands nombres (non connue en ces termes par les élèves), elle n'est jamais véritablement attendue, c'est un théorème utilisé en acte. La seule propriété supposée disponible est celle sur le lien entre aire et intégrale (pas utile dans toutes les *activités*), qui est une connaissance récente.

Ce que nous pouvons dire, c'est que le référentiel théorique visé par ces *activités* ne semble finalement pas construit (ou que partiellement). Dans tous les cas (excepté pour le manuel *Transmath*), la courbe est donnée dans le manuel. Cela ne permet pas de véritablement se questionner sur cette fonction de densité et les propriétés qu'elle doit vérifier. L'absence de réflexion sur le choix d'un modèle crée cette idée de fonction unique, sans savoir pour autant ce que cette fonction doit vérifier, mis à part qu'elle « *lisse* » ou « *épouse* » l'histogramme. Nous avons pu mettre en avant que dans aucune des activités la propriété d'aire sous la courbe égale à 1 ne ressort véritablement comme une propriété nécessaire de la courbe de densité.

d) Un appui sur une notion qui semble mal connue

Sur les cinq manuels, trois rappellent des éléments sur l'histogramme, sur le lien entre fréquence et aire des rectangles. On pourrait penser que l'histogramme, étant étudié plusieurs années entre la classe de cinquième et la classe de seconde, qu'il fait partie intégrante de l'ETM personnel des élèves et donc qu'il soit disponible chez les élèves. Cependant, les remarques que nous avons faites sur les manuels montrent déjà que les auteurs des manuels eux-mêmes, experts en mathématiques, introduisent des erreurs sur cet objet mathématique : nom de l'axe des ordonnées erroné, confusion entre fréquence et densité, confusion entre diagramme en bâtons et histogramme. Toutes ces remarques contredisent le fait que l'histogramme soit une connaissance disponible dans l'ETM personnel des élèves, vu qu'il ne l'est même pas dans les ETM idoines potentiels proposés par les manuels. Nous avons mené une étude complémentaire sur l'histogramme dans les manuels de seconde (Derouet & Parzysz, 2016a), dernière année du secondaire durant laquelle l'histogramme est explicitement au programme, pour essayer d'apprécier la raison pour laquelle de telles erreurs se trouvent dans les manuels de terminale. La section 5.2.5 présente les résultats de ce travail.

Les représentations erronées des histogrammes ajoutent un obstacle au changement de sous-domaine. De plus, nous avons remarqué que, dans aucun cas, le mot « densité » n'est justifié : cela semble être seulement un nom donné à la fonction

sans raison particulière. Alors qu'avec une bonne représentation et compréhension de l'histogramme, cette terminologie pourrait être justifiée : l'axe des ordonnées d'un histogramme n'est autre que la densité de fréquences.

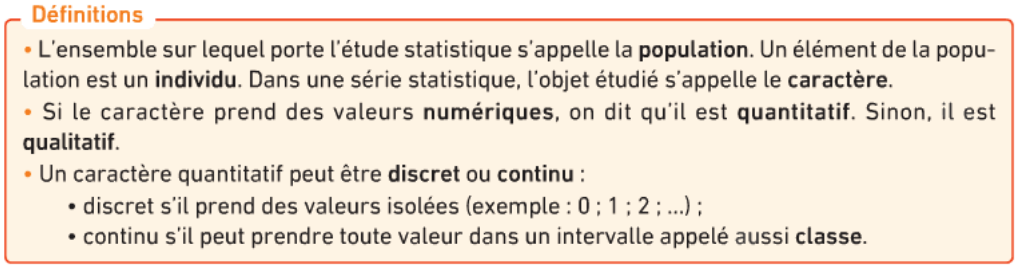
5.2.5 Compléments sur la notion d'histogramme

L'analyse des *activités* d'introduction de la notion de fonction de densité nous a permis de remarquer la présence de beaucoup d'erreurs sur la notion d'histogramme, qui n'aide pas à la compréhension de la notion de fonction de densité. Cette présence d'erreurs dans les manuels nous a poussée à aller voir du côté de l'enseignement de cet objet « histogramme ».

L'histogramme figure explicitement au programme du collège, à partir de la classe de cinquième, et le document *Ressources* de collège signale que « *les programmes précisent que les exemples étudiés se limitent au cas de classes d'égale amplitude* », et que « *l'histogramme se lit alors comme un diagramme en bâtons* » (MEN & DGESCO, 2007, p. 6). Au lycée, le document *Ressources* de seconde rappelle que « *Histogrammes, représentations graphiques,* [...] *sont autant d'outils que les élèves pourront utiliser pour faire émerger des questions sur les données, dégager des problématiques d'étude ou résumer l'information essentielle...* » (MEN & DGESCO, 2009, p. 4). Le même document mentionne « *l'aide incontournable de l'outil informatique* » (p. 3). On voit donc que l'histogramme est introduit très tôt dans l'enseignement secondaire, mais sous une forme particulière qui ne le différencie pas du diagramme en bâtons (accolés), et que le seul travail demandé à l'élève est, dans le meilleur des cas, de savoir en extraire de l'information numérique. Ce type de graphique n'est donc pas particulièrement valorisé.

Pour approcher l'espace de travail personnel des élèves de terminale scientifique, nous nous sommes intéressée à leurs connaissances supposées disponibles relativement à la notion d'histogramme, qui figure au programme de la classe de seconde : « *Représenter une série statistique graphiquement (nuage de points, histogramme, courbe des fréquences cumulées)* » (MEN & DGESCO, 2009). Nous avons pour cela entrepris de consulter onze manuels, édités entre 2009 et 2014, correspondants au programme actuel (entré en application en 2009), dont la liste se trouve dans la rubrique dédiée de la bibliographie.

La première question qui se pose est celle du type de variables statistiques associées à l'histogramme : il s'agit des variables quantitatives dont les valeurs, trop nombreuses pour être représentées individuellement dans un graphique, sont regroupées en classes contiguës qui constituent une partition d'un intervalle fermé borné de $\mathbb{R}$. Il est question alors de représenter les effectifs (ou les fréquences) de ces classes. Cependant, nous constatons que la question du regroupement en classes, ainsi que celle du choix de l'amplitude et des bornes de ces classes ne sont que rarement (2 cas) prises en compte (cf. figure 5.21).


Définitions

- L'ensemble sur lequel porte l'étude statistique s'appelle la **population**. Un élément de la population est un **individu**. Dans une série statistique, l'objet étudié s'appelle le **caractère**.
- Si le caractère prend des valeurs **numériques**, on dit qu'il est **quantitatif**. Sinon, il est **qualitatif**.
- Un caractère quantitatif peut être **discret** ou **continu** :
  - discret s'il prend des valeurs isolées (exemple : 0 ; 1 ; 2 ; ...) ;
  - continu s'il peut prendre toute valeur dans un intervalle appelé aussi **classe**.

FIGURE 5.21 – Extrait du cours du manuel *Mathématiques 2de* p. 158

Pourtant, une question cruciale pour le graphique est liée au choix de ces limites : se limite-t-on à des classes de même amplitude, ou traite-t-on le cas des classes d'amplitudes différentes ? Si l'on veut faire apparaître la notion de densité, indispensable pour l'étude des variables aléatoires à densité, on se doit de considérer le cas général. Notre étude fait apparaître que, si la majorité des manuels signalent les histogrammes à classes d'amplitudes différentes, ils se contentent le plus souvent de donner la définition générale, ou d'en traiter un exemple en exercice. Dans la plupart des cas, le cas des classes d'amplitudes différentes apparaît de fait comme un cas particulier du cas des classes de même amplitude, plutôt que l'inverse. De la même façon, les conversions (données → tableau → graphique) sont souvent absentes. En ce qui concerne la définition de l'histogramme, le plus souvent (6 cas) on indique que « *pour chaque classe, l'aire de chaque rectangle est proportionnelle à l'effectif, ou à la fréquence, de la classe* » (*Mathématiques 2de*, p. 162), mais on trouve aussi (2 cas) que « *l'effectif de chaque classe est proportionnel à l'aire du rectangle correspondant* » (*Travailler en confiance*, p. 233). C'est-à-dire que les premiers donnent une définition procédurale dans le sens de l'« écriture », permettant de construire le graphique à partir du tableau, tandis que les seconds donnent une grille de « lecture » du graphique permettant de reconstituer le tableau. Mais, en fait, très peu d'exercices proposent aux élèves de travailler réellement le passage d'un registre à l'autre. La nature même du graphique pose parfois problème elle aussi : un manuel donne la définition suivante : « *l'aire des rectangles est proportionnelle à l'effectif ou à la fréquence de la valeur* » (*Odyssée 2de*, p. 265). Ceci correspond en fait à la définition d'un diagramme en bâtons, c'est-à-dire un « histogramme » au sens du tableur Excel (qui n'en est pas un comme nous l'avons déjà signalé), et l'illustration du manuel le confirme (cf. figure 5.22).

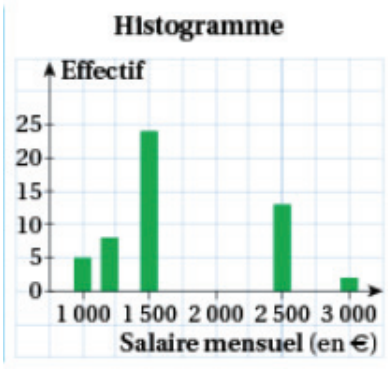


Figure 5.22 – Manuel *Odyssée 2de*, p. 265

Le risque de confusion entre histogramme et diagramme en bâtons existe réellement, puisqu'un manuel, au lieu de graduer, comme il est de règle, l'axe des abscisses comme l'axe des réels, dénomme parfois les classes par le symbole des intervalles correspondants (comme on le fait pour les modalités d'un caractère qualitatif (cf. figure 5.23)). Nous avons déjà remarqué cette erreur dans des manuels de terminale S (cf. figure 5.13, p. 152).

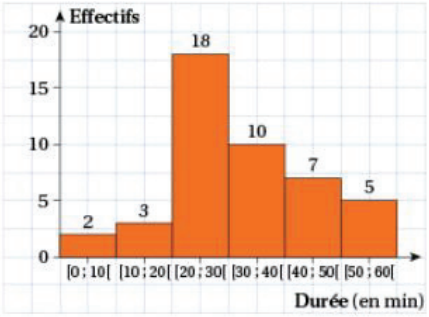


Figure 5.23 – Exemple d'erreur de graduation de l'axe des abscisses (*Odyssée 2de*, p. 281)

La question de l'axe des ordonnées est aussi un élément important de la démarche car il correspond justement à la densité de fréquence, notion cruciale dans le passage aux probabilités. La façon dont est traité cet axe dans les manuels témoigne d'un malaise chez les auteurs et certainement chez les enseignants aussi. Ainsi, dans les manuels observés on trouve à la fois des histogrammes à pas constant, avec un axe vertical ou sans, et des histogrammes à pas variable, avec ou sans axe vertical (cf. table 5.8).

	Pas constant	**Pas variable**
Avec axe vertical	*Déclic 2de* *Hyperbole 2de* *Indice 2de* *Math'x 2de* *Odyssée 2de* (cf. figures 5.22 et 5.23) *Repères 2de* *Travailler en confiance* (cf. figure 5.24)	*πxel* *Symbole 2de* (cf. figure 5.24) *Travailler en confiance*
Sans axe vertical	*Mathématiques 2de* *Transmath 2de*	*Mathématiques 2de* (cf. figure 5.25) *Math'x 2de* *Travailler en confiance*

TABLE 5.8 – Les différentes approches de la notion de fonction de densité dans les manuels de terminale S

Dans le cas où l'amplitude des classes est constante, il n'est pas incorrect de faire figurer un axe des ordonnées où sont portées les fréquences, car dans ce cas particulier la fréquence est égale à la densité (à coefficient près). Cependant, dans le cas contraire, c'est impérativement la densité de fréquence qu'il faudrait indiquer sur cet axe. Or, le mot « densité » n'apparaît jamais, même lorsque le quotient de la fréquence par l'amplitude de la classe est explicité sous la forme d'une formule du type $\frac{f_i}{b_i - a_i}$ (*Symbole 2de*, *Hyperbole 2de*). On trouve cependant (cf. figure 5.24) « *fréquence (en % par unité)* », et on peut, en comparant avec le manuel *Symbole* qui dénomme cet axe « *effectif par unité* » , évoquer une erreur de frappe ayant malencontreusement déplacé une parenthèse alors qu'on voulait écrire « *fréquence (en %) par unité* ». Mais parler de fréquence, ou d'effectif, « *par unité* » ne fait certainement pas sens pour les élèves.

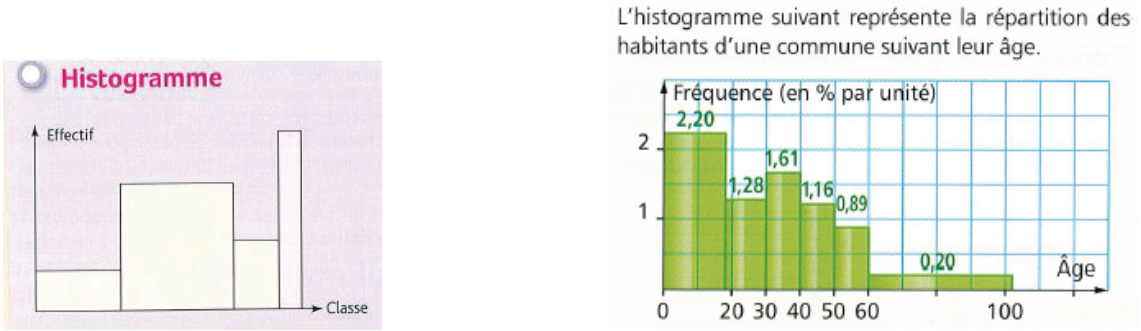


FIGURE 5.24 – Exemples d'erreurs de nom de l'axe des ordonnées (*Travailler en confiance*, p. 233 (à gauche) et *Symbole 2de*, p. 153 (à droite)

Cette difficulté explique sans doute pourquoi un certain nombre de manuels

préfèrent ne pas aborder la question de l'axe vertical et définir un carré dont l'aire correspond à une fréquence ou un effectif donné.e (cf. figure 5.25), comme le manuel de terminale S *Transmath*.

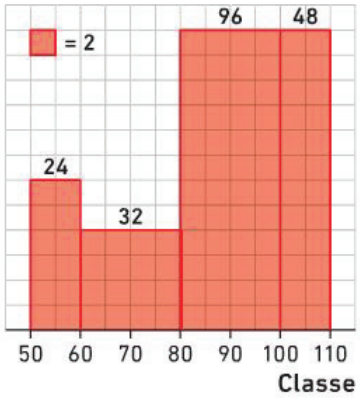


Figure 5.25 – Exemple de manuel sans axe des ordonnées, avec une unité d'aire (*Mathématiques 2de*, p. 162)

Enfin, sans doute pour faciliter l'identification de l'effectif des classes, la plupart des manuels (7 sur 11), superposent aux rectangles l'indication de la fréquence ou de l'effectif correspondant (cf. figures 5.24 et 5.25). Une autre grande absente de cette recension est la raison d'être de cette procédure de quotient, qui peut paraître *a priori* étrange aux élèves mais qui est absolument transparente pour les auteurs. On peut retenir de cette étude que la proportionnalité entre aires et fréquences est généralement mentionnée, mais non justifiée, et la signification de la dimension verticale (densité) fait totalement défaut. Faire travailler les élèves sur les conversions de registres tableau de données $\longleftrightarrow$ graphique n'est pris en compte que de façon marginale.

En définitive, les manuels de seconde ne s'intéressent pas vraiment à la notion d'histogramme. Trois raisons au moins sont susceptibles d'expliquer cette situation. La première est que, si l'on se limite aux classes d'égale amplitude, cette représentation fait en quelque sorte double emploi avec le diagramme en bâtons, plus facile à mettre en œuvre. La deuxième est que, si l'on veut traiter le cas des classes d'amplitudes variées, on va avoir à opérer un traitement de proportionnalité sur les données (technique que certains élèves maîtrisent encore mal à ce niveau). Enfin, la troisième est que les tableurs-grapheurs les plus courants ne le fournissent pas directement. Nous retrouvons des résultats similaires à une étude de Régnier (1998).

L'histogramme apparaît donc comme un objet marginal, qui n'a d'autre utilité qu'une fonction de représentation de certaines séries (quant à savoir lesquelles, c'est loin d'être toujours clair). Rien ne laisse présager, à ce niveau, qu'il puisse être réinvesti ultérieurement en probabilités dans le but de faciliter l'introduction aux variables aléatoires continues. Une étude récente (Roditi, 2009) fait apparaître que les professeurs, au lycée comme au collège, ne consacrent que très peu de temps à l'histogramme (environ une heure par an), et que « *près de la moitié* [d'entre eux] *estiment que l'enseignement de l'histogramme ne contribue pas à la formation*

mathématique des élèves » (p. 131).

Alors que nous estimons que l'histogramme, avec la notion de densité de fréquence, est un point clé pour aborder les lois à densité, l'étude montre que les ETM idoines proposés dans le manuels de seconde sont très pauvres et généralement mathématiquement faux (ou tout du moins mal présentés). Au vu des manuels de seconde, il est compréhensible de retrouver des erreurs dans les manuels de terminale S.

Ceci renforce l'idée que la notion d'histogramme ne fait vraisemblablement pas partie de l'ETM personnel des élèves et qu'il serait donc bon de ne pas considérer cet objet mathématique comme disponible pour les élèves (d'ailleurs cela explique que les manuels fassent des rappels). Bien au contraire, la notion doit être retravaillée et, notamment, l'axe des ordonnées doit être explicité pour permettre ensuite une meilleure compréhension de la fonction de densité de probabilité. Si l'on est amené à l'utiliser en terminale, il semble donc incontournable d'organiser une nouvelle « première rencontre ».

5.2.6 Conclusion

Nous pouvons conclure en disant que les manuels de terminale S proposent des *activités* d'introduction qui, cependant, manquent de richesse et sont parfois inadéquats mathématiquement. Notamment, ces *activités* ne permettent pas réellement la construction du référentiel théorique sur la notion de fonction de densité.

Dans les *activités* proposées, beaucoup de passages semblent aller de soi alors qu'ils cachent de véritables problèmes. Par exemple, le travail dans la dimension sémiotique sur l'objet histogramme, semble être considéré comme transparent pour les auteurs, alors que cette notion peut vraiment être problématique pour les élèves.

Alors que les enseignants ont besoin de support pour l'enseignement de ce nouveau savoir, les manuels ne fournissent finalement peut-être pas un support assez riche. Ce résultat va dans le sens de Ortiz, Cañizares, Batanero et Serrano (2002) qui ont déjà mis en évidence que les manuels et les documents institutionnels n'offrent pas un support assez important aux enseignants du secondaire. De plus, une autre étude (Ortiz, 1999) a montré que les manuels peuvent comporter des erreurs conceptuelles et une approche didactique inadéquate des concepts probabilistes (et statistiques), comme nous avons pu le voir.

5.3 Des connexions au niveau du calcul d'aires

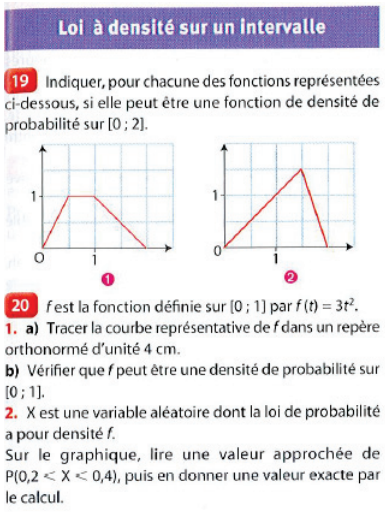
Nous nous sommes, dans la section précédente, intéressée plus particulièrement à l'introduction de la notion de fonction de densité dans les manuels. Maintenant, nous allons essayer d'identifier des tâches, dans les manuels, qui exploitent les liens entre probabilités à densité et calcul intégral.

Tout d'abord, comme dans les documents institutionnels, nous pouvons observer que peu de tâches exploitent réellement la relation entre probabilités à densité et calcul intégral. Nous retrouvons des tâches assez similaires à celles du baccalauréat. Nous ne reviendrons pas dessus. Nous allons ici mettre en évidence un point

de connexion qui existe entre les deux sous-domaines, au sujet des calculs d'aires rencontrés.

5.3.1 Mise en évidence de différents niveaux de calculs d'aires

Dans les manuels, le choix est fait de ne pas se cantonner à l'étude des trois lois au programme, mais de commencer par une partie plus générale. Excepté le manuel *Math'x*, tous les autres proposent des tâches en début de chapitre (sur les lois à densité) avec des fonctions de densité différentes des trois au programme. La figure 5.26 est un exemple des tâches que l'on peut trouver. Nous tenons tout de même à préciser que ces tâches sont peu nombreuses et que très vite, les exercices proposés se focalisent sur les lois uniformes, exponentielles et normales.



Loi à densité sur un intervalle

19 Indiquer, pour chacune des fonctions représentées ci-dessous, si elle peut être une fonction de densité de probabilité sur [0 ; 2].

❶ ❷

20 f est la fonction définie sur [0 ; 1] par $f(t) = 3t^2$.

1. a) Tracer la courbe représentative de f dans un repère orthonormé d'unité 4 cm.

b) Vérifier que f peut être une densité de probabilité sur [0 ; 1].

2. X est une variable aléatoire dont la loi de probabilité a pour densité f.

Sur le graphique, lire une valeur approchée de $P(0{,}2 < X < 0{,}4)$, puis en donner une valeur exacte par le calcul.

Figure 5.26 – Exercice résolu 1 p. 379 du manuel *Hyperbole*

Les différentes fonctions de densité rencontrées sont toujours définies sur un intervalle borné, jamais sur un intervalle non borné. Les tâches proposées demandent des calculs d'aires sous une courbe, comme l'on peut retrouver en calcul intégral. En regardant de plus près, il apparait que les différents calculs d'aires qui apparaissent sont similaires à ceux rencontrés en calcul intégral.

Nous pouvons classer les calculs d'aires en trois niveaux :

- Les calculs d'aires élémentaires (niveau A1). Ces types d'exemples peuvent être repérés au tout début du chapitre sur le calcul intégral avec l'intégrale de fonctions affines par morceaux où il s'agit alors de calculer des aires de figures élémentaires (rectangles, triangles, trapèzes).
- Les calculs d'aires par « primitivation » (niveau A2). Il s'agit des exemples où la connaissance des primitives est indispensable, pour déterminer l'intégrale (ou l'aire).
- Les calculs approchés d'aires (niveau A3). Il s'agit des calculs approchés d'aires, dans les cas où aucune primitive n'est connue (tout du moins au moment de la résolution de l'exercice).

Niveau	
A1	Calcul d'aires élémentaires
A2	Calcul d'aires par « primitivation »
A3	Calcul approché d'aires

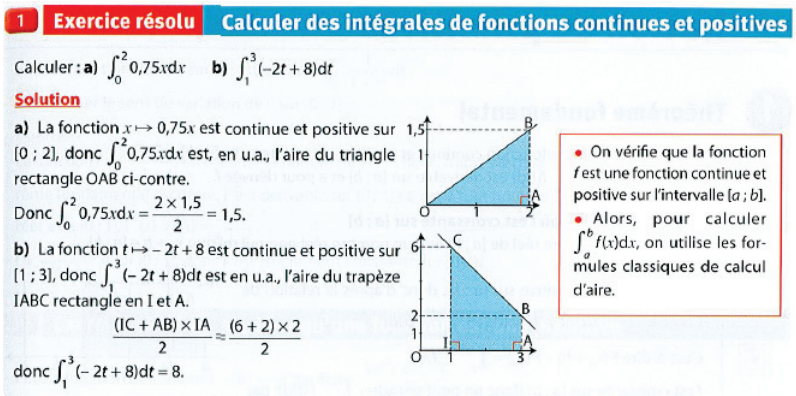
TABLE 5.9 – Les trois niveaux de calcul d'aire

Bien entendu les tâches relevant des probabilités à densité et celles du calcul intégral ne sont pas identiques. Les tâches sont données dans des registres différents : le registre symbolique probabiliste pour les probabilités et le registre symbolique de l'intégrale pour le calcul intégral. Cependant, après une adaptation, qui est d'opérer un changement de domaine des probabilités à l'analyse accompagné d'un changement de registres, les activités attendues pourraient *a priori* être les mêmes.

La question est alors : restent-elles vraiment les mêmes ? Le changement de domaines n'est pas à négliger, mais ces tâches ont, à première vue, un potentiel pour relier effectivement les deux sous-domaines. Pour avoir une idée des activités attendues, nous nous sommes attardée sur les exercices corrigés des manuels.

5.3.2 Le niveau A1

Nous rappelons que le niveau A1 désigne des calculs d'aires élémentaires, où seules les formules d'aires telles que celles du triangle, du rectangle ou du trapèze sont nécessaires. Pour illustrer, nous proposons ici deux exemples, un premier extrait du chapitre sur le calcul intégral (cf. figure 5.27), l'autre du chapitre sur les lois à densité (cf. figure 5.28), du manuel *Hyperbole*. À travers ces exercices corrigés, proposant des études avec des lois autres que celles du programme, on peut effectivement voir que les activités attendues des élèves, après un changement de domaines, sont bien les mêmes en probabilités à densité et en calcul intégral : il s'agit, comme il est écrit dans l'encadré de la figure 5.27, d'utiliser les formules classiques de calcul d'aire (de rectangles, triangles...).


1 **Exercice résolu** **Calculer des intégrales de fonctions continues et positives**

Calculer : **a)** $\int_0^2 0{,}75x\,dx$ **b)** $\int_1^3 (-2t+8)\,dt$

Solution

a) La fonction $x \mapsto 0{,}75x$ est continue et positive sur [0 ; 2], donc $\int_0^2 0{,}75x\,dx$ est, en u.a., l'aire du triangle rectangle OAB ci-contre.

Donc $\int_0^2 0{,}75x\,dx = \frac{2 \times 1{,}5}{2} = 1{,}5$.

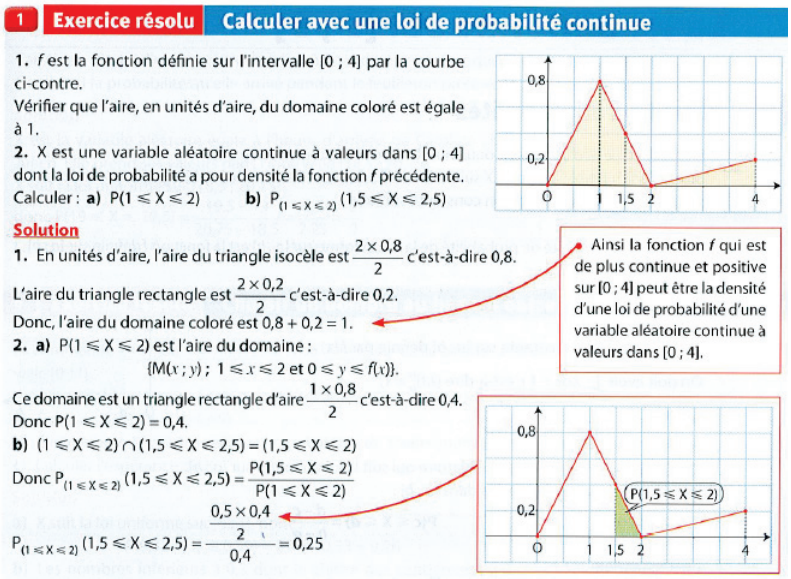
b) La fonction $t \mapsto -2t+8$ est continue et positive sur [1 ; 3], donc $\int_1^3 (-2t+8)\,dt$ est en u.a., l'aire du trapèze IABC rectangle en I et A.

$\frac{(IC + AB) \times IA}{2} = \frac{(6+2) \times 2}{2}$

donc $\int_1^3 (-2t+8)\,dt = 8$.

• On vérifie que la fonction f est une fonction continue et positive sur l'intervalle $[a\,;b]$.
• Alors, pour calculer $\int_a^b f(x)dx$, on utilise les formules classiques de calcul d'aire.

FIGURE 5.27 – Exercice résolu 1 p. 201 du manuel *Hyperbole*



1 Exercice résolu — Calculer avec une loi de probabilité continue

1. f est la fonction définie sur l'intervalle [0 ; 4] par la courbe ci-contre.
Vérifier que l'aire, en unités d'aire, du domaine coloré est égale à 1.
2. X est une variable aléatoire continue à valeurs dans [0 ; 4] dont la loi de probabilité a pour densité la fonction f précédente.
Calculer : **a)** $P(1 \leqslant X \leqslant 2)$ **b)** $P_{(1 \leqslant X \leqslant 2)}(1{,}5 \leqslant X \leqslant 2{,}5)$

Solution

1. En unités d'aire, l'aire du triangle isocèle est $\frac{2 \times 0{,}8}{2}$ c'est-à-dire 0,8.
L'aire du triangle rectangle est $\frac{2 \times 0{,}2}{2}$ c'est-à-dire 0,2.
Donc, l'aire du domaine coloré est 0,8 + 0,2 = 1.

Ainsi la fonction f qui est de plus continue et positive sur [0 ; 4] peut être la densité d'une loi de probabilité d'une variable aléatoire continue à valeurs dans [0 ; 4].

2. a) $P(1 \leqslant X \leqslant 2)$ est l'aire du domaine :
$\{M(x\,;y)\,;\ 1 \leqslant x \leqslant 2 \text{ et } 0 \leqslant y \leqslant f(x)\}$.
Ce domaine est un triangle rectangle d'aire $\frac{1 \times 0{,}8}{2}$ c'est-à-dire 0,4.
Donc $P(1 \leqslant X \leqslant 2) = 0{,}4$.
b) $(1 \leqslant X \leqslant 2) \cap (1{,}5 \leqslant X \leqslant 2{,}5) = (1{,}5 \leqslant X \leqslant 2)$
Donc $P_{(1 \leqslant X \leqslant 2)}(1{,}5 \leqslant X \leqslant 2{,}5) = \frac{P(1{,}5 \leqslant X \leqslant 2)}{P(1 \leqslant X \leqslant 2)}$

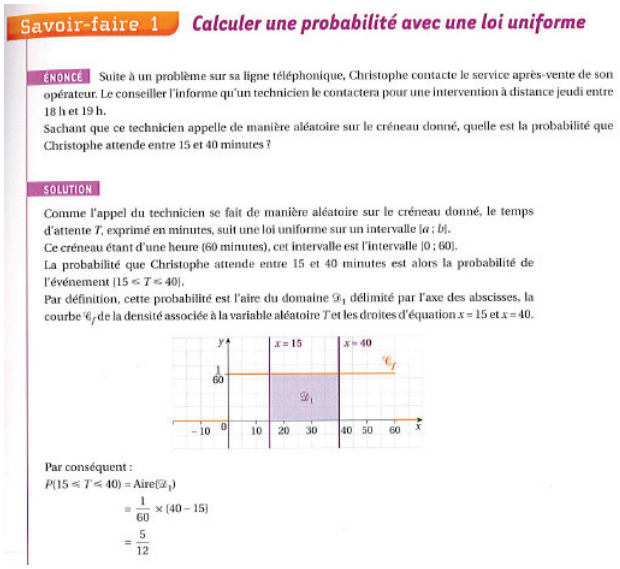
$$P_{(1 \leqslant X \leqslant 2)}(1{,}5 \leqslant X \leqslant 2{,}5) = \frac{\frac{0{,}5 \times 0{,}4}{2}}{0{,}4} = 0{,}25$$

Figure 5.28 – Exercice résolu 1 p. 379 du manuel *Hyperbole*

Bien entendu les registres utilisés dans un domaine ou dans l'autre ne sont pas les mêmes au départ, cependant les techniques à utiliser ensuite sont identiques. D'un côté, il s'agit du registre symbolique probabiliste, de l'autre, du registre symbolique de l'intégrale, mais dans tous les cas, nous sommes ramenés à un calcul d'aires élémentaires.

Dans ce niveau, nous retrouvons aussi le cas particulier des lois uniformes. Dans ce cas, il s'agit alors du calcul d'aires de rectangles. Suivant les manuels, on peut cependant trouver des méthodes différentes pour déterminer les probabilités :

– Soit, effectivement, par le calcul d'aire de rectangles (figure 5.29),



Savoir-faire 1 — *Calculer une probabilité avec une loi uniforme*

ÉNONCÉ Suite à un problème sur sa ligne téléphonique, Christophe contacte le service après-vente de son opérateur. Le conseiller l'informe qu'un technicien le contactera pour une intervention à distance jeudi entre 18 h et 19 h.
Sachant que ce technicien appelle de manière aléatoire sur le créneau donné, quelle est la probabilité que Christophe attende entre 15 et 40 minutes ?

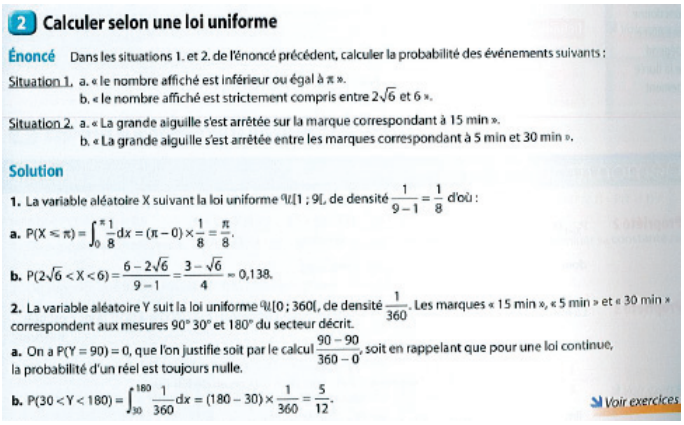
SOLUTION

Comme l'appel du technicien se fait de manière aléatoire sur le créneau donné, le temps d'attente T, exprimé en minutes, suit une loi uniforme sur un intervalle $[a\,;b]$.
Ce créneau étant d'une heure (60 minutes), cet intervalle est l'intervalle [0 ; 60].
La probabilité que Christophe attende entre 15 et 40 minutes est alors la probabilité de l'événement $[15 \leqslant T \leqslant 40]$.
Par définition, cette probabilité est l'aire du domaine $\mathscr{D}_1$ délimité par l'axe des abscisses, la courbe $\mathscr{C}_f$ de la densité associée à la variable aléatoire T et les droites d'équation $x = 15$ et $x = 40$.

Par conséquent :
$P(15 \leqslant T \leqslant 40) = \text{Aire}(\mathscr{D}_1)$
$= \frac{1}{60} \times (40 - 15)$
$= \frac{5}{12}$

Figure 5.29 – *Savoir-faire* 1 p. 405 du manuel *Odyssée*

– Soit en calculant une intégrale par « primitivation » (figure 5.30).



2 Calculer selon une loi uniforme

Énoncé Dans les situations 1. et 2. de l'énoncé précédent, calculer la probabilité des événements suivants :

Situation 1. a. « le nombre affiché est inférieur ou égal à π ».
b. « le nombre affiché est strictement compris entre $2\sqrt{6}$ et 6 ».

Situation 2. a. « La grande aiguille s'est arrêtée sur la marque correspondant à 15 min ».
b. « La grande aiguille s'est arrêtée entre les marques correspondant à 5 min et 30 min ».

Solution

1. La variable aléatoire X suivant la loi uniforme $\mathcal{U}[1\,;9[$, de densité $\frac{1}{9-1}=\frac{1}{8}$ d'où :

a. $P(X \leqslant \pi)=\int_0^{\pi}\frac{1}{8}dx=(\pi-0)\times\frac{1}{8}=\frac{\pi}{8}$.

b. $P(2\sqrt{6}<X<6)=\frac{6-2\sqrt{6}}{9-1}=\frac{3-\sqrt{6}}{4}\approx 0{,}138$.

2. La variable aléatoire Y suit la loi uniforme $\mathcal{U}[0\,;360[$, de densité $\frac{1}{360}$. Les marques « 15 min », « 5 min » et « 30 min » correspondent aux mesures 90° 30° et 180° du secteur décrit.

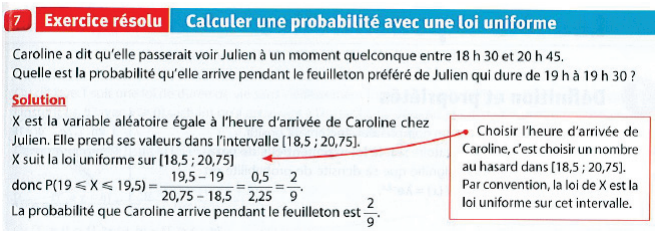
a. On a $P(Y=90)=0$, que l'on justifie soit par le calcul $\frac{90-90}{360-0}$, soit en rappelant que pour une loi continue, la probabilité d'un réel est toujours nulle.

b. $P(30<Y<180)=\int_{30}^{180}\frac{1}{360}dx=(180-30)\times\frac{1}{360}=\frac{5}{12}$.

Voir exercices

FIGURE 5.30 – *Exercice résolu* 1 p. 403 du manuel *Math'x*

– Soit, encore, en appliquant directement la formule du cours $\dfrac{d-c}{b-a}$ (figure 5.31).



7 Exercice résolu Calculer une probabilité avec une loi uniforme

Caroline a dit qu'elle passerait voir Julien à un moment quelconque entre 18 h 30 et 20 h 45.
Quelle est la probabilité qu'elle arrive pendant le feuilleton préféré de Julien qui dure de 19 h à 19 h 30 ?

Solution

X est la variable aléatoire égale à l'heure d'arrivée de Caroline chez Julien. Elle prend ses valeurs dans l'intervalle [18,5 ; 20,75].
X suit la loi uniforme sur [18,5 ; 20,75]
donc $P(19 \leqslant X \leqslant 19{,}5)=\frac{19{,}5-19}{20{,}75-18{,}5}=\frac{0{,}5}{2{,}25}=\frac{2}{9}$.
La probabilité que Caroline arrive pendant le feuilleton est $\frac{2}{9}$.

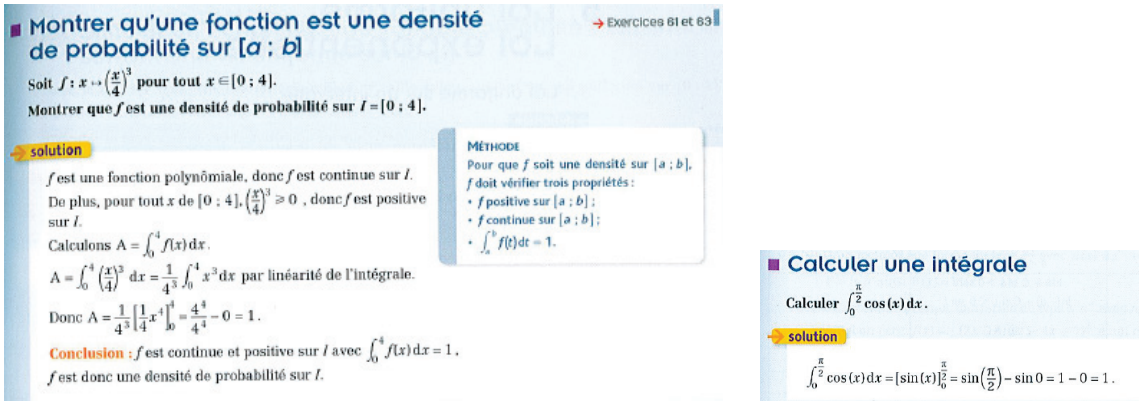
Choisir l'heure d'arrivée de Caroline, c'est choisir un nombre au hasard dans [18,5 ; 20,75]. Par convention, la loi de X est la loi uniforme sur cet intervalle.

FIGURE 5.31 – *Exercice résolu* 7 p. 381 du manuel *Hyperbole*

La deuxième méthode relève du niveau A2. Cependant, ce basculement vers le niveau A2 se retrouve aussi dans le chapitre sur le calcul intégral, une fois que les primitives sont connues. La dernière démarche quant à elle ne se retrouve pas en calcul intégral, car, justement dans ce chapitre c'est la démarche pour arriver au résultat qui importe et non le résultat. D'ailleurs, ce n'est jamais exactement les mêmes fonctions qui sont à intégrer donc aucune formule de ce type n'est connue.

5.3.3 Le niveau A2

Les calculs d'aires de niveau A2 correspondent aux calculs d'aires (et d'intégrales) par « primitivation ». Nous présentons ici deux exemples, le premier (à gauche) issu du chapitre sur les lois à densité et le second (à droite) issu du chapitre sur le calcul intégral du manuel *Repères* (cf. figure 5.32).



Montrer qu'une fonction est une densité de probabilité sur [a ; b] → Exercices 61 et 63

Soit $f : x \mapsto \left(\frac{x}{4}\right)^3$ pour tout $x \in [0 ; 4]$.

Montrer que f est une densité de probabilité sur $I = [0 ; 4]$.

solution

f est une fonction polynômiale, donc f est continue sur I.

De plus, pour tout x de $[0 ; 4]$, $\left(\frac{x}{4}\right)^3 \geqslant 0$, donc f est positive sur I.

Calculons $A = \int_0^4 f(x)\,dx$.

$A = \int_0^4 \left(\frac{x}{4}\right)^3 dx = \frac{1}{4^3}\int_0^4 x^3 dx$ par linéarité de l'intégrale.

Donc $A = \frac{1}{4^3}\left[\frac{1}{4}x^4\right]_0^4 = \frac{4^4}{4^4} - 0 = 1$.

Conclusion : f est continue et positive sur I avec $\int_0^4 f(x)\,dx = 1$, f est donc une densité de probabilité sur I.

MÉTHODE

Pour que f soit une densité sur $[a ; b]$, f doit vérifier trois propriétés :

- f positive sur $[a ; b]$;
- f continue sur $[a ; b]$;
- $\int_a^b f(t)\,dt = 1$.

Calculer une intégrale

Calculer $\int_0^{\frac{\pi}{2}} \cos(x)\,dx$.

solution

$\int_0^{\frac{\pi}{2}} \cos(x)\,dx = [\sin(x)]_0^{\frac{\pi}{2}} = \sin\left(\frac{\pi}{2}\right) - \sin 0 = 1 - 0 = 1$.

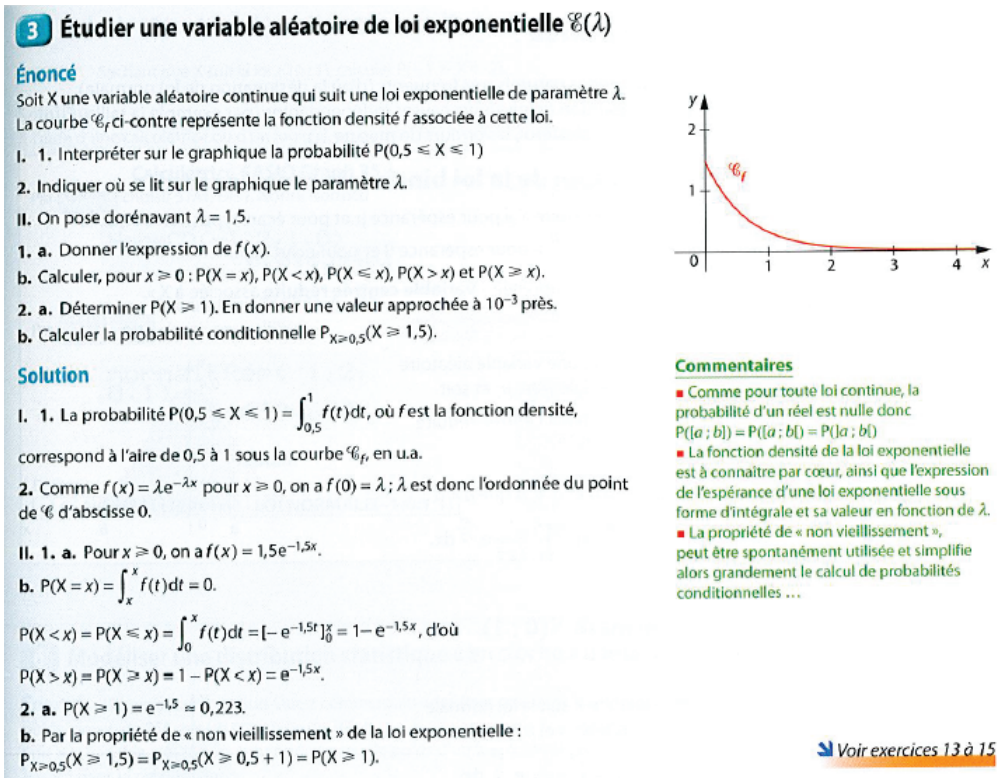
FIGURE 5.32 – Application p. 337 (calcul intégral) et Application p. 175 (lois à densité) du manuel *Repères*

Dans les deux, nous retrouvons le même type de démarche.

Dans la figure 5.32, la méthode encadrée précise que pour montrer qu'une fonction est une fonction de densité, il faut calculer une intégrale. Cela illustre bien la connexion entre les deux sous-domaines.

Pour les lois exponentielles aussi, nous sommes dans le cas où nous avons besoin de primitives pour déterminer les probabilités. Deux méthodes seulement sont envisageables dans ce cas :

– Le calcul de l'intégrale par « primitivation » (cf. figure 5.33) :



3 Étudier une variable aléatoire de loi exponentielle $\mathscr{E}(\lambda)$

Énoncé

Soit X une variable aléatoire continue qui suit une loi exponentielle de paramètre λ.

La courbe $\mathscr{C}_f$ ci-contre représente la fonction densité f associée à cette loi.

I. 1. Interpréter sur le graphique la probabilité $P(0{,}5 \leqslant X \leqslant 1)$

2. Indiquer où se lit sur le graphique le paramètre λ.

II. On pose dorénavant $\lambda = 1{,}5$.

1. a. Donner l'expression de $f(x)$.

b. Calculer, pour $x \geqslant 0$: $P(X = x)$, $P(X < x)$, $P(X \leqslant x)$, $P(X > x)$ et $P(X \geqslant x)$.

2. a. Déterminer $P(X \geqslant 1)$. En donner une valeur approchée à 10^{-3} près.

b. Calculer la probabilité conditionnelle $P_{X \geqslant 0,5}(X \geqslant 1{,}5)$.

Commentaires

- Comme pour toute loi continue, la probabilité d'un réel est nulle donc $P([a ; b]) = P([a ; b[) = P(]a ; b[)$
- La fonction densité de la loi exponentielle est à connaître par cœur, ainsi que l'expression de l'espérance d'une loi exponentielle sous forme d'intégrale et sa valeur en fonction de λ.
- La propriété de « non vieillissement », peut être spontanément utilisée et simplifie alors grandement le calcul de probabilités conditionnelles …

Solution

I. 1. La probabilité $P(0{,}5 \leqslant X \leqslant 1) = \int_{0,5}^{1} f(t)dt$, où f est la fonction densité, correspond à l'aire de 0,5 à 1 sous la courbe $\mathscr{C}_f$, en u.a.

2. Comme $f(x) = \lambda e^{-\lambda x}$ pour $x \geqslant 0$, on a $f(0) = \lambda$; λ est donc l'ordonnée du point de $\mathscr{C}$ d'abscisse 0.

II. 1. a. Pour $x \geqslant 0$, on a $f(x) = 1{,}5e^{-1,5x}$.

b. $P(X = x) = \int_x^x f(t)dt = 0$.

$P(X < x) = P(X \leqslant x) = \int_0^x f(t)dt = [-e^{-1,5t}]_0^x = 1 - e^{-1,5x}$, d'où

$P(X > x) = P(X \geqslant x) = 1 - P(X < x) = e^{-1,5x}$.

2. a. $P(X \geqslant 1) = e^{-1,5} \approx 0{,}223$.

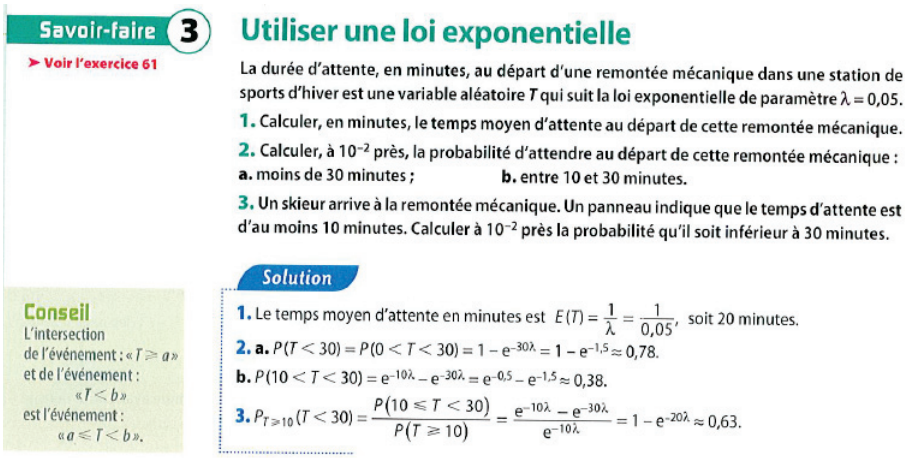
b. Par la propriété de « non vieillissement » de la loi exponentielle :

$P_{X \geqslant 0,5}(X \geqslant 1{,}5) = P_{X \geqslant 0,5}(X \geqslant 0{,}5 + 1) = P(X \geqslant 1)$.

Voir exercices 13 à 15

FIGURE 5.33 – Exercice résolu 3 p. 405 du manuel *Math'x*

– L'utilisation directe de la formule du cours (cf. figure 5.34) :



Savoir-faire 3 — **Utiliser une loi exponentielle**

➤ Voir l'exercice 61

La durée d'attente, en minutes, au départ d'une remontée mécanique dans une station de sports d'hiver est une variable aléatoire T qui suit la loi exponentielle de paramètre $\lambda = 0{,}05$.

1. Calculer, en minutes, le temps moyen d'attente au départ de cette remontée mécanique.

2. Calculer, à 10^{-2} près, la probabilité d'attendre au départ de cette remontée mécanique :
a. moins de 30 minutes ; **b.** entre 10 et 30 minutes.

3. Un skieur arrive à la remontée mécanique. Un panneau indique que le temps d'attente est d'au moins 10 minutes. Calculer à 10^{-2} près la probabilité qu'il soit inférieur à 30 minutes.

Conseil
L'intersection de l'événement : « $T \geqslant a$ » et de l'événement : « $T < b$ » est l'événement : « $a \leqslant T < b$ ».

Solution

1. Le temps moyen d'attente en minutes est $E(T) = \frac{1}{\lambda} = \frac{1}{0{,}05}$, soit 20 minutes.

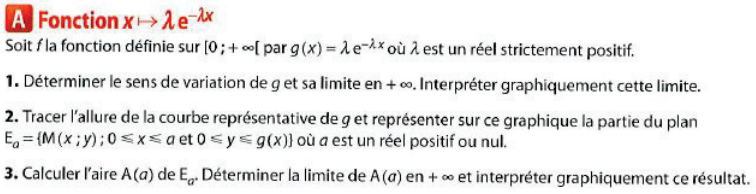
2. a. $P(T < 30) = P(0 < T < 30) = 1 - e^{-30\lambda} = 1 - e^{-1{,}5} \approx 0{,}78$.

b. $P(10 < T < 30) = e^{-10\lambda} - e^{-30\lambda} = e^{-0{,}5} - e^{-1{,}5} \approx 0{,}38$.

3. $P_{T \geqslant 10}(T < 30) = \frac{P(10 \leqslant T < 30)}{P(T \geqslant 10)} = \frac{e^{-10\lambda} - e^{-30\lambda}}{e^{-10\lambda}} = 1 - e^{-20\lambda} \approx 0{,}63$.

FIGURE 5.34 – *Savoir-faire* 3 p. 327 du manuel *Indice*

La loi exponentielle a, cependant, la particularité (par rapport aux autres fonctions rencontrées en calcul intégral ou aux autres exemples de fonctions de densité du niveau A2) d'être définie sur un intervalle non borné. À ce sujet, nous pouvons dire que, comme pour le baccalauréat, nous retrouvons quelques exercices sur des calculs de limites de fonctions ou de suites définies par une intégrale. Cependant, le fait qu'il est possible d'avoir une aire finie d'un domaine infini est peu questionné. Le manuel *Math'x* dans la partie *Travaux pratiques* p. 244 (cf. figure 5.35) du chapitre sur le calcul intégral est l'un des rares manuels à essayer de mettre en avant cette spécificité. Ce TP a pour objectif de « *rencontrer des calculs et des raisonnements sur les aires qui seront utilisés en probabilités* ».



A Fonction $x \mapsto \lambda e^{-\lambda x}$

Soit f la fonction définie sur $[0 ; +\infty[$ par $g(x) = \lambda e^{-\lambda x}$ où λ est un réel strictement positif.

1. Déterminer le sens de variation de g et sa limite en $+\infty$. Interpréter graphiquement cette limite.

2. Tracer l'allure de la courbe représentative de g et représenter sur ce graphique la partie du plan $E_a = \{M(x ; y) ; 0 \leqslant x \leqslant a \text{ et } 0 \leqslant y \leqslant g(x)\}$ où a est un réel positif ou nul.

3. Calculer l'aire $A(a)$ de E_a. Déterminer la limite de $A(a)$ en $+\infty$ et interpréter graphiquement ce résultat.

FIGURE 5.35 – Partie A du TP1 p. 244 du manuel *Math'x*

5.3.4 Le niveau A3

Le niveau A3, rappelons-le, se rapporte aux calculs approchés d'aires, qui sont présents dans le cas où aucune primitive explicite n'est connue (des élèves tout du moins) et donc aucun calcul exact n'est possible. Nous pouvons tout de même remarquer que même si les primitives sont connues, un calcul approché, par les méthodes que nous allons expliciter, reste tout de même possible.

En probabilités, le seul exemple présent dans les manuels relevant de ce niveau est celui des lois normales. Nous sommes bien dans le cas où nous ne connaissons pas de primitives explicites bien qu'elles existent. Nous voyons l'intérêt (conformément au programme) de parler, dans le chapitre sur le calcul intégral, de l'existence

de fonctions continues qui admettent des primitives mais que l'on ne connait pas explicitement (en réalité, fait que dans deux manuels).

Dans le cas de la loi normale, la technique préconisée pour déterminer des probabilités associées est l'utilisation de la calculatrice ou de logiciels, comme le présente clairement le manuel *Math'x* (cf. figure 5.36).

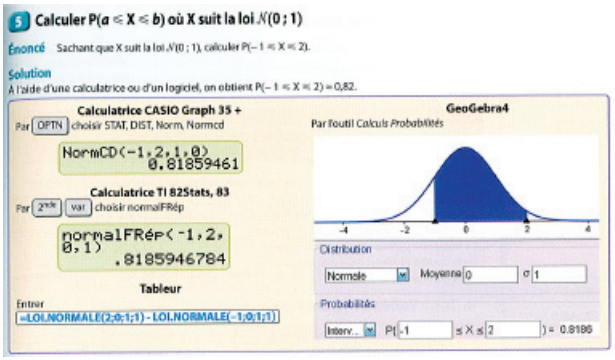


Figure 5.36 – Exercice résolu 5 p. 407 du manuel *Math'x*

Dans le chapitre dédié au calcul intégral, il est aussi possible d'utiliser des logiciels ou la calculatrice pour calculer des aires du niveau A3, mais ce ne sont pas les mêmes commandes qui sont utilisées (cf. figure 5.37).

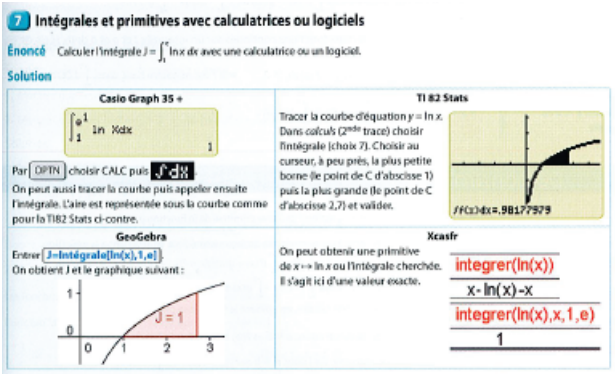


Figure 5.37 – Exercice résolu 7 p. 237 du manuel *Math'x*

Mis à part la calculatrice ou autre logiciel, la méthode la plus souvent utilisée en calcul intégral est la méthode des rectangles, et exceptionnellement la méthode des trapèzes.

On voit ici une rupture entre probabilités à densité et calcul intégral, dans les activités à mettre en œuvre dans le niveau A3. Cependant, pour nuancer cela, nous pouvons trouver dans différents manuels des tâches, en calcul intégral, sur la fonction $x \mapsto e^{-\frac{x^2}{2}}$. Par exemple, dans le manuel *Math'x* dans le même TP que cité au-dessus, la fonction $x \mapsto e^{-\frac{x^2}{2}}$ est étudiée à l'aide du logiciel *GeoGebra*. L'objectif est

de déterminer que $\lim_{t\to+\infty} \int_O^t e^{-\frac{-x^2}{2}}\,\mathrm{d}x = \sqrt{\frac{\pi}{2}}$ et de calculer des aires en jouant sur les symétries de la courbe (cf. Annexe E.3). Ici, nous pouvons voir un vrai travail pour créer du lien entre calcul intégral et probabilités à densité sur lequel le manuel pourrait s'appuyer lors du chapitre dédié aux lois à densité. Cependant, ce retour n'est pas fait dans le manuel. Cet exemple de tâche avec pour objectif explicite de rencontrer des notions qui serviront en probabilités à densité est une exception dans les manuels.

Nous remarquons que le calcul approché d'aire (niveau A3) n'est pas traité de la même manière dans les deux sous-domaines. Nous pouvons même voir ici que les outils technologiques à utiliser, tels que la calculatrice ou les logiciels, peuvent être les mêmes mais avec des commandes différentes s'il s'agit de probabilités ou d'analyse. Nous avons, dans la figure 5.38, représenté trois façons d'obtenir l'aire avec le logiciel *GeoGebra*. Les TICE, dans ce cas, sont responsables d'une rupture entre les deux sous-domaines.

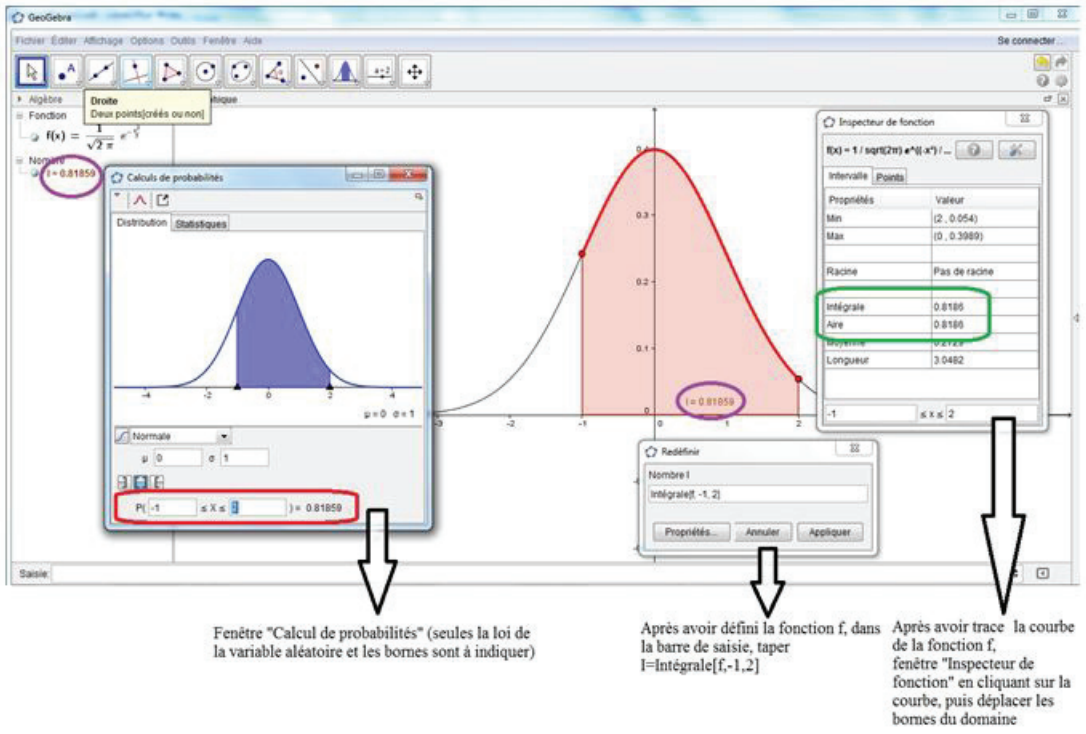


FIGURE 5.38 – Trois commandes pour calculer une probabilité ou une intégrale avec *GeoGebra*

5.3.5 Récapitulatif

Pour conclure cette section, cette étude a laissé apparaître trois niveaux différents de calculs d'aires qui se retrouvent dans les deux sous-domaines. La figure 5.39 rappelle les trois niveaux ainsi que la loi de probabilité associée à chaque d'eux. Chaque niveau se rattache à une des trois lois au programme.

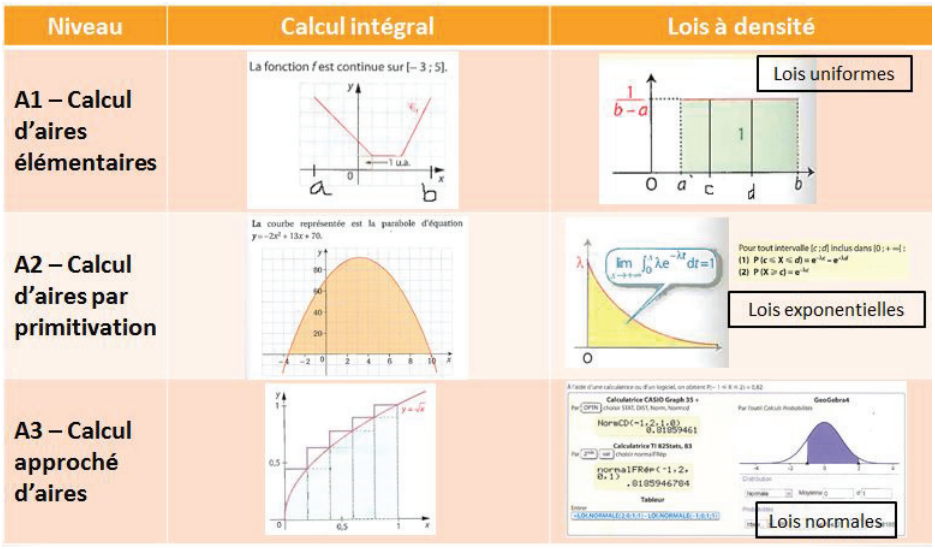


Figure 5.39 – Les trois niveaux de calculs d'aires en calcul intégral et en probabilités à densité

Dans le tableau 5.10, nous avons, pour chaque niveau, précisé les types de fonctions en jeu. Ensuite, nous avons rappelé les différents méthodes de calcul d'intégrale ou de probabilité (suivant le sous-domaine) possibles mises en évidence dans les sections précédentes. Cela nous a permis de repérer les articulations ou les ruptures existantes entre probabilités à densité et calcul intégral.

Nous pouvons remarquer que des articulations existent dans les niveaux A1 et A2, tandis qu'une rupture claire est présente dans le niveau A3.

Pour les lois uniformes et exponentielles, on peut retrouver les mêmes activités attendues des élèves (après adaptation : changement de domaines) qu'en calcul intégral, mais d'autres activités sont aussi possibles telles que l'application directe d'une formule, s'agissant toujours du même calcul.

Pour les lois normales, l'activité attendue en majorité est l'utilisation de la calculatrice. Cette activité est différente de celle attendue en calcul intégral, lorsqu'il s'agit de déterminer l'intégrale d'une fonction dont l'on ne connait pas de primitive.

Niveau	Type de fonctions	CI - Calcul d'intégrales (de fonctions positives)	PaD - Calcul de probabilités	
A1	Affines par morceaux	Application des formules d'aires	Application des formules d'aires	Articulation
		Passage par le niveau A2	Passage par le niveau A2	Articulation
			Application de la formule : $P(c \leqslant X \leqslant d) = \dfrac{d-c}{b-a}$ (dans le cas de la loi uniforme)	Rupture
A2	Fonctions de référence ou composées de fonction de référence	Application du théorème fondamental de l'analyse $F(b) - F(a)$	Application du théorème fondamental de l'analyse $F(b) - F(a)$	Articulation
			Application de la formule : $P(a \leqslant X \leqslant b) = e^{-\lambda a} - e^{-\lambda b}$ (dans le cas de la loi exponentielle)	Rupture
A3	Fonctions dont les élèves ne connaissent pas de primitives	Méthode des rectangles		Rupture
		Calculatrice ou logiciels (pour le calcul intégral)		Rupture
			Calculatrice ou logiciel (pour la loi normale)	Rupture

TABLE 5.10 – Tableau récapitulatif des niveaux et des méthodes correspondantes

5.4 Un point de vue sur les choix des enseignants par rapport aux choix des manuels

Pour compléter l'étude des manuels, des entretiens avec quatre enseignants nous permettront de mettre en lumière quelques éléments des ETM idoines qu'ils ont choisis, non pas dans un but de les détailler précisément mais pour voir s'ils semblent avoir des points communs avec ceux proposés notamment dans les manuels.

Nous allons présenter la méthodologie relative aux entretiens des enseignants. Puis, nous présenterons les éléments clés.

5.4.1 Méthodologie relative aux entretiens

a) Des entretiens semi-directifs

Pour obtenir des informations sur les ETM idoines proposés par des enseignants, sans pour autant aller dans les classes, nous avons fait le choix de procéder à des entretiens semi-directifs. Les thèmes abordés lors de ces entretiens se trouvent en annexe F.1.

Ces entretiens ayant eu lieu tôt dans l'avancement de notre travail de recherche, en juin 2014, les points qui ont été abordés lors de ces échanges étaient assez larges et les fins n'étaient pas ciblées sur nos questions de recherche actuelles. Pour cette raison, nous n'allons pas ici détailler précisément nos choix de thèmes. Le but de ces entretiens étaient pour nous de se faire une idée assez générale sur les choix faits par l'enseignant au niveau de l'enseignement du calcul intégral et des probabilités à densité. Les questions des entretiens se voulaient très larges, pour laisser à l'enseignant la liberté de répondre et de dire ce qui lui passait alors à l'esprit.

Nous ne présenterons pas l'ensemble des réponses apportées mais seulement celles qui permettent de nous éclairer notamment sur l'ordre des chapitres, l'introduction de la fonction de densité et les connexions éventuelles faites entre les probabilités à densité et le calcul intégral au sein de tâches données aux élèves.

Nous avons bien conscience qu'un entretien est déclaratif, ceci est cependant suffisant ici pour ce qui nous intéresse vu qu'il s'agit de connaître les choix des enseignants. Lors des entretiens, les enseignants disposaient de documents relatifs à leurs cours (photocopies des notes de cours d'un élève, documents de l'enseignant, polycopiés distribués aux élèves...). Cela permettait notamment à l'enseignant de pouvoir s'appuyer sur ces documents lors de l'entretien.

b) Profil des enseignants

En juin 2014, à la fin de la première année de la thèse, nous avons procédé à quatre entretiens semi-directifs individuels de quatre enseignants de terminale S. Les quatre enseignants ont été choisis pour leurs profils différents tant du point de vue des années d'expérience, du lieu d'exercice que du milieu social des élèves de leur classe :

– L'enseignant A a plus de 30 ans d'expérience d'enseignement, dont environ 20 ans d'expérience en terminale S. Il exerce dans un lycée, dit prévention

violence, du Val-de-Marne, de milieu social plutôt défavorisé, dans une classe jugée plutôt « faible » par l'enseignant.
– L'enseignant B exerce dans le même lycée que l'enseignant A, cependant ils ne travaillent pas ensemble pour préparer leurs enseignements respectifs. Il a plus de 10 ans d'expérience d'enseignement, dont 7 ans de terminale S. Il travaille quant à lui dans une classe qu'il juge plutôt de bon niveau (elle est constituée d'élèves de section européenne anglais).
– L'enseignante C a plus de 10 ans d'expérience dans l'enseignement et seulement 2 ans d'expérience en terminale S. Elle travaille dans un lycée au Mans, de milieu social plutôt rural, dans une classe hétérogène.
– Enfin, l'enseignante D, qui a plus de 20 ans d'expérience d'enseignement et a été formatrice en IUFM pendant 5 ans. Elle enseigne en terminale S depuis également une vingtaine d'années. Elle exerce dans un lycée parisien du 16^{e} arrondissement, de milieu social plutôt favorisé, dans une classe avec une bonne tête de classe.

Ces quatre enseignants de profils différents sont pour nous un moyen de toucher un panel de pratiques qui se veulent hétérogènes.

c) Les données

Les quatre entretiens ont duré entre 40 minutes et 1h05. Ils ont tous les quatre été enregistré sur dictaphone. Seul l'entretien avec l'enseignante D (Marie), qui sera à nouveau travaillé plus tard, a été transcrit (cf. Annexe F.2). Comme expliqué ci-dessus, nous disposons aussi en complément d'éléments de cours de chaque enseignant.

5.4.2 Élements des entretiens

Tout d'abord, comme dans les manuels, les quatre enseignants interrogés traitent séparément les deux chapitres de calcul intégral et de lois à densité. Le calcul intégral est suivi quelques semaines ou mois plus tard par les lois à densité. Aucun d'entre eux ne fait référence aux probabilités dans le chapitre relatif au calcul intégral.

Au niveau de l'introduction de la fonction de densité, les deux enseignants A et B ont décidé comme le manuel *Odyssée* de ne pas proposer d'activité pour introduire la notion de fonction de densité.

L'enseignant A explique qu'il introduit « *brutalement* » la fonction de densité en donnant la définition et que les premiers exemples (généraux) sont un moyen pour les élèves d'utiliser leurs « *connaissances sur le calcul intégral pour prouver des choses pour le moment sans lien avec la réalité, c'est-à-dire quelque chose d'abstrait* ». C'est donc au départ, un « prétexte » pour faire du calcul intégral. Aucun sens ne semble être donné aux lois à densité. Il explique que la « *concrétisation* » se fait ensuite avec la loi normale. Il s'interroge tout de même sur le problème que peut poser aux élèves cette introduction « *brutale* » de la définition de la fonction de densité, en concluant que comme les élèves ne posent pas de questions, cela ne semble finalement pas leur poser de problème.

L'enseignant B de la même façon impose la définition de la fonction de densité, mais avec un niveau de formalisme encore plus poussé. Son cours a été inspiré no-

tamment de cours destinés aux étudiants de classes préparatoires aux grandes écoles (CPGE) Économique. Il propose aussi quelques exemples généraux pour appliquer la définition, mais passe très vite à la suite.

Pour introduire la notion de fonction de densité, l'enseignante C, quant à elle, propose à ses élèves l'*activité* 1 p. 322 du manuel *Indice* sur l'*Éco-point*, que nous avons présentée page 145. Elle ne propose pas de modification particulière par rapport à l'énoncé du manuel. Les remarques que nous avons pu mettre en avant dans notre analyse (comme le fait que l'histogramme ne permettent pas directement d'obtenir la fonction de densité contrairement à ce que dit l'énoncé, le fait que probabilité et fréquence soient assimilées...) n'ont pas été abordées.

Enfin, l'enseignante D propose une *activité* qu'elle a construite elle-même. On peut la retrouver en annexe F.3.2. L'enseignante présente cette activité entre les lignes 308 et 352 de la transcription en annexe F.2. L'énoncé introductif est le suivant :

> *On s'intéresse aux clients d'une banque qui possèdent un crédit à la consommation. On choisit un client au hasard parmi ces derniers et on note X la variable aléatoire égale à la mensualité de son crédit à la consommation. Les mensualités de remboursement sont comprises entre 100 et 400 euros par mois.*
>
> *Pour définir la loi de probabilité de X, on a effectué une enquête auprès des clients de la banque dont les mensualités de remboursement de crédit à la consommation sont entre 100 et 400 euros par mois. On a tracé l'histogramme des fréquences obtenues.*

La variable aléatoire considérée correspond à la mensualité du crédit à la consommation d'un client d'une banque choisi au hasard. La probabilité que la mensualité soit entre a et b est donc égale à la fréquence des clients ayant une mensualité dans cet intervalle. Aucune phase de modélisation n'est attendue par l'enseignante. L'*activité* est découpée en trois parties (elles sont séparées par des sauts de lignes dans l'annexe), distribuées séparément au fur et à mesure de l'avancement. Tout d'abord, un histogramme d'amplitude 50 est proposé. Puis en demandant de déterminer $P(180 < X < 220)$, les élèves demandent à avoir les fréquences dans des intervalles plus petits (un histogramme d'amplitude 10 est alors distribué aux élèves). Ceci relance le travail de la question 2, et à nouveau surgit un problème pour déterminer $P(122,50 < X < 143,70)$. Ensuite, en prenant des amplitudes de plus en plus petites, les élèves voient « apparaître » une fonction. Pour passer à la fonction de densité, dans la troisième partie de l'*activité*, il est écrit :

> *En utilisant des intervalles d'amplitude de plus en plus petite, on construit des histogrammes d'aire 1 et on peut déterminer une fonction continue f telle que les histogrammes seront de plus en plus proches du domaine sous la courbe de la fonction f.*

Cette *activité* fait apparaître des points communs avec celle de l'*Éco-point* (*activité* 1 p. 322) du manuel *Indice*. On retrouve la même idée de prendre une amplitude des classes plus petite, mais ici il s'agit de diminuer de plus en plus cette amplitude. Cependant, il n'y a pas *a priori* de prise en compte du fait que les données statistiques sont en nombre fini et fixé et qu'il n'est donc pas possible en diminuant

les amplitudes autant que l'on veut de déterminer une telle fonction (de densité). Il s'agit justement à ce moment-là de choisir un modèle pour traiter ces données.

Dans cette *activité*, l'enseignante a conçu ces tâches de telle sorte qu'elles suscitent chez les élèves le besoin d'aménager les données, notamment en diminuant les amplitudes. Malgré cet effort d'introduction, la fonction de densité finit par être imposée par l'énoncé et ses propriétés ne sont pas construites, notamment le fait que l'aire sous la courbe doit impérativement valoir 1.

Chez les quatre enseignants, deux sur quatre donnent directement la définition de la fonction de densité et considèrent finalement les probabilités comme un « prétexte » pour faire du calcul intégral. Nous pouvons nous étonner de ce constat. Les deux autres préfèrent l'introduire avec une *activité* passant par les histogrammes, en ayant choisi une situation de catégorie 1.2, décrite dans la section 5.2.4. Ces *activités* ne permettent cependant pas réellement de construire l'objet fonction de densité avec ses propriétés.

Chacun des quatre enseignants proposent en début de chapitre des fonctions de densité « générales », notamment pour appliquer la définition (et faire du calcul intégral). Ils insistent ensuite particulièrement sur les lois normales.

Nous pouvons conclure que ces quatre entretiens confortent ce que nous avons pu repérer dans les manuels, dans le chapitre 5. Aucune tâche profondément différente de celles proposées dans les manuels n'a été rencontrée, ce qui permet de donner d'autant plus de valeur à l'analyse de manuels, tout en gardant à l'esprit qu'il ne s'agit que de quatre entretiens d'enseignants.

À travers l'entretien avec l'enseignante Marie (la transcription se trouve en Annexe F.2), nous allons essayer de comprendre son point de vue sur l'enseignement du calcul intégral et des probabilités à densité, leurs enjeux, ainsi que sur l'apprentissage des élèves sur ces notions, et aussi sur les contraintes qui peuvent peser sur l'enseignement de ces notions. Il s'agit simplement pour nous de présenter quelques éléments de l'entretien avec Marie, révélateur de quelques points de l'ETM idoine proposé par l'enseignante dans sa classe et surtout sur son ETM personnel relatif aux probabilités à densité et les contraintes institutionnelles qui pèsent sur ces choix. Dans cet entretien, nous avons pu dégager quelques aspects qui nous semblent intéressants.

5.5 Conclusion sur les ETM idoines

Notre analyse des manuels, ainsi que les entretiens des quatre enseignants, montrent que le calcul intégral et les lois à densité sont deux chapitres peu connectés. Le calcul intégral est un prérequis qui est ensuite utilisé comme outil en probabilités.

Les manuels ont fait des choix par rapport aux différents documents officiels et proposent des ETM idoines potentiels légèrement différents de ceux que laissaient entrevoir les documents officiels. Le premier choix est sur le niveau de généralité de l'étude des lois à densité. Dans tous les manuels, les lois à densité sont dans un premier temps définies dans un cadre général. Ils font donc le choix de considérer les trois lois au programme comme des exemples particuliers et non de faire trois études de cas.

Un effort est fait, excepté dans un manuel, pour essayer d'introduire la notion de fonction de densité. Nous avons pu repérer des essais avec, dans une majorité des cas, un passage par la statistique descriptive, et en particulier via la notion d'histogramme. Ce passage nous paraît essentiel comme nous l'avons mis en évidence dans le chapitre 3. Ces essais ne sont finalement pas si concluants si l'on se place du point de vue de la construction de la notion de fonction de densité, de son utilité et de ses propriétés.

Nous avons pu mettre en évidence plusieurs points :

- Il y a un appui essentiel sur la visualisation de l'histogramme et de la courbe, qui sont déjà tous les deux donnés.
- On trouve peu d'appui sur le référentiel théorique : les connaissances mises en jeu viennent peu des élèves (nous pensons tout particulièrement aux propriétés de l'histogramme qui sont pratiquement toujours rappelées).
- Il y a un fort appui sur l'histogramme, qui est une notion non disponible chez les élèves. Nous avons mis en évidence qu'elle ne fait pas partie de l'ETM personnel des élèves, via l'observation des ETM idoines des manuels de seconde.
- Des choix différents sont faits dans les différents manuels au niveau des catégories de démarches (nous en avons exhibé quatre) et au niveau des comparaisons (parfois il s'agit de comparaisons SD/PaD avec CI, d'autres fois SD avec PaD/CI...).
- Dans la catégorie 1 (cf. page 167), les données rencontrées, supposées « réelles », sont toujours « fabriquées ». Elles ont en fait été simulées par les auteurs des manuels suivant une loi prédéfinie, d'où des données trop « parfaites ». Cela enlève les questions liées à la modélisation, démarche qui n'est en fait jamais laissée aux élèves.
- La fonction de densité n'arrive pas comme une réponse à un problème car finalement il n'y a pas de problème posé au départ. Les problèmes n'ont pas le temps d'émerger que des réponses sont déjà données.
- Les différentes *activités* font ressortir que la courbe de densité doit « lisser » ou « épouser » l'histogramme, mais la nécessité de l'aire sous la courbe égale à 1, par exemple, ne ressort jamais explicitement comme une condition nécessaire.

Le choix d'introduire la notion de fonction de densité par une *activité*, plutôt que de l'imposer directement aux élèves, se retrouve chez deux enseignants sur les quatre interviewés. Cela semble peu, mais sur quatre enseignants nous ne pouvons pas tirer de conclusions plus générales.

Après l'introduction de la notion de fonction de densité pour créer le lien entre probabilité à densité et calcul intégral, l'exploitation de ce lien est finalement peu présenté. Cependant, comme dans les sujets de baccalauréat, mis à part quelques exemples exceptionnels, il y a encore des potentialités à exploiter.

En revanche, nous avons fait ressortir trois niveaux de calculs d'aires que nous retrouvons dans les deux sous-domaines, notamment à travers les trois lois au programme. Ces différents niveaux ont laissé apparaître des articulations possibles mais aussi parfois des ruptures entre les deux sous-domaines, notamment le niveau A3, relatif au calcul approché d'aire, avec la loi normale du côté PaD : tandis qu'en calcul intégral, on attend majoritairement une démarche avec la méthode des rectangles (et ponctuellement l'utilisation de la calculatrice ou d'un logiciel avec certaines com-

mandes relatives au calcul intégral), pour le calcul de probabilités associées à une loi normale, on souhaite une détermination à l'aide de la calculatrice avec les commandes relatives à la loi normale (exceptionnellement l'utilisation d'un logiciel).

Pour les deux autres niveaux, des articulations existent. La rupture se crée si les élèves utilisent directement les formules de probabilités liées aux lois uniformes ou exponentielles.

Cette étude des ETM idoines potentiels proposés dans les manuels ou encore ceux choisis par des enseignants laisse ressortir que malgré les efforts de la plupart des manuels et de certains enseignants, la relation entre probabilités à densité et calcul intégral est finalement imposée et ensuite très peu exploitée.

Conclusion de la Partie II

Dans cette deuxième partie, nous avons cherché à répondre à la deuxième question de recherche, qui, rappelons-la, est la suivante :

> **QR2** : En terminale S, quels sont les enjeux didactiques derrière la relation sémiotique $(*)$?
> Quelles sont les tâches proposées et les activités attendues des élèves pour l'introduction de la notion de fonction de densité ?
> Quelles sont les tâches qui permettent d'exploiter les articulations entre l'ETM_{PaD} et l'ETM_{CI} ?

L'étude des documents officiels et celle des manuels ont permis d'approcher l'ETM de référence ainsi que les ETM idoines potentiels proposés par les manuels.

Le programme, le document *Ressources* et les sujets du baccalauréat mettent en exergue deux domaines bien distincts, reliés par des écritures sémiotiques (écriture $(*)$ ou encore formule de l'espérance), qui ne sont jamais justifiées. Seule la loi normale centrée réduite est introduite par un travail spécifique. Ensuite, dans les tâches du baccalauréat, malgré des potentialités, les liens entre probabilités à densité et calcul intégral ne sont jamais exploités, mis à part pour le calcul de probabilités de la loi exponentielle ou autres démonstrations du cours.

Il ressort de l'analyse du programme et du document *Ressources* que les contours du programme en probabilités sont flous et laissent une certaine liberté de choix aux enseignants, quant la justification et à l'exploitation de la relation entre probabilités et analyse.

Cette marge de manœuvre laissée par le programme est exploitée par les manuels, qui essaient de donner un peu de sens à la relation entre les deux sous-domaines. Ils proposent notamment des *activités* d'introduction de la notion de fonction de densité, ce qui est inexistant dans le document *Ressources* par exemple.

Nous avons mis en évidence que la plupart des manuels font le choix de passer par le sous-domaine de la statistique descriptive pour amener la notion de fonction de densité, ce qui nous semble être une approche intéressante. Nous avons pu le voir dans la Partie I, la place de la statistique est importante dans la naissance historique de la notion de fonction de densité et de lois à densité.

Les *activités* en question s'appuient sur la notion d'histogramme qui est, nous semble-t-il, un bon levier pour faire émerger la fonction de densité, comme nous avons pu le trouver dans le cours de géodésie de Faye en 1881. Ce passage nous

semble être intéressant pour comprendre que le calcul intégral est nécessaire pour considérer les questions de probabilités. Cependant, nous ne retrouvons pas dans les manuels l'idée d'ajustement de la courbe par rapport aux données comme c'était le cas chez Faye.

Nous avons identifié quatre catégories de démarche utilisant la notion d'histogramme pour introduire la fonction de densité :

- Une première catégorie partant de données statistiques pour aboutir aux probabilités, divisée elle-même en deux suivant l'appartenance des données à un échantillon (catégorie 1.1) ou à une population totale (catégorie 1.2).
- La catégorie 2 est une démarche où l'on connait (élèves ou auteurs de manuel), des lois de probabilités qui nous permettent de faire des simulations, ce qui ensuite permet de déduire une loi inconnue des élèves.

Nous avons distingué la catégorie 2.1, où la simulation est proposée par le manuel car la loi sous-jacente est connu au départ par les auteurs de manuel, et la catégorie 2.2 où la loi inconnue est combinaison de lois connues par les élèves. Le tableau 5.7 des catégories se trouve page 167. C'est finalement la catégorie 1.1 qui correspond à la logique historique, ce qui ne discrédite pas les autres démarches, à première vue.

Pour proposer une introduction de la notion de fonction de densité basée sur la notion d'histogramme, il semble important que les connaissances sur l'histogramme soient disponibles pour les élèves. Or, nous avons mis en évidence que cet objet mathématique est mal présenté dans beaucoup de manuels (de terminale S et de seconde, classe où est étudiée cette représentation graphique), ce qui laisse penser que la notion d'histogramme n'est pas maîtrisée par les élèves. La plupart des manuels rappellent d'ailleurs la relation entre fréquence et aire des rectangles. La notion d'histogramme étant mal connue, cela empêche de faire véritablement émerger une problématique autour de la recherche d'une fonction de densité. D'ailleurs, la courbe de densité est toujours tracée (sauf une fois) et introduite comme une courbe qui « *épouse* » l'histogramme. Cela nous semble être un frein à la compréhension de la nécessité des propriétés que vérifient la fonction de densité. Le travail des manuels s'arrête à une validation par visualisation de ce qu'ils imposent.

Les logiciels sont parfois utilisés mais le but n'est pas, par exemple, d'obtenir rapidement de nouveaux graphiques, avec de nouveaux échantillons, ou des amplitudes de classes différentes...

Pour conclure, dans les introductions des manuels, il n'y a pas de véritable construction du référentiel théorique relatif aux fonctions de densité. Il reste finalement imposé.

Nous avons notamment montré que sur quatre enseignants interrogés, deux imposent directement la définition de fonction de densité et deux proposent une *activité* d'introduction du type 1.2 (d'un manuel ou conçu personnellement). Nous retrouvons donc que, malgré les efforts de certains enseignants à essayer de donner du sens à cette introduction, la notion de fonction de densité finit par être imposée, faute de mieux.

Au niveau des tâches proposées, peu exploitent réellement la relation entre probabilités et analyse. Nous avons cependant montré que des articulations, mais aussi parfois des ruptures, existent au niveau des différents calculs d'aires rencontrés dans les deux domaines. Nous avons mis en évidence que nous pouvons identifier trois

niveaux de calculs d'aires (A1/A2/A3), qui se retrouvent en probabilités et en calcul intégral. Cependant, les démarches de calculs associées ne sont pas toujours les mêmes que l'on soit dans un domaine ou dans l'autre.

Dans l'annexe F.4, nous revenons sur l'entretien de l'enseignante D (Marie) et mettons en avant des citations qui recoupent des éléments que nous venons de dégager sur les documents institutionnels ainsi que sur les manuels. Nous avons aussi mis en évidence d'autres contraintes, institutionnelles ou personnelles, qui pèsent sur l'enseignement de ces notions.

La Partie II de la thèse a permis de constater que la question de l'introduction de la fonction de densité dans la sous-partie des lois à densité n'est soulevée ni dans le programme, ni même dans le document *Ressources* (chapitre 4). Le lien entre les probabilités à densité et le calcul intégral est imposé sémiotiquement. Aucune proposition n'est faite pour l'introduire.

Les manuels quant à eux tentent d'introduire la notion de fonction de densité, dans les *activités* proposées en début de chapitre sur les lois à densité. Cependant, nous avons pu remarquer que ces *activités* n'amènent pas réellement à la construction de l'objet fonction de densité, notamment elles ne font pas émerger les propriétés relatives à une telle fonction. Il existe donc un saut épistémologique. Le fait qu'une partie reste imposée aux élèves peut empêcher une bonne appropriation de cet objet mathématique (chapitre 5). Dans la construction historique de la notion (chapitre 3), nous avons pu repérer que la statistique descriptive est un passage intéressant pour amener la notion de fonction de densité, notamment par le biais de la notion d'histogramme. Cette approche est aussi adoptée par la majorité des manuels. Cependant, comme nous l'avons aussi montré, l'histogramme reste un objet mathématique mal maîtrisé qui mérite d'être repris en amont si l'on souhaite s'appuyer dessus (chapitre 5).

Bien que n'ayant trouvé aucune tâche remplissant réellement notre objectif de construction conceptuelle de la notion de fonction de densité dans les manuels et les différents documents institutionnels, les Partie I et Partie II nous permettent tout de même de dégager des pistes à explorer pour concevoir de telles tâches.

Nous proposons plusieurs éléments qui nous semblent indispensables ou possibles pour créer la connexion entre les probabilités à densité et le calcul intégral dans l'introduction de la fonction de densité :

- La démarche de catégorie 1.1 (à savoir la démarche partant de données statistiques d'un échantillon (SD) pour inférer des résultats sur la population entière (PaD)) et la démarche de catégorie 2.2 (démarche où l'on part de variables aléatoires suivant des lois connues (PaD) pour en déduire des résultats sur une loi inconnue, combinaison des lois connues (PaD) par le biais de simulations (SD)) nous semblent pertinentes, si les problèmes associés font bien ressortir le besoin de choisir un modèle. La catégorie 2.1 nous semble trop artificielle et la catégorie 1.2 peut amener des confusions en assimilant probabilité et fréquence.
- L'utilisation de données réelles plutôt que des données « fabriquées » (pour la catégorie 1) nous semble indispensable, pour ne pas avoir des données trop « parfaites ». Des données réelles permettent de vraiment se poser la question

du choix d'un modèle. La modélisation semble donc un point important à prendre en compte.
– Ne pas fournir la courbe de densité nous semble être le meilleur moyen pour que les élèves se posent véritablement des questions sur ce qu'ils cherchent et qu'ils fassent émerger les propriétés nécessaires à cette courbe.
– La maîtrise de la notion d'histogramme, qui est notamment porteuse de sens pour le terme « densité », est nécessaire en amont de l'introduction.

Les niveaux de calcul d'aires A1/A2/A3, que nous avons pu repérer en calcul intégral et en probabilités à densité, nous semblent eux aussi à exploiter pour créer les liens entre les deux sous-domaines.

Ces premières remarques nous font entrevoir des pistes nous permettant de proposer des introductions de la notion de fonction de densité plus riches afin de permettre de faire émerger l'articulation qui existe entre les probabilités à densité et le calcul intégral. L'objectif de la Partie III sera donc de proposer, d'expérimenter et d'analyser des tâches d'introduction.

Troisième partie

Élaboration d'une ingénierie didactique collaborative. Des problèmes de modélisation articulant l'introduction des probabilités à densité et du calcul intégral

Bien que n'ayant trouvé aucune tâche remplissant réellement notre objectif de construction conceptuelle de la notion de fonction de densité dans les manuels et les différents documents institutionnels, les Partie I et Partie II nous ont tout de même permis de dégager des pistes à explorer pour concevoir de telles tâches (cf. conclusion de la Partie II).

Nous gardons donc l'hypothèse de recherche que des tâches permettant réellement de construire la notion de fonction de densité existent et qu'elles permettent de donner un véritable sens à la relation sémiotique $(*)$.

Nous pensons que la notion de fonction de densité peut être introduite comme réponse à un problème, appelée notion RAP (Robert, 2008d). Nous avons d'ailleurs identifier, dans la Partie I, à quels problèmes la notion a répondu historiquement. De plus, même si la notion de fonction de densité n'est pas une véritable extension de la notion d'histogramme, nous pouvons identifier une certaine analogie entre ces deux notions. Nous pouvons penser, dans un certain sens, que la notion de fonction de densité peut être une extension avec accident de la notion de fonction de densité (Robert, 2008d). Ces arguments renforcent notre idée que des problèmes peuvent effectivement permettre d'introduire efficacement la notion de fonction de densité, bien que les propositions des manuels ne nous aient pas pour autant convaincues.

L'objectif de cette troisième partie est de répondre à la troisième question de recherche qui est la suivante :

> **QR3** : Quels éléments didactiques peut-on prendre en compte pour concevoir et mettre en œuvre effectivement des tâches qui permettent aux élèves de construire la notion de fonction de densité ?

Dans le cadre d'une méthodologie de type ingénierie didactique, les objectifs de cette partie sont donc de :

- Déterminer des éléments didactiques (au niveau des contenus mathématiques mais aussi des artefacts par exemple, ou de la gestion des activités) qui semblent *a priori* à prendre en compte dans la conception de tâches ayant l'objectif visé.
- Expérimenter dans une classe puis analyser les déroulements de ces séances pour valider ou non les éléments retenus et ajouter des conditions supplémentaires qui ont semblé nécessaires, *a posteriori*, notamment sur la gestion des activités des élèves, pour que l'objectif soit atteint.

Nous ajouterons, de plus, la condition que ces tâches soient viables. Par viable, nous reprenons la définition proposée par Bednarz, Poirier et Desgagné (2001) :

> Les enseignants jugeront viable une action, une manière de faire tant et aussi longtemps qu'elle servira à l'accomplissement d'une tâche, à l'atteinte du but visé, et cela dans leur contexte, en tenant compte des contraintes qui sont les leurs. (p. 45)

Il semble donc important pour respecter cette condition de viabilité de prendre en compte les contraintes qui pèsent sur l'enseignant et sur l'enseignement des différentes notions en jeu. Pour cela, nous avons décidé de travailler en collaboration

avec l'enseignante Marie (enseignante D du chapitre 5). Nous avons présenté, en annexe F.4, quelques éléments de son ETM personnel et idoine ainsi que les différentes contraintes qui pèsent sur ses choix. Il ressort de son entretien, notamment, une insatisfaction de son *activité* d'introduction de la notion de fonction de densité, ainsi qu'un besoin et une envie de formation sur cette question d'introduction.

Les tâches d'introduction conçues sont deux problèmes impliquant une démarche de modélisation. Elles sont intégrées dans une séquence entière d'enseignement prenant appui sur cette introduction. Notre proposition de séquence sort de l'enseignement traditionnel proposé par les manuels et dans les classes (cf. chapitre 5), mais reste entièrement dans le cadre du programme. Cette séquence a la particularité de coupler l'enseignement des probabilités à densité avec celui du calcul intégral. Il ne s'agit plus, comme il est fait traditionnellement, de découvrir le calcul intégral avant les probabilités à densité, mais que les deux sous-domaines se nourrissent mutuellement et, notamment, que les probabilités motivent le besoin du calcul intégral. Les tâches d'introduction en plus d'avoir pour objectif d'introduire la fonction de densité auront aussi l'objectif d'introduire le calcul intégral.

Dans cette partie, contrairement à la Partie II, les chapitres seront dépendants les uns des autres. Les trois chapitres de cette partie répondent finalement à des phases différentes de notre méthodologie.

Dans le **chapitre 6**, nous préciserons dans un premier temps la méthodologie que nous avons mise en place pour répondre à la question QR3. Nous définirons ce que nous entendons par une méthodologie de type ingénierie didactique collaborative. Ce chapitre sera aussi l'occasion de présenter et justifier les choix que nous avons faits pour la conception des tâches incorporées dans la séquence entière. De la trame générale du scénario de la séquence, nous extrairons les choix explicites dans les tâches d'introduction que nous souhaitons tester avec notre méthodologie de type ingénierie didactique collaborative. Cela nous permettra alors de détailler nos questions de recherche relatives à la question de recherche QR3.

Dans le **chapitre 7**, nous aborderons le cœur de cette Partie III, à savoir la présentation détaillée de nos analyses didactiques des deux problèmes d'introduction. Il s'agira tout d'abord d'expliciter les outils méthodologiques choisis pour répondre à nos questions de recherche, puis de les mettre en œuvre pour les analyses *a priori* et *a posteriori* des séances d'introduction et enfin de présenter nos résultats de recherche, qui constituerons la validation interne.

Pour finir, dans le **chapitre 8**, nous apporterons une ouverture sur nos résultats de recherche. Nous essaierons en effet d'élargir la portée des résultats obtenus. Nous aborderons notamment des points non visibles dans l'expérimentation, comme les élèves individuels, et nous questionnerons la reproductibilité de ces tâches, dans d'autres classes, avec d'autres enseignants.

Chapitre 6

Présentation de l'ingénierie : conception et problématique

Dans la conclusion de la Partie II, nous avons présenté certaines pistes qui se sont dégagées des analyses des Partie I et Partie II. Dans le but de répondre à la question de recherche QR3, nous souhaitons tester ces pistes pour en extraire des éléments didactiques pouvant être pris en compte pour la conception et la mise en œuvre effective dans une classe de terminale S de tâches permettant réellement de construire la notion de fonction de densité. Cette troisième partie de la thèse a pour objectif de présenter l'ensemble du processus à savoir la conception, l'expérimentation, l'analyse et la validation.

Dans ce chapitre, nous commencerons par détailler notre méthodologie (section 6.1), qui s'inscrit dans une méthodologie de type ingénierie didactique. Cependant, nous ajouterons à cette méthodologie une dimension collaborative qui nous paraît ici nécessaire. Ensuite, nous nous focaliserons sur la phase de conception des tâches. Dans le cadre de cette méthodologie, nous exposerons les contraintes et les pistes à prendre en compte dans la conception (section 6.2) pour arriver à la description de la trame générale du scénario de la séquence conçue (section 6.3). Pour finir, les principaux choix retenus seront exhibés pour permettre de préciser la problématique et de nouvelles questions de recherche permettant de répondre à la question de recherche QR3 (section 6.4).

6.1 Une méthodologie de type ingénierie didactique collaborative

Notre travail s'inscrit dans une méthodologie de recherche de type ingénierie didactique (Artigue, 1988 ; 2011). Cependant, nous l'avons adaptée en ajoutant une dimension de recherche collaborative (Desgagné, 1997). Dans cette section, nous allons présenter la méthodologie que nous avons mise en place. Nous exposerons les éléments caractéristiques de ces méthodologies et ensuite nous montrerons comment nous les combinons.

6.1.1 La méthodologie de type ingénierie didactique

L'ingénierie didactique a été conçue comme une méthodologie de recherche (Artigue, 1988). Elle possède deux caractéristiques principales :

> – être une méthodologie basée sur des réalisations didactiques en classes, c'est-à-dire sur la conception, la réalisation, l'observation et l'analyse de séquences d'enseignement,
> – être une méthodologie dont la validation est essentiellement interne, fondée sur la confrontation entre analyse *a priori* et analyse *a posteriori* (et non externe basée sur la comparaison des performances de groupes expérimentaux et témoins [...]). (Artigue, 2011, p. 20)

Cette méthodologie est un processus expérimental constitué de quatre phases (Artigue, 1988) :

1. les analyses préalables (analyse épistémologique des contenus visés, analyse de l'enseignement usuel et de ses effets, analyse des conceptions des élèves...), sur lesquelles s'appuie la conception de l'ingénierie ;
2. la conception et l'analyse *a priori* de l'ingénierie dans lesquelles le chercheur agit sur un certain nombre de variables permettant, *a priori*, d'engager le processus de validation ;
3. l'expérimentation permettant de recueillir diverses données ;
4. l'analyse *a posteriori* s'appuyant sur ces données et la validation qui se fonde essentiellement sur la confrontation des analyses *a priori*/*a posteriori*.

Perrin-Glorian (2011) s'interroge sur les rapports entre recherche et enseignement ordinaire en « *regardant l'ingénierie didactique non seulement comme un moyen de faire apparaître des phénomènes didactiques pour la recherche mais aussi comme un moyen pour étudier et faire évoluer l'enseignement ordinaire* » (p. 57). La question, pour nous, ne sera pas d'adapter une ingénierie didactique existante pour la rendre viable dans une classe ordinaire mais bien de concevoir une ingénierie didactique déjà adaptée à une classe « ordinaire » (si tant est qu'elle puisse l'être si elle rentre dans un tel processus expérimental). Nous avons bien conscience, comme le dit Perrin-Glorian (2011) que « *l'ingénierie didactique sera la manière d'enseigner la notion concernée cette année là, elle entre dans le temps didactique de la classe* » (p. 57).

C'est la raison pour laquelle nous pensons que certaines conditions sont à respecter pour sa bonne intégration dans la classe. Il nous a donc semblé insuffisant de concevoir notre ingénierie sans la participation de l'enseignant ayant en charge la classe ; d'autant plus qu'il s'agit d'une séquence entière de plus de trois semaines, qui mélange les deux chapitres traditionnels du calcul intégral et des lois à densité. Pour cette raison, nous avons donc fait le choix de procéder à une coconstruction de l'ingénierie avec une enseignante. Nous avons donc adapté la méthodologie de l'ingénierie didactique en y ajoutant une approche collaborative de recherche.

6.1.2 Une recherche collaborative

Une des contraintes de notre expérimentation était qu'elle se fasse dans une classe de terminale S, sans qu'elle ne change les conditions de fonctionnement habituelles de cette classe, à savoir sans aménagement du temps de classe, sans utilisation de

matériel nouveau, sans modification des objectifs de la classe (par exemple, l'objectif du baccalauréat reste omniprésent)... La viabilité de la séquence dans la classe est donc centrale dans notre méthodologie.

Certains chercheurs (Bednarz, Poirier & Desgagné, 2001) proposent d'aborder la question de la production de séquences d'enseignement dans une autre perspective que celle de l'ingénierie didactique. Pour eux, une telle production ne peut se passer du point de vue des praticiens, c'est-à-dire des enseignants. Ils parlent de recherche collaborative. Desgagné (1997) précise que :

> Le point de vue privilégié par cette approche est ici celui de l'intervenant (l'enseignant) et des connaissances qui peuvent être construites et mises au service de son intervention de praticien, elle-même définie sur la base de l'ultime responsabilité à laquelle elle renvoie, soit l'apprentissage des élèves. (p. 372)

Cette approche accorde une place toute importante à l'association chercheurs et enseignants.

Bednarz *et al.* (2001) expliquent l'importance de cette collaboration :

> Il ne s'agit pas seulement pour nous de développer des situations d'enseignement riches et pertinentes sur le plan des apprentissages, contribuant à une appropriation d'un certain savoir par les élèves (ce qu'une analyse didactique peut bien sûr permettre d'éclairer), mais de produire des situations qui soient aussi viables en contexte (que valent en effet des scénarios s'ils ne rencontrent aucun écho dans l'expérience ?). Dans ces conditions, les scénarios d'enseignement s'ils sont retenus par les différents partenaires, vont apparaître comme une façon légitime de donner un sens à une expérience d'enseignement qui s'est avérée plausible et féconde sur le plan des apprentissages et qui va alors guider les actions futures. (p. 45)

Ces conditions permettent de donner plus de poids aux scénarios d'enseignement considérés. Ce travail d'équipe entre enseignants et didacticiens « *sera ainsi appelé à construire des activités, des interventions non seulement fécondes au plan des apprentissages mais aussi viables dans la pratique* » (*ibid.*, p. 46).

D'autre part, nous faisons l'hypothèse que les pratiques sont complexes et cohérentes (de Montmollin, 1984) et sont stables chez les enseignants expérimentés (Robert, 2008a). Il n'est donc pas raisonnable de proposer à un enseignant une séquence trop éloignée de ses pratiques habituelles. Cela renforce l'idée qu'il est important de prendre en compte son point de vue d'enseignant dans la construction de la séquence.

Dans notre méthodologie de type ingénierie didactique, il nous parait donc inévitable de considérer une dimension collaborative au sens notamment de la coconstruction de la séquence avec (au moins) un enseignant, prenant en compte les pratiques de l'enseignant et ses contraintes, ainsi que les connaissances du chercheur.

La recherche collaborative est caractérisée par une double dimension : une dimension formation (du côté de l'enseignant qui cherche à se « perfectionner ») et une dimension recherche (du côté du chercheur). Au cœur du modèle de recherche collaborative (Desgagné, Bednarz, Lebuis, Poirier & Couture, 2001), se trouve une activité réflexive, qui peut être aménagée de diverses façons suivant les projets : « *Cette activité prend forme à travers des rencontres régulières entre chercheurs et praticiens, rencontres qui permettent ainsi de créer une « zone interprétative » autour de la pratique qui est objet d'exploration* » (*ibid.*, p. 37). A l'intérieur de

ces échanges, les enseignants développent les dimensions de leur expérience et les croisent avec les interprétations des chercheurs.

La recherche collaborative évolue suivant plusieurs étapes. La première étape du projet est la négociation entre chercheurs et praticiens d'un objet de réflexion commun. Il existe plusieurs amorces possibles d'un projet : « *Pour les uns* [...], *la sollicitation vient des enseignants et s'exprime par un intérêt de perfectionnement sur un aspect de la pratique qui les préoccupe. Pour d'autres* [...], *la sollicitation vient des chercheurs et s'exprime par un intérêt de recherche sur un aspect de la pratique* » (Desgagné *et al.*, 2001, p. 54). Ce qui est sûr, c'est que chacun des côtés doit y trouver un intérêt :

> Elle prend d'abord le sens d'un échange de services entre des acteurs qui font partie de cultures différentes et qui n'ont pas à répondre aux mêmes finalités. La culture scolaire incite le praticien enseignant à se donner les moyens de développer et d'améliorer sa pratique, plus spécifiquement la qualité de son intervention auprès des élèves dont il a la responsabilité. La culture scientifique incite le chercheur universitaire à contribuer à la production des connaissances dans un domaine donné. L'entente collaborative consiste à faire en sorte qu'une même activité réflexive aménagée autour d'un projet d'exploration négocié fasse que ces attentes respectives soient comblées à la satisfaction des deux parties. [...] Le chercheur vise une production de connaissances qui inclut et tienne compte du point de vue du praticien et des contraintes de son contexte d'action. De même, on pourrait dire que le praticien vise un développement de pratique qui soit éclairé par le point de vue du chercheur et par les repères conceptuels qui guident sa production de connaissances. (Desgagné *et al.*, 2001, p. 39)

Dans ce modèle, contrairement à ce que Sensevy (2011) appelle ingénierie coopérative, les fins du chercheur et du praticien ne sont pas nécessairement communes et le rôle de chacun n'est pas obligatoirement symétrique : il n'est pas nécessaire que les praticiens participent aux tâches formelles de recherche. Il est important, dans cet esprit collaboratif, que les intérêts spécifiques de chacun soient respectés.

Enfin, il est important dans ce modèle que « *les produits de la démarche et leur diffusion aient des retombées tout autant pour la communauté de pratique que pour la communauté de recherche* » (*ibid*, p. 40).

Le modèle de la recherche collaborative est schématisé dans la figure 6.1.

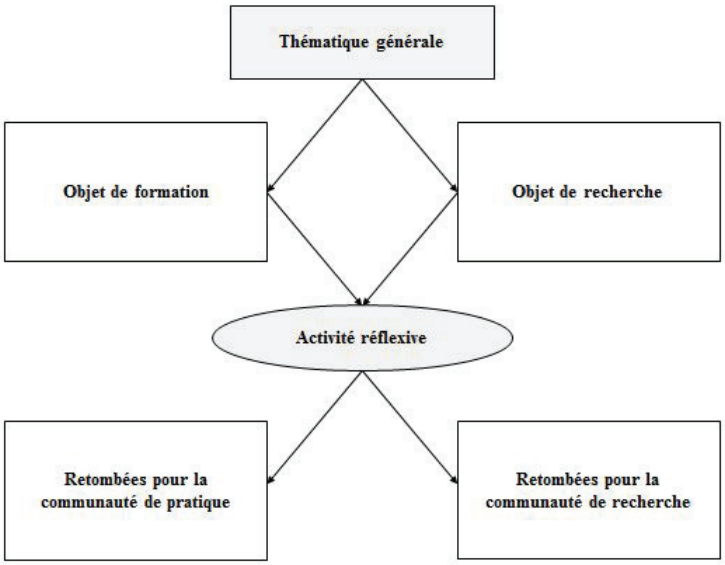


Figure 6.1 – Schéma du modèle de recherche collaborative (inspiré de Desgagné *et al.*, 2001, p. 41)

6.1.3 Vers une méthodologie de type ingénierie didactique collaborative

Dans le cadre de notre travail, nous avons développé ce que nous appelons une méthodologie de type ingénierie didactique collaborative. Il s'agit d'une ingénierie didactique, dans le sens que nous avons donné dans la section 6.1.1, dans laquelle nous avons ajouté une dimension collaborative, notamment avec la coconstruction de la séquence avec un enseignant. En revanche, l'activité réflexive de la recherche collaborative sera présente mais pas centrale dans cette méthodologie.

Cette méthodologie d'ingénierie didactique collaborative est structurée par la méthodologie de l'ingénierie didactique, nous retrouvons donc les mêmes phases caractéristiques. Cependant, à l'intérieur de ces phases, des nuances sont apportées du fait que le travail soit collaboratif (entre chercheur et enseignant). Nous donnons quelques précisions ci-dessous. Nous présenterons ensuite, dans la section suivante, la manière dont nous l'avons mise en œuvre dans notre recherche.

a) La phase 1 : Analyses préalables

Les analyses préalables regroupent l'analyse épistémologique et historique des contenus visés, l'analyse curriculaire, l'analyse de l'enseignement usuel et de ses effets. Il s'agit aussi de l'analyse des conceptions des élèves.

Ces analyses préalables sont faites par le chercheur. Cependant, il est tout à fait envisageable que l'enseignant impliqué dans le travail collaboratif participe aussi à cette phase ; par exemple, il peut s'investir dans les recherches épistémologiques ou encore, il est possible qu'il contribue à l'analyse de l'enseignement usuel ou des

conceptions des élèves, à travers des entretiens ou des questionnaires.

b) La phase 2 : La conception et l'analyse *a priori*

La phase de conception est nécessairement collaborative. En effet, la coconstruction de la séquence d'enseignement est centrale dans cette méthodologie. Il doit y avoir une prise en compte par le chercheur des contraintes de l'enseignant, mais aussi de ses pratiques habituelles pour veiller à ne pas trop s'en écarter.

La collaboration peut se décliner sous plusieurs formes. La conception peut se faire uniquement pendant des temps communs ou alterner des réflexions individuelles et des mises en commun régulières. Les formes des échanges peuvent être sous forme de réunions, par mails...

En un certain sens, une analyse *a priori* est faite conjointement lors de la coconception. Cependant, d'un point de vue de la recherche, l'analyse *a priori* est plutôt prise en charge par le chercheur.

c) La phase 3 : L'expérimentation

L'expérimentation est à nouveau une phase collaborative, la plupart du temps. L'enseignant a le rôle principal. Le chercheur est, quant à lui, observateur ; il peut, cependant, arriver que le chercheur ne soit pas présent lors de l'expérimentation.

d) La phase 4 : L'analyse *a posteriori* et la validation

Tout comme l'analyse *a priori*, l'analyse *a posteriori* est assumée par le chercheur, ainsi que la validation interne, qui se fonde essentiellement sur la confrontation des analyses *a priori*/*a posteriori* dans le cadre des questions de recherche posées.

Cependant, en gardant à l'esprit la dimension collaborative, la validation de l'ingénierie par l'enseignant après l'expérimentation est, elle aussi, à prendre en compte. Nous ne pouvons être pleinement satisfait de l'ingénierie si l'enseignant ne l'est pas.

6.1.4 La mise en œuvre de la méthodologie

Nous allons maintenant décrire la mise en oeuvre de cette méthodologie, dans le contexte de notre recherche particulière.

a) La phase 1 : Analyses préalables

Cette phase est certainement celle dans laquelle la collaboration est la moins présente, tout du moins au départ. Dans notre cadre théorique, les analyses préalables sont la mise en relief de la notion fonction de densité (Robert, 2007). Cela correspond donc à notre analyse historico-épistémologique (cf. chapitre 3), à notre analyse des ETM de référence (cf. chapitre 4) et à notre analyse des ETM idoines proposés par les manuels (cf. chapitre 5).

Ces analyses ont permis de dégager quelques pistes, que nous avons déjà mentionnées (cf. conclusion de la Partie II), à tester dans le cadre de l'ingénierie.

Le projet de collaboration est né lors de l'entretien avec l'enseignante Marie, en juin 2014 (cf. Annexes F.2 et F.4). Notre intérêt de chercheur et l'envie de Marie se sont rejoint.

Notre objectif de recherche est de travailler sur l'enseignement des probabilités à densité en lien avec le calcul intégral en terminale S et notamment sur l'introduction de la notion de fonction de densité. Cet intérêt se retrouve aussi chez l'enseignante, Marie, qui est en demande d'*activités* d'introduction plus « *convaincantes* », dans un but de « perfectionner » son enseignement.

La collaboration s'est limitée à deux partenaires : Marie, enseignante, et nous, chercheuse. La conception avec une seule enseignante nous paraissait déjà une bonne façon d'aborder notre question de recherche, en prenant réellement en compte les pratiques de l'enseignante.

L'objectif de formation est bien présent dans la méthodologie, cependant il n'en est pas l'objet de recherche.

b) La phase 2 : La conception et l'analyse *a priori*

La conception s'est faite de façon conjointe. La conception de la séquence favorisant les articulations entre probabilités à densité et calcul intégral a constitué l'activité pivot permettant de joindre l'intérêt de l'enseignante et notre intérêt de chercheuse.

La construction conjointe (coconstruction) de la séquence, et en particulier des problèmes d'introduction, a permis de prendre en compte à la fois, d'une part, les connaissances que nous avons pu constituer avec nos analyses préalables et, d'autre part, les pratiques de l'enseignante et les contraintes imposées par l'enseignante, qui sont pour certaines apparues lors de l'entretien (cf. Annexe F.2) ou bien par la suite.

La démarche conjointe de conception de la séquence s'est déroulée sur environ sept mois, de septembre 2014 à mars 2015, mois au cours duquel a débuté l'expérimentation en classe. La construction de la séquence ne s'est pas faite à deux tout le temps. Il y a eu des allers-retours entre les deux partenaires. La première étape a été de prendre en compte les pratiques habituelles de l'enseignante, que nous avons pu identifier à travers son entretien (cf. chapitre 5 et Annexe F.2), ses notes de cours de l'année scolaire 2013-2014 (cf. Annexe F.3) et quelques discussions informelles. Cette prise en compte est primordiale pour ne pas faire de propositions trop éloignées de ses pratiques.

Une fois ces éléments pris en compte, nous avons présenté à l'enseignante nos premières idées lors d'une première réunion, le 16 octobre 2014. Ensuite le travail conjoint de conception s'est réparti en différents modes d'échanges :

- des discussions informelles,
- trois réunions programmées d'une demi-journée chacune avant la séquence,
- des échanges par mail et par téléphone,
- des discussions au cours de l'expérimentation pour faire des modifications « sur le tas ».

Une analyse *a priori* des tâches d'introduction s'est effectuée conjointement, dans un langage courant. À deux, nous avons essayé de mettre en avant les différentes résolutions possibles, les blocages pouvant intervenir, les interventions que l'ensei-

gnante devra/pourra mettre en place... En revanche, nous avons, personnellement, assumé pleinement l'analyse *a priori*, d'un point de vue recherche.

c) La phase 3 : L'expérimentation

L'expérimentation a été menée dans la classe de terminale S de Marie de l'année scolaire 2014-2015. Il s'agit d'une classe de 35 élèves d'un lycée parisien du 16e arrondissement de milieu social plutôt favorisé (Marie n'a pas changé de lycée entre l'année scolaire 2013-2014 et 2014-2015). Selon Marie, sa classe est d'un bon niveau, mis à part quelques cas particuliers.

L'expérimentation a eu lieu du 13 mars au 4 avril 2015, soit sur plus de trois semaines de cours (13 séances d'une ou deux heures). Nous avons assisté à l'ensemble des séances de la séquence (mis à part quelques séances en demi-groupe) en tant qu'observatrice dans le fond de la classe. Nous nous déplacions parfois dans les rangs lors des phases de recherche.

Le chef d'établissement n'ayant pas accepté que la classe soit filmée, nous avons opté pour un enregistrement audio à l'aide d'un dictaphone. L'intégralité des séances en classe entière de la séquence, soit au total 13 séances de 55 minutes ou 1h50, ont été enregistrée. Étant donné que les gestes de l'enseignante et des élèves nous importaient peu, l'enregistrement audio a été une bonne option. De plus, le fait que nous étions présente en observateur nous a permis de prendre des notes et aussi de prendre des photos du tableau tout au long des séances, pour garder trace de ce que l'enseignante décide de prendre en note. Nous avons parfois complété nos données par des productions d'élèves, ainsi que des prises de notes d'élèves. Nous détaillerons précisément nos données au moment de l'analyse.

d) La phase 4 : L'analyse *a posteriori* et la validation

Nous avons pris entièrement en charge cette phase, les questions qui guident ce travail étant des questions de recherche. Cependant, la dimension collaborative reste présente à l'esprit dans les analyses et un regard particulier est mis sur la gestion du travail mathématique dans la classe.

Dans la phase de validation, le point de vue de l'enseignant est aussi pris en compte. La satisfaction de l'enseignante fait aussi partie de la réussite ou non de l'ingénierie.

Nous avons essayé de rendre compte de l'imbriquement méthodologique de l'ingénierie didactique et de la recherche collaborative dans cette recherche, dans la figure 6.2. Il ne faut pas voir une répartition bien distincte entre les phases de l'ingénierie didactique et les éléments de la recherche collaborative, mais c'est un moyen pour nous de présenter la globalité de la méthodologie mise en place. Nous tenons à pointer le fait que, dans notre travail de thèse, les retombées pour la communauté de pratique ne font pas partie de nos objectifs de recherche visés. Nos objectifs sont centrés sur le point de vue de la recherche, que nous avons encadré en rouge.

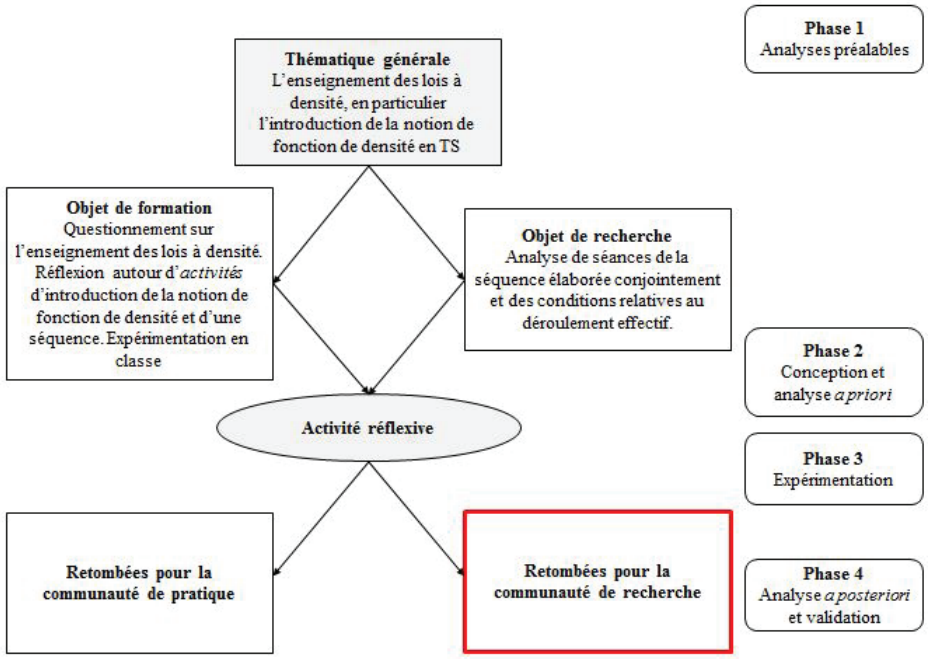


Figure 6.2 – Schéma de notre modèle d'ingénierie didactique collaborative

e) Chronologie du travail collaboratif

La chronologie du travail collaboratif mis en place avec l'enseignante Marie se trouve en figure 6.3. Nous n'avons pas précisé les échanges « informels » que nous avons eu, par mail ou par téléphone par exemple, qui ont eu lieu sur la période de septembre à avril 2015.

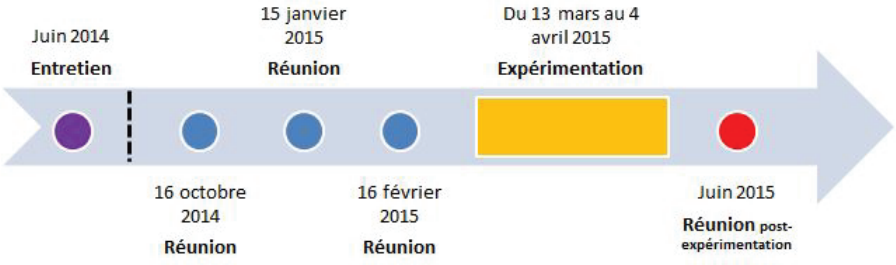


Figure 6.3 – Chronologie du travail collaboratif

f) Plan de l'étude des différentes phases

Les différentes phases de la méthodologie seront présentées successivement dans la Partie III. Le tableau 6.1 précise à quelles sections de la thèse se reporter pour le développement de chaque phase.

Phase	Chapitre/Section
Conception	Chapitre 6 Sections 6.2 et 6.3
Analyse *a priori*/ Analyse *a posteriori*	Chapitre 7 Sections 7.2 et 7.3
Validation interne	Chapitre 7 Section 7.4
Validation par l'enseignante	Chapitre 8 Section 8.1

TABLE 6.1 – Plan de l'étude des différentes phases

La phase de conception sera présenté dans la suite de ce chapitre et fera apparaître les choix faits et les raisons de ces choix. Cela permettra ensuite de préciser nos questions de recherche en fonction de ces choix.

Le chapitre 7 regroupera l'ensemble de nos analyses *a priori* et *a posteriori* des deux problèmes d'introduction conçus, prenant en compte les déroulements ayant eu lieu lors de l'expérimentation. Ensuite, nous présenterons les résultats de la validation interne.

Dans le chapitre 8 sera, entre autres, présenté la validation par l'enseignante.

Dans les sections suivantes, nous allons nous attarder sur la phase de conception de la séquence. Dans la section 6.2, nous allons donner les éléments importants qui ont constitué des étapes de la phase de conception. Ensuite, dans la section 6.3, nous exposerons la trame générale du scénario conçu.

6.2 Description de la phase de conception

Dans cette section, nous allons présenter des étapes importantes de la phase de conception de l'ingénierie didactique collaborative. Tout d'abord, nous exposerons les contraintes imposées par l'enseignante, puis les éléments de ses pratiques habituelles pris en compte pour arriver aux grandes lignes de notre scénario. Enfin, nous justifierons les principaux choix que nous avons retenus.

6.2.1 Des contraintes à respecter

Le fait qu'il s'agisse d'un travail collaboratif nécessite une prise en compte par le chercheur des contraintes de l'enseignante. Ces contraintes permettent notamment d'assurer la viabilité des tâches dans sa classe, ce sont tout du moins des conditions nécessaires. Lors de l'entretien avec Marie (cf. Annexe F.2) ainsi que lors des différents échanges que nous avons pu avoir, des contraintes non négociables sont apparues, liées aux différentes composantes du métier (Robert, 2008a). Nous voulons les citer ici.

a) Des contraintes institutionnelles

- La séquence doit respecter le programme. Cela fait aussi partie de nos contraintes.
- Le temps imparti à la séquence ne doit pas être modifié (ou tout du moins pas amplifié) par rapport aux années précédentes, soit environ trois semaines pour la séquence entière de calcul intégral et trois semaines pour l'ensemble des lois à densité, d'où un total de six semaines pour l'ensemble des deux chapitres.
- La partie introductive de la séquence ne doit pas excéder le temps qui lui était imparti les autres années, à savoir trois heures pour le calcul intégral et deux heures pour les lois à densité, soit un total de cinq heures.

b) Des contraintes cognitives

- Les élèves doivent savoir résoudre les mêmes tâches (sinon plus) que les années précédentes.
- L'objectif principal de l'enseignante reste de préparer ses élèves à l'épreuve du baccalauréat.

c) Des contraintes personnelles et médiatives

Marie n'acceptera, au final, de mener l'expérimentation dans sa classe que si elle adhère à notre scénario et qu'elle se sent capable de le mener à bien, du point de vue des contenus notamment.

d) Des contraintes matérielles

Marie est favorable à l'utilisation de logiciels, cependant la salle informatique de son établissement est difficilement accessible et seulement en demi-groupe. Pour cette raison la seule possibilité qu'elle envisage est d'utiliser l'ordinateur du bureau de l'enseignant en classe entière, donc un seul ordinateur pour toute la classe. L'écran peut être vidéoprojeté.

Pour Marie, l'essentiel est avant tout la réussite de ses élèves. Les contraintes qu'elle formule vont toutes dans ce sens.

6.2.2 Prise en compte des pratiques habituelles de Marie

Comme nous l'avons précisé dans la méthodologie, il n'est pas viable de proposer à un enseignant une séquence ou une tâche trop éloignée de ses pratiques habituelles. Pour cette raison, il a été important de partir des pratiques habituelles de Marie et de bien les prendre en compte dans la conception, pour ne pas faire des propositions trop éloignées de ses habitudes, en termes d'activités des élèves et de gestion de ces activités.

Nous allons présenter quelques éléments clés concernant ses enseignements sur le calcul intégral et sur les lois à densité de l'année 2013-2014. Les documents personnels de Marie sur sa séquence de calcul intégral de 2013-2014 se trouvent en Annexe F.3.1 et ceux sur les lois à densité (sans la loi normale) en annexe F.3.2.

a) Éléments sur les enseignements de Marie de l'année 2013-2014

Comme nous l'avons dit dans le chapitre 5, dans la section 5.4, Marie sépare les chapitres sur le calcul intégral et les lois à densité. Le calcul intégral est un prérequis pour les lois à densité. Le schéma de la figure 6.4 est une façon de résumer ces éléments.

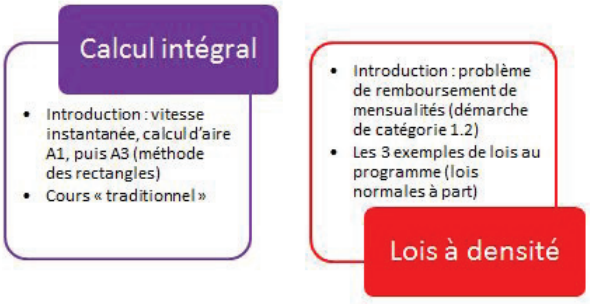


FIGURE 6.4 – Les deux chapitres

Les deux chapitres sont deux blocs bien distincts. Un des éléments importants sur lequel nous nous sommes appuyée est le fait qu'elle propose des *activités* d'introduction, dans les deux chapitres.

Pour l'introduction du calcul intégral, Marie commence par un problème sur les vitesses moyennes (nous ne disposons d'aucun document de l'enseignante) qui a pour but de motiver le besoin du calcul d'aire sous une courbe.

Elle propose ensuite une *activité* qui amène les élèves à calculer des aires sous une courbe de niveau A1 (fonctions affines par morceaux), puis ensuite à calculer l'aire sous une courbe proposée par leur voisin pour finalement arriver à un calcul d'aire sous la courbe de la fonction carré de niveau A3 (pour le moment le niveau A2 étant inconnu des élèves). Les élèves doivent trouver une méthode pour approcher ces aires (méthode des rectangles, méthode des trapèzes...), tout ceci dans un but de motiver ensuite l'introduction des primitives pour faire des calculs d'aires dans le niveau A2.

La suite de la séquence est assez traditionnelle, avec un début de cours sur les fonctions continues positives sur un intervalle borné, puis sur les fonctions continues de signe quelconque.

En ce qui concerne le chapitre sur les lois de probabilités à densité, l'*activité* d'introduction proposée par Marie a été présentée de façon assez détaillée dans la section 5.4. Pour rappel, il s'agit d'un problème sur les mensualités de remboursement. Le cours est ensuite dédié aux lois uniformes et exponentielles. Les lois normales ne sont abordées que dans le chapitre suivant.

b) Des introductions non satisfaisantes pour Marie

Il est important pour Marie de prévoir des *activités* d'introduction pour les chapitres dédiés au calcul intégral et aux lois à densité et pourtant, dans l'entretien et

d'autres échanges, nous avons pu identifier qu'elle n'était pas satisfaite des introductions qu'elle propose. Il lui semble difficile de motiver l'intérêt de calculer des aires sous une courbe. Elle utilise comme point de départ un problème de physique sur la vitesse moyenne ; cependant, elle trouve que ce type de problème ne fait pas écho chez les élèves, du fait notamment que ce soit des notions qui ne sont pas étudiées en physique.

Dans son entretien (Annexe F.2), Marie explique qu'elle n'est pas satisfaite de son problème introduction de la notion de fonction de densité, problème sur les mensualités de remboursement. Elle trouve notamment que le sujet n'intéresse pas les élèves. Au sujet de son *activité* d'introduction, elle dit :

> *Il y a juste à trouver un exemple intelligent, ce que je n'ai pas encore réussi à faire.* (lignes 353-354)

Elle mentionne notamment son manque de formation dans ce domaine. En parlant de l'ensemble de son enseignement sur les lois à densité, elle confie :

> *Je dois avouer que* [...] *je sais pas trop comment monter tout ça encore, soit dit en passant. Mais bon, ça va sûrement venir... [Rires]. Faut pas désespérer.* (lignes 618-620)

Au vu de son insatisfaction sur ses *activités* d'introduction, qui vont dans le sens de nos résultats de la Partie II, il nous a donc semblé raisonnable de lui proposer de nouvelles introductions de ces notions, en gardant toutefois à l'esprit certains éléments de ce qu'elle propose habituellement.

6.2.3 Une première proposition : une réorganisation et quelques modifications

Une fois ces différentes contraintes en tête, nous avons proposé nos premières idées à Marie pour savoir si nous continuons ou non dans cette direction. Nous voulons ici donner les éléments de la première ébauche proposée pour comprendre les étapes de la phase de conception.

Comme nous l'avons mis en évidence dans la section 5.9 du chapitre 5, nous trouvons des points de convergence dans les calculs d'aires rencontrés dans les deux chapitres, notamment parfois non exploités. L'idée de fusionner les deux chapitres nous a donc paru intéressante et notamment pour les deux introductions. Il s'agit de motiver l'introduction du calcul d'aire sous une courbe et donc le calcul intégral par le biais des probabilités.

Le schéma en figure 6.5 donne l'idée générale de la proposition faite.

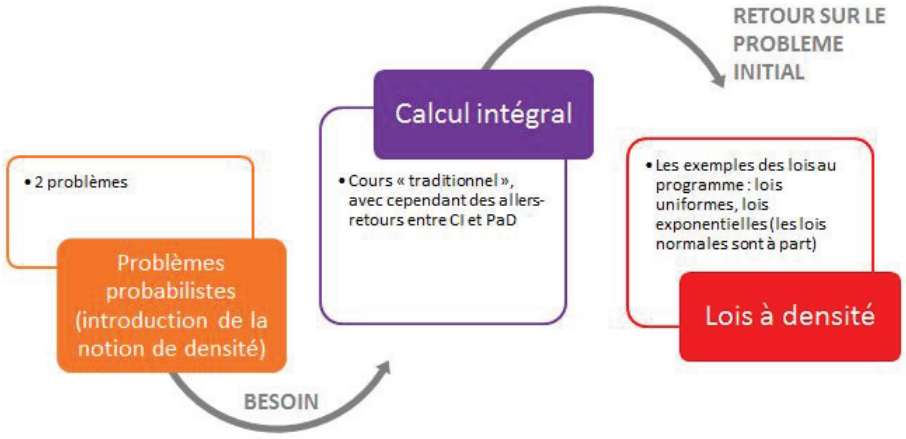


FIGURE 6.5 – Idée générale de la séquence

Nous cherchons à construire une séquence avec des situations permettant réellement de faire émerger la notion de fonction de densité, son utilité et ses propriétés. Le but est de créer des liens entre les probabilités à densité et le calcul intégral et non de les imposer.

Il n'est donc pas question de remplacer l'intégralité de ce qui se fait habituellement dans l'enseignement de ces deux chapitres mais plutôt de les combiner et d'apporter des compléments, en particulier sur l'introduction. Nous visons une ingénierie longue, dans le sens où elle constitue dans son ensemble toute une séquence couplant à la fois les lois à densité (sans aborder particulièrement les lois normales) et le calcul intégral (dans le cas des fonctions continues positives). L'idée, cependant, est de s'appuyer sur l'enseignement « traditionnel » et en particulier sur l'enseignement habituel de l'enseignante Marie. Nous avons proposé à Marie non pas de changer l'ensemble de ses deux ou trois chapitres mais essentiellement de les réorganiser, en mutualisant les séquences pour n'en faire qu'une. Il s'agit, plus précisément, de fusionner la partie du calcul intégral sur les fonctions continues et positives avec les lois à densité (le cours particulier sur les lois normales reste à part).

Tout d'abord, nous avons décidé d'introduire le calcul intégral comme outil nécessaire pour résoudre des problèmes de probabilités à densité : le calcul intégral n'est plus un prérequis pour aborder les lois à densité mais au contraire les probabilités sont un levier pour introduire le calcul intégral. L'idée est que l'introduction de la notion de densité se fasse sur plusieurs problèmes différents.

Dans un premier problème, nous cherchons à aboutir à une fonction à densité dont le calcul d'aire est de niveau A1, pour arriver à des calculs comme dans la première question de l'*activité* d'introduction de Marie en calcul intégral. Dans un second problème, la fonction de densité sera associée à un calcul d'aire de niveau A3.

L'enchaînement de ces deux problèmes permettra, dans un premier temps, d'introduire uniquement la notion de fonction de densité et dans un second temps, de consolider les connaissances dégagées sur la fonction de densité et de motiver le besoin du calcul d'aire sous une courbe (non associée à une fonction affine par mor-

ceaux). Ce second problème permettra, dans un second temps, de faire émerger des méthodes comme la méthode des rectangles, ensuite créera le besoin d'un outil mathématique pour calculer des aires sous les courbes, ce qui permettra d'introduire l'intégrale. On retrouvera dans ce problème, le même type d'activités que dans la question 3 de l'*activité* de Marie pour l'introduction du calcul intégral.

Ensuite, nous proposons à Marie de garder son cours habituel, en y incorporant cependant au fur et à mesure des exemples (ainsi que des définitions, des propriétés) issus des probabilités. Dans la séquence, le cours de calcul intégral sera restreint aux fonctions continues positives pour rester dans le même cadre que les fonctions de densité. Une partie du cours sera ensuite dédiée aux lois à densité (lois uniformes et exponentielles) pour mettre en avant les connaissances que les élèves doivent connaître en probabilités. Le seul problème de cohérence qui peut être souligné sur le fait de restreindre le calcul intégral ici aux fonctions continues positives concerne l'espérance de la loi uniforme sur $[a; b]$ dans le cas où a est négatif (ce qui implique que l'intégrale n'est pas celle d'une fonction positive sur $[a; b]$). Cependant, dans tous les exemples de lois uniformes rencontrés dans les manuels, les bornes de l'intervalle sont toujours positives, cela nous semble donc raisonnable de nous limiter à ce cas.

Au niveau des exercices, il n'est pas question d'y toucher mais simplement de mixer, encore ici, les exercices de probabilités et de calcul intégral tout au long de l'avancement. Enfin, un retour sur les lois à densité semble tout de même indispensable pour que les élèves identifient clairement ce qu'ils doivent savoir en probabilités.

Lors de la première réunion du 16 octobre 2014, cette proposition générale a été exposée à Marie, qui n'avait pas encore accepté de participer au projet. La proposition de mutualiser les deux séquences a été bien accueillie et les premières idées, encore vagues, sur les problèmes introductifs ont été validées.

Finalement à l'issue de la première réunion, le projet a fait consensus entre Marie et nous, et nous nous sommes alors lancées réellement dans le travail collaboratif. Le principal du travail de conception s'est donc essentiellement concentré sur les séances d'introduction.

Suite à nos propositions laissant parfois plusieurs choix (de logiciels notamment), Marie nous faisait des retours pour améliorer, reformuler, valider ou invalider les tâches et les scénarios prévus, soit lors de réunions, soit dans des échanges de mails. Nous n'allons pas présenter ici l'évolution des choix des tâches. Nous allons cependant présenter quelques éléments de justification des choix que nous avons fait pour les tâches d'introduction.

6.2.4 Principaux choix retenus pour la conception des tâches et justification de ces choix

Au vu de ce qui ressort de nos Partie I et Partie II, nous avons dégagé plusieurs éléments que nous voulons mettre en œuvre dans nos tâches d'introduction. Nous souhaitons, dans cette section, présenter les choix retenus et les justifier. Tout d'abord, le point central est de donner du sens à la notion de fonction de densité et, dans ce but, de proposer des problèmes d'introduction qui fassent réellement

émerger le besoin de cette fonction et les propriétés qu'elle doit vérifier.

Pour permettre à la fois l'introduction de la fonction de densité et du calcul intégral, il nous a semblé important de concevoir deux problèmes d'introduction, pour ne pas avoir trop de nouveautés à gérer en même temps. Le premier problème a pour objectif d'introduire la fonction de densité en lui donnant du sens, en aboutissant à une fonction dont le calcul d'aire sous la courbe est du niveau A1, ce qui ne pose pas de problème aux élèves. Ensuite, le second problème aboutit à une fonction de densité non affine par morceaux, ce qui motive le besoin de savoir calculer ce type d'aire. Le but final de ce problème est de déterminer une valeur approchée de la probabilité et donc de l'aire sous la courbe de niveau A3 en utilisant la méthode souhaitée.

Un autre choix est de proposer une loi sur un ensemble non borné, dès l'introduction des notions, pour ne pas se restreindre à des calculs sur des intervalles bornés et ensuite faire des calculs sur un ensemble non borné sans justifier pourquoi nous pouvons le faire. Cependant, pour cela, il nous semble indispensable d'anticiper ces questions d'aire finie d'un domaine infini en amont de la séquence.

Nous voulons aussi utiliser une approche passant par la statistique qui nous semble un bon levier pour construire le lien entre probabilités et analyse. Nous voulons garder les deux catégories des démarches rencontrées dans les *activités* des manuels, c'est-à-dire proposer un problème avec une démarche partant des probabilités pour arriver aux probabilités en faisant un crochet en statistique par le biais de la simulation et un autre problème partant de la statistique pour arriver aux probabilités, par inférence (cf. section 5.2.4). Dans ces catégories, nous avons repéré des sous-catégories, certaines plus pertinentes que d'autres. Pour cette raison, l'un des problèmes d'introduction sera de la catégorie 2.2 et l'autre de la catégorie 1.1 (cf. table 5.7, p. 167), pour avoir une approche par la simulation et une autre par la statistique inférentielle. Dans tous les cas, il y aura un passage par la notion d'histogramme. Comme nous l'avons vu dans le chapitre 5, l'histogramme n'est pas un objet mathématique disponible pour les élèves. Pour cette raison, il nous semble incontournable en amont de la séquence de faire à nouveau une « première rencontre » avec cette représentation graphique, qui est indispensable pour faire émerger la fonction de densité et ses propriétés. Il est aussi primordial que les élèves aient réellement à trouver la courbe de densité et donc l'expression de la fonction de densité, pour qu'ils soient confrontés à ses propriétés.

Il nous a aussi semblé important de ne pas « fabriquer » des données, mais bien de prendre des données réelles (qui semblent suivre des lois dont les fonctions de densité sont abordables par les élèves de terminale S), ce qui n'est pas forcément simple à trouver.

Pour le choix des lois, nous avons écarté la loi uniforme, qui ne nous semble pas pertinente pour introduire la notion de densité, comme nous l'avons montré dans la section 2.2.2. Nous avons donc pensé, pour le premier problème, à une loi associée à une fonction de densité affine sur un ensemble fermé borné. Au vu de nos critères exposés avant, la loi du second problème doit donc être à valeurs sur un ensemble non borné, dont l'aire sous la courbe est de niveau A3 et dont les élèves peuvent déterminer l'expression de la fonction. Nous avons choisi une loi exponentielle. Il est donc important que les fonctions de la famille $x \mapsto e^{-kx}$ soient étudiées plus tôt

dans l'année, pour qu'elles soient connues des élèves.

La démarche de modélisation va alors être un moteur dans ces introductions, ce que nous n'avons pas du tout rencontré dans les *activités* des manuels. Ce choix de proposer des problèmes de modélisation va avoir pour avantage de ne pas guider les activités des élèves et de faire en sorte de construire la notion au fur et à mesure par le collectif classe. Le collectif va jouer un rôle très important. Nous ne nous attendons pas à ce que chaque élève réussisse individuellement à faire émerger la notion de fonction de densité mais que le collectif le fasse. Les échanges entre élèves et les interactions avec l'enseignante vont pouvoir influencer leurs activités (Robert, 2008c). Bien que les élèves n'aient pas rencontré de problèmes de modélisation dans l'année, Marie est habituée à laisser parler ses élèves, sans trop intervenir. Ce type de tâches convient donc à ses pratiques.

Enfin, les logiciels nous semblent être des outils indispensables, notamment pour la construction rapide d'histogrammes. Il a donc fallu choisir, notamment entre *GeoGebra* et *R*. Bien que le second soit fortement préconisé dans le document *Ressources*, il est pour autant peu connu des enseignants et demande un investissement certain pour se l'approprier. Pour cette raison, Marie n'a pas jugé nécessaire de s'investir dans l'utilisation de ce logiciel, même s'il est vrai qu'il offre des possibilités plus intéressantes que *GeoGebra* telles que le fait de pouvoir faire des histogrammes avec des amplitudes non constantes. Le logiciel *GeoGebra* a l'avantage d'être plus « transparent » pour les élèves. Ils comprennent, même sans manipuler, ce que fait le logiciel, ce qui n'est pas le cas avec *R*. Nous avons donc opté pour le logiciel *GeoGebra*.

Après avoir présenté tous les points importants pris en compte dans la conception de la séquence, nous allons maintenant présenter la trame générale du scénario de la séquence.

6.3 Présentation de la trame générale du scénario de la séquence

Nous venons de présenter le cheminement qui a mené à nos choix en matière de contenus mathématiques. Nous allons, dans cette section, décrire la trame générale du scénario de la séquence finalement conçue, après avoir pris en considération les choix retenus. Il s'agit ici du « produit fini » des réflexions conjointes. Il n'est pas question ici de faire une analyse *a priori* des tâches mais simplement de donner les idées générales et les objectifs des problèmes d'introduction, en les insérant dans la séquence entière. Les analyses approfondies se trouvent dans le chapitre 7.

6.3.1 Avant la séquence

a) Rencontre des fonctions de référence probabilistes

Il s'agit, avant la séquence, de faire étudier aux élèves les différentes fonctions de référence probabilistes, sans les nommer comme telles. Nous entendons par là les fonctions $x \mapsto e^{-kx}$ et $x \mapsto e^{-x^2/2}$, à constante multiplicative près. Ceci se fait notamment dans la séquence sur la fonction exponentielle. L'objectif est que les

élèves aient connaissance de ces fonctions afin de les reconnaître en temps voulu. Cela servira lors de l'introduction des lois à densité.

b) Première rencontre d'une aire finie d'un domaine infini

Nous avons vu, dans la Partie II, que dans les manuels, par exemple, l'intégrale est définie sur un intervalle fermé borné. Seuls quelques exercices traitent ensuite de la limite d'une intégrale quand la borne supérieure tend vers $+\infty$ mais cela est toujours travaillé comme une étude de fonction ou de suite sans que la question de l'aire finie d'un domaine infini soit véritablement abordée. Le chapitre sur les lois à densité voit alors ensuite apparaître des limites d'intégrales sans notamment que l'on se pose la question de l'existence. Nous avons précisé notre volonté d'introduire une loi définie sur un intervalle non borné, c'est pour cette raison que nous voulons faire ce travail en amont.

L'objectif ici est de se poser la question de l'existence de domaines infinis qui ont une aire finie en amont de la séquence, afin que les élèves puissent s'appuyer sur cette connaissance lors de l'introduction de la séquence. Une telle réflexion est possible à travers une étude de suite, les suites étant abordées en début d'année. Le choix a été fait de proposer ce travail en devoir maison. L'exercice en question est l'exercice 2 du devoir maison n° 5 (DM5) qui se trouve en annexe G.1. Il s'agit, dans un premier temps, d'étudier la suite $u_n = \sum_{i=0}^{n} \frac{1}{2^i}$ et, dans un second temps, de faire le lien avec la somme totale des aires de la suite de rectangles de base 1 et de hauteur 1, 1/4, 1/9, 1/16... Il s'agit ensuite d'interpréter graphiquement cette suite.

c) Nouvelle première rencontre de l'histogramme et découverte de la loi uniforme

Le devoir maison n° 6 (DM6) (cf. Annexe G.2) est, nous semble-t-il, crucial pour le bon déroulement des séances d'introduction. Le premier exercice a pour but de réintroduire la notion d'histogramme de fréquences. Les histogrammes permettront l'introduction de la fonction de densité, c'est pourquoi ils ne doivent pas constituer un obstacle dans leur compréhension au moment du début de la séquence. Les élèves – bien que l'histogramme soit au programme de la cinquième à la seconde – n'auront peut-être jamais véritablement travaillé cette notion. De plus, la notion de densité de fréquence n'aura certainement pas été vue. Ici le but est d'introduire l'histogramme en passant par la compréhension de la densité de fréquence. Il est important que les élèves comprennent que l'axe des ordonnées d'un histogramme représente la densité de fréquence. Cela sera un appui pour la suite. Le deuxième exercice est plus orienté sur la lecture d'un histogramme. Il permet d'exploiter ce qui vient d'être fait dans le premier exercice.

La partie B du troisième exercice est dédié à la loi uniforme sur $[0; 7]$. Cet exercice est inspiré de situations proposées dans Barbazo *et al.* (2003). Comme nous l'avons dit plus tôt, la loi uniforme ne semble pas être le meilleur exemple pour introduire la notion de fonction de densité, car les élèves savent résoudre intuitivement des problèmes faisant intervenir des lois uniformes. C'est la raison pour laquelle nous pensons intéressant que les élèves traitent seuls cette loi en amont de la séquence.

Dans la partie B, il s'agit d'une puce électronique qui se déplace à vitesse constante sur le « pourtour » d'un triangle (on parlera du problème de la puce). Il s'agit alors de s'intéresser à la variable aléatoire correspondant à la position d'arrêt de la puce. Les élèves doivent déterminer des probabilités, dans des cas particuliers de triangles, notamment la probabilité que la puce s'arrête sur un sommet. Après avoir ramené cette expérience aléatoire à celle du choix au hasard d'un nombre entre 0 et 7, il est proposé de faire des simulations de cette nouvelle expérience aléatoire. Des histogrammes de fréquences sont proposés, notamment quatre histogrammes pour un échantillon de taille 10 000. C'est l'occasion de parler de fluctuations d'échantillonnage, tout en faisant remarquer qu'il « *semble y avoir une « courbe de tendance », c'est-à-dire une courbe qui « lisse » l'histogramme (la même pour les quatre)* ». Les objectifs de cet exercice sont multiples :

- rencontrer la loi uniforme sur un segment (sans pour autant la nommer),
- faire émerger le « problème » de la probabilité d'une valeur isolée (probabilité nulle d'un évènement possible),
- parler de fluctuation d'échantillonnage à travers les histogrammes et introduire une première idée de « courbe de tendance », qui tend à « stabiliser » le haut des histogrammes.

6.3.2 Introduction de la notion de fonction de densité (1) : le problème de la rencontre

Le problème de la rencontre correspond au premier problème introductif de la séquence. Nous rappelons qu'ici nous ne présentons que l'idée générale du problème et ses objectifs. Nous avons voulu commencer par un problème qui aboutit à une fonction de densité affine, permettant alors des calculs de probabilités, à l'aide des formules classiques de calculs d'aires (niveau A1). Nous avons donc décidé de débuter avec la loi de Xénakis, associée à la fonction affine décroissante $f(x) = -2x + 2$. La variable aléatoire associée correspond à la différence absolue (écart) de deux variables aléatoires indépendantes de loi uniforme sur $[0; 1]$. Plutôt que d'introduire cette loi directement comme une différence absolue de deux variables aléatoires uniformes (sachant que cette terminologie n'a pas encore été introduite) de façon abstraite, ou dans le cadre géométrique, comme la distance entre deux points choisis au hasard sur le segment $[0; 1]$, il nous a semblé plus motivant pour les élèves de trouver un problème issu de la réalité, ou tout du moins « concret ». L'énoncé est le suivant :

> Karine et Olivier décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Ils peuvent arriver à tout moment entre 7h et 8h. Que peut-on dire du temps d'attente du premier arrivé ?

L'énoncé est très ouvert, pour laisser aux élèves l'initiative de leur démarche. Il s'agit ici de mettre en place une démarche de modélisation. L'objectif implicite est de déterminer la loi de la variable aléatoire correspondant à la différence absolue (distance) entre les deux temps d'arrivée, que l'on peut modéliser par deux lois uni-

formes sur [7; 8] ou [0; 1], et donc de déterminer la fonction de densité associée à cette variable aléatoire. Ce problème est inspiré d'un exercice du document *Ressources* (MENJVA & DGESCO, 2012) qui en propose cependant une approche totalement différente. Après avoir choisi un modèle pour les temps d'arrivée, l'idée ici est de faire des simulations du temps d'attente et de construire des histogrammes pour ensuite déterminer l'expression de la fonction de densité. Il s'agit d'une démarche de type 2.2 (cf. table 5.7, p. 167) : la loi est inconnue mais est une combinaison de lois connues, il est donc possible à l'aide de simulations de déterminer la loi cherchée. Une fois seulement la fonction de densité déterminée, une feuille d'exercice est proposée aux élèves, dans laquelle il est demandé de déterminer différentes probabilités (cf. Annexe G.3).

Les objectifs de ce problème sont de :
- mettre en place une démarche de modélisation,
- passer par la simulation et l'histogramme de fréquences pour faire émerger le besoin d'une courbe de tendance associée à une fonction, que l'on appellera plus tard fonction de densité,
- faire émerger les propriétés d'une telle fonction, pour pouvoir déterminer son expression,
- faire des calculs de probabilités à l'aide de calculs d'aires sous une courbe (de niveau A1).

Une phase d'institutionnalisation est prévue à la fin de ce problème pour introduire le vocabulaire notamment variable aléatoire continue et fonction de densité de probabilité, ainsi que les propriétés de la fonction de densité qui sont censées émerger de ce travail.

6.3.3 Introduction de la notion de fonction de densité (2) et du calcul intégral : le problème du volcan Aso

Le problème du volcan Aso est la deuxième tâche d'introduction. Il permet, dans un premier temps, de réinvestir les connaissances nouvelles qui ont émergées du problème de la rencontre et donc de consolider le référentiel théorique en construction. C'est ce qui nous intéresse tout particulièrement. Cependant, la démarche de modélisation est différente, elle est ici de la catégorie 1.1 (cf. table 5.7, p. 167) : nous disposons des données d'un échantillon (SD) et nous cherchons à faire une estimation des probabilités sur la population entière, ou plutôt ici sur le futur (PaD). Il s'agit ici de partir de données réelles, les années de début d'éruption du volcan Aso entre le XIII^{e} et le XIX^{e} siècle (cf. Annexe G.4), pour essayer d'estimer la probabilité que la prochaine éruption ait lieu dans 3 ans et que la suivante ait lieu au cours de l'année 2030. Ici le temps d'attente entre deux éruptions peut être modélisé par une loi exponentielle dont le paramètre peut varier légèrement autour de 0,1. À partir de l'histogramme, en prenant en considération les différentes propriétés que doit vérifier la fonction de densité, les élèves doivent déterminer l'expression d'une fonction de densité qui semble convenir, ne connaissant bien entendu pas la loi exponentielle. C'est ici que l'on voit l'intérêt d'avoir déjà étudié ces fonctions en amont. Ne sachant pas calculer l'aire sous une courbe non affine, l'outil informatique est aussi indispensable. Il s'agit d'aires sur un domaine infini. Le logiciel ne donnera que des valeurs

approchées mais cela peut tout de même suffire ici pour invalider des fonctions et en valider *a priori*. Les calculs de probabilités s'effectuent ensuite aussi avec l'outil informatique.

Dans un second temps, apparaît le second objectif de ce problème, qui est de motiver le besoin du calcul d'aire sous une courbe à la main cette fois-ci. Ici, il devient nécessaire de savoir calculer l'aire sous une courbe (représentant une fonction non affine par morceaux). Cela permet alors ensuite de lancer un travail de calcul d'aires sous la courbe de la fonction de densité pour déterminer la probabilité, ou tout du moins une valeur approchée, que le temps d'attente entre deux éruptions soit entre 0 et 20 ans. Le but est alors que les élèves cherchent une méthode pour approcher la valeur de l'aire en question : méthode des rectangles, méthode des trapèzes... Il s'agit ici du même type d'*activités* que Marie propose habituellement.

6.3.4 Un cours et des exercices articulant calcul intégral et probabilités à densité

Comme nous l'avons annoncé dans la section 6.2.2, le but n'est pas de changer le contenu du cours ni des exercices mais de les réorganiser. En annexe E.1 se trouve la prise de notes de cours se rapportant à cette séquence d'un élève de la classe. Les fiches d'exercices se trouvent quant à elles en annexe G.5. En comparant avec les cours et les exercices de Marie de l'année précédente (cf. Annexes F.3.1 et F.3.2), nous pouvons vraiment remarquer le peu de différences au niveau des contenus mais l'énorme différence en termes d'organisation des notions. Maintenant il ne s'agit plus que d'une seule séquence si l'on s'intéresse au calcul intégral de fonctions continues positives et aux lois à densité. Au niveau des évaluations, nous pouvons préciser qu'il s'agit d'exercices extraits de différents sujets de baccalauréat, car comme nous l'avons précisé dans les contraintes, le but reste de préparer les élèves à l'examen.

6.3.5 Un autre devoir maison

Un autre devoir maison (DM7) a été conçu pour permettre une nouvelle fois de passer par un histogramme afin d'introduire une nouvelle loi : la loi normale (cf. Annexe G.6). Dans ce devoir maison, il s'agit d'étudier les tailles des pères de 1078 familles à la fin du XIX^{e} siècle. Ce sont des données réelles, récoltées par le statisticien Pearson. Le but de ce devoir n'est pas d'utiliser le logiciel *GeoGebra* pour construire l'histogramme (il est déjà donné) ni de déterminer l'expression de la fonction de densité, étant donné qu'elle est difficile à trouver. Pour pallier cette difficulté, trois propositions de fonctions sont faites et les élèves doivent choisir celle qui leur semble la plus appropriée. C'est surtout ici l'occasion de rencontrer une nouvelle loi, la loi normale, et de déterminer des probabilités associées, sachant que nous sommes dans le cas de calculs d'aire de niveau A3 (même après avoir travaillé le niveau A2). Cela permet d'ailleurs de mettre en évidence que certaines fonctions continues positives peuvent ne pas avoir de primitives explicites. Il est donc proposé d'utiliser la méthode des rectangles pour déterminer des probabilités. Un algorithme de la méthode des rectangles est aussi à étudier. Et enfin, des valeurs de quelques intégrales (données par la calculatrice) sont à utiliser pour déterminer

d'autres probabilités. Il s'agit bien ici d'un exercice qui essaie de lier probabilités à densité et calcul intégral autour du niveau A3 et notamment plusieurs méthodes sont proposées, que l'on retrouve habituellement dans la séquence de calcul intégral. C'est aussi l'occasion de travailler sur la symétrie de la fonction de densité associée à la loi normale. Ce devoir maison, donné en fin de séquence, permet d'introduire la loi normale (généralisée) en amont de la loi normale centrée réduite, bien que ce ne soit pas l'approche préconisée dans le programme.

6.4 Précisions sur la question de recherche

Dans ce chapitre, nous avons tout d'abord présenté notre méthodologie que nous appelons une ingénierie didactique collaborative. Cela nous a ensuite amené à prendre différents points en compte, à savoir les analyses préalables des deux premières parties ainsi que les contraintes liées au travail collaboratif, pour retenir des choix au niveau des contenus mathématiques mais aussi de la gestion des activités pour la conception des tâches. Après avoir justifié nos choix, nous avons proposé la trame générale du scénario de la séquence entière. Bien entendu, nous allons restreindre notre étude aux problèmes d'introduction, mais il nous a semblé important de les considérer dans un tout. À travers les deux problèmes conçus, nous voulons tester les choix retenus.

Rappelons la question de recherche QR3 :

> **QR3** : Quels éléments didactiques peut-on prendre en compte pour concevoir et mettre en œuvre effectivement des tâches qui permettent aux élèves de construire la notion de fonction de densité ?

Les choix que nous avons retenus et que nous avons présentés dans ce chapitre sont donc des éléments didactiques, principalement au niveau des contenus mathématiques mais aussi au niveau des artefacts par exemple, que nous voulons tester au travers des analyses *a priori* et *a posteriori*.

6.4.1 Éléments théoriques complémentaires

Compte tenu du fait que les deux problèmes conçus sont des problèmes de modélisation, dans lesquels le collectif classe va jouer un rôle important, il nous semble nécessaire de faire un complément théorique avant de revenir aux questions de recherche. Dans le chapitre 2, nous avons distingué trois niveaux d'ETM : ETM de référence, ETM idoine et ETM personnel. Dans cette troisième partie, nous souhaitons étudier l'ETM idoine proposé dans la classe et plus particulièrement l'ETM idoine effectif, qui s'est effectivement construit dans la classe. Cependant, cet ETM idoine effectif se construit petit à petit au fur et à mesure des tâches données et des activités effectives des élèves. Il se met donc en place à l'aide de l'ETM idoine prévu par l'enseignant, par l'ETM personnel de l'enseignant et des ETM personnels de chaque élève. En observant une classe notamment par le biais des mises en

commun, nous n'avons pas accès aux ETM personnels des élèves (ou seulement de manière très ponctuelle) mais à un ETM collectif, où chaque élève peut apporter des éléments de son ETM personnel, qui permet de faire avancer l'activité et petit à petit de façonner l'ETM idoine effectif. Bien entendu, ce n'est pas seulement l'ETM collectif de la classe qui permet de construire l'ETM idoine, l'enseignant aussi y contribue par ses aides, ses rappels...

6.4.2 Nos nouvelles questions de recherche

À travers les choix de tâches que nous avons faits, nous faisons l'hypothèse de recherche suivante :

Il existe des tâches qui permettent réellement aux élèves de construire la notion de fonction de densité grâce à des circulations riches au sein des ETM des trois sous-domaines (PaD, CI et SD).

Pour aborder la question de recherche QR3 et l'hypothèse de recherche que nous venons de formuler, nous avons décidé de considérer deux groupes de sous-questions. Le premier groupe de questions est relatif au travail mathématique entre les trois ETM (ETM_{PaD}, ETM_{CI} et ETM_{SD}). Il se situe au niveau du savoir mathématique mais aussi des activités dans la classe. Le but de ces questions est de vérifier que l'objectif d'apprentissage est bien rempli et de valider les choix retenus. Le deuxième groupe de questions tourne quant à lui autour de la répartition des rôles entre élèves et enseignante dans ce travail mathématique et du rôle particulier de l'enseignante. Son objectif est de vérifier que l'objectif d'apprentissage visé est bien atteint (par les élèves) et de déterminer d'éventuelles conditions supplémentaires, qui paraissent nécessaires après expérimentation, pour assurer le bon déroulement des activités.

Les questions relatives au travail mathématique sont les suivantes :

- QR3a : Est-ce que les activités, attendues et possibles, des élèves associées à ces tâches (les deux problèmes introductifs) permettent réellement d'introduire la notion de fonction de densité, et en plus le calcul intégral ?

- QR3b : Est-ce que le référentiel théorique lié à la fonction de densité se construit effectivement à l'aide de circulations riches entre les trois ETM en jeu ? Cela engendre-t-il un travail complet ?

- QR3c : Peut-on identifier des éléments (passages obligés, référence à l'ancien...) qui semblent indispensables pour la bonne réalisation de ce travail mathématique ?

Celles relatives plutôt à la gestion et à la répartition des rôles entre élèves et enseignante sont les suivantes :

- QR3d : Est-ce que l'ETM collectif permet seul de construire l'ETM idoine

effectif ? Qui, des élèves ou de l'enseignante, a la charge des changements de domaine et de dimensions dans les ETM dans la construction du référentiel théorique relatif aux probabilités à densité ?

– QR3e : Quel est le rôle de l'enseignante dans le déroulement des activités des élèves ?

– QR3f : Quelles contraintes et conditions a-t-on pu observer qui permettent à la classe d'arriver à l'objectif d'apprentissage visé ?

Après avoir précisé nos questions de recherche en lien avec les tâches conçues, nous allons passer aux analyses relatives à ces deux problèmes introductifs, dans le chapitre suivant.

Chapitre 7

Analyses des scénarios et des déroulements

Dans le chapitre précédent, nous avons défini notre méthodologie de type ingénierie didactique collaborative, qui reprend notamment les phases de l'ingénierie didactique, en y ajoutant un travail collaboratif avec une enseignante. Ce chapitre a aussi été l'occasion de justifier les choix retenus pour la conception des tâches d'introduction et de présenter la trame générale du scénario d'enseignement qui a été conçu en collaboration avec l'enseignante Marie.

Nous sommes ensuite arrivée à de nouvelles questions de recherche focalisées sur le problème de la rencontre et le problème du volcan Aso, tous deux ayant pour objectif d'introduire la notion de fonction de densité et de motiver le besoin du calcul d'aire sous une courbe, objectif qui s'est ajouté par rapport aux objectifs de départ. Comme nous l'avons présenté à la fin du chapitre précédent, les deux groupes de questions qui précisent notre question de recherche QR3 sont de deux types : d'un côté, des questions sur le travail mathématique entre les trois ETM, celui des probabilités à densité (PaD), celui du calcul intégral (CI) et celui de la statistique descriptive (SD), et de l'autre, des questions sur la répartition des rôles entre élèves et enseignante et sur la gestion de ce travail mathématique.

À travers cette méthodologie d'ingénierie didactique collaborative, nous voulons prouver un « théorème » d'existence et, de ce fait, valider les choix retenus pour la conception de ces tâches. Nous voulons montrer que les deux problèmes sont effectivement réalisables en classe et permettent la mise en place d'activités mathématiques qui amènent à la construction de connaissances sur la fonction de densité, sans que l'enseignante « n'en dise trop ». En plus de ce « théorème » d'existence, nous voulons mettre en évidence par quels moyens ces activités sont possibles, du point de vue du travail mathématique en jeu mais aussi du point de vue de la gestion du déroulement par l'enseignante.

L'objectif de ce chapitre est de construire une méthodologie d'analyses, des problèmes et de leur déroulement, capable de nous donner des renseignements sur les points qui nous intéressent dans nos questions de recherche pour ensuite la mettre en œuvre. Ensuite, nous présenterons les analyses *a priori* et *a posteriori* de ces problèmes suivant cette méthodologie, ce qui nous permettra en conclusion de dégager des résultats. Bien entendu, les analyses *a priori* sont solidaires de la conception

des tâches. Cependant, pour faciliter la lecture des analyses, il nous a semblé plus adapté de regrouper les analyses *a priori* et *a posteriori*.

Dans ce chapitre, nous adapterons l'analyse de tâches à des tâches de modélisation, ce qui nous permettra d'analyser les deux problèmes d'un point de vue global. Plus localement, nous essaierons de mettre en évidence les sous-activités en jeu et comment se font les circulations entre les ETM associés à la statistique descriptive, aux probabilités à densité et au calcul intégral et suivant quelles dimensions du modèle des ETM ; ceci dans un but de se demander si ce travail mathématique permet effectivement la construction du référentiel théorique de la fonction de densité. Nous tâcherons en parallèle d'identifier *a posteriori* qui, des élèves ou de l'enseignante, prend réellement en charge les activités et comment l'enseignante gère la mise en place de ce travail mathématique.

Dans la section 7.1, nous présenterons la méthodologie construite pour nos analyses *a priori* et *a posteriori* pour ensuite détailler nos analyses et donner nos premiers résultats relatifs au problème de la rencontre (section 7.2) et au problème du volcan Aso (section 7.3). Nous pourrons, suite à ces analyses, présenter les résultats relatifs à nos questions de recherche (section 7.4). Enfin, nous finirons ce chapitre par une discussion méthodologique autour des outils d'analyses utilisés (section 7.5).

7.1 Méthodologie des analyses

Notre méthodologie d'ingénierie didactique implique la nécessité de faire une confrontation entre analyse *a priori* et *a posteriori* dans le but de valider (ou non) les propositions faites donc, ici, les choix retenus dans les tâches conçues. Nous cherchons donc à construire une méthodologie avec des grilles d'analyses adaptées à nos questions de recherche et à notre cadre théorique afin de répondre à nos questions.

Dans le cadre de la théorie de l'activité dans lequel nous nous situons (cf. chapitre 2), Robert (dans Vandebrouck, 2008) propose une méthodologie pour étudier *a priori* les scénarios[24] et *a posteriori* les déroulements. Plusieurs exemples sont ensuite présentés. Il nous paraît adapté, ici, de nous placer dans ce contexte méthodologique. Cependant, des adaptations seront nécessaires. Nous allons présenter et justifier nos choix.

7.1.1 Précisions sur l'expérimentation

Avant de commencer, nous voulons donner des précisions sur l'expérimentation, pour la resituer dans son contexte, et sur les données dont nous disposons pour nos analyses.

24. Un scénario est l'ensemble ordonné de cours, d'exercices, d'évaluations (Robert & Vandebrouck, 2014). Nous pouvons ensuite parler des éléments du scénarios. Parfois, pour aller plus vite, nous utilisons scénario pour élément de scénario.

a) L'expérimentation

Le planning de la séquence entière se trouve en annexe H. Nous y trouvons les dates des séances ainsi que les exercices, problèmes ou cours abordés lors de ces séances. Les deux problèmes d'introduction qui nous intéressent ont été traités sur quatre séances consécutives (les quatre premières de la séquence). Le problème de la rencontre s'est déroulé sur deux séances :

- La séance 1, d'une durée de 1h50 ;
- La séance 2, sur une durée de 35 min (le reste de la séance a porté sur un ancien chapitre).

Le second problème, celui du volcan Aso, s'est aussi déroulé sur deux séances :

- La séance 3, d'une durée de 55 minutes ;
- La séance 4, sur une durée d'environ 1h20 (le reste de la séance a porté sur l'écriture du cours).

Le tableau 7.1 présente de façon plus détaillée ces séances et les séances antérieures nous concernant.

En amont		
Novembre/ Décembre 2014		Études des fonctions de référence probabilistes ($x \mapsto e^{-kx}$, $x \mapsto e^{-x^2/2}$)
Janvier 2015	DM5	Éléments de correction : Réflexion autour de l'aire d'un domaine infini, qui peut être finie
11 mars 2015	DM6	Éléments de correction : Révisions sur les histogrammes : densité de fréquence, classes... Première rencontre de la loi uniforme sur un segment (problème de la puce)
Séances d'introduction		
14 mars 2015	Séance 1 (1h50)	Problème de la rencontre : Loi de Xénakis (différence de deux lois uniformes)
17 mars 2015	Séance 2 (35min)	Institutionnalisation : Fonction de densité, ses propriétés. Exercice 1 du problème de la rencontre
18 mars 2015	Séance 3 (50min)	Problème du volcan Aso : Loi exponentielle
20 mars 2015	Séance 4 (1h30)	Problème du volcan Aso : Recherche d'une méthode d'approximation d'aire sous une courbe

Table 7.1 – Présentation des séances en amont et d'introduction de la séquence

Dans nos analyses, nous allons donc nous focaliser sur les séances numérotées de 1 à 4. Nous ne nous intéresserons pas forcément aux séances entières. Par exemple, lors de la séance 2, seule la partie relative au problème de la rencontre sera étudiée et même plus précisément seulement la phase d'institutionnalisation.

b) Les données sur les séances

Ces quatre séances ont été enregistrées et ont ensuite été transcrites dans leur quasi intégralité, pour celles qui traitent des deux problèmes. Seule la correction de l'exercice 1 de la situation de la rencontre n'a pas été transcrite. Les transcriptions de ces séances se trouvent en annexe I. Pour la transcription, nous n'avons pas suivi de codage particulier au niveau des temps de pause dans le discours par exemple car nous ne voulons pas faire une analyse poussée du discours. Nous avons en revanche incorporé les photos du tableau dans la transcription. Nous avons numéroté les élèves suivant leur positionnement dans la classe. Cela permettra, dans un second temps, de considérer les élèves en individuel si cela s'avère pertinent. Quand un élève n'a pas pu être identifié, il est noté « ? » ou « un élève ». Nous n'avons accès, à travers ces enregistrements et ces transcriptions, en grande majorité, qu'aux échanges des phases collectives. Nous avons parfois pu extraire des enregistrements des moments de réflexion en petits groupes – mais ces moments sont rares – et nous avons aussi pu les enrichir de nos notes personnelles. Cependant, au vu de nos questions de recherche, ce sont les échanges en classe entière qui attirent particulièrement notre attention. Nous avons complété nos données en prenant des photos de productions d'élèves. Enfin, nous disposons des notes de cours de deux élèves de la classe, qui se trouvent en annexe I.5.1.

7.1.2 Base de notre méthodologie

Dans le chapitre précédent (chapitre 6), nous avons présenté la trame générale du scénario de la séquence conçue. Nos analyses vont seulement se focaliser sur des éléments du scénario, à savoir les tâches du début de la séquence : le problème de la rencontre et celui du volcan Aso.

La méthodologie d'analyse de scénarios (ou ici d'éléments de scénarios) comporte plusieurs aspects (Robert & Vandebrouck, 2014). Tout d'abord, il s'agit de l'analyse du scénario prévu (à proposer). On peut alors y associer l'itinéraire cognitif prévu, c'est-à-dire la conversion par le chercheur de la tâche (au sens large) du scénario prévu en activités attendues, compte tenu des tâches et du déroulement prévus. Il s'agit finalement d'une analyse *a priori* des tâches qui permet de cerner ce qui est en jeu du point de vue des mathématiques à utiliser et des activités à mettre en œuvre. Ensuite, une analyse *a posteriori* permet de décrire le scénario effectif (reconstitué) grâce au déroulement. De même, on peut y associer l'itinéraire cognitif effectif (possible), compte tenu cette fois du déroulement effectif. Cette étape est facilitée par la description du scénario prévu, en mettant en regard les deux scénarios.

Cependant, cette méthodologie est bien adaptée pour des tâches simples et isolées ou encore des tâches complexes, mais qui restent fermées. Dans notre cas, nous avons affaire à des tâches de modélisation, qui sont donc des tâches ouvertes et qui

nécessitent une démarche particulière : celle de modélisation. Comme le précisent Cazes & Vandebrouck (2014) :

> Analyser une tâche de modélisation mathématique du point de vue d'un chercheur n'est pas une chose aisée. Les élèves doivent avoir des connaissances mathématiques disponibles (Robert, 1998) et les utiliser bien souvent en les adaptant : mélanger des connaissances, introduire des intermédiaires, des étapes... Nous sommes bien loin d'un exercice traditionnel d'application immédiate de connaissances. (p. 8)

Le problème de la rencontre et le problème du volcan Aso sont des éléments du scénario prévu, au sens de ce qui va être proposé, qui ne sont pas constitués de tâches précises prévues comme la plupart des tâches rencontrées habituellement. Tout d'abord, ce sont des tâches longues (qui demandent un temps important de résolution) qui nécessitent l'identification de nombreuses étapes, des étapes dans le raisonnement mais aussi des étapes liées à la modélisation. De plus, ces tâches peuvent donner lieu à de véritables choix, à des alternatives en termes d'activités. Autrement dit, nos deux problèmes de modélisation, correspondant au début de notre scénario prévu, peuvent engendrer finalement plusieurs scénarios possibles, scénarios dans le sens classique que nous avons défini ci-dessus. Dans ce type de problème, il existe une marge d'incertitude du déroulement beaucoup plus grande que pour d'autres tâches. Beaucoup de tâches et d'activités sont envisageables mais elles ne sont pas toutes attendues (comme c'est souvent le cas habituellement). Il nous semble important et incontournable de prendre en compte dans notre analyse le caractère particulier de ces tâches, lié à la démarche de modélisation, pour adapter la méthodologie « classique » d'analyse de scénarios et de tâches.

7.1.3 Prise en compte de la démarche de modélisation en probabilités

Les deux problèmes, celui de la rencontre et celui du volcan Aso, ont la particularité de mettre en jeu une démarche de modélisation. Nous allons donc ici présenter des éléments sur la démarche de modélisation et essayer ensuite d'en dégager une structure qui nous servira de support à une analyse globale de scénario et de tâche adaptée aux problèmes de modélisation.

a) Définitions

Henry (2001b) explique que le terme de « modèle » est utilisé dans des sens variés, parfois même contraires. Il arrive finalement à la définition suivante :

> Un modèle est une interprétation abstraite, simplifiée et idéalisée d'un objet du monde réel, ou d'un système de relations, ou d'un processus évolutif issus d'une description de la réalité. (p. 151)

Cette définition nous semble adaptée pour notre utilisation. Pour ce qui est de la modélisation, la définition qui apparaît dans l'introduction de la 14e étude ICMI, *Modelling and applications in mathematics education* est la suivante :

> A mathematical model consists of the extramathematical domain, D, of interest, some mathematical domain M, and a mapping from the extra-mathematical to the mathematical domain. Objects, relations, phenomena, assumptions,

> questions, etc. in D are identified and selected as relevant for the purpose and situation and are then mapped – translated – into objects, relations, phenomena, assumptions, questions, etc. pertaining to M. Within M, mathematical deliberations, manipulations and inferences are made, the outcomes of which are then translated back to D and interpreted as conclusions concerning that domain. This so-called modelling cycle may be iterated several times, on the basis of validation and evaluation of the model in relation to the domain, until the resulting conclusions concerning D are satisfactory in relation to the purpose of the model construction. The term modelling refers to the entire process, and everything involved in it. [25] (Niss, Blum & Galbraith, 2007, p. 4)

Il ne faut pas confondre le (ou les) modèle(s) et le processus de modélisation.

b) Les cycles de la modélisation

Le processus de modélisation a été étudié plusieurs fois dans le contexte de la didactique des mathématiques. Tout d'abord, dans les années 1990, une distinction entre *Real model* et *Mathematical model* a amené Kaiser (1995) et Blum (1996) à présenter un premier cycle de modélisation (cf. figure 7.1), qui est aujourd'hui notamment repris par PISA.

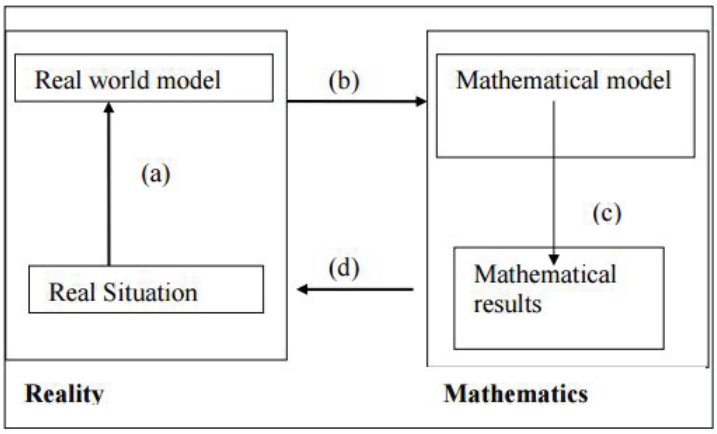


FIGURE 7.1 – Le cycle de modélisation de Kaiser (1995) et Blum (1996)

Comme l'explique Cabassut, dans un ouvrage consacré à la modélisation dans l'enseignement des mathématiques (Kuzniak & Vivier, 2011) :

25. Traduction proposée par Burgermeister & Dorier (2013) : « Un modèle mathématique comprend le domaine extra-mathématique D en jeu, un domaine mathématique M et une correspondance du domaine extra-mathématique sur le domaine mathématique. Les objets, relations, phénomènes, hypothèses, questions, etc. dans D sont identifiés et sélectionnés comme pertinents pour le but suivi et la situation et sont ensuite traduits en objets, relations, phénomènes, hypothèses, questions, etc. appartenant à M. A l'intérieur de M, des considérations, manipulations et inférences sont effectuées, dont les résultats sont à leur tour traduits et interprétés dans D comme conclusions concernant ce domaine. Ce cycle de modélisation peut être itéré plusieurs fois, en fonction de la validation et de l'évaluation du modèle par rapport au domaine, jusqu'à ce que les conclusions concernant D soient satisfaisantes, par rapport au but de construction du modèle. Le terme de modélisation se réfère à l'ensemble du processus et à tout ce qu'il implique. »

> Pour Kaiser, il s'agit de donner une représentation standard (*ideal-type*) du processus de modélisation. « Un processus de modélisation suit la procédure « ideal-type » suivante :
>
> – Une situation du monde réel est le point de départ du processus. Puis la situation est idéalisée (a), c'est-à-dire simplifiée ou structurée de façon à obtenir un « real world model ».
> – Puis ce « real world model » est mathématisé (b), c'est-à-dire traduit en mathématiques pour obtenir un modèle mathématique de la situation de départ.
> – Des considérations mathématiques dans le modèle mathématique entraînent des résultats mathématiques (c) qui doivent être interprétés (d) dans la situation réelle (*Real situation*). L'adéquation des résultats doit être vérifiée, c'est-à-dire validée. Dans le cas d'une solution non satisfaisante pour le problème, ce qui arrive fréquemment en pratique, le processus doit être recommencé ». (Kaiser, 2005, p. 101)

Il poursuit en indiquant que cette distinction entre *Real model* et *Real situation* s'avère trop simple lorsque le propos n'est plus seulement de traiter les problèmes (*word-problems*) calibrés et déjà épurés que proposent les manuels de mathématiques, mais de traiter des situations réelles. Pour cette raison, Blum et Leiss (2007) ont introduit une distinction entre la situation réelle et la situation modèle (*Situation model*) (cf. figure 7.2). Au terme « situation modèle », nous préférerons situation retenue.

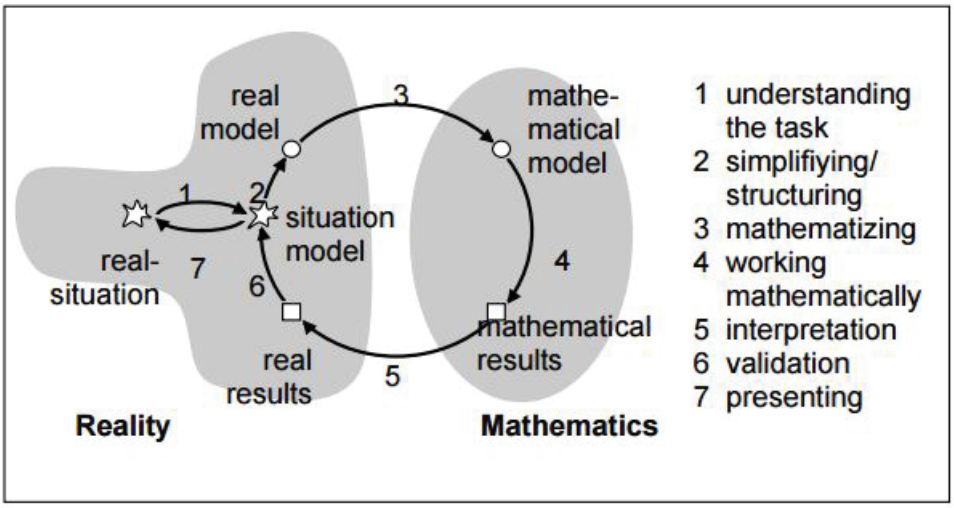


Figure 7.2 – Le cycle de modélisation de Blum et Leiss (2007)

Coulange (1998) propose, dans la démarche de modélisation, de séparer le domaine réel (domaine extra-mathématique) du domaine mathématique par un autre « domaine » : le domaine « pseudo-concret ». Nous pourrions cependant nous demander s'il s'agit effectivement d'un domaine au sens où nous l'entendons. Il semblerait que le « pseudo-concret » puisse faire partie à part entière d'un domaine mathématique, cependant dans un paradigme différent. Nous y reviendrons plus loin. Nous retrouvons aussi cette idée de « pseudo-concret » chez Henry (2001b), qui s'intéresse

lui plus particulièrement à la modélisation en probabilités. Il propose six étapes dans le processus de modélisation, qu'il nomme :

- Réalité,
- Modèle pseudo-concret,
- Modèle mathématique,
- Étude mathématique,
- Confrontation modèle-réalité,
- Généralisation et prévisions.

Il explique que la dernière étape n'est plus l'affaire du mathématicien car elle nécessite une connaissance spécialisée de la situation étudiée. Contrairement au cycle de Blum et Leiss (2007), toutes les étapes de Henry n'ont pas le même statut, parfois ce sont des actions et d'autres fois des états, ce qui peut laisser planer des confusions. Cependant, comme Coulange (1998) et Henry (2001b), il nous semble important de ne pas distinguer seulement la réalité et les mathématiques, ou tout du moins de ne pas considérer les premiers choix faits sur le modèle dans la réalité. C'est la raison pour laquelle, nous travaillerons avec un cycle schématisé en figure 7.3, qui garde la structure du cycle de Blum et Leiss (2007), tout en prenant en compte cette dernière remarque. Les étapes numérotées de 1 à 7 correspondent à des actions, tandis que les états sont représentés par les lettres de A à F.

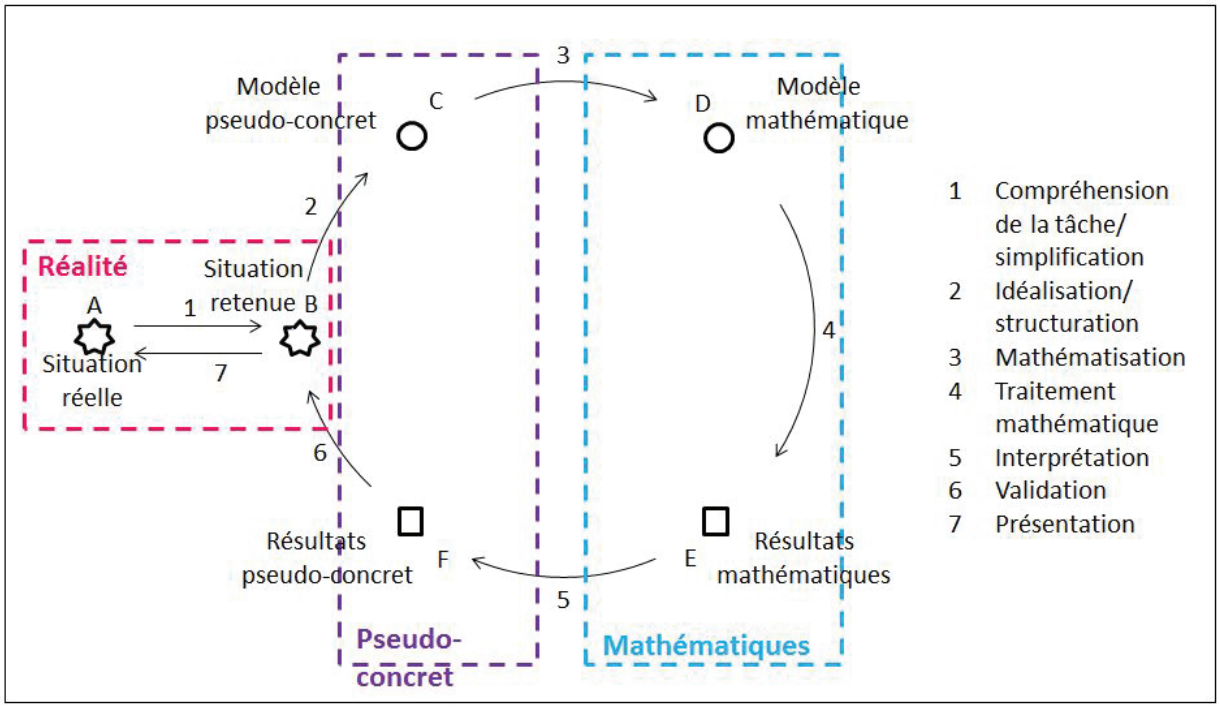


FIGURE 7.3 – Cycle de modélisation retenu

Entre la situation réelle (que l'on ne peut décrire dans toute sa complexité) et la situation retenue, il y a un processus de simplification, guidé par ce qu'Henry (1999) appelle « *un regard théorique* ». Ensuite, une idéalisation et une structuration des objets de cette situation retenue permettent de déboucher sur le modèle pseudo-

concret, avec, si l'on veut, une situation pseudo-concrète (c'est « l'habillage » dans la pratique des problèmes scolaires). Dans ce modèle, Henry parle d'hypothèses de travail qui se traduiront en hypothèses de modèle dans le modèle mathématique (Henry, 2003).

Il faut bien garder à l'esprit que les articles cités permettent une analyse théorique et didactique du processus de modélisation « décomposé ». Dans la réalité des activités des élèves, les différentes phases peuvent être imbriquées, surtout si le travail est collectif, avec des allers-retours entre modèle pseudo-concret et modèle mathématique, entre hypothèses de travail et hypothèses de modèle, à cause du statut ambigu de certaines affirmations et la présence trop prégnante parfois de la situation réelle.

c) Un lien avec les paradigmes probabilistes

Pour finir, dans ses travaux, Parzysz (2011, 2014) définit les paradigmes probabilistes suivants :

- celui de la « réalité », c'est-à-dire de l'expérience concrète effectivement réalisée à l'aide d'objets matériels (dés, pièces de monnaie, roue de loterie, jetons sortis d'une urne, etc.) ;
- un paradigme « pseudo-concret » P1 constituant une première simplification et théorisation, relativement peu formalisée, de la situation réelle (Henry, 2001b) ;
- un paradigme probabiliste plus formel P2 fondé sur la notion d'espace probabilisé fini, dont l'horizon théorique est une axiomatique de type Kolmogorov (paradigme P3).

Parzysz restreint ici P2 comme paradigme des probabilités discrètes finies, et étend ce paradigme à P2+ pour les probabilités à densité (Derouet & Parzysz, 2016a). Nous croyons que nous pouvons penser les choses différemment et envisager P1 et P2 comme des paradigmes associés à chacun de ces deux sous-domaines des probabilités. Par exemple, nous parlerons de PaD-P1 et PaD-P2. Parzysz (2014) précise aussi, ce qui semble aller dans le sens de notre remarque, que le paradigme P2 est :

> obtenu en posant un « regard probabiliste » (Henry, 1999) sur l'expérience, dans lequel on définit l'expérience aléatoire générique, ainsi que la notion de probabilité. On entreprend alors d'étudier les propriétés de cette probabilité, c'est-à-dire qu'on construit une algèbre des événements, illustrée sur des modèles classiques (en particulier des « modèles d'urne ») et les principales lois de probabilité (binomiale, exponentielle, géométrique, normale...). Les outils associés sont : la démonstration de type mathématique, les techniques de calcul usuelles, divers registres de représentation (tableau à double entrée, arbre pondéré, diagramme ensembliste, graphiques statistiques variés...), ainsi que les outils du cadre de la statistique descriptive (SD), transposés à celui des probabilités (lien entre fréquence et probabilité, entre moyenne et espérance mathématique...). (p. 68)

Dans le cas des probabilités, ici à densité, un parallèle est tout à fait envisageable entre les paradigmes et ce que Coulange (1998) appelle « domaines ». Le paradigme P1 est à relier au pseudo-concret et P2 aux mathématiques (cf. figure 7.4). Cela n'empêche pas que P1 et P2 fassent bel et bien partie du domaine mathématique, au sens où nous l'entendons.

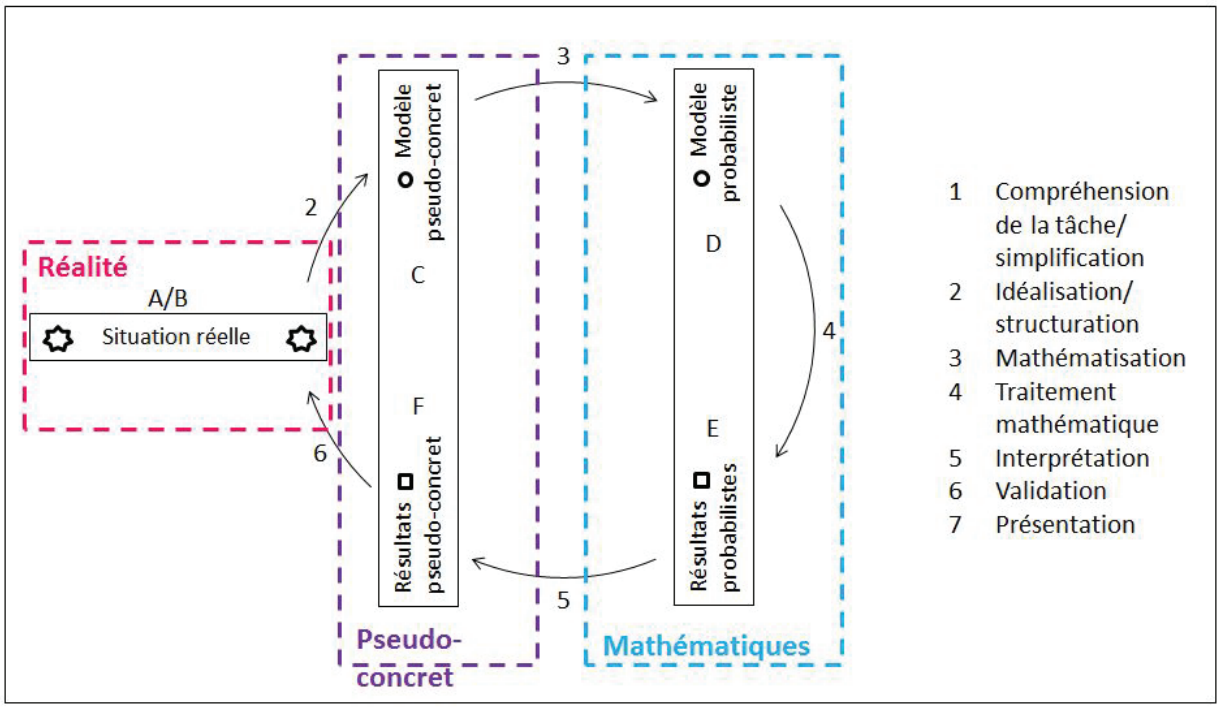


FIGURE 7.4 – Autre présentation du cycle de modélisation retenu

En outre, Parzysz (2014) fait référence au fait que pour travailler dans P2, il est possible parfois de faire appel à la statistique descriptive. Il propose, de plus, un diagramme présentant l'articulation des différents paradigmes dans les activités de simulation introduites dans les programmes de l'enseignement secondaire français (cf. figure 7.5). Dans ce cas, nous remarquons que le paradigme P2 est absent et que les articulations se font seulement entre la réalité, le paradigme P1 et la statistique descriptive. Ceci s'explique par le fait que les modèles probabilistes en jeu dans son étude ne sont pas au programme du niveau en question. Nous pourrons, dans nos exemples particuliers, enrichir le cycle retenu (cf. figure 7.3) en précisant les articulations entre les paradigmes P1 et P2 et la statistique descriptive.

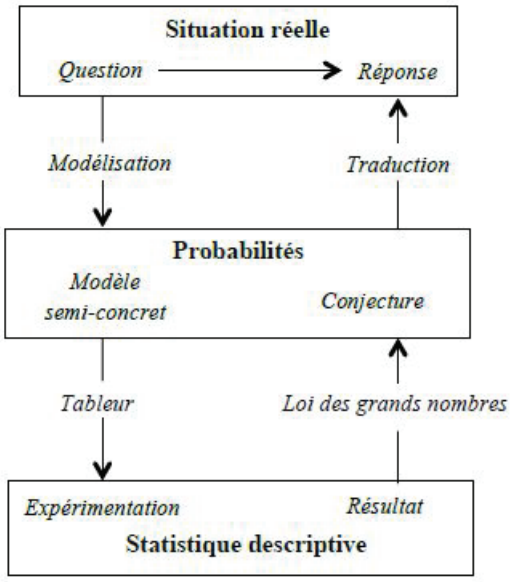


Figure 7.5 – Articulations entre les paradigmes dans les activités de simulation (Parzysz, 2014)

d) Une structuration pour l'analyse de tâches

Étant donné le caractère particulier des tâches de modélisation, il nous semble important de prendre en compte le cycle de modélisation comme cadre structurant nos analyses de tâches *a priori* et *a posteriori*. Ce cadre nous permettra d'assurer une certaine chronologie de raisonnement dans l'analyse *a priori*. En revanche, nous tenons à repréciser que l'objectif de cette thèse n'est pas d'analyser le processus de modélisation pour lui-même ou des étapes particulières de ce processus en tant que tel.

En prenant exemple sur l'exemple de Parzysz (cf. figure 7.5), nous proposerons pour chaque problème un enrichissement du cycle de modélisation, en lien avec la catégorie (catégorie 1.1 ou 2.2, cf. tableau p. 167) dans laquelle se situe le problème. Cela nous permettra de mettre en évidence des sous-étapes dans le cycle et en particulier de repérer celles où se joue l'objectif visé.

7.1.4 Adaptation de la méthodologie aux tâches de modélisation

Comme nous l'avons dit dans la section 7.1.2, les tâches de modélisation entraînent nécessairement des analyses différentes. Plusieurs raisons expliquent cela :

- l'ensemble des sous-tâches n'est pas connu à l'avance,
- les activités qui seront effectivement déclenchées ne sont pas toutes prévisibles ou ne sont qu'envisageables et non attendues, car des choix sont nécessaires au cours de la résolution du problème,
- aucun ordre dans les activités n'est réellement prévu,

– des activités qui peuvent paraître « moins mathématiques » apparaissent (c'est le cas des activités relatives aux étapes 1 et 2 du cycle de modélisation)...

Le « scénario » prévu des deux problèmes peut finalement engendrer, nous l'avons dit plus tôt, plusieurs scénarios différents (au sens classique). Nous voulons tout de même rendre compte de cette multiplicité dans nos analyses *a priori* pour garder la possibilité de reconstituer, *a posteriori*, les tâches que les élèves ont finalement rencontrées et d'analyser les activités possibles. Finalement nous aboutissons à une analyse en deux niveaux comme habituellement mais avec quelques adaptations. Le premier niveau concerne ce qui est prévu, donnant lieu à une analyse *a priori* un peu ouverte, qui comporte plusieurs chemins possibles, en particulier liés à des choix stratégiques pour les élèves. Ces choix et ces alternatives sont donc prévisibles mais pas leur issue. Du fait que cette analyse soit dans un certain sens moins précise et comporte plusieurs scénarios potentiels différents, nous avons décidé d'appeler un tel élément de scénario, canevas. Nous utilisons ce mot dans le sens où il s'agit d'un plan général dans lequel une marge de manœuvre existe, mais aussi parce qu'un canevas est un tissage peu serré et que cet enchevêtrement de fils donne cette idée de plusieurs alternatives possibles. Le second niveau donne lieu à une analyse *a posteriori* habituelle après déroulement, attachée à la reconstitution de ce qui a été choisi, cherché et résolu par les élèves. Cela permet d'identifier à travers le canevas de l'analyse *a priori* quel scénario effectif a été choisi par la classe. Une analyse *a priori* classique de tous les chemins envisageables serait très fastidieuse, d'où le choix de ne pas être trop fin dans l'analyse globale des problèmes, d'autant que l'intéressant réside aussi dans les choix à effectuer au fil des activités des élèves. Nous allons maintenant préciser la grille d'analyse relative à ces deux niveaux (*a priori* et *a posteriori*), en prenant notamment en compte les étapes du cycle de modélisation.

Nous rappelons qu'il n'est pas attendu que ces deux problèmes soient résolus individuellement par les élèves. Au contraire, il est prévu que le travail collectif de la classe fasse avancer la réflexion. De plus, l'important est la commande du collectif c'est-à-dire l'identification d'une tâche à effectuer plutôt que l'activité de traitement à mettre en œuvre pour résoudre cette tâche, sauf à certains moments spécifiques où ce travail est clairement attendu. Notre analyse ne peut donc pas se voir comme l'analyse des activités des élèves (individuellement) mais doit plutôt se voir comme celle des activités de la classe (sans penser de toute la classe).

7.1.5 Grilles d'analyses globales des deux problèmes

a) Grille d'analyse globale *a priori*

Pour la description du canevas des deux problèmes, nous allons prendre en compte plusieurs points. Tout d'abord, les étapes du cycle de modélisation (actions et états, cf. figure 7.4) vont nous permettre d'assurer une certaine « chronologie » dans notre analyse. Cette chronologie est, nous le savons bien, artificielle du fait que le cycle de modélisation n'est pas figé et que des allers-retours peuvent exister. Pour cette raison, il n'est pas évident qu'elle se retrouve effectivement dans le déroulement. Les différentes étapes seront à prendre comme une structuration de la démarche, pour le chercheur. Nous ne pensons absolument pas que ce soit le cas pour les élèves de la classe, qui n'ont jamais traité ce type de problème de mo-

délisation, encore moins en probabilités à densité. Cela nous permet d'ailleurs de dire, dès maintenant, que l'enchaînement des tâches à mettre en place relatives à ce type de problème, en lien avec les étapes du cycle de modélisation, n'est pas une connaissance disponible chez les élèves, tout du moins lors du premier problème.

À l'intérieur de ces étapes du cycle de modélisation, nous allons essayer de dégager les mathématiques en jeu, les tâches qui peuvent apparaître, les choix à faire par les élèves, les activités qui peuvent en découler. Ce ne sera pas une analyse aussi détaillée qu'une analyse de tâches habituelles. Dans le cycle de modélisation, les étapes numérotés de 1 à 7 sont des actions, nous essaierons donc de décrire des actions possibles en termes d'activités. Les étapes numérotées de A à F représentant des états à des moments du cycle, nous leur ferons correspondre des contenus mathématiques (ou pseudo-concret) arrêtés à ce stade du cycle.

Les problèmes de modélisation entraînent (normalement) des activités dans lesquelles beaucoup d'adaptations (Robert, 2008d) sont à mettre en œuvre. À ce stade de l'analyse, nous n'irons pas jusqu'à ces considérations. Nous prendrons en compte seulement les activités que nous appellerons (sous)-activités stratégiques. Il s'agit des activités liées spécifiquement aux tâches de modélisation où les élèves doivent faire des choix stratégiques entre plusieurs méthodes ou plusieurs réponses possibles (conscients ou inconscients). Ici ce sont les moments qui constitueront des alternatives dans le canevas. Nous identifierons dans le canevas les moments qui semblent *a priori* incontournables. Il sera aussi indiqué les références possibles à de l'ancien (notamment les devoirs maison antérieurs).

Le canevas sera détaillé sous forme d'un tableau mettant en évidence les étapes du cycle de modélisation, ainsi que les tâches et les activités envisageables, notamment les activités stratégiques (indiquées par le signe $\star$). Les moments jugés incontournables seront mis en gras. Nous repréciserons la légende relative aux éléments présentés au moment des analyses.

b) Grille d'analyse globale des déroulements

L'analyse *a posteriori* prend appui sur les transcriptions et éventuellement les notes de certains élèves. Il s'agit de repérer le chemin qui a été emprunté par la classe dans le canevas de l'analyse *a priori*. Il s'agit donc d'identifier dans les déroulements les sous-tâches qui apparaissent, ainsi que les activités possibles, aiguillées par notre analyse des activités envisageables. Cela nous permettra de décrire le scénario choisi par la classe.

En ce qui concerne les déroulements eux-mêmes, nous nous intéresserons simplement au travail collectif en classe entière. Cependant, nous noterons quand il s'agit d'autres formes de travail des élèves (individuel ou en groupe) car bien que nous ne nous intéressons qu'aux phases de travail collectives, nous avons bien conscience que les recherches individuelles ou en petits groupes vont jouer sur les mises en commun collectives. Nous compléterons par des indications de durées, qui elles aussi nous donnent des renseignements sur les activités possibles. Un travail de recherche individuel ne va pas avoir les mêmes conséquences en termes d'activités selon sa durée (Robert & Vandebrouck, 2014). Il en est de même pour un travail collectif.

Nous donnerons aussi des indications sur la répartition des rôles entre les élèves

(le collectif classe) et l'enseignante dans l'avancée de la résolution. Nous préciserons par un codage qui est responsable (enseignante ou élèves) de chaque activité, réflexion...

Tous ces renseignements seront donnés dans un tableau similaire à celui de l'analyse *a priori*, à l'intérieur duquel sera ajouté la durée des étapes ainsi que les lignes de la transcription. Le canevas proposé en amont sera repris comme base, dans laquelle les éléments qui ne sont pas apparus dans le déroulement seront grisés. Cela nous permettra visuellement de repérer le chemin choisi par la classe et donc le scénario choisi. Bien entendu, certaines tâches ou activités peuvent ne pas se trouver dans la « chronologie » des étapes du cycle de modélisation, elles seront donc déplacées. Un élément du canevas grisé suivi d'une flèche $\curvearrowright$ signifiera que cet élément est apparu mais plus tard dans le déroulement effectif. On le retrouvera plus loin en noir précédé par la même flèche. Une flèche dans l'autre sens, signifiera que l'élément s'est retrouvé plus tôt dans le déroulement. Nous repréciserons ces légendes au moment voulu. Dans cette analyse *a posteriori*, c'est l'ordre chronologique effectif qui prévaut sur le cycle de modélisation. Les moments jugés incontournables seront laissés en gras. Nous pourrons donc repérer s'ils se sont bien, tous ou non, retrouvés dans l'analyse *a posteriori*. Dans le scénario choisi, seront indiqués en italique les éléments (propositions, tâches, activités) pris en charge par le collectif classe (donc les élèves) et soulignés en vague ceux pris en charge par l'enseignante. Nous rappelons que nous parlons du collectif classe pour désigner les élèves. Il n'est pas attendu que tous les élèves fassent un proposition. Il suffit qu'un seul élève en fasse une pour que l'on parle du collectif. Des éléments complémentaires pourront être ajoutés comme les citations de certains passages de la transcription. Si des éléments inattendus apparaissent, ils seront précédés d'un point d'exclamation.

c) D'une analyse globale à une analyse locale

Les tâches de modélisation engendrent un travail mathématique conséquent avec beaucoup de sous-activités. La description du canevas et du scénario choisi par la classe présente des analyses globales sur le problème mais avec un grain d'analyse qui peut encore être affiné. Ces analyses sont nécessaires, mais nous allons le voir ne sont pas suffisantes. Elles sont nécessaires car elles nous permettront tout d'abord de savoir si globalement ces tâches sont effectivement réalisables en classe de terminale S, dans de bonnes conditions, c'est-à-dire sans que l'enseignante prenne trop de choses à sa charge. Nous pourrons aussi déterminer certains éléments qui semblent indispensables à la mise en place du travail mathématique, par exemple les activités incontournables ou les références à l'ancien. Par rapport à nos questions de recherche, ces analyses donnent des éléments de réponses aux questions QR3a, QR3c et QR3d.

Même si cette analyse apporte certaines réponses à nos questions de recherche, certaines restent en suspens. Il s'agit des questions spécifiques au passage relatif à la construction de la fonction de densité et de ses propriétés, qui est l'objectif d'apprentissage visé. Nous voyons donc apparaître la nécessité d'une analyse plus fine, mais pas de l'ensemble du scénario, seulement des étapes du cycle de modélisation en lien avec la construction de la fonction de densité, à savoir l'étape de mathématisation (étape 3) et l'étape de traitement mathématique (étape 4). L'analyse plus précise

du cycle de modélisation associé à chaque problème ainsi que les analyses globales nous permettront d'identifier les sous-étapes au sein de l'étape 3 à étudier plus précisément. Pour pouvoir aborder les autres questions de recherche une analyse locale plus fine de la partie mathématique du cycle de modélisation s'impose donc.

7.1.6 Grilles d'analyses locales

Nous rappelons que nos deux problèmes de modélisation ont une particularité. Il ne s'agit pas de reconnaître le modèle en jeu mais de construire ce modèle, notamment en faisant émerger le besoin pour définir ce modèle d'une fonction densité, qui doit respecter des contraintes qui seront à intégrer. C'est ici tout le véritable enjeu que nous voulons analyser. Bien entendu la globalité de la démarche de modélisation est importante pour notamment étudier le déroulement en classe (d'où l'analyse globale) mais les points clés qui nous intéressent dans notre problématique sont la construction du modèle, plus précisément la recherche de la fonction de densité du modèle en jeu. Donc, bien que l'ensemble des étapes du cycle de modélisation présente un intérêt, en ce qui concerne notre problématique, c'est essentiellement la partie mathématique du cycle qui va nous intéresser. Une analyse plus fine s'impose pour avoir accès plus précisément aux activités en jeu lors de l'étape de mathématisation (étape 3) et celle de traitement mathématique (étape 4). Nous nous focaliserons même plus spécifiquement sur des sous-étapes que nous identifierons dans l'analyse du cycle de modélisation de chaque problème.

Ces étapes « plus mathématiques » nous paraissent cette fois plus appropriées à une analyse de tâches plus traditionnelle. Même si des activités relatives aux tâches de modélisation peuvent apparaître au cours de ces étapes, comme ce que nous avons appelé les activités stratégiques, elles n'ont des conséquences que minimes sur les activités. Nous procéderons donc à une analyse *a priori*/*a posteriori* plus approfondie de ces passages. Il s'agit finalement d'un zoom sur un passage du canevas (*a priori*) et du scénario choisi (*a posteriori*).

a) Grille d'analyse locale *a priori*

Pour l'analyse locale *a priori*, nous allons prendre en compte à peu près les mêmes éléments que pour l'analyse des *activités* des manuels que nous avons déjà présentés dans la section 5.1.3. Il est conseillé de se reporter à la section 5.1.3 pour avoir des précisions mais, rapidement, nous prendrons en compte :

1. les sous-domaines mathématiques en jeu (et éventuellement les paradigmes), et pour le calcul intégral, nous pourrons préciser le niveau de calcul d'aire en jeu,
2. les connaissances à mettre en fonctionnement (ancien/nouveau, niveau de mise en fonctionnement),
3. les adaptations à mettre en place, ou plutôt les sous-activités en jeu (cela est précisé ci-dessous),
4. les registres en jeu,
5. les dimensions de l'ETM mobilisées et les articulations entre les différents ETM.

Nous ajouterons les références attendues à l'ancien.

Au niveau des domaines mathématiques, nous avons bien conscience que vu les choix que nous avons faits pour notre séquence, le calcul intégral n'est plus un prérequis pour l'introduction de la fonction de densité. Cependant, nous considérerons quand même le domaine CI en prenant en compte l'aire sous la courbe, qui fera pour nous partie du domaine CI et notamment aussi les fonctions associées à ces courbes.

Au niveau des connaissances en jeu, nous ajouterons le caractère connaissances en construction, qui sont justement les connaissances à faire émerger de ces problèmes. Dans ce cas, le caractère disponible ne signifie pas disponible pour l'ensemble des élèves de la classe. Il s'agit de disponible pour le collectif classe, donc que la connaissance est supposée disponible pour au moins un élève, ce qui permettra de faire avancer le travail de la classe.

En lien avec les adaptations, nous distinguerons plusieurs sous-activités. Nous reprendrons les trois sous-activités, liées aux connaissances en jeu et aux différentes adaptations des connaissances attendues, proposées par Robert et Vandebrouck (2014) :

- les (sous)-activités de reconnaissance d'outils ou d'objets mathématiques à mettre en fonctionnement : « *ce sont les théorèmes ou propriétés à mettre en fonctionnement. Cela peut comprendre des choix de connaissances, forcés ou non, selon les alternatives existantes* ». (*ibid.*, p. 256)
- les (sous)-activités d'organisation du raisonnement global : « *il s'agit de repérer les différents raisonnements précis à mener, avec les étapes éventuelles et leur ordre, les reprises de questions précédentes, les interprétations* ». (*ibid.*, p. 256)
- les (sous)-activités de traitement interne : « *il s'agit des constructions de figure, des calculs à effectuer, du travail sur les formules, simple remplacement des données par leurs valeurs ou transformations, équivalences, implications, mais aussi de l'introduction d'intermédiaires, notations ou expressions, des changements de registres ou de points de vue (imposés ou choisis) et des mélanges de cadres éventuels (imposés ou choisis)* ». (*ibid.*, p. 256)

Nous ajouterons les activités stratégiques que nous avons définies avant.

À l'aide de ces outils méthodologiques, nous allons pouvoir effectuer une analyse *a priori* des étapes 3 et 4 du cycle de modélisation, dans le but de pouvoir présenter les activités attendues dans la classe. Nous avons bien conscience que ces problèmes doivent être traités en collectif donc que tous les élèves ne vont pas tout faire. Cependant, nous nous arrêterons au niveau du collectif et dans l'analyse *a priori*, nous présenterons les activités attendues du collectif.

b) Grille d'analyse locale des déroulements

L'analyse des déroulements, *a posteriori*, essaiera de mettre en évidence les activités possibles dans la classe, en s'appuyant sur l'analyse *a priori*.

À l'aide des transcriptions et des étapes du cycle de modélisation, nous essaierons de procéder à un découpage des étapes 3 et 4 (plus particulièrement des sous-étapes qui nous intéressent). Ce découpage en phases se fera en prenant en compte des changements de sous-activités, en lien avec les domaines en jeu notamment, que nous identifierons. Dans chaque phase, nous signalerons les différentes tâches qui

apparaissent et nous indiquerons les interventions représentatives et les activités des élèves d'une part et de l'enseignante d'autre part. Cela nous permettra notamment d'identifier qui, des élèves ou de l'enseignante, est à l'initiative ou prend en charge les différentes sous-activités ainsi que les changements de domaines ou de registres associés. Nous reviendrons un peu plus tard sur les interventions de l'enseignante. Nous identifierons les mêmes éléments que dans l'analyse *a priori* : les changements de registres, les changements de domaines, les références à l'ancien et les sous-activités en jeu (de traitement, de reconnaissance, de raisonnement, stratégiques).

Pour essayer de répondre à la question relative au rôle de l'enseignante dans la gestion du travail mathématique, il nous semble important de regarder spécifiquement ses interventions. Pour cela, nous identifierons certains interventions particulières, dont les mathématiques en sont l'objet, c'est-à-dire les interventions qui correspondent à des mises en relation, des informations et des demandes complémentaires et des bilans (Pariès, Robert & Rogalski, 2009). Nous nous inspirons des indicateurs retenus dans l'article cité, à savoir :

- Répétition : l'enseignante redit la même phrase que l'élève, sans ajouts.
- Reformulation : l'enseignante répète ce que vient de dire l'élève mais de façon différente.
- Justification : l'enseignante ajoute des compléments d'informations pour justifier ce qui vient d'être dit.
- Évaluation des réponses des élèves (validation ou réfutation) : l'enseignante émet un simple avis sur la validité de réponses d'élèves ;
- Bilan : cet indicateur peut marquer la récapitulation. L'enseignant reprend ce qui a déjà été fait, fait appel à la mémoire en créant explicitement des liens, conclut.
- Rappel : l'enseignante donne des rappels de propriétés, de formules ou autres.

Nous identifierons aussi les demandes de justification, de précision ou d'opinion que peut faire l'enseignante auprès des élèves.

Nous ajouterons aussi les types d'aides apportées par l'enseignante. Nous distinguerons deux types d'aides : les aides procédurales et les aides constructives (Pariès *et al.*, 2009). Les aides appelées « procédurales » sont les aides qui jouent directement sur la tâche prescrite. Elles peuvent modifier les activités mathématiques des élèves par rapport à celles prévisibles. Dans les aides procédurales, nous distinguerons les aides directes, qui donnent une indication explicite dans la résolution, et les aides indirectes, par exemple l'enseignant qui pose une question sans induire aucune réponse (Robert, 2005). Les aides sont appelées constructives si elles ont une portée plus large que seulement celle de résoudre la tâche actuelle. Ces aides « *ajoutent quelque chose entre le travail de l'élève et la construction (espérée) de la connaissance qui pourrait en résulter* » (Pariès *et al.*, 2009). Pour plus de justesse, par aide procédurale il faudra entendre « aides à fonction procédurale » et par aide constructive, « aides à fonction constructive », car cet étiquetage est en fait relatif à chaque élève. Une aide procédurale pourrait être considérée comme une aide constructive pour certains élèves, prenant une valeur de méthode par exemple, alors qu'à l'inverse, certains élèves peuvent ne pas voir la fonction constructive d'une aide et rester du côté procédural (*ibid.*).

Ces critères vont nous permettre d'appréhender la gestion des activités par l'en-

seignante. Nous verrons si ces interventions guident les activités des élèves et comment.

Avec ces différentes informations, nous pourrons illustrer à l'aide des diagrammes des ETM les éventuelles circulations entre les ETM dans chaque phase. De plus nous rendrons compte du fait que les dimensions de l'ETM sont pris en charge par les élèves ou par l'enseignante. Les schémas que nous proposerons permettront de visualiser les dynamiques en jeu entre les différentes dimensions à des temps fixés, correspondant à chaque phase. Ces schémas nous permettront de nous rendre compte de la richesse ou non des circulations et notamment de voir si un travail complet ou quasi complet a eu lieu. Nous reprécisons que les trois dimensions de l'ETM de gauche à droite sont la dimension sémiotique, la dimension instrumentale et enfin la dimension discursive. Les schémas seront trop petits pour permettre de lire les différentes composantes.

Nous regrouperons l'ensemble de ces informations sous forme d'un tableau où figureront le minutage et la ligne correspondante dans la transcription de chaque phase et seront résumées les interventions et activités des élèves, d'une part, et de l'enseignante Marie, d'autre part. Dans les activités des élèves et de l'enseignante, nous identifierons les sous-activités de traitement, de reconnaissance ou d'organisation en jeu, ou éventuellement stratégiques, les changements de registres et la disponibilité des connaissances, les éventuelles références à l'ancien et les aides de l'enseignante. De plus, les diagrammes des ETM illustreront chaque phase.

c) Objectifs de ces analyses locales

Ces analyses locales nous permettront effectivement d'aborder les questions de recherches restées en suspens par les analyses globales. Nous présentons dans le tableau 7.2 les correspondances entre les analyses et les sous-questions de recherche de QR3.

QR3	Analyse globale	Analyse locale
a priori	a, c	a, b, c
a posteriori	a, c, d	a, b, c, d, e, f

TABLE 7.2 – Correspondance entre les analyses *a priori* et *a posteriori* et les questions de recherche

Bien que des réponses puissent être apportées pour les différentes questions, il nous semble que pour vraiment se rendre compte du véritable impact sur la construction du référentiel théorique lié à la fonction de densité, il ne faut pas s'arrêter à l'analyse des problèmes, mais aller aussi regarder la phase d'institutionnalisation.

7.1.7 Dernier niveau d'analyse : l'institutionnalisation

L'analyse des problèmes permet de voir si effectivement la notion de fonction de densité émerge et, de ce fait, si le référentiel commence à se construire. Cependant, l'aboutissement de sa construction dans l'ETM idoine effectif ne se fait que

lors de la phase d'institutionnalisation (Brousseau, 1998). Pour cette raison, il nous semble important de terminer l'analyse par celle de l'institutionnalisation proposée par l'enseignante. L'idée est de voir comment l'enseignante va chercher les connaissances mises en place avec les problèmes pour les institutionnaliser.

Ces analyses ne se feront qu'*a posteriori*. Nous regarderons plusieurs indicateurs :
- le vocabulaire introduit, les définitions données et les propriétés exhibées,
- qui a la charge de donner le vocabulaire, les définitions et les propriétés ci-dessus,
- le degré de généralité employé dans les définitions et propriétés (du contextualisé au plus décontextualisé),
- le niveau de rigueur employé dans les définitions, explications, justifications,
- les aides éventuelles.

Cela nous permettra de voir quelles connaissances sont extraites des problèmes de la rencontre et du volcan, dans les synthèses qui les suivent, et toujours de se demander par qui, par l'enseignante ou par les élèves. Ce dernier niveau d'analyse est nous semble-t-il indispensable pour clôturer la véritable introduction et construction du nouveau référentiel théorique de l'ETM idoine.

Après avoir présenté avec précision les détails de notre méthodologie, nous allons la mettre en œuvre sur les déroulements de notre expérimentation. Les analyses seront séparées par problèmes, avec au départ les analyses *a priori* suivies des analyses *a posteriori*.

7.2 Le problème de la rencontre

Dans cette section, nous allons présenter le détail de nos analyses *a priori* et *a posteriori* relatives au problème de la rencontre. Cela permettra notamment de voir la mise en œuvre effective de nos outils méthodologiques présentés dans la section 7.1. Nous essaierons de dégager de ces analyses quelques éléments de réponses à nos sous-questions de recherche. Avant de commencer les analyses, nous préciserons quelques éléments de contexte de l'expérimentation pour situer la séance et repréciser les prérequis (section 7.2.1). Ensuite, dans la section 7.2.2, nous présenterons nos analyses *a priori* globales puis locales qui serviront de base pour nos analyses *a posteriori* qui feront l'objet de la section 7.2.3. Nous aborderons ensuite l'analyse de l'institutionnalisation qui a suivi le problème de la rencontre (section 7.2.4). Enfin, nous donnerons les résultats de nos analyses relatives à ce problème.

7.2.1 Éléments de contexte

Pour rappel, le problème de la rencontre est la première situation d'introduction de la séquence. C'est la première fois que les élèves de la classe de terminale S de Marie rencontrent un problème de modélisation. De plus, il s'agit d'une quasi première rencontre avec les probabilités à densité, amorcée rapidement par le devoir maison n° 6. Rappelons-le, ce problème a pour objectif d'introduire les probabilités à densité et plus particulièrement de faire émerger la notion de fonction de densité.

Pour resituer dans son contexte, le devoir maison n° 5 (cf. annexe G.1), sur les aires finies d'un domaine infini, était à rendre le 24 janvier 2015, soit presque deux mois avant l'expérimentation. Une correction a ensuite été proposée en demi-groupes. L'exercice 2, celui qui nous intéresse, a posé des problèmes à certains élèves. La partie B était ouverte et certains élèves n'ont pas vu le lien avec la partie A. La correction a donc été le moment de discussions nécessaires pour mettre à plat le fait qu'un domaine infini puisse effectivement bien avoir une aire finie, ce qui n'est pas naturel pour les élèves.

Le devoir maison n° 6 (cf. annexe G.2) était, quant à lui, à rendre le 3 mars 2015 et a été corrigé en classe lors de la séance qui a précédé le problème de la rencontre. Il était clairement précisé sur le sujet que ce devoir préparait un futur chapitre, sans pour autant préciser lequel. Près d'une heure a été consacrée à la correction de ce devoir maison. Dans l'exercice 1, la notion de densité de fréquence semble bien comprise pour la majorité des élèves, cependant certains n'ont pas réussi à démontrer que la hauteur des rectangles d'un histogramme est proportionnelle à la densité. Nous interprétons cela plutôt comme une difficulté liée à la notion de proportionnalité. Le véritable point qui a posé problème à la majorité des élèves (ce qui était prévisible) et qui a persisté malgré la correction apportée, est le fait que la probabilité que la variable aléatoire prenne une valeur isolée soit nulle, dans l'exercice 3.B. Pour beaucoup d'élèves, il est inconcevable que la probabilité d'un tel évènement soit nulle. Pour finir, la question 5.b) qui demande de tracer une « courbe de tendance » identique pour les quatre histogrammes a pour certains élèves été mal traitée : ils ont proposé des courbes différentes pour chaque cas.

Pour résumer, au moment de la séance 1, les élèves ont donc *a priori* les idées plus claires sur les histogrammes, par exemple ils devraient savoir que :

- l'aire des rectangles d'un histogramme représente la fréquence sur les classes respectives ;
- l'axe des ordonnées d'un histogramme correspond à la densité de fréquence et non à la fréquence elle-même.

Ils ont de plus rencontré la loi uniforme, loi que suit la position de la puce électronique dans l'exercice 3 du DM6, qui n'a cependant pas encore été nommée comme telle. Dans ce problème, des simulations suivies de la construction des histogrammes correspondants ont été proposées et les élèves ont ensuite cherché une « courbe de tendance » (vocabulaire introduit) qui « lisse » les histogrammes. La « courbe de tendance » a été un moyen de confronter leurs calculs de probabilités effectués de façon intuitive à cette autre « méthode ». Les élèves vont, dans le problème de la rencontre, réutiliser cette démarche cependant non plus dans un souci de vérification des résultats mais bel et bien par nécessité. La fonction de densité est l'outil, accompagné ensuite de l'outil aire sous la courbe, qui permet de répondre au problème posé.

7.2.2 Analyse *a priori* du problème de la rencontre

a) L'énoncé

Nous avons déjà présenté succinctement le problème de la rencontre dans la section 6.3.2. Ici il s'agit de rentrer dans le détail des analyses. Nous repartons donc de

l'énoncé de ce problème, qui est le suivant :

> Karine et Olivier décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Ils peuvent arriver à tout moment entre 7h et 8h. Que peut-on dire du temps d'attente du premier arrivé ?

Une fois ce problème résolu, un exercice d'application qui mobilise la fonction de densité trouvée est distribué aux élèves (cf. Annexe G.3).

L'énoncé du problème est simple à comprendre. Ce problème ne présente pas tout à fait une situation réaliste. On peut sentir que des hypothèses sont déjà incorporées dans le problème, comme le fait que Karine et Olivier semblent ne pouvoir arriver qu'entre 7h et 8h, ni avant ni après. Cependant, il reste tout de même dans un contexte concret, qui laisse encore une place à la modélisation. C'est donc bien un problème de modélisation (probabiliste). La tâche prescrite est très ouverte : « Que peut-on dire du temps d'attente du premier arrivé ? ». C'est un moyen de laisser les élèves rentrer dans le problème et se l'approprier. Tous les élèves peuvent dire quelque chose dès le début. Des premières réponses peuvent être apportées, puis l'objectif est de reformuler cette question large en la question finale suivante, en termes pseudo-concret : quelle est la probabilité que le temps d'attente du premier arrivé soit entre le temps a et le temps b (avec a et b entre 0 et 1) ?

L'exercice 1 est finalement un exercice d'application de ce qui aura été trouvé dans le problème de façon générale. Pour cette raison, nous ne nous attarderons pas sur cette seconde partie.

b) Gestion prévue

Le temps prévu pour ce problème est d'1h30 environ, mais nous savons que la séance 1 d'une durée de 1h50 peut permettre quelques ajustements. Ce travail ayant lieu en classe entière, dans une salle équipée seulement d'un ordinateur, le travail de manipulation des logiciels n'est pas attendu. L'enseignante se chargera de manier les logiciels (*Excel* et *GeoGebra*). Cependant, elle agira sous la commande des élèves, sauf si la technique n'est pas supposée disponible pour eux. Les logiciels sont considérés comme des outils pour obtenir des objets mathématiques. Il est attendu que les élèves soient source d'idées mais qu'ils n'aient pas forcément tout à faire. Il est prévu que l'enseignante laisse parler les élèves, qu'elle ne les influence pas (trop).

Différentes phases de travail sont prévues, tout d'abord une phase de recherche individuelle des élèves, puis une phase collective de mise en commun, pour faire avancer le travail et amener de nouvelles phases de recherche individuelles. Les mises en commun sont des moments importants car ce sont elles qui vont permettre de construire petit à petit le référentiel théorique de l'ETM idoine relatif aux probabilités à densité.

c) Cycle de modélisation associé au problème de la rencontre

Le problème de la rencontre suppose une démarche de modélisation. Du fait que la situation ne soit pas tout à fait réelle, nous pouvons considérer que le cycle de modélisation (cf. figure 7.4) commence à l'étape B directement ou bien assimiler les étapes A et B. Ensuite, des choix vont être à faire notamment en ce qui concerne les temps d'arrivée de Karine et Olivier. L'énoncé indique que les deux amis peuvent arriver à tout moment entre 7h et 8h. Le « *à tout moment* » n'implique pas nécessairement une répartition uniforme sur la plage horaire 7h-8h. Nous faisons l'hypothèse que les temps d'attente seront modélisés par des lois uniformes (même loi que dans le DM6, problème de la puce), seule loi disponible en acte chez les élèves.

Ensuite, plusieurs démarches de résolution sont envisageables. Une résolution à l'aide d'une configuration géométrique comme proposée dans le document *Ressources* (MENJVA & DGESCO, 2012, en annexe C.3) est possible. Il suffit d'adapter la démarche au cas général. La probabilité que le temps d'attente du premier arrivé soit compris entre a et b correspond à la somme des aires colorées sur la figure 7.6.

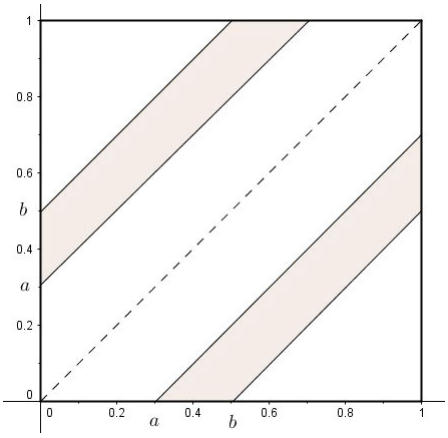


FIGURE 7.6 – Représentation géométrique du problème de la rencontre

À l'aide de calculs d'aires de trapèze, on obtient :

$$P(a \leqslant X \leqslant b) = 2 \times (b - a + \frac{a^2}{2} - \frac{b^2}{2}) = a^2 - b^2 + 2b - 2a.$$

L'objectif, dans notre cas, étant de faire émerger la notion de fonction de densité, ce n'est pas ce type de résolution qui est espéré (bien que possible). Nous proposons au lecteur de se reporter à l'analyse détaillée de cette résolution « géométrique » dans la thèse de Nechache (2016). Nous allons maintenant étudier seulement la seconde démarche, celle attendue.

Nous avons conçu ce problème dans un but de proposer une démarche de catégorie 2.2 (cf. section 5.2.4). Il s'agit donc, une fois les hypothèses faites sur le modèle, de se retrouver avec deux lois « connues » (tout du moins pouvant être simulées) pour en déduire une nouvelle loi, combinaison des deux premières. Nous partons donc des probabilités à densité (lois de départ) pour arriver aux probabilités à densité

(loi d'arrivée, inconnue), en passant par la simulation donc la statistique descriptive. Cette démarche espérée répond au cycle de modélisation, repris et complété, en figure 7.7.

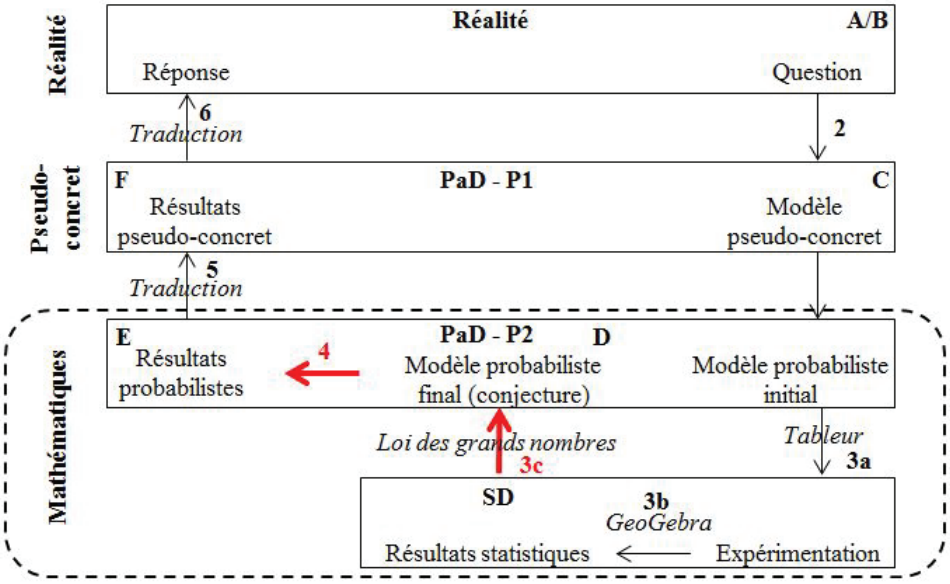


Figure 7.7 – Cycle de modélisation complété du problème de la rencontre

Nous voyons apparaître un passage par la statistique descriptive, via la simulation qui permet de construire le modèle probabiliste de la variable aléatoire du temps d'attente, inconnu au départ. Dans la partie mathématique, une fois que les lois connues et la combinaison de ces lois sont identifiées, l'utilisation du tableur permet de simuler des expérimentations. Ces données peuvent ensuite être représentées par un histogramme. L'artefact *GeoGebra* sera préconisé pour ce travail. Les amplitudes de cet histogramme pourront être modifiées à souhait notamment pour permettre un retour aux probabilités à densité par le biais de la construction de la courbe de tendance. La détermination d'une telle courbe va nécessiter de prendre en compte les contraintes qu'elle doit vérifier, notamment le fait que l'aire sous la courbe doit être égale à 1 (passage par CI). Cette courbe permettra ensuite de déterminer les résultats cherchés. Les passages en rouge sont principalement ceux qui nous intéressent dans la connexion entre probabilités à densité (PaD) et calcul intégral (CI). C'est pour cette raison que nos analyses s'attarderont plus particulièrement sur ces étapes 3c et 4.

Ce cycle de modélisation adapté à ce problème va nous permettre de cadrer notre analyse *a priori* et *a posteriori*.

d) Analyse globale *a priori* : description du canevas du problème

Dans le tableau ci-dessous, nous présentons le canevas du problème de la rencontre, à savoir l'ensemble des scénarios envisageables (*a priori*). Nous identifions les tâches et les activités envisageables, dans un sens large. Nous mettons une étoile

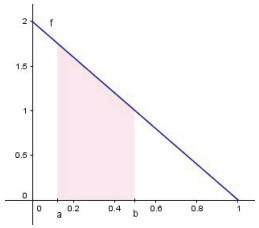
($\star$) devant les activités dites stratégiques, où des choix relatifs à la modélisation sont à faire (choix de modèles par exemple). Nous essayons aussi de repérer des passages qui nous semblent incontournables. Ces éléments sont représentés en caractères gras. Nous ajoutons, en rouge, les références possibles à l'ancien (par exemple, le DM6). Nous précisons aussi les blocages envisagés. Les étapes numérotées par des lettres sont des états, à un moment du cycle. Nous leur faisons donc correspondre des contenus mathématiques (ou pseudo-concret) arrêtés à ce stade du cycle (nous mentionnons aussi les énoncés distribués).

Comme cela est précisé dans la méthodologie (section 7.1), cette première analyse n'est pas fine. Il s'agit de présenter le problème dans sa globalité. Les étapes 3c et 4, que nous venons d'identifier dans le paragraphe précédent comme étant les étapes où se jouent la construction de la notion de fonction de densité, seront ensuite analysées plus finement.

Canevas du problème de la rencontre		
Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
A/B	Situation réelle / Situation retenue	Enoncé (donné aux élèves) : Karine et Olivier décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Ils peuvent arriver à tout moment entre 7h et 8h. Que peut-on dire du temps d'attente du premier arrivé ?
1/2	Compréhension de la tâche/ Simplification Idéalisation/Structuration	Un questionnement sur le « réalisme » du problème peut être abordé : sur l'impossibilité d'arriver avant 7h et après 8h, sur le rendez-vous sur une plage horaire aussi grande,... Les termes « temps d'arrivée » et « temps d'attente » doivent être évoqués. **Le caractère aléatoire du problème doit ressortir.** **Une discussion autour du caractère continu ou non des temps d'arrivée et du temps d'attente doit émerger, pour ensuite choisir de les modéliser par du continu.** ⋆ **Une simplification de la réalité doit être choisie. Le choix le plus probable est le suivant : Karine et Olivier arrivent à n'importe quel moment au hasard entre 7h et 8h. De plus, ils ne peuvent pas arriver avant 7h ou après 8h.** Des dessins peuvent être faits pour essayer de données des résultats qualitatifs à la question posée, comme par exemple une représentation géométrique de la situation en traçant un segment sur $[0; 1]$ et placer des points au hasard sur ce segment.
C	Modèle pseudo-concret	Les hypothèses faites par rapport au modèle ne seront peut-être pas clairement explicitées. Cependant, implicitement, le modèle doit donc vérifier plusieurs hypothèses simultanément : H_1 : Si Karine ou Olivier arrivent avant 7h ou après 8h, l'expérience est annulée. H_2 : Le temps est considéré comme continu. H_3 : Karine et Olivier arrivent au hasard entre 7h et 8h. Il doit être identifié que **le temps d'attente est la différence entre le temps d'arrivée du second et le temps d'arrivée du premier, c'est-à-dire la différence absolue entre les deux temps d'arrivée. Le temps d'attente est compris entre 0 et 1 (si l'on parle en heure) ou 0 et 60 (si l'on parle en minute).** Cela permet déjà de proposer une réponse à la question posée.

Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
		Il faudra, à un moment, reformuler la question de départ et arriver au fait que l'on cherche à déterminer la probabilité que le temps d'attente soit entre a et b (a, b $\in$ $[0;1]$ deux temps d'attente).
3	Mathématisation Modèle probabiliste initial	**Deux variables aléatoires qui représentent le temps d'arrivée de Karine et Olivier, par exemple T_K et T_O, doivent être introduites. $\star$ Elles sont à valeurs dans $[7;8]$, ou encore elles peuvent être considérées sur $[0;1]$.** **Il doit être identifié que T_K et T_O suivent le même type de loi que la variable aléatoire de position de la puce du DM6 (implicitement loi uniforme continue).** **Il doit être introduit une nouvelle variable aléatoire, par exemple X, la variable aléatoire qui représente le temps d'attente du premier arrivé.** **Il est nécessaire qu'émerge le fait que $X = \lvert T_K - T_O \rvert$** **$X$ est à valeurs dans $[0;1]$.** La tâche peut être reformulée comme ceci : On cherche à déterminer $P(X \in [a;b])$), pour a et b appartenant à $[0;1]$.
3a	↓	À cet endroit, un blocage est envisagé. Il faut qu'arrive l'idée de faire des simulations de l'expérience aléatoire pour débloquer la situation. **Simulation(s) de l'expérience aléatoire**. $\star$ Les deux choix d'artefacts envisageables par les élèves semblent être le tableur et la calculatrice. Le tableur semble être l'outil le plus simple pour présenter à la classe des simulations (vidéoprojecteur).
	Expérimentation	L'expérimentation doit fournir alors un ou des échantillon(s) de grande taille de temps d'attente.

Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
3b	↓	Il peut être proposé de calculer des fréquences d'apparition sur un intervalle. Cependant, cela n'aboutira que sur des approximations de la probabilité sur des intervalles fixés, à moins d'envisager de mettre des bornes a et b qui varient. Cette dernière démarche semble très compliquée à mettre en place sur un tableur par exemple. **Il est nécessaire qu'émerge l'idée de construire un histogramme de fréquences** associé à l'échantillon. L'enseignante proposera d'utiliser le logiciel *GeoGebra* pour le construire.
	Résultats	Un histogramme de fréquences doit être obtenu.
3c	[Analyse détaillée dans la section suivante p. 252]	⋆ Il peut ensuite être proposé de diminuer l'amplitude des classes de l'histogramme. Nous rappelons que le logiciel *GeoGebra* présente le désavantage de ne pas permettre d'obtenir, avec la fenêtre histogramme, des histogrammes à pas non constant. Il pourra être observé que si l'on diminue trop l'amplitude des classes, l'histogramme apparaît avec des irrégularités (des creux et des pics). **Il faut ensuite que soit proposé de déterminer une courbe de tendance** (terme utilisé dans le DM6) c'est-à-dire une courbe qui « lisse » l'histogramme (quelque soit l'amplitude choisie). **Après observation, il doit émerger que la courbe de tendance semble visuellement être une droite donc l'équation de la courbe est de la forme** $y = ax + b$ Si on appelle f la fonction représentée par la courbe de tendance, on a $f(x) = ax + b$. Il faut ensuite trouver deux conditions que cette fonction doit vérifier. Il semblerait que seulement deux conditions liées au problème puisse émerger. Il est attendu que soit proposé que f **doit vérifier** $f(1) = 0$. Cela peut se justifier mais non rigoureusement : la probabilité que le temps d'attente se rapproche de 1 est de plus en plus faible (il y a de moins en moins de possibilités pour les temps d'arrivée des deux amis) et finalement pour que le temps d'attente soit exactement de 1, il n'y a qu'une possibilité, que le premier arrive à 7h et le second à 8h. De plus, il doit ressortir la condition que **l'aire sous la courbe doit valoir 1**, car la probabilité que $X \in [0; 1]$ est doit être de 1. Il peut aussi être faîte une analogie entre l'aire totale des rectangles de l'histogramme qui vaut 1 et donc l'aire sous la courbe qui doit valoir 1.

Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
		Il faut alors ensuite résoudre un **système de deux équations à deux inconnues, qui donne** $f(x) = -2x + 2$.
D	Modèle probabiliste final	La démarche permet d'arriver au fait que **la courbe de tendance de la variable aléatoire** X **a pour équation :** $y = -2x + 2$
4	Traitement mathématique [Analyse détaillée dans la section suivante p. 252]	Par une analogie entre histogramme et courbe, il peut alors émerger que **cette courbe permet alors de déterminer** $P(a \leqslant X \leqslant b)$ **: il s'agit de l'aire sous la courbe entre** a **et** b**.** 
E/F	Résultats mathématiques et pseudo-concrets	$P(X \in [a; b]) = Aire_{ABCD}$ (aire du trapèze). La probabilité que le temps d'attente soit entre a et b est égale à l'aire de la figure $ABCD$

Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
A ce stade, on peut considérer le problème de la rencontre « résolu ». Un exercice (exercice 1) est alors distribué aux élèves.		
C	Modèle pseudo-concret	**Exercice 1** (questions posées en termes pseudo-concrets) En utilisant le modèle mis en place, répondez aux questions suivantes. a) Quelle est la probabilité que le premier arrivé attende exactement 40 minutes ? b) Quelle est la probabilité que le premier arrivé attende plus de 40 minutes ? c) Quelle est la probabilité que le premier arrivé attende moins de 40 minutes ? d) Quelle est la probabilité que le premier arrivé attende entre 20 et 40 minutes ? e) Quelle est la probabilité que Karine et Olivier arrivent en même temps ? f) Finalement, ils ont décidé de ne pas attendre plus d'un quart d'heure. Quelle est la probabilité qu'ils se retrouvent ?
D	Modèle probabiliste	Il s'agit d'utiliser le modèle probabiliste précédent.
4	Traitement mathématique	cf. Analyse de tâches en annexe J.2
E/F	Résultats mathématiques et pseudo-concrets	Réponses simultanées en termes mathématiques et en termes pseudo-concret.
Aucun retour sur le problème réel n'est attendu car aucun moyen de vérification n'est envisageable.		

Cette analyse globale nous permet de repérer des passages qui nous semblent incontournables pour le bon déroulement du scénario effectif. En résumé, il s'agit des points suivants :

- choix du modèle associé aux temps d'arrivée,
- simulation,
- construction d'un histogramme,
- recherche de la fonction de densité.

Ces « passages obligés » sont indépendants des issues des activités stratégiques, à une exception près : si les élèves ne choisissent pas que les temps d'arrivée suivent une loi uniforme. Nous repérons des appuis sur l'ancien, à savoir le DM6 (le problème de la puce). Nous allons maintenant nous focaliser sur les étapes du cycle de modélisation qui nous intéressent tout particulièrement : les étapes 3 et 4.

e) Analyse locale *a priori* de la mathématisation (étape 3) et du traitement mathématique (étape 4)

Comme nous l'avons précisé dans la présentation du cycle de modélisation associé à ce problème (cf. section c), p. 244), ce sont dans les étapes 3 et 4 que se passe la construction de la notion de fonction de densité, même plus spécifiquement dans les passages 3c et 4 (dans le tableau p. 249). Nous proposons, ici, une analyse *a priori* plus fine de ces passages, en prenant en compte la méthodologie présentée dans la section 7.1.6.

Une fois le modèle pseudo-concret arrêté pour les temps d'arrivée (étape C) dans le paradigme PaD-P1, on commence à entrer dans le paradigme PaD-P2 avec le vocabulaire de variable aléatoire, connaissance ancienne supposée disponible. Ici, il s'agit de reconnaître que les temps d'arrivée des deux protagonistes peuvent être modélisés par des variables aléatoires qui suivent la même loi que celle de la position de la puce du DM6. Il s'agit donc d'une connaissance en construction (loi uniforme) supposée disponible. L'activité attendue se situe ici alors dans le registre symbolique probabiliste.

La mobilisation de la dimension instrumentale est ensuite attendue pour procéder à des simulations de l'expérience aléatoire. Les artefacts que les élèves pourraient envisager d'utiliser sont la calculatrice et le tableur. On passe alors au registre associé à l'artefact. L'enseignante va ensuite utiliser le tableur. Elle pourra, par la suite, prendre ces données pour construire un histogramme. Les connaissances en jeu sont donc la connaissances du langage associé à l'artefact ; ici par exemple la formule =ABS(ALEA()-ALEA()) pour le tableur *Excel*. Ce passage par l'outil technologique permet de basculer dans le domaine de la statistique descriptive (SD). Le résultat obtenu est dans le registre numérique.

Ces simulations pourraient ensuite être utilisées pour approcher la probabilité d'appartenir à un intervalle par la fréquence d'apparition sur l'intervalle en question. Cette méthode ne permet pas de connaître la probabilité pour tous les intervalles de $[0; 1]$. L'activité attendue est de passer dans le registre histogramme, démarche qui a été rencontrée dans le DM6. Ici, le logiciel *GeoGebra* va entraîner le changement de registre. Cette technique instrumentale étant inconnue des élèves, elle sera prise

en charge par l'enseignante.

C'est une fois toutes ces étapes accomplies que nous arrivons à la sous-étape 3c qui va enfin faire émerger la fonction de densité. Ce changement de registre permet ensuite un travail de visualisation. Un traitement dans le registre histogramme peut être effectué, en modifiant l'amplitude et/ou l'échantillon simulé, ce qui permet de remarquer une certaine stabilité des histogrammes. On peut aussi remarquer qu'un pas d'amplitude trop petit donne un histogramme irrégulier et ne va pas permettre de déterminer la courbe de tendance, contrairement à ce que certains manuels laissent entendre (cf. section 5.2.3). Des histogrammes bien choisis (amplitude ni trop grande, ni trop petite) vont normalement permettre un basculement dans le domaine des probabilités à densité (PaD) par la recherche d'une courbe de tendance. Ce changement de domaine par le biais de la courbe de tendance, connaissance en construction supposée disponible, tout du moins pour une partie des élèves, grâce au DM6 qui a introduit la courbe de tendance comme la courbe qui « lisse » les histogrammes. Dans la gestion de classe, il est clairement prévu de laisser les élèves tracer la courbe de tendance comme bon leur semble. Il semble important que les élèves produisent des courbes comme ils le souhaitent, pour la plupart voire la totalité erronées, pour ensuite s'appuyer sur ces courbes de tendance pour mettre en avant que la contrainte de lissage du haut des rectangles de l'histogramme ne suffit pas et que d'autres contraintes sont à prendre en compte.

Ici, avec ce choix de modèle uniforme continu pour les temps d'arrivée, le modèle du temps d'attente est unique. Une seule courbe de densité (appelée courbe de tendance pour le moment) est possible (à un nombre fini de points près). C'est la raison pour laquelle il va falloir faire émerger des contraintes sur cette fonction pour que chaque élève ait bien la même courbe de tendance. Un passage dans le domaine de l'analyse est indispensable, en passant de la représentation graphique supposée de la courbe (c'est-à-dire une droite) à l'expression algébrique de la fonction dont on cherche la courbe. Les connaissances sur les fonctions affines, notamment que l'expression algébrique d'une telle fonction est de la forme $f(x) = ax + b$, sont des connaissances anciennes supposées disponibles. Pour déterminer a et b, il faut donc trouver deux conditions sur cette fonction, conditions qui doivent être indépendantes de la courbe que chacun a tracée. Une des conditions est que $f(1) = 0$. Une justification intuitive possible est que la densité de probabilité est proportionnelle à la longueur $1 - t$, où t est le temps d'attente du premier arrivé considéré. Quand le temps d'attente est de 1, c'est-à-dire 1h, il n'y a pas le choix : nécessairement le premier est arrivé à 7h et le second à 8h. De ce fait, la densité est nulle, autrement dit $f(1) = 0$. Cette remarque confirme aussi le fait que f est une fonction affine décroissante. Ensuite, la seconde contrainte est que l'aire sous la courbe doit être égale à 1. Il s'agit d'une connaissance en construction, qui est supposée disponible car déjà rencontrée dans le DM6. Elle peut être justifiée de deux façons : soit que la probabilité sur $[0; 1]$ doit obligatoirement être de 1, donc il en est de même pour l'aire sous la courbe, soit en faisant une analogie entre statistique et probabilités, en remarquant que l'aire totale des rectangles de l'histogramme vaut 1 et donc que l'aire sous la courbe doit être aussi de 1. La connaissance de la formule de l'aire d'un triangle, connaissance ancienne, est supposée disponible pour l'ensemble des élèves.

Ce travail entre les trois sous-domaines (probabilités à densité, statistique des-

criptive et calcul intégral) permet de faire émerger les contraintes liées à la courbe de tendance et le lien entre probabilités à densité et calcul intégral. Cela permet du coup de dévoiler l'utilité de cette courbe de tendance, c'est-à-dire que l'aire sous la courbe permet de déterminer la probabilité (étape 4). Nous proposons un diagramme en figure 7.8 présentant un cheminement possible entre les différentes notions, dans lequel les strates horizontales font référence aux sous-domaines en jeu.

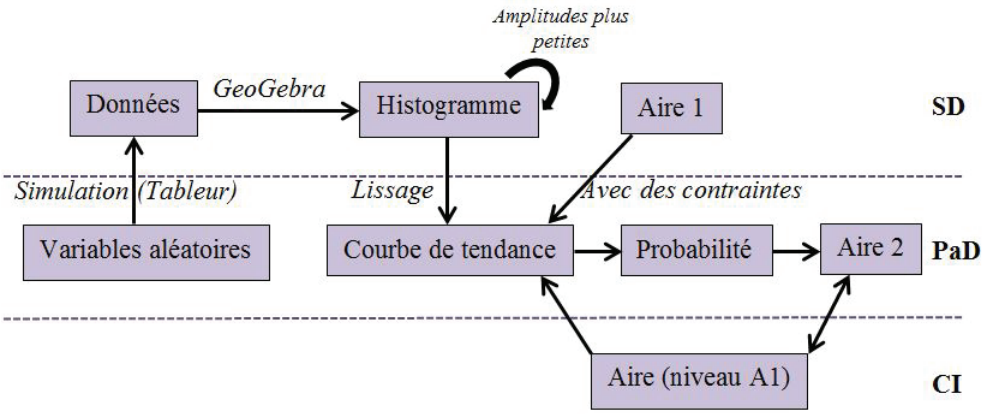


FIGURE 7.8 – Cheminement dans le problème de la rencontre

Les passages 3c et 4 sont vraiment le cœur de la construction de la fonction de densité. C'est à ces moments que nous pensons identifier des dynamiques entre les ETM des trois sous-domaines, qui construisent la notion non pas seulement par le biais de la dimension sémiotique comme c'est le cas dans les manuels mais aussi en s'appuyant sur les deux autres dimensions (instrumentale et discursive) pour permettre véritablement la construction de la notion de fonction de densité.

Nous pouvons aussi repérer un appui important sur les notions du DM6 (comme la notion d'histogramme, de densité de fréquence, de simulation, de courbe de tendance) qui nous semblent indispensables pour faire avancer le travail dans la classe. Une de nos hypothèses est que sans un travail sur ces notions, on ne peut pas arriver à cette construction.

Cette analyse *locale* a permis de préciser les activités envisageables (même attendues nous pouvons dire ici) dans la classe. À l'aide de ces différentes analyses, nous allons maintenant pouvoir aborder l'analyse du déroulement.

7.2.3 Analyse du déroulement

Après ces analyses *a priori* du problème de la rencontre, nous allons maintenant analyser le déroulement effectif. Nous allons en décrire les analyses *a posteriori*, enrichies par des compléments. Nous toucherons des points qui échappent à l'analyse *a priori* comme la gestion des activités par l'enseignante, notamment via ses interventions.

a) Déroulement général

Prévu initialement d'une durée d'environ 1h30, le travail sur le problème de la rencontre, en tenant compte de la synthèse et de la correction de l'exercice 1, a finalement duré environ 2h25 en classe. La démarche de modélisation autour du problème a eu lieu exclusivement lors de la séance 1 (un peu moins de 1h50). Nous allons donc, dans un premier temps, nous consacrer seulement à la séance 1 pour les analyses. La correction de l'exercice 1 (séance 2) ne sera pas abordée dans les analyses car elle ne concerne pas la construction de la notion de fonction de densité. Il s'agit d'une application une fois celle-ci déterminée, donc plutôt de l'exploitation de la fonction de densité. Nous traiterons ensuite, dans la section 7.2.4, de la phase de synthèse et d'institutionnalisation qui a eu lieu en début de séance 2.

b) Analyse globale *a posteriori* : scénario choisi

Après avoir travaillé sur la transcription de la séance 1 (cf. Annexe I.1), nous avons pu construire le tableau suivant, présentant le scénario choisi par la classe. Nous rappelons que nous avons gardé comme base le canevas construit pour l'analyse *a priori*. Nous avons ensuite mis en gris les éléments qui n'apparaissent pas à l'endroit prévu, accompagné d'une flèche si ce moment est repéré à un autre endroit, sans rien sinon. Nous avons ajouté en italique les propositions explicites des élèves et nous avons souligné par une vague les éléments donnés par l'enseignante. Nous avons mis un point d'exclamation aux endroits inattendus. De plus, une colonne précise les durées correspondants à chaque étape et les numéros de ligne, pour pouvoir se repérer dans la transcription.

Voici la légende :
italique : activités liées à l'intervention d'un élève
vague : activités liées à l'intervention de l'enseignante
gras : passages jugés incontournables lors de l'analyse *a priori*
gris seulement : activités qui ne sont pas apparues dans le déroulement
gris avec ↷ : activités qui sont apparues plus tard dans le déroulement
↷ noir : nouvel emplacement des activités ci-dessus
gris avec ↶ : activités qui sont apparues plus tôt dans le déroulement
↶ noir : nouvel emplacement des activités ci-dessus
rouge : référence à l'ancien
⋆ astérique : activités stratégiques

Analyse *a posteriori* du problème de la rencontre - Séance 1			
Étape du cycle de modélisation		Durée/ lignes	Tâches et activités effectives
Distribution de l'énoncé. Dévolution : lecture individuelle.			
A/B	Situation réelle / Situation retenue		Énoncé (distribué) : Karine et Olivier décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Ils peuvent arriver à tout moment entre 7h et 8h. Que peut-on dire du temps d'attente du premier arrivé ?
Phase de recherche individuelle ($\sim$ 5min)			
1/2	Compréhension de la tâche/ Simplification Idéalisation/Structuration	$\sim$ 4min l.56-126 $\sim$ 8min30 l.127-281	Un questionnement sur le « réalisme » du problème est apparu : *Proposition : C'est un peu bizarre comme rendez-vous.* Temps d'arrivée, temps d'attente. ↷ **Le caractère aléatoire du problème est ressorti** : *Autant de chance*, *Au hasard.* ↶ **Karine et Olivier ne peuvent pas arriver avant 7h ou après 8h.** ↷ *Temps d'arrivée, temps d'attente.* **Discussion autour du caractère continu ou non des temps d'arrivée et du temps d'attente** : *Proposition : Karine et Olivier ont autant de chance d'arriver sur chaque minute de 7h-8h.* *Proposition : C'est comme la puce.* *Proposition : On est en maths.* *Proposition : Ce n'est pas le même modèle que la puce (différence entre distance et temps).* **Simplification de la réalité choisie : *Karine et Olivier arrivent au hasard à n'importe quel moment entre 7h et 8h.* Ils ne peuvent pas arriver avant 7h ou après 8h.** ↶ Des dessins peuvent être faits pour essayer de données des résultats qualitatifs à la question posée, comme par exemple une représentation géométrique de la situation en traçant un segment sur $[0; 1]$ et placer des points au hasard sur ce segment.
C	Modèle pseudo-concret		« H_1 » : Si Karine ou Olivier arrivent avant 7h ou après 8h, l'expérience est annulée. « H_2 » : Le temps est considéré comme continu : *Proposition : On considère l'axe du temps comme l'axe des réelles. le temps d'arrivée entre 7h et 8h comme continu.*

Étape du cycle de modélisation		Durée/ lignes	Tâches et activités effectives
			« H_3 » : *Karine et Olivier arrivent au hasard entre 7h et 8h.* **Le temps d'attente est la différence entre le temps d'arrivée du second et le temps d'arrivée du premier, c'est-à-dire la différence absolue entre les deux temps d'arrivée. Le temps d'attente est compris entre 0 et 1 (si l'on parle en heure) ou 0 et 60 (si l'on parle en minute).** Reformulation de la question : déterminer la probabilité que le temps d'attente soit entre a et b (a, $b \in [0;1]$). ↷
3	Mathématisation Modèle probabiliste initial	∼ 20min l.282-632	↶ X ***la variable aléatoire qui représente le temps d'attente du premier arrivé.*** ! Une discussion (inattendue) a surgi sur *le problème de la possibilité ou non de l'ex aequo.* T_K ***et*** T_O ***deux variables aléatoires qui représentent le temps d'arrivée de Karine et Olivier***, **à valeurs dans [7 ;8] ou encore [0 ;1].** ↶ $X = \lvert T_K - T_O \rvert$ *Discussion sur le choix de l'unité de temps : minutes, secondes, heures.* ↶ X **prend toutes les valeurs de** $[0;1]$ (le choix de $[0;1]$ est soutenu par l'enseignante, les élèves proposent plutôt $[0;60]$). *Le temps d'attente maximal est d'une heure.* ! *Discussion sur évènement possible et probabilité nulle/Retour sur la discussion de l'ex aequo.* ! ⋆ Discussion sur la différence avec les probabilités discrètes : *les probabilités discrètes ne marchent plus.* ⋆ T_K **et** T_O **suivent le même type de loi que la variable aléatoire de position de la puce du DM6 (loi uniforme continue).** ↷ X **la variable aléatoire qui représente le temps d'attente du premier arrivé.** ↶ $X = \lvert T_K - T_O \rvert$ ↶ X **est à valeurs dans** $[0;1]$**.** ↶ On cherche à déterminer $P(X \in [a;b])$, pour a et b appartenant à $[0;1]$ ↷ ∼ 4min ! Une élève propose que la variable aléatoire X soit modélisée par le modèle suivant :

Étape du cycle de modélisation		Durée/ lignes	Tâches et activités effectives
		l.633-710	*Proposition du modèle 1 :* $p(X \in [a;b]) = \dfrac{b-a}{1-0}$ ↷ Autre proposition : T_K **et** T_O **suivent le même type de loi que la variable aléatoire de position de la puce du DM6 :** *C'est comme s'il y avait deux puces.*
3a	↓	∼ 20min l.711-892	Tâche : Comment peut-on valider ou invalider le modèle 1 ? ! *Recherche de contre-exemple.* Longue discussion pour argumenter sur ce modèle. Le travail semble s'enliser. (environ 8min) Relance de Marie : Que peut-on faire pour avoir une idée de la valeur de la probabilité que le temps d'attente soit dans l'intervalle $[0; 0,25]$? *Proposition :* ***Simulation de l'expérience aléatoire*** sur le tableur avec la formule : =ABS(ALEA()-ALEA()) ou sur la calculatrice. ↷ Tâche : Proposer une simulation du temps d'attente
			Phase de recherche individuelle (∼ 7min)
			↷ *Proposition : Utilisation du logiciel Excel, avec la formule* =ABS(ALEA()-ALEA()) Manipulation du logiciel Excel par l'enseignante
	Expérimentation		Échantillon de temps d'attente sur Excel.
3b	↓	∼ 2min l.892-919	Question : Que pourrait-on faire avec ces données statistiques ? *Proposition :* ***Construction de l'histogramme de fréquences*** *associé à l'échantillon* L'enseignante change de logiciel et ouvre *GeoGebra*, où un échantillon de taille 10 000 a été mis dans la fenêtre tableur. Elle manipule le logiciel *GeoGebra* pour construire l'histogramme, en expliquant aux élèves ce qu'elle fait.
	Résultats		Histogramme de fréquences

Étape du cycle de modélisation		Durée/ lignes	Propositions effectives
3c	[Analyse détaillée dans la section suivante, p 263]	$\sim$ 28min l.919-1274	*Réfutation de la proposition de modèle 1 à l'aide de l'histogramme.* Reformulation de la tâche principale : Calculer la probabilité $p(X \in [a;b])$, a et $b \in [0;1]$. ⋆ *Proposition :* ***Diminuer de l'amplitude des classes.*** *Proposition :* ***Détermination de la courbe de tendance*** (terme utilisé dans le DM6) : ***La courbe qui « lisse » l'histogramme semble visuellement être une droite*** **donc l'équation de la courbe est de la forme** $y = ax + b$ ↷ Tâche : Tracer votre courbe de tendance.
			Phase de recherche individuelle ($\sim$ 1min)
			Problème : Tout le monde a une courbe différente. Il faut essayer que tout le monde se mette d'accord sur la même courbe. ↷ ***La courbe de tendance est une droite donc l'équation de la courbe est de la forme*** $y = ax + b$. Tâche : Trouver les conditions pour déterminer a et b.
			Phase de recherche individuelle ($<$ 1min)
			Si on appelle f la fonction représentée par la courbe de tendance, *Proposition : La droite passe par* $(0;2)$*, car « ça le fait plutôt ».* ***Proposition : La droite passe par*** $(1;0)$**.** car la probabilité que le temps d'attente soit proche de 1 est très faible (il y a très peu de possibilités pour les temps d'arrivée des deux amis) L'enseignante introduit la fonction f. $f(0) = 1$ car *« ça veut dire qu'il est 8h donc on ne peut plus attendre ».* ***L'aire sous la courbe doit valoir 1.*** Tâche : Déterminer l'expression de la fonction f.
			Phase de recherche individuelle ($\sim$ 9min) Dès qu'un élève a trouvé l'expression de la fonction, l'enseignante lui distribue l'exercice 1 et il commence sa résolution.
			Tous les élèves finissent par trouver que **les deux équations à deux inconnues donnent** $f(x) = -2x + 2$
D	Modèle probabiliste final		**La courbe de tendance de la variable aléatoire** X **a pour équation :** $y = -2x + 2$

Étape du cycle de modélisation		Durée/ lignes	Propositions effectives
4	Traitement mathématique [Analyse détaillée dans la section suivante, p 263]	$\sim$ 5min l.1274- 1361	Question : A quoi sert cette courbe ? *Avec un autre échantillon, la courbe de tendance aurait été la même.* *Cette courbe permet alors de déterminer $P(a \leqslant X \leqslant b)$: il s'agit de l'aire sous la courbe entre a et b (aire d'un trapèze).*
E/F	Résultats mathématiques et pseudo-concrets		$P(X \in [a;b]) = Aire_{ABCD}$. La probabilité que le temps d'attente soit entre a et b est égale à l'aire du trapèze.
Cette étape commence en décalé suivant les élèves. À ce moment, tous les élèves ont l'énoncé de l'exercice 1.			
C	Modèle pseudo-concret		Exercice 1 En utilisant le modèle mis en place, répondez aux questions suivantes. a) Quelle est la probabilité que le premier arrivé attende exactement 40 minutes ? b) Quelle est la probabilité que le premier arrivé attende plus de 40 minutes ? c) Quelle est la probabilité que le premier arrivé attende moins de 40 minutes ? d) Quelle est la probabilité que le premier arrivé attende entre 20 et 40 minutes ? e) Quelle est la probabilité que Karine et Olivier arrivent en même temps ? f) Finalement, ils ont décidé de ne pas attendre plus d'un quart d'heure. Quelle est la probabilité qu'ils se retrouvent ?
Recherche individuelle ($\sim$ 10min)			
D	Modèle probabiliste		Modèle probabiliste précédent.
4	Traitement mathématique		
E/F	Résultats mathématiques et pseudo-concrets		Les deux se font simultanément
Aucun retour sur le problème réelle n'est vraiment attendu, notamment au niveau de la vérification			

Dans cette séance, environ 35 minutes ont été consacrées à une recherche individuelle. Le reste de la séance a été dédié à une discussion collective. Ceci n'est pas étonnant au vu des prévisions faites. L'objectif de la séance était de construire collectivement la notion de courbe de tendance, en comptant sur la « disponibilité collective » de certaines connaissances en construction. Dans la première phase de mise en commun, le passage de la réalité au modèle pseudo-concret prend environ 8 minutes. Le véritable enjeu de la séance est l'étape de mathématisation (étape 3) qui, elle, prend au total 1h10 de la séance, dont 18 minutes environ de travail de recherche individuel.

Le scénario choisi par la classe est compatible avec le canevas proposé dans l'analyse *a priori*. Tous les moments incontournables se sont effectivement retrouvés dans la séance, mais pas nécessairement aux moments envisagés. Ceci ne nous a pas étonné, nous l'avions anticipé dans l'analyse *a priori*. Cependant, deux éléments n'avaient pas été « prévus ». Le premier est la discussion autour de la possibilité ou non de l'ex aequo. Certains élèves ne veulent pas que les deux protagonistes arrivent en même temps. L'enseignante décide de laisser les élèves en débattre et, sans y répondre, propose d'y revenir plus tard (ce qui n'a pas été fait explicitement, mais seulement en partie dans la correction de l'exercice 1). Le second imprévu est lorsqu'une élève propose un modèle non prévu dans l'analyse *a priori*. L'élève propose que le temps d'attente du premier arrivé suive la même loi que la position de la puce du DM6 (loi uniforme continue). Dans ce second cas, l'enseignante rebondit sur cet imprévu et décide, pour continuer le travail collectif, de s'appuyer sur cette proposition en cherchant à ce que les élèves valident ou invalident ce modèle. Après un moment de flottement et une relance de l'enseignante, des élèves proposent de faire des simulations, ce qui relance les activités.

Nous pouvons remarquer chez les élèves des confusions entre le temps d'arrivée de Karine ou d'Olivier et le temps d'attente. Cazes et Vandebrouck (2014) ont montré dans leur travail sur les tâches complexes que la différence entre le temps qui représente une date (ici le temps d'arrivée) et temps qui représente une durée (ici le temps d'attente) est une difficulté pour les élèves. Nous pensons que parfois un des termes est utilisé à la place de l'autre, ce qui peut créer des quiproquos. Une autre source de confusion, qui laisse parfois planer des doutes sur les interprétations des discussions dans la classe, est l'existence d'une probabilité nulle d'un évènement possible. Nous pensons d'ailleurs que derrière les discussions sur la possibilité ou non de l'ex aequo se cache une compréhension erronée de la situation, de cet ordre. N'ayant pas plus d'informations, nous préférons de pas insister sur ce point.

Au début de la séance (de l'étape A à l'étape C), l'enseignante n'intervient pas auprès des élèves mis à part pour donner la parole, faire répéter ou faire préciser des éléments dits par les élèves. À la seule exception près, pour clôturer une discussion, elle impose que Karine et Olivier ne peuvent pas arriver avant 7h ou après 8h. C'est à partir de la sous-étape 3a qu'elle se met à intervenir pour structurer le débat, notamment en posant des questions plus ciblées, que l'on interprète comme des (sous)-tâches. Nous reviendrons sur les étapes 3c et 4 dans la section suivante donc nous ne nous y attardons pas ici. Nous nous demanderons notamment si Marie guide finalement plus le travail ensuite et si oui, comment.

Au niveau des logiciels, comme prévu, c'est Marie qui les manipule à la demande

des élèves pour *Excel* ou encore à sa propre initiative pour *GeoGebra*.

Le problème de la rencontre est donc une tâche qui est effectivement réalisable en classe, sans que l'enseignante ne semble trop guider les activités. Dans le tableau du scénario, on repère beaucoup de propositions venant des élèves. Pour renforcer ce constat, nous avons identifié qui, des élèves ou de l'enseignante, introduit le vocabulaire relatif aux objets mathématiques en jeu au fur et à mesure. Cela permet de voir qui est à l'initiative de l'avancement des activités de la classe. C'est un moyen de repérer dans la séance entière si ce sont bien les élèves qui sont acteurs dans la résolution du problème. Si le vocabulaire est toujours prononcé en premier par l'enseignante, nous pouvons penser qu'elle influence énormément le travail de la classe. Dans le tableau 7.7, nous indiquons la première fois où est utilisé le vocabulaire relatif aux probabilités (aléatoire, variable aléatoire, probabilité...), à la statistique (fréquence, histogramme...) et au calcul intégral (aire) et par qui. Pour les probabilités, nous ciblerons aussi les expressions moins « rigoureuses » liées à l'aléatoire comme chance, au hasard... car elles sont déjà signe d'une entrée dans le domaine des probabilités.

Élèves	Enseignante
l. 88 : aucune chance	
l. 98 : plein de chance	
	l. 119 : aléatoirement
l. 131 : autant de chance	
l. 132 : 1 chance sur 60	
	l. 136 : au hasard
l. 220 : continu / discret	
l. 246 : probabilités	
l. 282 : P de X	
l. 285 : variable aléatoire	
	l. 521 : histogramme (mais en référence au cours précédent)
l. 826 : faire l'expérience	
l. 828 : simulation	
l. 860 : tirage sur *Excel*	
	l. 894 : distribution de fréquences
l. 895 : histogramme	
	l. 911 : classes
	l. 969 : amplitudes
l. 1003 : courbe de tendance	
	l. 1009 : lisser l'histogramme
l. 1019 : fréquence	
l. 1036 : densité	
l. 1189 : aire	

TABLE 7.7 – Répartition entre les élèves et l'enseignante de l'introduction du vocabulaire

A travers ce tableau, on peut se rendre compte que Marie prend finalement peu à sa charge l'introduction du vocabulaire et nous supposons les activités qui s'en rapportent et laisse le soin aux élèves de le faire. Cela montre une volonté de l'enseignante de ne pas trop intervenir dans l'avancement de la construction du savoir ou tout du moins de ne pas trop l'influencer. Les élèves sont acteurs dans la résolution du problème qui leur est dévolu.

c) Analyse locale du déroulement de l'étape de mathématisation (3) et du traitement mathématique (4)

Nous allons donc maintenant nous focaliser sur les étapes 3 et 4 et plus spécifiquement sur la sous-étape 3c et l'étape 4 (visibles dans le tableau p. 259).

Rapide coup d'œil sur les sous-étapes 3a et 3b

L'analyse de « tâches » faite à la section e) (analyse locale *a priori*) se retrouve dans les activités des mises en commun. Cependant, nous pouvons remarquer qu'un temps assez long (presque 10 minutes) est laissé aux élèves avant que certains d'entre eux aient l'idée de simuler l'expérience aléatoire. Pendant un long moment, quelques élèves s'embourbent dans la recherche d'un contre-exemple. Puis l'enseignante (ligne 824) décide de clôturer cette discussion : « *On va essayer d'avancer* » et demande : « *Qu'est-ce qu'on peut faire puisqu'on a une proposition ? Que peut-on faire pour avoir une idée de cette valeur... de ça ?* ». Alors un élève propose de « *faire l'expérience* ». Une fois la formule trouvée par les élèves, Marie simule un échantillon. Puis ensuite elle passe sur le logiciel *GeoGebra* pour faire un histogramme à la demande des élèves.

Analyse détaillée des étapes 3c et 4

Nous allons maintenant analyser en détail l'étape 3c (après la réfutation du modèle 1) et l'étape 4, qui sont au cœur de notre problématique.

Nous nous situons juste après que l'enseignante, à l'aide de *GeoGebra*, ait tracé l'histogramme d'amplitude 0,25 en figure 7.9. Cet histogramme a notamment permis de réfuter le modèle 1, proposé par une élève.

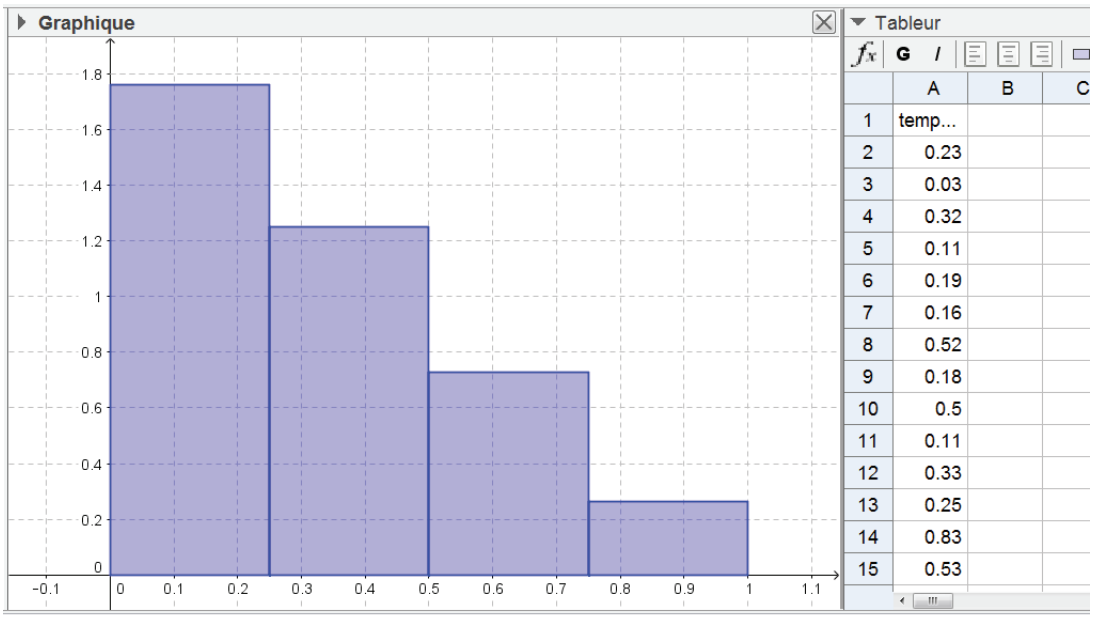
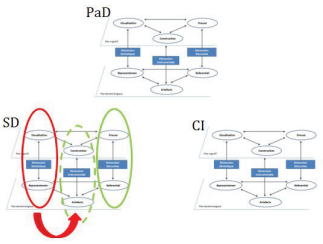


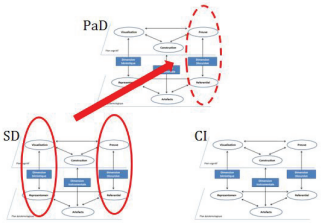
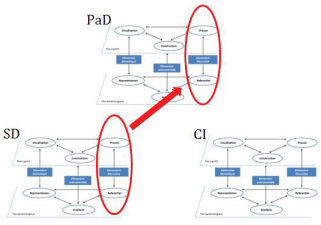
FIGURE 7.9 – Histogramme vidéoprojeté à la fin de la sous-étape 3b

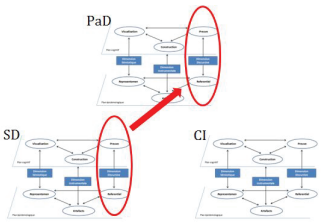
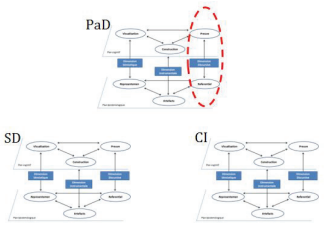
Nous allons ici proposer un zoom sur les deux étapes 3c et 4, que nous allons découper en phases. Dans chaque phase, nous essaierons de mettre en évidence les points forts et la répartition des interventions entre les élèves et l'enseignante. Comme nous l'avons précisé dans la section 7.1.6, du côté des élèves, nous repérerons les propositions faites et les activités. Du côté de l'enseignante, nous identifierons les types d'intervention qu'elle fait (reformulation, demande de précisions...) ainsi que les aides qu'elle apporte (souligné par une vague). Nous identifierons les changements de registre et la disponibilité des connaissances en jeu (en gras) ainsi que les types de (sous)-activités (gras italique). Nous indiquerons aussi les références à l'ancien, notamment aux notions du DM6 (en rouge). Enfin, pour décrire chaque phase, nous proposerons un schéma des trois ETM en jeu montrant les circulations repérées. Les dimensions sollicitées (activées) par les élèves sont entourées en rouge, celles par l'enseignante sont entourées en vert. Nous entourerons en rouge pointillé quand l'intervention des élèves a été sollicitée par une aide de l'enseignante et enfin en vert pointillé lorsque qu'une manipulation par exemple de l'enseignante a été préalablement commandée par un élève.

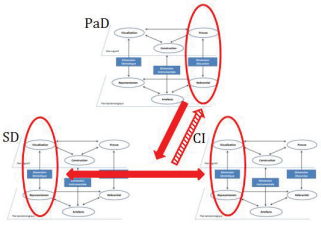
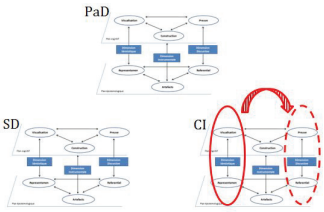
Voici la légende :
italique : extraits de transcription
vague : aides de l'enseignante
gras : changement de registres, disponibilité des connaissances
Gras italique : sous-activités
rouge : référence à l'ancien

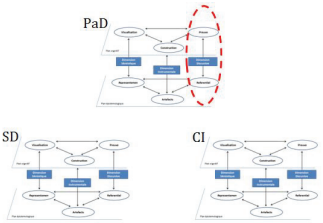
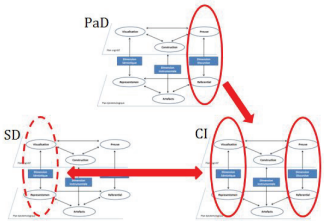
Analyse du déroulement des étapes 3c et 4 (lignes 952-1363) - Séance 1

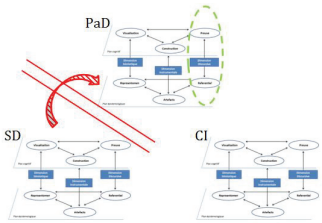
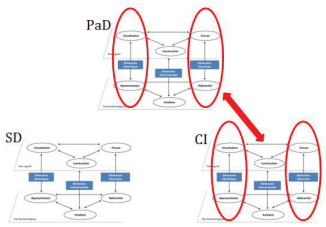
Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
Phase 3c1 - Travail sur l'histogramme			
1 :09 :28 l.952		Reformulation explicite de la tâche principale : – Calculer la probabilité que X appartienne à un intervalle $[a; b]$ inclus dans $[0; 1]$. – Ecrit au tableau.	
		Aide procédurale indirecte : « *Puisqu'on a une jolie petite machine, qu'est-ce qu'on pourrait lui demander de faire ?* »	
	Proposition : Fragmenter.	Demande de précisions.	
	Précision : Prendre des intervalles plus petits. ***Activité stratégique***		
		Demande de précision, pour que l'élève utilise le vocabulaire relatif aux histogrammes.	
		Donne le vocabulaire : Amplitudes, classes plus petites.	Travail exclusivement dans SD. Visualisation de l'histogramme à l'aide du logiciel *GeoGebra*. L'instrumentation est à la charge de l'enseignante. L'enseignante en profite pour rappeler du vocabulaire relatif aux histogrammes (référentiel théorique).
		Question (pour impliquer les élèves) : Que voulez-vous comme largeur [de classes] ?	
	Propositions.	Choix de 0,05.	
		Manipulation du logiciel *GeoGebra*, pour obtenir un histogramme d'amplitudes 0,05. ***Activité de traitement prise en charge par l'enseignante.*** *Registre histogramme sur GeoGebra.*	
		Aide constructive : reformulation, reprécision de ce qui vient d'être fait.	
		Confirmation de la réfutation du modèle 1.	
Phase 3c2 - Phase d'organisation du raisonnement			
1 :11 :44 l.995		Reformulation explicite de la tâche principale, sur un exemple générique : – Évaluer la probabilité que le temps d'attente soit entre 0,2 et 0,3. –	
		Aide procédurale indirecte : Comment feriez-vous avec cet histogramme ? Précision : On ne cherche pas le résultat, mais la méthode pour y arriver.	
	Proposition : Plusieurs étapes.		

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
	1. Tracer une courbe de tendance / 2. Calculer l'aire du triangle / 3. Prendre une amplitude qui correspond à 0,2-0,3. Réf. au DM6 (**connaissances en construction**). ***Activité d'organisation du raisonnement.***		
		Demande de répéter pour que toute la classe entende bien.	
	1. Tracer la courbe de tendance.	Reformulation : Essayer de lisser notre histogramme. Reprise implicite du vocabulaire du DM6. Réf. au DM6	
	2. Calculer l'aire de ce triangle.	Demande de précisions.	
	L'élève se reprend : Ce qui pourrait ressembler à un triangle.		
	L'aire est égale à 1.	Reformulation : Tout.	Lors de cette phase d'activité d'organisation du raisonnement, la notion de courbe de tendance apparaît (par un élève) et des contraintes sur cette courbe émergent : courbe qui « lisse » l'histogramme (par l'enseignante), aire totale sous la courbe égale à 1 (par un élève, qui justifie en s'appuyant sur le référentiel de SD). La construction de ce référentiel théorique est exactement ce qui nous intéresse.
		Reformulation/validation : On n'a pas besoin de calculer l'aire, ça fait 1.	
	Justification : Ça indique la fréquence.	Demande de justification.	
	3. Prendre l'ordonnée, grâce à la courbe de tendance, qui donne la hauteur. Puis multiplier la hauteur par l'amplitude de l'intervalle.	Validation.	
		Demande d'opinions des autres élèves.	
	Proposition : Tracer un trapèze.		
		Réoriente le débat : Ça dépend de votre courbe de tendance.	
Phase 3c3 - Première approche de la courbe de tendance			
l.1027		Question à l'élève : Comment vois-tu cette courbe de tendance ?	
	Proposition : Une droite.		
Phase 3c4 - Aparté : apport théorique			
l.1034	Question d'un élève : Je ne comprends pas ce qu'il y a en ordonnées.		
		Répétition.	
	Réponse d'un autre élève : La densité. Référence implicite au DM6. **Connaissance disponible sur l'histogramme.**		
	Validation de l'élève qui posait la question.		Convocation du référentiel théorique lié aux histogrammes (SD) pour construire implicitement le référentiel théorique des PaD.

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
		Validation.	
Reprise de la phase 3c3			
1 :14 :29 l.1038		Reprise de la proposition précédente. Demande à l'élève de venir tracer la courbe qu'il propose.	
	L'élève trace « dans l'air » une droite, au niveau du haut des rectangles de l'histogramme. **Changement de registre : passage du registre histogramme au registre graphique, par un changement de domaine : passage de SD à CI (analyse).** ***Activité de traitement***		Mobilisation du référentiel théorique de PaD en construction : courbe de tendance. Seule la contrainte de lissage de l'histogramme (implicite) semble être prise en compte dans ce travail de visualisation, qui cherche une correspondance visuelle entre l'histogramme et la courbe.
Phase 3c5 - Bilan sur l'organisation du raisonnement			
1 :15 :40 l.1051		Demande à un autre élève de faire un bilan.	
	1. Tracer une courbe de tendance.	Validation et ajout : L'élève précédent pense que c'est une fonction affine.	
	2. Calculer l'aire.	Répétition.	Lors de cette phase de bilan de l'organisation du raisonnement, la notion de courbe de tendance est reprise par un élève et la notion d'aire apparaît mais c'est l'enseignante qui reprécise que l'aire totale sous la courbe est égale à 1.
	3. Calculer l'amplitude. ***Activité d'organisation du raisonnement (reprise)***	Réfutation : L'aire sous la courbe vaut forcément 1 car c'est la somme des fréquences. L'élève précédent dit qu'il faut calculer l'aire entre 0,2 et 0,3.	
Phase 3c6 - Mise en place de la contrainte de lissage de l'histogramme			
1 :16 :28 l.1061		Sous-tâche explicite : – Tracer votre courbe de tendance (sur l'histogramme distribué). –	
	Phase individuelle. ***Activité de traitement.***		
1 :17 :26 l.1065		Question : Qu'avez-vous tracé ? Prise en compte du fait que certains n'ont pas encore tracé de courbe : demande aux élèves s'ils ont compris ce qu'il fallait faire. Relance le travail. Demande à nouveau s'ils ont tracé leur courbe de tendance.	
	Plusieurs élèves acquiescent.	Demande d'opinion : Avez-vous tous la même courbe ?	
	Réponse : Non.		

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
		Répétition/justification : Non, car vous l'avez tracée au jugé.	
		Question : Par où passe la droite ?	
		Aide procédurale indirecte : Est-ce que vous l'avez tracée ici ? [montre une droite qui n'est pas proche de l'histogramme].	
	Réponse : Non.		
		Relance : On ne peut pas se retrouver avec une courbe de tendance différente.	
		Question : Qu'est-ce qu'on cherche comme genre de fonction ?	Travail individuel similaire à la phase 3c4. Mobilisation du référentiel théorique de PaD en construction : courbe de tendance. Seule la contrainte de lissage de l'histogramme est prise en compte dans ce travail de visualisation, qui cherche une correspondance visuelle entre l'histogramme et la courbe. Cette contrainte est explicitée : construction du référentiel théorique.
	Proposition : Tracer la courbe de sorte qu'il y ait une partie qui soit en haut.		
		Reformulation/validation : Une courbe qui passe au milieu des hauts des rectangles.	
		Relance/demande d'opinion : Même en faisant au mieux, êtes-vous d'accord que chacun va avoir une droite différente ?	
	Réponse : Oui.		
Phase 3c7 - Changement de registre sur la courbe			
l.1115	Proposition : Il faut qu'elle passe par deux points.		
		Relance : On cherche une droite.	
		Aide procédurale indirecte : « *Qu'est-ce qu'on cherche comme genre de chose ?* » (question ambigüe)	
	Proposition : Une équation de droite.		
		Validation/demande de précisions.	Travail dans le sous-domaine CI (ici plus particulièrement sur les fonctions). Changement de registre (passage du registre graphique au registre algébrique), déclenché par l'enseignante, s'appuyant sur le référentiel théorique sur les fonctions affines.
	$y = ax + b$ **Changement de registre dans CI : passage du registre graphique au registre algébrique.** ***Activité de traitement***		
		Ecrit au tableau.	
Phase 3c8 - Vers la deuxième contrainte sur la courbe de tendance			
1 :21 :17 l.1132		Sous-tâche explicite : – Trouver l'équation de la droite. Trouver les conditions pour trouver a et b. –	
	Phase individuelle.		
1 :22 :06 l.1138		Tracé d'une droite « au jugé » sur le logiciel. **Registre graphique.**	
	Proposition : On peut faire passer la droite par des points.		

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
		Question : Y a-t-il des points dont nous sommes sûrs ?	
	Proposition : $(0; 2)$.		
		Demande de justification.	
	Précision : « *Ça a l'air arbitrairement* ».		
		Réfutation : On cherche l'équation de la droite, il nous faut de vrais arguments.	
	Proposition : $(1; 0)$.		
		Introduction de notation : $f(x) = ax + b$. **Changement de registre : registre fonctionnel.** Reformulation : En 1, ça vaut 0. Ecrit : $f(0) = 1$. Demande de justification.	
	Réponse : Parce qu'en 1, ça veut dire qu'il est 8h, donc on peut plus attendre.		
		Répétition/validation.	
		Bilan : On a une valeur.	
		Question : Pour trouver a et b, que faut-il ?	Construction du référentiel théorique de PaD en construction : courbe de tendance. Émergence de la contrainte : l'aire sous la courbe vaut 1, avec l'appui de l'enseignante.
	Proposition : Une deuxième valeur.		
		Aide procédurale directe : Reformulation : Il faut deux renseignements.	
		Relance : On cherche donc un autre renseignement.	
	Proposition : Le point $(0\,;\,2)$ car si on trace le rectangle qui a une aire de 2 et on trace la diagonale du rectangle...		
		Demande de précision : Quel renseignement utilises-tu sans le dire ?	
	L'aire du rectangle.		
		Réfutation.	
		Relance : Tu sais que l'aire sous la courbe doit valoir...	
	1.		
		Répétition/reformulation : L'aire du triangle rectangle doit valoir 1.	
Phase 3c9 - Détermination de l'équation de la courbe de tendance			
1 :25 :46 l.1204		Sous-tâche explicite : – Trouver l'expression de la fonction, par le calcul. –	
	Phase individuelle. ***Activité de traitement.***		
		Prise en compte des élèves : Demande de rappel de la formule de l'aire d'un triangle. **Connaissance non disponible pour l'ensemble des élèves.**	
	Réponse.		

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
		Passage dans les rangs pour valider individuellement les réponses des élèves. Une fois leur fonction validée, elle leur distribue l'exercice 1.	Mobilisation du référentiel théorique des PaD (aire sous la courbe de tendance égale à 1) pour déterminer l'équation de la courbe de tendance. La correspondance visuelle entre la courbe déterminée et l'histogramme reste un critère implicite (vérification).
1 :34 :56 l.1269		Question/relance : Est-ce que tout le monde a l'équation de la droite ?	
	Proposition : $y = -2x + 2$		
		Validation. S'assure que tout le monde l'a tracée sur son histogramme.	
Phase 3c10 - Bilan : vers la rupture entre la statistique descriptive et les probabilités à densité			
1 :35 :30 l.1278		Bilan : Si j'avais relancé le générateur aléatoire, j'aurais obtenu des histogrammes. Ils auraient eu presque la même ... ?	
	Propositions : Forme, allure.		
		Répétition. Question : Peut-on utiliser les histogrammes pour calculer des probabilités ?	
	Proposition : Non car cela changerait à chaque fois.		
		Répétition/reformulation. Question : A quoi ça sert d'avoir trouvé cette droite ?	Mise en avant du lien entre SD et PaD (implicitement loi des grands nombres), tout en créant la rupture entre les deux ETM. Caractère non spécifique de la courbe de tendance par rapport à l'histogramme de l'échantillon particulier choisi.
	Proposition : On a un repère.		
		Reformulation : On a une droite qui modélise tous nos histogrammes en quelque sorte.	
	Proposition : C'est un modèle théorique.		
		Validation.	
Phase 4.1 - Utilité de la courbe de densité			
l.1296		Relance : Que va-t-on en faire ?	
		Rappel de la tâche principale (modification de l'exemple générique) : Calculer la probabilité que X soit entre 0 et 0,25.	
	Proposition : On trace la droite $x = 0,25$, on obtient un petit triangle.		
		Reformulation en traçant la courbe au tableau.	
	Répétition de la proposition.		
		Question : Un triangle ?	
	Se reprend : Trapèze. Précise que l'on peut quand même utiliser l'aire du triangle en faisant la différence entre 1 et l'aire du triangle.		
		Question/bilan : Comment va-t-on calculer la probabilité que X soit entre 0 et 0,25 ?	Construction du lien entre PaD et CI par la construction du référentiel théorique : pour déterminer une probabilité, on calcule l'aire sous la courbe de tendance sur l'intervalle qui nous intéresse (CI niveau A1 ici). Appui important sur la visualisation graphique.
	Proposition : L'aire.		
		Répétition : L'aire.	
	Précision : L'aire du trapèze.		
		Reformulation : Tracé du trapèze.	

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
		Rappel de l'aire d'un trapèze au tableau. **Connaissance non disponible pour l'ensemble des élèves.** Aide procédurale directe : Autre méthode pour calculer l'aire.	
1 :40 :25 l.1363			

Ce tableau détaillant les interventions dans les étapes 3c et 4 permet de mettre en avant plusieurs points. Tout d'abord, nous pouvons remarquer que la tâche principale donnée dans l'énoncé de départ « *Que peut-on dire du temps d'attente du premier arrivé ?* » est reformulée, pour la première fois de la séance, au début de la phase 3c1 par l'enseignante par : « *Calculer la probabilité que X appartienne à un intervalle* $[a; b]$ *inclus dans* $[0; 1]$ ». Elle est à nouveau reformulée par Marie au début de la phase 3c2 sur un exemple générique : « *Évaluer la probabilité que le temps d'attente soit entre 0,2 et 0,3* ». Ensuite, des sous-tâches explicites apparaissent (cf. figure 7.10).

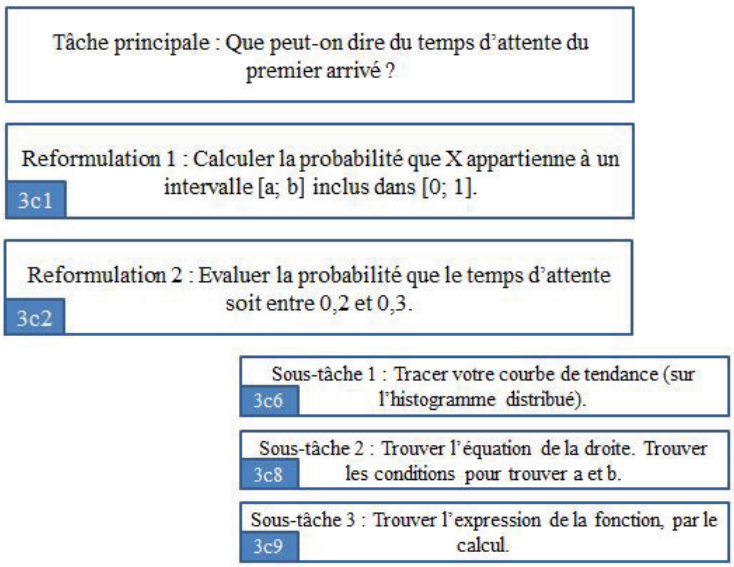


Figure 7.10 – Tâches explicites identifiées dans les étapes 3c et 4

L'analyse nous permet aussi de présenter une chronologie par phases de l'avancée du travail mathématique de la classe au niveau des dynamiques entre les trois ETM en jeu. Ces diagrammes sont des moyens de synthétiser une partie du travail mathématique en jeu dans la phase en prenant en compte la répartition des rôles entre élèves et enseignante. Dans la figure 7.11, nous avons regroupé sous forme de « comics » les schémas des ETM de chaque phase.

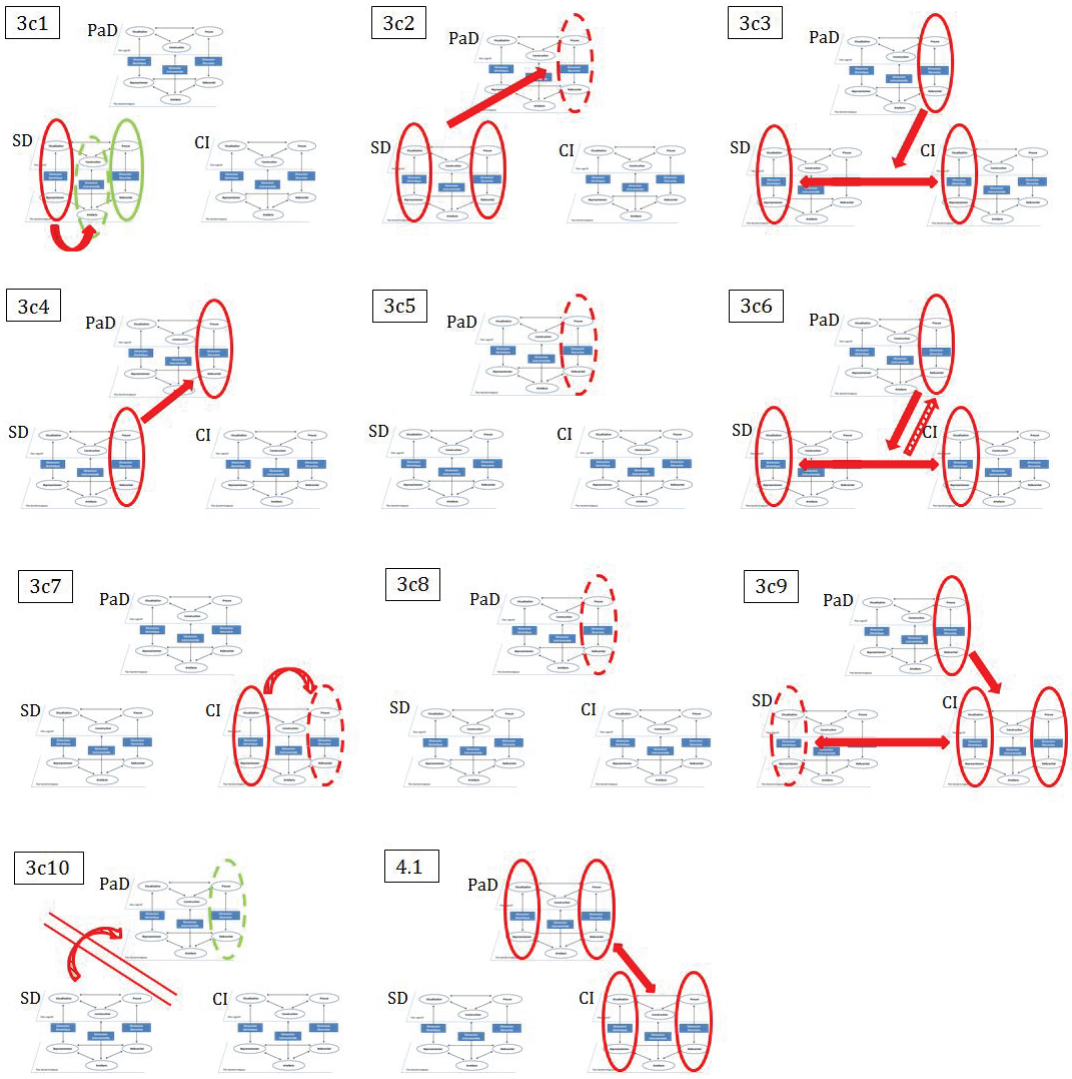


FIGURE 7.11 – Dynamiques entre les ETM dans les étapes 3c et 4

Ce « comics » nous permet de repérer la présence d'un jeu entre les trois sous-domaines que sont la statistique descriptive (SD), les probabilités à densité (PaD) et le calcul intégral (CI). Chaque sous-domaine a son propre rôle dans la construction des connaissances. En partant de la statistique descriptive, avec l'histogramme, il y a une construction au fur à mesure du référentiel des probabilités à densité et du lien entre probabilités et calcul intégral. On remarque que les sous-domaines SD/PaD sont au départ travaillés en analogie (phase 3c2), mais qu'ensuite une phase de bilan (phase 3c10) est proposée pour ne pas aller jusqu'à l'assimilation des deux sous-domaines et bien identifier la rupture existant entre les deux.

Nous pouvons aussi repérer que le travail est en majorité à la charge du collectif classe. Il faut garder à l'esprit que ce ne sont pas tous les élèves qui sont acteurs de l'ensemble des différentes activités, mais en tout cas, ces activités sont principalement à l'initiative des élèves et non de l'enseignante. On peut remarquer, dans certaines phases, que certaines dimensions sont bien convoquées par les élèves, mais

seulement après une intervention de l'enseignante (dimensions entourées en rouge pointillé). C'est notamment le cas parfois au niveau de la dimension discursive relative aux probabilités à densité, où l'enseignante met en place des aides procédurales indirectes pour aiguiller les élèves. Le référentiel théorique lié à la fonction de densité étant en construction, des interventions de l'enseignante étaient prévisibles. Cependant, l'enseignante n'intervient que pour donner des aides procédurales indirectes (sous forme de question), qui ne sont jamais des aides qui donnent des solutions aux élèves, ce sont plutôt des questions qui orientent le débat. Par exemple, à la phase 3c7, Marie oriente les élèves pour changer de registre avec une aide procédurale : « *Qu'est-ce qu'on cherche comme genre de chose ?* ». Bien que la question soit ambiguë, elle permet aux élèves de proposer ce changement de registre pour passer à l'expression algébrique de la courbe. La réponse est cependant donnée par un élève et non par l'enseignante.

Comme nous venons de le dire, la plupart des activités sont menées par les élèves (pas nécessairement tous les élèves) et non l'enseignante. Seules les activités de traitement qui sont liées à une utilisation d'un logiciel sont entièrement prises en charge par l'enseignante, comme prévu. Autrement dit, la dimension instrumentale, particulièrement celle du domaine de la statistique, est prise en charge par l'enseignante, cependant elle est souvent déclenchée par une demande des élèves (entouré en vert pointillé). Les autres activités de traitement, qui se reportent aux sous-tâches repérées dans la figure 7.10, sont laissées à la charge de l'ensemble des élèves, comme le tracé de la courbe de tendance (phase 3c6) et la détermination de l'équation de la courbe de tendance (phase 3c9).

En revanche, nous pouvons repérer que l'activité d'organisation du raisonnement en phase 3c2 est prise en charge uniquement par un seul élève. Cela n'est pas étonnant. Nous avions prévu que les connaissances soient disponibles pour le collectif mais pas pour chaque élève. L'enseignante s'efforce dans cette phase à reformuler les étapes et elle relance, dans la phase 3c5, un autre élève pour faire un bilan sur la démarche. Nous l'interprétons comme une envie de l'enseignante de voir cette activité d'organisation reprise pour qu'elle soit assimilée par tous les élèves. Il semble important pour l'enseignante de prendre du temps pour cela. Cette activité d'organisation du raisonnement pourrait aussi être vue comme une activité de reconnaissance car ici l'élève doit reconnaître la démarche utilisée lors du DM6 (simulation et histogramme), mais adaptée au problème.

Les activités des élèves s'appuient la plupart du temps sur la dimension sémiotique entre SD et CI, ce qui permet à la fois de construire le référentiel théorique et de l'utiliser pour relancer le travail.

La plupart des autres interventions de l'enseignante sont de la reformulation, de la reprise, de la demande de précision, de la validation et une fois de la réfutation (phase 3c5).

Chaque élève ayant construit sa propre courbe dans la phase 3c6, l'enseignante a pu prendre appui dessus pour faire émerger les contraintes que doit vérifier la courbe de densité (et, plus généralement, les courbes de tendance). Cette phase d'« erreur » est importante.

À l'aide de ce comics, il est facile de repérer des moments de blocage de l'avancement, il s'agit effectivement de tous les moments où les dimensions sont entourées

autrement qu'en trait plein rouge. Cependant ce qui est intéressant, c'est de voir que ces moments sont la plupart en rouge pointillé, ce qui veut dire que les dimensions ont été sollicitées par les élèves après une aide indirecte de l'enseignante. Les blocages sont donc la plupart du temps dépassés grâce à une intervention contrôlée de l'enseignante (sous forme d'aide procédurale indirecte). Les différents blocages ont été surmontés par les élèves, du fait d'une bonne anticipation de ces éventuels blocages lors de la phase de conception.

Comme nous l'attendions, nous pouvons remarquer plusieurs références explicites ou implicites aux notions du DM6. Les notions abordées dans le DM6, notamment l'histogramme, le modèle uniforme continue, la simulation, nous semblent ici être un appui indispensable, qui permet aux élèves de rentrer dans le problème et surtout d'avoir des arguments pour certaines justifications (comme par exemple dans la phase 3c4).

Enfin, nous pouvons terminer en disant que nous observons une circulation riche entre les ETM, avec de nombreux allers-retours entre les sous-domaines, prise en charge la plupart du temps par le collectif classe, mis à part pour la dimension instrumentale (dû à un choix de départ). La dimension discursive des probabilités à densité se construit bien au fur et à mesure à l'aide de la dimension sémiotique mais aussi de la dimension discursive de la statistique descriptive.

7.2.4 Analyse de la première institutionnalisation

Au début de la séance 2, l'enseignante propose de commencer par une synthèse de la séance 1, notamment pour permettre aux absents de la séance précédente d'être mis au courant de ce qui s'est fait. Cependant, cela donne lieu à une première institutionnalisation. Nous précisons que ce n'est pas le cours à proprement parlé mais simplement une synthèse écrite dans la partie dédiée aux exercices. La transcription de ce passage se trouve en annexe I.2.

Nous allons prendre appui sur la synthèse écrite au tableau par l'enseignante, en figure 7.12.

Synthèse

On note T_O *le moment d'arrivée d'Olivier*
T_K *le moment d'arrivée de Karine*
T_O *et* T_K *sont des variables aléatoires qui prennent toutes les valeurs réelles de* [7 ; 8] *que l'on peut ramener à* [0 ; 1].
On dit que ce sont des variables aléatoires continues qui suivent le modèle de la puce.
On note X *la variable aléatoire égale au temps d'attente du premier arrivé.*
X *prend toutes les valeurs de* [0 ; 1].
Pour décrire la loi de probabilité de X*, on introduit une fonction* f *définie sur* [0 ; 1] *par* $f(x) = -2x + 2$ *telle que la courbe de* f *« lisse le haut des rectangles de l'histogramme ».*
Ainsi pour calculer la proba $P(X \in [a\,;b])$ *où* $[a\,;b] \subset [0\,;1]$*, on calcule l'aire sous la courbe* C_f *entre* a *et* b.
f *s'appelle la fonction de densité de probabilité de* X.
f *possède les propriétés suivantes :* f *est positive sur* [0 ; 1]*, l'aire sous* C_f *entre* 0 *et* 1 *vaut* 1.

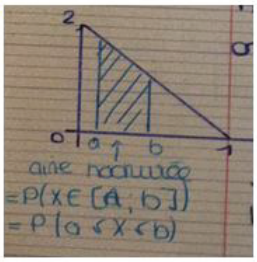
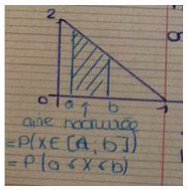


Figure 7.12 – Synthèse de la séance 2

Nous allons maintenant présenter sous forme de tableau le déroulement de cette institutionnalisation en mettant en avant le vocabulaire introduit, ainsi que les éventuelles définitions et propriétés (en gras) et les justifications apportées (en gras italique). Nous mettons aussi en avant qui, des élèves ou de l'enseignante, amène ces nouvelles connaissances. Les citations de la transcription sont écrites en italique. Nous faisons des commentaires pour préciser le degré de généralité employé à l'oral et à l'écrit (parfois souligné dans les citations) ainsi que le niveau de rigueur utilisé.

Analyse de la synthèse de la séance 2

Texte écrit au tableau	Interventions de l'enseignante	Interventions des élèves	Commentaires
Lignes 7-18 de la transcription			
On note T_O le moment d'arrivée d'Olivier *T_K le moment d'arrivée de Karine* *T_O et T_K sont des variables aléatoires qui prennent toutes les valeurs réelles de $[7\,;\,8]$ que l'on peut ramener à $[0\,;1]$.*	Rappels de l'énoncé du problème (reformulé), des notations T_O et T_K. **Vocabulaire :** **Variable aléatoire**		Reprise du problème de la rencontre. Le vocabulaire de variable aléatoire est connu des élèves depuis la classe de seconde.
Lignes 18-28			
On dit que ce sont des variables aléatoires continues qui suivent le modèle de la puce.	Une variable aléatoire qui prend toutes les valeurs réelles d'un intervalle donné, on l'appelle comment ? Précision : variable aléatoire continue.	**Vocabulaire :** **Variable continue**	Vocabulaire générale, dont la définition orale est donnée dans un contexte général, cependant l'écrit est quant à lui contextualisé au problème. Perte de généralité entre l'oral et l'écrit. Le vocabulaire avait déjà été introduit lors de la séance 1.
Lignes 32-80			
On note X la variable aléatoire égale au temps d'attente du premier arrivé. *X prend toutes les valeurs de $[0\,;1]$.* *Pour décrire la loi de probabilité de X, on introduit une fonction f définie sur $[0\,;1]$ par $f(x) = -2x + 2$*	Rappels sur la variable aléatoire X, sur la démarche mise en place. Les élèves donnent les idées générales (modéliser, simuler, trouver la courbe de tendance) et ensuite l'enseignante reprécise.		Les rappels à fonction procédurale ne sont pas écrits, cependant ils ont tout de même aussi une portée constructive (implicite).

Texte écrit au tableau	Interventions de l'enseignante	Interventions des élèves	Commentaires
Lignes 80-93 *fonction f définie sur $[0;1]$ par $f(x) = -2x + 2$ telle que la courbe de f « lisse le haut des rectangles de l'histogramme ».*	Demande de « définition » de la fonction f	**« Définition » :** *La courbe lisse le haut des rectangles de l'histogramme.*	Vocabulaire employé non rigoureux, mais cela est assumé par l'enseignante : *Est-ce que tout le monde comprend l'idée en tout cas à défaut de le dire de façon facile ?*
Lignes 99-148 *Ainsi pour calculer la proba $P(X \in [a;b])$ où $[a;b] \subset [0;1]$, on calcule l'aire sous la courbe C_f entre a et b.* 	*A quoi ça nous sert d'avoir trouvé cette jolie petite fonction ?* *Comment est-ce qu'on calcule cette probabilité ?* **Vocabulaire :** **Aire sous la courbe** Exemple graphique pour illustrer. Introduction des notations.	**Propriété :** Pour calculer des probabilités quand X appartient à un certain intervalle. Calcul de l'aire du triangle ou du trapèze.	Ce sont les élèves qui expliquent l'utilité de la fonction. On peut voir que les élèves et l'enseignante ne sont pas au même degré de généralité au niveau de l'aire sous la courbe. Les élèves restent sur l'exemple du problème de la rencontre, tandis que l'enseignante parle de l'aire sous la courbe de façon plus générale. Introduction du vocabulaire nouveau, aire sous la courbe, par l'enseignante. Avec un manque de rigueur dans la définition assumé par l'enseignante : *On formalisera ça dans le cours.* Fort appui sur le graphique.

Texte écrit au tableau	Interventions de l'enseignante	Interventions des élèves	Commentaires
Lignes 149-166			
f s'appelle la fonction de densité de probabilité de X.	**Vocabulaire :** **Fonction de densité de probabilité** L'enseignante reprécise en faisant implicitement référence au DM6.	***Justification des élèves :*** L'axe des ordonnées représente la densité.	Introduction du vocabulaire nouveau par l'enseignante, mais la justification de ce vocabulaire est donnée par les élèves.
Lignes 167-251			
f possède les propriétés suivantes : f est positive sur $[0;1]$*, l'aire sous* C_f *entre 0 et 1 vaut 1.*	Elle propose des contre-exemples graphiques pour faire apparaître la contrainte de positivité (aide à fonction constructive indirecte). [Détails plus loin]	**Propriété :** Plusieurs élèves trouvent cette première propriété (positivité). ***Justification :*** par rapport à l'histogramme. **Propriété :** La fonction doit être continue.	

Texte écrit au tableau	Interventions de l'enseignante	Interventions des élèves	Commentaires
	Débat sur la continuité. L'enseignante clôt le débat en disant : *Alors on va pas trancher ça maintenant mais effectivement les deux cas sont possibles. Mais dans le cadre du cours, nous, elles seront continues.*		Les propriétés de la fonction de densité sont dégagées par les élèves, mais l'enseignante proposent des aides indirectes pour que le maximum d'élèves les trouvent, notamment pour la positivité. Nous insisterons sur ces graphiques après ce tableau. Il semble à nouveau y avoir une différence de degré de généralité entre l'oral et l'écrit. L'enseignante appuie à l'oral sur la différence de statut du critère de continuité de la fonction de densité et d'ailleurs cette propriété n'apparait pas à l'écrit.
	Et alors par contre, une autre propriété qu'on a utilisé d'ailleurs ?		
		X supérieure à 0.	
	L'enseignante court-circuite cette proposition.		
		Propriété : *L'aire en-dessous de la courbe vaut 1.* ***Justification :*** *Parce que ça correspond à la fréquence.*	
	Recontextualise au problème. ***Justification :*** Analogie avec la somme des fréquences.		

Nous pouvons nous apercevoir que l'enseignante redonne rapidement l'énoncé du problème avec ses propres mots et introduit directement les notations qu'ils avaient choisies T_O et T_K, représentant les variables aléatoires associées aux temps d'arrivée d'Olivier et Karine, puis fait le choix de commencer la synthèse directement par l'étape 3 de mathématisation du cycle de modélisation. Les étapes précédentes ne sont pas du tout mentionnées dans la synthèse.

À l'aide du tableau, nous pouvons repérer que le vocabulaire « aire sous la courbe » et « fonction de densité de probabilité » est introduit dans la classe par l'enseignante. Ce vocabulaire étant entièrement nouveau (jamais il n'a été utilisé en amont), il est évident que seule l'enseignante pouvait l'introduire. Pour l'aire sous la courbe, Marie décide de ne pas la définir rigoureusement, mais la commente par un graphique. Cela est pleinement assumé devant la classe. Elle précise que la notion d'aire sous une courbe sera définie proprement dans le cours. Pour le vocabulaire « fonction de densité de probabilité », elle demande aux élèves de le justifier, ce qu'ils font :

> Marie : Ça a l'air affreux ça : fonction de densité... Pourquoi à votre avis ça s'appelle comme ça ou est-ce que c'est parce que ça fait bien de mettre un truc moche ?
> *Quatre/cinq élèves lèvent la main.*
> Marie : B31.
> Élève B31 : C'est parce que sur l'axe des...
> Marie : Oui, je suis en train d'essayer de les remettre les histogrammes.
> Élève B31 : C'est parce que sur l'axe des ordonnées, c'est la densité.
> Marie : Oui c'est ça le lien... dans le devoir à la maison, on a vu que lorsqu'on représentait un histogramme sur l'axe des ordonnées c'est la densité qui était portée, de chaque classe, donc forcément votre fonction $f(x)$ vous imaginez, vous l'avez fait, vous avez un histogramme avec la courbe dessinée dessus enfin sur l'histogramme, et bien sur l'axe des ordonnées c'est bien la densité des histogrammes et donc c'est pour ça que ça s'appelle la fonction de densité de probabilité. (lignes 155-166)

Un accent est mis ici pour donner du sens au terme « fonction de densité », comme cela était souhaité dans les objectifs de la séquence. Une première « définition » de cette fonction a été donnée par les élèves en amont (« *La courbe lisse le haut des rectangles de l'histogramme* »). Ensuite, Marie cherche à ce que les élèves exhibent les propriétés que doit vérifier une fonction de densité :

> Marie : Alors cette fonction f qu'est-ce qu'elle doit quand même un petit peu... Est-ce qu'elle est... Elle peut être quelconque ? Ou est-ce qu'il y a quand même des contraintes dessus ? Est-ce que toute fonction pourra être une fonction de densité de probabilité ou est-ce qu'il faut qu'elle possède tout de même quelques propriétés ? (lignes 166-169)

Malgré le fait que plusieurs élèves lèvent la main, elle préfère ne pas leur donner la parole tout de suite pour que plus d'élèves réussissent à les identifier. Elle propose alors, ce que nous nommons une aide à fonction constructive indirecte, des contre-exemples graphiques (cf. figures 7.13 et 7.14) pour impliquer un maximum d'élèves.

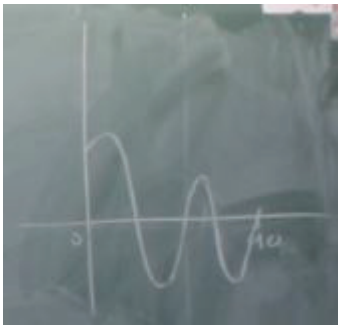


Figure 7.13 – Premier contre-exemple

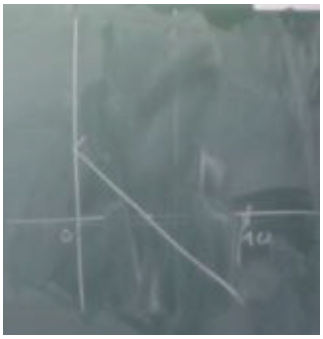


Figure 7.14 – Deuxième contre-exemple

Le premier contre-exemple a le désavantage de proposer une courbe de niveau A3 (à l'heure actuelle les élèves ne savent pas calculer ce type d'aires). De ce fait, cela bloque l'élève interrogé : « *On peut pas trop calculer d'aire* ». Marie lui répond qu'il a raison et propose alors la deuxième courbe pour faire ressortir ce qui l'intéresse : le fait qu'une fonction de densité doit nécessairement être positive. Une discussion suit ensuite sur la continuité nécessaire ou non d'une telle fonction, où les élèves proposent des arguments à l'aide d'histogrammes. Marie conclut en disant que la continuité n'est pas nécessaire mais qu'en terminale, tous les exemples rencontrés seront continus. Enfin, la condition que l'aire sous la courbe soit égale à 1 vient en dernier avec une justification par analogie avec la fréquence.

De cette analyse, nous pouvons conclure que, pour cette première synthèse, l'enseignante souhaite poser le vocabulaire rencontré lors de la séance 1, comme variable aléatoire continue, et introduire le nouveau vocabulaire, tel que fonction de densité. C'est aussi le moment pour faire expliciter aux élèves les propriétés d'une telle fonction. Cette institutionnalisation se déroule sur un jeu de questions/réponses entre l'enseignante et les élèves.

Nous pouvons remarquer que cette première institutionnalisation reste contextualisée au problème de la rencontre, même si de la généralité est parfois introduite (surtout à l'oral). Au niveau des propriétés liées à la fonction de densité, nous pouvons repérer une réelle généralité, car mis à part la propriété de l'aire sous la courbe égale à 1 qui a été utilisée lors du DM6 et du problème de la rencontre, les deux

autres n'ont pas été utilisées en action. Il s'agit donc ici de les faire émerger directement (d'où les contre-exemples proposés par Marie). Nous pouvons remarquer le statut différent de la propriété de continuité. En particulier, dans cette première synthèse, l'enseignante décide de ne pas écrire la propriété de continuité de la fonction de densité. Ce décalage entre écrit et oral peut s'expliquer par le fait que le but avéré au départ par l'enseignante est de faire une synthèse du problème de la rencontre pour les élèves absents, c'est pourquoi l'écrit reste contextualisé. Cependant, une ambition plus grande est de commencer la décontextualisation du savoir. Le problème de la rencontre prend finalement le statut d'exemple générique. Pour autant, l'enseignante a bien conscience qu'il ne peut pas y avoir un trop grand écart entre ce qui a été fait et le général visé, pour qu'il y ait une mise en relation entre ce qu'ils ont fait à la séance 1 et l'institutionnalisation ici. On peut d'ailleurs voir qu'un trop grand écart entre ce qui a déjà été fait et ce qui est institutionnalisé peut créer un blocage. C'est le cas lorsque Marie propose un premier contre-exemple avec une fonction de niveau A3, les élèves ne voient pas la généralité ici mais simplement une impossibilité de calculer l'aire, ce qui bloque alors la poursuite de la réflexion. Le passage se débloque quand elle propose un deuxième contre-exemple de niveau A1.

Nous pouvons aussi ajouter un dernier détail au niveau de la différence de contenu entre écrit et oral. Il n'est écrit que les éléments relatifs au vocabulaire, aux définitions et aux propriétés. Alors qu'à l'oral, des justifications sont données et il est fait référence aussi à la méthode mise en place pour résoudre le problème.

Cette synthèse a donc permis d'introduire plusieurs termes de vocabulaire, tous initiés et/ou justifiés par les élèves. Il n'est pas ici question d'imposer du vocabulaire ou des propriétés sans justification. La statistique descriptive est très importante pour permettre des justifications par les élèves. Nous pouvons remarquer que l'enseignante fait le choix d'introduire le vocabulaire, tout en restant tout de même en contexte. Le but de cette synthèse est finalement de tirer du problème de la rencontre les connaissances mises en place, plus ou moins implicitement. Nous pouvons tout de même dire que l'objectif final est de dégager les propriétés relatives à la fonction de densité, même si la décontextualisation n'est pas encore totale. L'enseignante ménage l'entrée dans la généralité, certains points sont totalement généralisés, d'autres non, de plus l'écrit reste contextualisé au problème de la rencontre.

Finalement, la construction du référentiel théorique de l'ETM idoine, à la fois relatif aux probabilités à densité mais aussi relatif au calcul intégral avec une première amorce, se fait en grande partie par les élèves, qui s'appuient sur le problème de la rencontre et les notions du DM6. L'ETM idoine effectif se construit donc en connectant les trois sous-domaines qui nous intéressent (un appui sur la statistique descriptive est remarqué) et permet de construire réellement des éléments du référentiel théorique et des liens notamment entre calcul intégral et probabilités à densité, et non de les imposer. Cependant, le niveau de généralité maximal n'est pas encore atteint.

7.2.5 Conclusion sur le problème de la rencontre

Dans cette conclusion, nous voulons mettre en évidence les points importants de ces trois niveaux d'analyse (globale, locale et institutionnalisation) qui permettent de donner des premiers éléments de réponse à nos six sous-questions relatives à la question de recherche QR3 (cf. p. 221).

Tout d'abord, nous voulons commencer en disant que le problème de la rencontre est effectivement réalisable en classe, avec le type de démarche espéré dans l'analyse *a priori*. Le scénario choisi par la classe s'est effectivement révélé inclus dans le canevas réalisé dans l'analyse globale *a priori*. Pour réellement convenir de la viabilité de ce problème dans la classe, il nous semble important de préciser que le déroulement s'est passé dans de « bonnes conditions », à savoir que globalement les élèves étaient investis et qu'ils ont contribué à l'avancée du travail. Le problème a bien débouché sur une institutionnalisation, ce qui laisse à première vue penser que l'objectif d'introduire la notion de fonction de densité a bien été atteint.

Cependant, cela ne suffit pas. C'est la raison pour laquelle nous avons prolongé nos analyses. Notamment, nous avons analysé plus finement les passages concernés sur l'introduction de la notion de fonction de densité (étapes 3c et 4 du cycle de modélisation) et le moment d'institutionnalisation qui a suivi le problème. L'analyse locale permet de voir que le travail mathématique en jeu lors des passages relatifs à la fonction de densité repose sur des circulations entre les trois sous-domaines des probabilités à densité, du calcul intégral et de la statistique descriptive. Ceci est clairement visible dans le « comics » proposé en figure 7.11 (p. 272). Nous qualifions les circulations entre les trois ETM comme étant riches, car nous repérons des allers-retours importants entre les sous-domaines. Ces circulations sont justement un appui pour faire apparaître au fur et à mesure des éléments du référentiel théorique des probabilités à densité, éléments introduits et utilisés. En plus de la présence de circulations entre les sous-domaines, nous pouvons repérer que les différentes dimensions de chaque ETM (dimension sémiotique, dimension instrumentale et dimension discursive) sont presque toutes sollicitées, sauf la dimension instrumentale des ETM des probabilités à densité et du calcul intégral. Pour les probabilités, il n'était effectivement pas dans la conception de la séquence de favoriser un passage par la dimension instrumentale, qui enlève justement toute la réflexion sur les objets en jeu. Pour le calcul intégral, l'utilisation de logiciel pour la recherche de la fonction n'est pas du tout utile ici. Le travail mathématique en jeu est quand même ce que nous appelons quasi complet. La dimension discursive n'est en tout cas pas mise de côté et bien au contraire. La construction (petit à petit) de la dimension discursive de l'ETM_{PaD}, qui est l'enjeu principal du problème de la rencontre, prend notamment appui sur le référentiel théorique de l'ETM_{SD}, ce qui n'était pas le cas dans les manuels par exemple. Le « comics » permet aussi de révéler que l'ETM collectif est en grande partie responsable de la construction de l'ETM idoine effectif et donc du référentiel théorique relatif aux probabilités. Cependant, l'enseignante joue tout de même un rôle dans cette construction. Mis à part le fait qu'elle prenne à sa charge la dimension instrumentale (ce qui était prévu), l'enseignante est parfois le déclencheur de certaines dimensions. Nous avons notamment repéré la présence d'aides. Cependant, ces aides procédurales sont indirectes, par le biais d'une question souvent. Ces questions arrivent comme une structuration pour l'avancée du travail. Elles sont

présentes pour orienter le débat, mais jamais aucune solution n'est apportée aux élèves. On peut remarquer d'ailleurs que certaines idées arrivent assez tardivement (par exemple, la proposition de faire une simulation). Ces aides permettent aux élèves de dépasser les blocages qui surviennent. Il est important de mentionner que c'est l'anticipation sur ces éventuels blocages, lors de la phase de conception, qui ont permis à l'enseignante de les gérer de cette façon.

Dans les étapes précédentes, nous pouvons identifier une différence dans la gestion de la classe par l'enseignante. En effet, au départ, Marie intervient très peu. Elle reprend simplement les idées et laissent les élèves débattre entre eux, sans apporter de réponse en cas de désaccord. L'enseignante laisse les élèves s'exprimer, sûrement pour permettre qu'ils se familiarisent avec le problème, qu'ils s'investissent, qu'ils se posent des questions... Au moment de la mathématisation, une gestion de la classe « plus guidée » semble nécessaire.

Notre étude de l'ETM idoine à l'aide de l'ETM collectif ne nous permet pas de différencier chaque élève. Cependant, nous remarquons que certaines activités sont attendues de tous les élèves. Par exemple, chaque élève doit avoir tracé la courbe de tendance (phase 3c6), qui sera différente pour chaque élève, pour notamment déclencher le besoin de trouver les propriétés que doit vérifier cette courbe. Puis chaque élève doit déterminer l'expression de la fonction de densité (qui ne porte pas encore ce nom), à la phase 3c9. En amont tout le monde ne participe pas nécessairement à toutes les activités. Nous avons pu remarquer que l'activité d'organisation du raisonnement (phase 3c2) est prise en charge par un élève. Cependant, elle est ensuite reprise par l'enseignante puis par un autre élève pour, nous semble-t-il, la rendre disponible à l'ensemble de la classe. Il nous semble que ces activités sont importantes pour que les élèves puissent essayer d'assimiler les propriétés, ici essentiellement que l'aire sous la courbe doit être égale à 1.

L'institutionnalisation nous permet enfin d'identifier quel savoir ressort effectivement de ce problème et notamment s'il y a effectivement une construction (ou au moins un début de construction) du référentiel théorique. L'analyse a mis en avant que les élèves ont bien construit des connaissances lors de la résolution du problème de la rencontre. Le vocabulaire nouveau est bien entendu amené par l'enseignante, mais toujours justifié par les élèves eux-mêmes. Par exemple, l'introduction du terme « fonction de densité de probabilité » est accompagné d'une justification donnée par les élèves eux-mêmes, ce qui donne un poids au sens qu'ils lui donnent. Ces justifications sont fortement appuyées sur le sous-domaine de la statistique descriptive, ce qui laisse entendre l'importance de ce domaine dans le déroulement du problème de la rencontre. De plus, on voit ici apparaître les propriétés que doit vérifier une telle fonction toujours à l'initiative des élèves. Cela montre une réelle implication des élèves dans le problème, mais aussi que ce problème, précédé par le DM6, permet de faire émerger la notion de fonction de densité. Cette première institutionnalisation est toutefois restée contextualisée au problème, et essentiellement à l'oral sont apparus quelques idées de généralisation. Nous pouvons donc conclure qu'à la fin de ce problème le référentiel théorique n'est pas encore totalement arrêté. Nous pouvons aussi remarquer qu'une première approche du calcul intégral est enclenchée avec la notion d'aire sous la courbe représentative d'une fonction. Bien que l'on reste dans le niveau A1, nous sommes déjà dans du calcul intégral.

Bien que non totalement achevée, nous venons de le voir, la construction du référentiel théorique par le biais de l'ETM collectif est bien enclenché. Les éléments que nous venons de pointer en sont la preuve.

Pour conclure, certains éléments et certaines conditions se sont révélées importantes pour aboutir à cet objectif. Tout d'abord, nous citerons les passages que nous avons présenté comme incontournables dans nos analyses globales. Il nous semble effectivement que dans le canevas de l'analyse globale *a priori* des passages semblent obligatoires et se sont d'ailleurs bien retrouvés dans le déroulement. Il s'agit des points suivants :

- choix du modèle associé aux temps d'arrivée,
- simulation,
- construction d'un histogramme,
- recherche de la fonction de densité.

Ensuite, cela était prévisible, mais a été fortement confirmé par le déroulement, le devoir maison n° 6 joue un rôle très important. L'appui sur la statistique descriptive notamment n'est réellement possible que grâce à ce devoir maison, ou tout du moins, un travail en amont sur l'histogramme. De plus, la rencontre avec la loi uniforme (problème de la puce) est essentiel pour ce problème.

Les autres contraintes et conditions qui permettent d'arriver à l'objectif visé sont liées à la gestion par l'enseignante. Il s'agit de laisser la classe, donc l'ETM collectif, construire l'ETM idoine tout en la guidant dans les moments importants. Les aides procédurales indirectes sont ici les moments qui permettent de relancer le travail de la classe, cependant sans donner la réponse. Les interventions autres que des aides jouent aussi un rôle comme les répétitions et les reformulations qui permettent à l'enseignante d'appuyer sur des points importants. Les demandes de justification ou de précision servent aussi à la construction de l'ETM idoine.

Cette conclusion intermédiaire, relative au problème de la rencontre, donne des éléments de réponse à nos questions de recherche. Nous pouvons repérer une construction en cours de la notion de fonction de densité. Les analyses du problème du volcan vont permettre de compléter nos résultats.

7.3 Le problème du volcan Aso

Dans cette section, nous allons présenter le détail de nos analyses *a priori* et *a posteriori* relatives au problème du volcan Aso, qui fait suite au problème de la rencontre. De la même manière que précédemment, nous présenterons nos analyses *a priori* globale puis locale (section 7.3.1), puis nos analyses *a posteriori* (section 7.3.2). Nous aborderons ensuite l'analyse de la deuxième institutionnalisation, qui a suivi le problème du volcan Aso (section 7.2.4). Enfin, nous donnerons les résultats de nos analyses relatives à ce problème dans le cadre de nos questions de recherche.

7.3.1 Analyse *a priori* du problème du volcan Aso

a) Rappels du contexte

Pour rappel, le problème du volcan Aso est le second problème d'introduction de la séquence. Il a lieu au cours des séances 3 et 4. Il arrive après le problème de la rencontre et la synthèse présentés antérieurement. L'objectif de ce problème est double. Dans un premier temps, le but est de consolider les propriétés que doit vérifier une fonction de densité et de renforcer le lien entre probabilités à densité et calcul intégral, notions qui ont été dégagées dans le premier problème. Ce problème permet de rencontrer un nouveau modèle, cette fois-ci à partir de données réelles. Dans ce cas, la fonction en question sera définie sur un intervalle non borné, de plus les calculs d'aires ne seront plus de niveau A1. Le second objectif est donc ensuite de motiver le besoin du calcul d'aire sous une courbe, ce qui débouchera sur une nouvelle problématique : déterminer une méthode pour approcher une aire sous la courbe (de niveau A3). Cette seconde tâche ne sera présentée que succinctement, dans la section 7.3.3.

b) L'énoncé

Le document distribué aux élèves (cf. annexe G.4) est un tableau regroupant les 73 années des différentes éruptions du volcan Aso (Japon) entre 1229 et 1897. La troisième colonne précise aussi le temps d'attente entre deux éruptions consécutives.

Les questions qui accompagnent ces données sont les suivantes :

> Le volcan Aso est actuellement en éruption.
> Comment évaluer la probabilité que la prochaine éruption :
>
> 1. ait lieu dans 5 ans ?
> 2. au cours de l'année 2030 ?

L'énoncé et les questions sont écrits avec des termes simples. Il s'agit, à partir de données réelles sur le passé, de « prédire » l'avenir. C'est une problématique de statistique inférentielle. Nous voulons à l'aide de données réelles connues, faire une estimation. La consigne de l'énoncé est ouverte. La méthode à utiliser n'est pas du tout explicitée, ni même suggérée. De plus, plusieurs réponses à ce problème seraient envisageables. Il est attendu des élèves qu'ils finissent par modéliser la situation par une loi exponentielle, mais d'autres lois pourraient être choisies. Nous rappelons que la loi exponentielle n'a pas encore été rencontrée par les élèves et le but va être de déterminer l'expression de la fonction de densité compatible ici. Il s'agit à nouveau ici de procéder à une démarche de modélisation. Mais elle sera légèrement différente de la précédente. Le terme « *Comment* » laisse bien penser que ce qui importe le plus, dans ce problème, est la manière d'y arriver plutôt que le résultat en lui-même.

c) Gestion prévue

Le temps prévu pour ce problème ainsi que la deuxième partie (méthode de calcul d'aire) est d'environ 2h. À nouveau, toutes les manipulations informatiques se feront par l'enseignante sur l'ordinateur de la salle. Il s'agira à nouveau d'utiliser le logiciel *GeoGebra*. Il va être utile non seulement pour la construction d'histogrammes mais aussi pour la recherche de la fonction de densité. Il va notamment permettre de déterminer des aires sous la courbe de niveau A3, que les élèves ne savent pas (encore) calculer. Un polycopié avec un histogramme d'amplitude 4 ans sera distribué pendant la séance. Nécessairement, le choix d'une amplitude de 4 ans dans ce polycopié va jouer dans le déroulement.

Comme pour le problème de la rencontre, il est prévu des phases de recherche individuelle (ou à deux) et des phases de mise en commun.

d) Cycle de modélisation associé au problème du volcan Aso

Comme le problème de la rencontre, le problème du volcan Aso nécessite une démarche de modélisation. Cependant, le cycle de modélisation du problème du volcan est différent du premier rencontré. Cela est dû au fait qu'ici nous avons affaire à des données réelles (et non truquées). Dans les données que nous avons, des choix ont déjà été faits, par exemple le fait que seules les années d'éruptions soient précisées (et non la date précise) et donc que les temps d'attente entre deux soient donnés en années. Il faut ensuite choisir comment interpréter ces données réelles pour proposer un modèle pseudo-concret (étape C). Il faut donc se demander comment considérer la donnée « k années ». Est-elle à voir comme une donnée isolée ou comme représentant un intervalle ? Lequel ? Il est assez naturel de considérer le temps d'attente entre deux éruptions comme continu. Ce qui est sûr c'est qu'en réalité une donnée k peut représenter n'importe quelle valeur de l'intervalle $]k-1;k+1[$, cependant cela entraîne des chevauchements. Il faut donc faire un choix arbitraire qui ne fasse pas de chevauchements. Plusieurs choix seraient possibles. Le choix que nous avons fait sur les polycopiés qui seront distribués aux élèves est de considérer k comme l'intervalle $]k-1;k[$. Nous nous situons donc dans le modèle pseudo-concret : nous avons déjà arrangé la réalité, en faisant des choix. Ce choix nous permet ensuite de représenter ces données statistiques par un histogramme de fréquences (SD), à l'aide du logiciel *GeoGebra*. Nous pouvons passer aux probabilités (PaD) en conjecturant un modèle probabiliste, par le biais d'une fonction de densité. Cette fois encore, il va falloir prendre en considération les contraintes d'une telle fonction. Ici, la statistique arrive en amont des probabilités. Cela est dû au fait que nous sommes dans une problématique de statistique inférentielle. À partir de données statistiques, nous devons faire des inférences. Le cycle de modélisation est donc nécessairement différent. La démarche attendue ici est celle présentée en figure 7.15. Une remarque, la limite entre la partie pseudo-concret et mathématique est moins claire que dans le premier problème, mais ce n'est pas important pour ce moment de l'analyse.

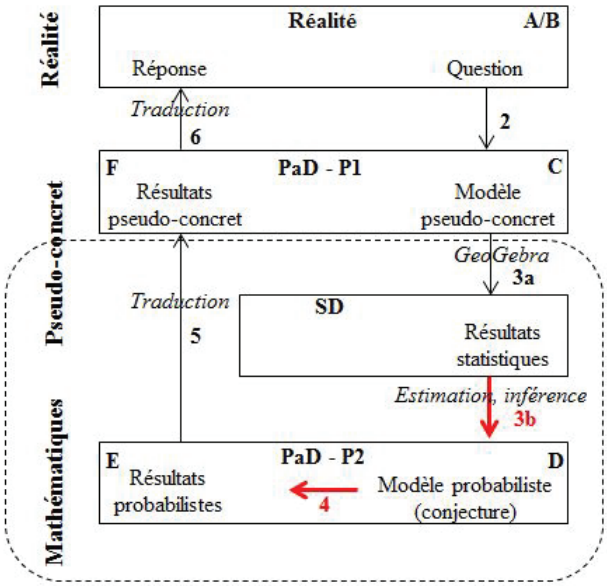


FIGURE 7.15 – Cycle de modélisation complété du problème du volcan Aso

Dans ce problème, ce sont les étapes 3b et 4 qui vont particulièrement nous intéresser.

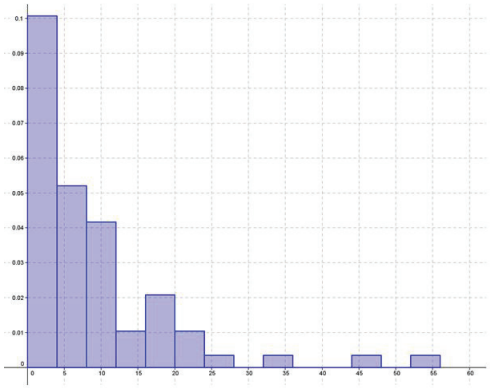
Contrairement au premier problème, plusieurs modèles seraient envisageables. Il faut garder en tête toutefois que le choix de modèle qui sera fait reste en fait une conjecture. Il faudrait vraiment revenir à la réalité et confronter le modèle à la réalité pour légitimer ou non ce modèle, ce qui ne nous intéresse pas ici. Nous tenons à souligner qu'un modèle n'est pas unique et son choix peut dépendre de multiples facteurs : simplicité et précisions attendues, facilité des calculs, possibilités d'en tirer des résultats (problème crucial par exemple pour les économistes), meilleure adéquation aux données statistiques...

e) Analyse globale *a priori* : description du canevas du problème

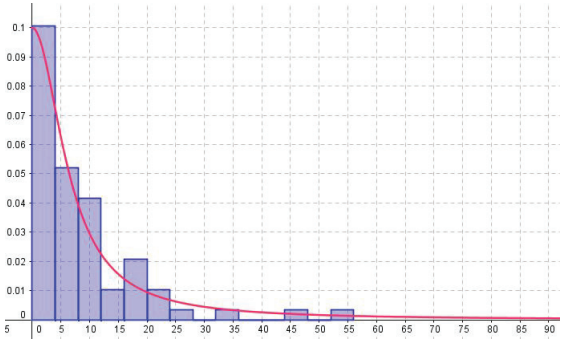
Nous allons dans le tableau ci-dessous, de la même façon que pour le problème de la rencontre, présenter le canevas du problème du volcan Aso. Pour rappel, nous identifions les tâches et les activités envisageables, dans un sens large. Nous mettrons une étoile ($\star$) devant les activités dites stratégiques, où des choix relatifs à la modélisation sont à faire (choix de modèles par exemple). Nous essayons aussi de repérer des passages qui nous semblent incontournables. Ces éléments seront représentés en caractères gras. Nous ajouterons, en rouge, les références possibles à l'ancien (par exemple, le DM6). Nous précisons aussi les blocages envisagés. Les étapes numérotées par des lettres sont des états à un moment du cycle, nous leur faisons donc correspondre des contenus mathématiques (ou pseudo-concret) arrêtés à ce stade du cycle (nous mettons aussi les énoncés distribués par exemple).

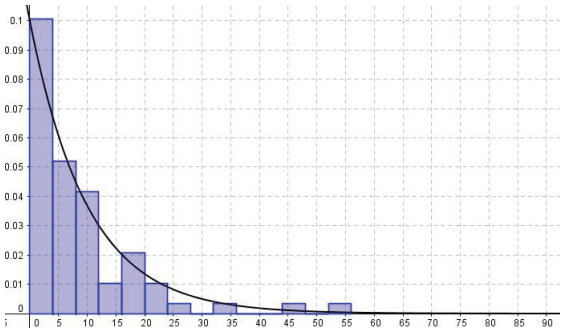
Seulement ensuite, nous nous focaliserons plus particulièrement sur les étapes 3b et 4 qui ont été identifiées comme les passages clés.

Canevas du problème du volcan Aso		
Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
A/B	Situation réelle / Situation retenue	Le volcan Aso est actuellement (le 18 mars 2015) en éruption. À partir des données des éruptions du volcan entre 1229 et 1897, comment évaluer la probabilité que la prochaine éruption : 1. Ait lieu avant 5 ans ? 2. Ait lieu au cours de l'année 2030 ?
1/2	Compréhension de la tâche/ Simplification Idéalisation/Structuration	⋆ **Un questionnement sur ce que représente un temps d'attente de k est nécessaire.** Étant donné qu'ici le temps d'attente est la différence entre deux années civiles, k est un entier qui peut en réalité désigner n'importe quelle valeur de l'intervalle $]k-1;k+1[$ (cela dépend si l'éruption se produit le 1er janvier ou le 31 décembre). **Le caractère aléatoire du problème doit ressortir.** **Le caractère continu du temps d'attente entre deux éruptions doit émerger, pour ensuite choisir de le modéliser par du continu.** ⋆ **Il faut absolument faire un choix sur la façon d'interpréter une donnée k.** Plusieurs choix sont envisageables : $]k-1/2;k+1/2[$, $]k-1;k]$, $[k;k+1[$ ou plus compliqué, il pourrait être envisagé pour chaque valeur k de décider, par exemple à pile ou face, si elle doit être associé à l'intervalle $]k-1;k]$ ou $[k;k+1[$. Bien entendu, aucune des propositions n'est véritablement le « bon » choix, mais pour continuer, il faut tout de même choisir. Dans le polycopié qui sera distribué plus tard aux élèves, il est fait le choix de l'intervalle $]k-1;k]$, ce qui va sûrement influencer le déroulement. [Remarque : Il faudrait en fait avoir des données plus précises pour être au plus proche de la réalité.]
C	Modèle pseudo-concret	Le temps d'attente entre deux éruptions est aléatoire et sera modélisé par du continu. Une des interprétations proposées au-dessus est à choisir.
3 3a	Mathématisation	 A la main, il peut être proposé un tableau d'effectifs puis de fréquences regroupées par classes. Ce tableau peut ensuite permettre de tracer un histogramme de fréquences des données.

Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
	↓	Cependant, il n'est pas attendu que les élèves fassent leur histogramme à la main (choix pour ne pas perdre trop de temps). **La construction des histogrammes sera ensuite prise en charge par l'enseignante, avec le logiciel *GeoGebra*** (les données ont été rentrées à l'avance dans le logiciel). ⋆ Plusieurs amplitudes peuvent être testées. **Il faut ensuite en choisir une pour continuer le travail**. A nouveau, le polycopié peut influencer les choix : l'histogramme prévu a une amplitude de 4 ans.
	Résultats	Un histogramme de fréquences doit être obtenu.  On peut remarquer que l'histogramme présente des trous, liés au fait que ce sont des données réelles et que nous ne disposons que de 73 données. [Remarque : Ce « problème » pourrait être supprimé si nous utilisions un logiciel permettant de construire des histogrammes à pas non constants. Cela appuie sur ce défaut de *GeoGebra*].
3b	[Analyse détaillée dans la section suivante, p. 295]	**La variable aléatoire X correspondant au temps d'attente entre deux éruptions doit être introduite.** **Il faut ensuite que soit proposé de déterminer une courbe de tendance** (ou une courbe de densité, le vocabulaire ayant été introduit en séance 2).

Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
	↓	**La fonction de densité est à considérer sur** $[0;+\infty[$**.** Cela peut poser des problèmes conceptuels. Cependant, il semble difficile de choisir une borne supérieure finie car pourquoi un temps supérieur ne pourrait-il pas survenir ? Pour le choix de cette courbe, **le premier critère reste une courbe qui « lisse » l'histogramme** (quelque soit l'amplitude choisie). ⋆ Au vu de l'allure de l'histogramme, plusieurs propositions peuvent apparaître, avec les fonctions de référence que les élèves connaissent : – La première idée qui semblerait pouvoir venir des élèves est une fonction de la famille de la fonction $x \mapsto \frac{1}{x}$, c'est-à-dire de la forme $x \mapsto \dfrac{a}{bx+c} + d$. – Il pourrait être proposé une fonction de la famille de $x \mapsto \frac{1}{x^2}$, bien que cette proposition nous paraisse déjà moins probable. – Enfin, une fonction de la famille $x \mapsto e^{-x}$ est à espérer. **Pour les fonctions de la forme** $x \mapsto \dfrac{a}{bx+c} + d$**, la contrainte de l'allure permet déjà de faire un premier tri, ensuite la contrainte de la positivité de la fonction en élimine d'autres. Il reste alors à prendre en compte l'aire sous la courbe. Il est nécessaire d'introduire des curseurs sur *Geogebra* pour pouvoir tester les différentes fonctions.** **Pour l'aire sous la courbe, il sera proposé aux élèves d'utiliser une commande sur *GeoGebra* pour obtenir l'aire sous la courbe. La commande est l'Inspecteur de fonction, qui permet de donner la valeur de l'aire sous la courbe notamment. Il suffit pour cela d'entrer les bornes de l'intervalle.** Ici on s'intéresse à l'aire d'un domaine infini, cela peut ouvrir le débat sur la possibilité au non que cette aire soit finie (à mettre en lien avec le DM5). Les valeurs des aires sous la courbe, toujours supérieures à 1, permettent de réfuter le choix des fonctions de la forme $x \mapsto \dfrac{a}{bx+c} + d$.

Étape du cycle de modélisation	Tâches et activités envisageables ou **incontournables**
↓	Pour les fonctions de la famille de $x \mapsto \frac{1}{x^2}$ le même type de raisonnement est possible. Des fonctions de ce type peuvent convenir mais elles sont difficiles à trouver par tâtonnement. Par exemple, la fonction $x \mapsto \dfrac{0,1}{0,024x^2+1}$ pourrait être proposée. Remarque : 0,024 est en fait une valeur approchée de $(0,1 \times \frac{\pi}{2})^2$.  Nous avons toutes les raisons de penser que cette fonction ne sera pas proposée. Enfin la troisième famille qui « doit » apparaître est la famille des fonctions $x \mapsto e^{-x}$, c'est-à-dire de la forme $x \mapsto \alpha e^{-\beta x}$. α correspondant à la valeur de la fonction en 0 peut être estimée à 0,1 (ou proche). Ensuite, un travail sur *GeoGebra*, notamment avec un curseur, peut permettre de déterminer β. La fonction $x \mapsto 0,1e^{-0,1x}$ vérifie bien tous les critères pour être une fonction de densité associée à ces données.

Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
	↓	 Il faut bien avoir conscience que ce n'est pas la seule. Le but ici est d'en trouver une. Il s'agit en fait de la loi exponentielle de paramètre $0,1$ (loi pas encore rencontrée par les élèves).
D	Modèle probabiliste	La démarche permet d'arriver au fait qu'une **fonction de densité de la variable aléatoire X, correspondant au temps d'attente entre deux éruptions, peut être :** $f(x) = 0,1e^{-0,1x}$. Cette proposition qui nous semble la plus probable de ressortir dans la classe, mais elle n'est effectivement pas unique.
4	Traitement mathématique [Analyse détaillée dans la section suivante, p. 295]	**La courbe de densité va permettre de déterminer $P(0 \leqslant X \leqslant 5)$: il s'agit de l'aire sous la courbe entre 0 et 5. Cependant, cette aire ne peut pas être déterminée à la main donc la valeur va être donnée par le logiciel, avec l'Inspecteur de fonction.** Pour la deuxième question, il faut savoir quand considérer le début de l'éruption actuelle pour savoir quelle probabilité calculer. **Si l'on prend le jour où est traité le problème (le 18 mars 2015) comme le début de l'éruption, on veut déterminer $P(15 - \frac{76}{365} \leqslant X \leqslant 16 - \frac{77}{365})$, c'est-à-dire l'aire sous la courbe entre 14,79 et 15,79 (arrondis à 10^{-2} près).**

Étape du cycle de modélisation		Tâches et activités envisageables ou **incontournables**
E/F	Résultats mathématiques et pseudo-concrets	$P(0 \leqslant X \leqslant 5) \simeq 0,39$ et $P(14,79 \leqslant X \leqslant 15,79) \simeq 0,0217$. La probabilité que la prochaine éruption soit dans moins de cinq ans peut être estimée à environ 0,39. La probabilité qu'elle ait lieu au cours de l'année 2030 peut être estimée à environ 0,0217.
Il pourrait y avoir un retour à la réalité, notamment en questionnant le modèle. On pourrait ajouter par exemple les données statistiques du XX^{e} et XXI^{e} siècles. Mais ici, ce n'est pas attendu.		

Cette analyse globale nous permet de repérer des passages qui nous semblent incontournables pour le bon déroulement du scénario effectif. En résumé, il s'agit des points suivants :

– choix sur l'interprétation des données,
– construction d'un histogramme,
– recherche de la fonction de densité.

Ces « passages obligés » sont indépendants des issues des activités stratégiques, qui dans ce problème se retrouvent au niveau du choix de l'interprétation des données et ensuite dans le choix de la fonction de densité. Nous repérons des appuis sur l'ancien, à savoir le DM5 pour justifier l'aire finie d'un domaine infini. implicitement ou explicitement, il peut aussi y avoir des références aux notions vues dans le DM6 et aux deux séances sur le problème de la rencontre. Nous allons maintenant nous focaliser sur les étapes du cycle de modélisation qui nous intéressent tout particulièrement : les étapes 3 et 4.

f) Analyse locale *a priori* de la mathématisation (étape 3) et du traitement mathématique (étape 4)

Comme nous l'avons précisé dans la présentation du cycle de modélisation associée à ce problème (cf. section d), p. 287), ce sont dans les étapes 3 et 4 que se passe la construction de la notion de fonction de densité, même plus spécifiquement ici dans les passages 3b et 4 (dans le tableau p. 290). Nous proposons ici une analyse *a priori* plus fine de ces passages, en prenant en compte la méthodologie présentée dans la section 7.1.6.

Une fois le modèle pseudo-concret arrêté dans le paradigme PaD-P1, il s'agit de construire, à partir des données réelles ainsi que des hypothèses faites sur le modèle, l'histogramme de fréquences correspondant (SD). On va retrouver les différentes étapes mises en œuvre dans le problème de la rencontre. Cette activité d'organisation du raisonnement (construire un histogramme, trouver une courbe de densité, puis déterminer l'aire sous la courbe) ayant déjà été rencontrée lors de la séance 1, va sûrement être disponible pour un bon nombre d'élèves.

Pour la construction de l'histogramme, c'est à nouveau la dimension instrumentale avec l'utilisation de *GeoGebra* qui sera activée. Cela n'empêche pas, dans un premier temps, de laisser les élèves eux-mêmes travailler à la main sur les données et pourquoi pas essayer de construire l'histogramme. Il a été fait le choix, en amont, avec Marie de ne pas attendre des élèves qu'ils construisent à la main l'histogramme, à la fois par gain de temps et aussi parce que le logiciel permet de changer les amplitudes facilement et rapidement. Un histogramme d'amplitude 5 ans pourrait, par exemple, permettre de faire une première estimation de la probabilité que la prochaine éruption ait lieu dans moins de 5 ans.

Une fois un histogramme choisi, il est prévu de distribuer aux élèves un polycopié avec l'histogramme d'amplitude 4 ans. La recherche d'une fonction de densité peut démarrer. Nous nous situons à l'étape 3b. Il s'agit du passage de SD à PaD-P2. Ici, contrairement au problème de la rencontre, il ne va pas s'agir d'une fonction affine.

L'utilisation du logiciel *GeoGebra*, en d'autres termes la mobilisation de la di-

mension instrumentale, va dans ce cas jouer un rôle primordial pour la détermination d'une fonction de densité. Un jeu entre les dimensions sémiotiques de la statistique et du calcul intégral est à prévoir et va être possible grâce à la dimension instrumentale, ce qui n'était pas le cas pour le problème de la rencontre où l'utilisation du logiciel a été cantonné à la construction d'histogrammes. Ici, le logiciel va être un véritable outil pour la recherche d'une fonction satisfaisante. Le référentiel théorique relatif aux fonctions de densité va aussi permettre d'alimenter le travail sémiotique.

Un appui sur le problème précédent est attendu. Les connaissances à mobiliser sur les fonctions de densité sont toujours en construction, bien qu'une première institutionnalisation ait eu lieu à la séance précédente. Les connaissances sur les fonctions de référence sont des connaissances anciennes supposées disponibles, bien que récentes par exemple pour l'exponentielle.

Nous proposons un diagramme en figure 7.8 présentant un cheminement possible entre les différentes notions, dans lequel les strates horizontales font référence aux sous-domaines en jeu.

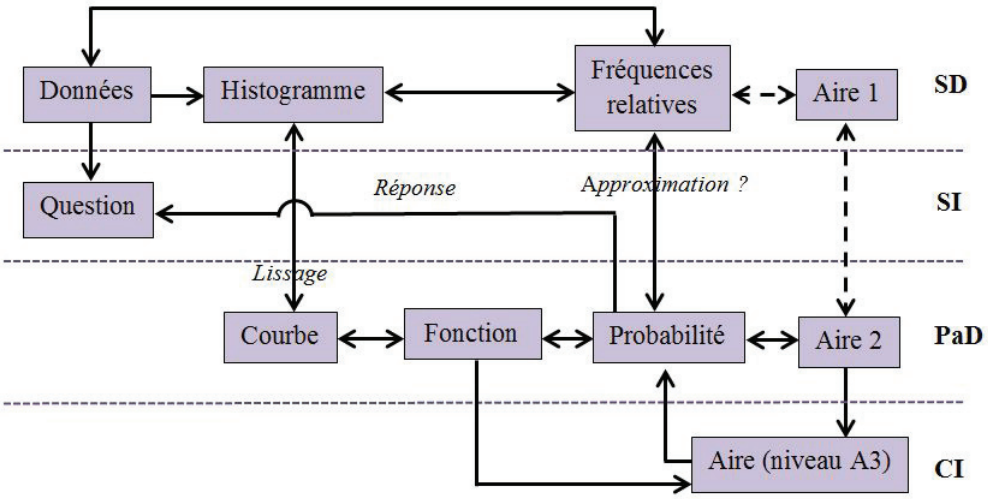


FIGURE 7.16 – Cheminement dans le problème du volcan Aso

Maintenant que les analyses *a priori* du problème du volcan Aso sont faites, nous allons nous intéresser au déroulement effectif des séances associées. Nous proposerons, comme pour le problème de la rencontre, une analyse *a posteriori* complétées par des analyses relatives notamment à la gestion de ces séances.

7.3.2 Analyse du déroulement

Après ces analyses *a priori* du problème du volcan Aso, nous allons maintenant analyser le déroulement effectif. Nous allons en décrire les analyses *a posteriori*, enrichies par des compléments. Nous toucherons des points qui échappent à l'ana-

lyse *a priori* comme la gestion des activités par l'enseignante, notamment via ses interventions. L'analyse sera d'abord globale puis locale.

a) Déroulement général

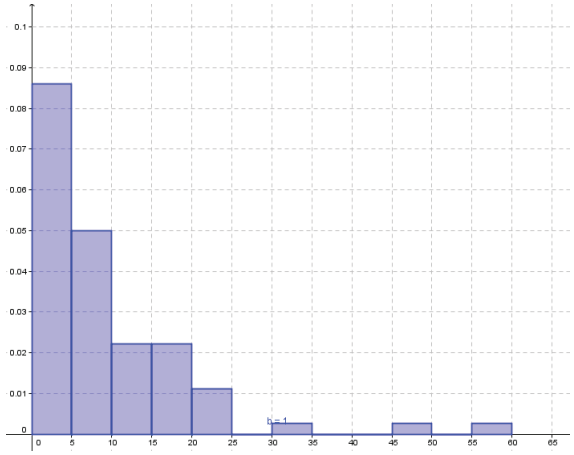
Le problème du volcan Aso s'est déroulé au cours de la séance 3 (55 minutes) et de la séance 4 (les 25 premières minutes), donc a duré environ 1h20. Avec la seconde partie dédiée à la recherche d'une méthode pour calculer les aires sous la courbe, le tout a pris environ 2h30, ce qui est légèrement plus que ce qui était prévu, tout en sachant que nous étions assez souples sur le déroulement. Nous allons ici nous intéresser à la séance 3 et aux 25 premières minutes de la séance 4, ce qui correspond au traitement du problème du volcan Aso. Nous aborderons ensuite rapidement, dans la section 7.3.3, l'institutionnalisation qui a suivi et ensuite la seconde tâche dévolue aux élèves.

b) Analyse globale *a posteriori* : scénario choisi

En nous appuyant sur les transcriptions de la séance 3 (cf. Annexe I.3) et de la séance 4 (cf. Annexe I.4), nous avons pu construire le tableau suivant, présentant le scénario choisi par la classe pour ce problème. Nous rappelons que nous avons gardé comme base le canevas construit pour l'analyse *a priori*. Nous tenons cependant à préciser que l'intégralité des éléments de l'analyse *a priori* n'a pas été gardée mais seulement les points principaux, dans le but d'alléger la lecture. Nous gardons les mêmes codages que pour le problème de la rencontre.

Nous rappelons la légende :
italique : activités dues à l'intervention d'un élève
vague : activités dues à l'intervention de l'enseignante
gras : passages jugés incontournables lors de l'analyse *a priori*
gris seulement : activités qui ne sont pas apparues dans le déroulement
gris avec ↷ : activités qui sont apparues plus tard dans le déroulement
↷ noir : nouvel emplacement des activités ci-dessus
gris avec ↶ : activités qui sont apparues plus tôt dans le déroulement
↶ noir : nouvel emplacement des activités ci-dessus
rouge : référence à l'ancien
⋆ astérique : activités stratégiques

Analyse *a priori* du problème du volcan Aso			
Étape du cycle de modélisation		Durée/ lignes	Tâches et activités effectives
Séance 3 (transcription en annexe I.3)			
Distribution de l'énoncé. Dévolution : lecture à voix haute d'un élève et commentaires. ($\sim 4min30$)			
A/B	Situation réelle / Situation retenue		Le volcan Aso est actuellement (le 18 mars 2015) en éruption. À partir des données des éruptions du volcan entre 1229 et 1897, comment évaluer la probabilité que la prochaine éruption : 1. Ait lieu avant 5 ans ? 2. Ait lieu au cours de l'année 2030 ?
Phase de recherche individuelle ou à deux ($\sim$ 14min, l. 46-136)			
1/2	Compréhension de la tâche/ Simplification Idéalisation/Structuration		**Un questionnement sur ce que représente un temps d'attente de k est nécessaire.** Étant donné qu'ici le temps d'attente est la différence entre deux années civiles, k est un entier qui peut en réalité désigner n'importe quelle valeur de l'intervalle $]k-1;k+1[$ (cela dépend si l'éruption se produit le 1er janvier ou le 31 décembre). **Le caractère aléatoire du problème doit ressortir.** **Le caractère continu du temps d'attente entre deux éruptions doit émerger, pour ensuite choisir de le modéliser par du continu.** **Il faut absolument faire un choix sur la façon d'interpréter une donnée k.** Plusieurs choix sont envisageables : $]k-1/2;k+1/2[$, $]k-1;k]$, $[k;k+1[$ ou plus compliqué, il pourrait être envisagé pour chaque valeur k de décider, par exemple à pile ou face, si elle doit être associé à l'intervalle $]k-1;k]$ ou $[k;k+1[$. Bien entendu, aucune des propositions n'est véritablement le « bon » choix, mais pour continuer, il faut tout de même choisir.
C	Modèle pseudo-concret		Le temps d'attente entre deux éruptions est aléatoire et sera modélisé par du continu. Une des interprétations proposées au-dessus est à choisir. ↷
3 3a	Mathématisation	 $\sim$ 2min l.137-150	 *Proposition : Regrouper les données et de les mettre dans un tableau d'effectifs et/ou de fréquences.*

Étape du cycle de modélisation		Durée/ lignes	Tâches et activités effectives
			Proposition : À partir de ce tableau, construire ensuite un histogramme de fréquences des données.
	↷ Étapes 1/2	∼ 3min l.151-192	Discussion engagée par l'enseignante pour dégager quels types de données permettent de construire un histogramme (différence entre diagramme en bâtons et histogramme). Le caractère continu des données est dégagé. *Proposition : Le temps d'attente est continu.* ⋆ Discussion sur l'interprétation d'un temps d'attente de un an.
	↓	∼ 4min l.193-252	⋆ *Proposition : Prendre des classes d'amplitude 5 ans.* ⋆ *Proposition : Prendre des intervalles du type* $]0;5]$*, car il n'y a pas de 0 dans les données.* **Construction de l'histogramme d'amplitude 5 ans, prise en charge par l'enseignante, avec le logiciel *GeoGebra*** (les données ont été rentrées à l'avance dans le logiciel). Plusieurs amplitudes peuvent être testées. **Il faut ensuite en choisir une pour continuer le travail.** ↷
	Résultats		Un histogramme de fréquences d'amplitude 5 ans est obtenu. 

Étape du cycle de modélisation		Durée/ lignes	Tâches et activités effectives
3b	[Analyse détaillée dans la section suivante, p. 305]	∼ 27min l.253-836	**La variable aléatoire X correspondant au temps d'attente entre deux éruptions doit être introduite.** ↷ **Il est proposé de trouver une courbe de tendance** (ou une courbe de densité, le vocabulaire ayant été introduit en séance 2). ! *Problème : Ça va être compliqué car il y a des trous (dans l'histogramme).* *Proposition : Prendre des amplitudes à pas non constant pour avoir un histogramme « plus harmonieux ».* *Proposition : Négliger les trois rectangles à droite (« valeurs aberrantes »).* *Un élève vient montrer, à main levée sur l'histogramme, la courbe qu'il envisage (de type exponentielle décroissante).* ↷ L'enseignante change l'amplitude a 4 ans. *On retrouve la « même » courbe de tendance.* (Distribution du polycopié avec l'histogramme d'amplitudes 4 ans). Rappels sur le vocabulaire lié aux fonctions de densité, lancés par l'enseignante. ***La fonction de densité est considérée sur*** $[0;+\infty[$, après qu'ait été considéré $[0;55]$ ou encore la question de la fin du monde. Tâche : On cherche une fonction f définie sur $[0;+\infty[$ qui approche « au mieux » l'histogramme.
			Phase de recherche individuelle (∼ 4min, l. 415-478)
			⋆ *Première proposition (reprise par l'enseignante à partir des résultats de la phase individuelle) : la fonction $x \mapsto \frac{1}{x}$, puis ensuite les fonctions de cette famille.* À l'aide d'un travail sur *GeoGebra* (manipulé par l'enseignante), cette proposition a fini par être rejetée en prenant en compte les propriétés de la fonction de densité. (Réf. au DM5) Il pourrait être proposé une fonction de la famille de $x \mapsto \frac{1}{x^2}$, bien que cette proposition nous paraisse déjà moins probable. ⋆ *Deuxième proposition : la fonction $x \mapsto e^{-x}$.* *Proposition : $f(0) = 0,1$.* *Proposition : $f(x) = e^{-x} - 0,9$.* Réfutée car la fonction devient négative.

Étape du cycle de modélisation		Durée/ lignes	Tâches et activités effectives
			Proposition : $f(x) = 0,1e^{-x}$. *Problème : Sa courbe représentative « descend trop vite ».*
Fin de la séance 3			
			Travail à faire à la maison : Tâche : À l'aide du logiciel *GeoGebra*, déterminer la fonction de densité à partir de l'histogramme, et répondre aux questions de départ.

Étape du cycle de modélisation		Durée/ lignes	Tâches et activités effectives
Séance 4 (transcription en annexe I.4)			
3b	[Analyse détaillée dans la section suivante, p. 305]	$\sim$ 9min l.10-176	Reprise du travail sur la fonction exponentielle, avec GeoGebra. *Proposition validée :* $f(x) = 0,1e^{-0,1x}$.
D	Modèle probabiliste		La démarche permet d'arriver au fait qu'une **fonction de densité** de la variable aléatoire X, correspondant au temps d'attente entre deux éruptions, **peut être :** $f(x) = 0,1e^{-0,1x}$.
4	Traitement mathématique [Analyse partielle dans la section suivante, p. 305]	$\sim$ 15min l.177-418	*Proposition :* ***Évaluer la probabilité en calculant l'aire sur l'intervalle considéré.*** $\curvearrowright$ **Introduction de la notation X variable aléatoire représentant le temps d'attente entre deux éruptions.** *Proposition :* $P(0 \leqslant X \leqslant 5)$ *correspond à l'aire sous la courbe sur l'intervalle* $[0;5]$. ! *Proposition : Utiliser la commande Integrale du logiciel GeoGebra (aide donnée dans l'énoncé du travail à faire à la maison) pour déterminer l'aire.* **Remarque : ce n'est pas la commande Inspecteur de fonction du logiciel *GeoGebra* qui a été donnée aux élèves.** **Pour la deuxième question, il est considéré que le début de l'éruption actuelle est le jour où a commencé à être traité le problème, c'est-à-dire le 18 mars 2015.** Cependant, des erreurs de calculs se sont glissés. *Proposition : Calculer* $P(14,75 \leqslant X \leqslant 15,75)$, au lieu de $P(14,79 \leqslant X \leqslant 15,79)$.
E/F	Résultats mathématiques et pseudo-concrets		$P(0 \leqslant X \leqslant 5) \simeq 0,39$ et $P(14,75 \leqslant X \leqslant 15,75) \simeq 0,02$ (au lieu de $P(14,79 \leqslant X \leqslant 15,79) \simeq 0,0217$) (invisible). La probabilité que la prochaine éruption soit dans moins de cinq ans peut être estimée à environ 0,39. La probabilité qu'elle ait lieu au cours de l'année 2030 peut être estimée à environ 0,0217.

Ces deux séances, réalisées en classe entière, voient alterner des phases de recherche individuelles (ou à deux) et des phases de mises en commun et de réflexion collective. Dans la première séance consacrée à ce problème (séance 3), 18 minutes ont été consacrées à la recherche individuelle, dont 14 minutes tout au départ. Les 40 minutes suivantes et les 25 premières minutes de la séance 4 sont des temps collectifs. On peut cependant ajouter un temps de travail à la maison entre les deux séances. Lors de la première phase de recherche individuelle, il est intéressant de remarquer que l'enseignante, au départ, propose 3-4 minutes de travail individuel (ligne 46) qui finalement se transforment en 14 minutes. Marie nous expliquera que les élèves étaient tous lancés dans une recherche, donc elle ne voulait pas les arrêter dans leur démarche. Elle aurait même finalement aimé les laisser aller jusqu'au bout de leurs idées, notamment les laisser construire leurs tableaux et histogrammes. Cependant, par manque de temps et par soucis de ne pas modifier la gestion prévue, elle décide à un moment de clôturer ce temps pour passer à une mise en commun (ligne 137).

Contrairement au problème de la rencontre, qui était le premier dans le genre, les activités des élèves commencent ici très rapidement. Alors que chaque étape de la démarche de modélisation arrivait petit à petit avec des temps plus ou moins longs à chaque fois pour le problème de la rencontre, ici on s'aperçoit qu'il y a directement un passage à la mathématisation (étape 3 du cycle de modélisation). Les éléments de l'étape 1/2 se retrouvent imbriqués dans l'étape 3 (mais sont présents). L'activité d'organisation du raisonnement global relatif au problème semble devenue disponible pour la plupart des élèves, peut-être devrions nous dire mobilisable car le fait que les deux problèmes se suivent dans le temps influence nécessairement l'activité des élèves. Nous repérons que 9 minutes sont consacrées à la sous-étape 3a (dont quelques minutes sont liées à l'étape 1/2), puis ensuite 27 minutes (séance 3) et 9 minutes (séance 4) pour la sous-étape 3b. Enfin 15 minutes, lors la séance 4, sont à mettre en lien avec l'étape 4. Il n'est pas étonnant de remarquer le peu de recherche individuelle lors de la sous-étape 3b, vu que la démarche est quant à elle différente de celle rencontrée lors du premier problème. Cela s'explique notamment par le fait qu'il y a besoin ici du logiciel *GeoGebra*, qui n'est accessible que par l'enseignante. Nous retrouvons à nouveau la manipulation du logiciel prise en charge par l'enseignante, la plupart du temps à l'initiative des élèves.

Le scénario choisi par la classe est bien inclus dans le canevas proposé en analyse *a priori*. Les fonctions du type $\frac{1}{x}$ ont d'abord été proposées pour, dans un second temps, laisser la place aux fonctions de type exponentiel. Les fonctions du type $x \mapsto \frac{1}{x^2}$ ne sont pas apparues, ce qui n'est pas étonnant. Nous pouvons remarquer que si elles étaient apparues, cela aurait alourdi considérablement le déroulement. Nous tenons à faire remarquer que le travail de mathématisation a débuté très rapidement mais que pour autant la variable aléatoire X représentant le temps d'attente entre deux éruptions n'est introduite explicitement qu'au cours de la séance 4. Les étapes se sont enclenchées sans que soit posé la variable aléatoire considérée, cependant au vu du déroulement, on peut supposer que cela est implicite.

Au niveau des « imprévus », le principal est dû à une mauvaise communication avec l'enseignante en amont de la séance. Au lieu d'utiliser la commande Inspecteur de fonction qui donne l'aire sous la courbe explicitement, l'enseignante propose aux élèves (dans le mail accompagnant leur travail à faire à la maison) d'utiliser la

commande Integrale, ce qui n'est pas une commande « naturelle » sachant que l'intégrale n'a pas été introduite en amont.

Nous pouvons remarquer plusieurs passages qui ont été sources de discussion. Nous pouvons citer la discussion sur l'ensemble de définition de la fonction de densité : le fait de considérer un ensemble non borné n'était pas naturel pour tous les élèves. Un autre blocage est survenu dans la classe : la traduction de la question b en termes probabilistes. Nous ne nous attardons pas sur ce point, qui nous semble-t-il n'est pas au cœur de notre questionnement.

À nouveau dans ce problème, l'enseignante prend finalement peu en charge les activités. Nous pouvons mentionner le moment où elle prend en main l'avancée du travail pour l'interprétation des données statistiques (lignes 151-192). Nous tenons à mettre en évidence l'aide constructive que l'enseignante apporte, à ce moment, pour que les élèves distinguent bien la différence entre diagramme en bâtons se rapportant à des données discrètes et histogramme se rapportant à des données continues. Nous étudierons plus en détail les interventions de l'enseignante dans les étapes 3 et 4 dans l'analyse locale.

Pour terminer sur l'analyse globale du problème, nous souhaitons présenter les différentes sous-tâches qui apparaissent au fur et à mesure des mises en commun et qui vont amener à la résolution de la tâche prescrite. Ces sous-tâches sont présentées en figure 7.17.

FIGURE 7.17 – Scénario choisi par la classe en termes de sous-tâches

À partir de notre analyse *a posteriori*, nous pouvons détailler cet organigramme en ajoutant les éléments de déroulement qui ont nourris ces différentes sous-tâches. Nous avons notamment identifié les étapes et même parfois des phases à l'intérieur de ces étapes.

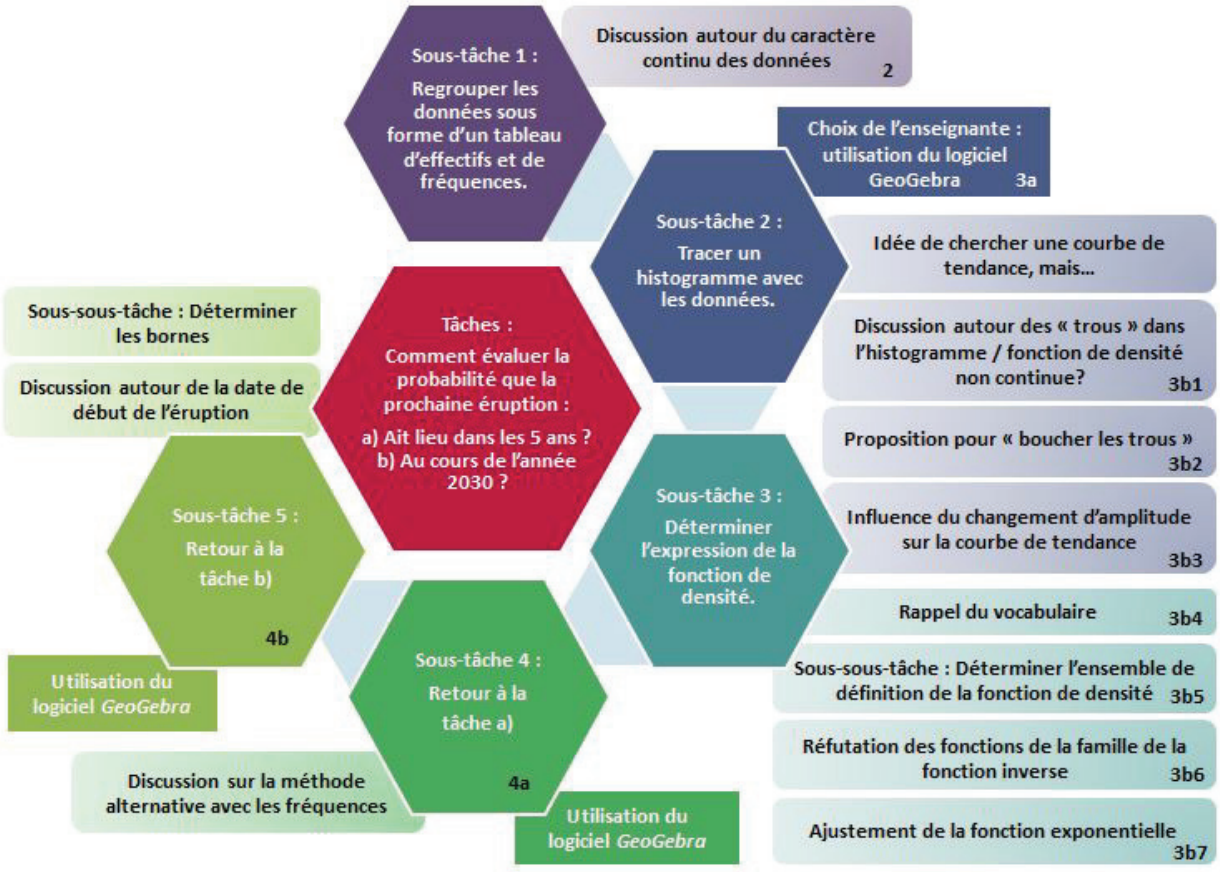


Figure 7.18 – Organigramme du déroulement du problème du volcan Aso (séances 3 et 4)

c) Analyse locale du déroulement de la mathématisation (3) et du traitement mathématique (4)

À la fin de l'étape 3a, une fois l'histogramme d'amplitude 5 ans tracé sur *GeoGebra* (séance 3, ligne 250), plusieurs phases s'enchaînent, comme le montre la figure 7.18. Nous n'allons pas revenir sur l'ensemble de ces phases, mais plutôt nous focaliser sur celles qui font ressortir les liens entre probabilités et analyse, autour de la notion de fonction de densité, à savoir les phases 3b6 et 3b7. Elles correspondent à la réfutation des fonctions de la famille $x \mapsto \frac{1}{x}$, pour la phase 3b6, et à l'ajustement de la fonction exponentielle, pour la phase 3b7. Au niveau du traitement mathématique, nous nous limiterons à la phase 4a, donc à la résolution de la première question. Ce choix s'explique car, d'une part, la seconde phase reprend le même raisonnement que la première au niveau du calcul d'aire et, d'autre part, elle présente une difficulté externe au calcul d'aire, que nous avons présenté dans le paragraphe précédent, qui n'a pas de réel intérêt dans notre analyse.

La recherche de la fonction de densité dans ce problème est beaucoup plus difficile que pour le problème de la rencontre. Les phases pour y arriver sont donc très longues et peuvent ensuite être à nouveau divisées en sous-phases. Ce sera le cas

dans le découpage des phases 3b6 et 3b7.

Analyse détaillée des phases 3b6, 3b7 et 4a

Après avoir déterminé l'ensemble de définition de la fonction de densité (phase 3b5), la classe se met d'accord et reformule la sous-tâche 3 par :

> On cherche une fonction f définie sur $[0;+\infty[$ qui approche « au mieux » l'histogramme.

L'histogramme en question est celui distribué sur polycopié et se trouve en figure 7.19.

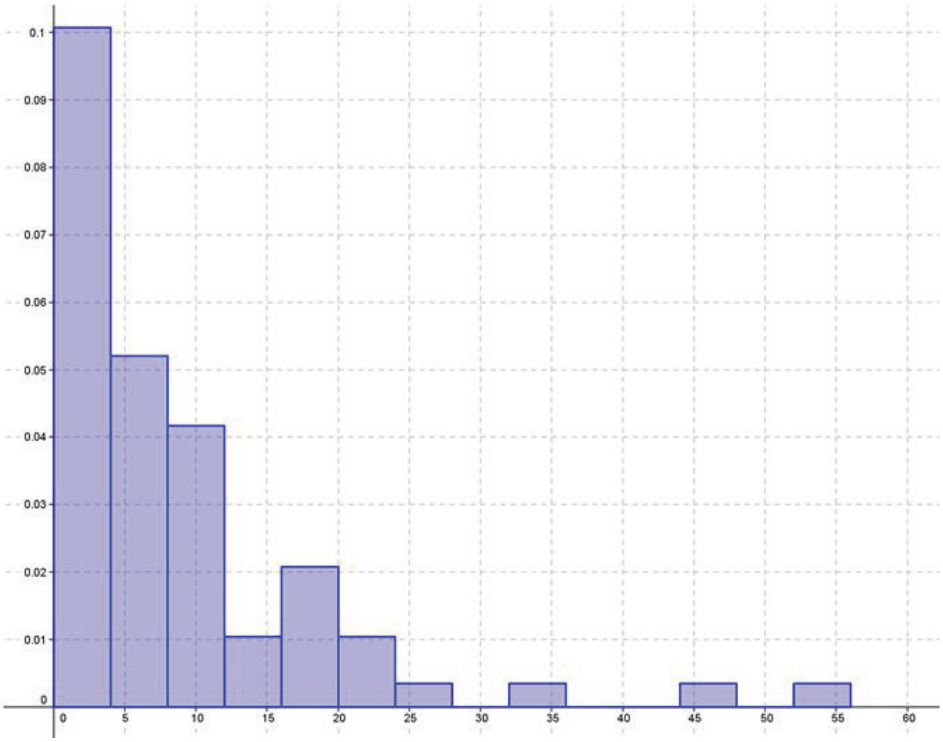


FIGURE 7.19 – Histogramme sur lequel repose la recherche d'une fonction de densité

Derrière la fonction en question dans l'énoncé, se cache donc la fonction de densité de la variable aléatoire du temps d'attente entre deux éruptions du volcan Aso. Une fois la question posée, une phase de recherche individuelle d'environ 4 minutes est laissée aux élèves. Certains tracent une courbe sur le graphique (cf. figure 7.20). L'enseignante demande explicitement une expression de la fonction. Beaucoup d'élèves proposent la fonction $x \mapsto \frac{1}{x}$. Nous pouvons voir en figure 7.21 qu'un élève a déjà éliminé cette fonction candidate.

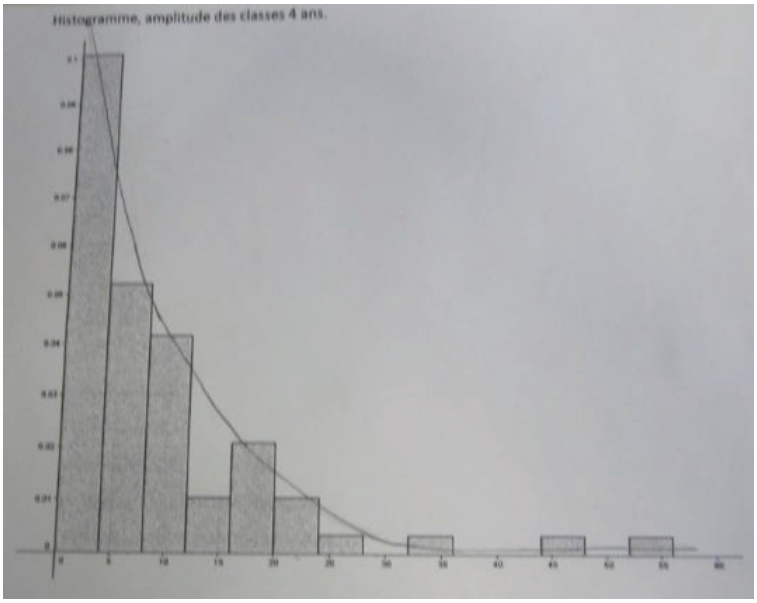


Figure 7.20 – Tracé de courbe d'un élève

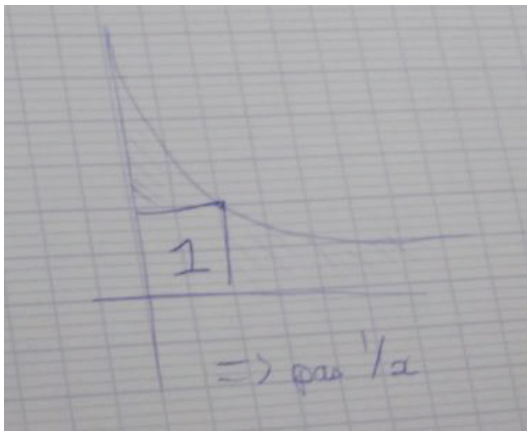


Figure 7.21 – Proposition d'un élève pour éliminer la fonction candidate $x \mapsto \frac{1}{x}$ lors de la phase de recherche individuelle

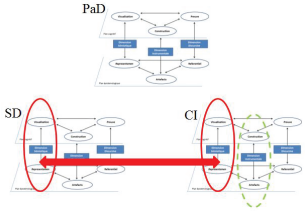
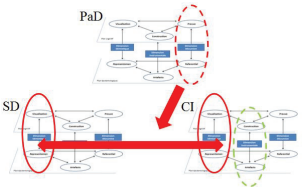
Nous allons maintenant étudier spécifiquement les phases qui suivent cette recherche individuelle.

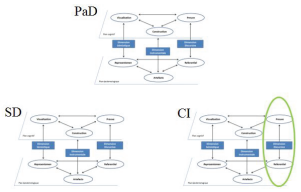
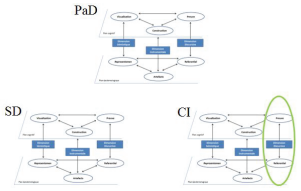
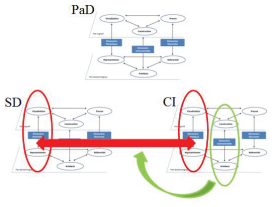
Nous reprenons la même grille d'analyse que celle utilisée pour le problème de la rencontre (cf. p. 308). Pour rappel, dans chaque phase, nous essaierons de mettre en évidence les points forts et la répartition des interventions entre les élèves et l'enseignante. Comme nous l'avons précisé dans la section 7.1.6, du côté des élèves, nous repérerons les propositions faites et les activités. Du côté de l'enseignante, nous identifierons les types d'intervention qu'elle fait (reformulation, demande de précisions...) ainsi que les aides qu'elle apporte (souligné par une vague). Nous identifierons les changements de registre et la disponibilité des connaissances en jeu (en gras) ainsi que les types de (sous)-activités (gras italique). Nous indiquerons aussi les références

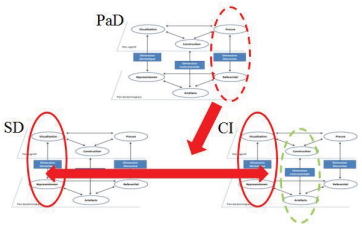
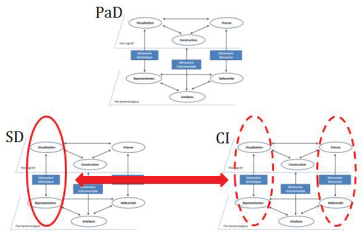
à l'ancien, notamment au DM6 (en rouge). Enfin, pour décrire chaque phase, nous proposerons un schéma des trois ETM en jeu montrant les circulations repérées. Les dimensions sollicitées par les élèves sont entourées en rouge, celles par l'enseignante sont entourées en vert. Nous entourerons en rouge pointillé quand l'intervention des élèves a été sollicitée par une aide de l'enseignante et enfin en vert pointillé lorsque qu'une manipulation par exemple de l'enseignante a été préalablement commandée par un élève.

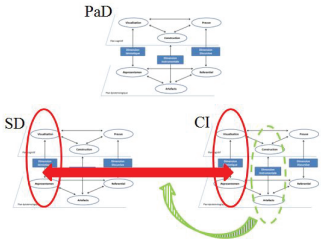
La légende est la suivante :
italique : extraits de transcription
vague : aides de l'enseignante
gras : changement de registres, disponibilité des connaissances
Gras italique : sous-activités
rouge : référence à l'ancien

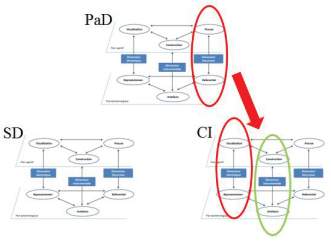
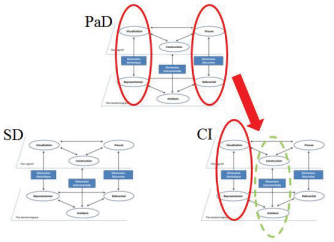
Analyse du déroulement des phases 3b6, 3b7 et 4a
Séance 3 (lignes 479-852) et séance 4 (lignes 12-225)

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
Séance 3 (transcription de l'annexe I.3)			
Phase 3b6 - Réfutation des fonctions de la famille $x \mapsto \frac{1}{x}$			
Sous-phase 3b6.1 - Recherche de correspondance entre l'histogramme et la courbe			
39 : 30 l.479		Manipulation du logiciel *GeoGebra* (inspirée du travail individuel des élèves lors de la phase de recherche précédente) : Tracé de la fonction $x \mapsto \frac{1}{x}$ dans la même fenêtre que l'histogramme.	
	Proposition : Décaler la courbe vers le bas.		
		Manipulation de *GeoGebra* : Déplacement de la courbe vers le bas à la demande des élèves. ***Activité de traitement prise en charge par l'enseignante.*** **Registre histogramme sur GeoGebra.**	Recherche de correspondance entre la dimension sémiotique de SD (histogramme) et celle du CI (courbe) à travers un travail de visualisation, alimenté par la dimension instrumentale (CI) à la charge de l'enseignante.
Sous-phase 3b6.2 - Prise en compte de la contrainte de positivité			
40 : 28 l.505		Aide procédurale indirecte : « *Alors oui mais est-ce que je peux la décaler... ? Regardez bien ce qui se passe partout parce que...* ».	
	Proposition : Non. Justification : Parce qu'elle est négative.		
		Demande de reformulation : Et on cherche des fonctions...	
	Positives.	Manipulation de *GeoGebra* : Déplacement de la courbe vers le haut.	
	Propositions de déplacements de la courbe, que doit ensuite prendre en charge l'enseignante.		
		Manipulation de *GeoGebra* : Déplacement de la courbe à la demande des élèves.	Le référentiel théorique lié à la fonction de densité (positivité de la fonction) est convoqué par les élèves et relance le travail de visualisation. La dimension instrumentale reste à la charge de l'enseignante, mais toujours sur initiative des élèves.
		Arrêt sur une courbe. Question : Est-ce une bonne candidate ?	
	Avis mitigés.		
		Manipulation de *GeoGebra* : Déplacement de la courbe à la demande des élèves.	
	Beaucoup d'élèves satisfaits.		
	Proposition : La courbe va peut-être couper l'axe des abscisses.		

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
		Manipulation de *GeoGebra* : Déplacement vers des abscisses plus grandes.	
	Rejet de la fonction considérée. ***Activité de traitement***		
Sous-phase 3b6.3 - Passage au registre algébrique			
42 : 53 l.575	Tentatives.	Reprise : On voudrait une fonction qui ne coupe pas l'axe des abscisses. Aide procédurale indirecte : *« C'est quoi la forme générique, là ? »* Aide procédurale directe : $f(x) = d + \frac{a}{bx+c}$ (donne la réponse). **Changement de registre : Passage du registre graphique au registre algébrique des fonctions.** ***Activité de traitement prise en charge par l'enseignante***	 Le changement de registre est pris en charge par l'enseignante, qui sollicite le référentiel théorique relatif aux fonctions inverse.
Sous-phase 3b6.4 - Travail sur les limites			
43 : 33 l.593	Propositions : $+\infty$, 0.	Aide procédurale indirecte : Quelle doit être la limite de la fonction en $+\infty$? Aide procédurale directe : 0. Aide procédurale directe : Que doit-on alors imposer à la fonction ? Prend en charge toutes les étapes du raisonnement et donne les réponses. $f(x) = \frac{a}{bx+c}$ ***Activité de traitement prise en charge par l'enseignante***	 La dimension discursive liées aux limites de fonctions (domaine de l'analyse) est entièrement prise en charge par l'enseignante.
Sous-phase 3b6.5 - Nouvelle recherche de correspondance entre l'histogramme et la courbe			
44 : 24 l.614	Propositions : Déplacer les curseurs pour modifier la courbe. Justifications (pour valider ou invalider) : Compensation (rectangles au-dessus de la courbe à certains endroits et vide à d'autres). Non ressemblance avec l'histogramme. ***Activité de traitement***	Aide procédurale directe : Utilisation des curseurs de *GeoGebra*. Manipulation de *GeoGebra* : Introduction des curseurs a, b et c.	 La dimension instrumentale (analyse) est mobilisé par le biais des curseurs, à la charge de l'enseignante. Cela relance le travail de visualisation, à la charge des élèves.

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
Sous-phase 3b6.6 - Prise en compte de la contrainte de l'aire			
46 : 28 l.663		Aide procédurale indirecte : « *Rappelez-moi, il y avait une autre contrainte sur la courbe qu'on cherche quand même. C'est quoi ?* »	
	Proposition : L'aire est 1.	Reprise.	
	Remarque : Donc ça ne marche pas si on met $+\infty$.	Demande d'explication.	
	« *Parce que c'est une forme qui n'est pas finie* ». Un autre élève : C'est un DM. Réf. au DM5 (**connaissance « ancienne »**).		
		Reformulation : On peut avoir un domaine infini et une aire finie.	
		Relance : manipulation de *GeoGebra*. Utilisation de la commande Integrale pour donner l'aire sous la courbe. Réponse : 5,74.	
	Proposition : Déplacer la courbe vers le bas.		
		Manipulation de *GeoGebra* : Déplacement de la courbe à la demande des élèves. ***Activité de traitement***	Mobilisation du référentiel théorique lié à la fonction de densité (aire sous la courbe égale à 1). La dimension instrumentale permet d'indiquer l'aire ce qui relance et guide à nouveau un travail de visualisation, qui permet finalement de rejeter ces fonctions.
	Proposition : Ne correspond plus à l'histogramme.		
		Aide procédurale indirecte : L'aire sera toujours...	
	Propositions : Trop grande.		
		Bilan : Rejet de cette famille de fonction, à cause d'un problème d'aire sous la courbe.	
Phase 3b7 - Ajustement de la fonction exponentielle			
Sous-phase 3b7.1 - Une nouvelle proposition			
48 : 52 l.721		Relance : Est-ce qu'il y a d'autres propositions ?	
	Proposition : $f(x) = \frac{1}{e^x} = e^{-x}$ Mais elle passe de l'autre côté des ordonnées.		
		Aide procédurale directe : On peut se restreindre à l'intervalle $[0; +\infty[$. [Problème informatique] Question : Quelle est l'allure de la courbe ? **Changement de registre**	
	Proposition : Tracé de la courbe de $f(x) = e^{-x}$ dans l'air.		Le référentiel théorique sur la fonction exponentielle permet de déterminer quelques conditions sur la fonction de densité, notamment en recherchant une ressemblance entre l'histogramme et la courbe.
		Question : Donner un point particulier de cette courbe.	
	Proposition : $f(0) = 1$		
		Aide procédurale directe : Rejet de cette fonction. Trace la courbe au tableau.	

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
		Aide procédurale directe : On veut que $f(0) = 0,1$. ***Activité stratégique (implicite) prise en charge par l'enseignante*** Question : « *Alors qu'est-ce que vous proposez ?* » Relance.	
	Proposition : $f(x) = e^{-x} - 0,9$. Problème : La courbe va descendre (implicitement : la fonction devient négative). Proposition : $f(x) = \frac{e^{-x}}{10}$.		
		Reformulation : $f(x) = 0,1 \times e^{-x}$.	
Sous-phase 3b7.2 - Prise en compte de la contrainte de lissage (interrompue)			
l.818			
		Proposition : Tracer la courbe. Manipulation de *GeoGebra* : saisie de la fonction.	
	Elle tend vers 0.		
		Reformulation : « Elle chute tragiquement ». Question : Sur quoi peut-on jouer ?	
Fin du cours (55 : 00)			
		Tâche à faire à la maison : À l'aide du fichier *GeoGebra* de l'histogramme, essayer d'affiner la fonction, avec $f(0) = 0,1$. Puis répondre aux questions de départ (cf. mail envoyé à la fin de l'annexe I.3).	
		Réexplication.	
Travail à la maison			
Séance 4 (transcription de l'annexe I.4)			
Reprise de la phase 3b7 - Ajustement de la fonction exponentielle			
Reprise rapide de la phase 3b7.1			
Sous-phase 3b7.3 - Prise en compte de la contrainte de lissage de l'histogramme			
l.43	Proposition : Prendre deux curseurs. Explication des curseurs.		
		Manipulation de *GeoGebra* : introduit les curseurs, comme indiqué. Reformulation : $f(x) = be^{-kx}$, avec b et k les curseurs.	
	Proposition : Tester des valeurs pour les curseurs.		
		Aide procédurale indirecte : Que vaut $f(0)$?	
	Proposition : $b = 0,1$.		
			La dimension instrumentale (analyse), à l'aide des curseurs (proposés par un élève), relance le travail de visualisation, à la charge des élèves.

Temps, Ligne	Interventions et activités élèves	Interventions et activités enseignante	Circulations entre les ETM
		Remarque : Curseur b non utile. Manipulation de *GeoGebra* : Place le curseur de b à 0,1 et fait varier le curseur k. ***Activité de traitement***	
	Validation de $k = 0,1$.		
Sous-phase 3b7.4 - Validation par la contrainte de l'aire			
l.120	Proposition : Il faut avoir l'aire sous la courbe.	Demande de justification. Reprise. Reformulation. Manipulation de *GeoGebra* : utilisation de la commande Integrale dans la barre de saisie pour trouver l'aire sous la courbe. Réponse : Ça fait 1. Mise en garde : Ce n'est pas une preuve que l'aire vaut bien 1. ***Activité de traitement*** Validation de la fonction $f(x) = 0,1e^{-0,1x}$, en précisant qu'ils démontreront dans le cours que l'aire vaut bien 1.	 Mobilisation du référentiel théorique de la fonction de densité (aire égale à 1) pour valider la fonction choisie. Cette vérification n'est possible que par la dimension instrumentale, prise en charge par l'enseignante.
Phase 4 - Traitement mathématique			
Phase 4a - Recherche de la probabilité			
	Proposition : On calcule l'aire sur $[0; 5]$. Proposition : Avec Integrale de 0 à 5. Réponse : Non.	Reformulation de la tâche : Le volcan est en éruption. Évaluer la probabilité que la prochaine éruption ait lieu au cours des 5 ans à venir. Demande de reformulation dans le registre probabiliste. Écrit $P(X \in [0; 5])) =$ Demande de précisions : Avec le logiciel ? Manipulation de *GeoGebra* : utilisation de la commande Integrale pour obtenir la valeur de l'aire. Réponse : 0,39. ***Activité de traitement*** Demande validation.	 Le lien est clairement identifié entre probabilité et aire sous la courbe par les élèves. La dimension instrumentale est ici nécessaire pour calculer cette aire.

Dans cette analyse, nous nous sommes focalisée sur les phases de recherche de la fonction de densité (phases 3b6 et 3b7) et sur la détermination de la probabilité une fois cette fonction choisie. Dans ces trois sous-phases, nous pouvons identifier clairement, comme dans le problème de la rencontre, un jeu entre les trois sous-domaines (SD, PaD, CI), bien visible dans le « comics » en figure 7.22.

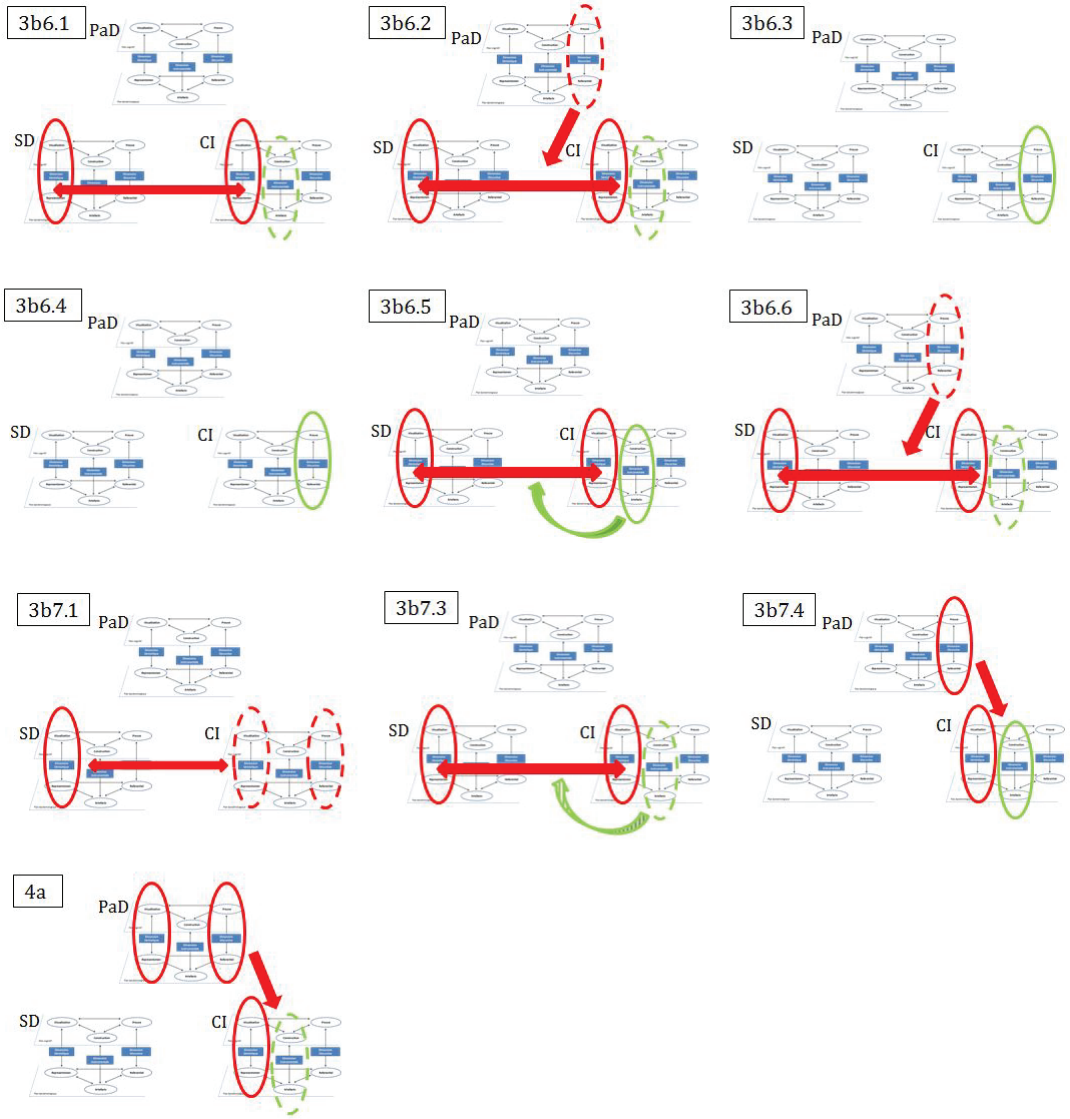


FIGURE 7.22 – Dynamiques entre les ETM dans les phases 3b6, 3b7 et 4a

De véritables dynamiques s'opèrent entre les différentes dimensions des trois ETM. La dimension instrumentale est ici beaucoup plus présente que dans le problème de la rencontre. Elle n'est pas seulement mobilisée dans le domaine de la statistique, pour la construction des histogrammes, mais aussi dans le domaine de l'analyse, pour la recherche d'une courbe adéquate (courbe de densité). Ceci est dû au fait que la fonction de densité en jeu ne peut pas se trouver à la main, du fait qu'il s'agit d'un calcul d'aire de niveau A3 (non calculable par les élèves, pour

le moment). L'activité de traitement se fait essentiellement dans cette dimension instrumentale (relative à CI), prise en charge par l'enseignante. Cependant, nous pouvons remarquer qu'elle est la plupart du temps sollicitée sur demande des élèves (en vert pointillé). Nous rappelons que l'objectif de la séance n'est pas la manipulation du logiciel. Le logiciel est ici un outil efficace pour permettre les tests successifs des fonctions. Cependant, cette activité n'est déclenchée que par la mobilisation des autres dimensions. C'est avant tout une activité de visualisation (dimension sémiotique SD/CI) qui déclenche une activité sur le logiciel (dimension instrumentale CI). Après les différents blocages, c'est le référentiel théorique (dimension discursive PaD), avec l'apparition au fur et à mesure des contraintes que doit vérifier la fonction de densité, qui est mobilisé pour réalimenter la visualisation et le travail informatique. Dans ce problème, plus que dans le précédent, on s'aperçoit de la véritable utilité de chaque propriété de la fonction de densité, notamment la positivité (et même la continuité, qui lance un débat dans la phase 3b1). Nous pouvons repérer une sorte de cycle qui permet au final d'aboutir à une fonction adéquate.

Nous tenons à préciser que cette analyse de séances de classe nous permet seulement de voir la dimension instrumentale travaillée par l'enseignante. Cependant nous rappelons qu'entre la sous-phase 3b7.2 et 3b7.3, un travail à la maison a été donné : les élèves devaient ajuster la fonction exponentielle $x \mapsto 0,1e^{-x}$ pour qu'elle puisse être considérée comme une courbe de densité pour les données dont ils disposent. Cette tâche devait explicitement être résolue à l'aide de *GeoGebra*. Cette activité de l'élève est inaccessible pour nous, mais il faut garder à l'esprit qu'elle a dû avoir lieu et donc qu'il ne faut pas penser que les élèves sont complètement exclus des activités dans la dimension instrumentale (CI).

Contrairement au problème de la rencontre, l'objectif n'est plus de faire émerger le référentiel théorique sur les fonctions de densité (PaD), celui-ci ayant été amené lors des deux séances précédentes. Le but ici est de consolider ce référentiel pour se l'approprier. Cela s'explique notamment par le fait qu'il n'y a plus de flèches arrivant vers la dimension discursive des probabilités. Nous aurions pu penser qu'ayant été introduites à la séance précédente, les propriétés relatives aux fonctions de densité étaient disponibles pour les élèves, cependant il faut être nuancé sur ce point. Nous pouvons remarquer qu'en grande partie, le référentiel théorique n'est mobilisé par les élèves qu'après une intervention de l'enseignante (en rouge pointillé). Il s'agit toujours d'aides procédurales indirectes. Nous donnons ici deux exemples. Pour que les élèves mobilisent la contrainte de la positivité de la fonction, l'enseignante pose la question suivante :

> [40 : 28] Marie : Alors oui mais est-ce que je peux la décaler... ? Regardez bien ce qui se passe partout parce que... (lignes 505-506)

Ensuite, pour que les élèves mobilisent la contrainte de l'aire sous la courbe égale à 1, elle demande :

> [46 : 28] Marie : Rappelez-moi, il y avait une autre contrainte sur la courbe qu'on cherche quand même. C'est quoi ? (lignes 663-664)

À travers ces interventions, nous nous rendons compte que l'enseignante se sent obligée de guider les élèves pour qu'ils mobilisent le référentiel théorique en construction. Nous tenons tout de même à préciser que, la plupart du temps, plusieurs élèves ont la réponse avant que l'enseignante n'intervienne. Cependant, l'enseignante ne les

interroge pas tout de suite et propose alors des aides procédurales indirectes pour que plus d'élèves soient en mesure de faire avancer le travail mathématique. C'est ici une stratégie de l'enseignante pour qu'un maximum d'élèves soit impliqué dans l'avancement de la résolution. On peut faire l'hypothèse qu'étant donné que ce n'est que la deuxième rencontre des élèves avec ce type de tâche (recherche d'une fonction de densité) et que la première institutionnalisation sur la fonction de densité et ses caractéristiques n'a eu lieu qu'au cours précédent, ces connaissances ne sont pas encore disponibles pour les élèves mais seulement mobilisables lorsque l'enseignante lance les élèves dans cette direction. Ce type d'activité des élèves n'est pas encore routinisé. L'enseignante les aiguille sans donner la réponse. Cela insiste sur le fait qu'un réinvestissement pour consolidation comme ici semble indispensable pour asseoir les propriétés.

Pour revenir sur le logiciel *GeoGebra*, celui-ci avait été choisi pour sa facilité d'utilisation (pour l'enseignant) mais aussi pour sa transparence vis-à-vis de l'élève. Même sans l'utiliser lui-même, il est simple, en regardant, de comprendre ce que fait le logiciel. Le fait que l'enseignante ait décidé d'utiliser la commande Integrale (qui donne la valeur de l'intégrale sur l'intervalle considéré) plutôt que Inspecteur de fonction (qui possède un item Aire) pour le calcul d'aires sous la courbe pose un problème vis-à-vis de cette « transparence » vis à vis des élèves. L'intégrale n'étant pas un objet mathématique connu des élèves, les élèves doivent croire sur parole l'enseignante que la valeur donnée par l'ordinateur est bien l'aire sous la courbe. Aucun malaise ne s'est fait ressentir dans la classe sur ce point, mais il a semblé important de le pointer.

Dans la phase 3b6, nous pouvons observer deux sous-phases entièrement à la charge de l'enseignante. Il s'agit des phases 3b6.3 et 3b6.4, au cours desquelles il y a un passage du registre graphique au registre algébrique (des fonctions), pour ensuite faire un calcul de limite. Marie prend ici beaucoup plus la main. Dans ces deux sous-phases, il s'agit d'un travail dans la dimension discursive du sous-domaine CI, et même plus largement dans le domaine de l'analyse : détermination de l'expression algébrique d'une fonction et travail sur la limite de la fonction trouvée. Après avoir sollicité les élèves, sans réponse convaincante, elle prend assez rapidement ce travail à sa charge, comme par exemple ici pour la sous-phase 3b6.3 :

> [42 : 53] Marie : C'est quoi la forme générique, là ? La forme canonique pour une hyperbole ?
> Une élève : Ah...
> Marie : Ah. Vous vous souvenez pas ? C'est quoi ? Vous savez pour les suites quand vous voulez... étudier le... B31 ?
> Élève B31 : C'est α...
> Marie : α. Oui.
> Élève B31 : $\alpha + \beta$ sur...
> Marie : Oui alors α et β, on va pas pouvoir faire... On va mettre des lettres... Alors α sur ? Tu te souviens pas ?
> Élève B31 : $x+$ ...
> Marie : On va mettre des lettres... $bx + c$ et d. Cette forme-là.
> *L'enseignante écrit au tableau en même temps :*
>
> $$f(x) = d + \frac{a}{bx + c}$$ *(lignes 575-588)*

Nous pouvons penser que l'enseignante fait ce choix (donner des aides procédurales directes et prendre à sa charge les activités) pour ne pas perdre de temps. Dans ce problème, l'important n'est pas dans ce travail sur les fonctions. L'analyse, ici, n'a pas le même statut que les probabilités, qui sont l'apprentissage visé. Les élèves sont censés savoir résoudre ces tâches, mais vu le temps de réponse, elle préfère ne pas s'attarder dessus. Il en est de même ensuite lorsque l'enseignante prend elle-même l'initiative d'introduire des curseurs dans GeoGebra. Pour ce dernier point, s'ajoute le manque de maîtrise du logiciel par les élèves.

Nous retrouvons cette même prise en main des activités par Marie dans la phase 3b7.1, quand elle décide d'imposer $f(0) = 0,1$. Il s'agissait ici d'une activité stratégique potentielle qui a été court-circuitée par son intervention. Cette intervention s'explique sûrement par le fait qu'on arrivait à la fin de la séance et qu'il était nécessaire d'arrêter des conditions en classe, pour pouvoir donner un travail à faire à la maison.

Dans ce problème, nous pouvons à nouveau repérer des moments de blocages qui sont dépassés par une aide de l'enseignante. Cette fois les aides sont indirectes quand il s'agit d'une blocage au niveau de la notion de fonction de densité, tandis qu'elles sont directes au moment du travail sur la fonction par exemple.

Contrairement au problème de la rencontre, il n'y a pas dans ce second problème de référence explicite aux notions du DM6, ni au premier problème, cependant il est assez raisonnable de penser que des références implicites existent. Nous avons en revanche remarqué une référence au DM5, par un élève, pour l'aire finie d'un domaine infini. Cette référence était attendue et est donc nous semble-t-il un point d'appui indispensable pour le bon déroulement de la séance. Cela permet de ne pas perdre de temps dans ce problème, sans laisser un blocage s'installer.

Ces phases montrent une circulation riche entre les différents ETM, où chacun a son propre rôle. Pratiquement toutes les dimensions de chaque ETM sont sollicitées, même si parfois elles sont prises en charge par l'enseignante. Nous remarquons que la dimension instrumentale PaD n'est pas sollicitée, ce qui est normal, étant dans la phase de construction des différentes lois. Pour la statistique descriptive, nous n'observons pas dans ces phases la mobilisation des dimensions instrumentale et discursive, qui ont été toutefois sollicitées mais en amont dans les phases 3a et 3b3. La dimension discursive (PaD) intervient plusieurs fois, comme espéré, et permet de relancer une dynamique dans le travail mathématique. Cette dimension est indispensable ici, alors qu'elle était totalement absente dans les manuels. Ces phases montrent un travail que l'on pourrait qualifier de quasi complet (voir complet si l'on prend en compte les phases en amont). Ce travail est guidé par l'ETM collectif de la classe et va permettre la construction d'un ETM idoine effectif nous semble-t-il assez solide. Cette construction n'aurait cependant pas été possible sans les interventions de l'enseignante, soit pour cadrer le débat soit cette fois-ci aussi pour faire gagner du temps dans les moments non cruciaux dans l'apprentissage visé, mais toujours en gardant le souci de l'adhésion du maximum d'élèves.

7.3.3 Analyse de la deuxième institutionnalisation

a) Une deuxième institutionnalisation : la généralisation

Nous rappelons que, lors de la séance 2, a eu lieu une première institutionnalisation qui a permis l'introduction du vocabulaire relatif notamment aux fonctions de densité. Cependant, l'institutionnalisation restait contextualisée au problème de la rencontre, tout du moins dans l'écrit. À la fin du problème du volcan Aso, Marie propose une nouvelle institutionnalisation qui, cette fois-ci, va être décontextualisée. Cette phase est identifiée de la ligne 426 à la ligne 497 de la transcription de la séance 4 (cf. Annexe I.4). Il s'agit à nouveau d'une synthèse, l'enseignante parle d'un résumé, et non du cours à proprement parlé. L'écrit se restreint au contenu de la figure 7.23.

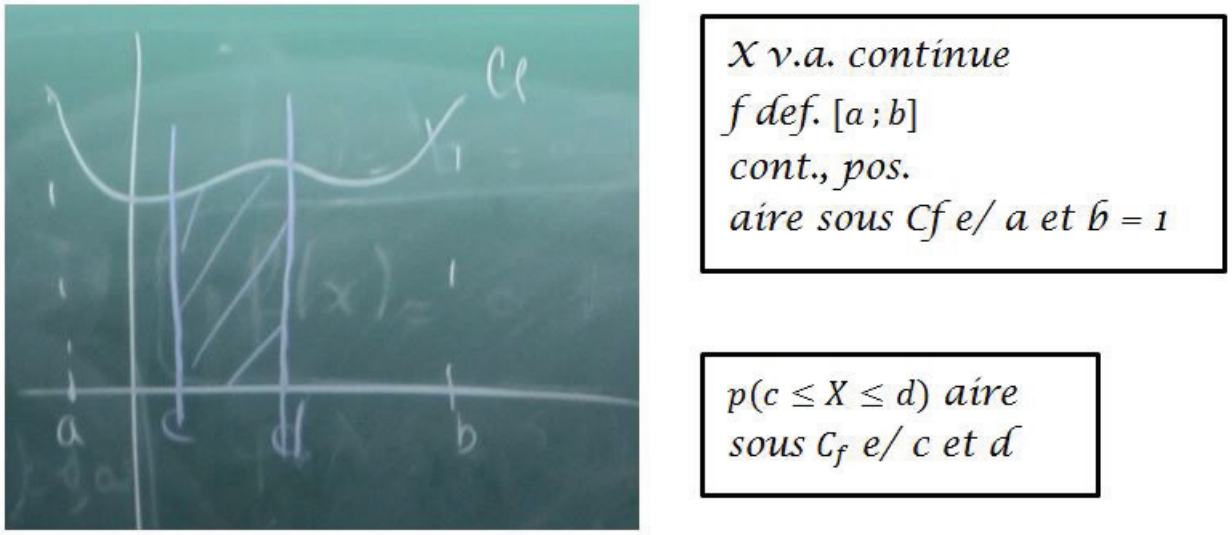


FIGURE 7.23 – Synthèse de la séance 4

Comme pour la première institutionnalisation, nous présentons sous forme de tableau le déroulement de l'institutionnalisation en mettant en avant le vocabulaire introduit, ainsi que les éventuelles définitions et propriétés (en gras) et les justifications apportées (en gras italique). Nous mettons aussi en avant qui, des élèves ou de l'enseignante, amène ces nouvelles connaissances. Les citations de la transcription sont écrites en italique. Nous faisons des commentaires pour préciser le degré de généralité employé à l'oral et à l'écrit (parfois souligné dans les citations) ainsi le niveau de rigueur utilisé.

Analyse de la synthèse de la séance 4

Interventions de l'enseignante	Interventions des élèves	Commentaires
Lignes 426-431 **Vocabulaire :** **Variable aléatoire continue** *Ça veut dire quoi ?*		Le vocabulaire, variable aléatoire continue, est réintroduit. Une définition est demandée aux élèves (seulement du terme continue, implicitement). La réponse erronée de l'élève est corrigée par l'enseignante mais aucune explication de son erreur n'est donnée. Degré de généralité élevé à l'oral : *d'un intervalle donné.*
	Qui prend toutes les valeurs. *Qui prend une infinité de valeurs entre...*	
Définition : Invalide : *Prend toutes les valeurs d'un intervalle donné.*		
Lignes 431-438 Démarche : On trouve la courbe. Dans un premier temps, l'enseignante demande ce que la fonction permet de calculer, puis se reprend pour demander comment s'appelle cette fonction.		L'enseignante commence par demander l'utilité de la fonction puis se reprend et demande le nom qu'on lui donne. Nous remarquons que le vocabulaire intermédiaire « courbe de tendance » reste très prégnant chez les élèves, mais le vocabulaire « officiel » est tout de même connu.
	Vocabulaire : Courbe de tendance, **Fonction de densité de probabilité**	
Lignes 439-454 Considère la fonction f sur l'intervalle $[a; b]$. Que doit avoir cette fonction comme contraintes ?		Décalage au départ entre le degré de généralité de l'enseignante et celui de l'élève, qui reste dans le cas du problème du volcan. L'enseignante explicite que l'on se situe dans le cas général. La continuité prend maintenant toute sa place dans les propriétés.
	$f(0) = 0,1$	
Non, dans le cas général.		

Interventions de l'enseignante	Interventions des élèves	Commentaires
	Propriété : *Strictement positive* (pas pris en compte).	
Propriété : *Continue* (a été proposé par des élèves en amont).		
	Propriétés : *Positive.* *Sur l'intervalle* $[a;b]$*, l'aire vaut 1.*	
Vocabulaire : **L'aire sous** $\mathcal{C}_f$ **entre** a **et** b**.**		
Lignes 456-469 Comment fait-on pour calculer la probabilité que X soit compris entre c et d ?		Il s'agit d'expliciter comment, à partir de la fonction de densité, calculer des probabilités. L'enseignante s'appuie sur le graphique d'une fonction. Ceci dans le but de faire émerger le problème de l'aire sous une courbe de niveau A3.
	Propriété : On calcule l'aire. Placer c et d et calculer l'aire entre c et d.	

À cours de cette institutionnalisation, l'enseignante amène la classe à réintroduire le vocabulaire « variable aléatoire continue », « fonction de densité de probabilité » et « aire sous la courbe ». Elle laisse d'ailleurs les élèves redonner eux-mêmes le vocabulaire de fonction de densité de probabilité. Les propriétés de la fonction de densité sont ensuite retravaillées. Ce sont les élèves eux-même qui rappellent les différentes contraintes qui pèsent sur cette fonction, toujours dans un jeu de questions/réponses avec l'enseignante. Cette fois-ci, l'enseignante n'a plus besoin de donner des aides pour que les propriétés ressortent.

Concernant le niveau de généralité utilisé, nous remarquons qu'il y a ici décontextualisation, à un point près : la variable aléatoire choisie est tout de même à valeurs dans un intervalle fermé borné $[a; b]$ et non sur un intervalle quelconque. Marie prend appui dans cette institutionnalisation sur un graphique d'une fonction non affine par morceaux (cf. figure 7.24).

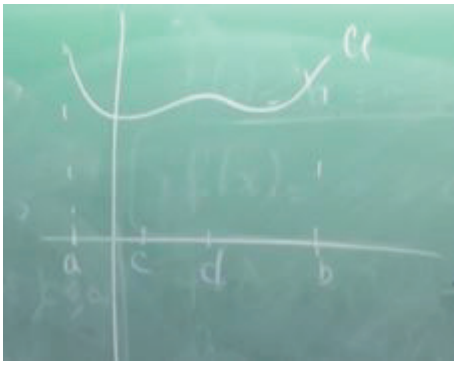


Figure 7.24 – Courbe tracée par Marie lors de la deuxième institutionnalisation

Le graphique est ensuite complété quand elle demande comment calculer la probabilité que la variable aléatoire soit entre c et d et qu'un élève répond directement qu'il s'agit de l'aire sous la courbe entre c et d (cf. figure 7.25).

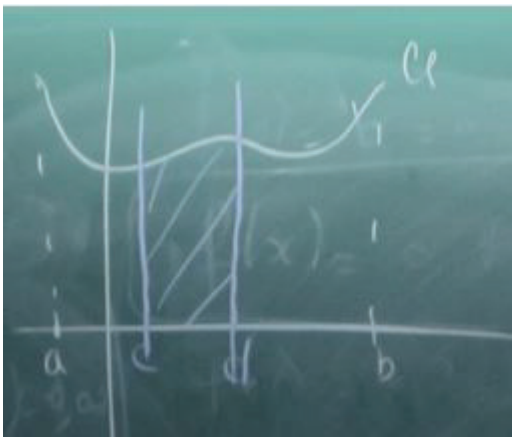


Figure 7.25 – Courbe complétée

Au cours de cette deuxième institutionnalisation, le texte écrit est beaucoup plus

succinct. C'est un moment qui permet de rappeler le vocabulaire introduit à la séance 2, ainsi que les propriétés données, pour le consolider et le généraliser. Contrairement à la première institutionnalisation, la caractère continue de la fonction de densité a maintenant le même statut que les autres propriétés. Il est maintenant écrit dans la synthèse. Nous remarquons qu'aucune justification du vocabulaire ou des propriétés n'est présentée lors de cette institutionnalisation. Enfin, nous voyons ici que cela ne pose plus de problème d'introduire une courbe de niveau A3, car une telle courbe a été rencontrée dans le problème du volcan Aso. Il y a donc une mise en relation avec le problème qui permet une certaine proximité avec les connaissances des élèves. Nous arrivons donc au niveau de généralisation attendu.

Pour conclure, il ne s'agit pas dans cette deuxième institutionnalisation de justifier à nouveau le vocabulaire, mais simplement de le rappeler pour permettre une meilleure mémorisation, et de le généraliser. Nous pouvons voir cet épisode comme une consolidation de la construction du référentiel théorique des probabilités à densité de l'ETM idoine. Pour finaliser l'institutionnalisation, un polycopié de synthèse des problèmes 1 et 2 est distribué aux élèves (cf. Annexe I.6), à compléter à la maison, pour fixer une nouvelle fois le vocabulaire et les propriétés. C'est ce polycopié qui sera inséré au cours écrit de la séquence. Nous redonnerons quelques informations complémentaires dans la section E.1.

b) Vers un nouveau problème

Cette institutionnalisation est aussi le point de départ pour permettre d'introduire un nouveau problème : celui du calcul d'aire de niveau A3. Nous n'allons pas ici proposer une analyse poussée de la tâche qui suit mais seulement montrer que les deux problèmes (de la rencontre et du volcan) permettent effectivement bien de déboucher sur une introduction du calcul intégral.

Dans l'institutionnalisation, Marie prend appui sur un graphique de niveau A3. Ce graphique lui permet alors de lancer une discussion :

> Marie : Alors. Quel est notre... problème ? En tout cas, on a eu à faire face à deux situations très différentes. C11 ?
> Élève C11 : La courbe, elle est un peu arrondie.
> Marie : La courbe, elle est un peu arrondie. Oui. Et ça t'aime pas ? Non. (lignes 471-474)

L'enseignante décide alors de revenir sur l'exemple de la puce, dont la fonction de densité était une constante, puis sur le problème de la rencontre, dont la fonction de densité était une droite oblique, pour conclure que dans ces deux cas, on était capable de calculer les aires car il s'agissait d'aires de trapèzes, de triangles, de rectangles... Il s'agissait effectivement dans ces deux cas de calculs d'aires de niveau A1. Par opposition, elle précise que dans le problème du volcan nous avons affaire à une fonction exponentielle décroissante qui n'est plus « *une figure simple* ». Elle fait ici émerger le problème du calcul d'aires de niveau A3 (les primitives n'étant pas encore connues), pour ensuite lancer la nouvelle tâche qui est la suivante : « *Je vous demande de me proposer un... moyen, un algorithme pour calculer l'aire sous notre courbe. Alors on va se fixer un intervalle, le même pour tout le monde, sinon on va pas s'en sortir. On va faire entre 0 et 20, pour l'instant* » (lignes 499-501).

Ce que nous voulons mettre en avant ici est le fait que le calcul d'aire sous une courbe est clairement motivé dans ce contexte et qu'il est alors naturel pour les élèves de se poser cette question. Nous pouvons d'ailleurs repérer, comme le mentionne Marie (lignes 493-494), que cette interrogation avait déjà émergé auprès d'une élève lors de la séance 3, lors d'une phase de recherche individuelle (lignes 454-471) : « *Mais comment calculer l'aire sous une courbe ?* ».

Nous n'allons pas analyser en détail la recherche d'une méthode pour calculer une valeur approchée de l'aire sous la courbe de la fonction de densité entre 0 et 20. Nous pouvons spécifier que, comme attendu, plusieurs méthodes ont émergé :

– la méthode des trapèzes,
– une méthode à base de tangente,
– la méthode des rectangles (cf. Annexe I.5.2).

Contrairement à ce que nous aurions pu penser, la méthode des rectangles n'a été proposée que par un seul groupe de deux élèves. Nous pensons que le quadrillage large du graphique n'a pas encouragé les élèves à utiliser cette méthode, qui peut sembler trop grossière dans un premier temps.

Pendant la mise en commun qui a suivi, les résultats de chaque méthode ont été présentés. Marie a ensuite poursuivi en insistant sur la méthode des rectangles, seule méthode au programme, qu'elle a illustrée grâce à des animations *GeoGebra* (lignes 866-904).

C'est donc bien ici le problème du volcan Aso qui a fait émerger le besoin de calculer des aires sous une courbe et a amené au calcul approché d'aire (niveau A3). Plusieurs méthodes ont été proposées par les élèves. Ensuite, il y a eu une focalisation sur la méthode des rectangles, méthode qui habituellement se retrouve exclusivement dans le chapitre sur le calcul intégral (cf. chapitre 5).

Cette partie est exclusivement dans le sous-domaine du calcul intégral. La tâche prescrite est ici dans un contexte mathématique. Le problème du volcan est l'élément déclencheur de cette nouvelle tâche.

7.3.4 Conclusion sur le problème du volcan Aso

Cette partie va être l'occasion pour nous de compléter les éléments mis en avant dans la conclusion sur le problème de la rencontre. Nous voulons mettre en évidence les points importants de ces trois niveaux d'analyse (globale, locale et institutionnalisation) qui permettent de donner de nouveaux éléments de réponse à nos six sous-questions relatives à la question de recherche QR3 (cf. 221).

Tout d'abord, nous voulons commencer en disant que le problème du volcan Aso, tout comme le précédent, est effectivement réalisable en classe, avec le type de démarche espéré dans l'analyse *a priori*. Le scénario choisi par la classe s'est effectivement révélé inclus dans le canevas réalisé dans l'analyse globale *a priori*, avec un bon investissement des élèves dans le travail mathématique. Le problème a de nouveau débouché sur une institutionnalisation, qui complète la première.

Nous avons complété notre analyse globale par une analyse locale plus fine sur les passages concernés sur l'introduction de la notion de fonction de densité (étapes 3b6, 3b7 et 4a du cycle de modélisation) et le moment d'institutionnalisation qui a suivi le problème. L'analyse locale permet de voir qu'à nouveau le travail ma-

thématique en jeu lors des passages qui nous intéressent repose sur des circulations riches entre les trois sous-domaines des probabilités à densité, du calcul intégral et de la statistique descriptive (visible dans le « comics » proposé en figure 7.22, p. 314). Cette fois-ci le référentiel théorique des probabilités est consolidé par une utilisation des différentes propriétés de la fonction de densité. Dans ce problème, les trois propriétés sont à utiliser, ce qui n'était pas le cas dans le problème précédent. Dans le passage sur la réfutation des fonctions inverses, de nombreux « blocages » voulus apparaissent successivement ce qui insiste sur la nécessité de ces propriétés. Ce passage par une « mauvaise » fonction est, nous semble-t-il, indispensable pour faire réellement prendre conscience aux élèves de la nécessité des propriétés. Le travail mathématique peut être jugé complet. La dimension discursive de la statistique descriptive n'est plus sollicitée pour faire émerger le référentiel théorique des probabilités à densité, ce qui est logique vu que cela a été fait lors du premier problème.

Le « comics » permet aussi de révéler que l'ETM collectif est en grande partie responsable de la construction de l'ETM idoine effectif et de la consolidation du référentiel théorique relatif aux probabilités, bien que nous puissions voir apparaître plusieurs activités prises en charge par l'enseignante. C'est le cas pour le travail sur les fonctions ou encore à la fin de la séance 3 dans le choix stratégique qu'elle prend à sa charge. Cependant, ce sont seulement pour des activités à des moments où le référentiel théorique en construction n'est pas à convoquer. Comme pour le problème de la rencontre, l'enseignante joue encore un rôle ici dans la construction du référentiel (consolidation), en proposant des aides procédurales indirectes. Nous pouvons conclure que les propriétés de la fonction de densité ne sont donc pas encore disponibles chez tous les élèves, mais les interventions de l'enseignante (aides procédurales indirectes) permettent de débloquer la situation pour beaucoup d'élèves et permet l'avancement du déroulement de la séance pour la classe. Les trois propriétés de la fonction de densité (continuité, positivité et aire égale à 1) sont mobilisées dans ce problème, ce qui permet vraiment de toutes les stabiliser.

Dans ce problème, l'activité centrale laissée aux élèves en autonomie, en individuel ou à deux, est celle au début de la séance 3. La dévolution faite, les élèves se lancent directement dans la recherche et entrent dans le cycle de modélisation. Nous avons pu repérer que les étapes 1/2 du cycle sont mises de côté et que les élèves se lancent directement dans la mathématisation. Tous les élèves sont en activité. Les premières étapes à mettre en œuvre dans ce type de problème de modélisation semblent devenues disponibles (ou au moins mobilisables) pour les élèves. Un second moment est réellement laissé en autonomie aux élèves, dû notamment au fait que cela tombe en fin de séance 3, c'est l'ajustement de la fonction exponentielle pour être fonction de densité. Aucune véritable trace de l'activité des élèves n'est possible, vu qu'il s'agit d'un travail à la maison, mais nous pouvons penser que bien que tous ne semblent pas l'avoir fait, cela a été l'occasion pour les autres de manipuler le logiciel et mettre en jeu à leur charge la dimension instrumentale.

La deuxième institutionnalisation met à nouveau en avant que les élèves sont acteurs de la construction du référentiel théorique, suite à la résolution des premiers problèmes. Après une instutionnalisation contextualisée à la suite du problème de la rencontre, ici il s'agit d'une institutionnalisation décontextualisée, qui est du niveau

de généralité attendu. La statistique descriptive a disparu dans cette institutionnalisation, cela est notamment dû au fait qu'ici aucune justification n'a été redemandée aux élèves (cela ayant été fait lors de la première institutionnalisation). Au vu de cette deuxième institutionnalisation, nous pouvons voir que l'objectif d'apprentissage visé est bien atteint, dans le niveau de généralité attendu.

Cette institutionnalisation est aussi à l'initiative de l'émergence d'un nouveau problème : celui du calcul d'aire sous la courbe d'une fonction non affine par morceaux, complètement motivé par le problème du volcan. Nous arrivons donc bien ici à une problématique de calcul intégral, motivée par les probabilités, qui était un des objectifs de nos problèmes.

Bien que non totalement achevée, nous venons de le voir, la construction du référentiel théorique par le biais de l'ETM collectif est bien enclenchée. Les éléments que nous venons de pointer en sont la preuve.

Pour conclure, certains éléments et certaines conditions se sont révélées importantes pour aboutir à cet objectif. Tout d'abord, nous citerons les passages que nous avons présentés comme incontournables dans nos analyses globales. Il nous semble effectivement que, dans le canevas de l'analyse globale *a priori*, des passages semblent obligatoires et se sont d'ailleurs bien retrouvés dans le déroulement. Il s'agit des points suivants :

- choix sur l'interprétation des données,
- construction d'un histogramme,
- recherche de la fonction de densité (avec une fausse route avec les fonctions inverses).

Ensuite, le devoir maison n° 5 permet de ne pas avoir des discussions qui s'installent dans la classe sur la problématique de l'aire finie d'un domaine infini, ce qui pourrait bloquer l'avancée du travail mathématique. Implicitement, les connaissances en jeu dans le devoir maison n° 6 ainsi que le problème de la rencontre sont aussi des prérequis qui semblent indispensables au bon déroulement.

Comme dans le problème de la rencontre, nous retrouvons des conditions dans la gestion de l'enseignante. Il s'agit à nouveau de laisser la classe, donc l'ETM collectif, construire l'ETM idoine tout en la guidant parfois. Les aides procédurales indirectes se sont à nouveau retrouvées dans les moments où doivent être sollicitées les propriétés de la fonction de densité. Cependant, dans ce problème, nous avons aussi repéré des aides procédurales directes et même des activités prises intégralement en charge par l'enseignante (non liées au logiciel). Nous pensons que ces moments ne sont pas nuisibles à la construction que l'on vise ici car à chaque fois, il s'agit d'activités ne portant pas sur le savoir en construction et au contraire cela peut s'avérer utile pour ne pas bloquer le travail sur des points non cruciaux dans le problème.

Avant de passer à la validation interne, rassemblant les conclusions faites sur les deux problèmes, nous voulons simplement présenter brièvement l'institutionnalisation finale, à savoir le cours qui a été écrit pour cette séquence.

Les dynamiques entre les ETM se sont révélées très riches et nous pouvons identifier un travail complet. Presque toutes les dimensions de chaque ETM sont effectivement sollicitées. La dimension discursive n'est pas mise de côté et bien au contraire, elle permet continuellement de relancer le travail, permettant d'invalider

ou valider une fonction candidate. Nous pouvons noter l'importance du DM5 dans ce problème pour débloquer le fait qu'il soit possible qu'un domaine infini ait une aire finie.

7.3.5 Éléments sur le cours associé à la séquence

Nous ne souhaitons pas présenter en détail et analyser le cours écrit de la séquence, cependant il nous semble tout de même intéressant de s'attarder sur la manière dont se sont imbriqués les probabilités à densité et le calcul intégral dans le cours écrit en classe, ainsi que sur le contenu mathématique arrêté en ce qui concerne les probabilités. Nous proposons de se reporter à l'annexe I.7, où se trouve le cours d'une élève de la classe. L'enseignante a intitulé la séquence « Calcul intégral et lois à densité ». Les problèmes d'introduction ont permis de déboucher sur un problème de calcul d'aires qui permet de lancer le cours sur le calcul intégral.

Il était prévu que le polycopié de synthèse des problèmes 1 et 2, qui se trouve au début de la partie V du cours, soit donné en introduction du cours. Cependant, l'enseignante, prise dans son élan lors de la séance 4, a oublié de distribuer le polycopié qui a alors été donné à compléter à la maison et sert ensuite de retour aux probabilités en début de la partie V (donc nettement plus loin dans le cours).

Nous pouvons remarquer que, comme prévu, des exemples en probabilités sont intégrés dès le début du cours. On trouve dans la deuxième page du cours un exemple de détermination d'un paramètre pour qu'une fonction soit une fonction de probabilité. La fonction de densité n'a pas encore été rappelée dans le cours (cela est notamment dû à l'oubli signalé plus haut), mais les synthèses des séances 2 et 4 sont un appui pour résoudre cette tâche. Cet exemple permet une première institutionnalisation sur la loi uniforme. L'exemple de la puce du DM6 est donné. Plus loin dans le cours, au moment du théorème fondamental, il y a un retour au problème du volcan, pour calculer la valeur exacte de la probabilité $P(0 \leqslant X \leqslant 20)$ qui avait été déterminée de façon approchée lors de la séance 4.

Ensuite, la partie V du cours est exclusivement dédiée aux lois à densité. Contrairement aux autres, cette partie est entièrement donnée sous forme de polycopiés à compléter. L'introduction des probabilités ayant été beaucoup travaillée, l'enseignante souhaite gagner du temps en donnant des polycopiés. Tout d'abord la synthèse des problèmes 1 et 2 redonne le vocabulaire en s'appuyant sur les trois exemples rencontrés (la puce du DM6, le problème de la rencontre et le problème du volcan). Cette synthèse a été complétée par les élèves à la maison et corrigée individuellement par l'enseignante. Ensuite, une fois les propriétés sur les probabilités rappelées, la suite du cours est dédiée à la loi uniforme et à la loi exponentielle, toujours avec des références aux problèmes rencontrés en amont. Seules les démonstrations et les exemples sont à compléter à la main.

Nous pouvons donc voir ici un cours imbriqué comme nous l'avions envisagé en amont. Nous pouvons aussi repérer des mises en relation régulières avec les trois problèmes rencontrés lors de l'introduction de la séquence.

7.4 Validation interne

Après avoir présenté nos analyses relatives au problème de la rencontre et celui du volcan Aso, nous voulons maintenant aborder la phase de validation interne de l'ingénierie didactique collaborative (cf. section 6.1.3, p. 203). Il s'agit, à partir de la confrontation des analyses *a priori* et *a posteriori*, de répondre à nos questions de recherche. Il s'agit d'évaluer l'ingénierie pour elle-même. Nous ajouterons, dans le chapitre 8 d'autres éléments de validation, mais d'un autre ordre.

Il s'agit finalement ici de faire une synthèse des conclusions sur les deux problèmes (sections 7.2.5 et 7.3.4), en reprenant les questions de recherche relatives à nos choix de tâches. Nous les rappelons ici :

- QR3a : Est-ce que les activités, attendues et possibles, des élèves associées à ces tâches (les deux problèmes introductifs) permettent réellement d'introduire la notion de fonction de densité, et en plus le calcul intégral ?
- QR3b : Est-ce que le référentiel théorique lié à la fonction de densité se construit effectivement à l'aide de circulations riches entre les trois ETM en jeu ? Cela engendre-t-il un travail complet ?
- QR3c : Peut-on identifier des éléments (passages obligés, référence à l'ancien...) qui semblent indispensables pour la bonne réalisation de ce travail mathématique ?
- QR3d : Est-ce que l'ETM collectif permet seul de construire l'ETM idoine effectif ? Qui, des élèves ou de l'enseignante, a la charge des changements de domaine et de dimensions dans les ETM dans la construction du référentiel théorique relatif aux probabilités à densité ?
- QR3e : Quel est le rôle de l'enseignante dans le déroulement des activités des élèves ?
- QR3f : Quelles contraintes et conditions a-t-on pu observer qui permettent à la classe d'arriver à l'objectif d'apprentissage visé ?

Nous pensons avoir, après ces analyses détaillées, des éléments pour répondre assez précisément à ces questions.

Tout d'abord, ces analyses ont permis de valider le fait que ces problèmes sont réalisables en classe de terminale S et, qui plus est, qu'elles ont bien l'effet escompté qui est de construire la notion de fonction de densité. Cela est visible dans les institutionnalisations progressives qui ont permis à la fin du problème du volcan d'aboutir au référentiel théorique dans le niveau de généralité souhaité (QR3a). Cependant, cela ne suffit pas pour parler d'une véritable construction.

Un premier indice qui nous permet de parler de construction est la participation active des élèves dans le jeu de questions/réponses dans les phases d'institutionnalisation, qui montre bien que les élèves sont acteurs dans la mise en place du référentiel théorique. Mais bien entendu, il faut revenir précisément aux passages concernant la recherche de la fonction de densité dans les problèmes pour voir le rôle des élèves dans l'émergence de la notion et dans la construction. Les analyses ont permis, dans les deux problèmes, d'observer des dynamiques dans le travail mathématique qui sont assez semblables. Nous avons observé des circulations riches entre les ETM, et

un travail que l'on a jugé complet. Ces dynamiques entre la statistique descriptive, les probabilités et le calcul intégral (plus largement l'analyse) permettent effectivement de faire émerger (dans le problème de la rencontre) et de consolider (dans le problème du volcan Aso) le référentiel théorique visé (QR3b).

Il est important de préciser que l'analyse de la répartition des rôles entre élèves et enseignante a permis de mettre en avant le fait que l'ETM collectif est bien acteur principal dans la construction (QR3d).

Nous avons cependant repéré le rôle important de l'enseignante dans les aides procédurales indirectes qu'elle peut apporter. Ces aides semblent indispensables, mais ne dénaturent pas les activités des élèves, ce qui est important pour l'objectif recherché. Son rôle est donc de permettre le bon déroulement en termes de travail mathématique, sans l'influencer trop, par exemple en faisant des aides procédurales indirectes. En revanche, il est possible qu'elle prenne parfois en charge des activités mathématiques lorsqu'il ne s'agit pas de moments clés dans la construction du savoir, comme par exemple le travail algébrique sur les fonctions dans le problème du volcan. Ces conditions de gestion du travail mathématique permettent à la classe d'arriver à l'objectif d'apprentissage visé, dans un temps raisonnable. Il est cependant indispensable de laisser du temps aux élèves de rentrer vraiment dans les deux problèmes, pour mettre en place une véritable démarche de modélisation (QR3e et QR3f).

Le travail de conception et d'analyse des tâches en collaboration avec l'enseignante a été important pour qu'elle puisse anticiper tous les blocages éventuels. Cette anticipation de l'enseignante a été nécessaire pour qu'ensuite elle réussisse à gérer ces blocages en classe, sans les prendre entièrement en charge (QR3f).

Il existe des « passages obligés » dans le scénario choisi pour que l'objectif soit atteint notamment le passage par l'histogramme pour ensuite déclencher la recherche de la fonction de densité. Si d'autres choix de résolution sont faits en amont (comme c'est possible pour le problème de la rencontre, avec la résolution géométrique), cela rend impossible le travail de construction du référentiel attendu (QR3c).

Enfin, un appui important sur la statistique descriptive est nécessaire, notamment avec l'histogramme. Le DM6 s'est donc révélé, comme prévu, indispensable pour le travail. Ce n'est le DM6 en lui-même, mais bien entendu les notions abordées qui sont nécessaires. Un vrai travail sur l'histogramme est indispensable. Cela permet aux élèves d'avoir un référentiel théorique, lié à la notion d'histogramme, disponible pour ensuite pouvoir l'utiliser dans la construction du référentiel théorique des probabilités. Le problème de la puce a aussi joué un rôle très important dans le problème de la rencontre, avec un travail sur une loi uniforme continue en plusieurs étapes et notamment une phase de simulation et de tracé d'une courbe de tendance, pour que cette démarche soit à l'esprit des élèves. Et pour finir, un travail sur l'aire finie d'un domaine infini (DM5) s'est aussi avéré efficace pour ne pas bloquer le travail mathématique dans le problème du volcan (QR3c).

Pour conclure, les résultats de l'expérimentation sont très positifs quant aux objectifs d'apprentissage visés. Nous n'avons pas mentionné qu'effectivement les deux problèmes, particulièrement le second, ont permis de motiver le besoin du calcul d'aire sous une courbe, et donc du calcul intégral. Le problème du volcan a même lancé une tâche de recherche d'une méthode pour approcher l'aire.

Nous pouvons dire que le travail mathématique mis en place lors de ces problèmes ont déclenché des circulations riches entre les trois sous-domaines (PaD, SD et CI) qui ont abouti effectivement à la construction du référentiel théorique lié à la fonction de densité. Le travail en question est quasi complet, ce qui nous le pensons favorise la construction attendue. Il semblerait donc que ces deux problèmes, en d'autres termes ces deux tâches, permettent effectivement de répondre positivement à notre question de départ, sur l'existence de tâches permettant d'introduire la relation sémiotique $P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x$, ceci par le biais d'une véritable introduction de la fonction de densité. La notion de fonction de densité semble donc pouvoir être une notion, à la fois, réponse à un problème, elle intervient pour résoudre des problèmes de modélisation, et une « extension » avec accident, où elle apparaît comme une extension de l'histogramme, cependant avec accident car il ne s'agit pas d'objets appartenant au même domaine et nous avons bien conscience qu'il ne s'agit pas d'une réelle extension. Pour cette raison, nous pensons que les deux parties sont importantes.

Du fait de notre méthodologie d'ingénierie didactique, cette validation interne à l'expérimentation permet de conclure positivement, sauf que ces résultats restent tout de même liés à notre méthodologie et à ce cas particulier. Nous essaierons donc, dans le chapitre 8, de donner une portée plus étendue à ces résultats. Cependant, avant, nous aimerions dans la section suivante prendre le temps de discuter la méthodologie choisie pour ces analyses *a priori* et *a posteriori*.

7.5 Discussion sur les choix méthodologiques

Dans ce chapitre, nous avons présenté la construction de notre méthodologie d'analyses globales et locales, *a priori* et *a posteriori*, et ensuite nous l'avons mise en œuvre pour l'analyse de nos deux problèmes, le problème de la rencontre et le problème du volcan. Ces analyses nous ont permis de montrer les résultats qui en découlent, dans la section précédente. Cependant, nous souhaitons maintenant discuter spécifiquement sur nos choix méthodologiques d'analyses. Nous voulons discuter de la pertinence et des conséquences de ces choix méthodologiques par rapport à nos questions de recherche et aussi mettre en avant les difficultés de sa mise en place dans les analyses. Nous aborderons quelques points qui nous semblent importants :

- la difficulté de la prise en compte de la démarche de modélisation dans les analyses,
- l'alliance entre les méthodologies de théories de l'activité et des ETM,
- la prise en compte des changements de domaine dans les ETM,
- l'analyse du collectif classe.

7.5.1 Prise en compte de la démarche de modélisation

Le fait que les problèmes conçus supposent une démarche de modélisation nous a nécessairement obligé à adapter la méthodologie « classique » d'analyse de scénarios

et de tâches (Vandebrouck, 2008) essentiellement appropriée pour des analyses de tâches fermées. Notre choix de prendre en compte les étapes du cycle de modélisation dans les analyses nous a permis de structurer nos analyses pour nous assurer une certaine « chronologie » (*a priori*). Cela nous a aussi servi pour identifier les étapes entre mathématiques et réalité et les étapes purement mathématiques. Cependant, il nous semble que nos analyses globales restent des essais d'adaptation qui peuvent encore être améliorés. L'analyse *a priori* en termes de canevas montre bien la difficulté de tout anticiper et surtout de tout détailler en termes d'activités, vu le fait que nous ne savons pas à l'avance quelles tâches vont réellement apparaître dans la classe. Cependant, pour une meilleure précision, il pourrait être intéressant de s'interroger sur ce que nous devons réellement indiquer dans ce canevas et dans le scénario choisi : est-ce que ce sont les tâches envisageables ou les activités envisageables ? Ou bien les deux ? Quel est le statut des propositions des élèves dans les discussions collectives que nous avons décidé de faire apparaître ? Il semblerait que nous considérons dans nos analyses à la fois les tâches et les activités envisageables mais aussi d'autres choses, difficilement identifiables. Ces points mériteraient d'être précisés, nous en avons conscience. Cela serait peut-être l'occasion d'identifier de nouveaux types de sous-activités, liées à la démarche de modélisation. Nous avons déjà introduit l'activité stratégique, qui relève de la prise de décision (consciente ou non) au moment d'un choix à faire sur le modèle. Peut-être que d'autres activités pourraient être prises en compte, comme les sous-activités relatives à des propositions en amont des prises de décisions. Ce sont des réflexions méthodologiques en suspens.

Vis à vis de nos analyses globales *a priori* sous forme de canevas, nous avons conscience que d'autres choix auraient pu être faits. Par exemple, nous aurions pu partir des déroulements pour identifier les tâches qui sont effectivement apparues dans la classe pour ensuite ne considérer que ces tâches et en faire une analyse *a priori*. Cela aurait eu l'avantage de limiter l'analyse *a priori* à un seul scénario, ce qui aurait peut-être permis d'être plus précis. Cependant, ce choix présenterait l'inconvénient de ne pas pouvoir servir de base à une analyse du même problème mais dans une autre classe qui choisirait un autre chemin.

Ces difficultés dues au fait qu'il s'agit de problèmes de modélisation expliquent aussi le fait que nous avons ensuite centré nos analyses plus localement, sur l'étape de mathématisation plus particulièrement. Ces passages ont permis des analyses plus « classiques », cependant des éléments sont tout de même à discuter, notamment en lien avec les deux cadres théoriques dans lesquels nous nous situons.

7.5.2 Alliance entre les méthodologies des théories de l'activité et des ETM

Nos questions de recherche portant sur le travail mathématique en jeu dans les deux problèmes, il nous a paru adapté de considérer des éléments méthodologiques des théories de l'activité (avec quelques adaptations comme nous venons de le souligner dans la section précédente). De même le modèle des ETM nous a semblé *a priori* utilisable pour rendre compte de certaines circulations entre les trois ETM considérés. Cependant, il s'agit d'un premier essai d'allier les deux cadres. Cette

alliance a été source de nombreuses réflexions pour réussir à les combiner de façon adéquate. Nous pensons avoir trouvé une façon de le faire en considérant les tâches et les activités en jeu dans les déroulements au centre de nos analyses pour ensuite en déduire des éléments sur le travail mathématique, vision plus globale sur les activités. Les analyses de tâches que nous avons faites, qui permettent d'identifier les activités précises des élèves, nous ont permis de porter ensuite un regard plus global sur le travail mathématique à l'aide du modèle des ETM. Ces deux analyses nous paraissent donc complémentaires.

La difficulté reste ensuite de savoir quelle précision ou quelle globalité doit rendre compte les ETM. Une de nos difficultés a notamment été le découpage en (sous) phases dans les analyses locales, où à chaque sous-phase est associée un diagramme. Il ne s'agissait pas de rendre compte de la réalité à travers le modèle des ETM mais d'en présenter une certaine représentation à des moments choisis qui nous paraissaient intéressants. C'est à nouveau des éléments de l'analyse des activités qui nous ont permis de choisir ces phases, notamment grâce aux changements de sous-activités, même parfois minimes. Le modèle des ETM permet d'avoir un point de vue global sur les activités engendrées par une tâche, mais il s'agit ensuite de trouver le bon niveau de globalité dont on veut rendre compte.

Pour continuer sur ces questions, les comics que nous avons proposés dans nos conclusions sont vraiment à prendre comme des photos à des temps fixés, qui montrent des circulations à un temps donné. Un comics ne rend pas compte de l'ensemble des circulations et des changements de dimension, et d'ailleurs ce n'est pas le but. Il s'agit de montrer des moments représentatifs qui pourront nous permettre de donner des résultats globaux sur le travail mathématique en jeu. Ce sont justement les choix de ces moments qui peuvent être difficiles à arrêter. D'autres choix de découpages auraient été possibles. Notre choix a l'avantage de montrer la richesse des circulations et le caractère complet du travail mathématique en jeu. De plus, nous y avons fait apparaître la répartition des rôles entre élèves et enseignante, ce qui ajoute encore d'autres éléments importants sur le travail mathématique. Cette analyse est aidée par l'analyse en termes d'aides procédurales directes ou indirectes. Nous pouvons d'ailleurs mentionner le fait que la prise en compte des aides permet de repérer les moments de blocages dans l'ETM et de voir par qui ils sont surmontés. Cela montre à nouveau la compatibilité des deux cadres théoriques.

Pour finir sur ce point, il nous semble donc intéressant de prendre la méthodologie d'analyse de scénarios et de tâches pour avoir un regard précis sur les activités des élèves et aussi ajouter la gestion des activités par l'enseignante, pour ensuite utiliser les ETM pour avoir un point de vue global sur l'activité. L'analyse avec les diagrammes du modèle des ETM permet finalement, rien que visuellement, d'en extraire des renseignements globaux sur le travail mathématique qui a eu lieu. Un retour sur l'analyse des tâches peut alors se faire si l'on veut étudier plus en détail un moment particulier repéré à l'aide des ETM. Ces analyses globales en termes d'ETM permettraient ensuite de faire des comparaisons avec d'autres déroulements, ce qui est beaucoup plus compliqué en termes d'analyse de tâches et d'activités du fait de la précision dans la description des activités.

7.5.3 Prise en compte des changements de domaine dans les ETM et retour sur l'articulation entre probabilités à densité et calcul intégral

Nous avons fait le choix de garder le regard sur les changements de domaine car nous voulions voir apparaître les articulations entre les probabilités à densité et le calcul intégral. Or, dans notre séquence proposée, contrairement au cas des manuels par exemple, le calcul intégral n'est pas un prérequis. De ce fait, l'ETM_{CI} peut être considéré comme vide *a priori*, tout comme l'ETM_{PaD}. Cependant, pour garder une cohérence dans nos analyses et en gardant en tête nos questions de recherche, il nous semblait nécessaire de garder les trois sous-domaines PaD, CI et SD. Nous nous reposons dans nos analyses sur le sous-domaine CI (calcul intégral) lorsque nous avons affaire à des aires sous une courbe et à la fonction associée à cette courbe. Ceci reste cohérent avec le programme et aussi avec nos analyses préalables. Cependant, bien entendu ici encore d'autres choix auraient pu être faits.

Cependant, nous pensons que le fait que les deux ETM soient « vides » au départ permet justement de mieux construire le lien entre les deux sous-domaines. Effectivement, le fait que les deux soient introduits simultanément permet de véritablement construire la notion de fonction de densité. C'est d'ailleurs ce qui fait l'originalité de cette séquence, le sous-domaine du calcul intégral n'est pas prioritaire par rapport à celui des probabilités à densité. Ils se coconstruisent. Le fait que l'ETM_{CI} ne soit pas déjà construit permet aux élèves de se focaliser au départ sur le sens plutôt que sur les calculs.

7.5.4 Analyse du collectif classe

Nous avons fait le choix dans nos analyses de ne pas considérer chaque élève mais bien le collectif classe, à travers l'ETM collectif, où chaque élève peut apporter des éléments de son ETM personnel, ce qui permet de faire avancer l'activité et petit à petit de façonner l'ETM idoine effectif. Dans nos analyses, il s'agissait de ne pas différencier les élèves, cependant ce choix ne nous permet pas d'avoir accès aux élèves particuliers et donc pas non plus réellement aux activités des élèves. Nous ne pouvons finalement juger du résultat de l'expérimentation que par le spectre du collectif et pas des apprentissages individuels des élèves. Nous en avons conscience et cela ne nous parait pas grave pour autant puisqu'ici les séances sont constituées essentiellement de phases collectives et dans ces phases, nous n'avons accès qu'à un collectif et pas à chaque élève individuellement, il suffit notamment de penser à une classe où tous les élèves lèvent la main, cependant l'enseignant ne va en interroger qu'un seul. Il est alors difficile de considérer chaque élève en soi. Cependant, nous avons conscience tout de même des limites que cela peut apporter. Il est possible, par exemple, que seul un élève dans la classe finalement ne participe réellement à l'avancée du travail mathématique et que cela ne soit pas visible à travers nos analyses. Pour cela, nous apporterons des compléments à nos analyses dans la section 8.2.

Cette discussion rend compte de réflexions sur nos choix méthodologiques, qui ont été faits sous certaines contraintes et qui peuvent être remis en cause ou questionnés, comme nous avons pu le montrer. Nous allons maintenant passer au chapitre 8 qui va essayer justement de donner une portée plus grande aux résultats que nous avons mis en avant dans ce chapitre. Nous allons essayer d'étendre la validité de ce processus d'ingénierie didactique collaborative et étendre la réflexion sur ces problèmes d'introduction au delà de la seule expérimentation faite dans cette classe.

Chapitre 8

Sur la portée de nos résultats

Dans le chapitre 6, nous avons présenté notre méthodologie d'ingénierie didactique collaborative et les choix retenus pour la conception notamment des tâches étudiées. Dans le chapitre 7, nous avons détaillé nos analyses *a priori* et *a posteriori* de nos deux problèmes d'introduction. Ces analyses nous ont conduit à des résultats qui ont permis une validation interne à notre ingénierie didactique. Dans ce chapitre, nous voulons élargir la portée de ces résultats. Tout d'abord, comme nous l'avons présenté dans le chapitre 6, la phase de validation de la méthodologie d'ingénierie didactique collaborative, du fait de l'ajout de la dimension collaborative, ne peut se passer d'une autre forme de validation : la validation de l'enseignante. Ensuite, nous proposerons d'autres éléments, qui enrichissent nos résultats. Il ne s'agira pas de faire des analyses poussées, mais seulement de donner des pistes sur une portée plus grande de nos résultats.

Ce chapitre est donc l'occasion de présenter l'évaluation de la séquence, et plus particulièrement des deux problèmes, par l'enseignante (section 8.1). Ensuite, nous reviendrons sur les limites soulevées dans la section 7.5 portant sur l'analyse du collectif. Nous apporterons donc des compléments à nos analyses pour donner des éléments sur les élèves de la classe, considérés individuellement (section 8.2). Nous ajouterons ensuite quelques éléments de validation complémentaire (section 8.3). Il s'agit des réponses des élèves de la classe de Marie à un questionnaire, plusieurs semaines après la séquence. Nous proposerons ensuite d'aborder la question de la reproductibilité de cette ingénierie (section 8.4).

8.1 Validation de l'enseignante

Nous avons mentionné, dans le chapitre 6, le fait qu'une recherche collaborative, au sens de Desgagné *et al.* (2001), peut être vue comme un échange de services entre le chercheur et l'enseignant. Notre méthodologie d'ingénierie didactique collaborative implique donc que la seule validation par la recherche ne peut suffire à la validation de l'ingénierie. Il est donc primordial d'interroger l'évaluation de la séquence par l'enseignante Marie, elle-même. Pour cela, il nous parait important de revenir sur les contraintes qu'elle nous avait imposées au départ pour vérifier si elles ont effectivement été prises en compte et ensuite de s'intéresser à son ressenti en tant qu'enseignante.

8.1.1 Respect des contraintes

a) Les contraintes institutionnelles

La première contrainte était le respect du programme, qui a été effectivement prise en compte. Ensuite, il s'agissait de différentes contraintes de temps. Tout d'abord, il ne fallait pas que les séances consacrées à l'introduction de la séquence excèdent cinq heures, temps qu'elle y consacrait avant. Si l'on se restreint effectivement au travail sur les deux problèmes (avec la recherche d'une méthode pour le calcul d'aire) au cours des quatre premières séances de la séquence, nous arrivons à une durée inférieure à cinq heures (entre 4h30 et 5h). De même, une contrainte avait été donnée sur la durée totale de la séquence, en ajoutant la partie sur le calcul intégral d'une fonction de signe quelconque et celle sur la loi normale, qui ne devait pas dépasser six semaines. Cette contrainte a également été respectée.

b) Les contraintes cognitives

Bien entendu, les contraintes cognitives sur le fait que les élèves doivent savoir résoudre les mêmes tâches (sinon plus) que les années précédentes ont été respectées, nous pouvons dire que les tâches proposées en introduction ont même enrichi les compétences des élèves, qui habituellement ne rencontraient jamais de problèmes de modélisation. Ensuite, l'objectif principal des élèves et donc de l'enseignante est resté le baccalauréat. L'ingénierie n'a pas empiété sur cet objectif.

c) Les contraintes matérielles

Nous avons effectivement pris en compte les contraintes matérielles liées à l'outil informatique. La salle informatique n'étant pas facilement accessible, nous n'avons évidemment pas choisi cette option dans notre ingénierie. Nous avons, par contre, exploité le fait qu'elle disposait d'un ordinateur dans sa salle.

Les différentes contraintes imposées par Marie, en amont de la conception, ont bien été prises en compte dans l'ingénierie, ce qui est déjà un premier point positif pour la validation de l'ingénierie par l'enseignante.

8.1.2 Son ressenti d'enseignante

Pendant et après la séquence, le ressenti de Marie sur les différentes séances est très positif. Nous avons cependant pu voir une légère déception de l'enseignante à la fin de la première séance. Elle n'était pas satisfaite du déroulement de la séance, ayant l'impression que les élèves n'étaient pas impliqués. Elle a finalement été agréablement surprise lors de la synthèse de la séance 2, durant laquelle la participation des élèves lui a semblé bonne et a laissé penser qu'ils avaient en réalité retenu des choses. Elle a ensuite été ravie de l'implication des élèves pour le problème du volcan Aso et a regretté de ne pas avoir laissé les élèves tracer leur propre histogramme, vu qu'ils étaient pour la majorité d'entre eux, lancés dans cette idée.

Marie est très satisfaite de l'implication de ses élèves qui, selon elle, ont très bien réagi lors de ces séances, plus particulièrement lors du problème du volcan.

Nous pensons, personnellement, qu'il est compréhensible que les élèves semblent plus investis dans le second problème. Lors du premier problème, il s'agit d'une première rencontre des élèves avec ce type de problème de modélisation. Il ne semble donc pas étonnant que les élèves soient un peu bloqués et partent dans différentes directions. Finalement, ils ont tout de même réussi à aboutir. La synthèse montre bien qu'ils ont assimilé des connaissances lors de cette première séance. Le problème du volcan est alors un moyen de réinvestir ce qui a pu être mis en place lors du premier problème. Cette deuxième rencontre est importante notamment pour que les élèves, qui la première fois ont peut-être été pour certains plutôt spectateurs, puissent à leur tour rentrer dans le problème. Il nous paraît donc assez naturel que cette deuxième rencontre laisse un sentiment de « meilleure » réussite à Marie, même si finalement il faut se dire que sans le premier problème, il n'en serait certainement pas de même.

Selon Marie, les deux problèmes d'introduction accompagnés des deux devoirs maison ont bien rempli les objectifs didactiques qu'ils avaient, à savoir donner du sens à la fonction de densité et faire émerger les propriétés de cette fonction, ainsi que de motiver le besoin du calcul d'aire sous une courbe. Elle souligne l'importance du devoir maison n° 6, grâce auquel les élèves se sont posés beaucoup de questions lors de la correction. Les différents problèmes ont, pour elle, énormément marqué les élèves.

L'ensemble de la séquence lui a paru adaptée et agréable. Le fait de mélanger les deux chapitres traditionnels ne lui a finalement pas posé de problème. Les élèves n'ont pas été affectés par ces choix, cela leur a même paru naturel.

Pour finir, Marie a décidé de garder cette séquence pour l'année suivante. Il n'est plus question pour elle de revenir à sa façon de faire antérieure. Cette adoption de l'ingénierie nous semble être la preuve même d'une validation de la part de l'enseignante.

L'évaluation de Marie, qui se résume parfaitement par la dernière phrase, est très positive et donne d'autant plus de poids aux résultats exposés dans le chapitre précédent. C'est aussi un constat important pour la recherche, cela valide la viabilité de cette séquence dans une classe. De plus, cette engouement de l'enseignante ouvre de nouvelles perspectives pour relancer la recherche.

8.2 Limites de l'analyse du collectif

Nous voulons maintenant revenir sur notre analyse du collectif, qui en efface les élèves individuels. Nous avons bien conscience que d'avoir fait le choix d'analyser les activités du collectif classe, plutôt que les activités de chaque élève, a ses limites pour valider l'effet des séances. Bien que le déroulement des séances semble s'être bien passé d'un point de vue collectif, il pourrait s'avérer que seul un ou deux élèves aient fait avancer le travail et que finalement cela ne soit bénéfique que pour peu d'élèves. Pour cette raison, il nous semble intéressant pour affiner nos résultats de se pencher rapidement sur la question de l'ETM collectif de la classe. Cet ETM collectif que nous avons analysé est-il réellement « représentatif » de l'ETM de la classe ? Le travail mathématique de la classe est-il réellement collectif ? Ces questions sont pour

nous un moyen de questionner nos résultats, pour peut-être en estimer les limites. Étudier l'ETM collectif de la classe était-il finalement un choix judicieux ici pour approcher le travail de la classe ? Il parait assez difficile de répondre précisément à ces questions. Pour se faire une petite idée, nous allons nous intéresser à la participation et l'engagement dans ces deux problèmes d'introduction. Combien d'élèves participent en classe entière lors de ces séances et combien de fois ? Nous avons bien conscience que ne pas participer ne veut pas dire ne pas prendre part à l'avancée de la classe : certains élèves ne prennent que rarement la parole, d'autres peuvent lever la main mais ne pas être interrogés... Cela nous semble tout de même un bon indicateur pour se faire une idée de l'investissement général et de l'adhésion à ce type de travail.

Pour mesurer la participation et l'engagement dans ces deux problèmes d'introduction, nous rassemblons dans le tableau 8.1 le nombre de participants dans chaque séance.

	Durée	Nombre de participants	Pourcentage
Séance 1	1h50	**19** / 31 présents	61%
Séance 2 (institutionnalisation)	$\sim$ 20min	**16** / 34 présents	47%
Séance 3	55min	**13** / 34 présents	38%
Séance 4 (sans le cours)	$\sim$ 1h25	**25** / 34 présents	74%

TABLE 8.1 – Nombres d'élèves participants oralement lors des problèmes d'introduction

Ce tableau nous montre des taux de participation des élèves plutôt élevés. Il faut bien entendu prendre en compte que les durées considérées ne sont pas les mêmes, ce qui peut parfois laisser moins de place à la prise de parole des élèves. Pour la séance 3, nous pouvons aussi rappeler que des temps de recherche individuel ont pu contribuer à une diminution de la participation générale. Ce tableau ne suffit pas à mettre en lumière la participation de la classe. Dans le tableau 8.2 qui suit, nous avons comptabilisé le nombre de prises de parole de chaque participant dans les différentes séances. Il faut, bien entendu, mettre cela en rapport avec la durée de la séance considérée (cela peut être seulement une partie de la séance). Nous ne prenons en compte que les prises de parole en lien avec les problèmes.

Nombre de prises de parole du participant	Entre 1 et 4	Entre 5 et 9	Entre 10 et 19	Entre 20 et 49	Plus de 50
Séance 1	5	3	6	3 (A22 ; B31 ; B52)	2 (A52 ; C41)
Séance 2	14	1 (B11)	1 (A52)		
Séance 3	2	5	3	3 (A42 ; A52 ; B31)	
Séance 4	10	8	5	2 (A42 ; B11)	

Table 8.2 – Prises de parole des participants

Nous pouvons remarquer qu'il y a un groupe de sept élèves, qui prennent plus la parole que les autres, mais ce ne sont pas nécessairement les mêmes à chaque séance. Ce tableau montre aussi une participation assez bien répartie entre les élèves, même si certains parlent certes plus que d'autres. Il faut garder à l'esprit que, sauf rares exceptions, les élèves prennent la parole de leur propre volonté. Ceci montre une adhésion des élèves aux problèmes et un investissement certain. De plus, comme nous l'avons précisé plus tôt, il ne faut pas considérer que les élèves ne participant pas du tout à l'oral ne sont pas impliqués dans la démarche de la classe.

Pour conclure, nous pensons que ces résultats montrent à leur échelle que, dans notre cas, considérer le collectif classe à travers les interventions des différents élèves a bien un sens. Ce collectif, à travers les interventions des différents élèves, nous semble bien représentatif de la classe.

Nous allons maintenant continuer sur la prise en compte de chaque élève en proposant une analyse des réponses des élèves à un questionnaire, quelques semaines après la séquence.

8.3 Validation complémentaire : questionnaire de fin d'année

Après des résultats très positifs du point de vue de l'analyse didactique, ainsi que du point de vue de l'enseignante, il nous semble intéressant de donner quelques éléments sur les élèves. Les séances analysées ont montré un investissement du collectif dans la construction du savoir sur les fonctions de densité. Comme nous l'avons déjà fait remarquer, nous avons dans nos analyses pu approcher seulement l'ETM idoine effectif et l'ETM collectif de la classe. Il nous semble important pour valider réellement ce travail, d'aller voir du côté de l'ETM personnel des élèves.

La question que nous allons aborder dans cette section est la suivante : Quel référentiel théorique sur la fonction de densité leur reste-il et quel sens donne-t-il à cet objet mathématique, après la séquence ? Nous allons présenter ici quelques résultats sur les réponses des élèves de la classe de Marie à un questionnaire qui leur

a été donné en classe quelques semaines après la fin de la séquence.

8.3.1 Méthodologie relative au questionnaire

a) Le questionnaire

Le questionnaire en question est composé de dix questions. L'objectif de ce questionnaire est beaucoup plus vaste que ce que nous allons présenter ici.

Le but du questionnaire entier est de voir comment les élèves sont capables ou non de prendre en charge des tâches non habituelles (qu'ils ne rencontrent pas en classe) qui nécessitent des changements de sous-domaines entre les probabilités à densité et le calcul intégral. Le questionnaire se trouve en annexe J.1.

Ici, nous allons nous focaliser exclusivement sur les questions portant sur la notion générale de fonction de densité car ce sont les questions qui rentrent plus particulièrement dans notre problématique. Nous allons donc considérer les trois premières questions du questionnaire qui sont les suivantes :

Pour les trois premières questions :
Soit X une variable aléatoire continue à valeurs dans un intervalle I de $\mathbb{R}$. On note f la fonction de densité de probabilité associée à X.

Question 1 Cette fonction de densité de probabilité f vérifie-t-elle des conditions particulières ? Si oui, lesquelles ?

Si la question 1 vous pose problème, levez la main pour être débloqué et cochez cette case ☐

Question 2 Pouvez-vous expliquer, avec vos mots, pourquoi la fonction de densité f doit vérifier les conditions que vous avez cité à la question 1 ?

Question 3 A quoi sert la fonction de densité f ? Pourquoi s'intéresse-t-on à l'aire sous la courbe représentative de cette fonction ?

Ces trois questions sont donc relatives à la fonction de densité et peuvent faire écho aux quatre séances d'introduction que nous avons analysées dans le chapitre 7. La première question porte sur les propriétés que doit vérifier une fonction de densité. Elle peut être vue comme une question de cours, cependant elle n'est pas formulée de cette manière. Il est attendu que les élèves mentionnent les trois propriétés suivantes : la propriété de continuité de la fonction sur I (propriété de terminale S), la propriété de positivité sur I et la propriété que l'aire sous la courbe vaut 1 (ou sous forme d'intégrale). Cette question nous permet de voir le référentiel théorique qui reste chez les élèves après la séquence. Il est proposé aux élèves qui ne sauraient pas répondre à cette question de demander une aide, pour ne pas être bloqué pour la suite du questionnaire.

La deuxième question n'est pas une question de cours. Il s'agit de laisser s'exprimer les élèves sur le sens qu'ils donnent à ces différentes propriétés. Ce sont des

idées qui ont été abordées notamment lors de la synthèse de la séance 2, mais rien n'a été écrit par l'enseignante sur ces points. Il n'est rien attendu de précis sur cette question : les élèves peuvent faire référence à la statistique descriptive en faisant une analogie entre les probabilités et la fréquence, ou encore des justifications peuvent être données dans le cadre même des probabilités en faisant référence à l'aire.

Enfin, la troisième question porte sur l'utilité de la fonction de densité et sur la justification du fait que ce soit l'aire sous la courbe de cette fonction qui nous intéresse. Cette question sert notamment à voir si un lien est effectivement fait entre les probabilités et l'aire sous la courbe de la fonction de densité. La seconde partie de la question laisse la possibilité aux élèves de parler pourquoi pas de l'analogie entre l'aire sous la courbe et l'histogramme. Ici encore aucune réponse précise n'est attendue, c'est surtout l'occasion de laisser parler les élèves pour voir ce qui leur restent.

À travers ces questions, nous pouvons avoir des éléments sur la relation sémiotique $(*)$: est-elle disponible ? Est-elle limitée à une simple écriture sémiotique ou est-ce que les élèves lui donnent un sens ?

b) Les données

Le questionnaire entier a été donné en classe le 5 juin 2015, soit deux mois après la fin de la séquence et donc plus de deux mois et demi après les séances d'introduction. Les élèves ont disposé d'une heure pour y répondre mais la plupart l'ont rendu au bout de 40 minutes. Nous disposons de 29 questionnaires. En cette fin d'année scolaire, six élèves étaient absents pour cause d'épreuves anticipées du baccalauréat.

8.3.2 Résultats de ce questionnaire

Nous allons donc ici présenter les résultats des élèves de la classe de Marie aux trois premières questions de ce questionnaire.

a) Question 1

Aucun élève de la classe n'a demandé d'aide pour cette question.

72% de bonnes réponses

21 élèves sur les 29 (72%) ont mentionné les trois propriétés attendues (cf. figure 8.1).

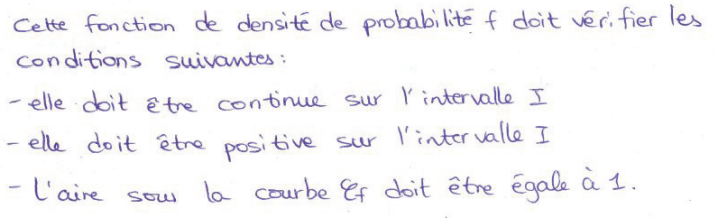

Cette fonction de densité de probabilité f doit vérifier les conditions suivantes :
- elle doit être continue sur l'intervalle I
- elle doit être positive sur l'intervalle I
- L'aire sous la courbe Cf doit être égale à 1.

Figure 8.1 – Réponse de l'élève A11

Plusieurs élèves ont demandé la positivité stricte de la fonction (cf. figure 8.2), ce qui est vrai pour toutes les fonctions de densité qu'ils ont rencontrés. Cette réponse est donc valide.

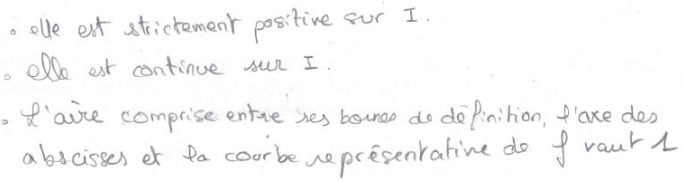


FIGURE 8.2 – Réponse de l'élève A21

Enfin, dans les différentes réponses la majorité des élèves parle de l'aire sous la courbe, seuls trois élèves introduisent l'intégrale (cf. figure 8.3). L'élève B14 donne les propriétés pour une fonction de densité associée à une variable aléatoire à valeurs dans $\mathbb{R}$ et précise aussi la limite de f en $+\infty$ et $-\infty$ (cf. figure 8.4).

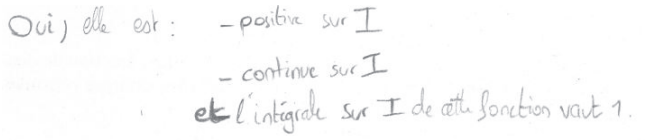


FIGURE 8.3 – Réponse de l'élève C42

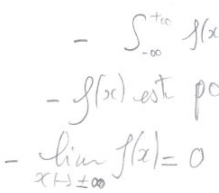


FIGURE 8.4 – Réponse de l'élève B14

Les autres réponses

Sur les 8 élèves restants, deux ont oublié de citer une des trois propriétés. L'élève B34 n'a pas mentionné la positivité de la fonction de densité tandis que l'élève B24 a oublié la continuité. Cinq élèves ont, quant à eux, donné des réponses ambiguës ou confuses, comme par exemple les réponses des figures 8.5, 8.6 et 8.7. On remarque cependant que ces réponses « erronées » ne sont tout de même pas totalement éloignées de la réponse attendue. On retrouve des idées sur ce qu'est une fonction de densité. L'élève B12, par exemple, plutôt que de donner les propriétés de la fonction de densité donne son utilité.

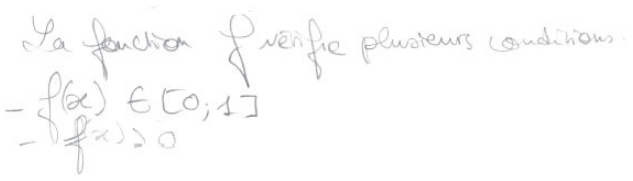


Figure 8.5 – Réponse de l'élève A41

Figure 8.6 – Réponse de l'élève B23

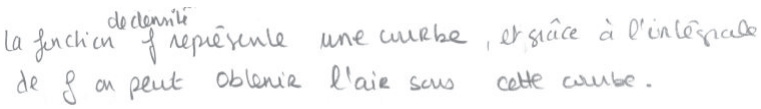


Figure 8.7 – Réponse de l'élève B12

Une seule réponse (cf. figure 8.8) semble à l'écart.

Figure 8.8 – Réponse de l'élève C21

Les résultats à cette première question sont assez bons. 72% des élèves ont su rappeler les trois propriétés que la fonction de densité doit vérifier. Un autre résultat encourageant est que même les élèves qui n'ont pas mentionné les trois propriétés ne semblent pas étrangers à la notion de fonction de densité.

Ce questionnaire a aussi été donné à 11 élèves volontaires, de la classe de l'enseignant B (cf. chapitre 5) de l'année 2014-2015, classe jugée de bon niveau par l'enseignant (les élèves sont en classe européenne). Pour comparer, seuls trois élèves sur les 11 ont mentionné les trois propriétés de la fonction de densité, soit 27% des élèves. Même si nous ne pouvons tirer de conclusions sur la justification de cette différence de résultats, nous pouvons tout de même admettre que dans la classe de Marie les résultats sont nettement meilleurs.

b) Question 2

Il est plus difficile de dégager des catégories dans l'analyse de cette deuxième question. Nous nous proposons de faire seulement quelques commentaires.

Seuls deux élèves n'ont pas répondu à cette question, ces élèves avaient pour autant répondu correctement ou partiellement à la question 1. L'élève B12 qui avait répondu à la question 1 de façon erronée (cf. figure 8.7) s'intéresse à nouveau à l'utilité de cette fonction (cf. figure 8.9).

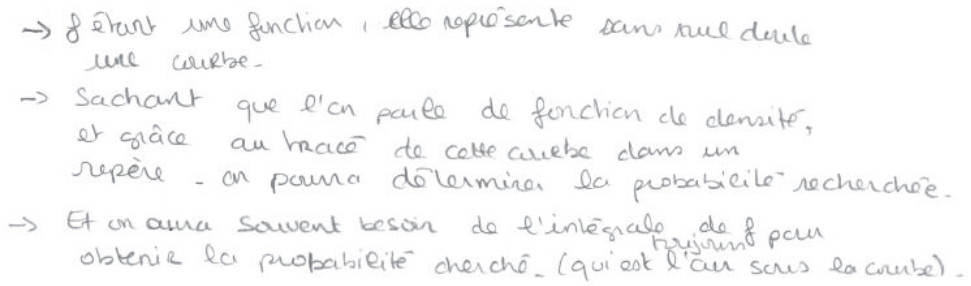


FIGURE 8.9 – Réponse de l'élève B12

Sur les 26 élèves restants, 23 ont donné des justifications en lien avec l'aire sous la courbe et/ou les probabilités. En voici quelques exemples dans les figures 8.10, 8.11, 8.12 et 8.13.

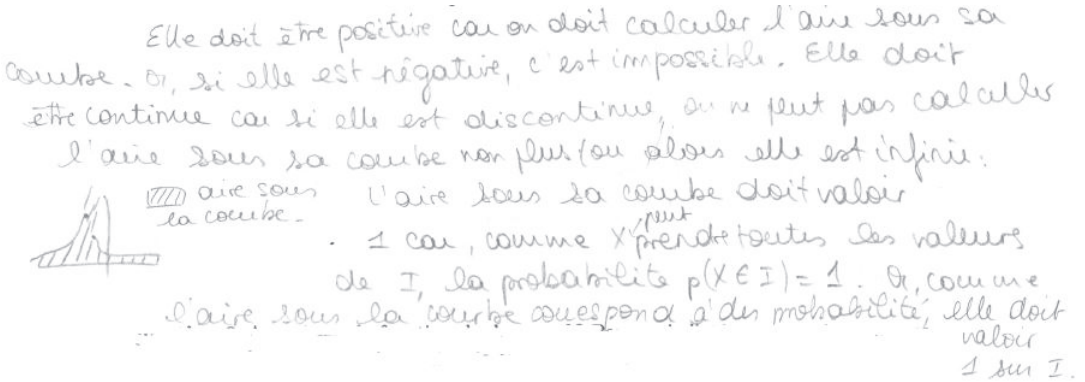


FIGURE 8.10 – Réponse de l'élève B31

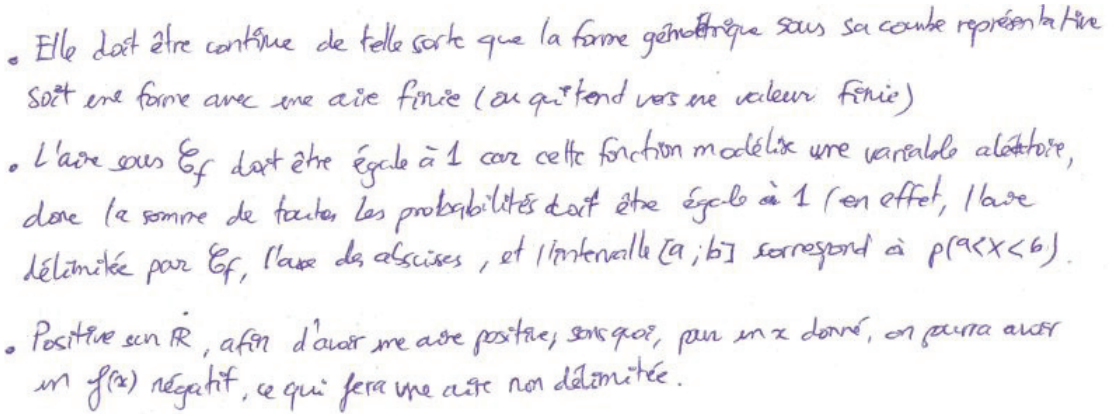


Figure 8.11 – Réponse de l'élève A52

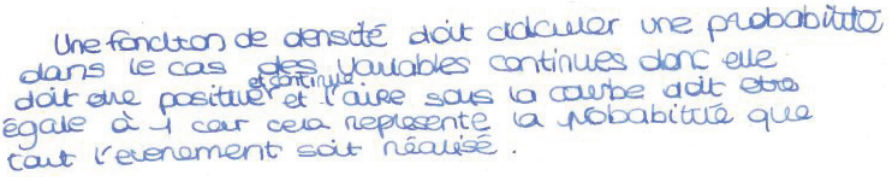


Figure 8.12 – Réponse de l'élève B53

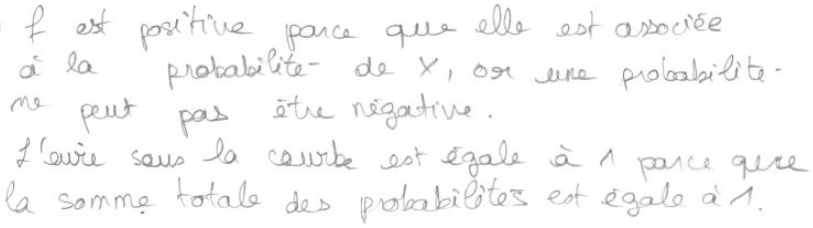


Figure 8.13 – Réponse de l'élève B24

Ensuite un élève a justifié en se référant à l'intégrale et aux probabilités (cf. 8.14).

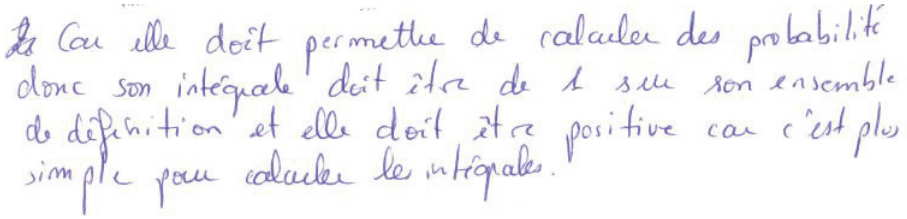


Figure 8.14 – Réponse de l'élève B11

Enfin les deux dernières réponses nous intéressent particulièrement car la première fait référence à la densité (cf. figure 8.15) et la seconde fait référence, et de façon détaillée, à l'histogramme (cf. figure 8.16). Ce sont les deux seules réponses, et particulièrement la dernière, qui font appel à la statistique et donc à la manière dont la fonction de densité a été amenée lors de la séquence.

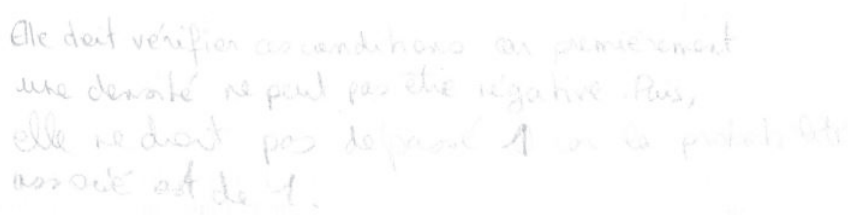


FIGURE 8.15 – Réponse de l'élève C12

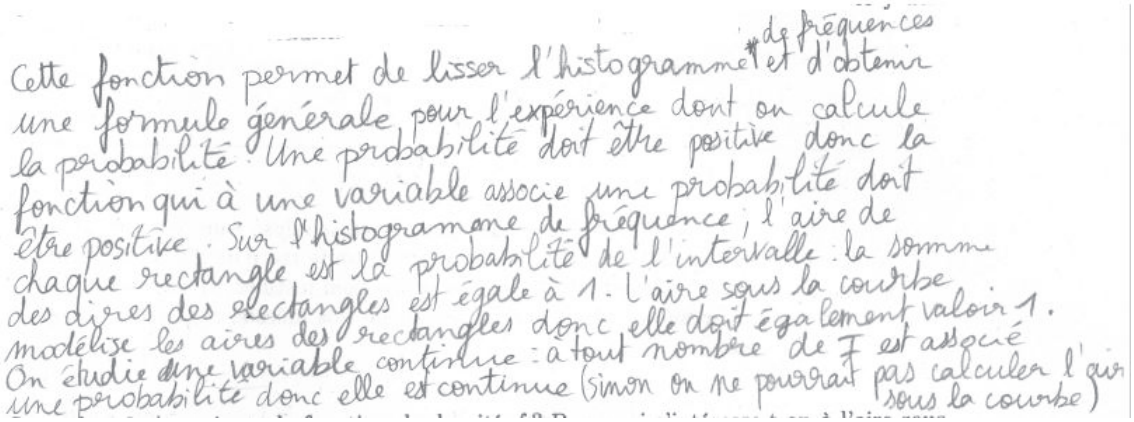


FIGURE 8.16 – Réponse de l'élève B32

Le fait que seulement une élève fait directement référence à l'introduction de la fonction de densité qu'ils ont faite en classe peut paraître très surprenant. Cependant, que les élèves n'en parlent pas ne veut pas dire qu'ils n'en sont pas influencés. Peut-être simplement que les connaissances qu'ils ont construites lors de la séquence entière les ont tournés vers de nouvelles justifications, mettant de côté le sous-domaine de la statistique descriptive. Cela peut parfaitement s'expliquer vu que ce sous-domaine n'est effectivement présent que dans l'introduction.

Nous tenons à préciser que les réponses ne sont pas toujours précises voire parfois contiennent des erreurs, cependant il y a un réel souci de justification chez les élèves. Finalement, très peu donnent des justifications vraiment claires et totalement justes. Cependant mis à part deux élèves, tous ont fait des tentatives intéressantes. Nous pouvons repérer essentiellement des maladresses au niveau du discours.

Nous pensons que les résultats présentés ici sont tout de même assez positifs avec un réel souci d'essayer de justifier les différentes propriétés.

c) Question 3

Dans cette question 3, 21 élèves sur 29 expriment un lien entre la fonction de densité et les probabilités. Cependant seuls 14 d'entre eux précisent explicitement que les probabilités se calculent à l'aide de l'aire sous la courbe de la fonction de densité (cf. figures 8.17 et 8.18). Six élèves s'arrêtent seulement à dire que la fonction de densité sert à calculer des probabilités. Cela ne veut pas dire qu'ils ne savent pas que c'est spécifiquement l'aire sous la courbe qui nous intéresse. Le dernier élève de cette catégorie explique que la fonction de densité modélise une variable aléatoire continue.

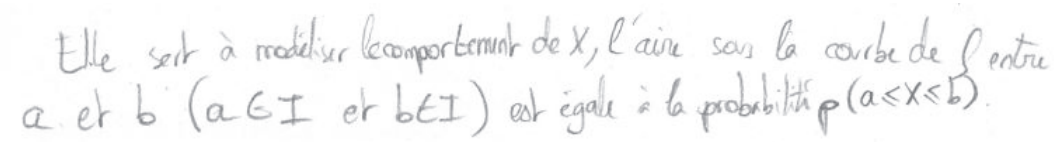


Figure 8.17 – Réponse de l'élève C42

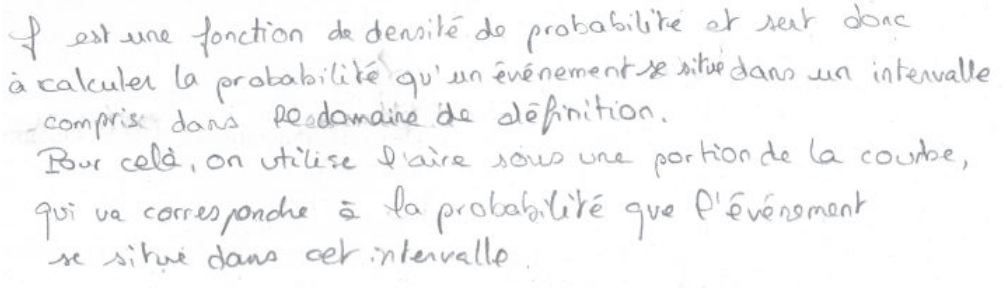


Figure 8.18 – Réponse de l'élève A21

Sur les 8 élèves restants, trois n'ont pas répondu à la question et cinq ont dit que la fonction de densité sert à calculer l'aire sous la courbe. Certains d'entre eux avaient pourtant dans la question précédente parlé de calculs de probabilités.

Vu la tournure de la question, qui ne voulait pas trop influencer les élèves, et vu les réponses précédentes de certains élèves, il est certain que nous ne pouvons pas conclure en disant que seulement 14 élèves sur 29 savent que l'on calcule la probabilité $P(a \leqslant X \leqslant b)$ en déterminant l'aire sous la courbe de la fonction de densité sur l'intervalle $[a; b]$. Pour aller plus loin, nous pensons que la majorité des élèves de la classe le sait.

Pour ajuster nos résultats, il faudrait que nous analysions les questions suivantes du questionnaire, ce que nous ne ferons pas ici.

La seconde partie de la question 3, sur la justification du fait que ce soit l'aire qui nous intéresse et non par exemple l'image, n'a été réellement prise en compte que par une seule élève, qui a utilisé l'histogramme (cf. figure 8.19).

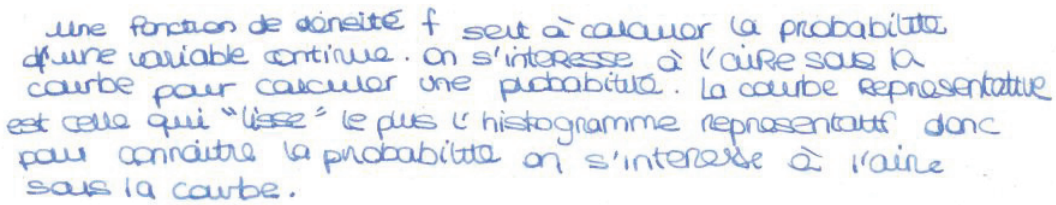


FIGURE 8.19 – Réponse de l'élève B53

8.3.3 Conclusion sur le questionnaire

Les résultats présentés sont encourageants, notamment sur le fait que 72% des élèves sont capables de donner les propriétés de la fonction de densité. De plus, nous avons pu remarquer que même si certains donnent des réponses erronées ou imprécises, la fonction de densité ne semble être une notion inconnue pour personne, ce qui est bon signe. La question 2 a montré la volonté de l'ensemble des élèves (excepté deux) à essayer de justifier les propriétés de cette fonction. La majorité des explications a été donnée en faisant référence à l'aire sous une courbe et/ou aux probabilités. Enfin, à la dernière question, les réponses sont plus mitigées si l'on s'arrête à ce qui est purement écrit dans la réponse des élèves. Elles ont montré que la moitié des élèves fait le lien entre probabilité et aire sous la courbe de la fonction de densité. Cependant nous avons précisé que les élèves faisant le lien entre probabilité et fonction de densité avaient certainement en tête l'aire sous la courbe (et non seulement la fonction). De plus pour les autres élèves, en regardant leurs réponses précédentes, nous pouvons entrevoir chez certains ce lien.

Un élément qui peut surprendre est le fait que seulement deux élèves sur les 29 ont parlé à un moment donné de l'histogramme. Comme nous l'avons dit précédemment cela ne veut pas nécessairement dire que l'articulation entre les probabilités à densité et la statistique descriptive a totalement disparu de leur esprit mais que, tout du moins, il ne constitue plus pour eux un moyen de justification même si ça a été le cas dans l'introduction de la séquence.

Il faut tout de même moduler ces résultats en prenant en compte que la formulation des questions a pu déstabiliser certains élèves. Le fait de ne pas vouloir influencer les élèves nous a fait choisir des tournures de phrases qui peuvent ne pas être claires pour certains, par exemple dans la question 1 avec « **cette** fonction de densité » qui peut introduire une confusion entre le général et le particulier. La question peu formelle qui serait « Quelles sont les contraintes (ou les propriétés) que doit vérifier une fonction de densité ? » aurait peut-être donné encore de meilleurs résultats.

Pour conclure, ces trois questions du questionnaire ont pu approcher des éléments de l'ETM personnel des 29 élèves de la classe de Marie présents ce jour-là. Il est certain que pour l'ensemble des élèves, à l'exception peut-être de l'élève C21, la fonction de densité fait partie de leur ETM personnel. Après, le référentiel théorique associé n'est pas nécessairement assimilé par tous, mais tout de même par une majorité. Mis à part pour deux élèves, aucune analogie entre probabilités à densité et statistique descriptive ne sert d'appui à la justification, comme cela a été le cas

pourtant dans l'introduction de la séquence. Les élèves préfèrent s'appuyer sur les probabilités et le calcul intégral (aire sous la courbe). Nous expliquons cela par le fait que les connaissances qu'ils ont construites au cours de la séquence entière les ont tournés vers d'autres justifications, mettant de côté le sous-domaine de la statistique descriptive ; ce qui ne veut pas dire qu'ils n'ont plus cette introduction à l'esprit.

Les résultats sont encourageants même plutôt positifs et constituent, nous semble-t-il, une validation complémentaire de l'ingénierie ou tout du moins ne permet pas de l'invalider.

8.4 La question de la reproductibilité

Enfin, pour finir, nous voulons aborder la question de la reproductibilité de cette séquence. Le choix d'une méthodologie de type ingénierie didactique pose la question de la transmission et de la reproductibilité de la séquence expérimentée (Artigue, 1988). Nous avons choisi d'ajouter une dimension collaborative à cette recherche en particulier pour assurer une certaine viabilité de la séquence d'enseignement. Notre travail collaboratif se limitant à deux personnes (une chercheuse et une enseignante), nous pouvons nous questionner sur la véritable portée de la validation de notre ingénierie. Certes elle est viable dans cette classe avec cette enseignante, mais cela ne peut peut-être pas suffire à sa reproductibilité. Bien que l'expérimentation ait eu lieu dans une classe dans des conditions habituelles, nous avons bien conscience que des conditions particulières existaient, comme par exemple notre présence, en tant que chercheuse, dans la préparation des séances et aussi en observateur dans la classe. De plus, la motivation et l'implication de l'enseignante dans le projet jouent forcément un rôle. Cependant, la séquence et plus particulièrement les problèmes de modélisation en introduction de la séquence restent-ils viables dans d'autres classes ? Avec d'autres enseignants ?

8.4.1 Le cas de Laurent

a) Une reproduction concluante

Nous disposons de quelques éléments pouvant attester d'une possible reproductibilité de cette ingénierie. Il s'agit du cas d'un collègue de Marie, du même établissement scolaire. Lors de notre travail collaboratif sur l'année 2014-2015, Marie a discuté dès le début avec son collègue, que nous appellerons Laurent, des idées que nous avions sur la séquence. Au fur et à mesure de nos séances de travail, Marie a continué de partager nos réflexions et les différents devoirs maison et problèmes avec son collègue, qui finalement a décidé de faire de même dans sa classe. Personnellement, nous n'avons pas rencontré Laurent en amont de sa séquence. Nous l'avons rencontré et avons discuté avec lui seulement après qu'il ait traité dans sa classe le problème de la rencontre et le problème du volcan Aso. Laurent s'est donc emparé de cette ingénierie (n'incluant ici que le devoir maison n° 6, les problèmes d'introduction et le fait de coupler les probabilités à densité et le calcul intégral), par l'intermédiaire de sa collègue Marie qui lui a fourni les documents que nous avions conçu ainsi que certaines explications bien entendu. Comme nous n'avions

pas travaillé ensemble avant qu'il mette en place la séquence, Laurent a fait quelques adaptations, ce qui est entièrement naturel. Par exemple, par manque d'anticipation, le devoir maison n° 5 n'a pas été donné en amont, ce qui a nécessité des adaptations. De plus ayant travaillé les problèmes de son côté, il a fait certains choix différents sur l'interprétation des données sur le volcan Aso. Par exemple, il a décidé d'attribuer à la moitié des données de valeur k la valeur $k - \frac{1}{2}$ et à l'autre moitié $k + \frac{1}{2}$. Ce choix a nécessairement influé sur l'itinéraire choisi par la classe, mais ne modifie en rien la validité du travail. Ces adaptations sont naturelles et même indispensables, cela montre l'implication de l'enseignant dans la mise en place de ces problèmes. Nous pouvons de plus préciser que ces adaptations ne font pas perdre l'objectif de ces problèmes, nous pouvons donc dire que l'ingénierie a effectivement été reproduite ici. Il faudrait cependant des analyses plus fines pour réellement justifier ce résultat.

b) Ingénierie testée dans quatre classes

Au final, Laurent a lui aussi décidé de recommencer la séquence l'année suivante. Les deux enseignants ont maintenant clairement intégré l'ingénierie dans leurs pratiques. À l'heure actuelle, l'ingénierie a donc été testée par deux enseignants sur deux années consécutives, ce qui fait donc quatre classes différentes. Il s'avère que ces quatre classes ont, d'après les deux enseignants, quatre profils différents, que nous détaillons dans le tableau 8.3.

	Marie	Laurent
2014-2015	Forte	Faible
2015-2016	Très hétérogène, peu active	Très forte

TABLE 8.3 – Profils des classes de Marie et Laurent sur les deux dernières années

Pour ces quatre classes, les retours de Marie et Laurent sont positifs. Bien que nous n'ayons pas tous les éléments pour conclure, il semblerait que cette ingénierie soit reproductible dans des classes à profils différents. Nous tenons cependant à préciser que l'établissement dans lequel travaille Marie et Laurent est de milieu social plutôt favorisé et qu'aucun problème de discipline ne perturbe la classe, ce qui effectivement est à prendre en compte dans ces éléments de discussion. Ce qui est sûr c'est que dans ces quatre classes, les deux problèmes d'introduction ont réellement suscité l'intérêt des élèves. Un petit bémol est ressorti chez certains élèves dits « très bons » qui ont trouvé cette introduction très longue alors qu'il suffisait de leur donner directement la définition. C'est justement le problème des élèves très scolaires qui ne voient pas nécessairement l'intérêt de construire une notion et de faire émerger des liens, mais qui préfèrent appliquer ce qu'on leur apprend. Ce ressenti de quelques élèves est cependant nuancé par Laurent qui dit que tous les élèves étaient tout de même investis lors de ces séances et que c'est peut-être plus un ressenti après-coup.

Ces éléments apportent des conjectures quant à la reproductibilité possible de cette ingénierie, tout d'abord par d'autres enseignants et ensuite dans des classes de profils différents. Il nous semble cependant important de se demander quelles conditions seraient nécessaires pour une reproduction concluante de cette ingénierie.

8.4.2 Des conditions à la reproductibilité

Le travail collaboratif avec Marie a assuré la viabilité de l'ingénierie dans sa classe. À travers l'exemple de Laurent, nous venons de voir qu'il est possible pour un autre enseignant de se l'approprier. Cependant, il nous semble que des conditions sont indispensables, notamment du côté de l'enseignant, pour la réussite de cette appropriation de l'ingénierie. Voici les facteurs qui nous paraissent indispensables pour la reproductibilité :

- La motivation de l'enseignant : La première chose primordiale est que l'enseignant ait réellement envie de modifier ses pratiques, pour qu'ensuite il puisse réellement s'impliquer dans l'appropriation de l'ingénierie.
- La formation : La motivation seule ne peut suffire, il faut effectivement que l'enseignant soit formé pour qu'il comprenne les enjeux de ces deux problèmes, notamment pour ne pas passer à côté de l'essentiel. Il ne suffit évidemment pas de donner à l'enseignant l'énoncé du problème et des éléments de réponses. Il faut notamment que les enseignants aient identifié les blocages éventuels qui peuvent apparaître, pour pouvoir les anticiper.
- La proximité avec les pratiques habituelles : Comme c'était le cas pour Marie, pour que l'ingénierie soit possible, il faut effectivement que ce qui est proposé dans l'ingénierie reste proche des pratiques habituelles de l'enseignant. Pour cette raison, elle ne peut évidemment pas convenir à tous les enseignants. Par exemple, un enseignant qui n'a pas l'habitude de laisser s'exprimer les élèves, sans intervenir et donner son avis, risque en s'emparant du problème de la rencontre par exemple de fermer au fur et à mesure les questions et donc de faire disparaitre l'objectif de départ qui est de laisser les élèves construire la notion de fonction de densité.

Voici trois points qui nous semblent nécessaires pour permettre la reproductibilité. Un dernier point, qui finalement pourrait être moteur de la motivation des enseignants, est le temps. Les probabilités à densité sont une partie récente du programme qui n'a pas été spécialement bien accueillie par les enseignants, il faut donc laisser du temps aux enseignants pour « digérer » cette nouveauté pour ensuite pouvoir s'impliquer dans ce contenu.

Nous voudrions revenir sur l'aspect formation qui nous semble effectivement indispensable pour la reproductibilité et mettre ceci en relation avec une partie du travail collaboratif que nous avons mis entre parenthèses dans notre thèse. Reprenons le schéma de notre modèle d'ingénierie didactique collaborative (cf. figure 6.2, p. 207). Nous avions repris à la méthodologie de type recherche collaborative, les finalités qui sont, d'un côté, des retombées pour la communauté de recherche, ce qui a fait l'objet de notre Partie III et, d'un autre côté, les retombées pour la communauté de pratique, que nous avons décidé de mettre de côté dans la thèse. C'est justement

à ce niveau que se situe la diffusion et la formation autour de cette séquence d'enseignement à la communauté enseignante. À l'heure actuelle, aucune diffusion n'a été commencée, mis à part pour le cas de Laurent. La première diffusion dans ce cadre se fera avec Marie lors d'un atelier-communication aux journées nationales de l'APMEP (Association des Professeurs de Mathématiques de l'Enseignement Public) en octobre 2016. La production d'une ressource serait à réfléchir et à envisager.

Il faudrait ensuite tester la séquence sur de nouveaux enseignants pour conclure sur sa réelle reproductibilité. Le cas de Laurent nous donne tout de même bon espoir.

8.5 Impacts et portée sur nos résultats

Pour compléter la validation interne du chapitre 7, qui est relative à nos choix méthodologiques d'analyses, nous avons essayé d'étendre la portée de ces résultats. Pour répondre à la méthodologie de type ingénierie didactique collaborative, nous avons pris en compte l'évaluation de l'enseignante, qui s'est avérée très positive. Elle a décidé d'adopter définitivement la séquence conçue. Ce résultat, positif pour l'enseignante, l'est aussi pour la recherche en ce qui concerne la viabilité mais aussi la reproductibilité éventuelle.

Ensuite, nous avons, par une petite étude quantitative, essayé d'approcher qui se cache réellement derrière le collectif classe que nous avons pris en compte dans nos analyses. Est-ce tous les élèves ? Quelques élèves ? Une majorité d'élèves ? Dans notre cas, nous avons montré une vraie implication de la classe, ce qui nous permet de dire que l'ETM collectif était nourri par un bon nombre d'élèves. Cela pose tout de même la question de la prise en compte du collectif, plus généralement. Est-ce qu'il suffit qu'un élève prenne la parole en classe entière pour considérer le collectif classe ? Ceci amène de nouvelles questions.

Nous avons ajouté une analyse complémentaire pour essayer de voir l'impact de l'ingénierie sur les élèves. Ces résultats présentent une autre forme de validation. Nous repérons quelques points mitigés, mais tout de même encourageants sur la notion de fonction de densité et le sens que les élèves lui donne. Il est certain que pour la plupart des élèves de la classe, la relation sémiotique $(*)$ ne se limite pas à une simple écriture sémiotique.

Enfin, nous avons abordé la question de la reproductibilité de la séquence conçue. Le cas de l'enseignant Laurent, qui lui aussi a mis en place les tâches dans sa classe, est encourageant. Nous avons essayé de dégager les points nécessaires à la reproductibilité. L'étude de la reproductibilité constitue des perspectives de poursuite de la recherche.

Conclusion générale

Cette conclusion générale va tout d'abord nous permettre de synthétiser les différents résultats de la thèse (section 9.1) et en présenter les limites (section 9.2). Dans la section 9.3, nous mettrons en évidence les apports de cette recherche, au-delà des résultats. Enfin, nous terminerons en proposant des prolongements à ce travail et des perspectives de recherche (section 9.4).

9.1 Synthèse des résultats

Les probabilités et l'analyse sont deux domaines mis en relation par l'écriture sémiotique :

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x. \qquad (*)$$

Nous avons orienté notre recherche autour de cette relation sémiotique. Après avoir mis en évidence le rôle central de la fonction de densité dans la relation entre les probabilités à densité et le calcul intégral, notre question de recherche principale s'est précisée :

> Dans le cadre de la terminale scientifique, est-il possible de concevoir et de mettre en œuvre des tâches dont les activités, attendues et possibles, des élèves leur permettent de construire conceptuellement la notion de fonction de densité ?

L'objectif central a donc été de concevoir et d'expérimenter dans une classe de terminale S des tâches permettant *a priori* de construire les connaissances relatives à la notion de fonction de densité. Les différentes analyses menées dans la thèse ont permis d'aboutir à une réponse positive à notre question principale qui a pu être complétée. Nous allons ici rappeler les différents résultats obtenus.

Dans le cadre de la Partie I, nous nous sommes focalisée sur une étude historique et épistémologique. Nous avons abordé la naissance du concept de lois à densité, et plus particulièrement de fonction de densité (QR1), dans un but de mieux comprendre cette notion mathématique. Cette étude nous a montré le rôle important de la théorie des erreurs dans la naissance des lois à densité. Ce sous-domaine est donc apparu en réponse à des problèmes concrets d'erreurs de mesures. La statistique descriptive, à partir des données statistiques des mesures, a donc joué un rôle central dans la construction historique de la relation sémiotique $(*)$ entre les probabilités à densité et le calcul intégral. Des documents historiques nous ont aussi montré que la notion d'histogramme de fréquences a été un moyen utilisé pour enseigner la théorie des erreurs à des non-mathématiciens. Ces considérations épistémologiques plus

fines nous ont ensuite servies pour aborder nos questions didactiques, notamment en se demandant le rôle laissé à la statistique descriptive dans l'introduction de la fonction de densité, en terminale S.

En prenant en compte les résultats de la partie précédente, la Partie II a été centrée sur la question QR2 :

> En terminale S, quelles sont les tâches proposées et les activités attendues des élèves pour l'introduction de la notion de fonction de densité ? De plus, quelles sont les tâches qui permettent d'exploiter la relation sémiotique (∗) ?

L'objectif était de comprendre quel sens est donné à la relation sémiotique (∗) dans les classes de terminale S. Comment émerge-t-elle ? Comment évolue-t-elle dans les classes ? Nous avons d'abord cherché à aborder ces questions au niveau de l'ETM de référence, c'est-à-dire à travers les documents institutionnels (programme, document *Ressources* et baccalauréat). Ce premier niveau d'analyse nous a montré un programme avec des ambiguïtés et notamment une absence de recommandations sur l'approche à mettre en place pour introduire les lois à densité, mis à part dans le cas de la loi normale centrée réduite pour laquelle l'introduction par la visualisation du théorème de Moivre-Laplace est demandée et fortement commentée. La place importante prise dans ces documents par la loi normale ne permet pas de révéler de véritable exploitation du lien entre probabilités à densité et calcul intégral ; on remarque que l'utilisation de la calculatrice est fortement présente. Seuls les quelques calculs liés aux lois uniformes et exponentielles présentent une application de la relation (∗). Les sujets du baccalauréat montrent même une opacité entre les deux domaines, qui restent clairement séparés malgré des potentialités existantes.

L'analyse s'est poursuivie à un second niveau, celui des ETM idoines potentiels. Dans l'étude des manuels, nous avons mis en évidence des tentatives d'introduction à la notion de fonction de densité. Les manuels plus spécifiquement étudiés proposent un passage par la statistique descriptive, par le biais de la notion d'histogramme. Il y a donc une interaction entre la statistique descriptive, les probabilités à densité et le calcul intégral. Tout comme nous l'avons remarqué dans notre étude épistémologique, l'histogramme ici aussi va jouer un rôle important. Cependant, nous avons montré que ces *activités* d'introduction ne permettent pas réellement de construire la notion de fonction de densité, notamment à cause du fait que la courbe de la fonction de densité considérée est toujours donnée par les auteurs des manuels, ce qui ne permet pas de susciter un véritable questionnement sur cette fonction et ses propriétés. Ces introductions reposent essentiellement sur la visualisation ; aucun appui théorique n'est réellement pris en compte, ce qui entraîne finalement à imposer les propriétés de la fonction de densité. Un autre point important à souligner concerne l'histogramme. Cet objet sur lequel s'appuie les manuels est, dans pratiquement tous les cas, mal défini ou mal représenté, ce qui questionne sur la véritable aide que peut être le passage par cette représentation graphique, mal connue. Une étude plus poussée sur l'utilisation de cet objet mathématique dans les manuels de seconde a confirmé des approches didactiques et même mathématiques erronées de la notion d'histogramme, ce qui confirme d'autres recherches (Régnier, 1998 ; Roditi, 2009).

Nous avons tout de même mis en avant un point de recollement possible autour de ce que nous avons appelé les niveaux de calculs d'aires. Effectivement, que ce soit en probabilités à densité ou en calcul intégral, différents niveaux de calcul d'aires sont rencontrés : le niveau A1, le calcul d'aire élémentaire (loi uniforme), le niveau A2, le calcul par « primitivation » (loi exponentielle), et le niveau A3, le calcul approché d'aire (loi normale). Ce point non exploité vu la présence de méthodes parfois différentes suivant le domaine considéré nous a paru une piste à explorer pour la suite. Nous avons terminé cette partie en confirmant quelques résultats avec la présentation d'éléments d'entretiens de quatre enseignants de terminale S.

L'analyse des manuels sur l'ensemble des deux chapitres en jeu a confirmé le fait que la relation entre probabilités et analyse est très peu exploitée même parfois inexistante, si ce n'est dans les démonstrations de cours. La relation sémiotique $(*)$ semble réduite à une écriture sémiotique, formule que les élèves appliquent sans (pouvoir) lui donner véritablement de sens. Nous avons pu même repérer que parfois la relation sémiotique peut disparaitre avec l'utilisation de la calculatrice pour la loi normale par exemple.

Cette partie nous a permis de conclure que les manuels, et de ce fait certainement de nombreux enseignants, ne proposent pas de tâches permettant réellement de construire la notion de fonction de densité. Cela confirme donc l'hypothèse que les élèves ne peuvent pas réellement donner de sens à la relation sémiotique $(*)$. Cependant, ces analyses nous ont aussi permis de dégager des pistes, comme l'utilisation de la notion d'histogramme, mais en proposant un véritable travail sur cet objet en amont pour qu'il soit disponible chez les élèves, ou encore de ne pas imposer la fonction de densité en jeu dans l'*activité* mais au contraire la laisser chercher aux élèves.

Les Partie I et Partie II nous ont alors servi d'analyses préalables dans la méthodologie d'ingénierie didactique dans laquelle nous avons voulu tester les pistes soulevées. Dans la Partie III, nous avons donc considéré la question QR3 :

Quels éléments didactiques peut-on prendre en compte pour concevoir et mettre en œuvre effectivement des tâches qui permettent aux élèves de construire la notion de fonction de densité ?

Pour la conception de ces tâches, nous avons travaillé conjointement avec l'enseignante Marie et nous avons pris en compte ses contraintes. Cette prise en compte de l'enseignant dans notre méthodologie a été clairement définie dans ce que nous avons appelé une ingénierie didactique collaborative.

Pour la phase de conception, nous avons décidé de privilégier des tâches favorisant la circulation entre les probabilités à densité, le calcul intégral et les lois à densité. Nous avons donc abouti sur la proposition de deux problèmes d'introduction complémentaires s'inscrivant dans une séquence d'enseignement complète articulant à la fois les probabilités à densité et le calcul intégral, dans laquelle le calcul intégral n'est pas un prérequis aux probabilités, ce qui en fait son originalité. Le problème de la rencontre et celui du volcan Aso sont deux problèmes de modélisation dont l'objectif est la construction de la notion de la fonction de densité et l'introduction du calcul intégral. Des analyses poussées de ces deux problèmes ont permis de conclure sur la validité de cette ingénierie didactique. Les deux problèmes ont effectivement

mené les élèves à la construction de la notion de densité par le biais de l'ETM collectif de la classe. Certains éléments de gestion du travail mathématique se sont révélés importants pour accompagner les activités des élèves en permettant de mener à bien le travail mathématique, sans perdre de vue l'objectif de construction du référentiel théorique. Nous pensons par exemple à la présence d'aides procédurales indirectes, qui ont permis de relancer le travail dans les moments de blocages.

Le chapitre 8 a présenté quelques éléments permettant d'élargir la portée des résultats révélés par l'expérimentation. Notamment, nous avons présenté une analyse des réponses des élèves de la classe de Marie à un questionnaire donné quelques semaines après l'expérimentation ou encore, nous avons questionné la possible reproductibilité des problèmes de la rencontre et du volcan Aso dans le cadre d'une séquence couplée.

En conclusion, pour donner du sens à la relation sémiotique $(*)$, il nous a semblé important de nous centrer sur la construction de la notion de fonction de densité. Pour permettre cette construction, nous avons montré que de permuter l'ordre d'apparition des notions liées aux probabilités à densité et au calcul intégral incite à donner réellement un sens à la fonction de densité. En effet, le fait que les élèves ne connaissent pas l'intégrale ne donne pas la priorité au calcul, mais permet au contraire de réellement se questionner sur la notion de fonction de densité et permet de faire apparaitre le besoin de calculer des aires sous une courbe. La démarche de modélisation joue un rôle important dans cette introduction.

9.2 Limites de cette recherche

Nous avons conscience que les choix méthodologiques faits dans cette recherche présentent des limites.

Une première limite se trouve dans les analyses préalables. Nous avons fait le choix de décrire les ETM idoines par le spectre des ETM idoines potentiels proposés par les manuels. Ce choix nous a permis de disposer d'un panel assez grand de possibles, que nous avons supposé représentatif des choix que peuvent faire les enseignants. Cependant, tout d'abord, les manuels proposent des tâches qui, même si elles sont reprises par un enseignant, peuvent être modulées, adaptées, ce qui fait que les activités attendues en sont modifiées. De plus, les échanges en classe nécessitent encore des modifications non présentes dans les manuels. Nous avons, tout de même, pris en considération les choix de quatre enseignants, dont deux qui ne font pas d'*activité* d'introduction. Pour les deux enseignantes qui en proposent une, nous n'avons cette fois encore pas accès à tous les paramètres. Pour compléter notre étude, il serait intéressant d'assister au déroulement de séances d'introduction « classiques » dans différentes classes.

Une deuxième limite que nous avons déjà évoquée dans la section 7.5.2 de la discussion méthodologique (chapitre 7) est la limite de nos analyses du collectif. Dans les analyses des deux problèmes, nous avons fait le choix de considérer le collectif classe plutôt que les élèves individuellement. Bien entendu, dans ces conditions, nous perdons des informations sur les activités de chaque élève. Cependant, il est indéniable que, dans ces deux problèmes, c'est le collectif qui permet la construction du référentiel théorique et non tous les élèves individuellement. La section 8.2 du

chapitre 8 permet de montrer que de considérer le collectif classe, dans notre cas, reste assez représentatif de la classe du fait de la bonne participation des élèves. Il n'empêche que cela ne renseigne pas sur les activités précises de chaque élève.

Une autre limite de ce travail est que l'expérimentation n'a été menée que dans une seule classe. Dans le cadre de la thèse, l'analyse d'une seule classe avec un seul enseignant est déjà un premier pas, mais il est évident que cela limite la portée des conclusions. Nous avons cependant évoqué dans la section 8.4 (chapitre 8), le cas de Laurent, collègue de Marie, qui a lui aussi mis en place les deux problèmes d'introduction. Bien entendu, il ne s'agit pas là de résultats de recherche mais cela peut donner des idées de poursuite de la recherche. Nous y reviendrons dans les perspectives.

Nous avons aussi dû dans la thèse nous restreindre au niveau du contenu mathématique. Nous aurions évidemment pu nous pencher sur la question de l'introduction de la loi normale à l'aide du théorème de Moivre-Laplace, ce qui n'a pas été le cas.

9.3 Apports de ce travail

D'un point de vue plus général, nous pouvons nous demander quels sont les apports de ce travail de thèse. Nous en voyons trois directions.

Le premier est un apport disciplinaire. Les travaux en didactique des probabilités en particulier sur les lois à densité sont encore très peu développés, notamment en France (cf. section 2.2). Nous pensons donc que ce travail spécifiquement centré sur la fonction de densité apporte une nouveauté dans le domaine de la didactique. Nous avons fait le choix de ne pas nous focaliser sur une loi de probabilité particulière mais sur la notion générale de fonction de densité. De plus, nous avons adopté un point de vue original en nous intéressant à la question de la mise en relation entre les probabilités et l'analyse à travers l'introduction de cette notion de fonction de densité. Ce travail a débouché sur la conception et la mise en œuvre de tâches d'introduction, dont les résultats sont concluants.

Le deuxième apport est cette fois-ci plus théorique. Il s'agit d'un premier travail de recherche essayant d'allier le cadre des Espaces de Travail Mathématique et la théorie de l'activité. Nous pensons avoir montré une compatibilité des deux cadres théoriques et même une complémentarité. Il nous semble que les analyses de tâches et des activités permettent effectivement une analyse fine, qui permettent ensuite de porter un regard global sur le travail mathématique en jeu dans le cadre des ETM (en prenant conscience que, du fait de ce point de vue plus global, les détails des activités échappent au modèle des ETM). Cependant, rien n'empêche, une fois un passage clé repéré grâce aux ETM, de revenir à l'analyse fine de la tâche (cf. section 7.5.2).

Cette alliance théorique a nécessité des adaptations méthodologiques qu'il a fallu construire et mettre en œuvre. Nous pensons notamment aux analyses locales des déroulements que nous avons proposées, qui ont notamment débouchées sur les « comics » décrivant les différentes phases étudiées. Notre troisième apport est donc méthodologique. Nous venons d'en préciser des éléments.

Nous ajouterons aussi une autre adaptation méthodologique. L'analyse de tâches « classique » a été jusqu'à maintenant utilisée pour décrire des tâches simples ou

complexes, cependant toujours fermées. Ici nous avons dû l'adapter pour analyser des tâches de modélisation. Nous avons donc proposé une adaptation utilisant le cycle de modélisation comme structure de l'analyse. Nous avons discuté des points qui restent à approfondir dans la section 7.5.1 (chapitre 7).

Enfin, toujours d'une point de vue méthodologique, nous avons développé ce que nous appelons une ingénierie didactique collaborative, qui est une adaptation de la méthodologie de type ingénierie didactique prenant en compte un travail collaboratif entre chercheur et enseignant.

9.4 Prolongements et perspectives

9.4.1 Autour de l'ingénierie

Comme nous l'avons dit dans le chapitre 8, un deuxième enseignant, collègue de Marie, a lui aussi expérimenté de son côté les problèmes de la rencontre et du volcan Aso pour introduire la séquence articulant probabilités et calcul intégral. Nous avons récupéré tous les documents de cet enseignant Laurent ainsi que les notes de plusieurs élèves de sa classe de l'année 2014-2015, ainsi que de l'année 2015-2016. De même, pour Marie, en plus des données de l'expérimentation, nous disposons des documents distribués ainsi que des notes de plusieurs élèves pour l'année 2015-2016. Ces données pourraient permettre une première analyse pour mesurer la stabilité et la variabilité dans le déroulement des deux problèmes pour un même enseignant entre les deux années, et ensuite entre les deux enseignants. Le fait que les deux enseignants n'aient pas eu le même contact avec nous, chercheuse, peut aussi influencer l'adaptation que chacun en fait dans sa classe, ce qui pourrait être intéressant de repérer.

Cependant, ces données ne pourront pas nous permettre une analyse aussi fine que celle proposée dans le chapitre 7, du fait que nous ne disposons pas des enregistrements audio des séances. Nous avons pour projet de retourner dans la classe de Marie et aller dans celle de Laurent observer et enregistrer les séances relatives aux deux problèmes pendant l'année scolaire 2016-2017. Nous disposerons alors des mêmes types de données. Nous pourrons, de ce fait, utiliser la même méthodologie ; ce qui nous permettra de faire des comparaisons. Nous pourrons nous questionner sur l'évolution de ces tâches au bout de la troisième année en classe et, de façon précise, nous interroger sur la stabilité et la variabilité des déroulements dans les différentes classes. Ces analyses nous permettront aussi de discuter de la reproductibilité.

Une des autres perspectives de ce travail est de le diffuser auprès des enseignants. Nous nous interrogeons cependant sur la manière de le faire. La qualité d'une ressource n'implique pas nécessairement la qualité de ses usages. Pour cela, il est important de prendre en compte des éléments d'utilité, d'utilisabilité et d'acceptabilité dans la conception d'une ressource (Tempier, 2013). En considérant qu'une ressource s'enrichit à travers ses usages (Trouche, 2005), nous pourrions envisager de travailler à l'élaboration de cette ressource dans le cadre d'un groupe IREM avec la collaboration de plusieurs enseignants. Cela nous permettrait notamment de voir la mise en place de ces problèmes par d'autres enseignants et donc d'améliorer la ressource.

Pour poursuivre notre travail, nous pouvons aussi nous poser la question de l'enseignement des probabilités continues dans l'enseignement supérieur. Sans aller jusqu'à la théorie de la mesure, beaucoup de filières, scientifiques ou non, abordent ces contenus dans le cadre de l'intégrale de Riemman, comme par exemple les Classes Préparatoires aux Grandes Écoles (CPGE) HEC ou encore les premières années de licence de mathématiques de certaines universités. Nous pourrions aussi envisager une adaptation des deux problèmes de modélisation dans l'enseignement supérieur, en prenant en compte le fait que les étudiants connaissent le calcul intégral et aussi que les logiciels utilisés ne sont pas les mêmes.

9.4.2 Sur les connaissances des élèves

Du côté de l'apprentissage des élèves, nous disposons aussi de plus de 200 questionnaires remplis par des élèves de terminale S en fin d'année scolaire et des étudiants de première année de CPGE scientifique en tout début d'année. Ce questionnaire vise à tester la capacité des élèves à faire des changements de sous-domaines entre calcul intégral et probabilités à densité à travers des tâches non habituelles en terminale S. Ces données n'ont pas encore été traitées, mais elles offrent des opportunités d'approfondissement. Nous pourrons aussi essayer de voir si l'on peut repérer des différences entre les élèves de Marie et ceux d'autres enseignants par exemple.

9.4.3 Vers de nouvelles notions à étudier

Les probabilités et plus spécifiquement les lois à densité, thème qui nous intéresse tout particulièrement, offrent de nombreuses perspectives de recherche. Nous pensons par exemple à l'introduction de la loi normale par le théorème de Moivre Laplace en terminale. Nous avons déjà commencé à réfléchir sur cette question (Derouet & Parzysz, 2016b). D'autres prolongements seraient d'étudier la transition entre la loi normale et la statistique inférentielle à ce niveau scolaire. Il nous semblerait aussi intéressant d'étudier les conceptions des élèves sur l'égalité $P(X = a) = 0$, bien que l'événement puisse se réaliser. Nous avons déjà pu repérer des éléments au travers de nos transcriptions de séances.

9.4.4 Un questionnement plus théorique

Suite à ce travail, des questions de recherche liées au modèle des ETM et de la théorie de l'activité apparaissent.

Nous pensons, tout d'abord, à la question de la modélisation. Comment étudier ce type de travail mathématique bien spécifique au sein de ces cadres théoriques ? Nous avons proposé quelques premières idées dans ce travail, qui mériteraient d'être approfondies (cf. section 7.5.1). Il nous semblerait intéressant notamment d'identifier des types d'activités des élèves spécifiques à la démarche de modélisation (comme les activités stratégiques que nous avons introduites, par exemple). Nous pouvons aussi nous demander comment considérer l'ensemble de la démarche de modélisation au sein du modèle des ETM. Peut-être faudrait-il considérer des ETM différents suivant l'étape du cycle de modélisation ?

Nous pensons aussi qu'il serait nécessaire de creuser la question autour de l'ETM collectif de la classe. Dans le futur, il faudra peut-être penser les analyses du travail collectif, dans le modèle des ETM, de façon un peu différente. Considérer l'ETM collectif dès qu'un élève prend la parole, comme nous l'avons fait, n'est peut-être pas toujours représentatif du véritable collectif. Il serait intéressant de se demander réellement ce qu'est l'ETM collectif de la classe et ce qui le constitue. Est-ce la réunion des ETM personnels de chaque élève de la classe ? Ou plutôt l'intersection ? Ou, finalement, un « ETM moyen » des ETM de chaque élève ? Nous pencherions plus vers la dernière solution, mais il s'agirait ensuite de définir ce qu'est réellement cet ETM moyen.

De même, il serait important de réfléchir à la place du collectif dans la théorie de l'activité et surtout à approfondir une méthodologie pour analyser les activités du collectif, plutôt que d'assimiler les activités du collectif aux activités des élèves particuliers, comme nous avons pu le faire.

Bibliographie

Armatte, M. (2004). La théorie des erreurs (1750-1820) : enjeux, problématiques, résultats. In E. Barbin & J.-P. Lamarche (Eds.), *Histoires de probabilités et de statistiques* (pp. 141–160). Paris : Ellipses Édition.

Artigue, M. (1988). Ingénierie didactique. *Recherches en Didactique des Mathématiques*, *9*(3), 281–308.

Artigue, M. (2002). Learning mathematics in a CAS environment : the genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning*, *7*, 245–274.

Artigue, M. (2011). L'ingénierie didactique comme thème d'étude. In *En amont et en aval des ingénieries didactiques* (pp. 15–25). Grenoble : La pensée sauvage.

Bansilal, S. (2012). Understanding student engagement with normal distribution problems. *Pythagoras*, *33*(1), 37–49.

Bansilal, S. (2014). Using an APOS framework to understand teachers' responses to questions on the normal distribution. *Statistics Education Research Journal*, *13*(2), 42–57.

Barbazo, E., Bouscasse, J.-M., Pomès, R., Puyou, J., Puyou, M., & Terracher, P.-H. (2003). *L'esprit des lois continues ou Quelques aspects du calcul de probabilités au lycée*. Bordeaux : Irem d'Aquitaine.

Batanero, C., Godino, J., & Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education*, *12*(1).

Batanero, C., Tauber, L., & Meyer, R. (1999). From data analysis to inference : a research project on the teaching of normal distributions. In *Bulletin of the international statistical institute : Proceedings of the fifty-second session of the international statistical institute (tome lviii, book 1)* (pp. 57–58). Helsinki.

Batanero, C., Tauber, L. M., & Sánchez, V. (2004). Students' reasoning about the normal distribution. In B.-Z. D. & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 257–276). Dordrecht : Springer.

Bautier, E. (2006). Le rôle des pratiques des maîtres dans les difficultés scolaires des élèves. Une analyse des pratiques intégrant la dimension des difficultés socialement différenciées. *Recherche et Formation*, *51*, 105–118.

Bednarz, N., Poirier, L., & Desgagné, S. (2001). Conception de séquences d'enseignement en mathématiques : une nécessaire prise en compte des praticiens. In *Le génie didactique. usages et mésusages des théories de l'enseignement* (pp. 43–69). Bruxelles : De Boeck Université.

Blum, W. (1996). Anwendungsbezüge im Mathematikunterricht – Trends und perspektiven. In G. Kadunz (Ed.), *Trends und perspektiven. schriftenreihe didaktik der mathematik (vol. 23)* (pp. 15–38). Vienne : Hölder-Pichler-Tempsky.

Blum, W., & Leiss, D. (2007). How do students and teachers deal with modelling problems. In G.-P. B. W. Haines C. & S. Khan (Eds.), *Mathematical modelling. education, engineering and economics* (pp. 222–231). Chichester : Horwood Publishing.

Bosch, M., & Chevallard, Y. (1999). La sensibilité de l'activité mathématique aux ostensifs : Objet d'étude et problématique. *Recherches en didactique des mathématiques*, *19*(1), 77–123.

Brousseau, G. (1988). Le contrat didactique : le milieu. *Recherches en Didactique des Mathématiques*, *9*(3), 309–336.

Brousseau, G. (1989). Les obstacles épistémologiques et la didactique des mathématiques. In C. G. Nadine Bednarz (Ed.), *Construction des savoirs , Obstacles et Conflits* (p. 41-63). CIRADE Les éditions Agence d'Arc inc.

Brousseau, G. (1998). *Théorie des Situations Didactiques*. La pensée sauvage.

Bru, B. (2006). La courbe de Gauss ou le théorème de Bernoulli raconté aux enfants. *Mathématiques & sciences humaines. Mathematics & social sciences*, *175*, 5–23.

Burgermeister, P., & Dorier, J. (2013). La modélisation dans l'enseignement des mathématiques en Suisse romande. *Petit x*, *91*, 5–24.

Burrill, G., & Biehler, R. (2011). Fundamental Statistical Ideas in the School Curriculum and in Training Teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 57–69). Dordrecht : Springer Netherlands.

Cazes, C., & Vandebrouck, F. (2014). Vil Coyote rattrapera-t-il Bip - Bip ? *Repères-IREM*, *95*, 5–22.

Chevallard, Y. (1999). L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en didactique des mathématiques*, *19*(2), 221–265.

Cléro, J. P. (1988). Un instrument de mesure des croyances. La Règle de Bayes. *Histoire & Mesure*, *3*(4), 491–513.

Cohen, S., & Chechile, R. A. (1997). Probability Distributions, Assessment and Instructional Software : Lessons Learned from an Evaluation of Curricular Software. In J. B. Gal, I. & Garfield (Ed.), *The assessment challenge in statistics education* (pp. 253–262). Amsterdam : IOS Press.

Colmez, P. (2012). Faut-il arrêter d'enseigner les statistiques au lycée ? *Images des Mathématiques*. Consulté sur `http://images.math.cnrs.fr/Faut-il-arreter-d-enseigner-les-1321.html`

Coulange, L. (1998). Les problèmes « concrets » à « mettre en équation » dans l'enseignement. *Petit x*, *47*, 33–58.

Dantal, B., & Henry, M. (2003). Variables aléatoires et lois de probabilité : utilisation de l'analyse en probabilités. In *Probabilités au lycée* (APMEP éd., pp. 125–142). Paris : APMEP.

DelMas, R. C., Garfield, J., & Chance, B. (1999). A model of classroom research in action : Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education*, *7*(3).

de Montmollin, M. (1984). *L'intelligence de la tâche*. Berne : Peter Lang.

Derouet, C., & Parzysz, B. (2016a). How can histograms be useful for introducing continuous probability distributions ? *ZDM - Mathematics Education*, *48*(6).

Derouet, C., & Parzysz, B. (2016b). Le théorème de Moivre-Laplace : une illustration d'un travail centré dans le plan sémiotico-instrumental. In *Symposium espace de travail mathématique etm5.*

Desgagné, S. (1997). Le concept de recherche collaborative : l'idée d'un rapprochement entre chercheurs universitaires et praticiens enseignants. *Revue des sciences de l'éducation*, *23*(2), 371–393.

Desgagné, S., Bednarz, N., Lebuis, P., Poirier, L., & Couture, C. (2001). L'approche collaborative de recherche en éducation : un rapport nouveau à établir entre recherche et formation. *Revue des sciences de l'éducation*, *27*(1), 33–64.

Dorier, J. (2000). Recherches en histoire et en didactique des mathématiques sur l'algèbre linéaire - Perspectives théoriques sur leurs interactions. *Les cahiers du laboratoire Leibniz*, *12*.

Douady, R. (1986). Jeux de cadres et dialectique outil-objet. *Recherches en didactique des mathématiques*, *7*(2), 7–31.

Droesbeke, J.-J., & Tassi, P. (1990). *Histoire de la statistique*. Paris : Presses Universitaires de France.

Dutarte, P., Kern, C., Nougues, M.-F., Dhers, C., Arbre, D., Delzongle, F., ... Mazo, L. (2003). *Le nouveau programme de statistique et probabilités au lycée. Carnet de stage*. Villetaneuse : Université Paris-Nord.

Duval, R. (1993). Registres de représentation sémiotique et fonctionnement cognitif de la pensée. *Annales de Didactique et de Sciences cognitives*, *5*, 37–65.

Duval, R. (1996). Quel cognitif retenir en didactique des mathématiques ? *Recherches en didactique des mathématiques*, *16*(3), 349–382.

Escoffier, J. (2010). *Probabilités et statistiques pour le CAPES et l'Agrégation interne - 2ème édition*. Paris : Ellipses Édition.

Faye, H. (1928). *Cours d'astronomie et de géodésie de l'école polytechnique*. Paris : Gauthier Villars.

Gal, I. (2002). Adult's statistical literacy. Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1–25.

Garet, O., & Kurtzmann, A. (2011). *De l'intégration aux probabilités*. Paris : Ellipses.

Gattuso, L., & Ottaviani, M. G. (2011). Complementing mathematical thinking and statistical thinking in school mathematics. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 121–132). Netherlands : Springer.

Giaquinto, M. (2005). Mathematical activity. In P. Mancosu (Ed.), *Visualization, explanation and reasoning styles in mathematics* (pp. 75–87). Netherlands : Springer.

Girard, J.-C. (2003). Difficultés et obstacles dans l'enseignement des probabilités. In APMEP (Ed.), *Probabilités au lycée* (pp. 35–48). Paris : APMEP.

Girard, J.-C. (2004). La liaison statistique-probabilités dans l'enseignement. *Repères-IREM*, *57*, 83–91.

Gueudet, G., & Trouche, L. (2010). Des ressources aux documents, travail du professeur et genèses documentaires. In G. Gueudet & L. Trouche (Eds.), *Ressources vives. le travail documentaire des professeurs en mathématiques* (pp. 57–74). Rennes / Lyon : Presses Universitaires de Rennes / INRP.

Guin, D., & Trouche, L. (2002). Mastering by the teacher of the instrumental genesis in CAS environments : necessity of intrumental orchestrations. *Zentralblatt für Didaktik der Mathematik*, *34*(5), 204–211.

Haddad, S. (2012). *L'enseignement de l'intégrale en classe terminale de l'enseignement tunisien* (Thèse de doctorat non publiée). Université Paris Diderot et Université virtuelle de Tunis.

Hawkins, A., Jolliffe, F., & Glickman, L. (1991). *Teaching statistical concepts*. Londres : Longman.

Hawkins, T. (2001). *Lebesgue's theory of integration : its origins and development* (Vol. 282). Madison : American Mathematical Soc.

Henry, M. (1999). L'introduction des probabilités au lycée : un processus de modélisation comparable à celui de la géométrie. *Repères-IREM*, *36*, 15–34.

Henry, M. (2001a). *Autour de la modélisation en probabilités*. Besançon : Presses Universitaires Franc-Comtoises.

Henry, M. (2001b). Notion de modèle et modélisation dans l'enseignement. In C2I Statistique et probabilités (Ed.), *Autour de la modélisation en probabilités* (pp. 149–159). Besançon : Presses Universitaires Franc-Comtoises.

Henry, M. (2003). Des lois de probabilité continues en terminale S, pourquoi et pour quoi faire ? *Repères-IREM*, *51*, 5–25.

Henry, M. (2004). La démonstration par Jacques Bernoulli de son théorème. In É. Barbin & J.-P. Lamarche (Eds.), *Histoires de probabilités et de statistiques* (pp. 121–140). Paris : Ellipses Édition.

Henry, M. (2010). Evolution de l'enseignement secondaire français en statistique et probabilités, contributions des IREM à la formation continue des enseignants. In *Statistique et enseignement* (Vol. 1, pp. 35–45). Société Française de Statistique.

Holmes, P. (1980). *Statistics in your world*. Slough : Foulsham Educational.

Houdement, C., & Kuzniak, A. (2006). Paradigmes géométriques et enseignement de la géométrie. *Annales de Didactique et de Sciences cognitives*, *11*, 175–193.

IREM de Besançon. Groupe Probabilités et Statistique. (2007). *Lois continues, test d'adéquation. Une approche pour non spécialiste*. Besançon : Presses universitaires de Franche-Comté.

Jacob, J., & Protter, P. (2003). *L'essentiel en théorie des probabilités*. Paris : Cassini.

Kaiser, G. (1995). Realitätsbezüge im Mathematikunterricht – Ein Überblick über die aktuelle und historische Diskussion. In G. Graumann (Ed.), *Materialien für einen realitätsbezogenen mathematikunterricht* (pp. 66–84). Bad Salzdetfurth : Franzbecke.

Kaiser, G. (2005). Mathematical modelling in school – Examples and experiences. In G. Kaiser & H.-W. Henn (Eds.), *Mathematikunterricht im spannungsfeld von evolution und evolution* (pp. 99–108). Hildesheim : Franzbecke.

Kolmogorov, A. N. (1933). *Foundations of Probability*. Berlin.

Kuzniak, A. (2011). L'espace de Travail Mathématique et ses genèses. *Annales de didactique et de sciences cognitives*, *16*, 9–24.

Kuzniak, A. (2014). Travail mathématique et domaines mathématiques. *Revista Latinoamericana de Investigación en Matemática Educativa*, *Especial 2*(Tome II), 385–400.

Kuzniak, A., Nechache, A., & Drouhard, J.-P. (2016). Understanding the development of mathematical work in the context of the classroom. *ZDM - Mathematics Education*, *48*(6).

Kuzniak, A., & Richard, P. R. (2014). Espaces de travail mathématique. Points de vue et perspectives. *Revista Latinoamericana de Investigación en Matemática Educativa*, *Especial 2*(Tome I), 29–40.

Kuzniak, A., & Vivier, L. (2011). La modélisation dans l'enseignement des mathématiques. Mise en perspective critique. *Cahiers du laboratoire de didactique André Revuz*, *3*.

Lanier, D. (1992). *Huygens : l'espérance et l'infini*. Brest : IREM de Brest.

Lautier, F. (1999). *Ergotopiques. Sur les espaces des lieux de travail*. Toulouse : Octarès.

Lefort, X., & Boyé, A. (1996). De Cassini à Gauss : du calcul d'erreurs aux probabilités. In *Contribution à une approche historique de l'enseignement des mathématiques* (pp. 239–258). Besançon : IREM de Franche-Comté.

Leontiev, A. N. (1984). *Activité, conscience, personnalité*. Moscou : Editions du Progrès.

Liagre, J. (1852). *Calcul des probabilités et théorie des erreurs*. Alexandre Jamar.

Métin, F. (2000). Buffon et le problème de l'aiguille : un Mémoire sur le jeu de Franc-Carreau de 1733. In É. Barbin & J.-P. Le Goff (Eds.), *Si le nombre m'était conté...* (pp. 175–192). Paris : Ellipses.

Ministère de la Jeunesse de l'Éducation nationale et de la Recherche (MJENR). (2002). *Mathématiques. Classe terminale. Série économique et sociale. Série scientifique.* Paris : Centre national de documentation pédagogique.

Ministère de l'Éducation nationale de la Jeunesse et de la Vie associative (MENJVA), & DGESCO. (2012). *Ressources pour la classe terminale générale et technologique. Probabilités et statistique.*

Ministère de l'Education nationale de l'Enseignement supérieur et de la Recherche (MENESR), & DGESCO. (2014). *Ressources pour la classe de terminale générale et technologique. Exercices de mathématiques. Classes de terminale S, ES, STI2D, STMG.*

Ministère de l'Education nationale (MEN). (2001). Programme de l'enseignement des mathématiques en classe terminale de la série scientifique. *Le B.O.*, *Hors-série*.

Ministère de l'Education nationale (MEN). (2008). Programmes du collège. Programmes de l'enseignement de mathématiques. *Le BO. Bulletin officiel spécial*, *6*, 1–38.

Ministère de l'Education nationale (MEN). (2011). Programme de l'enseignement spécifique et de spécialité de mathématiques. Classe terminale de la série scientifique. *Le B.O.*, *Bulletin o*.

Ministère de l'Education nationale (MEN), & DGESCO. (2007). *Ressources pour les classes de 6e, 5e, 4e, et 3e du collège. Organisation et gestion des données au collège.*

Ministère de l'Education nationale (MEN), & DGESCO. (2009). *Ressources pour la classe de seconde. Probabilités et statistique.*

Ministère de l'Education nationale (MEN), & Direction de l'Enseignement Scolaire. (2001). *Mathématiques. Classe de première des séries générales.* Paris : Centre national de documentation pédagogique.

Montoya Delgadillo, E., & Vivier, L. (2014). Les changements de domaine dans le cadre des Espaces de Travail Mathématiques. *Annales de Didactique et de Sciences cognitives*, *19*, 73–101.

Nechache, A. (2015). Comparaison de la démarche de validation dans les espaces de travail idoines en géométrie et en probabilités. In M. I. Gómez-Chacón, J. Escribano, A. Kuzniak, & P. Richard (Eds.), *Espace de travail mathématique, actes quatrième symposium etm* (pp. 51–68). Madrid : Publicaciones del Instituto de Matemática Interdisciplinar, Universidad Complutense de Madrid.

Nechache, A. (2016). *La validation dans l'enseignement des probabilités au niveau du secondaire* (Thèse de doctorat non publiée). Université Paris Diderot.

Niss, M., Blum, W., & Galbraith, P. (2007). Introduction. In W. Blum, P. Galbraith, H. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education. the 14th icmi study* (pp. 4–32). New-York : Springer.

Nowacki, D., & Milliard, H. (2013). Le jeu du « franc carreau ». *Le Bulletin Vert*, *491*, 687–694.

Ortiz, J. J. (1999). *Significado de conceptos probabilísticos en los textos de Bachillerato* (Thèse de doctorat non publiée). Universidad de Granada.

Ortiz, J. J., Cañizares, M. J., Batanero, C., & Serrano, L. (2002). An experimental study of probabilistic language in secondary school textbooks. In *Contributed paper to the international conference on teaching statistics 6 (vol. 6).*

Ouvrard, J.-Y. (2009). *Probabilités 2. Master - Agrégation.* Paris : Cassini.

Pariès, M., Robert, A., & Rogalski, J. (2009). Comment l'enseignant de mathématiques, en classe, met ses élèves sur le chemin des connaissances : un point de vue méthodologique en didactique des mathématiques. *Travail et Apprentissages*, *3*, 95–123.

Parzysz, B. (2003). L'enseignement de la statistique et des probabilités en France : évolution au cours d'une carrière d'enseignant (période 1965-2002). In APMEP (Ed.), *Probabilités au lycée* (pp. 9–34). Paris : APMEP.

Parzysz, B. (2011). Quelques questions didactiques de la statistique et des probabilités. *Annales de Didactique et de Sciences cognitives*, *16*, 127–147.

Parzysz, B. (2013). La longue genèse de la loi normale. *Le Bulletin Vert*, *502*, 29–40.

Parzysz, B. (2014). Espaces de travail en simulation d'expérience aléatoire au lycée : une étude de cas. *Revista Latinoamericana de Investigación en Matemática Educativa*, *17*(4 (I)), 65–82.

Peirce, C. S. (1931). *Collected papers. Vols 1-6.* Cambridge : Harvard University Press.

Perrin, D. (2015). Remarques sur l'enseignement des probabilités et de la statistique au lycée. *Statistique et Enseignement*, *6*(1), 51–63.

Perrin-Glorian, M.-J. (2011). L'ingénierie didactique à l'interface de la recherche avec l'enseignement. Développement des ressources et formation des enseignants. In *En amont et en aval des ingénieries didactiques* (pp. 57–78). Grenoble : La pensée sauvage.

Pfannkuch, M., & Reading, C. (2006). Reasoning about distribution : A complex process. *Statistics Education Research Journal*, *5*(2), 4–9.

Piaget, J. (1971). *Le jugement et le raisonnement chez l'enfant.* Neuchâtel.

Piaget, J. (1985). *The equilibration of cognitive structures : The central problem of intellectual development.* Chicago : University of Chicago Press.

Pichard, J.-F. (2005). Théorie des erreurs, courbes en cloche et normalité. In B. Chaput & M. Henry (Eds.), *Statistique au lycée. volume 1 : Les outils de la statistique* (pp. 219–246). Paris.

Rabardel, P. (1995). *Les hommes et les technologies. Une approche cognitive des instruments contemporains.* Paris : Arman Colin.

Raoult, J.-P., & Arnoux, P. (2013). Pourquoi enseigner les probabilités et la statistique dans les cours de mathématiques. *Images des Mathématiques*. Consulté sur `http://images.math.cnrs.fr/+Pourquoi-enseigner-les+`

Reading, C., & Canada, D. (2011). Teachers' knowledge of distribution. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics : Challenges for teaching and teacher education* (pp. 223–234). Dordrecht : Springer.

Régnier, J.-C. (2002). A propos de la formation en statistique. Approches praxéologiques et épistémologiques de questions du champ de la didactique de la statistique. *Revue du Centre de Recherche en Education*, *22-23*, 157–201.

Régnier, J.-C. (1998). Histogramme. In J.-C. Girard, D. Gros, P. Planchette, J.-C. Régnier, & R. Thomas (Eds.), *Enseigner la statistique du cm à la seconde pourquoi ? comment ?* (pp. 21–42). Villeurbanne : IREM de Lyon.

Richard, P. R. (2004). L'inférence figurale : un pas de raisonnement discursivo-graphique. *Educational studies in Mathematics*, *57*(2), 229–263.

Robert, A. (1998). Outils d'analyse des contenus mathématiques à enseigner au lycée et à l'université. *Recherches en didactique des mathématiques*, *18*(2), 139–189.

Robert, A. (2005). De recherches sur les pratiques aux formations d'enseignants de mathématiques du second degré : un point de vue didactique. *Annales de Didactique et de Sciences cognitives*, *10*, 209 – 249.

Robert, A. (2007). Stabilité des pratiques des enseignants de mathématiques (second degré) : une hypothèse des inférences en formation. *Recherches en Didactique des Mathématiques2*, *27*(3), 271–312.

Robert, A. (2008a). La double approche didactique et ergonomique pour l'analyse des pratiques d'enseignants de mathématiques. In F. Vandebrouck (Ed.), *La classe de mathématiques : activités des élèves et pratiques des enseignants* (pp. 59–68). Toulouse : Octarès Editions.

Robert, A. (2008b). Le cadre général de nos recherches en didactique des mathématiques. In F. Vandebrouck (Ed.), *La classe de mathématiques : activités des élèves et pratiques des enseignants* (pp. 11–22). Toulouse : Octarès Editions.

Robert, A. (2008c). Sur les apprentissages des élèves : une problématique inscrite dans les théories de l'activité et du développement. In F. Vandebrouck (Ed.), *La classe de mathématiques : activités des élèves et pratiques des enseignants* (pp. 33–43). Toulouse : Octarès Editions.

Robert, A. (2008d). Une méthodologie pour analyser les activités (possibles) des élèves en classe. In F. Vandebrouck (Ed.), *La classe de mathématiques : activités des élèves et pratiques des enseignants* (pp. 45–58). Toulouse : Octarès Editions.

Robert, A., & Rogalski, J. (2002). Le système complexe et cohérent des pratiques des enseignants de mathématiques : une double approche. *Canadian Journal of Science, Mathematics and Technology Education*, *2*(4), 505–528.

Robert, A., & Vandebrouck, F. (2014). Proximités-en-actes mises en jeu en classe par les enseignants du secondaire et ZPD des élèves : analyses de séances sur des tâches complexes. *Cahiers du laboratoire de didactique André Revuz*, *10*.

Robinet, J. (1984). *Ingénierie didactique de l'élémentaire au supérieur* (Thèse de doctorat non publiée). Université Paris VII.

Roditi, E. (2009). L'histogramme : à la recherche du savoir à enseigner. *Spirale-Revue de recherches en éducation*, *43*, 129–138.

Rogalski, J. (2008). Le cadre général de la théorie de l'activité. Une perspective de psychologie ergonomique. In F. Vandebrouck (Ed.), *La classe de mathématiques : activités des élèves et pratiques des enseignants* (pp. 23–30). Toulouse : Octarès Editions.

Schneider, M. (1988). *Des objets mentaux « aire » et « volume » au calcul des primitives* (Thèse de doctorat non publiée). Université catholique de Louvain.

Sensevy, G. (2011). *Le sens du savoir. Éléments pour une théorie de l'action conjointe en didactique.* Bruxelles : Groupe De Boeck s.a.

Stigler, S. M. (1986). *The history of statistics : The measurement of uncertainty before 1900*. Harvard University Press.

Tempier, F. (2013). *La numération décimale de position à l'école primaire. Une ingénierie didactique pour le développement d'une ressource* (Thèse de doctorat non publiée). Université Paris Diderot.

Thurston, W. P. (1995). On proof and progress in Mathematics. *For the Learning of Mathematics*, *15*(1), 29–35.

Tran Luong, C. K. (2006). *La notion d'intégrale dans l'enseignement des mathématiques au lycée : une étude comparative entre la France et le Vietnam* (Thèse de doctorat non publiée). Université Joseph Fourier, Grenoble.

Troesch, A. (2007). *Cours de probabilités. Classe préparatoire HEC, option scientifique. Première année.* Consulté sur `http://alain.troesch.free.fr/2006/Fichiers/cours{_}proba.pdf`

Troesch, A. (2012). *Cours de probabilités, ECS deuxième année.* Consulté sur `http://alain.troesch.free.fr/2011/Fichiers/coursproba.pdf`

Trouche, L. (2005). Des artefacts aux instruments, une approche pour guider et intégrer les usages des outils de calcul dans l'enseignement des mathématiques. In *Actes de l'université d'été de st flour : le calcul sous toutes ses formes.*

Vallecillos, A. (1999). Some empirical evidence on learning difficulties about testing hypotheses. In *Bulletin of the international statistical institute : Proceedings of the fifty-second session of the international statistical institute* (Vol. 58, pp. 201–204).

Vandebrouck, F. (2008). *La classe de mathématiques : activités des élèves et pratiques des enseignants.* Toulouse : Octarès Editions.

Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review*, *63*(1), 3–23.

Vergnaud, G. (1991). Langage et pensée dans l'apprentissage des mathématiques. *Revue Française de Pédagogie*, *96*, 79–86.

Vygotsky, L. (1997). *Pensée et langage.* Paris : La dispute.

Wessels, H., & Nieuwoudt, H. (2013). Teachers' reasoning in a repeated sampling context. *Pythagoras*, *34*(1).

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal*, *5*(2), 10–26.

Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 221–248.

Wilensky, U. (1995). Learning probability through building computational models. In D. Carraher & L. Meira (Eds.), *Proceedings of the nineteenth international conference on the psychology of mathematics education* (pp. 152–159). Recife, Brazil.

Wilensky, U. (1997). What is normal anyway ? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, *33*(2), 171–202.

Liste des manuels de terminale S

Barros, J.-M., Bénizeau, P. & Morin, J. (2012). *Transmath. Term S. Enseignement spécifique.* Paris : Nathan.

Cante, S., Feid, M., Lebert, C., Longin, P.-M., Massiera, Y., Meny, J.-M., ... Poncy, M. (2012). *Indice. Maths. Term S. Enseignement spécifique.* Paris : Bordas.

Choquer, A., Hanouch, B., Cocault, M., Simon, D., Joffredo, T., Bruneau, F., ... Ferre, F. (2012). *Maths Repères. Term S. Enseignement spécifique et de spécialité.* Paris : Hachette Éducation.

Chrétien, B., Desrousseaux, P.-A., Duquesnoy, J.-M., Eynard, D., Lambert, F., Lample, H., ... Malaval, J. (2012). *Hyperbole Terminale S spécifique.* Paris : Nathan.

Deschamps, C. (2012). *Symbole Maths Term S.* Paris : Belin.

Le Yaouanq, M.-H. (2012). *Math'x. Term S. Enseignement spécifique.* Paris : Didier.

Sigward, E. (2012). *Odyssée. Mathématiques. Tle S. Enseignement spécifique.* Paris : Hatier.

Talamoni, C., Brun, V., Beltramone, J.-P., Labrosse, J., Truchan, A., Sidokpohou, O., & Merdy, C. (2012). *Déclic Maths Tle S spécifique et spécialité. Livre élève. Format compact (édition 2012).* Paris : Hachette Éducation.

Liste des manuels de seconde

Antibi, A. (2010). *Math 2de. Travailler en confiance.* Paris : Nathan.

Barbazo, E. (2014). *Mathématiques 2de.* Paris : Hachette Education.

Barra, R. (2010). *Transmath. 2de.* Paris : Nathan.

Beltramone, J.-P. (2014). *Déclic mathématiques 2de.* Paris : Hachette Education.

Bouvier, J.-P. (2010). *πxel. Maths. 2de.* Paris : Bordas.

Chesné, J.-F., & Le Yaouanq, M.-H. (2010). *Math'x. 2de.* Paris : Didier.

Choquer-Raoult, A., Cocault, M., Hanouch, B., & Joffrédo, T. (2010). *2de maths Repères.* Paris : Hachette Education.

Deschamps, C. (2010). *Symbole maths 2e.* Paris : Belin.

Gauthier, R., & Poncy, M. (2009). *Indice. Maths. 2de.* Paris : Bordas.

Malaval, J., & Courbon, D. (2010). *Hyperbole. Mathématiques. 2de.* Paris : Nathan.

Sigward, E. (2010). *Odyssée. Mathématiques. 2de.* Paris : Hatier.

Annexes

Table des annexes

A Documents historiques **383**
A.1 Tableau de données expérimentales (Bessel, 1818) 383
A.2 Table de calcul d'intégrales citée dans la figure 3.5 (Liagre, 1852) . . . 384
A.3 Extrait de *Cours de Géodésie*, Faye, 1881 385

B Le programme de terminale S **389**
B.1 Extraits du programme de terminale S 389
B.2 Extraits du programme de terminale S 391

C Le document *Ressources* de terminale sur les probabilités et la statistique **393**
C.1 Justification « visuelle » de la définition de l'espérance d'une variable aléatoire de loi uniforme (pp. 42-43) 394
C.2 Premier exemple d'exercice sur la loi uniforme (p. 43) 395
C.3 Deuxième exemple d'exercice sur la loi uniforme (pp. 43-44) 395

D Les sujets de baccalauréats **396**
D.1 Liste des exercices de baccalauréat traitant de probabilités à densité ou calcul intégral entre 2013 et 2015 396
D.2 Analyse des exercices de calcul intégral 400
D.3 Analyse des exercices de lois à densité 407
D.4 Exercices de baccalauréat 418
D.4.1 Métropole - Réunion – Septembre 2015 418
D.4.2 Polynésie – Septembre 2015 420
D.4.3 Nouvelle-Calédonie – Novembre 2015 422
D.4.4 Métropole – Juin 2015 424
D.4.5 Amérique du Nord – Mai 2013 425
D.4.6 Antilles - Guyane – Juin 2013 426
D.4.7 Liban – Mai 2014 427

E Les manuels **428**
E.1 Cours sur les lois à densité dans les manuels – Généralités 428
E.1.1 Cours du manuel *Déclic* (p. 366) 428
E.1.2 Cours du manuel *Hyperbole* (p. 378) 429
E.1.3 Cours du manuel *Indice* (p. 324) 430
E.1.4 Cours du manuel *Math'x* (p. 402) 431
E.1.5 Cours du manuel *Odyssée* (pp. 394-399) 432

E.1.6 Cours du manuel *Repères* (p. 336) . . . 437
E.1.7 Cours du manuel *Symbole* (p. 390) . . . 438
E.1.8 Cours du manuel *Transmath* (p. 380) . . . 439
E.2 *Activités* d'introduction de la notion de fonction de densité . . . 440
E.2.1 *Activité* 2 pp. 386-387 du manuel *Symbole* . . . 440
E.2.2 *Activité* 1 pp. 398 du manuel *Math'x* . . . 441
E.2.3 *Activité* 3 p. 329 du manuel *Repères* . . . 442
E.3 Exercices . . . 443

F Les entretiens 444
F.1 Les questions de l'entretien . . . 444
F.2 L'entretien avec Marie . . . 446
F.3 Documents de cours de Marie en 2013-2014 . . . 465
F.3.1 Documents relatifs au chapitre « Calcul intégral » . . . 465
F.3.2 Documents relatifs au chapitre « Variables aléatoires continues à densité » . . . 480
F.4 Analyse de l'entretien de l'enseignante Marie . . . 487
F.4.1 Méthodologie relative à l'analyse de contraintes institutionnelles et personnelles . . . 487
F.4.2 Des points qui vont dans le sens de nos résultats précédents . 487
F.4.3 Des contraintes institutionnelles . . . 491
F.4.4 Éléments de la composante cognitive . . . 492
F.4.5 Des difficultés professionnelles relevant de la composante personnelle . . . 492
F.4.6 L'émergence d'un besoin . . . 494

G La séquence conçue 495
G.1 Devoir maison n° 5 traitant d'aire finie d'un domaine infini . . . 495
G.2 Devoir maison n° 6 sur l'histogramme et la loi uniforme . . . 498
G.3 Problème de la rencontre . . . 504
G.4 Problème du volcan Aso . . . 507
G.5 Les exercices . . . 512
G.6 Problème des tailles . . . 519

H Planning de la séquence 524

I Les données relatives à la séquence 526
I.1 Transcription de la séance 1 . . . 526
I.2 Transcription de la séance 2 . . . 561
I.3 Transcription de la séance 3 . . . 568
I.4 Transcription de la séance 4 . . . 595
I.5 Documents complémentaires . . . 620
I.5.1 Notes d'élèves sur le problème de la rencontre et le problème du volcan Aso . . . 620
I.5.2 Méthodes de calcul d'aire . . . 633
I.6 Synthèse donnée à compléter par les élèves . . . 634
I.7 Le cours . . . 636

J Analyses complémentaires **651**
J.1 Questionnaire 651
J.2 Analyse de tâches de l'Exercice 1 658

Annexe A

Documents historiques

A.1 Tableau de données expérimentales (Bessel, 1818)

Range (seconds)	Frequency of errors: Observed	Calculated	Recalculated
	A. 300 Observations of declinations		
0.0–0.4	66	65	64.63
0.4–0.8	58	60	60.01
0.8–1.2	55	53	51.72
1.2–1.6	28	41	41.39
1.6–2.0	27	30	30.75
2.0–2.4	23	21	21.21
2.4–2.8	10	13	13.58
2.8–3.2	15	8	8.08
3.2–3.6	8	5	4.46
3.6–4.0	4	2	2.28
Above 4.0	6	2	—
	B. 300 Observations of right ascensions		
0.0–0.1	114	107	107.22
0.1–0.2	84	87	86.80
0.2–0.3	53	57	56.87
0.3–0.4	24	30	30.14
0.4–0.5	14	13	12.91
0.5–0.6	6	5	4.47
0.6–0.7	3	1	1.25
0.7–0.8	1	0	0.28
0.8–0.9	1	0	0.05
	C. 470 Observations of right ascensions		
0.0–0.1	94	95	94.39
0.1–0.2	88	88	88.50
0.2–0.3	78	78	77.81
0.3–0.4	58	64	64.15
0.4–0.5	51	49	49.58
0.5–0.6	36	35	35.94
0.6–0.7	26	24	24.42
0.7–0.8	14	16	15.56
0.8–0.9	10	9	9.30
0.9–1.0	7	5	5.21
Above 1.0	8	5	—

FIGURE A.1 – Comparaisons des distributions de valeurs absolues d'erreurs résiduelles de trois groupes d'observations, par Bessel (Bessel, 1818, pp. 19-20, dans Stigler, 1986, p. 204)

A.2 Table de calcul d'intégrales citée dans la figure 3.5 (Liagre, 1852)

Table n° 3.

VALEURS DE L'INTÉGRALE DÉFINIE $P_z = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$, POUR DES VALEURS DE t EXPRIMÉES EN FONCTION DE ρ PRIS POUR UNITÉ.

$\frac{t}{\rho}$	$\frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt.$	Différences	$\frac{t}{\rho}$	$\frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt.$	Différences
0,0	0,000	54	2,5	0,908	13
0,1	0,054	53	2,6	0,921	10
0,2	0,107	53	2,7	0,931	10
0,3	0,160	53	2,8	0,941	9
0,4	0,213	51	2,9	0,950	7
0,5	0,264	50	3,0	0,957	6
0,6	0,314	49	3,1	0,963	6
0,7	0,363	48	3,2	0,969	5
0,8	0,411	45	3,3	0,974	4
0,9	0,456	44	3,4	0,978	4
1,0	0,500	42	3,5	0,982	3
1,1	0,542	40	3,6	0,985	2
1,2	0,582	37	3,7	0,987	3
1,3	0,619	36	3,8	0,990	1
1,4	0,655	33	3,9	0,991	2
1,5	0,688	31	4,0	0,993	1
1,6	0,719	29	4,1	0,994	1
1,7	0,748	27	4,2	0,995	1
1,8	0,775	25	4,3	0,996	1
1,9	0,800	23	4,4	0,997	1
2,0	0,823	20	4,5	0,998	0
2,1	0,843	19	4,6	0,998	0
2,2	0,862	17	4,7	0,998	1
2,3	0,879	16	4,8	0,999	0
2,4	0,895	13	4,9	0,999	0
2,5	0,908		5,0	0,999	

Cette table est indépendante de la précision des observations : elle donne la probabilité que l'erreur, pour une espèce quelconque d'observations, ne dépasse pas une certaine valeur exprimée en fonction de l'erreur probable.

Elle montre que, sur 1000 erreurs, il en reste 54 au-dessous de 0,1 de l'erreur probable ; 107 au-dessous de 0,2, etc. En d'autres termes, on peut parier 54 contre 946 que l'erreur que l'on commettra, dans une espèce quelconque d'observations, sera moindre que 0,1 de l'erreur probable ; 107 contre 893 qu'elle sera moindre que 0.2 de l'erreur probable, etc.

A.3 Extrait de *Cours de Géodésie*, Faye, 1881

droites. Bradley a observé 470 fois, à la lunette méridienne de l'Observatoire de Greenwich, la différence d'R entre le Soleil et une étoile écliptique. Bessel a réduit avec beaucoup de soin ces 470 observations, dont la moyenne donne la position du point vernal d'une manière très exacte. En comparant cette moyenne avec chaque observation, on obtient 470 écarts, qui représentent à très peu près les erreurs réellement commises. On a obtenu ainsi, pour l'erreur moyenne $\varepsilon_1 = \pm\sqrt{\frac{[\varepsilon\varepsilon]}{470-1}}$, la valeur $\pm 0^s,39$. C'est l'erreur à craindre sur une mesure isolée. Quant à l'erreur à craindre sur la moyenne, elle se réduit, comme nous l'avons vu, à

$$\pm \frac{0^s,39}{\sqrt{470}} = 0^s,017.$$

On voit donc que cette moyenne est à très peu près l'expression définitive, par suite, les 470 écarts ci-dessus mentionnés représentent de très près les erreurs réellement commises.

En examinant le Tableau de ces 470 erreurs, on constate, quant aux signes, qu'elles sont indifféremment et en pareil nombre positives ou négatives; quant à la grandeur, on trouve que :

Entre 0,0 et 0,1 il y a 91 erreurs, tant positives que négatives.

»	0,1	0,2	88 »
»	0,2	0,3	78 »
»	0,3	0,4	58 »
»	0,4	0,5	51 »
»	0,5	0,6	36 »
»	0,6	0,7	26 »
»	0,7	0,8	14 »
»	0,8	0,9	10 »
»	0,9	1,0	7 »
Au-dessus de	1,0		8 »

Ainsi, à en juger par cette série, la probabilité qu'une erreur soit comprise entre 0″ et 0″,1, c'est-à-dire le rapport

$$\frac{\text{nombre des erreurs de cet ordre}}{\text{nombre total des erreurs}}$$

est $\frac{91}{470}$, tandis que la probabilité des erreurs plus fortes, c'est-à-dire $\frac{88}{470}$, va en décroissant.

Évidemment ces fractions ou ces probabilités suivent une cer-

p. 271

« Cours de Géodésie »

Faye 1881

Recherche théorique de cette loi.

Les théoriciens ont cherché à l'établir *a priori* au moyen d'une hypothèse sur l'essence des erreurs accidentelles. Supposons, ont-ils dit, qu'il n'y ait pas de distinction essentielle à faire entre celles-là et les systématiques, et que les premières soient dues à des causes constantes, comme les secondes, mais en très grand nombre. Admettons, de plus, que chaque cause produise des erreurs ne pouvant dépasser une très petite quantité α, toujours la même, et affectant indifféremment les deux sens, positif et négatif. Cette hypothèse nous replace dans le cas très particulier que nous avons considéré, page 69, dans l'emploi des tables de logarithmes. Admettons, enfin, que les erreurs accidentelles observées résultent de toutes les combinaisons possibles d'un nombre très grand n de ces petites erreurs élémentaires. Le Calcul des probabilités montre que ces combinaisons seront exprimées par les coefficients du binome $(1-x)^n$, et une analyse fort simple établit que la différence entre deux coefficients successifs devient, en passant à la limite où $n = \infty$, la dérivée de la loi de probabilité des erreurs qu'une simple intégration fera dès lors connaître. Nous nous bornerons à signaler cette conception *a priori* et nous considérerons les choses d'un nouveau point de vue. Nous allons appliquer ici la méthode expérimentale, c'est-à-dire chercher, par expérience, s'il existe une relation entre la probabilité d'une erreur et la grandeur de cette erreur, puis opérant par des mesures réellement faites, sur des écarts dûment constatés, nous tâcherons de découvrir graphiquement cette loi.

Recherche empirique de la loi de probabilité des erreurs fortuites.

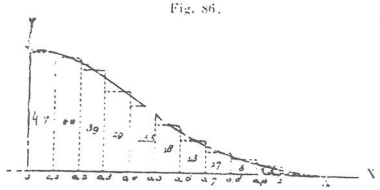
Prenons, en premier lieu, des observations astronomiques. Pour déterminer la position du point vernal, origine des ascensions

taine loi : cherchons-la empiriquement à l'aide d'une construction graphique.

Si l'on choisit un système d'axes des x et des y, et que l'on porte les erreurs sur l'axe des x, on ne saurait porter les probabilités correspondantes sur l'axe des y, car la probabilité d'une erreur x déterminée, c'est-à-dire comprise entre x et $x+dx$, doit être infiniment petite; elle n'est finie que si les limites de l'erreur diffèrent elles-mêmes d'une quantité finie. On est donc conduit à chercher graphiquement une courbe $y=f(x)$ telle que les probabilités ci-dessus soient représentées par les aires comprises entre les couples d'ordonnées qui répondent aux abscisses 0^s et $0^s,1$, $0^s,1$ et $0^s,2$, ... Comme il y a autant d'erreurs négatives que d'erreurs positives, cette courbe devra être symétrique de part et d'autre de l'axe des y; en second lieu, entre les ordonnées relatives à $x=0$ et $x=+0^s,1$, l'aire devra être la même qu'entre les ordonnées de $x=0$ et $x=-0^s,1$, c'est-à-dire $\frac{94}{2}$ ou 47. On obtiendra cette courbe en divisant l'axe des x en parties proportionnelles à $0^s,1$, $0^s,2$, $0^s,3$, ... et en construisant sur les intervalles de 0^s à $0^s,1$, de $0^s,1$ à $0^s,2$, ... des rectangles proportionnels aux nombres 47, 44, ..., puis en traçant un trait continu qui laisse, en dedans, de petites aires triangulaires équivalentes à celles qu'il laissera en dehors.

La courbe suivante, dont nous ne donnons que la moitié située

Fig. 86.



du côté des x positifs, représentera, par ses aires, la loi de probabilité des erreurs comptées sur les abscisses.

Elle se rapproche évidemment du genre de celles qui ont pour équation $y=ae^{-h^2x^2}$, a et h étant des paramètres à déterminer. Si, en effet, on les détermine par deux de ces points, on verra que la

courbe $y=47,365\,e^{-(1,764)^2x^2}$ passe à peu près par tous les autres.

La seule différence appréciable consiste en ce que la courbe représentée par cette équation devrait atteindre l'axe des x, comme le fait la courbe que nous venons de tracer, en deux points symétriques par rapport à l'origine o, dont les abscisses $x=\pm k$ représenteraient les erreurs limites, impossibles à commettre dans ce genre d'observations. Au lieu de cela, la courbe

$$Y=47,365\,e^{-(1,764)^2x^2}$$

est en réalité asymptote à l'axe des x. Mais les lignes de cet ordre s'approchent si rapidement de l'axe des x, qu'au delà de l'abscisse k leur aire devient entièrement négligeable, comme on le verra plus tard.

Nous regarderons donc comme empiriquement établi avec toute l'exactitude désirable, que, pour le genre d'observations ici considéré, la probabilité P d'une erreur comprise entre les limites x et $x+\Delta x$, ou, en d'autres termes, le rapport

$$\frac{\text{aire de la courbe comprise entre les ordonnées correspondantes}}{\text{aire totale de la courbe}},$$

aura pour expression

$$P=\frac{\int_x^{x+\Delta x} ae^{-h^2x^2}dx}{\int_{-\infty}^{+\infty} ae^{-h^2x^2}dx},$$

car, d'après ce qui a été dit tout à l'heure, il est permis de remplacer, sans erreur sensible,

$$\int_{-k}^{+k} ae^{-h^2x^2}dx \quad \text{par} \quad \int_{-\infty}^{+\infty} ae^{-h^2x^2}dx.$$

Or, on sait que l'intégrale définie

$$\int_{-\infty}^{+\infty} ae^{-h^2x^2}dx=\frac{a\sqrt{\pi}}{h};$$

donc on a finalement

$$P=\frac{h}{\sqrt{\pi}}\int_x^{x+\Delta x} e^{-h^2x^2}dx,$$

et la probabilité infiniment petite de l'erreur déterminée x (c'est-à-dire comprise entre x et $x + dx$) sera

$$\frac{h}{\sqrt{\pi}} e^{-h^2 x^2}\, dx.$$

Telle est, d'après les excellentes observations de Bradley, la loi de probabilité des erreurs purement accidentelles, c'est-à-dire dégagées de toute erreur systématique. Et nous sommes ici en droit de considérer les choses ainsi, car, dans le système d'observations de Bradley, les erreurs systématiques dues à l'instrument étant sensiblement les mêmes pour l'étoile et pour le Soleil, ces erreurs disparaissent de la différence conclue des R.

Mais il faut voir encore si, dans un genre de mesure tout différent, nous retrouverons la même loi de probabilité des erreurs accidentelles.

Le général Didion donne, dans son *Traité des probabilités appliquées au tir*, la série suivante d'épreuves faites avec un pistolet de cavalerie. Ce sont les écarts des balles tirées sur une cible où nous supposerons tracée une ligne verticale. Ces écarts sont donc mesurés dans le sens horizontal. L'arme employée n'ayant pas de déviation constante et les coups ayant été tirés avec le même soin, ces écarts, soit à droite, soit à gauche de la cible, doivent être considérés comme des erreurs accidentelles. En voici le Tableau :

cm	cm	cm	cm	cm	cm	cm	cm	cm
— 5	— 10	— 5	— 31	+ 8	— 14	+ 8	0	— 23
— 5	— 13	+ 29	— 25	— 22	— 27	+ 16	— 3	— 26
— 29	— 22	— 40	0	— 15	— 13	+ 8	— 31	+ 9
— 7	— 25	— 7	— 7	— 10	— 28	— 2	— 12	+ 19
— 4	— 1	— 13	— 27	— 2	+ 28	— 37	— 7	+ 36
+ 11	+ 18	0	— 17	— 33	+ 3	— 2	— 23	— 2
+ 6	— 3	— 4	— 10	— 1	— 28	— 2	— 37	— 10
— 6	+ 3	— 3	— 2	— 14	0	— 19	— 18	— 11
— 24	— 25	— 27	— 11	— 12	— 21	— 9	— 1	+ 17
— 17	+ 5	— 12	— 13	— 16	— 1	— 2	0	— 8

En faisant la somme des carrés des écarts, divisant par 90 — 1 et extrayant la racine carrée, on trouve l'erreur moyenne d'un coup

$$\varepsilon_1 = \pm 18^{cm},3.$$

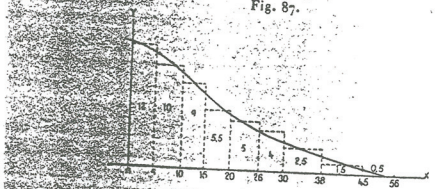
Il y a d'ailleurs 42 erreurs positives, 43 négatives et 5 erreurs

nulles, indifféremment positives ou négatives. Si l'on range ces écarts par ordre de grandeur, sans faire acception des signes, on trouve

Entre 0 cm et 5 cm	24	écarts positifs ou négatifs.	
» 5 » 10	20	»	»
» 10 » 15	18	»	»
» 15 » 21	11	»	»
» 21 » 26	10	»	»
» 26 » 31	8	»	»
» 31 » 38	5	»	»
» 38 » 45	3	»	»
Et au delà de 56	1	»	»

En traçant une courbe dont les abscisses seront 0, 5, 10, 15, ... et telle que les aires comprises entre les ordonnées correspondantes soient proportionnelles aux nombres 12, 10, 9, $5\frac{1}{2}$, 5, 4, $2\frac{1}{2}$, $1\frac{1}{2}$, $\frac{1}{2}$, on obtient la figure 87.

Fig. 87.



Cette courbe est représentée de très près par l'équation

$$y = 11,31\, e^{-(0,0369)^2 x^2}.$$

Elle est donc de même espèce que la précédente.

On trouvera le même résultat dans tous les genres de mesures ou d'observations, pourvu qu'elles ne soient pas affectées d'erreurs systématiques. La Statistique en fournit de nombreux exemples. C'est ainsi que les anomalies accidentelles de la taille humaine suivent la même loi de probabilité; on s'en est assuré aux États-Unis, à l'occasion de la guerre de sécession, en examinant les documents statistiques qu'on y a recueillis sur la taille des contingents militaires. En voici le Tableau :

Sur 10000 *hommes.*

Taille.	Nombre réel.	Formule.
pouces		
61.....	105	100
62.....	169	171
63.....	369	368
64.....	686	675
65.....	1044	1051
66.....	1391	1399
67.....	1584	1584
68.....	1607	1531
69.....	1218	1260
70.....	852	881
71.....	467	531
72.....	277	267
73.....	139	118
74.....	92	61
.....	....	....

La taille moyenne, $67^{\text{pouces}}, 24$, semble être le type que la nature cherche à réaliser dans cette région, mais qu'elle n'atteint qu'avec des écarts attribuables à des causes multiples et purement accidentelles.

Nous considérerons donc l'intégrale

$$\frac{h}{\sqrt{\pi}} \int_{-x}^{+x} e^{-h^2x^2}\,dx$$

comme représentant très sensiblement la probabilité d'une erreur accidentelle comprise entre $-x$ et $+x$ dans tout sytème de mesure ou d'observation dont le degré de précision sera défini par le paramètre h. C'est la loi de Gauss.

Table de probabilités.

C'est la Table des valeurs de l'intégrale définie que nous venons de réduire en série. Pour calculer aisément cette série, remplaçons hx par son équivalent $h\eta\,\frac{x}{\eta}$, ou par $0,4769\,\frac{x}{\eta}$; puis, prenant pour argument $\frac{x}{\eta}$ et lui donnant successivement pour valeurs 0; 0,1; 0,2; 0,3, ..., formons un tableau des valeurs que prend la série, c'est-à-dire

$$P = \frac{2}{\sqrt{\pi}}\left[0,47\frac{x}{\eta} - \frac{1}{3}\left(0,47\frac{x}{\eta}\right)^3 + \frac{1}{10}\left(0,47\frac{x}{\eta}\right)^5 - \dots\right].$$

Valeurs de $\frac{x}{\eta}$.	P.	Différence.
0,0.....	0,000	
0,1.....	0,053	54
0,2.....	0,107	53
0,3.....	0,160	53
0,4.....	0,213	53
0,5.....	0,264	51
0,6.....	0,314	50
0,7.....	0,363	49
0,8.....	0,411	48
0,9.....	0,456	45
1,0.....	0,500	44
1,1.....	0,542	42
1,2.....	0,582	40
1,3.....	0,619	37
1,4.....	0,655	36
1,5.....	0,688	33
1,6.....	0,719	31
1,7.....	0,748	29
1,8.....	0,775	27
1,9.....	0,800	25
2,0.....	0,823	23
2,1.....	0,843	20
2,2.....	0,862	19
2,3.....	0,879	17
2,4.....	0,895	16
2,5.....	0,908	13

Valeurs de $\frac{x}{\eta}$.	P.	Différence.
2,5.....	0,908	
2,6.....	0,921	13
2,7.....	0,931	10
2,8.....	0,941	10
2,9.....	0,950	9
3,0.....	0,957	7
3,1.....	0,963	6
3,2.....	0,969	6
3,3.....	0,974	5
3,4.....	0,978	4
3,5.....	0,982	4
3,6.....	0,985	3
3,7.....	0,987	2
3,8.....	0,990	3
3,9.....	0,991	1
4,0.....	0,993	2
4,1.....	0,994	1
4,2.....	0,995	1
4,3.....	0,996	1
4,4.....	0,997	1
4,5.....	0,998	1
4,6.....	0,998	0
4,7.....	0,998	0
4,8.....	0,999	1
4,9.....	0,999	0
5,0.....	0,999255	0

Annexe B

Le programme de terminale S

B.1 Extraits du programme de terminale S : *Intégration*

Intégration Définition de l'intégrale d'une fonction continue et positive sur $[a,b]$ comme aire sous la courbe. Notation $\int_a^b f(x)\mathrm{d}x$. Théorème : si f est une fonction continue et positive sur $[a,b]$, la fonction F définie sur $[a,b]$ par $F(x)=\int_a^x f(t)\mathrm{d}t$ est dérivable sur $[a, b]$ et a pour dérivée f.		On s'appuie sur la notion intuitive d'aire rencontrée au collège et sur les propriétés d'additivité et d'invariance par translation et symétrie. On peut mener un calcul approché d'aire (parabole, hyperbole, etc.) pour illustrer cette définition. ▣ Il est intéressant de présenter le principe de la démonstration du théorème dans le cas où f est positive et croissante.

Contenus	Capacités attendues	Commentaires
Primitive d'une fonction continue sur un intervalle.	• Déterminer des primitives des fonctions usuelles par lecture inverse du tableau des dérivées. • Connaître et utiliser les primitives de $u'e^u$, $u'u^n$ (n entier relatif, différent de -1) et, pour u strictement positive, $\frac{u'}{\sqrt{u}}$, $\frac{u'}{u}$.	Une primitive F de la fonction continue et positive f étant connue, on a : $\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$.
Théorème : toute fonction continue sur un intervalle admet des primitives.		▣ Il est intéressant de démontrer ce théorème dans le cas d'un intervalle fermé borné, en admettant que la fonction a un minimum. On admet le cas général. On fait observer que certaines fonctions comme $x \mapsto \exp(-x^2)$ n'ont pas de primitive « explicite ».
Intégrale d'une fonction continue de signe quelconque. Linéarité, positivité, relation de Chasles. Valeur moyenne.	• Calculer une intégrale. • Utiliser le calcul intégral pour déterminer une aire. • Encadrer une intégrale. ◇ Pour une fonction monotone positive, mettre en œuvre un algorithme pour déterminer un encadrement d'une intégrale.	La formule $\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$, établie pour une fonction continue et positive, est étendue au cas d'une fonction continue de signe quelconque. L'intégration par parties n'est pas un attendu du programme. La notion de valeur moyenne est illustrée par des exemples issus d'autres disciplines. ⇆ [SPC] Mouvement uniformément accéléré. ⇆ [SI] Valeur moyenne, valeur efficace dans un transfert énergétique. (AP) *Calcul du volume d'un solide.*

B.2 Extraits du programme de terminale S : *Notion de loi à densité à partir d'exemples*

Notion de loi à densité à partir d'exemples Loi à densité sur un intervalle.		Les exemples étudiés s'appuient sur une expérience aléatoire et un univers associé Ω, muni d'une probabilité. On définit alors une variable aléatoire X, fonction de Ω dans R, qui associe à chaque issue un nombre réel d'un intervalle I de R. On admet que X satisfait aux conditions qui permettent de définir la probabilité de l'événement $\{X \in J\}$ comme aire du domaine : $\{M(x,y) ; x \in J \text{ et } 0 \leq y \leq f(x)\}$ où f désigne la fonction de densité de la loi et J un intervalle inclus dans I. Toute théorie générale des lois à densité et des intégrales sur un intervalle non borné est exclue.

Contenus	Capacités attendues	Commentaires
Loi uniforme sur $[a,b]$. Espérance d'une variable aléatoire suivant une loi uniforme.	• Connaître la fonction de densité de la loi uniforme sur $[a,b]$.	L'instruction « nombre aléatoire » d'un logiciel ou d'une calculatrice permet d'introduire la loi uniforme sur $[0,1]$. La notion d'espérance d'une variable aléatoire à densité f sur $[a,b]$ est introduite à cette occasion par $\mathrm{E}(X) = \int_a^b t f(t)\,\mathrm{d}t$. On note que cette définition constitue un prolongement dans le cadre continu de l'espérance d'une variable aléatoire discrète. (AP) *Méthode de Monte-Carlo.*
Lois exponentielles.	• Calculer une probabilité dans le cadre d'une loi exponentielle.	■ On démontre qu'une variable aléatoire T suivant une loi exponentielle vérifie la propriété de durée de vie sans vieillissement : pour tous réels t et h positifs, $P_{T\geq t}(T \geq t+h) = P(T \geq h)$.
Espérance d'une variable aléatoire suivant une loi exponentielle.	■ Démontrer que l'espérance d'une variable aléatoire suivant une loi exponentielle de paramètre λ est $\frac{1}{\lambda}$.	L'espérance est définie comme la limite quand x tend vers $+\infty$ de $\int_0^x t f(t)\,\mathrm{d}t$ où f est la fonction de densité de la loi exponentielle considérée. Cette partie du programme se prête particulièrement à l'étude de situations concrètes, par exemple sur la radioactivité ou la durée de fonctionnement d'un système non soumis à un phénomène d'usure.

Contenus	Capacités attendues	Commentaires
Loi normale centrée réduite $\mathcal{N}(0,1)$. Théorème de Moivre Laplace (admis).	• Connaître la fonction de densité de la loi normale $\mathcal{N}(0,1)$ et sa représentation graphique. ▣ Démontrer que pour $\alpha \in]0,1[$, il existe un unique réel positif u_α tel que $P(-u_\alpha \leq X \leq u_\alpha) = 1-\alpha$ lorsque X suit la loi normale $\mathcal{N}(0,1)$. • Connaître les valeurs approchées $u_{0,05} \approx 1{,}96$ et $u_{0,01} \approx 2{,}58$.	Pour introduire la loi normale $\mathcal{N}(0,1)$, on s'appuie sur l'observation des représentations graphiques de la loi de la variable aléatoire $Z_n = \frac{X_n - np}{\sqrt{np(1-p)}}$ où X_n suit la loi binomiale $\mathcal{B}(n,p)$, et cela pour de grandes valeurs de n et une valeur de p fixée entre 0 et 1. Le théorème de Moivre Laplace assure que pour tous réels a et b, $P(Z_n \in [a,b])$ tend vers $\int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ lorsque n tend vers $+\infty$. L'espérance d'une variable aléatoire suivant la loi $\mathcal{N}(0,1)$ est définie par $\lim_{x\to-\infty} \int_x^0 t f(t)dt + \lim_{y\to+\infty} \int_0^y t f(t)dt$ où f désigne la densité de cette loi. On peut établir qu'elle vaut 0. On admet que la variance, définie par $E((X - E(X))^2)$, vaut 1.
Loi normale $\mathcal{N}(\mu, \sigma^2)$ d'espérance μ et d'écart-type σ.	• Utiliser une calculatrice ou un tableur pour calculer une probabilité dans le cadre d'une loi normale $\mathcal{N}(\mu, \sigma^2)$.	Une variable aléatoire X suit une loi $\mathcal{N}(\mu, \sigma^2)$ si $\frac{X-\mu}{\sigma}$ suit la loi normale $\mathcal{N}(0,1)$. On fait percevoir l'information apportée par la valeur de l'écart-type. ⇆ [SI et SPC] Mesures physiques sur un système réel en essai.
	• Connaître une valeur approchée de la probabilité des événements suivants : $\{X \in [\mu-\sigma, \mu+\sigma]\}$, $\{X \in [\mu-2\sigma, \mu+2\sigma]\}$ et $\{X \in [\mu-3\sigma, \mu+3\sigma]\}$, lorsque X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$.	La connaissance d'une expression algébrique de la fonction de densité de la loi $\mathcal{N}(\mu, \sigma^2)$ n'est pas un attendu du programme. On illustre ces nouvelles notions par des exemples issus des autres disciplines.

Annexe C

Le document *Ressources* de terminale sur les probabilités et la statistique

(MENJVA & DGESCO, 2012)

C.1 Justification « visuelle » de la définition de l'espérance d'une variable aléatoire de loi uniforme (pp. 42-43)

On peut observer que la définition de l'espérance par la formule $E(X) = \int_a^b xf(x)dx$ prolonge celle de l'espérance d'une variable aléatoire discrète.

En effet, le terme $f(x)dx$ peut s'interpréter comme l'aire d'un rectangle de côtés dx et $f(x)$, fournissant en quelque sorte la probabilité que la variable X prenne la valeur x. Dans ces conditions, l'intégrale $\int_a^b xf(x)dx$ correspond à une « somme » de produits $x \times f(x)dx$.

La figure 1 ci-dessous présente la situation dans le cas où $a = 0$ et $b = 1$.

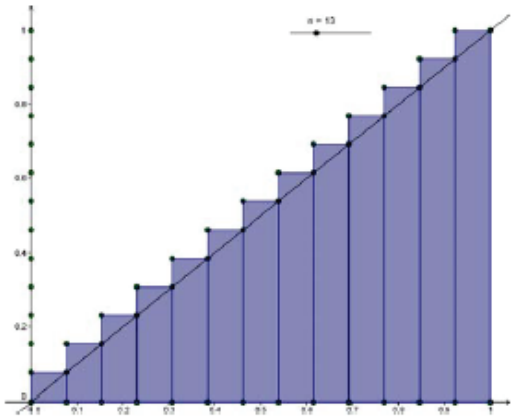


Figure 16

Document associé : espérance d'une variable uniforme.ggb

On a représenté les rectangles de base $\frac{1}{n}$ et de hauteur $\frac{k}{n}$, avec k entier variant de 1 à n.

La somme $S_n = \sum_{k=1}^{n} \frac{k}{n} \frac{1}{n}$ des aires de ces rectangles peut s'interpréter comme l'espérance d'une variable discrète équirépartie prenant les n valeurs $\frac{k}{n}$, pour k variant de 1 à n.

Elle vaut $\frac{n(n+1)}{2n^2}$ et a pour limite $\frac{1}{2}$.

Quand n tend vers l'infini, la somme des aires des rectangles tend vers l'aire située sous la droite d'équation $y = x$. On retrouve ainsi l'égalité $\int_0^1 xf(x)dx = \frac{1}{2}$.

C.2 Premier exemple d'exercice sur la loi uniforme (p. 43)

1. A partir de 7 heures le matin, les bus passent toutes les quinze minutes à un arrêt précis. Un usager se présente à cet arrêt entre 7h et 7h30. On fait l'hypothèse que l'heure exacte de son arrivée à cet arrêt, représentée par le nombre de minutes après 7h, est la variable aléatoire uniformément répartie sur l'intervalle [0, 30].

1) Quelle est la probabilité que l'usager attende moins de cinq minutes le prochain bus ?
2) Quelle est la probabilité qu'il attende plus de dix minutes ?

C.3 Deuxième exemple d'exercice sur la loi uniforme (pp. 43-44)

2. Partie A

Olivier vient tous les matins entre 7h et 7h 45 chez Karine prendre un café.

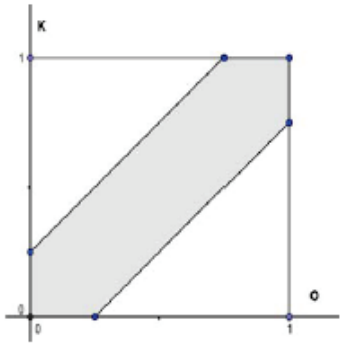
1) Sachant qu'Olivier ne vient jamais en dehors de la plage horaire indiquée et qu'il peut arriver à tout instant avec les mêmes chances, quelle densité peut-on attribuer à la variable aléatoire « heure d'arrivée d'Olivier » ?
2) Calculer la probabilité qu'Olivier sonne chez Karine :
   • Après 7h30 • Avant 7h10 • Entre 7h20 et 7h22 • A 7h30 exactement.

2. Partie B

Olivier et Karine décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Les instants d'arrivée d'Olivier et Karine sont assimilés à des variables aléatoires de loi uniforme sur $[0,1]$. Chacun attend un quart d'heure mais jamais au-delà de 8h. Quelle est la probabilité qu'ils se rencontrent ?

Éléments de solution

Pour la partie B, si on note O la variable aléatoire « instant d'arrivée d'Olivier » et K celle de Karine. La probabilité cherchée est $P(|O-K| \leq \frac{1}{4})$; en utilisant une représentation graphique, cette probabilité est l'aire de la zone grisée ci-dessous, ensemble des points de coordonnées (x, y) du carré tels que $|x-y| \leq 0{,}25$. (On trouve $\frac{7}{16}$).



Annexe D

Les sujets de baccalauréat

D.1 Liste des exercices de baccalauréat traitant de probabilités à densité ou calcul intégral entre 2013 et 2015

Pays/Date	Exercices faisant intervenir des probabilités à densité	Exercices faisant intervenir du calcul intégral
2013		
Pondichéry (avril 2013)	Exercice 4 Loi normale	
Amérique du Nord (mai 2013)	Exercice 3 Loi normale Loi exponentielle	Exercice 4 Aire finie d'un domaine infini
Liban (mai 2013)	Exercice 2 Loi normale	Exercice 3 Calcul d'intégrale, calcul d'aire
Polynésie (juin 2013)	Exercice 3 Loi normale	Exercice 1 Calcul approché/exact d'aire
Antilles-Guyane (juin 2013)	Exercice 2 Loi normale	Exercice 3 Calcul d'aire
Asie (juin 2013)		
Centres étrangers (juin 2013)	Exercice 1 Loi exponentielle Loi normale	Exercice 3 Lien avec l'aire
Métropole (juin 2013)		Exercice 2 Lien avec l'aire
Antilles-Guyane (septembre 2013)	Exercice 3 Loi normale	Exercice 2 Calcul d'aire
Métropole (septembre 2013)		Exercice 1 Lien primitive-dérivée-intégrale Aire approchée
Nouvelle-Calédonie (novembre 2013)	Exercice 3 Loi normale	
Amérique du Sud (novembre 2013)		
Nouvelle-Calédonie (mars 2014)	Exercice 2 Loi normale	Exercice 3 Calcul approché/exact d'aire
Total	9/13[1] Loi uniforme : 0/13 Loi exponentielle : 2/13 Loi normale : 9/13	9/13
2014		
Pondichéry (avril 2014)	Exercice 1 Loi exponentielle	Exercice 3 Calcul d'aires
Liban (mai 2014)	Exercice 1 Loi normale	Exercice 3 Fonction Aire
Amérique du Nord (mai 2014)	Exercice 1 Loi normale	Exercice 2 Fonction Aire
Centres étrangers (juin 2014)	Exercice 1 Loi exponentielle Loi normale	Exercice 3 Calcul d'aire
Polynésie (juin 2014)	Exercice 3 Loi exponentielle	Exercice 4 Calcul d'aire
Antilles-Guyane (juin 2014)	Exercice 1 Loi normale	Exercice 2 Calcul d'aire

[1] Il y a 9 sujets sur 13 qui traitent de questions portant sur les probabilités à densité. Ce décompte ne prend pas en compte le nombre d'exercices du sujet qui éventuellement traitent de ce sous-domaine.

Asie (juin 2014)	Exercice 2 Loi normale	Exercice 3 Calcul d'intégrale Exercice 4 Suite/intégrale, calcul approché d'intégrale
Métropole (juin 2014)	Exercice 2 Loi normale	Exercice 1 Suite/intégrale
Antilles-Guyane (septembre 2014)	Exercice 3 Loi exponentielle Loi normale	
Métropole (septembre 2014)	Exercice 2 Loi exponentielle Loi normale	Exercice 1 Calcul approché d'aire
Amérique du Sud (novembre 2014)	Exercice 1 Loi normale	Exercice 4 Calcul d'aire
Nouvelle-Calédonie (novembre 2014)	Exercice 1 Loi normale	
Nouvelle-Calédonie (mars 2015)	Exercice 2 Loi exponentielle	
Total	13/13 Loi uniforme : 0/13 Loi exponentielle : 6/13 Loi normale : 10/13	10/13
2015		
Pondichéry (avril 2015)	Exercice 3 Loi normale	Exercice 1 Aire fini domaine infini
Liban (mai 2015)		Exercice 2 Suite d'intégrale de suite de fonctions
Amérique du Nord (juin 2015)	Exercice 3 Loi normale	Exercice 4 Primitive, intégrale, interprétation
Centres étrangers (juin 2015)	Exercice 1 Loi normale	Exercice 4 Primitive, aire
Polynésie (juin 2015)	Exercice 3 Loi normale	
Asie (juin 2015)	Exercice 1 Loi exponentielle	Exercice 3 Aire sous une suite de courbes Cn
Antilles-Guyane (juin 2015)	Exercice 2 Loi exponentielle	
Métropole (juin 2015)	Exercice 1 Loi exponentielle Loi normale	Exercice 4 Aire, méthode des trapèzes
Métropole- La Réunion (septembre 2015)	Exercice 1 Loi exponentielle Loi normale	Exercice 2 Aire fini domaine infini (suite) Exercice 4 Primitive et aire
Polynésie (septembre 2015)	Exercice 2 Loi normale	Exercice 1 Aire entre deux courbes
Antilles - Guyane (septembre 2015)		Exercice 1 Suite définie par une intégrale de suite de fonctions
Nouvelle-Calédonie (novembre 2015)	Exercice 1 Loi normale	Exercice 1 Aire
Amérique du Sud (novembre 2015)	Exercice 3 Loi normale	Exercice 1 Aire (fonction négative)

Total	11/13 Loi uniforme : 0/13 Loi exponentielle : 4/13 Loi normale : 9/13	11/13
2013-2014-2015		
TOTAL	33/39 Loi uniforme : 0/39 Loi exponentielle : 12/39 Loi normale : 28/39	30/39

D.2 Analyse des exercices de calcul intégral

Pays/Date	Proportion/ pourcentage de l'exercice traitant du calcul intégral	Contexte	Notions abordées	Registres rencontrés dans l'énoncé	Artefact	Liens possibles avec les probabilités à densité
2013						
Pondichéry (avril 2013)						
Amérique du Nord (mai 2013)	Exercice 4 3/11[1] (27%)	Mathématique	Etude d'une suite définie par l'**aire** sous une courbe entre les droites d'équation $x = a$ (a donné) et $x = n$.	Langue naturelle Analyse	X	Aire finie (égale à e) d'un domaine infini.
Liban (mai 2013)	Exercice 3 2/11 (18%)	Mathématique	Calcul d'intégrale et interprétation. Calcul d'aire entre deux courbes.	Analyse Langue naturelle Graphique	X	
Polynésie (juin 2013)	Exercice 1 4/7 (57%)	Mathématique	Algorithme de la méthode des rectangles Calcul exact/approché d'aire	Analyse Langue naturelle Graphique Pseudo-code	X	
Antilles-Guyane (juin 2013)	Exercice 3 1/6 (17%)	Mathématique	Etude de la fonction $A(a)$ représentant l'aire entre deux courbes entre l'axe (Oy) et la droite d'équation $x = a$: détermination de son expression et étude de la limite	Analyse Langue naturelle Graphique	X	Aire finie (égale à 2 e) d'un domaine infini.
Asie (juin 2013)						
Centres étrangers (juin 2013)	Exercice 3 Tout mais surtout 4/7 (57%)	Mathématique	Partage d'aire (verticalement puis horizontalement) Calcul d'aires	Algébrique Analyse Langue naturelle Graphique	X	Partage des aires verticalement : peut être mis en lien avec les probabilités

[1] Il s'agit du rapport des questions portant sur les probabilités à densité sur le nombre total de questions de l'exercice.

			Egalité d'aires			(détermination de la valeur a telle que $P(X \leq a) = P(X \geq a) = 0{,}5$ (« médiane »)
Métropole (juin 2013)	Exercice 2 2/13 (15%)	Mathématique	Calculs d'aires	Analyse Langue naturelle Graphique	X	
Antilles-Guyane (septembre 2013)	Exercice 2 1/11 (9%)	Mathématique	Calcul d'aire sous une courbe	Langue naturelle Pas de graphique	X	La fonction en question $x \to -(x+1)exp(-x)$ est une primitive de $x \to x\ exp(-x)$ -- qui sert à calculer l'espérance d'une loi exponentielle de paramètre 1.
Métropole (septembre 2013)	Exercice 1 7/10 (70%)	Mathématique	Fonction primitive Interprétation d'une intégrale Calcul d'intégrale Méthode des rectangles	Langue naturelle Analyse Graphique Pseudo-code	X	
Nouvelle-Calédonie (novembre 2013)						
Amérique du Sud (novembre 2013)						
Nouvelle-Calédonie (mars 2014)	Exercice 3 6/10 (60%)	Mathématique	Méthode des rectangles Compréhension de l'algorithme Encadrement d'aire Calcul exact d'aire sous la courbe	Langue naturelle Analyse Graphique Pseudo-code	X	
2014						
Pondichéry (avril 2014)	Exercice 4 1/9 (11%)	Logo d'un créateur	Calcul de l'aire entre deux courbes	Langue naturelle Graphique	X	
Liban	Exercice 3	Mathématique	Fonction Aire sous la courbe	Langue naturelle	X	A peut être interprétée

(mai 2014)	7/9 (77%)		Lien primitive – dérivée Calcul d'aire	Analyse Graphique		comme la fonction de répartition associée à une variable aléatoire continue de fonction de densité f. De plus $f(x) = xe^{-x}$ sur $[0 ; +\infty[$. Or son intégrale sur ce domaine correspond à l'espérance de la loi exponentielle de paramètre 1. Aire finie d'un domaine infini. Beaucoup de liens possibles.
Amérique du Nord (mai 2014)	Exercice 2 4/9 (44%)	Mathématique	Etude de la fonction Aire entre deux courbes	Langue naturelle Analyse Graphique	X	
Centres étrangers (juin 2014)	Exercice 3 Calcul d'aire 3/10 (30%)	Image numérique Fonction de retouche Eclairci le plus l'image si l'aire est la plus petite	Calcul d'aires sous une courbe Interprétation	Langue naturelle Analyse Graphique	X	
Polynésie (juin 2014)	Exercice 4 1/10 (10%)	Mathématique	Calcul d'aire entre deux courbes	Langue naturelle Pas de graphique	X	
Antilles-Guyane (juin 2014)	Exercice 2 2/9 (22%)	Mathématique	Fonction primitive Calcul d'aire	Langue naturelle Analyse Pas de graphique	X	La fonction $x \to xexp(-x)$ est la fonction à intégrer pour déterminer l'espérance d'une loi exponentielle de

						paramètre 1. La primitive donnée permettrait alors le calcul de l'espérance.
Asie (juin 2014)	Exercice 3 1/6 (16%)	Chaine suspendue à deux points d'accroche	Calcul d'intégrale	Langue naturelle Analyse Pas de graphique	X	
	Exercice 4 10/10 (100%)	Mathématique	Etude d'une suite définie par une intégrale ($I_n = \int f_n$) Interprétation graphique de l'intégrale. Calcul d'intégrale. Algorithme de la méthode des rectangles.	Langue naturelle Analyse Graphique Pseudo-code	X	
Métropole (juin 2014)	Exercice 1 4/6 (66%)	Mathématique	Etude d'une suite définie par une intégrale du type $I_n = \int f_n$ Interprétation graphique de l'intégrale. Calcul d'intégrale.	Langue naturelle Analyse Graphique	X	
Antilles-Guyane (septembre 2014)						
Métropole (septembre 2014)	Exercice 1 2/9 (22%)	Mathématique	Expression de l'aire en fonction d'une intégrale Calcul d'intégrale	Langue naturelle Analyse Graphique	Non	La fonction en question est $f(x) = x + 1 + axe^{-x^2}$. On retrouve pratiquement $x \to xe^{-x^2/2}$, utile pour le calcul de l'espérance de la loi normale centrée réduite
Amérique du Sud (novembre 2014)	Exercice 4 4/7 (58%)	Aire d'un portail	Primitive Calcul d'aire Calcul d'aire de rectangles à l'aide d'un algorithme	Langue naturelle Analyse Schéma Graphique Pseudo-code		

Nouvelle-Calédonie (novembre 2014)						
Nouvelle-Calédonie (mars 2015)						
2015						
Pondichéry (avril 2015)	Exercice 1 4/8 (50%)	Mathématique	Primitive Interprétation graphique de l'intégrale Calcul d'intégrale Détermination de l'aire d'un domaine infini	Langue naturelle Analyse Graphique	X	Aire finie d'un domaine infini
Liban (mai 2015)	Exercice 2 8/8 (100%)	Mathématique	Etude d'une suite définie par une intégrale du type $I_n = \int f_n$ Calcul d'intégrale	Langue naturelle Analyse Pseudo-code	X	
Amérique du Nord (juin 2015)	Exercice 4 1/8 (13%)	Mathématique	Primitive Calcul d'intégrale Interprétation graphique de l'intégrale	Langue naturelle Analyse Pas de graphique	X	
Centres étrangers (juin 2014)	Exercice 4 4/7 (57%)	Logo d'entreprise	Primitive Calcul d'aire	Langue naturelle Analyse Figure	X	
Polynésie (juin 2015)						
Asie (juin 2015)	Exercice 3 1/6 (17%)	Mathématique	Aire sous une suite de courbes	Langue naturelle Graphique	X	
Antilles-Guyane (juin 2015)						
Métropole (juin 2015)	Exercice 4 4/8 (50%)	Piste de skateboard	Primitive Calcul d'aire (surface) Algorithme de la méthode des trapèzes	Langue naturelle Analyse Graphique Schéma	X	
Métropole – La Réunion	Exercice 2 8/8 (100%)	Mathématique	Etude d'une suite définie par une intégrale du type	Langue naturelle Analyse	Algo méth des rectangles	Limite d'une intégrale (intégrale impropre)

(septembre 2015)			$I_n = \int_0^n f$ Algorithme de la méthode des rectangles	Pseudo-code Graphique		Aire finie d'un domaine infini Présence de la fonction $x \to (-x-1)e^{-x}$, primitive de la fonction $x \to xe^{-x}$
	Exercice 4 3/3 (100%)	Mathématique	Lien primitive – fonction, à l'aide du graphique Calcul approché/exact de l'aire	Langue naturelle Analyse Graphique	X	Lien primitive/fonction : lien fonction de répartition et fonction de densité
Polynésie (septembre 2015)	Exercice 1 1/12 (8%)	Mathématique	Calcul d'aire entre 2 courbes	Langue naturelle Analyse Graphique		
Antilles - Guyane (septembre 2015)	Exercice 1 8/12 (67%)	Mathématique	Etude d'une suite définie par une intégrale du type $I_n = \int f_n$ Calcul d'intégral	Langue naturelle Analyse Pas de graphique	X	
Nouvelle-Calédonie (novembre 2015)	Exercice 1 2/9 (22%)	Etiquette de bouteille	Calcul d'aire : aire du disque et aire sous une courbe	Langue naturelle Algébrique Graphique	X	Dans le même exercice que les probabilités à densité mais dans des parties indépendantes
Amérique du Sud (novembre 2015)	Aire 2/8 (25%)	Mathématique	Calcul d'aire (fonction négative) Egalités d'aires	Langue naturelle Algébrique Graphique	X	

D.3 Analyse des exercices de lois à densité

Pays/Date	**Proportion/ pourcentage de l'exercice traitant de probabilités à densité**	**Contexte**	**Lois étudiées.** ***Types de tâches***	**Registres rencontrés dans l'énoncé**	**Artefact**	**Liens possibles avec le calcul intégral**
2013						
Pondichéry (avril 2013)	Exercice 4 1/9[1] (11%)	Nombre de salariés d'une entreprise absents pendant une épidémie de grippe	Approximation de la loi binomiale par une loi normale Loi normale : paramètres connus *Détermination de probabilité à l'aide de la loi centrée réduite*	Langue naturelle Probabiliste Tableau Pas de graphique	Table de la loi normale centrée réduite → Non nécessité de la calculatrice	
Amérique du Nord (mai 2013)	Exercice 3 6/8 (75%)	Boulangerie : poids d'un pain / temps de fonctionnement d'une balance électronique	Loi normale : paramètres connus, puis σ inconnu *Détermination de probabilité.* *Détermination de σ, connaissant une probabilité.* Loi exponentielle : λ inconnu, puis donné *Détermination de λ connaissant une probabilité.* *Calcul de probabilité conditionnelle et de probabilité.*	Langue naturelle Probabiliste Tableau Pas de graphique	Table de la loi normale en question → Non nécessité de la calculatrice	
Liban (mai 2013)	Exercice 2 4/8 (50%)	Entreprise de compotes « allégées »	Loi normale : paramètres connus, puis σ inconnu *Détermination de probabilités.* *Reconnaissance de la loi*	Langue naturelle Probabiliste Tableau Pas de graphique	Tables : de la loi en question, puis de la loi centrée réduite → Non nécessité de la calculatrice	

[1] Il s'agit du rapport des questions portant sur les probabilités à densité sur le nombre total de questions de l'exercice.

			centrée réduite associée à une loi normale. *Détermination de σ, connaissant une probabilité.*			
Polynésie (juin 2013)	Exercice 3 2/7 (28%)	Durée d'une chanson sur un MP3	Loi normale : paramètres connus *Détermination de probabilités.*	Langue naturelle Probabiliste Tableau Pas de graphique	Table de la loi normale en question → Non nécessité de la calculatrice	
Antilles-Guyane (juin 2013)	Exercice 2 2/7 (28%)	Nombre d'étudiants connaissant le sigle URSSAF	Approximation de la loi binomiale par une loi normale Loi normale : paramètres inconnus *Détermination des paramètres, connaissant la loi binomiale qu'elle approche.* *Détermination de probabilités.*	Langue naturelle Probabiliste Tableau Pas de graphique	Table de la loi normale en question → Non nécessité de la calculatrice	
Asie (juin 2013)						
Centres étrangers (juin 2013)	Exercice 1 5/10 (50%)	Industriel : commercialisation de vannes électroniques (demande mensuelle)	Loi exponentielle : paramètre connu *Calcul de l'espérance.* *Calcul d'une probabilité.* Loi normale : paramètres connus *Détermination de probabilités.*	Langue naturelle Probabiliste Pas de graphique	Calculatrice non obligatoire. Le théorème des intervalles 1, 2, 3 sigmas peut permettre de répondre (seulement apparait un problème au niveau de l'arrondi demandé)	
Métropole (juin 2013)						
Antilles-Guyane (septembre 2013)	Exercice 3 2/4 (50%)	Entreprise industrielle : fabrication de pièces cylindriques (longueur, diamètre)	Loi normale : paramètres connus *Détermination d'une probabilité.*	Langue naturelle Probabiliste Tableau	Calculatrice non obligatoire. Le théorème des intervalles 1, 2, 3 sigmas permet de	

				Pas de graphique	répondre, puis la table. Table de la loi normale en question	
Métropole (septembre 2013)						
Nouvelle-Calédonie (novembre 2013)	Exercice 3 1/8 (13%)	Usine : fabrication de billes sphériques (diamètre)	Loi normale : paramètres connus *Détermination d'une probabilité.*	Langue naturelle Probabiliste Tableau Pas de graphique	Table de la loi normale en question → Non nécessité de la calculatrice	
Amérique du Sud (novembre 2013)						
Nouvelle-Calédonie (mars 2014)	Exercice 2 6/12 (50%)	Mathématique pour la ROC Laboratoire : contenance des pipettes	Loi normale : paramètres connus *Détermination d'une probabilité.*	Langue naturelle Probabiliste Tableau Pas de graphique	Table de la loi normale en question → Non nécessité de la calculatrice	ROC (restitution organisée des connaissances) : Lien intégrale / probabilité avec $H(x) = P(-x \leq X \leq x)$ (X variable aléatoire suivant la loi normale centrée réduite)
Remarque					Il y pratiquement toujours une table fournie contrairement à ce que le programme laissait entendre (importance de la calculatrice) : nous pouvons penser que cet artefact a été donné, pour la transition.	
2014						
Pondichéry (avril	Exercice 1 5/6 (83%)	Durée de vie d'un moteur	Loi exponentielle : paramètre inconnu, puis	Langue naturelle Probabiliste	Calculatrice (seulement pour arrondir les	

2014)			connu *Détermination de λ connaissant une probabilité.* *Calcul d'une probabilité.* *Démonstration de la propriété de loi sans mémoire.* *Calcul de probabilité conditionnelle.* *Calcul de l'espérance et interprétation.*	Pas de graphique	résultats)	
Liban (mai 2014)	Exercice 1 5/9 (56%)	Temps de parcours en vélo/en bus, entre son domicile et son lycée	Loi normale : paramètres connus, puis σ inconnu *Détermination de probabilités.* *Détermination une valeur u telle que* $p(X \leq u) = a$ *(a donné).* *Reconnaissance de la loi centrée réduite associée à une loi normale.* *Détermination de σ, connaissant une probabilité.*	Langue naturelle Probabiliste Pas de graphique	Calculatrice	
Amérique du Nord (mai 2014)	Exercice 1 4/7 (57%)	Enseigne de cosmétiques : conditionnement des pots de crème	Loi normale : paramètres connus, puis σ inconnu *Détermination de probabilités.* *Reconnaissance de la loi centrée réduite associée à une loi normale.* *Détermination une valeur u telle que* $p(X \leq u) = a$ *(a donné)* *Détermination de σ, connaissant une*	Non	Calculatrice	

			probabilité.			
Centres étrangers (juin 2014)	Exercice 1 2/4 (50%) QCM	Hypermarché : Durée de vie d'une télé / masse des baguettes	Loi exponentielle : paramètre connu *Calcul d'une probabilité.* Loi normale : σ inconnu *Détermination d'une probabilité, connaissant une autre probabilité.*	Langue naturelle Pas de graphique	Calculatrice (pour arrondir pour la loi exponentielle) Calculatrice non obligatoire pour la loi normale (théorème des intervalles 1, 2, 3 sigmas)	
Polynésie (juin 2014)	Exercice 3 1/4 (25%) Vrai/Faux	Temps d'attente à un guichet	Loi exponentielle : paramètre connu *Calcul d'une probabilité.*	Langue naturelle Pas de graphique	Calculatrice (pour arrondir)	
Antilles-Guyane (juin 2014)	Exercice 1 2/8 (25%)	Ostréiculteur : masse des huitres	Loi normale : paramètres connus *Détermination de probabilités.*	Langue naturelle Probabiliste Pas de graphique	Calculatrice	
Asie (juin 2014)	Exercice 2 3/9 (33%)	Taux d'hématocrite	Loi normale : σ inconnu, puis paramètres connus *Reconnaissance de la loi centrée réduite associée à une loi normale.* *Détermination de probabilités à l'aide de la loi centrée réduite.* *Détermination d'une probabilité.*	Probabiliste Pas de graphique	Calculatrice non obligatoire (théorème des intervalles 1, 2, 3 sigmas)	
Métropole (juin 2014)	Exercice 2 2/7 (29%)	Laboratoire pharmaceutique : conformité des comprimés d'un médicament	Loi normale : paramètres connus *Détermination d'une probabilité.* *Détermination de h tel que* $P(\mu - h \leq X \leq \mu + h) \approx 0{,}99$.	Langue naturelle Probabiliste Pas de graphique	Calculatrice	
Antilles-Guyane (septembre 2014)	Exercice 1 4/7 (58%)	Entreprise de jouet en peluche – normes	Loi exponentielle : paramètre inconnu, puis connu *Détermination de* λ *connaissant une*	Langue naturelle Probabiliste Pas de graphique	Calculatrice	

			probabilité. *Calcul d'une probabilité.* *Calcul de probabilité conditionnelle.* Loi normale : σ inconnu *Reconnaissance de la loi centrée réduite associée à une loi normale.* *Détermination de σ, connaissant une probabilité.*			
Métropole (septembre 2014)	Exercice 2 5/6 (84%)	Mode de fonctionnement de restaurants : temps d'attente	Loi exponentielle : paramètre inconnu *Détermination de λ, connaissant l'espérance.* *Calcul d'une probabilité.* *Calcul de probabilité conditionnelle.* Loi normale : paramètres connus *Détermination de probabilités.*	Langue naturelle Probabiliste Pas de graphique	Calculatrice (explicite)	
Amérique du Sud (novembre 2014)	Exercice 1 2/7 (29%)	Taille de ballons de football	Loi normale : paramètre connus, puis σ inconnu *Détermination d'une probabilité.* *Détermination de σ, connaissant une probabilité.*	Langue naturelle Probabiliste Pas de graphique	Calculatrice non obligatoire (théorème des intervalles 1, 2, 3 sigmas et résultat donné pour une loi normale centrée réduite)	
Nouvelle-Calédonie (novembre 2014)	Exercice 1 1/4 (25%)	Masse de crème glacée	Loi normale : σ inconnu *Détermination de σ, connaissant une probabilité.*	Langue naturelle Pas de graphique	Calculatrice	
Nouvelle-Calédonie (mars	Exercice 2 3/9 (33%)	Durée de vie d'une puce	Loi exponentielle : paramètre inconnu *Détermination de λ,*			

2015)			*connaissant l'espérance. Calcul de probabilités. Interprétation.*			
Remarque					Plus de tables. Seulement calculatrice.	
2015						
Pondichéry (avril 2015)	Exercice 3 7/11 (64%)	Durée de vie d'un appareil électroménager	Loi normale : σ inconnu *Détermination d'une probabilité, à l'aide du graphique. Détermination de σ, à l'aide du graphique. Reconnaissance de la loi centrée réduite associée à une loi normale. Détermination de σ, connaissant une probabilité. Détermination de probabilités.*	Langue naturelle Probabiliste **Graphique**	Calculatrice	Jeu sur les aires du graphique.
Liban (mai 2015)						
Amérique du Nord (juin 2015)	Exercice 3 2/4 (50%)	Contrôle avant la mise sur le marché : masse d'une tablette de chocolat	Loi normale : paramètres connus, puis σ inconnu *Détermination d'une probabilité. Détermination de σ, connaissant une probabilité.*	Langue naturelle Probabiliste Pas de graphique	Calculatrice	
Centres étrangers (juin 2015)	Exercice 1 2/5 (40%)	Nombre de cadenas vendus	Loi normale : paramètres connus *Détermination d'une probabilité. Détermination le plus petit n vérifiant* $P(X \geq n) \leq$ 0,05.	Langue naturelle Probabiliste Pas de graphique	Calculatrice non obligatoire (théorème des intervalles 1, 2, 3 sigmas)	
Polynésie (juin 2015)	Exercice 3 2/3 (66%)	Taille homme/femme	Loi normale : paramètres connus	Langue naturelle	Calculatrice	

			Détermination de probabilités.	Pas de graphique		
Asie (juin 2015)	Exercice 1 3/7 (43%)	Durée de vie d'un panneau électrique affichant des scores Mathématique pour la ROC	Loi exponentielle : paramètre connu *Calcul d'une probabilité.* *Démonstration de la formule de l'espérance.* *Calcul de l'espérance.*	Langue naturelle Analyse Pas de graphique	Calculatrice (pour arrondir les résultats)	ROC (restitution organisée des connaissances) : lien probabilités et intégrale avec la démonstration de la formule de l'espérance de l'exponentielle
Antilles-Guyane (juin 2015)	Exercice 2 8/10 (80%)	Mathématique pour la ROC Durée de vie d'un composant électronique	Loi exponentielle : paramètre inconnu *Démonstration de la formule de l'espérance.* *Représentation d'une probabilité sur le graphique.* *Lecture de λ sur le graphique.* *Interprétation de l'espérance.* *Détermination de λ, connaissant l'espérance.*	Langue naturelle Analyse Probabiliste **Graphique**	X	ROC (restitution organisée des connaissances) : lien probabilités et intégrale avec la démonstration de la formule de l'espérance de l'exponentielle. Travail sur le graphique.
Métropole (juin 2015)	Exercice 2 7/10 (70%)	Mathématique	Loi exponentielle : paramètre inconnu, puis paramètre connu *Démonstration de la formule de la probabilité que X appartienne à* $[c\,;d]$. *Détermination de λ, connaissant une probabilité.* *Déterminer l'espérance.* *Calcul de probabilités.*	Langue naturelle Probabiliste Pas de graphique	Calculatrice (pour arrondir)	ROC (restitution organisée des connaissances) : nécessité de passer par l'intégrale.
Métropole – La Réunion (septembre 2015)	Exercice 1 2/5 (40%) QCM	Radioactivité Mathématique	Loi exponentielle : paramètre connu *Calcul d'une probabilité.* Loi normale : paramètres	Langue naturelle Probabiliste Pas de graphique	Calculatrice	QCM

			connus *Calcul d'une probabilité.*			
Polynésie (septembre 2015)	Exercice 2 2/6 (33%)	Taux d'une substance présente dans le sang	Loi normale : paramètres connus, puis μ et σ inconnus *Calcul d'une probabilité.* *Détermination de μ en interprétant l'énoncé.* *Détermination de σ, connaissant une probabilité.*	Langue naturelle Pas de graphique	Calculatrice	
Antilles-Guyane (septembre 2015)						
Nouvelle-Calédonie (septembre 2015)	Exercice 1 3/9 (33%)	Taux de calcium de l'eau que contient une bouteille d'eau minérale d'une usine (prélevée au hasard)	Loi normale : paramètres connus, puis σ inconnu *Détermination de probabilités.* *Détermination de σ, connaissant une probabilité.*	Langue naturelle Probabiliste Pas de graphique	Calculatrice	Dans le même exercice que le calcul intégral mais dans des parties indépendantes
Amérique du Sud (novembre 2015)	Exercice 3 4/12 (25%)	Temps d'incubation du virus	Loi normale : μ inconnu *Conjecture de la valeur de μ, à l'aide du graphique.* *Représentation d'une probabilité sur le graphique.* *Reconnaissance de la loi centrée réduite associée à une loi normale.* *Détermination de μ, connaissant une probabilité.*	Langue naturelle Probabiliste **Graphique**	Calculatrice	
Remarque					Plus de tables. Seulement calculatrice.	

D.4 Exercices de baccalauréat

D.4.1 Métropole - Réunion – Septembre 2015

EXERCICE 2 **7 POINTS**

Commun à tous les candidats

Soit f la fonction définie et dérivable sur l'intervalle $[0\,;+\infty[$ telle que :

$$f(x)=\frac{x}{\mathrm{e}^x-x}$$

On admet que la fonction f est positive sur l'intervalle $[0\,;+\infty[$.
On note $\mathscr{C}$ la courbe représentative de la fonction f dans un repère orthogonal du plan.
La courbe $\mathscr{C}$ est représentée en annexe, **à rendre avec la copie**.

Partie A

Soit la suite (I_n) définie pour tout entier naturel n par $I_n=\displaystyle\int_0^n f(x)\,\mathrm{d}x$.

On ne cherchera pas à calculer la valeur exacte de I_n en fonction de n.

1. Montrer que la suite (I_n) est croissante.
2. On admet que pour tout réel x de l'intervalle $[0\,;+\infty[$, $\mathrm{e}^x-x\geqslant\dfrac{\mathrm{e}^x}{2}$.
   a. Montrer que, pour tout entier naturel n, $I_n\leqslant\displaystyle\int_0^n 2x\mathrm{e}^{-x}\,\mathrm{d}x$.
   b. Soit H la fonction définie et dérivable sur l'intervalle $[0\,;+\infty[$ telle que :

   $$H(x)=(-x-1)\mathrm{e}^{-x}$$

   Déterminer la fonction dérivée H' de la fonction H.
   c. En déduire que, pour tout entier naturel n, $I_n\leqslant 2$.
3. Montrer que la suite (I_n) est convergente. On ne demande pas la valeur de sa limite.

Partie B

On considère l'algorithme suivant dans lequel les variables sont
— K et i des entiers naturels, K étant non nul ;
— A, x et h des réels.

Entrée :	Saisir K entier naturel non nul
Initialisation	Affecter à A la valeur 0 Affecter à x la valeur 0 Affecter à h la valeur $\dfrac{1}{K}$

Traitement	Pour i variant de 1 à K
	  Affecter à A la valeur $A+h\times f(x)$
	  Affecter à x la valeur $x+h$
	Fin Pour

Sortie	Afficher A

1. Reproduire et compléter le tableau suivant, en faisant fonctionner cet algorithme pour $K=4$. Les valeurs successives de A seront arrondies au millième.

i	A	x
1		
2		
3		
4		

2. En l'illustrant sur l'annexe **à rendre avec la copie**, donner une interprétation graphique du résultat affiché par cet algorithme pour $K = 8$.
3. Que donne l'algorithme lorsque K devient grand ?*

Exercice 4 **3 points**

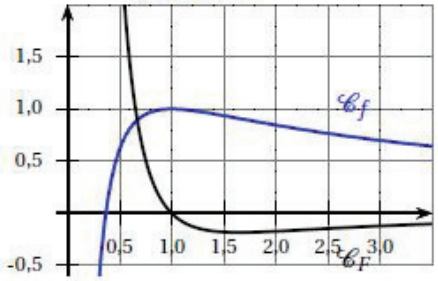
Commun à tous les candidats

On considère la fonction f définie sur $]0\,;\,+\infty[$ par

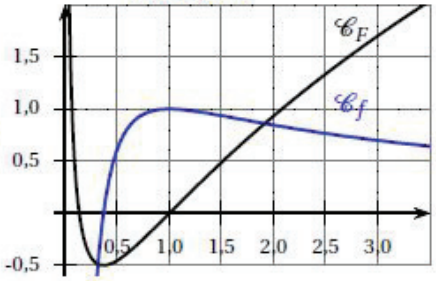
$$f(x) = \frac{1}{x}(1 + \ln x)$$

1. Dans les trois situations suivantes, on a dessiné, dans un repère orthonormé, la courbe représentative $\mathscr{C}_f$ de la fonction f et une courbe $\mathscr{C}_F$. Dans une seule situation, la courbe $\mathscr{C}_F$ est la courbe représentative d'une primitive F de la fonction f. Laquelle ? Justifier la réponse.

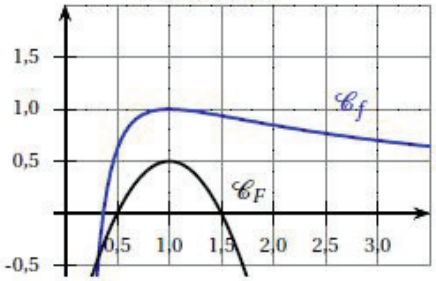
Situation 1



Situation 2



Situation 3



2. Dans la situation retenue à la question 1, on appelle :
   - K le point d'intersection de la courbe $\mathscr{C}_f$ et de l'axe des abscisses et $\mathscr{D}$ la droite passant par K et parallèle à l'axe des ordonnées ;
   - L le point d'intersection de $\mathscr{C}_F$ et de l'axe des abscisses, ayant une abscisse supérieure à $\frac{1}{2}$ et Δ la droite passant par L et parallèle à l'axe des ordonnées.

   a. Déterminer une valeur approchée de l'aire du domaine du plan délimité par les droites $\mathscr{D}$ et Δ, par la courbe $\mathscr{C}_f$ et par l'axe des abscisses.

   b. Peut-on déterminer la valeur exacte de cette aire ?

D.4.2 Polynésie – Septembre 2015

EXERCICE 2 **5 points**

Commun à tous les candidats

Partie A

On étudie une maladie dans la population d'un pays. On a constaté que le taux, en nanogrammes par millilitre $\left(\text{ng.mL}^{-1}\right)$, d'une substance Gamma présente dans le sang est plus élevé chez les personnes atteintes de cette maladie que chez les personnes qui n'en sont pas atteintes.

1. Le taux de cette substance Gamma dans la population des personnes qui ne sont pas atteintes par la maladie est modélisé par une variable aléatoire T qui suit la loi normale d'espérance $\mu = 40$ et d'écart-type $\sigma = 8$.

   On choisit au hasard une personne parmi celles qui ne sont pas atteintes par la maladie étudiée.

   Calculer la probabilité que le taux dans le sang de la substance Gamma soit supérieur à 60 ng.mL^{-1}.

2. Des études ont mis en évidence que le taux moyen de la substance Gamma chez les personnes atteintes par la maladie étudiée est de 50 ng.mL^{-1} et que 10 % d'entre elles ont un taux de substance Gamma inférieur à 43 ng.mL^{-1}.

   On appelle T' la variable aléatoire qui modélise le taux de la substance Gamma en ng.mL^{-1} chez une personne atteinte par la maladie étudiée.

   On admet que T' suit la loi normale d'espérance μ' et d'écart-type σ'.

   Préciser la valeur de μ' et déterminer la valeur de σ'.

Partie B

Pour dépister chez une personne la maladie étudiée, on effectue une prise de sang.
On considère que le dépistage est positif si le taux de la substance Gamma est supérieur ou égal à 45 $ng.mL^{-1}$.
Une personne étant choisie au hasard dans la population, on appelle :

- M l'évènement « le patient est atteint par la maladie étudiée » ;
- D l'évènement « le patient a un dépistage positif ».

On admet que :

- 82 % des personnes atteintes par la maladie étudiée ont un dépistage positif ;
- 73 % des personnes non atteintes par cette maladie ont un dépistage négatif.

On sait de plus que 10 % de la population étudiée est atteinte par cette maladie.

1. Démontrer que la probabilité qu'un patient ait un dépistage positif est de 0,325.
2. Calculer $P_{\overline{D}}(M)$. Interpréter ce résultat.
3. Un patient a un dépistage positif. Le médecin le rassure en lui indiquant qu'il n'a qu'une chance sur quatre d'avoir contracté la maladie. Qu'en pensez-vous ?

Partie C

Lors du dépistage précédent, la prise de sang est effectuée chez des sujets à jeun.
Les données montrent que 82 % des patients malades ont un dépistage positif.
Pour améliorer le confort des personnes susceptibles de subir cet examen sanguin, on souhaite vérifier si le fait d'être à jeun est une condition indispensable dans le protocole.
On considère un groupe de 300 personnes malades sur lesquelles la prise de sang n'est pas effectuée à jeun.
Le dépistage se révèle positif pour 74 % d'entre elles.
Ce dépistage peut-il être effectué sur des personnes qui ne sont pas à jeun ?*

D.4.3 Nouvelle-Calédonie – Novembre 2015

EXERCICE 1 **7 points**
Commun à tous les candidats

Une usine produit de l'eau minérale en bouteilles. Lorsque le taux de calcium dans une bouteille est inférieur à 6,5 mg par litre, on dit que l'eau de cette bouteille est très peu calcaire.

Dans cet exercice les résultats approchés seront arrondis au millième.

Partie A

L'eau minérale provient de deux sources, notées « source A » et « source B ».
La probabilité que l'eau d'une bouteille prélevée au hasard dans la production d'une journée de la source A soit très peu calcaire est 0,17. La probabilité que l'eau d'une bouteille prélevée au hasard dans la production d'une journée de la source B soit très peu calcaire est 0,10.

La source A fournit 70 % de la production quotidienne totale des bouteilles d'eau et la source B le reste de cette production.

On prélève au hasard une bouteille d'eau dans la production totale de la journée. On considère les évènements suivants :
A : « La bouteille d'eau provient de la source A »
B : « La bouteille d'eau provient de la source B »
S : « L'eau contenue dans la bouteille d'eau est très peu calcaire ».

1. Déterminer la probabilité de l'évènement $A \cap S$.
2. Montrer que la probabilité de l'évènement S vaut 0,149.
3. Calculer la probabilité que l'eau contenue dans une bouteille provienne de la source A sachant qu'elle est très peu calcaire.
4. Le lendemain d'une forte pluie, l'usine prélève un échantillon de 1 000 bouteilles provenant de la source A. Parmi ces bouteilles, 211 contiennent de l'eau très peu calcaire.
   Donner un intervalle permettant d'estimer au seuil de 95 % la proportion de bouteilles contenant de l'eau très peu calcaire sur l'ensemble de la production de la source A après cette intempérie.

Partie B

On note X la variable aléatoire qui, à chaque bouteille prélevée au hasard dans la production d'une journée de la source A, associe le taux de calcium de l'eau qu'elle contient. On suppose que X suit la loi normale de moyenne 8 et d'écart-type 1,6.
On note Y la variable aléatoire qui, à chaque bouteille prélevée au hasard dans la production d'une journée de la source B, associe le taux de calcium qu'elle contient. On suppose que Y suit la loi normale de moyenne 9 et d'écart-type σ.

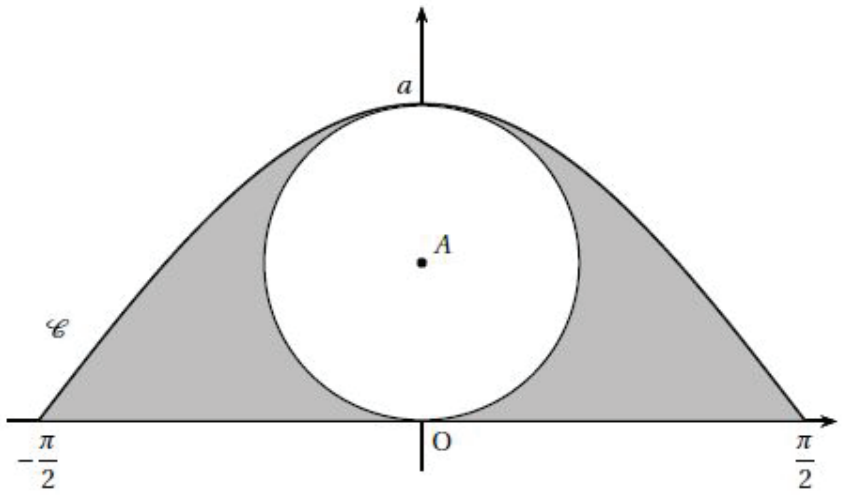
1. Déterminer la probabilité pour que le taux de calcium mesuré dans une bouteille prise au hasard dans la production d'une journée de la source A soit compris entre 6,4 mg et 9,6 mg.
2. Calculer la probabilité $p(X \leqslant 6,5)$.

3. Déterminer σ sachant que la probabilité qu'une bouteille prélevée au hasard dans la production d'une journée de la source B contienne de l'eau très peu calcaire est 0,1.

Partie C

Le service commercial a adopté pour les étiquettes des bouteilles la forme représentée ci-dessous dans un repère orthonormé du plan.
La forme de ces étiquettes est délimitée par l'axe des abscisses et la courbe $\mathscr{C}$ d'équation $y = a\cos x$ avec $x \in \left[-\frac{\pi}{2} ; \frac{\pi}{2}\right]$ et a un réel strictement positif.
Un disque situé à l'intérieur est destiné à recevoir les informations données aux acheteurs. On considère le disque de centre le point A de coordonnées $\left(0 ; \frac{a}{2}\right)$ et de rayon $\frac{a}{2}$. On admettra que ce disque se trouve entièrement en dessous de la courbe $\mathscr{C}$ pour des valeurs de a inférieures à 1,4.

1. Justifier que l'aire du domaine compris entre l'axe des abscisses, les droites d'équation $x = -\frac{\pi}{2}$ et $x = \frac{\pi}{2}$, et la courbe $\mathscr{C}$ est égale à $2a$ unités d'aire.
2. Pour des raisons esthétiques, on souhaite que l'aire du disque soit égale à l'aire de la surface grisée. Quelle valeur faut-il donner au réel a pour respecter cette contrainte ?



D.4.4 Métropole – Juin 2015

EXERCICE 1 **6 POINTS**

Commun à tous les candidats

Les résultats des probabilités seront arrondis à 10^{-3} près.

Partie 1

1. Soit X une variable aléatoire qui suit la loi exponentielle de paramètre λ, où λ est un réel strictement positif donné.
   On rappelle que la densité de probabilité de cette loi est la fonction f définie sur $[0\,;+\infty[$ par
   $$f(x) = \lambda \mathrm{e}^{-\lambda x}.$$

   a. Soit c et d deux réels tels que $0 \leqslant c < d$.
   Démontrer que la probabilité $P(c \leqslant X \leqslant d)$ vérifie $P(c \leqslant X \leqslant d) = \mathrm{e}^{-\lambda c} - \mathrm{e}^{-\lambda d}$.

   b. Déterminer une valeur de λ à 10^{-3} près de telle sorte que la probabilité $P(X > 20)$ soit égale à 0,05.

   c. Donner l'espérance de la variable aléatoire X

   Dans la suite de l'exercice on prend $\lambda = 0{,}15$.

   d. Calculer $P(10 \leqslant X \leqslant 20)$.

   e. Calculer la probabilité de l'évènement $(X > 18)$.

2. Soit Y une variable aléatoire qui suit la loi normale d'espérance 16 et d'écart type 1,95.

   a. Calculer la probabilité de l'évènement $(20 \leqslant Y \leqslant 21)$.

   b. Calculer la probabilité de l'évènement $(Y < 11) \cup (Y > 21)$.

Partie 2

Une chaîne de magasins souhaite fidéliser ses clients en offrant des bons d'achat à ses clients privilégiés. Chacun d'eux reçoit un bon d'achat de couleur verte ou rouge sur lequel est inscrit un montant.
Les bons d'achats sont distribués de façon à avoir, dans chaque magasin, un quart de bons rouges et trois quarts de bons verts.
Les bons d'achat verts prennent la valeur de 30 euros avec une probabilité égale à 0,067 ou des valeurs comprises entre 0 et 15 euros avec des probabilités non précisées ici.
De façon analogue, les bons d'achat rouges prennent les valeurs 30 ou 100 euros avec des probabilités respectivement égales à 0,015 et 0,010 ou des valeurs comprises entre 10 et 20 euros avec des probabilités non précisées ici.

1. Calculer la probabilité d'avoir un bon d'achat d'une valeur supérieure ou égale à 30 euros sachant qu'il est rouge.
2. Montrer qu'une valeur approchée à 10^{-3} près de la probabilité d'avoir un bon d'achat d'une valeur supérieure ou égale à 30 euros vaut 0,057.
   Pour la question suivante, on utilise cette valeur.

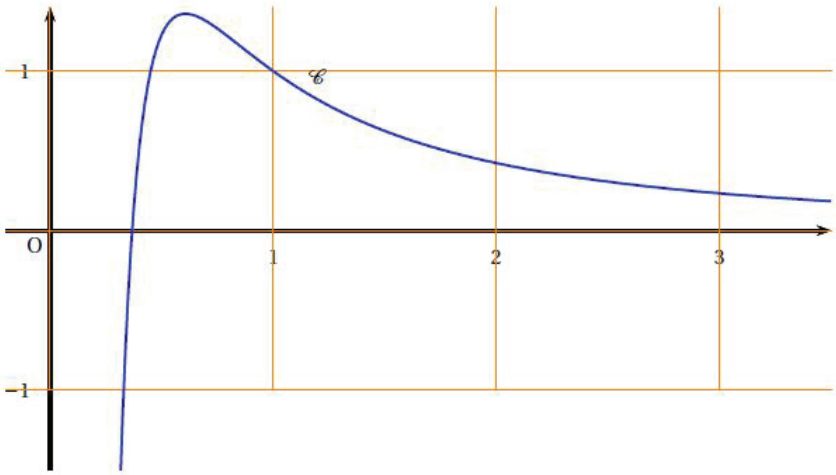
D.4.5 Amérique du Nord – Mai 2013

Exercice 4 **5 points**
Commun à tous les candidats

Soit f la fonction définie sur l'intervalle $]0\,;\,+\infty[$ par

$$f(x) = \frac{1+\ln(x)}{x^2}$$

et soit $\mathscr{C}$ la courbe représentative de la fonction f dans un repère du plan. La courbe $\mathscr{C}$ est donnée ci-dessous :



1. a. Étudier la limite de f en 0.
   b. Que vaut $\lim\limits_{x\to+\infty} \dfrac{\ln(x)}{x}$? En déduire la limite de la fonction f en $+\infty$.
   c. En déduire les asymptotes éventuelles à la courbe $\mathscr{C}$.
2. a. On note f' la fonction dérivée de la fonction f sur l'intervalle $]0\,;\,+\infty[$. Démontrer que, pour tout réel x appartenant à l'intervalle $]0\,;\,+\infty[$,

$$f'(x) = \frac{-1-2\ln(x)}{x^3}.$$

   b. Résoudre sur l'intervalle $]0\,;\,+\infty[$ l'inéquation $-1-2\ln(x) > 0$.
   En déduire le signe de $f'(x)$ sur l'intervalle $]0\,;\,+\infty[$.
   c. Dresser le tableau des variations de la fonction f.
3. a. Démontrer que la courbe $\mathscr{C}$ a un unique point d'intersection avec l'axe des abscisses, dont on précisera les coordonnées.
   b. En déduire le signe de $f(x)$ sur l'intervalle $]0\,;\,+\infty[$.
4. Pour tout entier $n \geqslant 1$, on note I_n l'aire, exprimée en unités d'aires, du domaine délimité par l'axe des abscisses, la courbe $\mathscr{C}$ et les droites d'équations respectives $x = \dfrac{1}{\mathrm{e}}$ et $x = n$.
   a. Démontrer que $0 \leqslant I_2 \leqslant \mathrm{e} - \dfrac{1}{2}$.
   On admet que la fonction F, définie sur l'intervalle $]0\,;\,+\infty[$ par $F(x) = \dfrac{-2-\ln(x)}{x}$,est une primitive de la fonction f sur l'intervalle $]0\,;\,+\infty[$.
   b. Calculer I_n en fonction de n.
   c. Étudier la limite de I_n en $+\infty$. Interpréter graphiquement le résultat obtenu.

D.4.6 Antilles - Guyane – Juin 2013

EXERCICE 3 **5 points**

Commun à tous les candidats

Dans tout ce qui suit, m désigne un nombre réel quelconque.

Partie A

Soit f la fonction définie et dérivable sur l'ensemble des nombres réels $\mathbb{R}$ telle que :

$$f(x) = (x+1)\mathrm{e}^x.$$

1. Calculer la limite de f en $+\infty$ et $-\infty$.
2. On note f' la fonction dérivée de la fonction f sur $\mathbb{R}$.
   Démontrer que pour tout réel x, $f'(x) = (x+2)\mathrm{e}^x$.
3. Dresser le tableau de variation de f sur $\mathbb{R}$.

Partie B

On définie la fonction g_m sur $\mathbb{R}$ par :

$$g_m(x) = x + 1 - m\mathrm{e}^{-x}$$

et on note $\mathscr{C}_m$ la courbe de la fonction g_m dans un repère $\left(\mathrm{O}, \vec{\imath}, \vec{\jmath}\right)$ du plan.

1. **a.** Démontrer que $g_m(x) = 0$ si et seulement si $f(x) = m$.
   b. Déduire de la partie A, sans justification, le nombre de points d'intersection de la courbe $\mathscr{C}_m$ avec l'axe des abscisses en fonction du réel m.
2. On a représenté en annexe 2 les courbes $\mathscr{C}_0$, $\mathscr{C}_\mathrm{e}$, et $\mathscr{C}_{-\mathrm{e}}$ (obtenues en prenant respectivement pour m les valeurs 0, e et $-$e).
   Identifier chacune de ces courbes sur la figure de l'annexe en justifiant.
3. Étudier la position de la courbe $\mathscr{C}_m$ par rapport à la droite $\mathscr{D}$ d'équation $y = x + 1$ suivant les valeurs du réel m.
4. **a.** On appelle D_2 la partie du plan comprise entre les courbes $\mathscr{C}_\mathrm{e}$, $\mathscr{C}_{-\mathrm{e}}$, l'axe $(\mathrm{O}y)$ et la droite $x = 2$. Hachurer D_2 sur l'annexe 2.
   b. Dans cette question, a désigne un réel positif, D_a la partie du plan comprise entre $\mathscr{C}_\mathrm{e}$, $\mathscr{C}_{-\mathrm{e}}$, l'axe $(\mathrm{O}y)$ et la droite Δ_a d'équation $x = a$. On désigne par $\mathscr{A}(a)$ l'aire de cette partie du plan, exprimée en unités d'aire.
   Démontrer que pour tout réel a positif : $\mathscr{A}(a) = 2\mathrm{e} - 2\mathrm{e}^{1-a}$.
   En déduire la limite de $\mathscr{A}(a)$ quand a tend vers $+\infty$.

D.4.7 Liban – Mai 2014

Exercice 3 **5 points**

Soit f la fonction définie sur l'intervalle $[0\,;+\infty[$ par

$$f(x) = xe^{-x}.$$

On note $\mathscr{C}$ la courbe représentative de f dans un repère orthogonal.

Partie A

1. On note f' la fonction dérivée de la fonction f sur l'intervalle $[0;+\infty[$.
   Pour tout réel x de l'intervalle $[0\,;+\infty[$, calculer $f'(x)$. En déduire les variations de la fonction f sur l'intervalle $[0\,;+\infty[$.
2. Déterminer la limite de la fonction f en $+\infty$. Quelle interprétation graphique peut-on faire de ce résultat ?

Partie B

Soit $\mathscr{A}$ la fonction définie sur l'intervalle $[0;+\infty[$ de la façon suivante : pour tout réel t de l'intervalle $[0\,;+\infty[$, $\mathscr{A}(t)$ est l'aire, en unités d'aire, du domaine délimité par l'axe des abscisses, la courbe $\mathscr{C}$ et les droites d'équations $x = 0$ et $x = t$.

1. Déterminer le sens de variation de la fonction $\mathscr{A}$.
2. On admet que l'aire du domaine délimité par la courbe $\mathscr{C}$ et l'axe des abscisses est égale à 1 unité d'aire. Que peut-on en déduire pour la fonction $\mathscr{A}$?
3. On cherche à prouver l'existence d'un nombre réel α tel que la droite d'équation $x = \alpha$ partage le domaine compris entre l'axe des abscisses et la courbe $\mathscr{C}$, en deux parties de même aire, et à trouver une valeur approchée de ce réel.
   a) Démontrer que l'équation $\mathscr{A}(t) = \dfrac{1}{2}$ admet une unique solution sur l'intervalle $[0\,;+\infty[$
   b) Sur le graphique fourni en **annexe (*à rendre avec la copie*)** sont tracées la courbe $\mathscr{C}$, ainsi que la courbe Γ représentant la fonction $\mathscr{A}$.
   Sur le graphique de l'**annexe**, identifier les courbes $\mathscr{C}$ et Γ, puis tracer la droite d'équation $y = \dfrac{1}{2}$. En déduire une valeur approchée du réel α.
   Hachurer le domaine correspondant à $\mathscr{A}(\alpha)$.
4. On définit la fonction g sur l'intervalle $[0;+\infty[$ par $g(x) = (x+1)e^{-x}$.
   a) On note g' la fonction dérivée de la fonction g sur l'intervalle $[0;+\infty[$.
   Pour tout réel x de l'intervalle $[0;+\infty[$, calculer $g'(x)$.
   b) En déduire, pour tout réel t de l'intervalle $[0;+\infty[$, une expression de $\mathscr{A}(t)$.
   c) Calculer une valeur approchée à 10^{-2} près de $\mathscr{A}(6)$.

Annexe E

Les manuels

E.1 Cours sur les lois à densité dans les manuels – Généralités

E.1.1 Cours du manuel *Déclic* (p. 366)

Cours

1 Loi à densité sur un intervalle borné

EXEMPLE D'INTRODUCTION

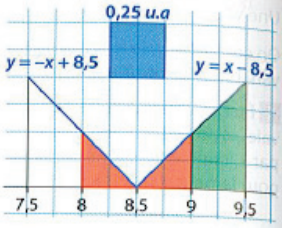
Un entrepôt accueille tous les matins des camions de livraison sur un créneau de deux heures d'ouverture, de 7 h 30 à 9 h 30. On s'intéresse à l'heure d'arrivée d'un camion qui se présente tous les matins à l'entrepôt aux heures d'ouverture. On admet que la probabilité que ce camion arrive dans un intervalle de temps donné $[t_1 ; t_2]$ est égale à l'aire du domaine compris entre l'axe des abscisses, **les deux segments tracés ci-contre**, et les droites d'équations $x = t_1$ et $x = t_2$ parallèles à l'axe des ordonnées.

Ainsi, la probabilité d'arrivée du camion entre 8 h 00 et 9 h 00 est égale à l'aire colorée en rouge : $P([8 ; 9]) = 0{,}25$,
celle qu'il arrive entre 9 h et 9 h 30 est :

$$P([9 ; 9{,}5]) = \int_{9}^{9{,}5} (x - 8{,}5)\,dx = 0{,}375.$$

Enfin, on vérifie (et c'est indispensable) que :

$$P([7{,}5 ; 9{,}5]) = \int_{7{,}5}^{9{,}5} |x - 8{,}5|\,dx = 1.$$



VOCABULAIRE

La fonction $x \mapsto |x - 8{,}5|$ est la **densité** permettant de définir la loi de probabilité, sur $[7{,}5 ; 9{,}5]$, de l'heure d'arrivée à l'entrepôt.

Définitions Soit $I = [a ; b]$ un intervalle borné de $\mathbb{R}$, avec $a < b$, et f une fonction positive sur I, telle que $\int_a^b f(t)dt = 1$. On définit une loi de probabilité sur l'intervalle I, modélisant le choix d'un réel x dans I, en posant pour tout intervalle J, d'extrémités c et d, avec $a \leqslant c \leqslant d \leqslant b$:

$$P(\{x \in J\}) = \int_c^d f(t)dt.$$

On dit que f est la **densité de cette loi de probabilité** sur I.

Remarque Les valeurs plus ou moins grandes prises par la fonction sur les différents intervalles donnent plus ou moins « de poids » à la probabilité de cet intervalle : d'où le nom de « **densité** » donné à la fonction.

Remarques

- On note plus simplement : $P(\{x \in J\}) = P(J)$.
- Pour tout réel c de I, $P(\{c\}) = \int_c^c f(t)dt = 0$; la probabilité que le nombre choisi soit c exactement est nulle.

E.1.2 Cours du manuel *Hyperbole* (p. 378)

Du cours ...

Loi à densité sur un intervalle

On considère une expérience aléatoire et un univers associé Ω, muni d'une probabilité.

1 Variable aléatoire continue

DÉFINITION Une variable aléatoire **continue** X est une fonction qui à chaque issue de Ω associe un nombre réel d'un intervalle I de ℝ.

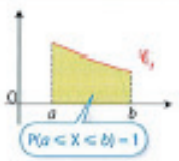
EXEMPLE : La variable aléatoire égale à la durée de bon fonctionnement d'un équipement produit en grande série est une variable aléatoire continue.

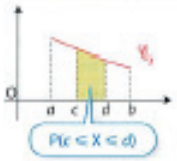
2 Loi de probabilité à densité

DÉFINITION X est une variable aléatoire continue à valeurs dans un intervalle I et f est une fonction continue, positive sur I telle que :

- $\int_a^b f(t)\,dt = 1$ si $I = [a\,;b]$
- $\lim\limits_{x\to+\infty} \int_a^x f(t)\,dt = 1$ si $I = [a\,;+\infty[$.

Dire que P est **la loi de probabilité de densité f** de X signifie que pour tout intervalle J inclus dans I, la probabilité $P(X \in J)$ est égale à l'aire du domaine $\{M(x\,;y)\,;x \in J \text{ et } 0 \leqslant y \leqslant f(x)\}$.





CONSÉQUENCES

(1) Pour tout nombre réel c de I, $P(X = c) = 0$.

En effet, $P(X = c) = P(c \leqslant X \leqslant c) = \int_c^c f(t)\,dt = 0$.

(2) On déduit immédiatement de **(1)** que :

$$P(c \leqslant X \leqslant d) = P(c \leqslant X < d) = P(c < X \leqslant d) = P(c < X < d)$$

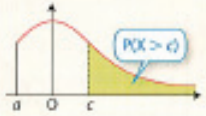
Aussi, note-t-on parfois $(c\,;d)$ un intervalle de bornes c et d.

(3) Si $J = (c\,;d)$, alors $P(x \in J) = \int_c^d f(t)\,dt$.

En effet, on sait que l'aire du domaine $\{M(x\,;y)\,;x \in (c\,;d) \text{ et } 0 \leqslant y \leqslant f(x)\}$ est égale, en unités d'aire, à $\int_c^d f(t)\,dt$.

(4) Si $I =]a\,;+\infty[$ et si c est un nombre réel tel que $c > a$:

$$P(X > c) = 1 - P(a < X < c)$$
$$= 1 - \int_a^c f(t)\,dt$$



Remarques : Les propriétés des probabilités d'événements rencontrées dans le cas discret s'étendent naturellement au cas continu. Par exemple :

- si J' est le complémentaire de J dans I, alors $P(J') = 1 - P(J)$;
- si $I' \subset I$ et $P(I') \neq 0$, si $J \subset I$, alors $P_{I'}(J) = \dfrac{P(I' \cap J)}{P(I')}$.

E.1.3 Cours du manuel *Indice* (p. 324)

Cours

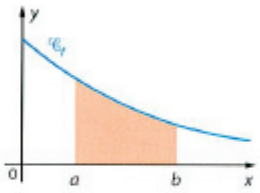
1 Lois de probabilité à densité et loi uniforme

Variable aléatoire à densité

Dans de nombreux domaines, on étudie des variables aléatoires pouvant prendre, **du moins théoriquement**, toute valeur d'un intervalle I de $\mathbb{R}$.
Ces variables aléatoires sont dites **continues**.

Vocabulaire
Une loi à densité est aussi appelée **loi continue**.

Exemple : On veut définir la variable aléatoire X qui, à tout téléviseur, associe sa durée de bon fonctionnement exprimée en heures. Cette durée peut prendre toute valeur de l'intervalle $I = [0 ; 50\,000]$.
On va chercher à calculer les probabilités $P(X \geqslant 10000)$ ou $P(0 < X < 20000)$. Pour cela, on utilise une fonction f définie sur I dont $\mathscr{C}_f$ est la courbe représentative dans un repère orthogonal : la probabilité $P(a < X < b)$ est définie comme l'aire (exprimée en unités d'aire) du domaine plan compris entre l'axe des abscisses, la courbe $\mathscr{C}_f$ et les droites d'équations $x = a$ et $x = b$.



Définitions
(1) On appelle **fonction de densité de probabilité** sur l'intervalle I, toute fonction f définie, continue et positive sur I telle que l'intégrale de f sur I soit égale à 1.
(2) Une **variable aléatoire à densité** X sur un intervalle I est définie par la donnée d'une fonction de densité de probabilité f définie sur I. La probabilité pour que X appartienne à un intervalle $[a ; b]$ de I est égale à l'aire sous la courbe de f sur $[a ; b]$, soit $\int_a^b f(t)\,dt$.

À noter
Une fonction de densité de probabilité est aussi appelée **fonction densité** ou **densité de probabilité**.

On en déduit que le domaine compris entre la courbe représentative de f et l'axe des abscisses a pour aire $P(X \in I)$, c'est-à-dire 1.

Propriétés Pour tous réels a et b appartenant à l'intervalle I :
(1) $P(X = a) = 0$
(2) $P(X > a) = 1 - P(X \leqslant a)$ et $P(a < X < b) = P(X < b) - P(X \leqslant a)$
(3) Pour une loi continue dans les calculs de probabilités, on peut remplacer des inégalités larges par des inégalités strictes et réciproquement. Par exemple : $P(X > a) = P(X \geqslant a)$.

À noter
$P(X = a) = \int_a^a f(t)\,dt = 0$

Espérance mathématique

À noter
La définition de l'espérance mathématique d'une variable aléatoire continue est à rapprocher de celle d'une

Définition Soit X une variable aléatoire continue de fonction de densité f sur l'intervalle $[a ; b]$, alors l'**espérance mathématique de X** est le réel défini par $E(X) = \int_a^b t f(t)\,dt$.

E.1.4 Cours du manuel *Math'x* (p. 402)

Cours

A Deux exemples de lois continues

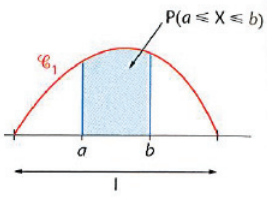
De nouveaux univers ...

Jusqu'à présent, une expérience aléatoire conduisait à un univers fini et une variable aléatoire X prenait un nombre fini de valeurs. La loi de probabilité de X se définissait alors par la donnée des probabilités de toutes ses valeurs ...

Mais il arrive aussi que les issues d'une expérience aléatoire ou les valeurs prises par une variable aléatoire puissent être n'importe quel nombre d'un intervalle I de $\mathbb{R}$ (par exemple, la durée T d'une communication).

Adaptation du modèle

Lorsqu'une variable aléatoire X prend pour valeurs les réels d'un intervalle I, sa loi de probabilité, dite « continue », n'est plus associée à la probabilité de chacune de ses valeurs mais à la probabilité de tout intervalle $[a\,;b]$ inclus dans I. C'est par un calcul d'aire sous la courbe d'une fonction propre à chaque loi, appelée « densité », que s'opèrent ces calculs. Une fonction de densité est une fonction continue, positive et dont l'aire totale sous sa courbe est égale à 1.



Note

La formule de l'espérance d'une variable aléatoire discrète :
$E(X) = \sum x_i p(X = x_i)$
s'adapte aussi au cas d'une variable aléatoire continue :
$E(X) = \int_a^b x f(x)\,dx$.

E.1.5 Cours du manuel *Odyssée* (pp. 394-399)

Cours

A. Variables aléatoires à densité

1 Variable aléatoire continue

DÉFINITION ET NOTATIONS

- L'ensemble des issues d'une expérience aléatoire s'appelle l'**univers** de l'expérience. On le note Ω.
- Une **variable aléatoire** est une fonction définie sur Ω et à valeurs dans $\mathbb{R}$. On la note X.

EXEMPLE 1

On lance deux dés cubiques équilibrés dont les faces sont numérotées de 1 à 6. On peut définir une variable aléatoire qui, à chaque lancer, associe la somme des nombres obtenus. Dans ce cas, l'univers Ω est l'ensemble des couples $(i\,;j)$ où i et j prennent des valeurs entières comprises entre 1 et 6. Comme cette variable aléatoire ne prend qu'un nombre fini de valeurs (entiers compris entre 2 et 12), elle est dite **discrète**.

DÉFINITION

Vocabulaire

Soit Ω l'univers d'une expérience aléatoire.
Une variable aléatoire X définie sur Ω, qui peut prendre comme valeur tous les nombres réels d'un intervalle I de $\mathbb{R}$, est dite **continue**.

EXEMPLE 2

Une entreprise produit des ampoules à basse consommation.
On peut définir une variable aléatoire X qui, à chaque ampoule produite par cette entreprise, associe sa durée de vie en heures. Cette durée n'est pas nécessairement un nombre entier d'heures et théoriquement, on ne connaît pas la durée maximale de vie d'une telle ampoule. Cette variable aléatoire X est donc continue et l'intervalle I est l'intervalle $[0\,;+\infty[$.

2 Fonction densité

Une fois la variable aléatoire définie, on s'intéresse à sa **loi de probabilité**.

- Dans le cas d'une variable aléatoire discrète, la loi de probabilité associée est généralement donnée sous la forme d'un tableau. Elle peut aussi être donnée par l'intermédiaire d'une formule générale ; c'est le cas pour les variables aléatoires qui suivent une loi binomiale.

EXEMPLE

Dans l'exemple 1, la loi de probabilité de la variable aléatoire discrète X est définie par le tableau suivant :

s	2	3	4	5	6	7	8	9	10	11	12
$P(X=s)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

On vérifie que la somme des probabilités des événements $(X=s)$ (avec $2 \leqslant s \leqslant 12$) est bien égale à 1.

■ Cours

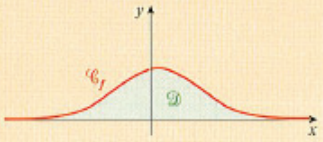
Vocabulaire

- Dans le cas d'une variable aléatoire continue, on a recours à une fonction définie sur $\mathbb{R}$ appelée **densité**.

NOTE
Dans le cas d'une variable aléatoire discrète, la somme des probabilités des événements $\{X = k\}$ est égale à 1. Dans le cas d'une variable aléatoire continue, cette propriété se traduit par l'aire du domaine $\mathcal{D}$ est égale à 1.

DÉFINITION

Une fonction f définie sur $\mathbb{R}$ est appelée fonction densité, ou **densité**, si :
- f est positive sur $\mathbb{R}$: pour tout nombre réel x, $f(x) \geqslant 0$;
- f est continue sur $\mathbb{R}$ (sauf éventuellement en un nombre fini de points) ;
- l'aire du domaine $\mathcal{D}$, domaine délimité par la courbe $\mathscr{C}_f$ représentative de la fonction f dans un repère orthogonal, et par l'axe des abscisses, est égale à 1.



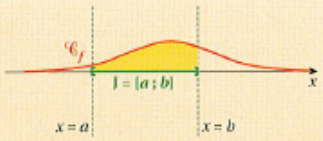
3 Probabilité d'un événement

NOTE
La fonction densité d'une variable aléatoire continue définit à elle seule la loi de probabilité de la variable aléatoire continue à laquelle elle est associée.

DÉFINITION

Soit Ω l'univers d'une expérience aléatoire et X une variable aléatoire définie sur Ω, continue de densité f.
La probabilité de l'événement $\{X \in J\}$, où J est un **intervalle** de $\mathbb{R}$, est notée $P(X \in J)$ et elle est définie comme l'aire du domaine suivant :

$$\{M(x\,;y)\,;\ x \in J \text{ et } 0 \leqslant y \leqslant f(x)\}$$



Aire du domaine jaune = $P(X \in J)$
$= P(a \leqslant X \leqslant b)$

REMARQUE

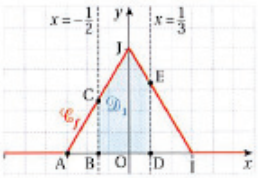
Par cette définition, la probabilité de l'événement $\{X \in \mathbb{R}\}$ correspond à l'aire du domaine $\mathcal{D}$, aire du domaine délimité par la courbe $\mathscr{C}_f$ et l'axe des abscisses. Par la troisième caractéristique de la densité f, cette aire est égale à 1. On a alors $P(X \in]-\infty\,;+\infty[) = 1$.

EXEMPLE

Soit X une variable aléatoire continue dont la densité f est définie sur $\mathbb{R}$ par :

$$f(x) = \begin{cases} 1+x & \text{si } -1 \leqslant x < 0 \\ 1-x & \text{si } 0 \leqslant x \leqslant 1 \\ 0 & \text{sinon} \end{cases}$$

On désire calculer la probabilité de l'événement $\left\{X \in \left[-\frac{1}{2}\,;\frac{1}{3}\right]\right\}$, ce qui se note $P\left(X \in \left[-\frac{1}{2}\,;\frac{1}{3}\right]\right)$ ou encore $P\left(-\frac{1}{2} \leqslant X \leqslant \frac{1}{3}\right)$. Par définition, cette probabilité correspond à **l'aire du domaine** $\mathcal{D}_1$ colorié en bleu sur la figure ci-contre où $\mathscr{C}_f$ est la courbe représentative de la densité f dans le repère orthogonal (O ; I, J).



En notant A, B, C, D et E les points de coordonnées respectives $(-1\,;0)$, $\left(-\frac{1}{2}\,;0\right)$, $\left(-\frac{1}{2}\,;\frac{1}{2}\right)$, $\left(\frac{1}{3}\,;0\right)$ et $\left(\frac{1}{3}\,;\frac{2}{3}\right)$, l'aire de ce domaine $\mathcal{D}_1$ est égale à l'aire du triangle AIJ à laquelle il faut soustraire l'aire du triangle ABC et celle du triangle DEI. Par suite, on a :

$$\text{Aire}(\mathcal{D}_1) = \text{Aire(AIJ)} - \text{Aire(ABC)} - \text{Aire(DEI)} = 1 - \frac{1}{8} - \frac{2}{9} = \frac{47}{72}$$

NOTE
Il est implicitement supposé que la densité f est continue sur l'intervalle $[a\,;\,b]$.

REMARQUES

- L'aire du domaine $\mathcal{D}_1$ est déterminée ici au moyen d'arguments géométriques. Mais on peut également déterminer cette aire à l'aide d'un calcul d'intégrale. En effet, comme la densité f est une fonction positive et **continue** sur l'intervalle $\left[-\frac{1}{2}\,;\,\frac{1}{3}\right]$, on a la relation suivante :

$$\text{Aire}(\mathcal{D}_1) = \int_{-\frac{1}{2}}^{\frac{1}{3}} f(x)\,\mathrm{d}x$$

Par linéarité de l'intégrale, il en découle :

$$\text{Aire}(\mathcal{D}_1) = \int_{-\frac{1}{2}}^{0}(1+x)\mathrm{d}x + \int_{0}^{\frac{1}{3}}(1-x)\mathrm{d}x = \left(\frac{1}{2}-\frac{1}{8}\right)+\left(\frac{1}{3}-\frac{1}{18}\right) = \frac{47}{72}$$

D'une manière générale, pour tous nombres réels a et b tels que $a \leqslant b$:

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x$$

- On constate graphiquement que la fonction f est bien positive et continue sur $\mathbb{R}$. En outre, l'aire du domaine $\mathcal{D}$, domaine délimité par la courbe $\mathcal{C}_f$ et par l'axe des abscisses, correspond ici à l'aire du triangle AIJ qui est égale à 1. On a ainsi vérifié que la fonction f est une densité.
- Par définition, la probabilité de l'événement $\{X \in [-1\,;\,1]\}$ est égale à l'aire du domaine délimité par la courbe $\mathcal{C}_f$, l'axe des abscisses et les droites d'équation $x = -1$ et $x = 1$. Ce domaine est le triangle AIJ dont l'aire est égale à 1. Ainsi, il est certain que X prend ses valeurs dans cet intervalle. On peut même d'ailleurs démontrer que l'intervalle $[-1\,;\,1]$ est le plus petit intervalle tel que la probabilité que X y prenne ses valeurs est égale à 1. En dehors de cet intervalle, on constate que la densité est nulle.

D'une manière générale, une variable aléatoire continue prend ses valeurs sur l'intervalle I où la densité ne s'annule pas. Cet intervalle I est appelé le **support de la densité.**

PROPRIÉTÉ

Soit k un nombre réel et X une variable aléatoire continue de densité f.
La probabilité que la variable aléatoire X prenne la valeur k est nulle : $P(X = k) = 0$.

DÉMONSTRATION

L'événement $\{X = k\}$ s'écrit aussi $\{X \in [k]\}$ ou encore $[X \in [k\,;\,k]]$. Par suite, $P(X = k) = P(X \in [k\,;\,k])$. Par définition, la probabilité de cet événement est alors l'aire du domaine délimité par l'axe des abscisses, la courbe $\mathcal{C}_f$ et *les droites* d'équation $x = k$. Par conséquent :

$$P(X = k) = \int_k^k f(x)\mathrm{d}x = 0 \quad \blacksquare$$

CONSÉQUENCE

Dans le cadre du calcul de la probabilité d'un événement $\{X \in \mathrm{J}\}$ où X est une **variable aléatoire continue** et J un intervalle de $\mathbb{R}$, les éventuelles inégalités larges peuvent être remplacées par des inégalités strictes et vice-versa.

EXEMPLE

On reprend l'exemple précédent. On désire calculer $P\left(X \in \left]-\frac{1}{2}\,;\,\frac{1}{3}\right]\right)$ et $P\left(-1 < X < -\frac{1}{2}\right)$.

- $P\left(X \in \left]-\frac{1}{2}\,;\,\frac{1}{3}\right]\right) = P\left(-\frac{1}{2} < X \leqslant \frac{1}{3}\right) = P\left(-\frac{1}{2} \leqslant X \leqslant \frac{1}{3}\right) = \text{Aire}(\mathcal{D}_1) = \frac{47}{72}$
- $P\left(-1 < X < -\frac{1}{2}\right) = P\left(-1 \leqslant X \leqslant -\frac{1}{2}\right) = \text{Aire(ABC)} = \frac{1}{8}$

Cours

4 Fonction de répartition

Vocabulaire

DÉFINITION

Soit X une variable aléatoire continue de densité f dont la courbe représentative est notée $\mathscr{C}_f$.
La **fonction de répartition** associée à la variable aléatoire X est la fonction F définie sur $\mathbb{R}$ par :

$$F(x) = P(X \leqslant x) = P(X \in]-\infty\,;\,x])$$

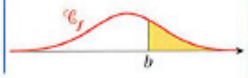
Aire du domaine jaune = $P(X \leqslant x)$
$= F(x)$

REMARQUES

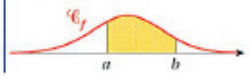
- Comme $P(X = x) = 0$ pour tout nombre réel x, on peut remplacer dans la définition précédente l'inégalité large par une inégalité stricte. Ainsi, pour tout réel x, $F(x) = P(X \leqslant x) = P(X < x)$.
- L'image d'un nombre réel x par la fonction de répartition F, $F(x)$, est naturellement un nombre positif (aire d'un domaine). En outre, comme l'aire du domaine délimité par la courbe $\mathscr{C}_f$ et l'axe des abscisses est égale à 1, ce nombre est inférieur ou égal à 1. Ainsi, pour tout nombre réel x, $0 \leqslant F(x) \leqslant 1$.
- La fonction de répartition est une fonction continue et croissante sur $\mathbb{R}$.
- La probabilité de l'événement $(X \in J)$ où J est un intervalle de $\mathbb{R}$ peut toujours s'exprimer à l'aide de la fonction de répartition. En effet :

- dans le cas où J est de la forme $]-\infty\,;\,b]$ ou $]-\infty\,;\,b[$, b étant un nombre réel, la probabilité correspondante est :
$P(X \leqslant b) = P(X < b) = F(b)$

- dans le cas où J est de la forme $[b\,;+\infty[$, la probabilité correspondante est :
$P(X \geqslant b) = P(X > b) = 1 - F(b)$



- dans le cas où J est de la forme $[a\,;\,b]$, la probabilité correspondante est :
$P(a \leqslant X \leqslant b) = P(a < X \leqslant b)$
$= P(a \leqslant X < b) = P(a < X < b)$
$= F(b) - F(a)$

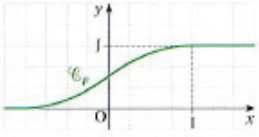


Ainsi, la loi de probabilité d'une variable aléatoire continue est entièrement déterminée **soit** par la connaissance de la fonction de densité f, **soit** par la connaissance de la fonction de répartition F. Cela s'explique par le lien fort entre ces deux fonctions. En effet, sur tout intervalle K où la fonction de répartition F est **dérivable**, on a pour tout nombre réel $x \in$ K, $F'(x) = f(x)$.

EXEMPLE

Dans le cas de l'exemple précédent, on peut démontrer que la fonction de répartition F est définie sur $\mathbb{R}$ par :

$$F(x) = \begin{cases} 0 & \text{si } x < -1 \\ \dfrac{(1+x)^2}{2} & \text{si } -1 \leqslant x < 0 \\ 1 - \dfrac{(1-x)^2}{2} & \text{si } 0 \leqslant x \leqslant 1 \\ 1 & \text{si } x > 1 \end{cases}$$



Sa courbe représentative $\mathscr{C}_F$ dans un repère orthogonal (O ; I, J) est donnée par la figure ci-dessus.
On constate que la fonction de répartition F est bien continue et croissante sur $\mathbb{R}$ et que pour tout nombre réel x, $0 \leqslant F(x) \leqslant 1$.

2 Espérance

Après la détermination de la loi de probabilité d'une variable aléatoire, s'ensuit généralement le calcul de certaines de ses caractéristiques telles que l'espérance.

- Dans le cas d'une variable aléatoire **discrète** Y, l'espérance notée $E(Y)$ est définie par :

$$E(Y) = \sum_{i=1}^{n} y_i P(Y = y_i)$$

où $(y_i)_{1 \leqslant i \leqslant n}$ désignent les n valeurs prises par la variable aléatoire Y.

EXEMPLE

Dans l'**exemple 1** de ce cours (le lancer de deux dés cubiques), l'espérance de la variable aléatoire considérée est égale à : $\frac{252}{36} = 7$.

L'interprétation de ce résultat est la suivante : si on lance les deux dés cubiques un grand nombre de fois, la somme des nombres obtenus est en moyenne égale à 7.

- Dans le cas d'une variable aléatoire **continue**, cette somme n'a pas de sens. En effet, une telle variable peut prendre comme valeur tous les nombres réels d'un intervalle I. On a alors recours à la définition qui suit, prolongement naturel du cas discret.

DÉFINITION

Soit a et b deux nombres réels distincts tels que $a < b$.
L'espérance d'une variable aléatoire continue X, **à valeurs dans l'intervalle** $\mathbf{I = [a\,;\,b]}$, dont la densité associée est notée f, est définie par :

$$E(X) = \int_a^b x f(x)\,\mathrm{d}x$$

REMARQUE

Cette définition de l'espérance est uniquement valable pour les variables aléatoires continues dont le support de la densité associée est un intervalle de la forme $[a\,;\,b]$, a et b étant deux nombres réels tels que $a < b$. C'est le cas pour les variables aléatoires continues qui suivent une loi uniforme.

E.1.6 Cours du manuel *Repères* (p. 336)

4. Lois de probabilités continues

Dans cette partie, l'univers Ω est un intervalle de $\mathbb{R}$, a et b sont deux réels tels que $a < b$.

Préliminaires

- Si $I = [a\,;\,b]$, alors $\int_I f(t)\,dt$ désigne $\int_a^b f(t)\,dt$.
- Si $I = [a\,;\,+\infty[$, alors $\int_I f(t)\,dt$ désigne $\lim\limits_{x \to +\infty} \int_a^x f(t)\,dt$.

définition

Soit I un intervalle. On appelle **densité de probabilité** sur I toute fonction f continue et positive sur I telle que :

$$\int_I f(t)\,dt = 1.$$

définition

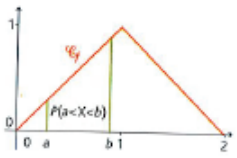
Soit I un intervalle. On dit qu'une variable aléatoire X, à valeurs dans I, suit une **loi de probabilité** de densité f sur I, si pour tout intervalle J inclus dans I :

$$P(X \in J) = \int_J f(t)\,dt.$$

Remarques

- Une telle variable X est dite **continue**.
- $P(X \in [a\,;\,b])$ pourra être noté $P(a \leqslant X \leqslant b)$.
- $P(a \leqslant X \leqslant b)$ est donc égale à l'aire du domaine compris entre $\mathscr{C}_f$, l'axe des abscisses et les droites d'équation $x = a$ et $x = b$.

On a représenté ci-contre la densité de probabilité de la variable aléatoire étudiée à la découverte 3..



propriétés

Soit X une variable aléatoire suivant une loi de probabilité de densité f sur I.

1. $P(X \in I) = 1$.
2. Quel que soit x_0 dans $\mathbb{R}$, $P(X = x_0) = \int_{x_0}^{x_0} f(t)\,dt = 0$.
3. $P(X \in [a\,;\,b]) = P(X \in]a\,;\,b]) = P(X \in [a\,;\,b[) = P(X \in]a\,;\,b[)$ et $P(X > a) = P(X \geqslant a)$.

Remarques

- $P(X \in [a\,;\,b]) = P(a \leqslant X \leqslant b)$, $P(X \in]a\,;\,b]) = P(a < X \leqslant b)$, etc.
- On admet pouvoir étendre certaines propriétés sur les lois de probabilités discrètes aux lois continues, en particulier :

$$P_{(X \in J)}(X \in K) = \frac{P((X \in J) \cap (X \in K))}{P(X \in J)}.$$

E.1.7 Cours du manuel *Symbole* (p. 390)

A Loi à densité

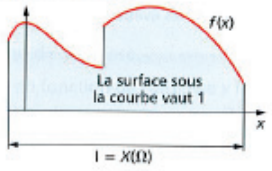
Considérons une expérience aléatoire dont l'univers associé est muni d'une loi de probabilité P.
Nous envisageons ici le cas d'une variable aléatoire prenant ses valeurs dans un intervalle I de $\mathbb{R}$.

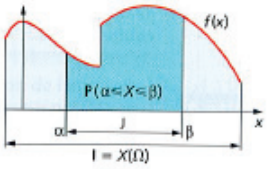
Définition 1

Note : *Les surfaces sont ici calculées en unités d'aire.*

Soit X une variable aléatoire définie sur un univers Ω, à valeurs dans l'intervalle I de $\mathbb{R}$. On admettra qu'il existe une fonction f définie sur $\mathbb{R}$ et nulle en dehors de I telle que pour tout intervalle J inclus dans I, la probabilité de l'événement « $X \in J$ » soit l'aire du domaine défini par : $\{M(x\,;y)\,;\, x \in J \text{ et } 0 \leqslant y \leqslant f(x)\}$.
On dira alors que la fonction f est une densité de la variable aléatoire X.

REMARQUES : On admet que f est une fonction densité si elle satisfait aux deux conditions suivantes : • la fonction f est positive ;
• la surface du domaine situé entre la courbe et l'axe des abscisses vaut 1 (car « $X \in I$ » est un événement certain).





Théorème 1

Note : *Attention, si $\alpha = \beta$, la probabilité pour que $X = \alpha$ est nulle.*

Si X est une variable aléatoire de densité f et si f est continue sur $J = [\alpha\,;\beta]$, alors :
$P(\alpha \leqslant X \leqslant \beta) = \int_\alpha^\beta f(t)\,dt$.

REMARQUE : Ce théorème est une conséquence des propriétés sur l'intégration vues au chapitre 7. Dans le cas de la figure 2, le découpage du segment J en deux segments sur lesquels la fonction f est continue permet le calcul de l'aire grâce aux intégrales.

Définition 2

Note : *Il est usuel d'appeler valeur moyenne de X l'espérance de X.*

Soit X une variable aléatoire définie sur Ω, de densité f, et $I = X(\Omega)$.
• Si $I = [\alpha\,;\beta]$ et f continue sur $[\alpha\,;\beta]$, alors X admet pour espérance :
$E(X) = \int_\alpha^\beta t \times f(t)\,dt$.
• Si $I = [\alpha\,;+\infty[$ et f continue sur $[\alpha\,;+\infty[$, alors X admet pour espérance :
$E(X) = \lim\limits_{x \to +\infty} \int_\alpha^x t \times f(t)\,dt$.
• Si $I =]-\infty\,;\beta]$ et f continue sur $]-\infty\,;\beta]$, alors X admet pour espérance :
$E(X) = \lim\limits_{x \to -\infty} \int_x^\beta t \times f(t)\,dt$.
• Si $I = \mathbb{R}$ et f continue sur $\mathbb{R}$, alors X admet pour espérance :
$E(X) = \lim\limits_{x \to -\infty} \int_x^0 t \times f(t)\,dt + \lim\limits_{x \to +\infty} \int_0^x t \times f(t)\,dt$.

Ces définitions ne valent que si les limites existent.

REMARQUE : Ces trois dernières limites sont souvent notées : $\int_\alpha^{+\infty}$, $\int_{-\infty}^\beta$ et $\int_{-\infty}^{+\infty}$.

E.1.8 Cours du manuel *Transmath* (p. 380)

COURS

1 Lois de probabilité à densité

1.1 Introduction

• Dans les situations étudiées jusqu'à présent, à une expérience aléatoire on associait un univers fini U muni d'une loi de probabilité P. Toute variable aléatoire X ne prenait alors qu'un nombre fini de valeurs $x_1, x_2, \ldots, x_k$.

• Cependant, certaines expériences aléatoires conduisent à utiliser des variables aléatoires qui prennent toutes les valeurs d'un intervalle I de $\mathbb{R}$.

Exemple

1. On tire sur une cible circulaire de 1 m de rayon, sans jamais la manquer. La variable aléatoire qui indique la distance, en mètre, du point d'impact au centre prend toutes les valeurs de $[0\,;1]$.

2. Lors d'un appel téléphonique, la durée de communication est une variable aléatoire qui prend ses valeurs dans un intervalle de temps.

• Dans ces conditions, il n'est plus possible de définir la loi de X en dressant le tableau des probabilités de chacun des événements « $X = x_i$ » puisqu'il y en a une infinité.

Une autre approche est alors nécessaire. On s'intéresse aux événements du type « **X prend ses valeurs dans un intervalle J** » noté « $X \in J$ ».

Il s'agit alors de définir la probabilité $P(X \in J)$.

La notation « $X \in J$ » est abusive car X n'est pas un nombre. Elle prolonge la notation « $X = a$ » déjà utilisée.

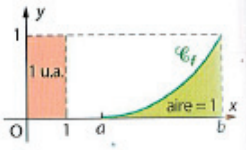
1.2 Variable aléatoire suivant une loi à densité

Dans ce paragraphe, I désigne un intervalle I (borné ou non) de $\mathbb{R}$.

Définition 1 On appelle **densité de probabilité sur I** toute fonction *f* définie sur I telle que :

• *f* est continue et positive sur I ;

• l'aire sous la courbe $\mathscr{C}_f$ est égale à 1 u.a.

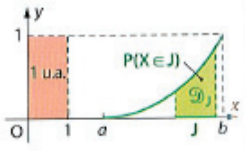


Définition 2 *f* désigne une densité de probabilité définie sur I.

Dire qu'**une variable aléatoire X suit la loi de densité *f*** signifie qu'à tout intervalle J inclus dans I, on associe la probabilité :

$$P(X \in J) = \text{aire}(\mathscr{D}_J)$$

où $\mathscr{D}_J$ est le domaine sous la courbe $\mathscr{C}_f$ sur l'intervalle J.



Remarque. Pour tous intervalles J et K inclus dans I :

• $P(X \notin J) = 1 - P(X \in J)$.

• $P(X \in J \cup K) = P(X \in J) + P(X \in K) - P(X \in J \cap K)$.

En particulier, lorsque J et K sont disjoints : $P(X \in J \cup K) = P(X \in J) + P(X \in K)$.

Conséquence. D'après la définition 2, si $J = [\alpha\,;\beta]$ avec $\alpha \leqslant \beta$, alors $P(X \in J) = \int_\alpha^\beta f(x)\,dx$.

Si $\alpha = \beta$, J est réduit à la seule valeur α et $P(X = \alpha) = \int_\alpha^\alpha f(x)\,dx = 0$.

Ainsi, la probabilité que X prenne une valeur isolée est nulle ;

d'où $P(X \leqslant \alpha) = P(X < \alpha)$.

« $X \in J$ » se note aussi « $\alpha \leqslant X \leqslant \beta$ ».

E.2 *Activités* d'introduction de la notion de fonction de densité

E.2.1 *Activité* 2 pp. 386-387 du manuel *Symbole*

Activité 2 **Approche de la fonction densité et de la fonction de répartition**

OBJECTIF
Découvrir : la fonction densité et la fonction de répartition dans le cas d'une variable aléatoire non discrète.

A. Construire le polygone des fréquences cumulées croissantes et la courbe « densité »

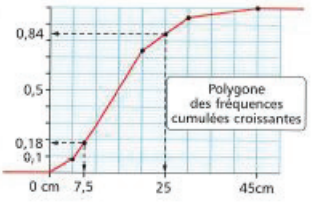
Un tireur effectue 50 tirs. La répartition des distances au centre de la cible des 50 impacts de balles est donnée ci-dessous :

Classe	de 0 à 5 cm	de 5 à 10 cm	de 10 à 20 cm	de 20 à 30 cm	de 30 à 45 cm
Effectifs	4	10	23	10	3

1 Copier et compléter le tableau suivant :

Numéro	k	1	2	3	4	5
Classe	$[a_k, b_k[$	[0, 5[	[5, 10[	[10, 20[	[20, 30[	[30, 45]
Amplitude	Δ_k	5 cm				
Effectifs	n_k	4				
Effectifs cumulés croissants	N_k	4				
Fréquences	f_k	0,08				
Fréquences cumulées croissantes	F_k	0,08				

2 Comment construit-on à partir de ce tableau le polygone des fréquences cumulées croissantes, représenté sur la figure ci-contre ?



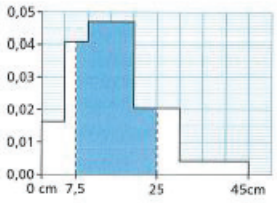
Note

Suivant l'usage, on admet qu'à l'intérieur d'une classe l'effectif est uniformément réparti ce qui permet de passer d'une fréquence cumulée à la suivante par un segment de droite.

3 On ajoute une nouvelle ligne au tableau de la question **1**. Pour chaque classe k, on introduit la **densité** d_k égale au quotient de la fréquence de la classe par son amplitude : $d_k = \frac{F_k}{\Delta_k}$.

a. Vérifier que cette densité est la pente du segment représentant la classe sur le polygone des fréquences cumulées croissantes. En quoi peut-on dire qu'il s'agit de la dérivée de la fonction représentée par la courbe des fréquences cumulées croissantes ?

b. La figure ci-contre représente la densité d_k pour chaque classe k. Quelle est la relation entre la fréquence de chaque classe et l'aire du rectangle ayant pour base la classe considérée et pour hauteur la densité ?



B. Calculer des probabilités

1 On choisit au hasard l'un des 50 tirs et l'on note $X(\omega)$ la distance de l'impact ω de ce tir au centre de la cible.

a. Comment déterminer la probabilité $P(7{,}5 \leqslant X \leqslant 25)$, à partir du polygone des fréquences cumulées croissantes ?

b. Calculer l'aire bleue du domaine situé sous l'histogramme et limité par les droites d'équation $x = 7{,}5$ et $x = 25$.

c. Quelle est la relation entre ces deux résultats ?

d. Calculer de même les probabilités $P(0 \leqslant X \leqslant 5)$ et $P(10 \leqslant X \leqslant 20)$.

2 Un tir ω est jugé mauvais si $X(\omega) > 20$ et très mauvais si $X(\omega) > 30$.

a. Quelle est la probabilité qu'un tir soit mauvais ?

b. Quelle est la probabilité qu'un mauvais tir soit très mauvais ?

Conclusion

La courbe représentative de la densité permet de calculer la probabilité de l'événement $a \leqslant X \leqslant b$ comme l'aire du domaine situé entre l'axe des abscisses et la courbe, limité à gauche et à droite par les droites d'équations $x = a$ et $x = b$. La courbe de répartition permet de calculer la probabilité de l'événement $a \leqslant X \leqslant b$ comme la différence $F(b) - F(a)$.

E.2.2 *Activité* 1 pp. 398 du manuel *Math'x*

1 Loi uniforme et densité

TICE

OBJECTIF
Prolonger la notion d'équiprobabilité au cas où l'univers des possibles est un intervalle.
Introduire la notion de densité de probabilité.

Le tirage au hasard d'un réel x dans l'intervalle [0 ; 10] paraît difficile à simuler !
Par exemple, la fonction « 10*ALEA() » d'un tableur renvoie un nombre dans [0 ; 10] qui est en fait un décimal, comportant 14 chiffres après la virgule.
On étudie ici quelle doit être la loi de probabilité P d'une variable aléatoire correspondant au choix au hasard d'un nombre réel dans [0 ; 10].

A. Probabilité d'un réel

1. Peut-on définir P en se donnant $P(X = x)$ pour chaque réel x de [0 ; 10] ?

2. Pour traduire « le choix **au hasard** d'un réel », on impose à $P(X = x)$ de prendre une même valeur α, avec $\alpha \geqslant 0$, quelque soit x.

a. Soit E_n l'événement : « x est un décimal non nul comportant au plus n décimales ».
Justifier qu'il existe 10^{n+1} tels décimaux.
En déduire $P(E_n)$ en fonction de n et α.

b. On suppose que le réel α n'est pas nul. En quoi la valeur de $P(E_n)$ pose-telle problème pour n suffisamment grand ? Quelle valeur le réel α doit-il prendre ?

B. Probabilité d'un intervalle

On simule sur un tableur 10 000 tirages d'un nombre x dans [0 ; 10[, à l'aide de la fonction « 10*ALEA() » et on fait afficher les fréquences des événements « $x \in [\sqrt{2} ; \pi]$ » et « $x \in [3 + \sqrt{2} ; 3 + \pi]$ ».

Télécharger le fichier sur le site pour d'autres simulations par F9.

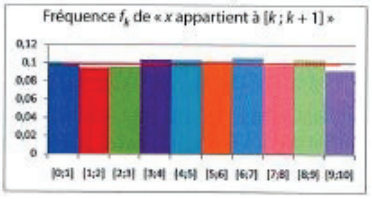
	A	B	C	D
1	événement	fréquence	événement	fréquence
2	rac2 <= x <= pi	0,17402	3+rac2 <= x <= 3+ pi	0,16790

a. Quelle observation peut-on faire sur les deux fréquences affichées ?

b. Quelle similitude des intervalles $[\sqrt{2} ; \pi]$ et $[3 + \sqrt{2} ; 3 + \pi]$ peut expliquer cela ?

C. Probabilité et calcul d'aire

L'histogramme ci-dessous représente les fréquences f_k des dix événements « $x \in [k ; k + 1]$ » pour 10 000 réalisations de la fonction « 10*ALEA() ».



1. Conjecturer $P(X \in [k ; k + 1])$, pour k entier, $0 \leqslant k \leqslant 9$.

2. Le tableur affiche une courbe rouge, appelée « courbe de tendance », qui « lisse » l'histogramme. Préciser quelle fonction f elle représente.

3. a. Que vaut nécessairement $P(X \in [0 ; 10])$? Comparer avec $\int_0^{10} f(x)\,dx$.

b. En traduisant que, pour $[a ; b] \subset [0 ; 10]$, la probabilité $P(X \in [a ; b])$ ne dépend que de la longueur $b - a$ de $[a ; b]$, à laquelle elle est proportionnelle, proposer une formule calculant $P(X \in [a ; b])$ en fonction de a et b. Comparer avec $\int_a^b f(x)\,dx$.

c. Appliquer ces deux démarches de calcul à $P(X \in [\sqrt{2} ; \pi])$. Indiquer si le résultat paraît en accord avec la fréquence observée dans le tableau de la partie B.

E.2.3 *Activité* 3 p. 329 du manuel *Repères*

3. TICE Vers la densité d'une loi de probabilité

Soit X et Y deux variables aléatoires indépendantes prenant des valeurs uniformément réparties dans l'intervalle [0 ; 1]. Soit S la variable aléatoire associant au couple (X ; Y) la somme X + Y.
On se propose de déterminer la probabilité des événements $(S \leqslant 1)$, $(S < 0{,}5)$.
Pour répondre à ces questions, nous allons utiliser une démarche expérimentale, puis un nouvel outil : la densité de probabilité d'une variable aléatoire continue.

1. Partie *Simulation*

a. Quel est l'ensemble des valeurs possibles de la variable aléatoire S ?

b. À l'aide du logiciel *XCAS*, utiliser la commande **EchX :=randvector(1000,0..1)** pour simuler un échantillon, appelé EchX, de 1 000 valeurs suivant la loi de X. Répéter l'opération pour simuler un échantillon de 1 000 valeurs suivant la loi de Y. Enfin, créer un échantillon de 1 000 valeurs suivant la loi de S en utilisant la commande **EchS :=EchX+EchY**.

c. Utiliser la fonction « classes » pour déterminer les effectifs d'une classe. En déduire les fréquences correspondantes pour compléter le tableau ci-dessous :

Événement	$S \leqslant 1$	$S < 0{,}5$	$S > 1{,}2$	$0{,}6 \leqslant S \leqslant 1{,}2$	$S \leqslant 2$
Fréquence					

d. Comparer ses résultats avec ceux de son voisin. Que remarque-t-on ?

2. Partie *Modélisation*

On souhaite approcher la loi de S graphiquement à l'aide d'un histogramme. Le programme *dec3()* ci-contre permet de représenter l'histogramme des fréquences d'un échantillon simulé par l'ordinateur.

RAPPEL
Dans un histogramme, l'aire est proportionnelle à l'effectif des classes.

```
Prog  Edit  Ajouter
dec3():= {
local taille,X,Y,S,G;
saisir(taille);
X:=randvector(taille,0..1);
Y:=randvector(taille,0..1);
S:=X+Y;
G:=histogram(S,0,0.1);
return(G);
}
:;
```

a. Recopier ce programme et le tester pour des tailles d'échantillon de 1 000, 10 000, 100 000. Que remarque-t-on ?

b. Modifier le programme précédent en remplaçant la ligne « G := ... », par « G :=histogram (S,0,0.1),plot(x,x=0..1),plot(2-x,x=1..2) ; » pour qu'il affiche en même temps l'histogramme et la courbe $\mathscr{C}_f$ représentative de la fonction f définie par :

$$f : x \mapsto \begin{cases} x & \text{si } 0 \leqslant x \leqslant 1 \\ 2 - x & \text{si } 1 \leqslant x \leqslant 2 \end{cases}$$

c. Que penser de cette courbe ?

3. Utilisation du modèle

D'après la partie 2., il paraît légitime d'utiliser la fonction f, appelée **densité** de la variable aléatoire S, pour calculer la probabilité de certains événements.

a. Calculer l'aire en unités d'aires, située sous cette courbe $\mathscr{C}_f$.

b. Rappeler l'interprétation géométrique de $\int_0^{0{,}5} f(t)\,dt$. En déduire un lien entre $\int_0^{0{,}5} f(t)\,dt$ et le calcul des fréquences de la partie 1. pour $S < 0{,}5$.
On admet que pour tous réels a et b de l'intervalle [0 ; 2], $p(a \leqslant S \leqslant b) = \int_a^b f(t)\,dt$.

c. Compléter le tableau suivant en utilisant l'égalité ci-dessus.

Événement	$S \leqslant 1$	$S < 0{,}5$	$S > 1{,}2$	$0{,}6 \leqslant S \leqslant 1{,}2$	$S \leqslant 2$
Probabilité					

d. Comparer les tableaux des questions 1. d. et 3.c. Conclure.

E.3 Exercices

B Fonction $x \mapsto e^{-\frac{x^2}{2}}$

On considère la fonction f définie sur $\mathbb{R}$ par $f(x) = e^{-\frac{x^2}{2}}$
et Γ sa courbe représentative dans le plan muni d'un repère orthogonal $(O ; \vec{i}, \vec{j})$.
Pour tout a réel, on pose $I(a) = \int_0^a e^{-\frac{x^2}{2}}\, dx$.

1. Étude de f

a. Donner le sens de variation de la fonction f et sa limite en $+\infty$.

b. Justifier que $I(a)$ existe et l'interpréter graphiquement.

On ne connaît pas d'expression explicite d'une primitive de f et donc de $I(a)$, mais des logiciels peuvent permettre d'explorer $I(a)$.

2. Exploration sur GeoGebra *(on peut utiliser en vidéo projection le fichier disponible sur le site)*

a. Dans le menu *Options, Arrondi*, choisir *5 décimales*.

Créer la fonction f.

Créer un curseur a de 0 à 50.

Créer l'intégrale $I(a)$, notée I, en entrant dans la barre de saisie la commande : `I=Intégrale[f,0,a]`

b. À l'aide du logiciel, donner une valeur approchée à 10^{-3} près de $I(0{,}5)$, $I(1)$, $I(3)$, $I(10)$.

c. Conjecturer le comportement de $I(a)$ quand a tend vers $+\infty$.

d. Un logiciel de calcul formel donne $\lim\limits_{a \to +\infty} I(a) = \sqrt{\frac{\pi}{2}}$.

Confronter ce résultat à la conjecture émise et interpréter graphiquement cette limite.

3. D'une aire à une autre

a. Justifier que Γ est symétrique par rapport à l'axe des ordonnées.

b. Déterminer une valeur approchée à 10^{-3} près de l'aire, en u. a., de chacune des parties suivantes du plan $\mathcal{P}$:

$E_0 = \{M(x ; y) \in \mathcal{P} ;\ 0 \leqslant y \leqslant f(x)\}$

$E_1 = \{M(x ; y) \in \mathcal{P} ;\ -1 \leqslant x \leqslant 0 \text{ et } 0 \leqslant y \leqslant f(x)\}$

$E_2 = \{M(x ; y) \in \mathcal{P} ;\ -1 \leqslant x \leqslant 1 \text{ et } 0 \leqslant y \leqslant f(x)\}$

$E_3 = \{M(x ; y) \in \mathcal{P} ;\ x \leqslant -1 \text{ et } 0 \leqslant y \leqslant f(x)\}$

$E_4 = \{M(x ; y) \in \mathcal{P} ;\ x \geqslant 1 \text{ et } 0 \leqslant y \leqslant f(x)\}$

$E_5 = \{M(x ; y) \in \mathcal{P} ;\ x \geqslant -1 \text{ et } 0 \leqslant y \leqslant f(x)\}$

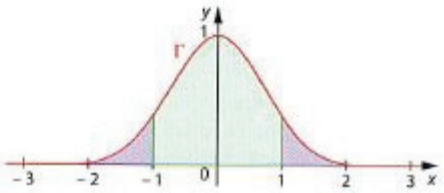


Figure E.1 – Partie B du TP1 p. 244 du manuel *Math'x*

Annexe F

Les entretiens

F.1 Les questions de l'entretien

Entretien enseignant

Informations personnelles
Depuis combien de temps enseignez-vous ? En TS ?

Calcul intégral

- Introduction de l'intégrale
- Place de l'aire
- Aire finie d'un domaine infini
- Les TICE pour le calcul exact d'intégrale
- Les TICE pour le calcul approché d'intégrale
- Exemples d'utilisation
- A savoir pour ce chapitre ?
- Applications du calcul intégral dans ce chapitre ?

Lois de probabilité à densité (lois uniformes, lois exponentielles, lois normales)

- Introduction lois à densité (histogramme?)
- Fonction de densité : introduction, définition générale
- Place des graphiques, aire sous la courbe
- Lois sur un intervalle non borné
- Que doivent savoir les élèves ? Formules par cœur ? Lesquelles ? Apprendre/voir ?
- Dans les preuves
- Choix des exemples : importance ?
- Autres lois à densité ? Lesquelles ?
- Les TICE pour le calcul de probabilité
- Exemples d'utilisation

Liens entre calcul intégral et lois à densité

- Liens entre calcul intégral et proba ?
- Contradiction/ Non cohérence entre CI et P ?
- Difficultés des élèves en lois à densité (des problèmes de calcul d'intégrales / des conceptions en proba / de la représentation graphique / de l'utilisation des TICE / des différences de « langages » entre les deux chapitres)
- Méthode de Monte-Carlo
- Autres

F.2 L'entretien avec Marie

Entretien avec l'enseignante Marie, le 26 juin 2014.
Avec ses notes personnelles de cours.

Marie : Bon alors le calcul intégral…
Chercheur : Voilà. J'ai regardé ce que tu avais fait… Mais… Toi comment tu vois l'introduction de l'intégrale ?
Marie : L'idée c'est vraiment de partir sur du calcul d'aire plane… Donc l'aire sous une courbe… J'avais une petite introduction qui ne figurait pas, qui était sur les vitesses moyennes… En partant d'un problème qu'ils avaient en physique. Et donc on a été amené à se poser la question de « comment on va calculer l'aire sous une courbe d'une fonction continue positive ? ». Donc voilà… J'ai posé le problème. Après on a, on a regardé les cas simples donc les fonctions affines, affines par morceaux. Ce qui permet de remettre au goût du jour les formules de calcul d'aires, les histoires d'unités d'aire, etc.… Et puis après, comme tu peux voir, et bien, je leur demande à chacun d'inventer, de dessiner une jolie fonction… définie, continue, positive sur un intervalle $[a; b]$, qu'on s'était fixé… Et puis, à charge… À leurs camarades de calculer l'aire sous la courbe qu'ils ont proposée. Donc, et puis après… Je passe dans les rangs et… Je pique les exemples qui m'intéressent, des courbes qu'ils ont proposées… Donc … Une majorité des élèves ont reproposé des fonctions affines par morceaux, mais avec plein de petits morceaux… Pour compliquer la tâche de leurs camarades. Certains vont faire des demi-cercles, enfin des arcs de cercle… Généralement c'est des exemples que je prends. Puis certains proposent la parabole, classiquement, la fonction exponentielle ou log suivant la dernière qu'on vient de faire. Et donc du coup on bascule sur ces exemples là… pour voir comment on peut s'y prendre… pour calculer l'aire sous… quand voilà on ne voit, reconnaît pas … un domaine, une surface qu'on connaît et dont on sait déjà calculer l'aire. Donc … Après ça dépend complètement d'eux. Donc moi j'ai plein de documents à disposition, la courbe que je choisis, c'est soit la parabole soit … cette année… Je sais pas si je me souviens… Ça doit être… Comme j'ai deux classes de terminale, j'ai pris une courbe et l'autre une autre… Enfin peu importe. Disons que je le fais avec la parabole sur $[0; 1]$ et donc j'ai plusieurs documents à… à la disposition des élèves, mais c'est eux qui viennent demander, je ne leur montre pas. Y'a des courbes qui sont déjà tracées sur papier blanc… uni, il y en a sur papier quadrillé qui sont déjà… Tu l'as pas là, mais qui sont déjà quadrillés. Et puis voilà… Ils font des propositions, qu'ils cherchent à deux et ils font des propositions et très, assez rapidement les encadrements arrivent très vite, des affinements d'encadrements… et très vite la méthode des rectangles en fait… Ça vient chaque année, de leur part… L'année dernière, j'avais aussi la méthode des trapèzes… Donc voilà. Donc après on s'y colle… à faire nos petits rectangles et donc voilà, ça donne le document que tu as là [*cf. documents de cours de Marie*] parce que une fois que tout le monde a… Une fois qu'un groupe d'élèves a proposé l'idée… Et que tout le monde l'a comprise et bien, là on y va ensemble avec une feuille qui est… Alors généralement, je ne leur donne pas tout de suite parce que ça dépend comment ils réagissent. S'ils se débrouillent tous seuls, on décide de subdiviser en 10 parties égales…

Et là, ils font à la main sans mon aide… Voilà… Après on fait pour… On voit comment ça se passe pour 100. Généralement, moi je l'ai calculé pour 1000, pour leur donner un nouvel encadrement, parce que là, ils en ont marre quand même. [*Rires*].
Chercheur : Ils le font à la main les… pour 100 ?
Marie : Pour 10. Non, pour 100, ils comprennent tous seuls comment ils vont le… Comment la formule… Parce qu'on l'écrit bien, de façon à ce qu'on puisse la généraliser en fait. Et du coup, ils comprennent tous seuls ce qui va se passer pour 100, ils voient aussi que entre les deux la somme des aires inférieures et des aires supérieures, ils voient bien que la différence… y a pas grand-chose à faire… Donc tout ça, ça vient d'eux, du coup, les autres sont… Enfin voilà… Tout le monde « pige » ça assez vite… Et du coup, après, par contre, je leur donne pour le cas général, généralement le polycopié en leur disant « ça c'est ce qu'on a déjà, c'est ce qu'on a déjà fait ». Et puis, du coup, là on passe à la limite des deux suites. Avec des notations qu'on garde après dans tout le cours, pour s_n et S_n, quand on va faire la méthode des rectangles dans d'autres cas. Alors, y a plus les suites adjacentes, voilà… Donc, donc c'est un peu plus… Enfin c'est un peu moins bien posé qu'avant à la fin, quand même… Mais ils comprennent très bien que la suite s_n, là, elle va être croissante, l'autre décroissante, que la différence va tendre vers zéro, donc qu'en gros sans le dire on a tout ça et on en déduit la valeur de l'aire…. Pour l'instant comme ça… Du coup, du coup voilà… On fait une synthèse du TP et… Comme je t'ai écrit là, en disant que l'on va admettre que dans tous les cas, quand la fonction et dans les bonnes conditions, ces deux suites de rectangles, enfin d'aires de rectangles, vont… vont bien encadrer notre aire et vont bien converger vers le même nombre…
Chercheur : OK…
Marie : Donc voilà…
Chercheur : Et puis après…
Marie : Après, après voilà… On part sur le cours…
Chercheur : On met en place les…
Marie : On définit… enfin plus, enfin voilà… La définition telle qu'elle est au programme pour une fonction continue positive : l'aire sous une courbe. Alors après ça prend pas mal de temps d'expliciter la notation. Ça, ça… ça les… ça les interroge toujours un petit… Une bonne semaine, quand même…
Chercheur : Oui le dx…
Marie : Le dx… Même si avec l'aire des rectangles, en fait, c'est assez… C'est assez… C'est assez évident pour eux mais voilà, c'est une notation qu'ils ont quand même au début du mal un petit peu… C'est-à-dire que le fait que cette chose, absolument barbare, soit en fait un nombre réel, ça les… ça les laisse un peu songeur, pendant un bon moment quand même…
Chercheur : Oui, ça…
Marie : Donc oui.
Chercheur : Ils prennent ça plutôt pour une fonction ? Ou même pas ?…
Marie : Non même pas… Ils prennent surtout pas ça pour une fonction. Non, ils prennent ça pour un truc qui a l'air comme ça d'être affreux en fait. Je crois que la notation les bloque… Et… que ce soient les ES ou S… puisqu'en fait les ES, le début du cours, c'est le même. C'est… Les ES acceptent plus la notation, en se disant: « Oh c'est encore un truc

bizarre du prof de maths… mais c'est pas grave ! On a su calculer les aires sous la courbe, on est content ». Et les TS s'interrogent un peu plus sur « pourquoi tout ça, puisque c'est une aire ? ».
Chercheur : Oui.
Marie : Ça, voilà… Mais bon… Ça passe de toute façon au bout d'un moment…
Chercheur : Et du coup, pour la méthode des rectangles, ils font tout au papier-crayon ?
Marie : Ah non non non non… Y a… y a, bien sûr, des… un truc sur TICE. Mais moi je suis pas dans la salle info tout le temps donc, là, moi j'avais préparé… la même chose sur GeoGebra, et bien sûr on a fait tendre le nombre des rectangles… Devenir très très grand… Et on voit que les… que les aires inférieures, qui n'étaient pas colorées de la même façon que les autres, ça se rejoint bien… Non, non… On a tout fait. Et là, je m'étais arrangée pour que cette séance-là soit le jour où j'étais avec un ordinateur et un vidéoprojecteur. Donc ça ils ont vu sur écran. Et les ES aussi, ils ont vu sur écran.
Chercheur : Et, tu penses que… c'est important ?
Marie : Ça reste dans la tête quand même. Oui… Tout ce... Cette partie-là, comme... le théorème de Moivre-Laplace… [Inaudible] Le fait d'avoir une animation, ça reste bien quand même visuellement, dans leur tête… Ce n'est pas eux qui l'ont fait. C'est moi qui avais tout fait. Mais voilà, c'est très parlant quand même pour eux.
Chercheur : Le fait de voir…
Marie : D'abord, ça bouge oui, c'est animé. On avait retrouvé les valeurs qu'ils avaient calculées… puisque les sommes, tu les fais afficher. Donc oui, oui… c'était… D'ailleurs, après on y fait référence à chaque fois. « Vous voyez bien... ». « Vous vous rappelez ce qui se passait pour les rectangles ? ». Ça, c'est… C'est des images, en fait.
Chercheur : Oui. Du coup, dans un premier temps, ils ont fait sur papier crayon, et puis, au moment de la synthèse, avec GeoGebra. Ou après ?
Marie : Oui… Non. À la fin de cette séance-là, on avait fait… Moi, j'avais… Une fois qu'ils avaient compris le truc, je leur montrais au vidéoprojecteur ce qu'il se passait… pour afficher les valeurs pour 1000. Même si, moi, je les avais calculées avant. Et puis pour leur montrer, qu'effectivement, quand on faisait tendre n vers plus l'infini, et voilà la différence, elle devenait minime. Les ES aussi, ça leur a… Ça leur a été utile d'avoir une image mentale comme ça… de bien voir les rectangles se resserrer… On voit bien ce qui se passe.
Chercheur : D'accord. Et au niveau de… Donc on voit bien que l'introduction, c'est avec de l'aire. Et après, tout au long du chapitre, c'est… quelle est la place de l'aire ?
Marie : Eh bien … Alors, comme tu peux le voir… On continue par faire plusieurs exemples, qui sont tous à base de calcul d'aire, déjà. Puisqu'on n'a pas encore… C'est-à-dire qu'on n'a pas les primitives, donc on a un moment on est… on reste dessus. Donc y a des exemples dans les exercices qui sont : « on applique la méthode des rectangles à d'autres courbes que celles qu'on a choisies dans l'introduction ». Nous, c'est avec un pas de $n = 4$. Enfin pas très gros, parce que là, il faut qu'ils le fassent à la main. Mais voilà… Donc la méthode des rectangles, elle est bien… Après, enfin… En tout cas, là, elle est bien fixée pour le coup, et puis on retrouve… Pour le théorème, je crois que dans le programme qu'il appelle théorème fondamental ou je sais pas quoi, là, on retrouve de nouveau des calculs d'aires… enfin, la démonstration qui est basée sur l'aire puisque la fonction là, c'est l'aire sous la courbe entre a et x, et que ton $f(x_0 + h) - f(x_0)$, c'est en termes d'aire qu'on va l'interpréter… La base enfin l'aire va apparaître dans les démonstrations par la suite. Donc ça reste là, quelque chose de fort à ce moment. En tout cas moi j'ai choisi de démontrer comme ça. Je sais que ce n'est pas le cas de tous mes collègues. Tu vois... On interprète géométriquement cette fonction. $F(x)$ c'est l'intégrale

de zéro à x de $f(t)dt$. Par un calcul d'aire, on va obtenir une expression pour des cas classiques pour eux et ils vont remarquer que ce qu'ils ont obtenu, c'est… Enfin que la dérivée de ce qu'ils obtiennent c'est la fonction f. Et du coup après on passe au cas général dans le cadre du programme, c'est-à-dire moi j'ai choisi le cas d'une fonction strictement croissante. Mais c'est pareil, c'est à base d'aires qu'on fait les premières… Voilà le $F(x_0 + h) - F(x_0)$, ça représente l'aire sous… d'un rectangle ou… Donc on en déduit l'encadrement… Après c'est plus de l'analyse, mais au début c'est… Tout est basé sur des choses qui sont en termes d'aires en fait. Alors après effectivement une fois qu'on a réussi à faire le lien à la fin, ici, entre l'intégrale et la primitive, après le cours, là, c'est sur les primitives. Donc ça, c'est pas du tout en termes d'aire. Là où on va le retrouver, c'est… C'est sur les propriétés de l'intégration. Cette année j'avais fait, je sais pas si je te l'avais donné, j'avais fait un polycopié pour gagner un petit peu de temps sur l'écriture des… Des propriétés. Non… Enfin, c'était juste que les énoncés étaient tapés et que on a tout, enfin ce qu'on pouvait, on l'a interprété en termes d'aire dans le cadre d'une fonction continue et positive. Donc c'est eux qui ont fait les dessins. Voilà, la partie dessin était vierge et dès qu'on a pu regarder sur les dessins ce qui se passait pour justifier ses propriétés, on l'a fait en termes d'aire en fait.

Chercheur : Et tu penses que ça reste plus pour eux, le fait d'interpréter en termes d'aire?

Marie : Je suis pas sûre là… En tout cas, cette année, j'ai pas eu l'impression. L'idée de départ que l'intégrale d'une fonction continue positive, ça oui… Maintenant là, pour les propriétés, je… Je suis pas sûre. J'avais l'impression qu'ils les avaient pas super bien apprises cette année les propriétés et une fois qu'ils ont les primitives… Ça, ça tombe un peu à l'eau… l'aire.

Chercheur : C'est ce que j'allais te demander parce que si par exemple, si après avoir fait les primitives, si on reprend les intégrales de fonctions affines, est-ce que…

Marie : Ah, si… il y en a qui restent sur du calcul d'aire quand même.

Chercheur : D'accord.

Marie : Particulièrement les ES d'ailleurs.

Chercheur : D'accord, donc y a deux groupes…

Marie : Mais pas beaucoup… Ce n'est pas des groupes équilibrés... C'est trois élèves qui sont… C'est un cours, un chapitre qui est très long. Il faut se rendre compte que c'est un chapitre où faire le cours, ça prend énormément de temps. Donc y a des élèves qui baissent leur attention et qui lâchent. Qui sont très très concentrés en début de chapitre sur le calcul d'aire. Et en plus ça marche très bien, c'est-à-dire que tous les élèves… s'y mettent à essayer de trouver comment calculer l'aire sous une courbe. Mais, au bout d'un moment, ils fatiguent et quand on arrive là, ça fait déjà plus d'une semaine qu'on est en train de bosser là-dessus… Et là, ça commence à les fatiguer un peu plus et il y en a qui restent sur les premières heures du chapitre et qui décollent pas forcément. Ça se voit après dans les exercices, sur l'utilisation des propriétés. Ils bloquent d'ailleurs.

Chercheur : Mais du coup… Toi, qu'est-ce que tu penses du fait qu'ils continuent à utiliser, enfin, par exemple pour une fonction affine, qui continuent à utiliser l'aire ?

Marie : Si elle est positive sur l'intervalle, moi ça m'est totalement égal.

Chercheur : Oui. A priori, ça te…

Marie : Oui. , je l'ai défini comme ça, alors je vais pas leur dire… qu'ils ont plus le droit une fois que j'ai fait…

Chercheur : Oui. Et, une fois qu'ils ont fait le calcul algébrique…

Marie : Avec les primitives…

Chercheur : Est-ce que tu penses que. Une fois les primitives, ils oublient le fait que c'était une aire ?
Marie : Non. Non. Et c'est même très prégnant parce que… ils restent sur l'idée qu'une intégrale c'est un nombre positif. Et l'idée que c'est… Enfin la première idée qu'ils ont reçu, c'est celle d'une aire sous la courbe, ça reste vraiment celle qui est captée par les élèves en fait. Et du coup, il est arrivé plus d'une fois qu'une fois que du bête calcul intégral, quelque chose de technique « calculer les intégrales suivantes », on trouve des résultats négatifs ou j'inverse les bornes, je mets, tu vois, et là d'un seul coup, ils me disent « y a une erreur dans l'énoncé », je dis « non, non, y a pas d'erreurs… ». « Madame, j'ai trouvé un résultat négatif, c'est bizarre, on avait dit que c'était toujours… ». Je dis « attention ! ». Donc, non, c'est vraiment la première idée. Mais c'est comme dans tout c'est la première rencontre qui reste… Le truc le plus fixé dans leur tête.
Chercheur : Oui, alors du coup, ça peut poser des petits problèmes après ?… Mais il faut…
Marie : Oui. Mais je pense que c'est plus dans le post bac que ça posera des problèmes. Pour nous, ça nous a pas posé de problème parce que tout d'abord, elles sont très souvent positives nos fonctions… [*Rires*] En gros, le… on peut remarquer que les fonctions qui interviennent le plus souvent là sont positives sur les intervalles en question. Suffit de regarder le sujet du bac de cette année. Donc en fait, on est quasiment toujours dans le cadre du début du cours. Je pense que ça posera plus de problème l'année prochaine.
Chercheur : Et est-ce que… donc tu dis que tu utilises les TICE pour… donc au niveau de la méthode des rectangles. Est-ce qu'il y a d'autres moments dans le chapitre où tu utilises les TICE ?
Marie : Alors… [Inaudible]
Chercheur : Par exemple, la calculatrice ?
Marie : [*elle tourne les pages de son cours*]. Tu vois, les calculs d'aire pour le calcul de la moyenne, ça revient à des interprétations graphiques et pour tous les théorèmes sur les inégalités. Ça a été in… instantanément traduit par les élèves en termes d'aire sous la courbe, pour expliquer les résultats en tout cas. Pas pour les démonstrations bien sûr, mais pour expliquer les résultats tout a été fait en termes d'aire. Alors je sais pas si l'algorithmique, tu le fais rentrer dans les TICE…
Chercheur : Ça dépend… déjà...
Marie : Mais en tout cas, oui, on fait de l'algorithmique dans ce chapitre-là.
Chercheur : Et vous l'avez mis sur calculatrice après ?
Marie : Oui eux ils le mettent. Moi je leur avais pas demandé, mais parce que c'était lié à la fonction…Après, je sais pas trop si… Non ça doit pas être un chapitre où on l'a beaucoup utilisé par la suite. Ils ont beaucoup de problèmes techniques… à calculer une intégrale déjà. Ça c'était très technique aussi… Attends, j'ai ramené mes trucs car je savais plus ce que je t'avais donné. Non, j'ai pas l'impression, à part la méthode des rectangles, je me souviens pas. Je regarde juste dans mes exos, parce que des fois… En plus, comme y avait les deux terminales, j'ai eu tendance à me prendre les pinceaux entre les deux classes. Parce que en fait le problème, c'est que la part technique est aussi importante dans le chapitre donc c'est les primitives et qu'ils ont quand même beaucoup de mal, soit dit en passant. Ça c'est un sujet du bac qui est tombé. Alors là y avait un peu d'algorithmique mais … Non. Moi j'ai pas souvenir, non.
Chercheur : D'accord.
Marie : Je vais te dire « non ». On a fait des exos de ce genre, tu vois… Donc une courbe machin… étude… aire… Ça c'est des fiches de technique, ça c'est des exos avec des

suites… Non, non. Ah oui, si, on avait fait un truc sur la méthode des trapèzes que j'avais mis à l'écran. Mais c'est pareil, c'est du calcul… comme la méthode des rectangles.
Chercheur : Et du coup… est-ce que vous avez vu des aires finis d'un domaine infini ?
Marie : Oui, dans les exercices.
Chercheur : Oui ?
Marie : Oui.
Chercheur : Et en utilisant les suites ? Enfin…
Marie : Hum. Exactement.
Chercheur : Et est-ce que ça, ça interpelle les élèves ? Ça pose des problèmes ?
Marie : Eh bien, ça dépend. En fait, c'est divers. Cette année, ça les a pas trop interpellés. Mais j'ai une classe extrêmement studieuse, qui est pas… comment dire… qui s'interroge pas forcément beaucoup, très profondément sur ce qu'on fait… qui le fait très bien mais … « oh oui, c'est vrai… Oui c'est vrai quand vous le dites, quand même… C'est vrai que c'est bizarre finalement qu'on trouve une aire finie pour ce domaine qu'avait pas l'air si fini que ça… ». [*Rires*]. Mais quand on commence à en discuter, ça … Je vois bien que ça les interrogeait pas tellement. La classe de l'année dernière, oui, elle, elle se posait de vraies questions. Avec elle, il y avait possibilité de travail. Et puis, je savais qu'on avait fait quelque chose sur le flocon de Koch en première S, donc on avait déjà parlé de ce genre de difficultés donc on a pu rebondir mais cette année, j'ai bien vu que ça les avait pas du tout passionné mon histoire. [*Rires*]. Il fallait que je change de sujet. [*Rires*]. Non, mais c'est des problèmes qui sont difficiles quand il y a pas de réceptions en face dans la classe, c'est contre-pro… ça sert pas à grand-chose de … j'insiste pas parce que… Et l'année dernière par contre, on avait fait plus de choses mais… Mais j'avais de très très brillants élèves en maths … Eux ça les interpellait. Ils se posaient des questions. Là ils voient ça comme une espèce de curiosité mais c'est tout, ça s'arrête là. [*Rires*]. Ça s'est arrêté là aussi pour moi.
Chercheur : Ils ont confiance…
Marie : Ah oui, oui, oui… ça avait l'air bien ce que je faisais au tableau alors ils se sont dit que ça devait être juste… [*Rires*].
Chercheur : Ok. Alors, si tu devais dire, qu'est-ce qu'il faut savoir à la fin de ce chapitre, tu … ?
Marie : Qu'est-ce qu'il faut savoir à la fin du chapitre… Tout bêtement savoir calculer une intégrale à l'aide d'une primitive. Il faut savoir, on a beaucoup travaillé dessus, que si on ne sait pas calculer l'intégrale avec une primitive, il serait bien de pouvoir l'encadrer, ça serait sympathique. Alors on peut utiliser les théorèmes qu'on avait. Qu'est-ce qui faut savoir à la fin de ce chapitre ? Si j'oublie le lien entre intégrale et dérivée ou primitive…mais finalement, ce n'est pas ce que retiennent vraiment les élèves. Oui, je dirais ça…
Chercheur : Puis du coup aussi avec les aires, enfin le lien avec les aires…
Marie : Oui, mais c'est-à-dire que une fois passé le … cette première rencontre,… Alors après, y a des questions toute bêtes « calculer l'aire sous la courbe », mais pour eux, c'est une formalité et c'est pas… ce n'est pas le calcul d'aire qu'est au centre de leur problématique quand même. Donc,…
Chercheur : Oui.
Marie : Ils font… Ils retiennent que l'intégrale c'était le calcul d'une aire, au départ. Après il s'est passé des choses bizarres : la prof, elle a changé les bornes, elle a mis des fonctions négatives, elle a fait des trucs pas clairs… Mais … La problématique du calcul d'aire, elle est très vite… A part la question classique d'un exo qui est « calculer l'aire

sous la courbe » où ça te sert strictement à rien d'ailleurs, soit dit en passant… c'est juste pour faire plaisir. Mais … Eux ils l'oublient quand même un petit peu.
Chercheur : Et là… Je pense à d'autres questions… Du coup, tu … Toi, tu… Pour toi, le calcul approché d'aire c'est… c'est important en terminale ?
Marie : Ah oui parce que… Oui, oui, oui. Parce que soit tu… ils savent calculer leur… alors on s'entend, calculer leur intégrale avec une primitive, soit ils savent pas parce qu'ils savent pas intégrer la fonction qui est à l'intérieur et auquel cas quand on sait pas calculer un nombre en maths et on essaie de l'approcher. Donc c'est… Et c'est quelque chose qu'ils font au travers de toute l'année de terminale donc là les inégalités, etc… Alors soit on sait et on fait, soit on essaie d'encadrer.
Chercheur : Et du coup… là, vous avez pas utilisé des… par exemple, des logiciels pour faire des calculs approchés d'aires.
Marie : Non.
Chercheur : Par exemple, enfin je sais, avec GeoGebra, on peut faire…
Marie : Ah non. Parce que sur cette partie-là, on n'a pas été en salle…
Chercheur : D'accord
Marie : On n'a pas été en salle équipée.
Chercheur : Alors. Est-ce que tu as fait des applications du calcul intégral dans le chapitre ?
Marie : A part le calcul d'aire ?
Chercheur : Oui. Enfin…
Marie : Je me souviens pas. Alors, le calcul d'aire, du calcul approché d'aires, ça c'est sûr. On a fait du calcul approché de nombres réels… et des applications à d'autres domaines tu veux dire ?
Chercheur : Par exemple.
Marie : Non. Non. J'ai pas fait.
Chercheur : Et du coup, si on passe aux lois de proba à densité…
Marie : Oui, chouette. [*Rires*]. Chapitre très long aussi.
Chercheur : Très long.
Marie : Je t'ai ramené mes… mes cours.
Chercheur : Et du coup, comment est-ce que toi tu introduis?
Marie : Alors, avec des histogrammes. On est parti… On en avait déjà parlé… Moi j'ai pas trouvé d'exemples plus intelligents que celui de l'année dernière. Donc, j'ai gardé mon exemple, j'ai juste réécrit mon truc au vu de ce qui s'était passé l'année dernière, quand même. Mais ça aussi ça marche assez bien. Donc on est parti de… Donc y avait un premier histogramme. Ils avaient que l'histogramme. Ils avaient pas le tableau, pour le premier exemple. Et donc on a cherché quelle est la probabilité d'intervalle à l'aide de l'histogramme. Et la question c'était « si c'est possible ou pas ». Donc il y en avait dans le lot qui étaient possibles, d'autres qu'étaient pas jouables avec le pas de l'histogramme que j'avais donné. Du coup… Alors en fait la feuille, elle était… tu vois y a des blancs parce que la feuille était coupée en trois morceaux. J'ai pas donné tout l'énoncé en même temps aux élèves. Donc, après la question c'était « comment on va faire pour celle qu'on n'a pas pu calculer pour l'instant ? ». Alors tous seuls, ils trouvent que si on avait pu avoir des petits rectangles un peu moins grands on aurait pu répondre pour [180 ; 220]. Il aurait suffi qu'on… comment dire… qu'on fasse des classes sur nos données statistiques plus étroites. Du coup, c'est là où je suis très forte, je leur dis « j'ai tout prévu ». [*Rires*]. Et donc, je leur donne le tableau avec un pas de 10 parce que c'était ce qu'il fallait et un deuxième histogramme. Alors, j'ai choisi de pas leur faire tracer les histogrammes parce que l'année d'avant, j'avais fait tracer les histogrammes et que ça

avait été un massacre et ça avait pris un temps considérablement fou. Et que j'étais… Je me suis dit « Plus jamais un histogramme en terminale S, c'est trop compliqué ». [*Rires*]. Même en leur donnant l'unité d'aire, ça avait été l'horreur, l'année dernière, j'avais même cru que j'allais les assassiner tellement c'était mauvais. [*Rires*]. Donc là, j'ai donné les deux histogrammes. Sur le deuxième, on peut tracer la courbe directement, ça a été fait exprès pour que l'échelle en ordonnées corresponde à la… à la fonction. C'est aussi pour ça que je l'ai fait parce que du coup, on pouvait après tracer la courbe directement dessus et du coup après voir que l'aire de notre histogramme, si on avait diminué le pas, ça leur rappelait tragiquement un truc qu'on avait fait en calcul intégral. Alors là par contre le lien, il est fait automatiquement en moins de trente secondes par tous les élèves. Ça … le calcul… enfin, les rectangles, une courbe et une aire, y a pas de doutes que ça, ça avait été bien assimilé. Donc après… donc une fois qu'on avait travaillé sur le deuxième, qu'on avait vu qu'il y avait des questions auxquelles on pouvait répondre mais d'autres auxquelles on pouvait toujours pas répondre, ils comprennent le principe et donc du coup, je donne la dernière partie du… de l'introduction. Alors je leur donne… une fonction parce que… enfin à l'oral on a beaucoup discuté, je leur ai demandé « donc on va prendre des rectangles… des bases de plus en plus petites… » du coup il y en a toujours qui te disent « on a qu'à prendre une fonction qui passe à peu près par les coins », ou « par le milieu », enfin bon voilà, ils voient bien une espèce de fonction. Donc on se met d'accord sur le fait que si on arrivait à trouver une fonction qui a l'air sympathique avec notre histogramme, finalement on fera du calcul intégral. Donc le tout, c'est de trouver cette fameuse fonction, que j'appelle une courbe de tendance au départ. Et donc là je leur ai fait la proposition… Donc ils ont tracé soigneusement sur l'annexe 2. Et donc vérifier que l'intégrale valait bien 1, … donc ça on avait déjà discuté à l'oral donc il y a pas eu de problème et puis calculer des probabilités d'évènements… dont certains qu'on avait calculé avant avec l'histogramme, d'autres nouveaux qu'on n'avait pas encore pu calculer et puis $P(X = 168)$ en l'occurrence. Voilà. Et donc là, le lien entre aire sous la courbe et calcul de probabilité, ça ne pose aucune difficulté. Il y a juste à trouver un exemple intelligent, ce que je n'ai pas encore réussi à faire. [*Rires*].

Chercheur : J'ai cherché… Je te dirais un petit peu où j'en suis arrivée…

Marie : Et c'est vrai que y a juste une partie... parce qu'en fait ce qui serait vraiment intéressant, c'est de… de… faire construire les données statistiques. Mais ça y a… il faut du temps… Donc voilà, ça peut être affiné et amélioré, la partie… Après ce qu'on voit dans le cours, c'est que quand on commence à leur dire que cette fameuse fonction f densité de probabilité, que la probabilité que X soit entre a et b ça sera l'intégrale… ça, ça ne pose aucune difficulté. Voilà, ils ont compris, ils ont fait eux-mêmes, ils ont compris que les histogrammes c'est ce qui leur permettait de calculer la probabilité d'un évènement et que du coup c'était l'aire sous la courbe. Donc ça, ça pose… c'est vrai que ça décharge toute la partie du début du cours qui ne pose aucune difficulté du coup…

Chercheur : Et du coup… Après quand apparaît la fonction densité…

Marie : Après pour toutes les lois… Ah complètement, ils ont compris que pour calculer la probabilité quand on a une variable aléatoire continue, enfin à densité ici, il suffit de… Si tu as ta fameuse fonction, il y a aucune difficulté. Et du coup pour les trois modèles qu'on a après dans le cours, tout se joue assez facilement parce que si tu as ta fonction et y a pas de… y a pas de soucis.

Chercheur : Et du coup, est-ce qu'après quand tu parles de fonction de densité, est-ce que tu donnes une… une définition générale ou…

Marie : C'est-à-dire ?

Chercheur : Est-ce que c'est adapté à… aux trois exemples du cours ou est-ce que tu fais quand même plus…
Marie : Non, non, j'ai fait un paragraphe de généralités. Alors c'est… […]. On avait discuté des variables aléatoires continues avant. C'est-à-dire, de parler du problème que on pouvait plus calculer ça, voilà… qu'on pourrait calculer dans un intervalle… donc là, j'ai pas mis de définition pour variable aléatoire continue. Après, j'en ai mis une pour fonction de densité de probabilité sur un intervalle I : toute fonction continue, positive telle que l'intégrale sur I est égale à 1. Et je note sur I parce qu'après on peut expliciter le problème. Quand I est un intervalle fermé borné, comme on avait fait dans l'exemple, c'était l'intervalle [200 ; 400], voilà ça on sait faire, on a notre fonction, c'est donc l'intégrale de 200 à 400, pas de problème, on a fait un chapitre qui servait à ça, y a pas de soucis… Après, il se passe quoi quand I est un intervalle qui est non borné ? Donc là je prends l'un des deux exemples de ce type là et voilà, je leur explique que ce n'est pas étudié en terminale, que ce qu'on va faire, c'est qu'on va mettre l'intégrale de a à X qu'on va chercher à calculer la limite lorsque X tend vers +infini, que si ça, ça admet une limite finie alors on va dire que ça se passe bien et on dira que c'est ce nombre qui est… Donc… en leur explicitant bien que voilà, ce n'est pas au programme de la terminale mais que je vais en avoir quand même besoin parce que j'en ai besoin pour définir des choses dans la suite du cours. Donc je sépare quand même les deux cas et ça les dérange pas plus que ça en fait… Bon, ils trouvent curieux que d'un seul coup, je fasse des trucs… que j'ai besoin de trucs pas dans le programme, ça par contre mais ils se disent « mais c'est quoi le programme… ce programme ? ». « Je sais pas ». C'est voilà… Voilà, on va pas vous demander vous de manipuler, quoi que c'est un peu faux mais il faudrait que vous compreniez que ces intégrales avec une borne infinie, c'est pas… voilà, c'est un petit peu particulier et qu'il faut faire un travail supplémentaire. Donc … donc c'est… je présume que c'était ça ta question, que je définis…
Chercheur : Et puis même, si une fois que tu as défini comme ça, tu fais des exemples?
Marie : Oui, y a des exemples mais c'est dans les feuilles d'exercices du coup pour vérifier que des fonctions que je leur ai donné sont des fonctions de densité de probabilité.
Chercheur : Et du coup…
Marie : J'en fais pas cent cinquante milles, j'en ai fait deux trois avec eux, ça a été assez vite…
Chercheur : Et dans les fonctions que tu as choisi, c'est des fonctions avec…
Marie : Alors y avait … y avait de tout, y avait un polynôme je suis sûre… et y avait… où là je crois même c'est eux qui devaient trouver les coefficients pour que ce soit une fonction de densité… Après, je me souviens pas parce que je crois que par contre je les ai pas là… Ah si, j'en ai une là. Exemple. oui parce que les exemples… […] Après il y en avaient qui étaient dans leur bouquin… Il me semble qu'il y en avait une… Je sais plus lesquelles. J'en ai fait deux ou trois avec eux, de ce genre-là.
Chercheur : Mais que où il fallait calculer des primitives quand même ou aussi avec…
Marie : Non pour montrer que… juste sur la question, je donne une fonction, vérifier que c'est une fonction de densité de proba sur un intervalle, j'en ai fait deux trois comme ça.
Chercheur : Oui, mais du coup c'est pour calculer que l'intégrale est égale à 1 en utilisant les primitives ?
Marie : Ah oui, oui…
Chercheur : D'accord. Et… Du coup, là, au départ…
Marie : Du coup, après, je… je dis aussi que X a pour densité de probabilité f si seulement… Je fais le lien entre les deux. [inaudible]. Du coup, après y a du calcul d'aire,

enfin y a des interprétations en termes d'aire, puisqu'on est revenu à notre chapitre d'avant donc… Alors, j'ai pas choisi… je crois que dans le programme, ils disent X a pour densité si c'est l'aire sous la courbe, je l'ai écrit comme ça et après je l'ai remis donc… l'aire sous la courbe.
Chercheur : D'accord.
Marie : Mais …
Chercheur : Tu as d'abord mis avec l'intégrale et…
Marie : Oui, je l'ai d'abord mis avec l'intégrale. Donc voilà. Et tu vois que tout de suite pour les évènements toute façon ça s'interprète en termes… enfin voilà l'interprétation d'aire sous la courbe, elle reste très prégnante dans toute… dans tout le chapitre en fait.
Chercheur : Et du coup, … au départ, on a vu que déjà tu as… tu as parlé du cas où y allait avoir l'infini et du coup quand arrivent les exemples donc… par exemple la loi exponentielle…
Marie : Alors, pour la loi uniforme, y a pas de problème, mais c'est pareil tu vois que le calcul intégral tout… tout est très interprétable en termes d'aire…
Chercheur : Oui.
Marie : Sous cette… sous cette fameuse fonction… donc ils retrouvent ça facilement sachant que l'aire vaut 1.
Chercheur : Oui.
Marie : Donc ils trouvent… et comme c'est un rectangle, puisque c'est une fonction constante, ils retrouvent la hauteur du rectangle donc f, enfin… ça le calcul d'aire pour le coup revient donc ça c'est… Alors ah voilà la loi exponentielle. Donc, donc on a le problème tout de suite pour montrer que cette fonction est une fonction de densité de probabilité sur $[0\,;+\infty[$. Donc là on y va, là ça va parce qu'ils peuvent intégrer sur $[0\,;X]$ et après on cherche la limite et … Donc on utilise ce qu'on a défini en amont, mais sans… sans le préciser au départ, je vois pas comment on peut… on peut s'en sortir…Donc … Je sais que les recommandations… enfin que dans le programme c'est pas, c'est pas excessivement clair… c'est on doit le faire sans le faire mais il y a des démonstrations obligatoires dans lesquelles par exemple pour l'espérance de la loi exponentielle, je vois pas comment tu fais si tu leur as pas expliqué qu'il fallait chercher la limite de…
Chercheur : Oui.
Marie : Sans leur expliquer complètement tout mais ils ont compris qu'il y avait un problème qu'il allait falloir travailler un petit peu plus.
Chercheur : Mais du coup, tu penses qu'après, ils… est-ce que ça peut rester que… que une intégrale même si y a + l'infini, ça posera pas de problème ?
Marie : Ah ! ça posera pas de problème… Ça ne posera pas de problème si c'est fini.
Chercheur : Oui, mais est-ce que tu penses que pour eux…
Marie : Après moi je sais pas parce que j'ai pas du tout de retour de…Je pense que ça quand même ils oublient très vite… Je pense que ce qu'ils retiennent des proba continues c'est les calculs de probabilité. Mais après… Parce que ces problèmes de limite d'intégrale infinie, c'est pour les démonstrations qu'on les a. Parce que par contre ce qu'on … je leur ai bien expliqué… ce qu'on a surtout expliqué aux ES, c'est que jamais on n'allait calculer l'intégrale de 0 à +… de a à + l'infini… qu'on allait toujours se ramener à celle qu'on savait faire … et que ce soit pour ça ou pour la loi normale… donc tout de suite on bascule tout notre probabilité pour que ne jamais avoir… d'intégrales enfin oui d'intégrales généralisées … eux, dans les exercices, ils l'ont pas utilisé. Le seul moment où on l'a utilisé, c'est en cours pour faire les démonstrations de cours.
Chercheur : Oui. Donc après c'est…

Marie : Après, ils savent que ils savent pas calculer l'intégrale… Ils vont pas aller s'embarquer pour calculer l'intégrale de 3 à + l'infini et qu'ils disent que c'est 1- la probabilité… on utilise les proba après donc… Donc voilà… et et très sincèrement quand j'ai fait la démonstration sur le… l'espérance, ça a été juste un mauvais moment à passer. Ils me l'ont bien fait comprendre… [*Rires*]. J'ai pourtant dit qu'elle était exigible celle-là mais voilà. Alors le professeur s'est un peu excité au tableau… voilà… elle est gentille alors on va pas trop l'embêter… Quand je leur ai demandé à la fin combien il y en avait qui avaient suivi, y a quatre ou cinq doigts qui se sont levés. J'ai dit « c'est bien, vous êtes dans la moyenne… tout va bien ». J'ai pas été plus mauvaise que d'habitude. [*Rires*]. Voilà. Après, on n'en a plus besoin en fait pour la loi exponentielle.

Chercheur : Et du coup, toi, tu… […] Toi du coup tu considères qu'ils doivent connaître leurs formules, que… ?

Marie : Alors oui, on en avait discuté. Oui, alors… Alors je suis restée sur un entre-deux très sympathique… [*Rires*] qui a été une espèce de consensus avec les élèves… je leur ai dit « alors pour l'examen, moi je vous conseille de refaire les calculs » parce que tu vois que là je les ai mis en propriétés, alors on les a démontré… mais parce que quand tu fais cinq exercices de suite, tu n'en peux plus de cette intégrale à calculer. Voilà. Mais voilà je leur ai dit … Je suis assez honnête avec eux… Je leur ai dit que je trouvais le programme flou sur ce qu'était vraiment… ce qu'on leur demandait… et je leur ai dit « le jour du bac, au pire, il y aura une question donc vous me refaites l'intégrale correctement ». Et en contrôle, ils m'ont refait… ils m'ont refait les intégrales enfin y avait un exercice et ils me l'ont refait. Ils ont bien compris la règle du jeu. Ils sont aussi habitué à ce que je leur dise « là c'est un petit peu pas très clair donc … » y a deux mesures : y a le cours au quotidien où on va pas se taper quinze fois l'intégrale $\lambda e^{-\lambda t}$. On n'en peut plus et puis, le jour du contrôle et de l'examen où là on fait des choses de plus… et ça leur a pas… ça leur pose pas de problème.

Chercheur : Y a les deux, ils ont les deux et ils choisissent un peu… enfin…

Marie : Ils choisissent…

Chercheur : Enfin tu proposes qu'en examen plutôt…

Marie : Ah non en examen, je leur ai dit « non, non, en examen, vous me refaites au moins… si y a trois questions, vous faites la première intégrale vous la faites bien, pour montrer que vous savez l'intégrer ce fameux truc. Après c'est bon… ». Voilà, donc… Et il y en avait… il y en a qui de toute façon préfère recalculer l'intégrale à chaque fois, ce qui est par exemple mon cas… parce que moi, je leur ai dit, je ne… d'ailleurs ils l'ont vu… je ne connais pas les formules… j'ai dit j'apprends pas ces formules-là, je vais me planter, je me connais, donc moi je recalcule l'intégrale à chaque fois… donc … voilà, ceux qui sont comme moi, des poissons rouges, et bien vous recalculez votre intégrale et puis les autres… c'est vrai que quand tu fais cinq exercices à la suite les uns des autres t'en as un peu marre quoi… Mais c'était pas très clair, enfin moi j'ai pas trouvé ça très clair dans le programme… et je suis pas la seule…

Chercheur : Et du coup, pour la loi uniforme, pareil ils refont le calcul soit avec l'intégrale soit… ?

Marie : Ah la loi uniforme, j'avais trouvé ça un tout petit peu plus clair mais j'ai eu le même discours pour les deux. Et ils se sont débrouillés.

Chercheur : Et du coup, ils utilisent plus le calcul intégral ? Ou ils font leurs petits rectangles ?

Marie : Ca dépend des élèves… Non, non… on sent… on sent les élèves qui apprennent le cours par cœur et … là, y avait la formule, on est content, on va l'appliquer… et puis les

élèves qui sont plus, plus autonomes et qui savent qu'il y a un truc à faire et qui recalcule leur intégrale à chaque fois.
Chercheur : D'accord. Et y a personne qui… qui… qui fait sa loi… enfin qui trace la loi uniforme et qui voient que c'est un rectangle…
Marie : Si. Y a des élèves… Oui. Y a des élèves plus… plus empiriques qui vont vraiment repasser par le calcul d'aires, qui vont… Et il y en avait en ES qui faisaient ça… plus facilement en fait.
Chercheur : Et du coup, tout le monde fait pas le graphique à chaque fois ?
Marie : Non, non.
Chercheur : Et qui voient…
Marie : Ce qui est pas le cas pour la loi normale par contre.
Chercheur : Oui. Du coup, si on passe à la loi normale…
Marie : Alors, la loi normale donc moi je sais que j'ai pas respecté les recommandations du programme puisque je n'ai pas commencé par le théorème de Moivre-Laplace parce que j'ai eu beaucoup de mal à comprendre comment je vais le vendre aux élèves. Je… je suis pas la seule, je le sais parce qu'on a beaucoup discuté au sein de l'équipe… Je… je vois pas comment vendre aux élèves le fait de repartir d'une loi binomiale et d'un seul coup de se dire « Oh, regardez . [*Rires*]. C'est rigolo »… pour… pour arriver à tracer la courbe d'une fonction qu'ils n'identifieront pas de toute façon puisque ils ne connaissent pas par cœur la courbe de Gauss… Donc je n'arrive pas à comprendre comment je peux le faire avec les élèves… du coup, je fais comme pour les autres… On était parti du principe que si on avait la fonction de… enfin la fonction de densité on pouvait calculer… donc c'est jamais que comme une loi uniforme et une loi exponentielle… donc on a une nouvelle fonction de densité, ce qui nous donne la loi normale… centrée réduite… Alors on a travaillé sur cette… sur cette fonction… on avait fait des choses un petit peu… j'avais anticipé quand on a fait la fonction exponentielle, quand on l'avait déjà rencontrée… voilà donc… y avait plein de choses… à dire…
Chercheur : Et du coup, pour voir que l'intégrale vaut 1, ça se passe… ?
Marie : Non. Ça se passe comme ça. [*Rires*]. Ça se passe très bien comme tu peux le voir. Regarde. [*Rires*]. D'accord. Et donc après oui… voilà… donc … on parle du problème de calculer puisque cette fois-ci, on n'a pas de primitives... parce que pour les deux autres, ils font du calcul intégral mais là ils pourront pas faire de calcul intégral… Donc on est très vite sur les machines… et par contre, voilà, on voit comment calculer toutes sortes de probabilités avec l'aire sous la… enfin en interprétant en termes d'aire sous la courbe… et en changeant les intervalles pour que ce soit bien comme la machine sait faire en fait. Donc ça, ça passe… on voit le principe, ça va très bien. Oui, alors ça, c'est pareil… Et alors après par contre, oui, voilà…. Ça c'est un choix qui était différent par exemple… donc on a pas la fonction de répartition, c'est ce que j'ai compris du programme en tout cas, et donc voilà… je… on va démontrer une série de petites propriétés pour éviter de retravailler sur… sur ce dessin trop souvent… Mais qu'on a tout interprété en termes d'aire… et en termes de propriétés de proba. Donc ce qui leur pose plus de problème, en fait, c'est ça… c'est l'intervalle associé à une probabilité donnée… Ils veulent calculer des probabilités d'évènements mais alors par contre, faire l'inverse, ils… enfin comme c'est la première fois qu'on le fait en gros… parce que quand ils font des proba avant, on ne fait jamais ça… Là c'était plus difficile pour eux de comprendre… enfin… c'est pas difficile de comprendre la démonstration… ça a été plus difficile de comprendre, ce qu'on faisait… donc… Bon après on a utilisé tout ça pour avoir les fameuses valeurs à connaître par cœur… qu'on sait pas encore pourquoi faut les apprendre par cœur… Parce que la machine, comme me l'ont fait remarquer les élèves,

mais la machine elle donne tout de suite… Du coup, oui… mais c'est pas grave… Et les lois normales non centrées réduites… enfin quelconques plutôt… donc là on était obligé de revoir les petites propriétés pour l'espérance d'une transformée affine et la variance… ça ils avaient complètement oublié… donc ça m'a un peu… enfin, j'avais pourtant anticipé un peu mais… pas suffisamment visiblement… Donc on fait des exemples tout de suite… Mais bon, après, une fois qu'on a la loi normale centrée réduite, ça s'est pas… Alors par contre ce qu'on… ce qu'on voit voilà, c'est les influences des paramètres… à m fixé… à σ fixé… donc des choses de ce genre… Donc tout ça c'était aussi à l'écran… avec des TICE. Et puis, le théorème de Moivre-Laplace.

Chercheur : Donc en…

Marie : En fait, qui me sert de lien… entre ça et les intervalles de fluctuations. C'est ce qui fait le lien. Et là, là par contre, il y avait… Là il y avait une animation TICE. Et on est parti aussi… J'arrive plus à me souvenir dans quel sens j'ai fait… Oui, non, on a fait des calculs. Je confonds avec mon intervalle de fluctuation dans l'introduction. On a fait des calculs de… Il y avait un exercice qu'était en base… enfin en support… avec un petit n et p… Après on a regardé ce qui se passait et là on a vu que ça commençait à se compliquer pour calculer certaines. Voilà. Et du coup, c'est là où on a… J'ai montré l'animation où voilà… on transforme les bâtons, les histogrammes… enfin les bâtons en histogrammes et puis on fait tendre… n vers + l'infini… et là tout de suite, ils… oui, ils reconnaissent bien la courbe de Gauss… On l'a centré, on l'a réduit… enfin bon, tout était magique… donc ça marche bien… Mais c'est moi qui fait tout. J'ai un collègue qui a choisi de faire faire aux élèves mais le TP est très guidé parce qu'en fait on leur donne tout ce qui faut taper donc ça représente pas trop d'intérêt. Ça prend beaucoup de temps… Donc… Là c'est moi qui ait animé en classe.

Chercheur : Et vous avez tous choisi finalement de le faire en … à la fin de…

Marie : Dans l'équipe, oui. Et pour la même… enfin moi j'en ai discuté avec mes camarades, pour la même raison, on arrive pas enfin… je ne suis pas convaincue d'arriver à le vendre aux élèves…Donc …

Chercheur : Et du coup, après…

Marie : Et ça me pose un vrai problème… enfin ça ne me pose pas de problème de le faire comme ça… ce n'est pas ce qui est demandé dans le … dans le programme, ça j'en suis consciente… on nous demande d'introduire la loi normale grâce au théorème de Moivre-Laplace… mais vraiment… je sens que je vais pas bien… je vais pas bien m'en sortir alors je… voilà je botte en touche tant que j'ai pas trouvé…

Chercheur : Un truc qui te convient.

Marie : Un truc qui me satisfait… parce que je sais que quand je suis pas convaincue, j'arrive vraiment pas à le faire passer pour le coup… Donc …

Chercheur : Et du coup après, tu fais des… tu fais des exercices d'application… enfin par exemple, tu prends une loi normale…

Marie : On a fait des… des grandes binomiales oui… ce qu'on appelait avec les élèves « des grandes binomiales »… oui, mais pas… pas tant que ça parce que en fait, c'est vrai que ce théorème de Moivre-Laplace… en lui-même, on va pas énormément l'utiliser…

Chercheur : Oui, oui…

Marie : Il nous sert pour construire notre intervalle de fluctuation par la suite, mais voilà… alors on en a fait quelques-uns mais une fois qu'ils en ont fait deux… de toute façon, comme dans tout le chapitre, les exos sont extrêmement formatés… donc au bout du deuxième cylindre de loi normale, ils en pouvaient plus… la troisième bouteille à remplir… ils voulaient plus en entendre parler. [*Rires*]. Et là c'est pareil… une fois qu'ils ont compris ce qu'ils vérifiaient, ils voient plus très bien l'intérêt de se taper ça encore

pendant je sais pas combien de temps donc… Je… je dois avouer que je… je… j'ai… y a des moments où je sais pas trop comment… monter tout ça encore… soit dit en passant. Mais bon, ça va sûrement venir… [*Rires*]. Faut pas désespérer !
Chercheur : Est-ce que… du coup, dans les démonstrations, on a vu en regardant… dans les choix des exemples, est-ce que…
Marie : Des exemples de cours ou des exemples...
Chercheur : Les exemples… par exemple, quand tu utilises la loi exponentielle ou la loi normale, je sais pas… est-ce que vous faites des exemples concrets ?
Marie : Ah oui. Oui, ils sont tous concrets.
Chercheur : Ou c'est des…
Marie : Non, c'est tous concrets. Enfin, quasiment tous… A part les exercices techniques, qui est… vérifier qu'une fonction est une fonction à densité… y a très peu d'exercices formels. Alors… y a une raison à ça, c'est que c'est le dernier chapitre et que le temps pressait légèrement en ce qui me concernait… donc … là j'ai plus travaillé pour l'examen, que pour le post bac sur ce chapitre, ça c'est sûr… ce qui n'est pas forcément le cas sur les autres chapitres… Mais là… Là, je me suis contentée des exercices pour qu'ils survivent jusqu'à la fin du mois de juin… J'espère qu'ils ont survécu. Et après, tous les exercices sont à support concret.
Chercheur : Mais, c'est des exercices fabriqués ou… enfin…ou est-ce que, je sais pas, c'est des…
Marie : Non. Ah non, enfin ça dépend… C'est des trucs « bateau » la plupart du temps… entre autres parce que je suis pas super à l'aise… Après, dès fois, il y en a qui sont… Qui sont construits… qui sont inventés pour les élèves… mais c'est plus des… comment dire… des blagues entre nous qu'autre chose… [*Rires*]. Donc, ça dépend des classes que j'ai quoi… Quand j'ai envie de leur parler de football, je vais leur parler de football… Quand j'ai envie de leur parler de la cantine, je leur parle de la cantine… et quand ils arrivent en retard en cours, ils ont des exercices sur les élèves qu'arrivent en retard en cours… [*Rires*]. Donc c'est pas très…
Chercheur : Oui, mais concrètement, la modélisation on sait pas si c'est… enfin c'est une modélisation choisie comme ça…
Marie : Ah non, mais ça oui, on en a discuté… c'est … d'ailleurs, ils le voient bien eux-même… Tous ces exos, c'est juste un habillage… On sait pas pourquoi… ça suit… enfin… on a discuté du fait qu'une loi normale quel genre de phénomène ça pouvait traduire, une loi exponentielle aussi… mais maintenant on n'a pas du tout fait de modélisation dans ce chapitre-là… La modélisation est prise en charge. C'est… d'ailleurs les élèves décryptent très bien les énoncés… tu as dix lignes de blabla… paf, une loi normale de paramètre trucmuche, chouette, je vais faire des calculs… [*Rires*]. Et ça n'a aucun sens… Et la plupart des exercices n'ont aucun sens pour eux… voire sont profondément débiles… « On va longtemps encore mesurer des diamètres de tringles, de machin, de poutres… ? »... Je dis « non, on va remplir des bouteilles »… [*Rires*]. Alors, on discutait un petit peu de… du fait que pour le coup, ça… ça prend le sens sur du… sur des chaînes de fabrication, des choses comme ça… Ils comprennent que quand on remplit des paquets de café, ça va pas faire 250 grammes et que évidemment on peut pas se permettre de faire tout et n'importe quoi… donc ça, ça prend un peu son sens… mais c'est pas eux qui prennent en charge la modélisation et honnêtement, je ne me suis pas plongée sur le sujet dans le chapitre… mais par manque de connaissance de ma part…
Chercheur : Et du coup… est-ce… donc finalement à part les trois lois et puis un petit peu au départ avec les fonctions… enfin fonctions de densité, vérifier que ça en est une ou pas, tu vois pas d'autres lois à densité ? Tu en introduis pas d'autres dans ton chapitre ?

Marie : Non, non. Là j'ai plus le temps.
Chercheur : C'est par manque de temps.
Marie : Le problème oui... C'est que c'est très très long de faire juste tout ça. C'est-à-dire, ça nous prend quand même trois semaines.
Chercheur : Oui.
Marie : Et j'ai pas encore fait les intervalles de fluctuation et de confiance. Et je sais qu'il va encore falloir faire ça après… et … alors c'est aussi une question de progression sur l'année… Moi j'avais envisagé de ne faire que le calcul intégral sur des fonctions positives continues, pour embrayer sur les lois continues… pour casser un peu ce chapitre qu'est beaucoup trop lourd… mais ça demandait de faire la fonction log plus tard… parce que la fonction exponentielle on en avait besoin… et on est dans une équipe, y a des bacs blancs donc… voilà, on peut pas toujours tout vendre… on peut pas obtenir tout ce qu'on veut donc… voilà. Et à la fin on se retrouve… voilà avec cet énorme pan proba continues suivi d'échantillonnage… Et c'est très très long… ça dure cinq semaines… Et y a beaucoup… y a beaucoup de cours parce que quand tu regardes le cours… tout ça, c'est nouveau pour eux… donc ça prend du temps… y a du vocabulaire spécifique, y a toutes les démonstrations quasiment… Donc ça prend beaucoup de temps pour… une batterie de savoir-faire très faible… Donc… donc c'est compliqué… On s'était dit avec des collègues, qu'on allait essayer de réduire le cours… et on n'a pas… enfin on n'a pas réussi à enlever des morceaux… c'est-à-dire on est déjà au strict minimum de ce qu'on peut…vendre pour que ça est… enfin pour que le cours se tienne… sans balancer tout…Moi je vois pas trop comment faire… Et y reste le problème de l'échantillonnage après donc…
Chercheur : Et du coup, est-ce que toi… toi tu vois les proba à densité comme une application du calcul intégral ou quelque chose de finalement totalement à part et… ?
Marie : Totalement à part, peut-être pas mais… pas comme une application du calcul intégral.
Chercheur : Mais pourquoi ? Parce que…
Marie : Non parce que dans ma tête, c'est pas la même chose… c'est tout… [*Rires*]. C'est pas le même domaine. Il y en a un, c'est des stat-proba… c'est à moitié des maths… C'est un truc là, on nous a demandé de faire ça, on se demande bien pourquoi… Non, je plaisante… oui, dans ma tête, c'est pas le même domaine en fait. Mais je sais que mon cours d'intégration, à la fac, c'était… c'était… les deux étaient dedans. Mais… mais pour l'instant, ça reste encore une espèce d'OVNI… c'est intéressant d'un autre côté mais… on fait des choses intéressantes avec les élèves quand même mais c'est… on voit bien le lien quand même avec le calcul intégral, il serait absolument hors de question de ne pas avoir fait le calcul intégral et calcul d'aire avant… et on voit quel soutien c'est aussi… mais … mais c'est vrai que dans ma tête, c'est des domaines séparés. […]
Marie : c'est deux énormes chapitres… et qu'on ne peut pas malgré tout… faire en début d'année… tu peux pas commencer le calcul intégral au mois de septembre… voilà… le problème, c'est qu'il faut absolument… il faut un peu construire l'analyse… ils ont pas les limites, ils ont… ils sont très faibles en dérivation donc… [inaudible]. Enfin il y a plein pleins de trucs à faire… on a besoin de suites, on a besoin… donc malgré toute la bonne volonté de vouloir faire une progression… il y a il y a vraiment… moi j'ai pas pu avancer plus que ça sur … sur le sujet… j'ai… et c'est vrai qu'en plus, à un certain moment… arrivé à la fin de l'année où… si jusqu'à Pâques on va dire cette année, début mai environ, on arrive à tenir les élèves sur le post bac… c'est-à-dire que l'objectif , dans le lycée dans lequel je suis, c'est pas le baccalauréat, c'est clair… A partir de début mai, ils ont des épreuves de bac toutes les semaines… des vraies épreuves de bac, pas de bac blanc…

donc c'est très compliqué… moi je n'ai jamais eu ma classe entière devant moi depuis début mai… […] Et eux, dans leurs têtes, ça y est le bac a commencé et donc là, ils écoutent pas… enfin… ils sont moins ouverts à des exercices plus sophistiqués, qu'on a pu faire sur d'autres chapitres que là maintenant où c'est « oui, mais alors au bac, ça sera comme ça ? ». Alors que dans le reste de l'année, ils m'ont pas posé la question. […]
Chercheur : Finalement, ça tombe trop tard dans l'année… enfin mais pas parce que… parce qu'il y a pas le choix.
Marie : Ça tombe trop tard dans l'année mais je vois pas comment le faire beaucoup plus tôt non plus…et de toute façon, ce que j'avais prévu au début, c'est-à-dire couper le chapitre d'intégration pour mettre des proba tout de suite… les proba continues, le début tout de suite… en fait c'est… ce qui est très long, c'est uniforme, exponentielle, loi normale… pour casser entre les deux … ça n'aurait pas fait avancer énormément les choses quand même parce que la fonction log, ils ont beaucoup de difficultés, on peut pas non plus la faire en tout dernière… donc …
Chercheur : Parce que du coup quand même… avant dans l'ancien programme, il y avait quand même les lois uniformes enfin… la loi uniforme sur [0 ; 1] et la loi exponentielle
Marie : La loi exponentielle… oui mais il y avait beaucoup moins…
Chercheur : Oui il y avait que ça…
Marie : Il y avait que ça… c'était beaucoup plus court et il y avait aucune démonstration et en fait oui on faisait de façon très technique et il y avait pas grand-chose… parce que là on nous demande quelque chose qui est un peu plus construit…
Chercheur : Oui du coup ça…
Marie : Et vraiment, le cours est long…
Chercheur : Oui il y a beaucoup de choses à écrire
Marie : Oui il y a beaucoup de choses à écrire… il y a beaucoup de vocabulaire pour eux à se mettre dans la tête tout simplement donc … au début ça démarre doucement quand même... donc…
Chercheur : Et tout à l'heure tu parlais de la fonction de répartition, tu penses que c'est un problème ou pas que…
Marie : Non moi ça me pose pas de problème parce que je la déteste. Mais … [*Rires*]. […] Je sais parce qu'on en a discuté avec mes collègues que ça les gène eux qu'y est pas la fonction de répartition. Moi ça m'a pas gêné qu'il n'y est pas la fonction de répartition parce qu'on peut très bien parler de la fonction de répartition sans qu'il y ait la fonction… il suffit d'écrire P de X plus petit que a… moi ça ne me pose pas de problème de pas écrire $F(a)$, voilà. Et je pense que l'on a déjà alors là pour le coup assez de trucs sans rajouter encore… et on a travaillé avec P de X inférieur ou égal à a, ça n'a pas empêché de démontrer des choses etc… donc moi ça ne me dérange absolument pas et je pense que ça serait même assez insupportable pour le coup la dose parce que ça ferait encore…
Chercheur : Une notion en plus
Marie : Une notion… des graphiques en plus…enfin c'est… ça ferait peut-être beaucoup là… on est déjà à une dose de probabilité que les élèves, soit dit en passant… trouvent que, pour des TS, c'est bon quoi… [*Rires*]. Et ce n'est pas moi qui leur ai soufflé… [*Rires*]. […] ils trouvent que faire tout ça de cours… où on leur demande de se concentrer sur des choses un petit peu bizarre etc… pour calculer avec sa machine P de X compris entre 200 et 210 … ils trouvent qu'on exagère un petit peu quoi…voilà… puisque après tout est réglé… et c'est ce que l'inspection l'année dernière il y avait eu une réunion pour présenter le nouveau programme particulièrement la partie proba continue … on va dire que c'est en off mais l'inspecteur avait fini par nous dire « oui on y reste pas des

plombes parce qu'il y a deux exercices à savoir faire ». Donc… je caricature pas ce qu'il a dit… en fait il y a que deux choses à savoir faire alors on n'y passe pas du temps… … malgré tout, on y passe du temps… donc… mais je pense aussi qu'on ne sait pas bien le gérer… parce que nous tous dans nos formations, certains d'entre nous plus âgés que moi n'ont pas du tout fait de stat-proba… moi j'en ai fait … mais c'est la première fois que je les enseigne…en gros, je veux dire comme ça…donc on n'est pas aguerri et on a encore beaucoup de mal à se poser les bonnes questions… c'est pas mon domaine de prédilection, ça c'est certain… donc je vois bien moi chaque été je travaille avec des bouquins du supérieur pour vérifier que je … j'ai pas loupé encore un truc ou que je suis sûre de pas raconter de bêtises… donc on n'a pas… en ce qui me concerne, je n'ai pas suffisamment de recul pour… tu vois, répondre à ta question de tout à l'heure… ou savoir comment je pourrais monter le truc autrement que… pour monter mon cours moi j'ai pris des livres du supérieur pour le monter de façon logique et qu'y est pas de… tu vois… de problèmes… donc on est loin du reste du programme quand même tu vois…

Chercheur : Je pense qu'il y a un problème de…

Marie : Je pense qu'il y a un problème de formation, c'est évident… et je suis pas… moi je le dis aisément… il y a pas de soucis là-dessus mais… j'ai besoin encore de travailler énormément la théorie pour être capable de monter un truc intelligent sur le sujet… et j'ai pas assez de recul dans le domaine… ça c'est sûr… tu vois quand tu es en analyse, en géométrie, en complexes et tout, tu sais quels sont les exercices du post bac, tu… tu sais ce vers quoi tu veux aller… là moi je l'ai pas du tout le recul… donc du coup ça donne un cours qui est très… très proche d'un truc on va dire magistral quand même, même si je me fends d'une activité au départ pour essayer de leur expliquer comment on va calculer ces proba un peu particulières mais voilà… j'ai … c'est ma mauvaise connaissance du domaine moi qui me… qui me brime.

Chercheur : Et tu penses que les difficultés des élèves sur ce chapitre… c'est…

Marie : Il n'y en a pas.

Chercheur : Il n'y en a pas… Enfin c'est juste que c'est lourd, c'est trop lourd pour eux ? Enfin en longueur…

Marie : C'est lourd en termes de cours… et il y a deux phénomènes qui se juxtaposent si tu veux. C'est un cours très long voilà… ou on… là j'ai fait le choix de faire un cours où on écrit au tableau, j'ai pas donné un polycopié de cours sur le sujet… sur la fluctuation d'échantillonnage, il y avait un polycopié de cours… mais parce que j'ai axé beaucoup sur les activités et que là j'étais plus à l'aise, très curieusement… donc ça ça prend beaucoup de temps, d'écrire tout ça au tableau, de le faire avec eux… de faire les dessins… ils mettent du temps à faire les dessins… c'est pour ça que les exemples sont polycopiés numérotés… ils avaient plein de courbes numérotées qu'il y avait à mettre dans le cours pour essayer de gagner du temps parce que leurs lois normales sont tellement moches que c'est pas possible de dessiner dessus des aires qui se ressemblent à peu près… non mais tout ça ça prend énormément de temps… et même en numérotant tous les polycopiés pour que voilà… […] ça prend énormément de temps et ils en ont marre… et on est à la fin de l'année… alors ils en ont marre… ils sont pressés de passer le bac… ils sont pressés que ça se termine… C'est lourd pour eux. […] Faut être honnête, c'est la seule fois dans l'année où ils m'ont demandé « Madame, c'est bientôt fini ? ». « Il y en a encore un après ? ». « C'est fini, c'est fini ». Et on sent que la fin, heureusement que la fin de l'année arrive donc il y a ça aussi qui se… Ça serait un autre chapitre, ils en auraient peut-être marre tout autant. J'aurais terminé par la géométrie dans l'espace peut-être qu'ils auraient détesté ça et qu'ils m'auraient dit « C'est bientôt fini ? ». Je pense qu'il y a

vraiment ce phénomène qui joue beaucoup sur le dernier chapitre de terminale quand même parce qu'ils sont vraiment pressés de passer le bac, ils en peuvent plus donc...
Chercheur : La pression de l'année.
Marie : Oui. Il est temps que ça se termine en fait. Donc c'est vrai que ce chapitre, il est… Moi j'ai pas… Il y a pas d'élèves sur l'année dernière ou cette année… alors j'ai eu deux fois des très bonnes classes soyons honnêtes… pour autant aucun élève n'arrive pas à faire les exercices de proba continue qui sont dans les manuels… parce que ça relève de l'application… tout simplement… et les TES c'est pareil… et ils étaient même plutôt contents à l'idée d'en avoir à l'examen en se disant « ça au moins… »… calculer avec sa machine… la seule question, c'est « on met sur la copie qu'on calcule avec la machine ou pas ? ».
Chercheur : Juste une dernière chose : est-ce que la méthode de Monte-Carlo ?
Marie : Non, cette année j'en ai pas parlé. Non. Mais ça dépend, ça dépend des années aussi. Moi j'ai été très stressée par ces histoires de jeudis [jours fériés] donc je me suis moins… on s'est moins amusés . Et je savais que pour la fluctuation d'échantillonnage j'avais quelque chose donc voilà, j'ai passé du temps avec eux sur une introduction. On peut pas non plus… Parce que là il y a moins de cours après, sur cette partie-là.

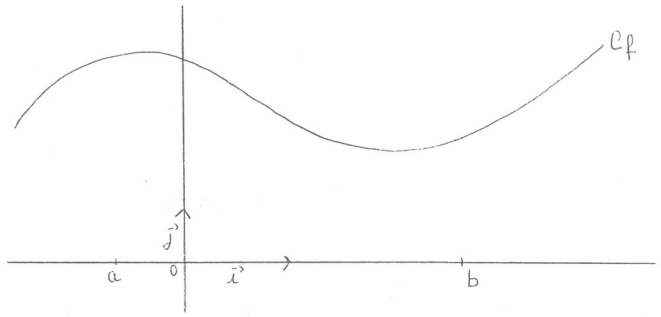
F.3 Documents de cours de Marie en 2013-2014

F.3.1 Documents relatifs au chapitre « Calcul intégral »

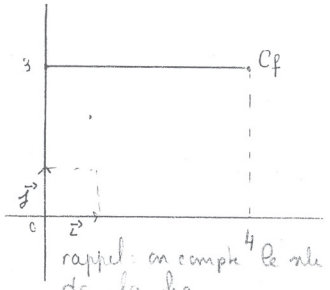
AIRE SOUS LA COURBE

Objectif : On considère une fonction f définie, positive et continue sur $[a,b]$ et C_f sa courbe représentative dans un repère orthogonal $(O,\vec{i},\vec{j})$.

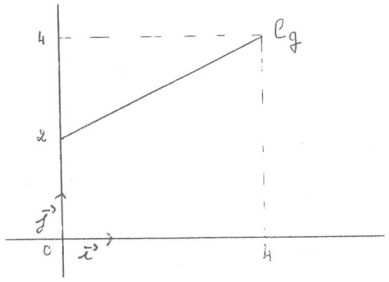
On appelle aire sous la courbe C_f, l'aire du domaine plan compris entre les droites d'équation x = a, x = b, la courbe C_f et l'axe des abscisses. Pour mesurer cette aire, on utilise comme unité d'aire (notée u.a.) l'aire du rectangle construit sur les vecteurs de base.

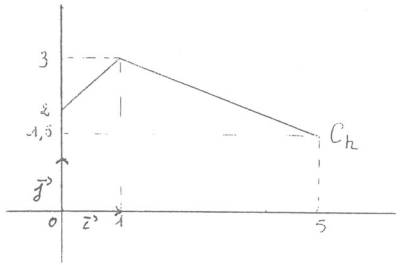


1. Pour chacune des courbes suivantes, calculer, en u.a. , l'aire sous la courbe.



rappel : on compte le nb d'u.a ds le pg.

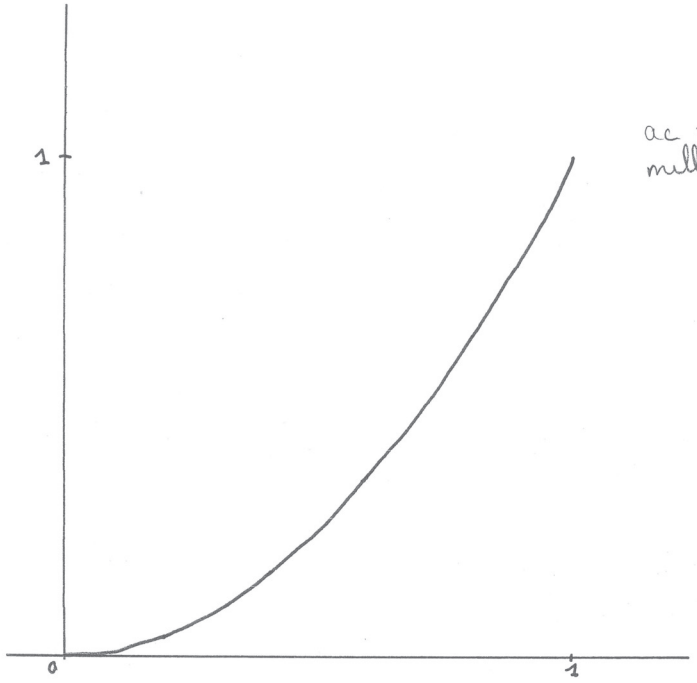




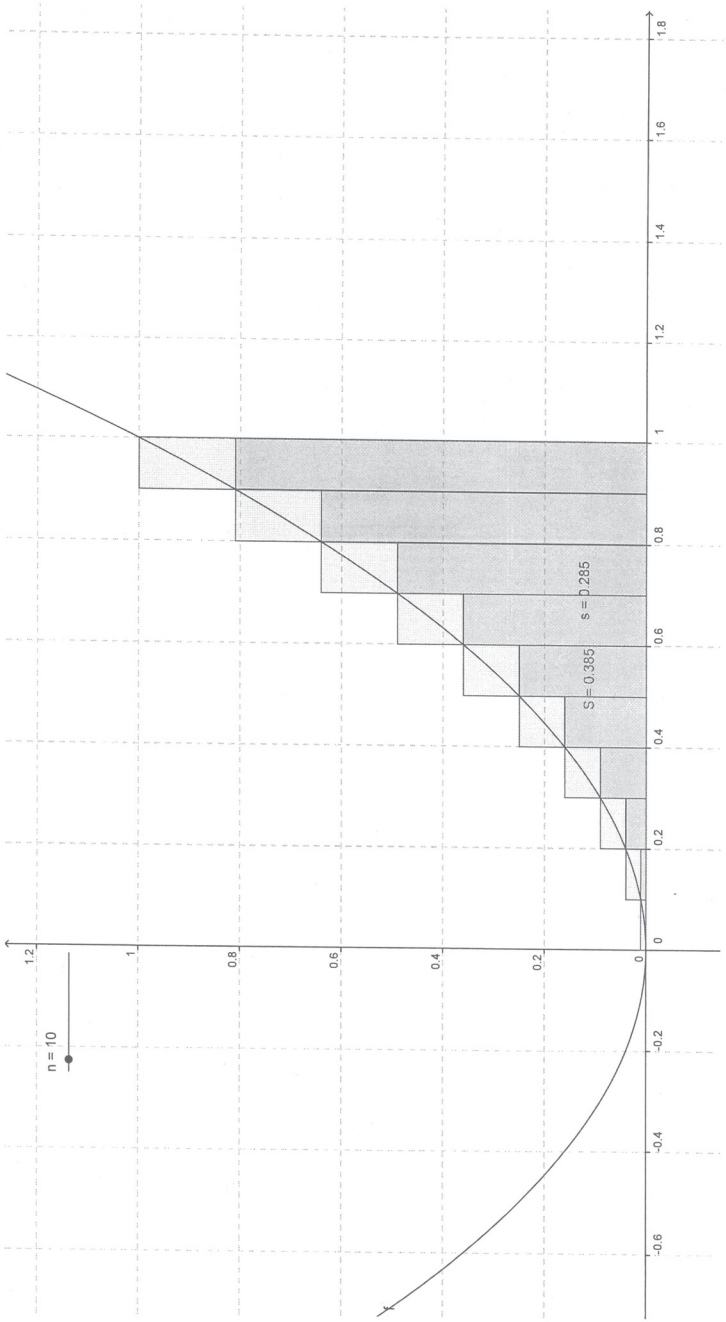
2. A vous ! Proposez la courbe représentative d'une fonction définie, continue, positive sur un intervalle [a ;b] dont nous calculerons l'aire sous la courbe.

3. On considère maintenant la fonction $f : x \mapsto x^2$ et C_f sa courbe dans un repère orthonormé.

On souhaite calculer l'aire du domaine D défini par : $M(x;y) \in D \Leftrightarrow \begin{cases} 0 \le x \le 1 \\ 0 \le y \le x^2 \end{cases}$.



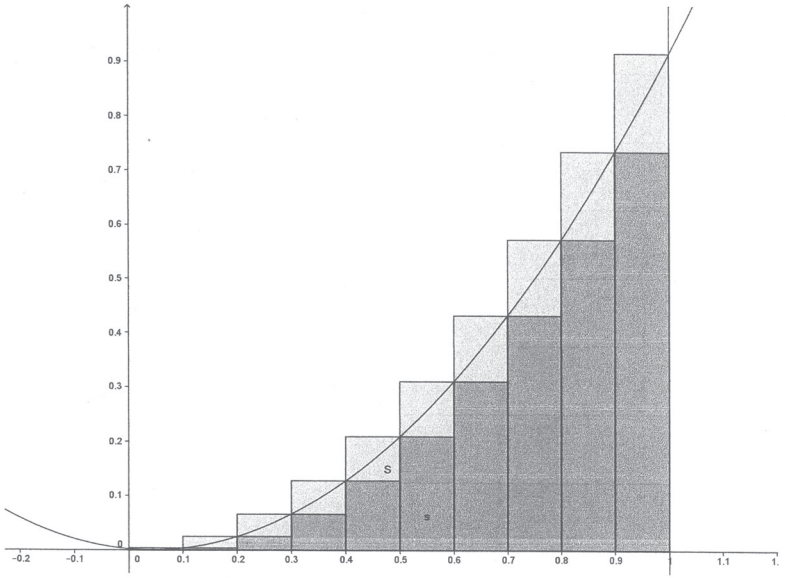
ac papier milli en dessous



Echelle en cm: 10:1

4) **La méthode des rectangles** : On considère toujours la fonction carrée sur $[0;1]$.

On subdivise l'intervalle $[0;1]$ en n intervalles de même longueur –. Sur chaque subdivision, on construit un « petit » rectangle r (hachuré sur le dessin) et un « grand » rectangle R.
L'aire que l'on cherche à calculer, est comprise entre la somme s des aires des « petits » rectangles et la somme S des aires des « grands » rectangles.
Figure dans le cas n = 10



1. On se place dans le cas où n = 10.
   a) Quelle est l'aire d'un « petit » rectangle r_k de base $\left[\frac{k}{10};\frac{k+1}{10}\right]$, k variant de 0 à 9 ? d'un « grand » rectangle R_k ?
   b) Calculer alors s_{10} et S_{10}. En déduire un encadrement de A.
2. Pour améliorer l'encadrement, on subdivise l'intervalle $[0;1]$ en 100 intervalles de même longueur. Calculer alors s_{100} et S_{100}. Quel encadrement obtient-on pour A ?
3. Cas général : On subdivise l'intervalle $[0;1]$ en n intervalles de longueur $\frac{1}{n}$.
   A est encadrée entre :

$$s_n = \frac{1}{n}\left(0^2 + \left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2 + \ldots\ldots\ldots + \left(\frac{n-1}{n}\right)^2\right) = \frac{1}{n^3}\left(0^2 + 1^2 + 2^2 + \ldots\ldots\ldots + (n-1)^2\right)$$

et $S_n = \frac{1}{n}\left(\left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2 + \ldots\ldots\ldots + \left(\frac{n}{n}\right)^2\right) = \frac{1}{n^3}\left(1^2 + 2^2 + \ldots\ldots\ldots + n^2\right)$.

Déterminer les limites des suites (s_n) et (S_n). En déduire la valeur de ~~A~~.

TP 2 : Une aire variable

Partie 1 :

Soit f une fonction continue et positive sur $\mathbb{R}$. On considère alors la fonction F définie sur $[0;+\infty[$ par $F(x)=\int_0^x f(t)dt$.

1. Interpréter géométriquement $F(x)$ pour x appartenant à $[0;+\infty[$. Que vaut $F(0)$?
2. Par un calcul d'aire, exprimer $F(x)$ en fonction de x dans les cas suivants :
   a) $f(t)=3$ b) $f(t)=2t+1$
3. Que remarquez-vous ?

Partie 2 :Soit f une fonction continue, positive et strictement croissante sur [a ; b]. On considère alors la fonction F définie sur $[a\,;b]$ par $F(x)=\int_0^x f(t)dt$. On souhaite montrer que F est dérivable en tout point x_0 de $[a;b]$ et que,
pour tout $x\in[a;b]$.....
Pour cela, il faut.....

Soit x_0 un réel de [a ; b] et h un réel non nul tel que $x_0+h\in[a;b]$.

1er cas : $h>0$.

$F(x_0+h)-F(x_0)$ représente .....

On en déduit que : ... ... $\leq F(x_0+h)-F(x_0)\leq \cdots$ ... ...

puis que : ... ... $\leq \frac{F(x_0+h)-F(x_0)}{h}\leq \cdots$ ... .

2 ième : $h<0$.

$F(x_0)-F(x_0+h)$ représente .....

On en déduit que : ... ... $\leq F(x_0)-F(x_0+h)\leq \cdots$ ... ...

puis que : ... ... $\frac{F(x_0)-F(x_0+h)}{h}$ ... .

Conclusion :

Calcul intégral

I Calcul d'aire sous la courbe. cf TP1

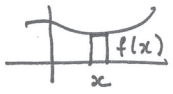
Synthèse : On admet que, si f est une fonction déf, continue et positive sur un intervalle $[a ; b]$ alors la méthode des rectangles permet de construire deux suites (s_n) et (S_n) tq. $s_n \leq \mathcal{A} \leq S_n$ et $\lim\limits_{n\to+\infty} s_n = \lim\limits_{n\to+\infty} S_n = L$ On pose $\mathcal{A} = L$.

II Intégrale d'une fonction continue positive sur $[a ; b]$.

Déf : Soit f une fonction continue et positive sur $[a ; b]$.
On appelle intégrale de f de a à b l'aire en u.a. du domaine situé sous la courbe $\mathcal{C}_f$. On note :
$$\int_a^b f(x)\,dx .$$

dessin

Rq : . $\int_a^b f(x)\,dx$ se lit "intégrale de a à b de $f(x)\,dx$"

. $f(x)\,dx$ est l'aire d'un pt rect 

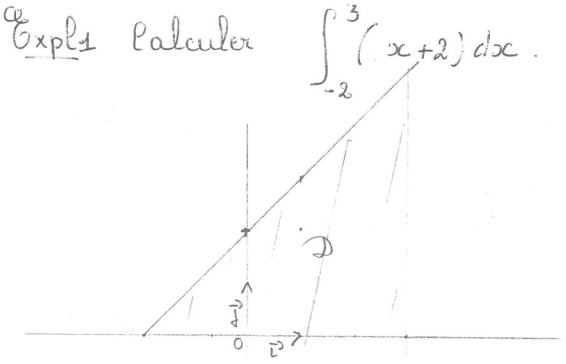
. $\int$ est un S stylisé qui rappelle la somme des aires

. a et b s'appellent les bornes de l'intégrale.

. x est appelée variable "muette" on peut la remplacer par ce qu'on veut.

Exemple : $\int_0^1 x^2\,dx = \frac{1}{3}$ (cf TP).

Expl 1 Calculer $\int_{-2}^{3}(x+2)\,dx$.



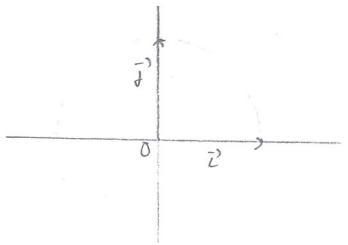
$x \longmapsto x+2$ est continue et ⊕ sur $[-2;3]$
donc $\int_{-2}^{3}(x+2)\,dx$ est l'aire, en u.a, du domaine D

$$\int_{-2}^{3}(x+2)\,dx = \frac{b\times h}{2} = \frac{5\times 5}{2} = 12{,}5 \text{ u.a.}$$

Expl 2: Calculer $\int_{-1}^{1}\sqrt{1-x^2}\,dx$

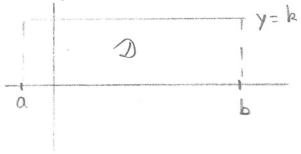
(A l'aide de la calculatrice, tracer C_f où $f : x \longmapsto \sqrt{1-x^2}$ ($x \in I-1;1[$, $y \in [0,1]$)

En déduire I)



$$\int_{-1}^{1}\sqrt{1-x^2}\,dx = \frac{\pi}{2}$$

Expl 3: St $k \in \mathbb{R}^+$ calculer $\int_a^b k\,dx$ $(a \le b)$.



$x \longmapsto k$ est continue et positive sur $[a,b]$
donc $\int_a^b k\,dx = \mathcal{A}_D = k\times(b-a)$.

TP 2.

Théorème 1: Soit f une fonction continue et positive sur $[a;b]$. La fonction F définie sur $[a;b]$ par $F(x) = \int_a^x f(t)\,dx$ est dérivable sur $[a;b]$ et pour tout $x \in [a;b]$ $F'(x) = f(x)$.

démo (principe) dans le cas où f est ↗ sur $[a;b]$.

Rem: Si on connaît une f de F tq $F' = f$ on peut calculer $\int_a^x f(t)\,dt$.

III Primitives d'une fonction.

~~TP: Lien intégrales - primitives~~

Définition : Soit f une fonction définie sur un int. I. On appelle primitive de f sur I toute fonction F déf. et dérivable sur I telle que $F' = f$

Rq : D'après le Th 1, si f est continue et ⊕ sur $[a;b]$ $F : x \mapsto \int_a^x f(t)dt$ est une primit de f sur $[a;b]$

Ce Expl : Donner une primitive de f sur I

a) $f(x) = x^2 + 3$ $\quad I = \mathbb{R}$

b) $f(x) = \cos x$ $\quad I = \mathbb{R}$

c) $f(x) = \frac{1}{x}$ $\quad I =]0;+\infty[$

Rq : Il existe plusieurs primitives

la prop suivante ~~donne~~ précise le lien e/ les ≠ primit

Prop 1 : St f une fonction définie sur un int I.
Si F est une primit de f sur I alors
l'ens. des primit de f sur I est
l'ens. des fonctions de la forme $F + k$, où $k \in \mathbb{R}$

démo : hyp : F est une primit de f sur I.
$(F+k)' = F' + 0 = f$ donc $F+k$ est une primit de f sur I.
Réciproque : Soit G une primit de f sur I.
$(-F+G)' = F' - G' = f - f = 0$
$-F + G$ est une primit de f sur I
$\exists k \in \mathbb{R} \quad G - F = k$
$\exists k \in \mathbb{R} \quad G = F + k$.

Cond : Si f admet des primit sur un int alors elle en admet une ∞ : Deux d'entre elles, différent d'une cst.

Prop 2 : Soit f une fonction déf. sur un int I et admettant sur I une primit. F.
Soit $x_0 \in I$ et $y_0 \in \mathbb{R}$.
Il existe une unique primitive de f sur I qui prend la valeur y_0 en x_0

dessin courbes intégrales

démo : D'après la prop 1, tt primit de f sur I est de la forme $F + k$ où $k \in \mathbb{R}$
$(F+k)(x_0) = y_0$
$F(x_0) + k = y_0$
$k = y_0 - F(x_0)$ (on trouve une unique valeur)
St $G : x \longmapsto F(x) + y_0 - F(x_0)$ est la primit de f sur I / $G(x_0) = y_0$.

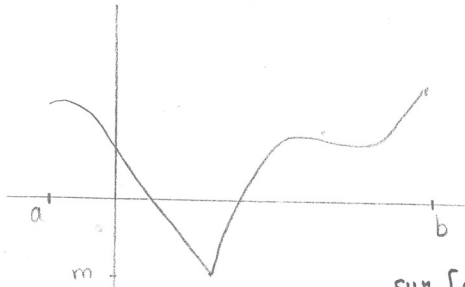
Calcul de primitives

Voir tableau.

Théorème : Toute fonction continue sur un intervalle I admet des primitives sur I.

Ce Th est donné avant la prop 1.

Principe de la démo pour f continue sur et admettant un minimum m sur



. $\forall x \in$, $f(x) \geqslant m$
$\Leftrightarrow \forall x \in$, $f(x) - m \geqslant 0$
La fonction $x \longmapsto f(x) - m$ est continue et positive sur . D'après le Th 1, ~~$F : x \mapsto \int_a^x (f(t) - m)\,dt$ est dérivable sur $[a;b]$ et $F'(x) = f(x) - m$.~~
elle admet une primit F sur $[a;b]$.

IV Intégrale d'une f^c de sgn qqc.

Théorème 3: Soit f une fonction continue et positive sur $[a;b]$ et F une primitive de f sur $[a;b]$. On a:

$$\int_a^b f(t)\,dt = F(b) - F(a)$$

démo: f est continue et positive sur $[a;b]$ donc on sait que $\Phi: x \longmapsto \int_a^x f(t)\,dt$ est une primitive de f sur $[a;b]$.
St F une primit qqc de f sur $[a;b]$
On sait qu'il existe $k \in \mathbb{R}$ tq
$F = \Phi + k$.

$$\begin{aligned} F(b) - F(a) &= \Phi(b) + k - \Phi(a) - k \\ &= \Phi(b) - \Phi(a) \\ &= \int_a^b f(t)\,dt - 0 \end{aligned}$$

cqfd

On admet que le résultat précedent se généralise à tte f^c f continue de signe qqc sur un intervalle I. } tmd

Définition: Soit f une fonction continue sur un intervalle I, a et b deux réels de I et F une primitive de f sur I l'intégrale de f entre a et b est le réel noté $\int_a^b f(t)\,dt$ défini par $\int_a^b f(t)\,dt = F(b) - F(a)$.

$[a;b]$ et on étend à a et b qq après Charles ?

Not: $\int_a^b f(x)\,dx = \left[F(x)\right]_a^b = F(b) - F(a)$

F calculée entre a et b

Exemples (donnés par les élèves)

Rq: Ds le cas où f n'est de plus ⊕, l'intégrale est un nbre de signe qqc.

III Propriétés ~~algébriques~~ de l'intégrale

Les prop sont admises. Pour des fonctions positives sur un int. $[a,b]$, l'interprétation de l'intégrale comme une aire rend ces prop évidentes.

a) Prop alg

Prop 1 : Si f une fonction continue sur un intervalle I.

a/ Pour tt réel $a \in I$ $\int_a^a f(x)\,dx = 0$

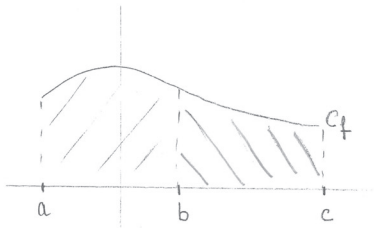
b/ Pour tt réel a, b, c de I

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx$$

(Relation de Chasles).

c/ $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$

Interprétation si f continue sur $a < b < c$.



$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \mathcal{A}_D + \mathcal{A}_{D'} = \mathcal{A}_{D \cup D'} = \int_a^c f(x)\,dx$

Prop 2 (linéarité) : f et g désignent 2 fonctions continues sur I

d/ Pour tout réel a, b de I

$$\int_a^b (\lambda f(x) + \mu g(x))\,dx = \lambda \int_a^b f(x)\,dx + \mu \int_a^b f(x)\,dx$$

où λ et μ sont des cst réels

(linéarité de l'intégrale)

Expl : 1/ Que vaut $\int_0^1 x^2\,dx$?

2/ Calculer $\int_0^1 (x+3)\,dx$.

3/ En déduire $\int_0^1 (5x^2 + 3x + 3)\,dx$ $\left(\frac{31}{6}\right)$

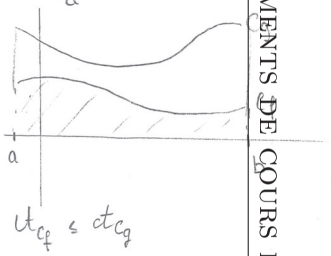
➡

b) Intégrales et inégalités

Prop 2 : Soit f et g deux fonctions continues sur $[a;b]$ $(a \leq b)$.

a/ Si $f \geq 0$ sur $[a,b]$ alors $\int_a^b f(x)\,dx \geq 0$ (Positivité)

b/ Si $f \leq g$ sur $[a;b]$ alors $\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$ (Croissance)



c/ Si, pour tout $x \in [a;b]$ $m \leq f(x) \leq M$

alors $m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a)$ (Inégalité de la moyenne)

dessin

dém

a/ cf $f \geq 0$ et cont. sur $[a,b]$. Par déf. $\int_a^b f(x)\,dx$ est une aire donc ①

b/ $\forall x \in [a;b]$ $\quad f(x) \leq g(x)$

$\iff \forall x \in [a;b]$ $\quad 0 \leq g(x) - f(x)$

$\overset{(a)}{\Rightarrow} \int_a^b [g(x) - f(x)]\,dx \geq 0$

$\overset{\text{linéarité}}{\Rightarrow} \int_a^b g(x)\,dx - \int_a^b f(x)\,dx \geq 0$

$\iff \int_a^b g(x)\,dx \geq \int_a^b f(x)\,dx$

(8

c/ $\forall x \in [a;b] \quad m \leq f(x) \leq M$

donc, d'après b,

$$\int_a^b m\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b M\,dx$$

$$m\int_a^b dx \leq I_a^b \leq M\int_a^b dx$$

$$m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a)$$

Expl: Montrer que $\int_1^3 \sin(t^2)\,dt \leq 2$.

$\forall t \in \mathbb{R} \quad \sin(t^2) \leq 1$

$$\Rightarrow \int_1^3 \sin(t^2)\,dt \leq \int_1^3 1\,dt$$

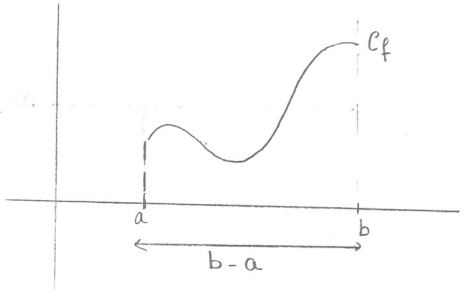
$$\Leftrightarrow \int_1^3 \sin(t^2)\,dt \leq 2.$$

Rq : Les Th sont utiles pr encadrer une intégrale que l'on sait pas calculer.

3. Valeur moyenne d'une fonction.

Déf: Soit f une fonction cont. sur $[a;b]$ $(a<b)$
la valeur moyenne de f sur $[a;b]$
est égale à $\mu = \frac{1}{b-a}\int_a^b f(t)\,dt$.

Interprétation graphique (cas où $f \oplus$)



Ci $f \oplus$ sur $[a;b]$ alors
$\int_a^b f(t)\,dt$ est $\mathcal{A}_D$
$\mu = \frac{\mathcal{A}_D}{b-a}$ = hteur du rectangle de b-a qui a la m aire que $\mathcal{D}$

Si f est de sgr qqc la valeur moyenne est l'unique $f \in$ cst sur $[a;b]$ qui a la m intégrale que f sur $[a;b]$

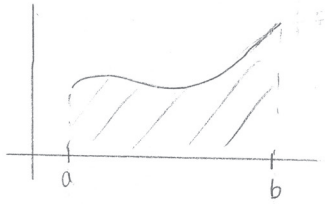
Expl 1: Calculer la valeur moyenne de f définie par $f(x) = 4x^2 - 5x + 2$ sur $[1;3]$

Appl. en cinématique:
Déterminer la vitesse moy Exemple

V Applications du calcul intégral.

1. Calcul d'aires planes

L'u.a. est aire du quad. construit sur les v. de base.

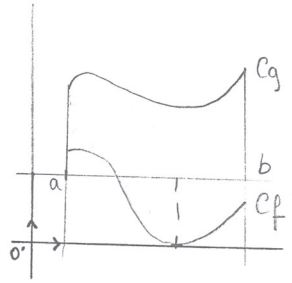


Th: 1/ Si f est cont. et $\oplus$ sur $[a;b]$
L'aire ss la courbe est, en u.a, égale à $\int_a^b f(t)dt$

2/ Si f est cont. et $\ominus$ sur $[a;b]$
l'aire de la surface définie par $a \leq x \leq b$ et $f(x) \leq y \leq 0$ est $-\int_a^b f(t)\,dt$

3/ Si f change de signe

4/ Si f et g sont 2 fonctions continues sur $[a;b]$ tq $f(x) \leq g(x)$
L'aire de la surface def par $a \leq x \leq b$ et $f(x) \leq y \leq g(x)$ est $\int_a^b [g(x) - f(x)]\,dx$.

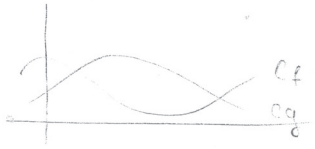


dém : 1/ cf déf. de l'intégrale
2/ Raison de symétrie
3/ On se ramène au cas où $0 < f(x) \leq g(x)$
en changeant le repère / transl $t_{-m\inf \vec{j}}$. L'aire reste inchangée et

$$\mathcal{A}_D = \mathcal{A}_C - \mathcal{A}_{C_f} = \int_a^b (g(x)+M)\,dx - \int_a^b (f(x)+M)\,dx$$

$$= \int_a^b [g(x) - f(x)]\,dx$$

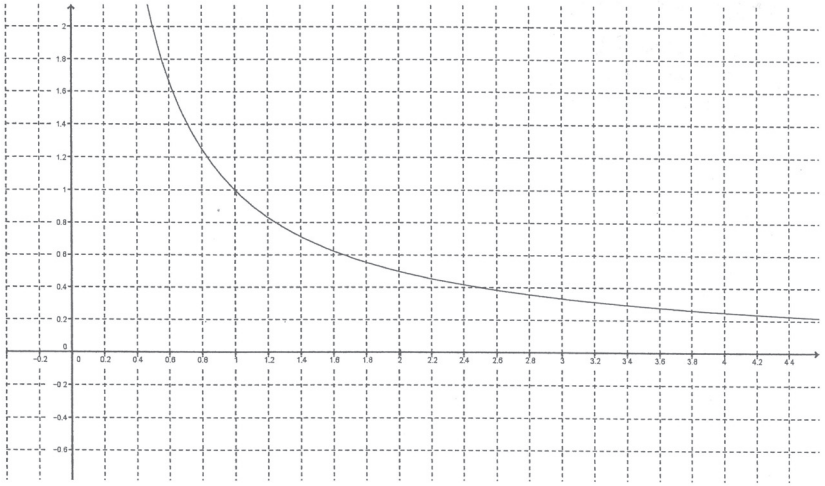
Cas où



Rem: Pour effectuer un calcul, il faut au préalable connaître la posⁿ relative des courbes

Intégrales **Feuille 1**

Exercice 1 :
Utilisez la méthode des rectangles sur l'intervalle $[1,2]$ avec une subdivision de largeur 0,2 pour encadrer l'aire sous la courbe représentative de la fonction inverse sur $[1;2]$.



Exercice 2 : En utilisant un calcul d'aire, calculer les intégrales suivantes (vous commencerez par faire un dessin) :

a) $\int_2^3 (x-2)\,dx$

b) $\int_{-7}^{-4} 5\,dx$

c) $\int_{-5}^{3} |x|\,dx$

d) $\int_{-3}^{8} |x+2|\,dx$

Intégrales **Feuille 2**

Exercice 3 : On considère la fonction carré et C sa courbe représentative dans un repère du plan. On souhaite encadrer l'aire sous la courbe sur un intervalle $[a;b]$ en utilisant la méthode des rectangles. Pour cela, on écrit un algorithme (le retour !).

Variables :
a, b, h, U et V sont des variables réelles
n est une variable entière
Entrées :
Lire a
Lire b
Lire n
Traitement :
Affecter 0 à U
Affecter 0 à V
Affecter $\frac{b-a}{n}$ à h
Pour k allant de ..... à..... faire
Affecter $U + h \times (a + k \times h)^2$ à U
Affecter $V + h \times (a + (k+1) \times h)^2$ à V
Fin de pour
Sortie :
Afficher U
Afficher V

a) Compléter cet algorithme.
b) Qu'affiche l'algorithme si on entre $a = 0, b = 1$ et $n = 5$?
c) Programmer cet algorithme sur votre calculatrice. Qu'affiche le programme pour $a = 0, b = 1$ et $n = 100$?

Intégrale — feuille 3

Ex 7 : Déterminer une primitive de f sur l'intervalle I.

a) $f(x) = -5x + \pi$ $\quad I = \mathbb{R}$

b) $f(x) = \dfrac{3}{x^2}$ $\quad I = \mathbb{R}_+^*$

c) $f(x) = \dfrac{-5}{x}$ $\quad I =]0;+\infty[$

d) $f(x) = \dfrac{x}{\sqrt{x^2+1}}$ $\quad I = \mathbb{R}$

e) $f(x) = x e^{x^2}$ $\quad I = \mathbb{R}$

f) $f(x) = \dfrac{x}{x^2+4}$ $\quad I = \mathbb{R}$

g) $f(x) = e^{-x}$ $\quad I = \mathbb{R}$

h) $f(x) = \cos(3x+1)$ $\quad I = \mathbb{R}$

i) $f(x) = \dfrac{e^x}{e^x+2}$ $\quad I = \mathbb{R}$

j) $f(x) = \dfrac{4}{(x+1)^2}$ $\quad I = \mathbb{R}$

k) $f(x) = e^x (e^x+3)^2$ $\quad I = \mathbb{R}$

Ex 8 : Déterminer les réels a et b pour que F soit une primitive de f sur I.

a) $F(x) = (ax+b)e^x$; $f(x) = 3xe^x$; $I = \mathbb{R}$

b) $F(x) = (a \ln x + b)x^2$; $f(x) = x \ln x$; $I =]0;+\infty[$

Ex 9 : Soit $f : x \longmapsto \dfrac{5x^2+21x+22}{(x-1)(x+3)^2}$ définie sur $]1;+\infty[$.

1. Montrer qu'il existe 3 réels a, b et c tels que $\forall x \in]1;+\infty[$, $f(x) = \dfrac{a}{x-1} + \dfrac{b}{x+3} + \dfrac{c}{(x+3)^2}$.
2. En déduire une primitive de f sur $]1;+\infty[$ puis, la primitive de f sur $]1;+\infty[$ qui s'annule en 2.

Intégrales — feuille 5

Exercice 10 : Calculer les intégrales suivantes.

$I_1 = \displaystyle\int_0^3 (x-4)\,dx$

$I_2 = \displaystyle\int_{-2}^{-1} \left(x^2 - \frac{1}{x^2}\right) dx$

$I_3 = \displaystyle\int_0^2 (2x+1)(x^2+x+1)\,dx$

$I_4 = \displaystyle\int_1^2 \frac{dt}{t^5}$

$I_5 = \displaystyle\int_{-\pi/6}^{3\pi/2} \cos t\,dt$

$I_6 = \displaystyle\int_1^3 \frac{du}{u+1}$

$I_7 = \displaystyle\int_{-3}^3 x e^2\,dx$

$I_8 = \displaystyle\int_1^2 e^{2x}\,dx$

$I_9 = \displaystyle\int_0^{\pi/2} \sin 2t\,dt$

$I_{10} = \displaystyle\int_{-7}^{-9} dt$.

$I_{11} = \displaystyle\int_1^2 \frac{dx}{3x+2}$

$I_{12} = \displaystyle\int_{-\pi/6}^{\pi/4} \sin^2 t\,dt$

$I_{13} = \displaystyle\int_0^1 t e^{t^2-1}\,dt$

$I_{14} = \displaystyle\int_e^{e^2} \frac{1}{t \ln t}\,dt$

Exercice 11 :

a) Déterminer les réels a, b, c tels que : pour tout $t \in \mathbb{R} \setminus \{\frac{1}{2}\}$, $\dfrac{t^2-1}{2t-1} = at + b + \dfrac{c}{2t-1}$

b) Calculer $\displaystyle\int_{-1/2}^0 \frac{t^2-1}{2t-1}\,dt$

Intégrales **feuille 6**

Exercice 12 :

1. Soit f la fonction définie sur $[1 ; +\infty[$ par $f(x) = \dfrac{x}{e^x - 1}$ et soit H la fonction définie sur $[1 ; +\infty[$ par $H(x) = \displaystyle\int_1^x f(t)\,dt$.

a) Justifier que f et H sont bien définies sur $[1 ; +\infty[$.

b) Quelle relation existe-t-il entre H et f ?

c) Soit C la courbe représentative de f dans un repère orthonormal $(O ; \vec{i}, \vec{j})$ du plan. Interpréter en termes d'aire le nombre $H(3)$.

2. On se propose, dans cette question, de donner un encadrement du nombre $H(3)$.

a) Montrer que, pour tout réel $x > 0$, $\dfrac{x}{e^x - 1} = x \times \dfrac{e^{-x}}{1 - e^{-x}}$.

b) En déduire que $\displaystyle\int_1^3 f(x)\,dx = 3 \ln\left(1 - \frac{1}{e^3}\right) - \ln\left(1 - \frac{1}{e}\right) - \int_1^3 \ln(1 - e^{-x})\,dx$.

c) Montrer que, si $1 \leqslant x \leqslant 3$, alors $\ln\left(1 - \dfrac{1}{e}\right) \leqslant \ln(1 - e^{-x}) \leqslant \ln\left(1 - \dfrac{1}{e^3}\right)$.

d) En déduire un encadrement de $\displaystyle\int_1^3 \ln(1 - e^{-x})\,dx$, puis de $\displaystyle\int_1^3 f(x)\,dx$.

Exercice 13 :

On considère la suite de terme général : $u_n = \sum_{k=1}^{k=n} \dfrac{1}{n+k}$ où n est un entier naturel non nul.

Il est conseillé……. d'écrire u_n sous forme étendue.

1. Montrer que, pour $n \geq 1$ et $k \geq 1$:
$$\int_{n+k}^{n+k+1} \frac{1}{t}dt \leq \frac{1}{n+k} \leq \int_{n+k-1}^{n+k} \frac{1}{t}dt$$
En donner une interprétation graphique.
2. En déduire un encadrement de u_n.
3. Etudier la convergence de la suite (u_n).

Intégrales **feuille** $|2 + i\sqrt{3}|^2$

Exercice 14 :

On considère la fonction f définie sur $[0 ; +\infty[$ par :
$$f(x) = \frac{\ln(x+3)}{x+3}.$$

Montrer que f est dérivable sur $[0 ; +\infty[$. Étudier le signe de sa fonction dérivée f', sa limite éventuelle en $+\infty$ et dresser le tableau de ses variations.

On définit la suite $(u_n)_{n \geqslant 0}$ par son terme général $u_n = \displaystyle\int_n^{n+1} f(x)\,dx$.

Montrer, sans chercher à calculer u_n, que, pour tout entier naturel n,
$$f(n+1) \leqslant u_n \leqslant f(n).$$
En déduire que la suite (u_n) est convergente et déterminer sa limite.

Soit F la fonction définie sur $[0 ; +\infty[$ par $F(x) = [\ln(x+3)]^2$.

Justifier la dérivabilité sur $[0 ; +\infty[$ de la fonction F et déterminer, pour tout réel positif x, le nombre $F'(x)$.

On pose, pour tout entier naturel n,
$$I_n = \int_0^n f(x)\,dx.$$
Calculer I_n.

On pose, pour tout entier naturel n,
$$S_n = u_0 + u_1 + \ldots + u_{n-1}.$$
Calculer S_n. La suite (S_n) est-elle convergente ?

→ Exo proba où la techn d'intégration est importante

F.3.2 Documents relatifs au chapitre « Variables aléatoires continues à densité »

Introduction à la notion de densité de probabilité

On s'intéresse aux clients d'une banque qui possèdent un crédit à la consommation. On choisit un client au hasard parmi ces derniers et on note X la variable aléatoire égale à la mensualité de son crédit à la consommation . Les mensualités de remboursement sont comprises entre 100 et 400 euros par mois.
Pour définir la loi de probabilité de X, on a effectué une enquête auprès des clients de la banque dont les mensualités de remboursement de crédit à la consommation sont comprises entre 100 et 400 euros par mois. On a tracé (annexe 1) l'histogramme des fréquences obtenues.

1) Quelle est l'aire totale de l'histogramme ?
2) En choisissant comme loi de probabilité de X la distribution des fréquences donnée en annexe 1 , déterminer si possible, la probabilité des événements suivants :
$(X \leq 200)$ $(200 \leq X \leq 350)$ $(X = 168)$ $(180 \leq X \leq 220)$.

3) On a regroupé les résultats de l'enquête dans des classes d'amplitude 10.
On note I_k l'intervalle $[10k; 10(k+1)]$ où k est un entier compris entre 10 et 39.
Exemple : $I_{10} = [100; 110]$. L'histogramme des fréquences est donné en annexe 2.

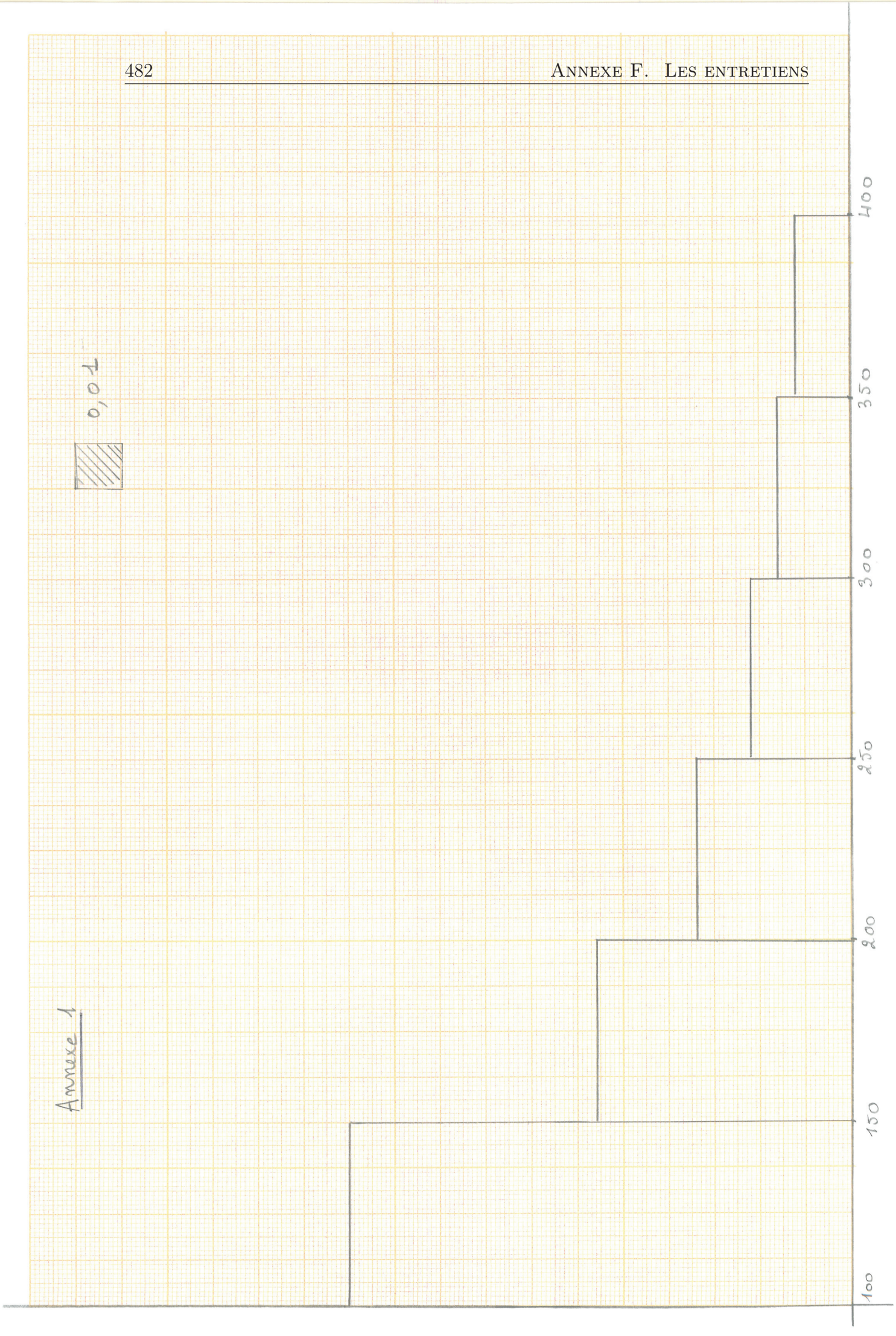
I_{10}	I_{11}	I_{12}	I_{13}	I_{14}	I_{15}	I_{16}	I_{17}	I_{18}	I_{19}
0,121	0,101	0,085	0,073	0,063	0,056	0,049	0,044	0,039	0,035
I_{20}	I_{21}	I_{22}	I_{23}	I_{24}	I_{25}	I_{26}	I_{27}	I_{28}	I_{29}
0,032	0,029	0,026	0,024	0,022	0,021	0,019	0,018	0,016	0,015
I_{30}	I_{31}	I_{32}	I_{33}	I_{34}	I_{35}	I_{36}	I_{37}	I_{38}	I_{39}
0,014	0,013	0,013	0,012	0,011	0,011	0,010	0,010	0,009	0,009

Calculer , si possible, les probabilités des événements :
$(180 \leq X \leq 220)$; $(122{,}50 \leq X \leq 143{,}70)$.

En utilisant des intervalles d'amplitude de plus en plus petite, on construit des histogrammes d'aire 1 et on peut déterminer une fonction continue f telle que les histogrammes seront de plus en plus proches du domaine sous la courbe de la fonction f.

4) Dans notre exemple, on peut prendre la fonction définie sur $[100; 400]$ $par\ f(x) = \frac{400}{3x^2}$.
a) Tracer soigneusement[1] la courbe représentative de la fonction f sur l'annexe 2.
b) Justifier que : $\int_{100}^{400} f(x)dx = 1$. De façon générale, pour tous rées a et b de$[100; 400]$, comment calculerez-vous P$(a \leq X \leq b)$?
c) Calculer, $P(100 \leq X \leq 350)$ et $P(180 \leq X \leq 220)$, puis $P(122{,}50 \leq X \leq 143{,}70)$. Que vaut $P(X = 168)$?

[1] Cette précision s'adresse tout particulièrement à schruch..schruch.oh un tunnel.....


0,01
400
350
300
250
200
150
100
Annexe 1

C_4
Annexe 2
250
200
150

Variables aléatoires continues à densité

Introduction

Synthèse n°1 : Auparavant les v.a étudiées prenaient un nbr fini de valeurs (v.a discrète). Il existe de nbr situations où la v.a considérée prend * toutes les valeurs d'un intervalle I (temps d'attente à la caisse, distance entre 2 pts ...)
On parle alors de v.a. continues.
Pour définir la loi de proba. de X, on ne peut plus calculer $p(X=x_i)$ où $x_i \in X(\Omega)$ puisque il y a une infinité, de x_i non dénombrable
On est amené à calculer $P(X \in J)$ pour tout intervalle $J \subset I$.

On considère une expérience aléatoire et Ω son univers sur lequel on a défini une proba.

* au moins théoriq^{t}

Synthèse n°2 :

Doc^{t} : • Les xpl du cours
• Courbe cloche
• normale σ
• moins daplace

I Densité de probabilité

Déf 1 : On appelle fonction de densité de probabilité sur un intervalle I toute fonction f définie, continue et positive sur I telle que $\int_I f(t)\,dt = 1$

Voca : On dit aussi plus rapid^{t} " densité de proba ".

Rq : I est un intervalle qqc. de $\mathbb{R}$:

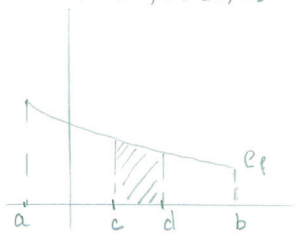
↳ Si I est un int. fermé borné
d.e. $I = [a;b]$ où $a<b$
alors $\int_I f(t)\,dt = \int_a^b f(t)\,dt$

↳ Si I est un int. non borné
par exemple : $I =]-\infty; a]$
alors $\int_I f(t)\,dt = \underbrace{\int_{-\infty}^{a} f(t)\,dt}$
non étudiée en term.
nous admettrons que
si $\int_x^a f(t)\,dt$ a une limite finie l
qd $x \to -\infty$, on pose $\int_{-\infty}^{a} f(t)\,dt = l$

Si $\lim\limits_{x\to-\infty} \int_x^a f(t)\,dt$ est un réel l,
on pose $\int_I f(t)\,dt = l$.

Déf 2: Soient X une v.a continue sur un intervalle I et f une densité de proba. sur I.
On dit que X a pour densité de proba. f si, et seulement si, pour tout intervalle $J \subset I$,
$p(X \in J) = \int_J f(t)\,dt$. ($p(X \in J)$ est l'aire sous $\mathcal{C}_f$ sur J).

$I = [a;b]$, $J = [c;d]$



$P(X \in [c;d]) =$ aire ss $\mathcal{C}_f$ entre c et d

$p(X \in [c;d]) = \int_c^d f(t)\,dt$

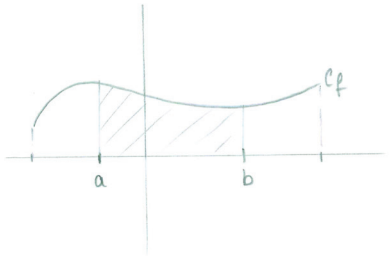
Prop 1: Soit X une v.a continue sur I de densité de proba f.

1) $\forall a \in I$, $p(X = a) = 0$
2) $\forall a \in I$, $p(X > a) = 1 - p(X \leq a)$
3) Pour tout réel a, b de I tels que $a < b$, on a:
$p(a \leq X < b) = p(X < b) - p(X < a)$

1) $p(X = a) = \int_a^a f(t)\,dt = 0$ Cons. $p(a < X < b) = p(a \leq X < b) = p(a < X \leq b) = p(a \leq X \leq b)$

2) Les événements $(X > a)$ et $(X \leq a)$ forment une partition de I.
$I = (X > a) \cup (X \leq a)$ où $(X > a) \cap (X \leq a) = \emptyset$
$p(I) = p(X > a) + p(X \leq a)$
$1 = p(X > a) + p(X \leq a)$
$\Leftrightarrow p(X > a) = 1 - p(X \leq a)$

3) $(X < b) = (X < a) \cup (a \leq X < b)$ union disjointe



Exemple 1: La prod quot X d'un pdt en tonnes est une v.a continue sur $[0;10]$ dt la densité de proba est
$f : x \longmapsto 0,006(10x - x^2) = -0,006x^2 + 0,06x$
a) Vérifier que f est une densité de proba
b) Calculer $p(X \leq 7)$
B: "la prod quot dépasse 6 t"
c) Calculer $E(X)$.

a) f est une fct poly de d° 2
f est continue sur $[0;10]$ et $\oplus$.
($\Delta > 0$ et $\exists$ 2 racines 0 et 10)

x	0		10
	0	+	0

$\int_0^{10} (-0,006x^2 + 0,06x)\,dx$
$= \left[-\frac{0,006}{3}x^3 + \frac{0,06}{2}x^2 \right]_0^{10}$
$= -0,002 \times 10^3 + 0,03 \times 10^2 = 1$

b) $p(X \leq 7) = \int_0^7 f(x)\,dx = 0,784$
$p(B) = p(X > 6) = 1 - p(X \leq 6)$
$= 0,352$.

c) $E(X) = \int_0^{10} x f(x)\,dx = \int_0^{10} (-0,006x^3 + 0,06x^2)\,dx$
$= \left[-\frac{0,006}{4}x^4 + \frac{0,06}{3}x^3 \right]_0^{10}$
$= 5$

Exemple 1 :

La production quotidienne d'un produit en tonnes est une variable aléatoire continue sur $[0; 10]$ dont la densité de probabilité est la fonction f définie sur $[0; 10]$ par $f(x) = 0{,}006(10x - x^2)$.

1. Vérifier que f est une densité de probabilité.
2. Calculer $p(X \leq 7)$ et la probabilité de l'événement B : « La production quotidienne dépasse 6 tonnes ».

Exemple 2 : Tous les matins, un élève dont je tairai le nom (N'insistez- pas !) arrive au hasard à l'arrêt de bus entre 7h15 et 7h35. A partir de 7h00, les bus passent exactement[1] toutes les 10 minutes. On note X la variable aléatoire égale à l'heure d'arrivée de cet élève à l'arrêt de bus.

1. Quelle est la loi de X ?
2. Quelle est la probabilité que cet élève attende moins de 5 minutes ?
3. S'il rate le bus de 7h30, cet élève est en retard au lycée. Quelle est la probabilité que cet élève soit en retard au lycée[2] ?

Exemple 3 : La durée de vie en heures d'un certain type d'ampoule est une variable aléatoire X qui suit une loi exponentielle de paramètre 0,002.

1. Calculer la probabilité pour qu'une telle ampoule ne tombe pas en panne avant 100h.
2. Calculer la probabilité pour qu'une telle ampoule fonctionne au bout de 600h sachant qu'elle fonctionne encore au bout de 500h.

[1] Oui, au pays des maths les bus passent à l'heure!
[2] Et, oui même au pays des maths ,les élèves eux sont en retard.....rien n'est parfait !

F.4 Analyse de l'entretien de l'enseignante Marie

Parmi les quatre enseignants interrogés, nous avons choisi d'analyser plus en détails l'entretien de l'enseignante D, que nous avons appelée Marie. Dans cette annexe, il ne s'agit de présenter l'ensemble de son entretien mais de dégager des éléments révélateurs de son discours, à mettre en relation avec les composantes relatives au métier d'enseignant (Robert, 2008a). Il nous semble que ces éléments peuvent être utiles à connaître notamment dans le but ensuite de construire une séquence d'enseignement viable en classe.

Parmi les quatre enseignants interviewés (cf. section 5.4.1), nous avons choisi de nous intéresser particulièrement à l'entretien avec l'enseignante Marie pour plusieurs raisons. Les enseignants A et B ont montré un point de vue assez « théorique » sur les lois à densité, voyant ce chapitre comme un prétexte pour faire du calcul intégral, sans essayer de donner réellement de sens à la fonction de densité ou même plus généralement aux probabilités. Le point de vue des enseignantes C et D est différent, elles ont notamment toutes les deux proposé une *activité* d'introduction à la fonction de densité, on sent donc chez ces deux enseignantes la volonté d'essayer de donner du sens à cet objet. Parmi ces deux enseignantes, l'enseignante D a donné beaucoup plus d'éléments sur les contraintes qui pèsent sur son enseignement.

Il s'avère que Marie est l'enseignante avec qui nous avons travaillé pour la conception et l'expérimentation de l'ingénierie, dans la partie Partie III.

F.4.1 Méthodologie relative à l'analyse de contraintes institutionnelles et personnelles

L'entretien avec Marie a duré environ 1h05. Il est entièrement transcrit en annexe F.2, dont les lignes sont numérotées.

À travers cet entretien, nous allons essayer de comprendre le point de vue de Marie sur l'enseignement du calcul intégral et des probabilités à densité, leurs enjeux, sur l'apprentissage des élèves sur ces notions et surtout sur les contraintes aussi qui peuvent peser sur l'enseignement. Il s'agit simplement pour nous de présenter quelques éléments de l'entretien avec Marie, révélateur de quelques points de l'ETM idoine proposé par l'enseignante dans sa classe et surtout sur son ETM personnel relatif particulièrement aux probabilités à densité. Nous les catégoriserons en fonction des composantes du métier d'enseignant de la double approche (Robert, 2008a).

F.4.2 Des points qui vont dans le sens de nos résultats précédents

Dans l'entretien avec Marie , nous pouvons retrouver des points qui confortent ce qui a pu être mis en évidence dans les chapitres 4 et 5. Tout d'abord, il ressort bien chez Marie l'importance des liens entre intégrale et aire et entre probabilité et aire. Pour le deuxième cas, par exemple, Marie dit :

> *On voit comment calculer toutes sortes de probabilités avec l'aire sous la... enfin en interprétant en termes d'aire sous la courbe...* (lignes 552-

554)

On a tout interprété en termes d'aire... et en termes de propriétés de proba. (lignes 559-560)

Dans les choix que fait Marie dans son enseignement, on retrouve de fortes similitudes avec ce qui est fait dans les manuels. Tout d'abord, comme nous l'avons dit dans la section 5.4, le calcul intégral est fait en amont des probabilités et Marie insiste en disant :

On voit bien le lien quand même avec le calcul intégral, il serait absolument hors de question de ne pas avoir fait le calcul intégral et calcul d'aire avant [les probabilités à densité]... (lignes 701-703)

Comme dans les manuels, dans un premier temps, elle propose dans son cours un paragraphe de généralités suivi par quelques exemples divers, avant de se concentrer sur les trois lois au programme :

J'ai pas mis de définition pour variable aléatoire continue. Après, j'en ai mis une pour fonction de densité de probabilité sur un intervalle I : toute fonction continue, positive telle que l'intégrale sur I est égale à 1. (lignes 378-381)

Elle fait directement ressentir son malaise vis à vis du programme :

Et je note sur I parce qu'après on explicite le problème. Quand I est un intervalle fermé borné, comme on avait fait dans l'exemple, c'était l'intervalle $[200; 400]$*, voilà ça on sait faire, on a notre fonction, c'est donc l'intégrale de* 200 *à* 400*, pas de problème, on a fait un chapitre qui servait à ça, y a pas de soucis... Après, il se passe quoi quand I est un intervalle qui est non borné ? Euh donc là je prends l'un des deux exemples de ce type là et voilà, je leur explique que c'est pas étudié en terminale, que ce qu'on va faire, c'est qu'on va mettre l'intégrale de a à X qu'on va chercher à calculer la limite lorsque X tend vers $+\infty$, que si ça, ça admet une limite finie alors on va dire que ça se passe bien et on dirait que c'est ce nombre qu'est... Donc... en leur explicitant bien que voilà, c'est pas au programme de la terminale mais que je vais en avoir quand même besoin parce que j'en ai besoin pour définir des choses dans la suite du cours. Donc je sépare quand même les deux cas et... ça les dérange pas plus que ça en fait... Bon, ils trouvent curieux que d'un seul coup, je fasse des trucs... que j'ai besoin de trucs pas dans le programme, ça par contre mais ils se disent « mais c'est quoi le programme... ce programme ? ». « Je sais pas ». C'est voilà... Voilà, on va pas vous demander vous de*

manipuler quoi que c'est un peu faux mais il faudrait que vous compreniez que ces intégrales avec une borne infinie, c'est pas... voilà, c'est un petit peu particulier et qu'il faut faire un travail supplémentaire. (lignes 381-398)

Ce malaise vis à vis des intégrales impropres, présentes en probabilités et absentes en calcul intégral, se fait sentir dès la définition de la fonction de densité et ensuite lors de l'étude de la loi exponentielle :

Donc on a le problème tout de suite pour montrer que cette fonction est une fonction de densité de probabilité sur $[0; +\infty[$. *Donc là on y va, là ça va parce qu'ils peuvent intégrer sur* $[0; X]$ *et après on cherche la limite et... Donc on utilise ce qu'on a défini en amont, mais sans... sans le préciser au départ, je vois pas comment on peut... on peut s'en sortir...Donc... Je sais que les recommandations...enfin que dans le programme c'est pas, c'est pas excessivement clair... c'est on doit le faire sans le faire mais il y a des démonstrations obligatoires dans lesquelles par exemple pour l'espérance de la loi exponentielle, je vois pas comment tu fais si tu leur as pas expliqué qu'il fallait chercher la limite de... [...] Sans leur expliquer complètement tout mais ils ont compris qu'il y avait un problème qu'il allait falloir devoir travailler un petit peu plus.* (lignes 444-456)

Cela confirme bien l'ambiguïté du programme qui précise que l'on ne fera pas de généralités sur les intégrales impropres mais qui laisse un flou quand à comment aborder tout de même cet objet en probabilités.

Lors de cet entretien, nous sommes aussi revenu sur le fait que les élèves doivent ou non directement appliquer les formules de probabilités données dans le cours dans le cas des lois uniformes et exponentielles. Marie laisse encore entendre que le programme est flou à ce sujet :

Alors je suis restée sur un entre-deux très sympathique... [Rires] qui a été une espèce de consensus avec les élèves... Je leur ai dit « alors pour l'examen, moi je vous conseille de refaire les calculs » parce que tu vois que là je les ai mis en propriétés, alors on les a démontré... mais, parce que quand tu fais cinq exercices de suite, tu n'en peux plus de cette intégrale à calculer. Voilà. Mais voilà je leur ai dit... Je suis assez honnête avec eux... Je leur ai dit que je trouvais le programme flou sur ce qu'était vraiment... ce qu'on leur demandait... et je leur ai dit « le jour du bac, au pire, il y aura une question donc vous me refaites l'intégrale correctement ». Et en contrôle, ils m'ont refait... ils m'ont refait les intégrales enfin il y avait un exercice et ils me l'ont refait. Ils ont bien compris la règle du jeu. Ils sont aussi habitué à ce que je leur dise « là c'est un petit peu pas très clair donc... » il y a deux mesures : il y a le cours au quotidien où on va pas « se taper » quinze fois l'intégrale $\lambda e^{-\lambda t}$, *on n'en peut*

> *plus et puis, le jeu du contrôle et de l'examen où là on fait des choses de plus... et ça leur a pas... ça leur pose pas de problème.* (lignes 484-497)

Nous pouvons repérer dans cette citation des éléments relevant de la composante sociale. Les choix que fait Marie, dans ses classes, sont révélateurs de la relation qu'elle a instauré avec ses élèves particuliers. Il n'est pas sûr que ce serait les mêmes choix dans un autre contexte, par exemple dans un lycée en zone d'éducation prioritaire (ZEP).

Il ressort aussi, ce que nous avions repéré dans l'analyse des sujets du baccalauréat, que les exercices sur les probabilités à densité sont extrêmement formatés. Seuls quelques types de tâches, toujours les mêmes, sont traités. Seul l'habillage change. Marie explique :

> *Comme dans tout le chapitre, les exos sont extrêmement formatés... donc... au bout du deuxième cylindre de loi normale, ils [les élèves] en pouvaient plus... la troisième bouteille à remplir... ils voulaient plus en entendre parler...[Rires]. Et là c'est pareil... une fois qu'ils ont compris ce qu'ils vérifiaient, ils voient plus très bien l'intérêt de se taper ça encore pendant je sais pas combien de temps donc...* (lignes 614-618)

> *Tous ces exos, c'est juste un habillage... On sait pas pourquoi... ça suit... enfin... On a discuté du fait qu'une loi normale quel genre de phénomène ça pouvait traduire, une loi exponentielle aussi... mais maintenant on n'a pas du tout fait de modélisation dans ce chapitre-là... La modélisation est prise en charge. C'est... D'ailleurs les élèves décryptent très bien les énoncés. T'as dix lignes de blabla... paf, une loi normale de paramètre trucmuche, chouette, je vais faire des calculs. [Rires]. Et ça n'a aucun sens... Et la plupart des exercices n'ont aucun sens pour eux... voire sont profondément débiles. « On va longtemps encore mesurer des diamètres de tringles, de machin, de poutres... ? ». Je dis « non, on va remplir des bouteilles ». [Rires]. Alors, on discutait un petit peu de... du fait que pour le coup, ça... ça prend le sens sur du... sur des chaînes de fabrication, des choses comme ça... Ils comprennent que quand on remplit des paquets de café, ça va pas faire 250 grammes et que évidemment on peut pas se permettre de faire tout et n'importe quoi... donc ça, ça prend un peu son sens... mais c'est pas eux qui prennent en charge la modélisation.* (lignes 649-662)

Exactement comme pour le baccalauréat – qui est le principal objectif des élèves en fin d'année scolaire précise Marie (le chapitre a été fait en avril-mai) – les exercices rencontrés sont stéréotypés et la modélisation est toujours prise en charge par l'énoncé.

Marie confirme aussi le fait que les élèves ne maîtrisent pas l'objet « histogramme » (bien qu'elle l'utilise dans son *activité* d'introduction) :

> *J'ai choisi de pas leur faire tracer les histogrammes parce que l'année d'avant, j'avais fait tracer les histogrammes et que ça avait été un massacre et ça avait pris un temps considérablement fou. Et que j'étais... Je me suis dit « Plus jamais un histogramme en terminale S, c'est trop compliqué ». [Rires]. Même en leur donnant l'unité d'aire, ça avait été l'horreur.* (lignes 324-328)

Nous retrouvons des éléments similaires à ceux que nous avons exposés dans les chapitres 4 et 5, mais aussi d'autres éléments apparaissent, notamment des contraintes institutionnelles liées au programme.

F.4.3 Des contraintes institutionnelles

Comme nous l'avons vu dans l'étude du programme, les deux chapitres représentent chacun un huitième du programme. Marie insiste bien sur le fait que le calcul intégral est conséquent :

> *C'est un cours, un chapitre qui est très long. Il faut se rendre compte que c'est un chapitre où faire le cours, ça prend énormément de temps.* (lignes 164-165)

Elle fait ensuite remarquer qu'il en est de même pour le chapitre sur les lois à densité :

> *C'est deux énormes chapitres... et qu'on ne peut pas malgré tout... faire en début d'année. Tu peux pas commencer le calcul intégral au mois de septembre... voilà... le problème, c'est que il faut absolument... il faut un peu construire l'analyse... Ils ont pas les limites, ils ont... ils sont très faibles en dérivation donc... [inaudible]. Enfin il y a plein plein de trucs à faire...* (lignes 705-709)

> *Il y a beaucoup de choses à écrire... Il y a beaucoup de vocabulaire pour eux à se mettre dans la tête tout simplement donc... au début ça démarre doucement quand même.* (lignes 740-742)

> *C'est que c'est très très long de faire juste tout ça. C'est-à-dire, ça nous prend quand même trois semaines.* (lignes 669-670)

En plus du fait que ces deux chapitres sont longs, les contenus du programme impose le fait que ces chapitres soient faits en fin d'année scolaire :

> *Alors c'est aussi une question de progression sur l'année... Moi j'avais envisagé de ne faire que le calcul intégral sur des fonctions positives continues, pour embrayer sur les lois continues... pour casser un peu*

ce chapitre qui est beaucoup trop lourd. Mais ça demandait de faire la fonction log plus tard... parce que la fonction exponentielle on en avait besoin. Et on est dans une équipe, il y a des bacs blancs donc... voilà, on peut pas toujours tout vendre. On peut pas obtenir tout ce qu'on veut donc... Voilà. Et à la fin on se retrouve... voilà avec cet énorme pan proba continues suivi d'échantillonnage... Et c'est très très long... Ça dure cinq semaines. Et il y a beaucoup... Il y a beaucoup de cours parce que quand tu regardes le cours... tout ça, c'est nouveau pour eux. Donc ça prend du temps. Il y a du vocabulaire spécifique, il y a toutes les démonstrations quasiment. Donc ça prend beaucoup de temps pour... une batterie de savoir-faire très faible. Donc... donc c'est compliqué. On s'était dit avec des collègues qu'on allait essayer de réduire le cours... et on n'a pas... enfin on n'a pas réussi à enlever des morceaux. C'est-à-dire on est déjà au strict minimum de ce qu'on peut... vendre pour que ça ait... enfin pour que le cours se tienne. Sans balancer tout. Moi je vois pas trop comment faire. Et il y reste le problème de l'échantillonnage après donc... (lignes 673-689)

Le fait que le chapitre sur les lois à densité soit aussi long entraîne une tension avec le fait que peu de choses soient attendues finalement des élèves en termes de types de tâches. Il semble ressortir un sentiment de « Tout ça pour ça ? ».

F.4.4 Éléments de la composante cognitive

D'après Marie, du côté des élèves, le chapitre des lois à densité ne semble pas représenter de problèmes particuliers, cela est notamment dû au fait que les exercices soient formatés.

Après ce qu'on voit dans le cours, c'est que quand on commence à leur dire que cette fameuse fonction f densité de probabilité, que la probabilité que X soit entre a et b ça sera l'intégrale... Ça, ça ne pose aucune difficulté. (lignes 358-361)

Aucun élève n'arrive pas à faire les exercices de proba continues qui sont dans les manuels... parce que ça relève de l'application... tout simplement. (lignes 818-820)

F.4.5 Des difficultés professionnelles relevant de la composante personnelle

Malgré le fait que les probabilités ne présentent pas de difficultés pour les élèves, cela n'enlève pas le sentiment de non satisfaction de l'enseignante à l'égard de ce chapitre. Au sujet de son *activité* d'introduction, elle dit :

Il y a juste à trouver un exemple intelligent, ce que je n'ai pas encore réussi à faire. (lignes 353-354)

En parlant de l'ensemble de son enseignement sur les lois à densité, elle confie :

Je dois avouer que je... j'ai... Il y a des moments où je sais pas trop comment monter tout ça encore, soit dit en passant. Mais bon, ça va sûrement venir... [Rires]. Faut pas désespérer. (lignes 618-620)

Elle précise qu'au niveau des exercices, elle ne propose rien de « *très original* » :

Entre autres parce que je suis pas super à l'aise. (ligne 639)

Elle ne s'est pas non plus attelée à des problèmes de modélisation. Elle explique :

Honnêtement je ne me suis pas plongée sur le sujet dans le chapitre... mais par manque de connaissance de ma part. (lignes 662-663)

Elle insiste fortement dans l'entretien sur son manque de connaissances sur le sujet, sur son manque de recul et plus généralement sur le manque de formation en probabilités. Elle dit :

Parce que nous tous, dans nos formations, certains d'entre nous plus âgés que moi n'ont pas du tout de stat-proba. Moi j'en ai fait... mais c'est la première fois que je les enseigne. En gros, je vais dire comme ça. Donc on n'est pas aguerri et on a encore beaucoup de mal à se poser les bonnes questions. C'est pas mon domaine de prédilection, ça c'est certain. Donc je vois bien, moi, chaque été je travaille avec des bouquins du supérieur pour vérifier que je... j'ai pas loupé encore un truc ou que je suis sûre de pas raconter de bêtises. Donc on n'a pas... En ce qui me concerne, je n'ai pas suffisamment de recul pour... Tu vois répondre à ta question de tout à l'heure... ou savoir comment je pourrais monter le truc autrement que... Pour monter mon cours moi j'ai pris des livres du supérieur pour le monter de façon logique et qu'il est pas de... tu vois... de problèmes. Donc on est loin du reste du programme quand même tu vois. (lignes 767-778)

Je pense qu'il y a un problème de formation, c'est évident. Et je suis pas... Moi je le dis aisément. Il y a pas de soucis là-dessus mais... J'ai besoin encore de travailler énormément la théorie pour être capable de monter un truc intelligent sur le sujet... et j'ai pas assez de recul dans le domaine... ça c'est sûr. Tu vois quand t'es en analyse, en géométrie, en complexes et tout tu sais quels sont les exercices du post bac, tu...

> *tu sais ce vers quoi tu veux aller. Là moi je l'ai pas du tout là le recul. Donc du coup ça donne un cours qu'est très... très proche d'un truc on va dire magistral quand même même si je me fends d'une activité au départ pour essayer de leur expliquer comment on va calculer ces proba un peu particulières mais voilà... J'ai... C'est ma mauvaise connaissance du domaine moi qui me... qui me brime.* (lignes 779-789)

À travers ces citations, nous pouvons remarquer, dans ce cas, les difficultés rencontrées par l'enseignante du fait de son manque de formation dans ce domaine.

F.4.6 L'émergence d'un besoin

Nous avons, au travers des extraits de l'entretien de Marie, essayé de présenter l'état d'esprit de cette enseignante, qui laisse transparaître le même type d'analyse que nous avons pu repérer dans les manuels, mais qui révèle aussi des points en plus. Tout d'abord, on sent le besoin de formation, même la demande, lié au fait que cet enseignement est nouveau au lycée mais aussi du fait d'un manque certain sur ces notions en formation initiale.

Il ne s'agit pas d'une enseignante débutante, mais d'une enseignante expérimentée (plus de 20 ans d'enseignement), qui plus est, qui a été formatrice au cours de sa carrière pendant 5 années, on peut donc raisonnablement penser que ses besoins sont partagés par un grand nombre de collègues. Cependant, ce besoin de formation se situe un peu sur le contenu disciplinaire, mais aussi et surtout en force de propositions.

Annexe G

La séquence conçue

G.1 Devoir maison n^{o} 5 traitant d'aire finie d'un domaine infini

DM n°5 : **A rendre le 17 janvier 2015**

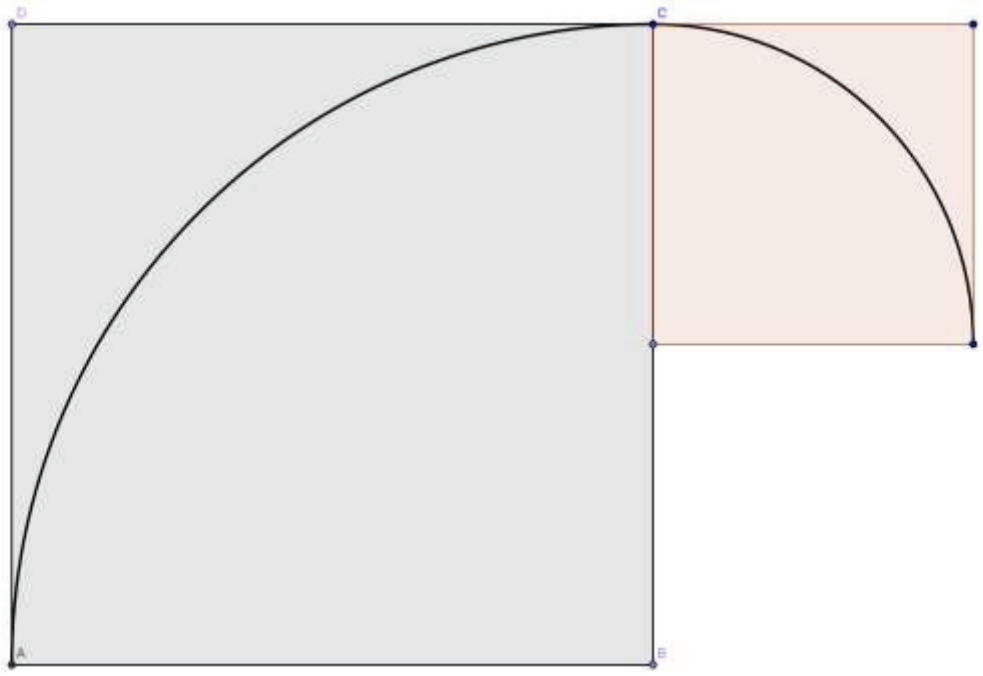
Exercice 1 : On considère la figure ci-dessous construite par itération. On construit un carré $A_0B_0C_0D_0$ de 10 cm de côté. A l'intérieur du carré, on trace le quart de cercle de centre B_0 et de rayon B_0C_0 (étape 0). On appelle B_1 le milieu du segment $[B_0C_0]$ et on trace, à l'extérieur du carré $A_0B_0C_0D_0$, le carré $C_0B_1C_1D_1$ puis le quart de cercle de centre B_1 et de rayon B_1C_1. L'itération du processus donne les carrés : $C_1B_2C_2D_2$, $C_2B_3C_3D_3$.....et la spirale $A_0C_0C_1C_2C_3$..... On note L_n la longueur totale de la spirale obtenue à l'étape n.

a) Compléter soigneusement la figure ci-dessous (étape 2 et 3).

b) Calculer la longueur L_3 de la spirale obtenue à l'étape 3.

c) On souhaite connaître la longueur de la spirale quand on itère le processus une infinité de fois. De façon intuitive, quelle est alors la longueur de la spirale ?

d) Exprimer en fonction de n, la longueur Ln de la spirale obtenue à l'étape n.
Déterminer la limite Ln quand n tend vers $+\infty$.



(Echelle non respectée)

Exercice 2 :

Partie A :

Pour tout $n \in \mathbb{N}^*$, on pose $u_n = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots \ldots \ldots \ldots \ldots + \frac{1}{n^2}$.

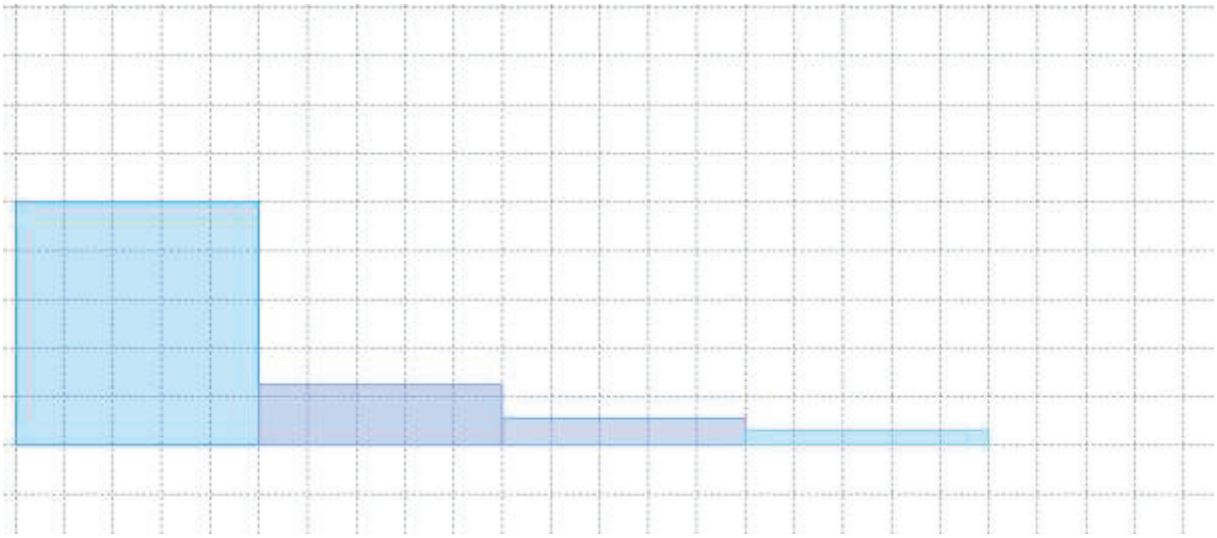
a) Etudier la monotonie de la suite (u_n).

b) Montrer que, pour tout $n \geq 2$, $\frac{1}{n^2} \leq \frac{1}{n-1} - \frac{1}{n}$.
En déduire que, pour tout n non nul, $u_n \leq 2 - \frac{1}{n}$.

c) Justifier que la suite (u_n) converge. On ne cherchera pas à déterminer cette limite.

Partie B :

On construit une suite de rectangles de base 1 et de hauteur 1, $\frac{1}{4}$ $\frac{1}{9}$ $\frac{1}{16}$ ……On note A_n la somme totale des aires des n premiers rectangles.



Que pouvez-vous dire de cette aire quand n tend vers $+\infty$?

Quel(s) commentaire(s) cela vous inspire-t-il ?

G.2 Devoir maison n° 6 sur l'histogramme et la loi uniforme

DM n°6. A rendre pour le mardi 3 mars **TS 2**

Attention ce devoir prépare un futur chapitre.

Exercice 1

Nous disposons des données statistiques suivantes :

Thème Population
Sous-thème Évolution et structure de la population
POP1A - Population par âge regroupé en 2011
Commune de Paris
© Insee

Population totale par âge regroupé

	Effectif
Moins de 3 ans	73946
3 à 5 ans	65679
6 à 10 ans	103810
11 à 17 ans	140886
18 à 24 ans	244026
25 à 39 ans	598565
40 à 54 ans	435966
55 à 64 ans	257600
65 à 79 ans	222118
80 ans ou plus	107378
Ensemble	2249975

Source : Insee, RP2011 exploitation principale.

1. Quelle est la population donnée ? Quelle est l'effectif total ?
2. Quel est le caractère étudié ?
3. Ce caractère statistique est-il qualitatif ou quantitatif ? Est-il discret ou continu ?
4. A l'aide du tableau ci-dessus, compléter le tableau de fréquences suivant (en arrondissant à 10^{-2} près) :

Tranche d'âge	$[0;4[$	$[4;6[$	$[6;11[$	$[11;18[$	$[18;25[$	$[25;40[$	$[40;55[$
Fréquence							

Tranche d'âge	$[55;65[$	$[65;80[$	$[80;105[$
Fréquence			

5. Nous voulons représenter graphiquement ces données à l'aide d'un histogramme de fréquences. Nous allons d'abord faire quelques commentaires et utiliser une analogie avec la géographie.

> En 2005, Monaco avait 32 543 habitants et le Japon, 127 417 244 (source : Institut national d'études démographiques). Bien sûr, les démographes diront que ces renseignements sont très largement insuffisants pour comparer la démographie des deux pays : il faut au minimum s'intéresser aux superficies de ces deux pays et calculer pour chacun d'entre eux la densité de population, c'est-à-dire le nombre d'habitants au kilomètre carré. Avec une superficie de 2,02 km² pour Monaco et de 378 000 km² pour le Japon, les densités sont respectivement $d_1 = 32\,543 / 2{,}02 = 16\,110{,}40$ h/km² pour Monaco et $d_2 = 127\,417\,244 / 378\,000 = 337$ h/km² pour le Japon. Autrement dit, alors que la population de Monaco est la moins importante en taille, sa densité de population est plus importante que celle du Japon.

Dans notre exemple, il n'est donc pas pensable de représenter de la même façon la fréquence 0,05 *de la tranche d'âge* $[6; 11[$ *et de la tranche d'âge* $[80; 105[$. *Ces deux classes ont la même fréquence mais l'amplitude de la première classe est inférieure à celle de la seconde. Par analogie avec la densité de population, on comparera les rapports de fréquence d'une classe d'âge sur l'amplitude de cette classe. On parlera de densité.*

Dans notre exemple, la densité de fréquence de la tranche d'âge $[6; 11[$ *est* $\frac{0{,}05}{5} = 0{,}01$ *et la densité de la tranche d'âge* $[80; 105[$ *est* $\frac{0{,}05}{25} = 0{,}002$. *Comme perçue, la densité de fréquence de la première tranche d'âge est plus grande que la densité de fréquence de la seconde tranche d'âge. La représentation graphique doit rendre visible ce phénomène.*

Définition

Un ***histogramme de fréquences*** *est un diagramme composé de rectangles collés dont les* ***aires*** *sont proportionnelles aux fréquences et dont les bases sont déterminées par les intervalles des classes.*

Dans le cas d'une variable quantitative continue, on définit la ***densité*** *de fréquence* d_i *d'une classe de fréquence* f_i *et d'amplitude* a_i *par :* $d_i = f_i / a_i$.

a) Justifier que, dans un histogramme, les hauteurs des rectangles sont proportionnelles à la densité. *Ainsi une classe de densité plus forte est représentée par un rectangle plus haut.*

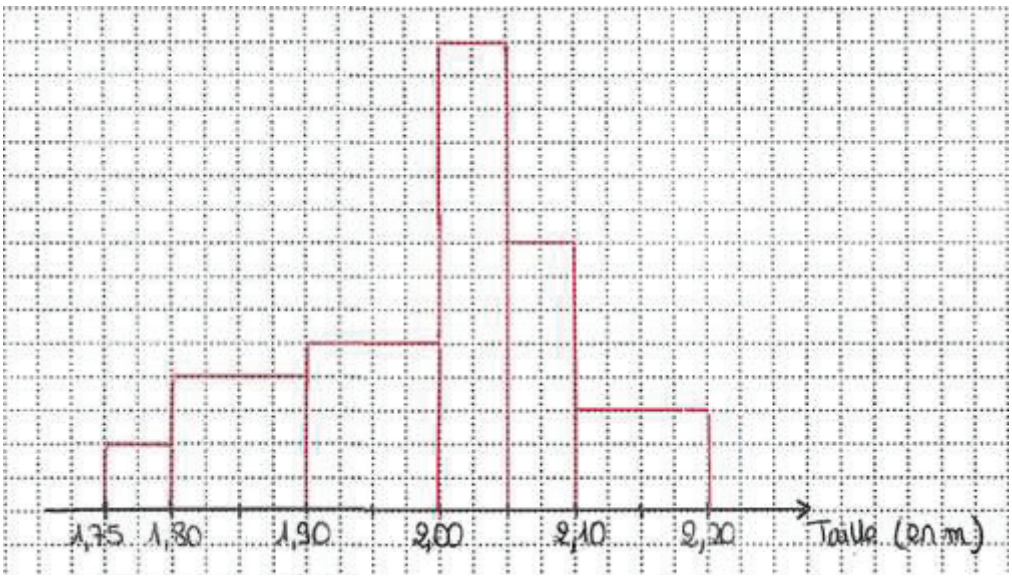
b) Compléter l'histogramme de fréquences de la répartition de la population parisienne en 2011 par âge (sur la feuille ci-après). On donne l'aire d'un carreau correspondant à une fréquence de 0,001.
Que représente l'axe des ordonnées ? Graduer l'axe.

6. Dans quelle tranche d'âge la densité est-elle la plus importante ? Cela vous semble-t-il raisonnable ?

7. Dans un histogramme, si un rectangle est plus haut qu'un autre, peut-on affirmer que la fréquence associée est la plus grande ?

Exercice 2

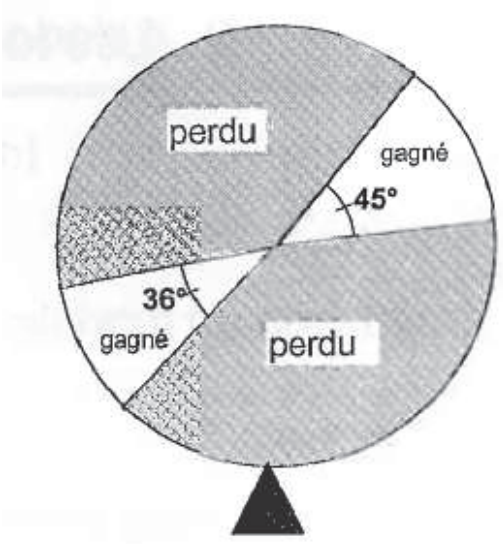
Voici l'histogramme des fréquences des tailles des joueurs de l'équipe de France basket présélectionnés pour la Coupe du Monde 2014.



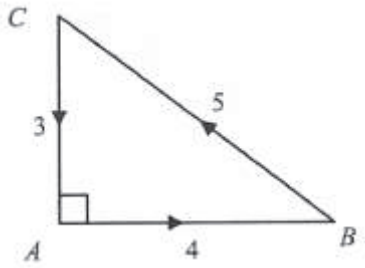
1. A quelle fréquence correspond la somme des aires des rectangles ?
   On dira que l'aire des rectangles est égale à 1 **unité d'aire**.
2. Construire le tableau des fréquences correspondant à cet histogramme. Arrondir les fréquences à 10^{-3} près.
3. Sachant que cette présélection comptait 24 joueurs.
   a) Calculer combien de joueurs mesuraient plus de 2,10 m.
   b) Combien de joueurs mesuraient entre 1,80 m et 2 m ?
   c) Combien de joueurs mesuraient entre 1,85 m et 1,95 m ? Quelle hypothèse avez-vous faite pour répondre à la question ?

Exercice 3

A. On lance une roue de la fortune traditionnelle. Quelle est la probabilité de gagner ?



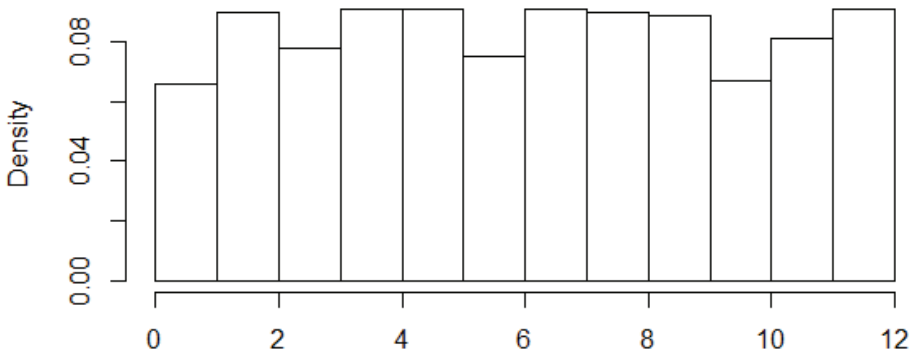
B. Une puce électronique se déplace à vitesse constante sur le « pourtour » d'un triangle, toujours dans le même sens (indiqué par la flèche). On suppose que la puce démarre au point C.



Une panne se produit subitement et la puce s'arrête instantanément.

1. Quelle est la probabilité p_1 que la puce s'arrête sur l'hypoténuse ?
2. Quelle est la probabilité p_2 que la puce s'arrête au sommet de l'angle droit ? Commenter la valeur de p_2 obtenue.
3. Quelle est la probabilité p_3 que la puce s'arrête au milieu du segment [AB] ? Commenter la valeur de p_3 obtenue.
4. On peut remarquer que cette situation se ramène à l'expérience aléatoire qui correspond à choisir un nombre réel au hasard dans l'intervalle $[0\,;12]$.
   A l'aide d'un tableur, on a simulé le tirage de 1000 nombres réels dans l'intervalle $[0\,;12]$.

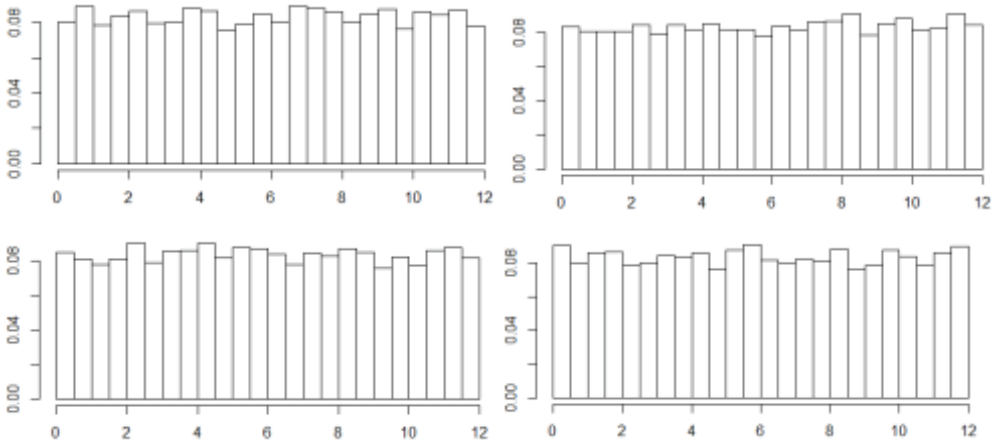
   Voici l'histogramme de fréquences :



a) Quel est l'amplitude de chaque classe ?

b) A quelle fréquence les puces s'arrêtent-elles sur l'hypoténuse ?

5. On a ensuite simulé, plusieurs fois, le tirage de 10 000 nombres réels dans l'intervalle $[0\,;12]$. Voici plusieurs histogrammes correspondants :



a) Les histogrammes ci-dessus sont différents. Justifier pourquoi.
b) Malgré ces différences ne semble-t-il pas y avoir une « courbe de tendance », c'est-à-dire une courbe qui « lisse » l'histogramme (la même pour les quatre) ? Tracer la sur chacun des histogrammes ci-dessus.
c) Que doit valoir l'aire sous cette courbe ?
d) En déduire l'équation de cette courbe.
e) A l'aide de cette courbe, retrouver la probabilité p_1 que la puce s'arrête sur l'hypoténuse.
f) De la même façon, retrouver la probabilité p_2 que la puce s'arrête au sommet de l'angle droit.

G.3 Problème de la rencontre

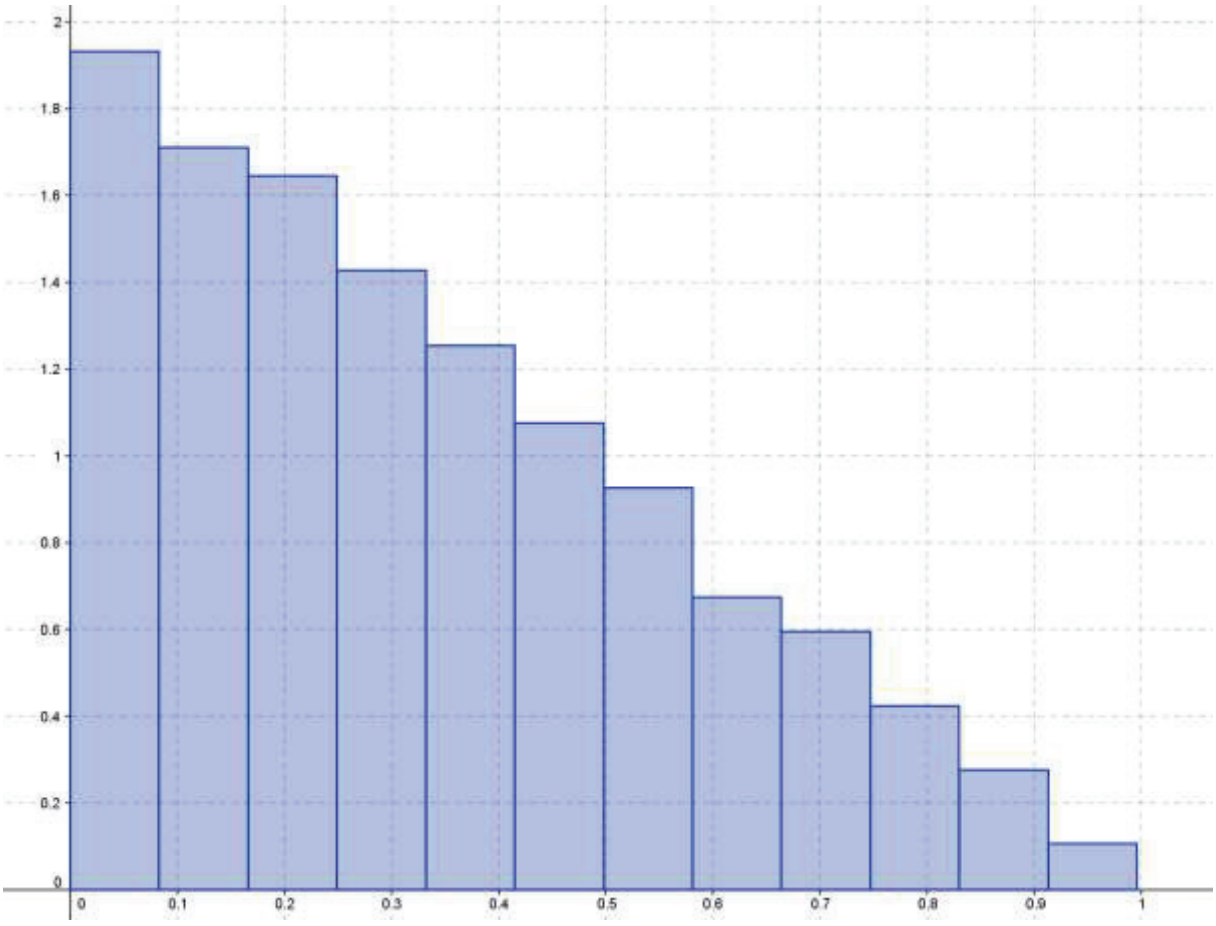
Enoncé

Problème 1 :

Karine et Olivier décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Ils peuvent arriver à tout moment entre 7h et 8h. Que peut-on dire du temps d'attente du premier arrivé ?

Document distribué une fois les histogrammes construits sur *GeoGebra*

Histogramme d'un échantillon de taille 10 000, classe d'amplitude 5 minutes.



Exercice distribué une fois la fonction de densité déterminée

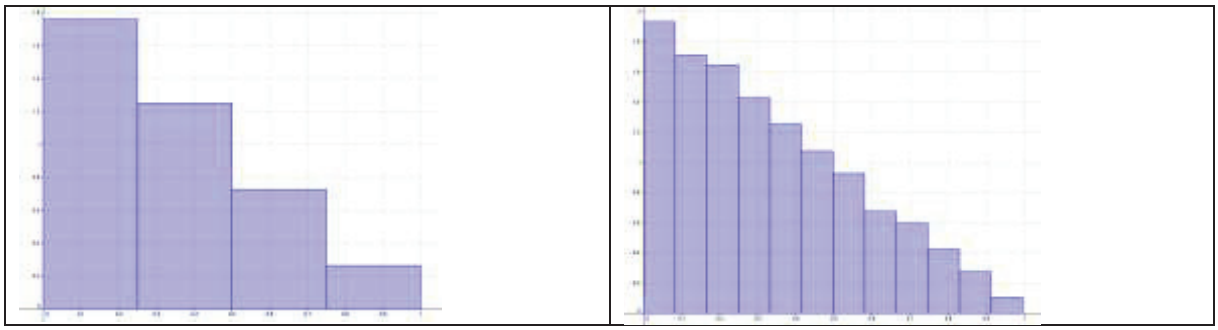
Exercice 1 : En utilisant le modèle mis en place, répondez aux questions suivantes.

a) Quelle est la probabilité que le premier arrivé attende exactement 40 minutes ?
b) Quelle est la probabilité que le premier arrivé attende plus de 40 minutes ?
c) Quelle est la probabilité que le premier arrivé attende moins de 40 minutes ?
d) Quelle est la probabilité que le premier arrivé attende entre 20 et 40 minutes ?
e) Quelle est la probabilité que Karine et Olivier arrivent en même temps ?
f) Finalement, ils ont décidé de ne pas attendre plus d'un quart d'heure. Quelle est la probabilité qu'ils se retrouvent ?

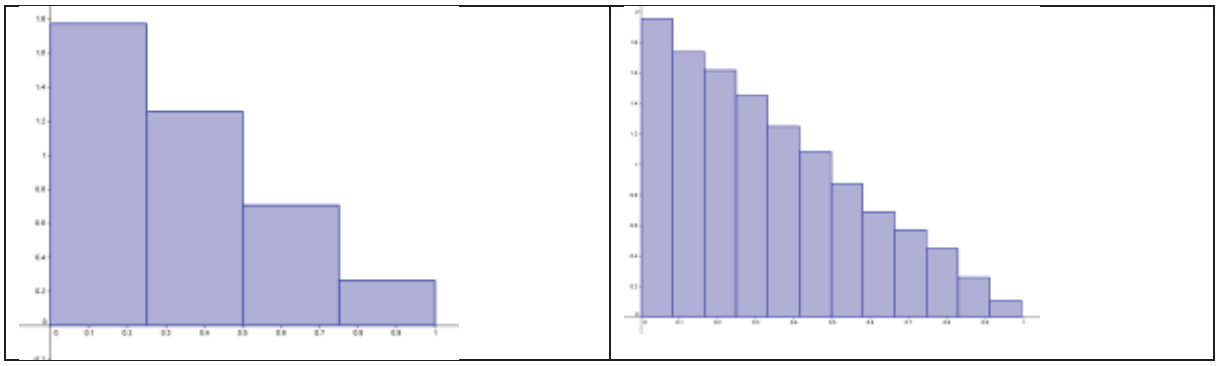
Document distribué lors de la séance 2

Voici trois échantillons du temps d'attente du premier arrivé simulés sur ordinateur. Nous avons tracé les histogrammes pour des classes d'amplitude 15 puis 5minutes.

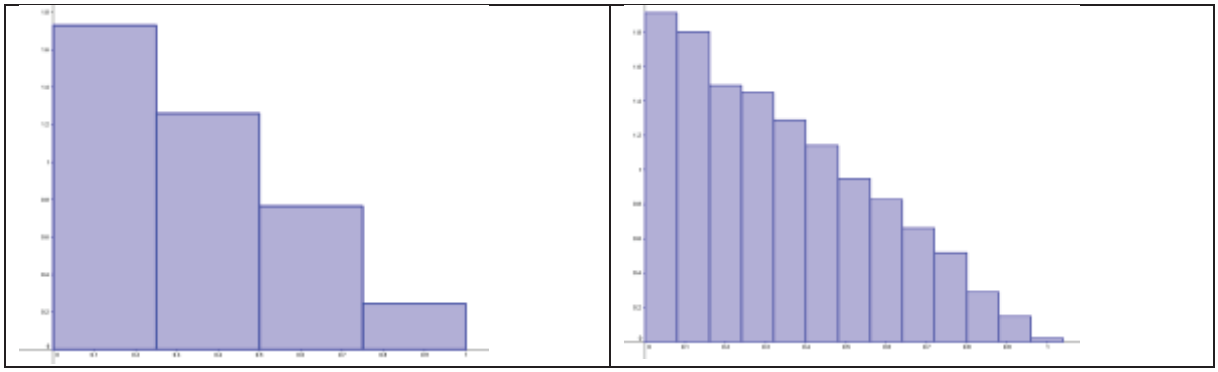
Echantillon 1 de taille 10 000:



Echantillon 2 de taille 10 000:



Echantillon 3 de taille 10 000 :



G.4 Problème du volcan Aso

Problème 2 :

Le volcan Aso, situé sur l'île de Kyushu au Japon, est l'un des plus actifs au monde. On possède les statistiques de ses éruptions, régulièrement tenues depuis le XIIIè siècle. Nous avons, dans le document tableur, les années d'éruptions jusqu'au XIXè siècle (à partir du XXè siècle, les données, d'une autre nature, ne sont pas compatibles(*)).On s'intéresse au temps écoulé, en années, entre deux éruptions successives.

Années des 73 éruptions du volcan Aso du XIIIe au XIXe siècle

Donnée n°	Année	Temps d'attente		Donnée n°	Année	Temps d'attente
1	1229			34	1562	4
2	1239	10		35	1563	1
3	1240	1		36	1564	1
4	1265	25		37	1576	12
5	1269	4		38	1582	6
6	1270	1		39	1583	1
7	1272	2		40	1584	1
8	1273	1		41	1587	3
9	1274	1		42	1598	11
10	1281	7		43	1611	13
11	1286	5		44	1612	1
12	1305	19		45	1613	1
13	1324	19		46	1620	7
14	1331	7		47	1631	11
15	1335	4		48	1637	6
16	1340	5		49	1649	12
17	1346	6		50	1668	19
18	1369	23		51	1675	7
19	1375	6		52	1683	8
20	1376	1		53	1691	8
21	1377	1		54	1708	17
22	1387	10		55	1709	1
23	1388	1		56	1765	56
24	1434	46		57	1772	7
25	1438	4		58	1780	8
26	1473	35		59	1804	24
27	1485	12		60	1806	2
28	1505	20		61	1814	8
29	1506	1		62	1815	1
30	1522	16		63	1826	11
31	1533	11		64	1827	1
32	1542	9		65	1828	1
33	1558	16		66	1829	1
				67	1830	1
				68	1854	24
				69	1872	18
				70	1874	2
				71	1884	10
				72	1894	10
				73	1897	3

Le volcan Aso est actuellement en éruption.

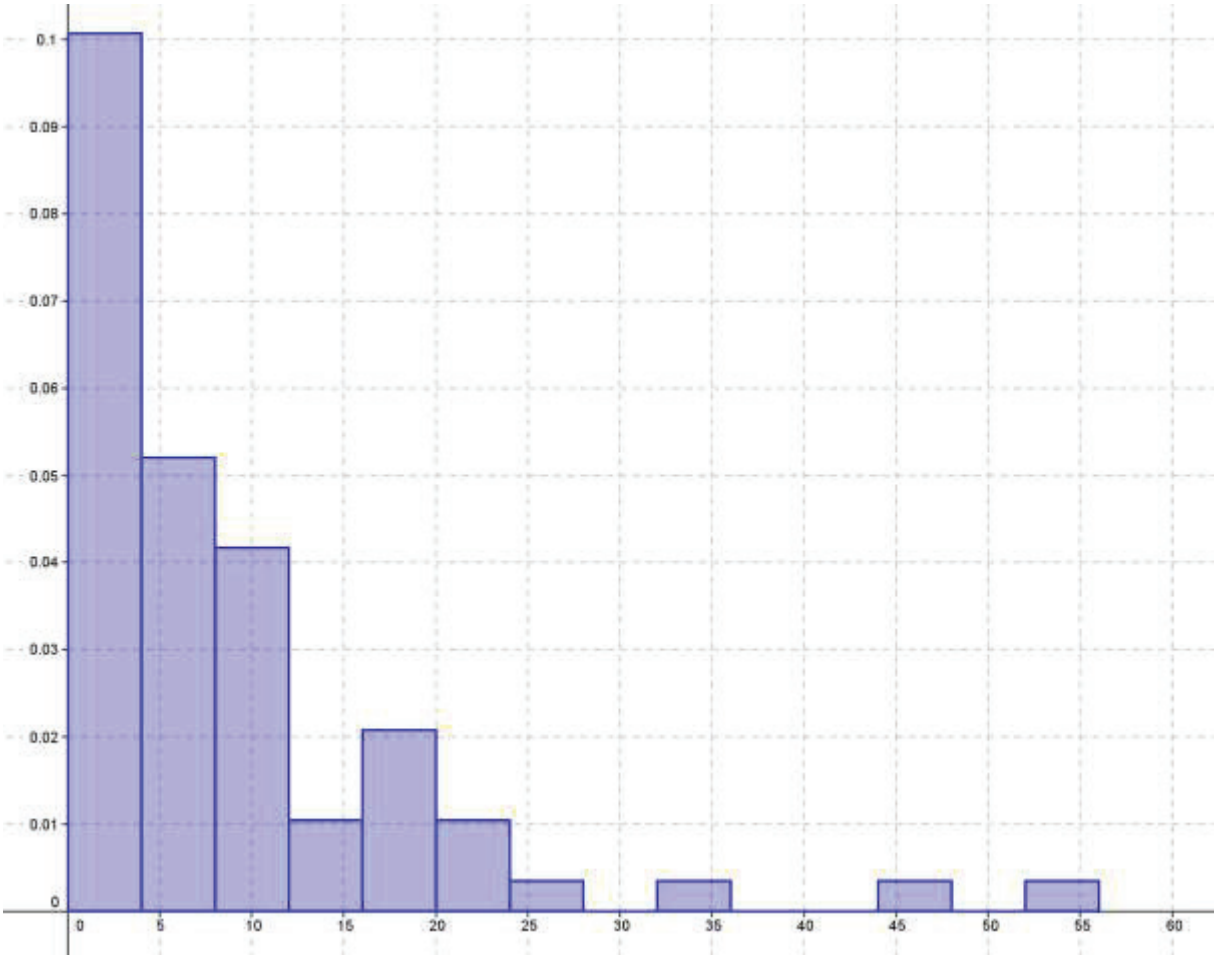
Comment évaluer la probabilité que la prochaine éruption:

a) ait lieu dans les 5 ans ?

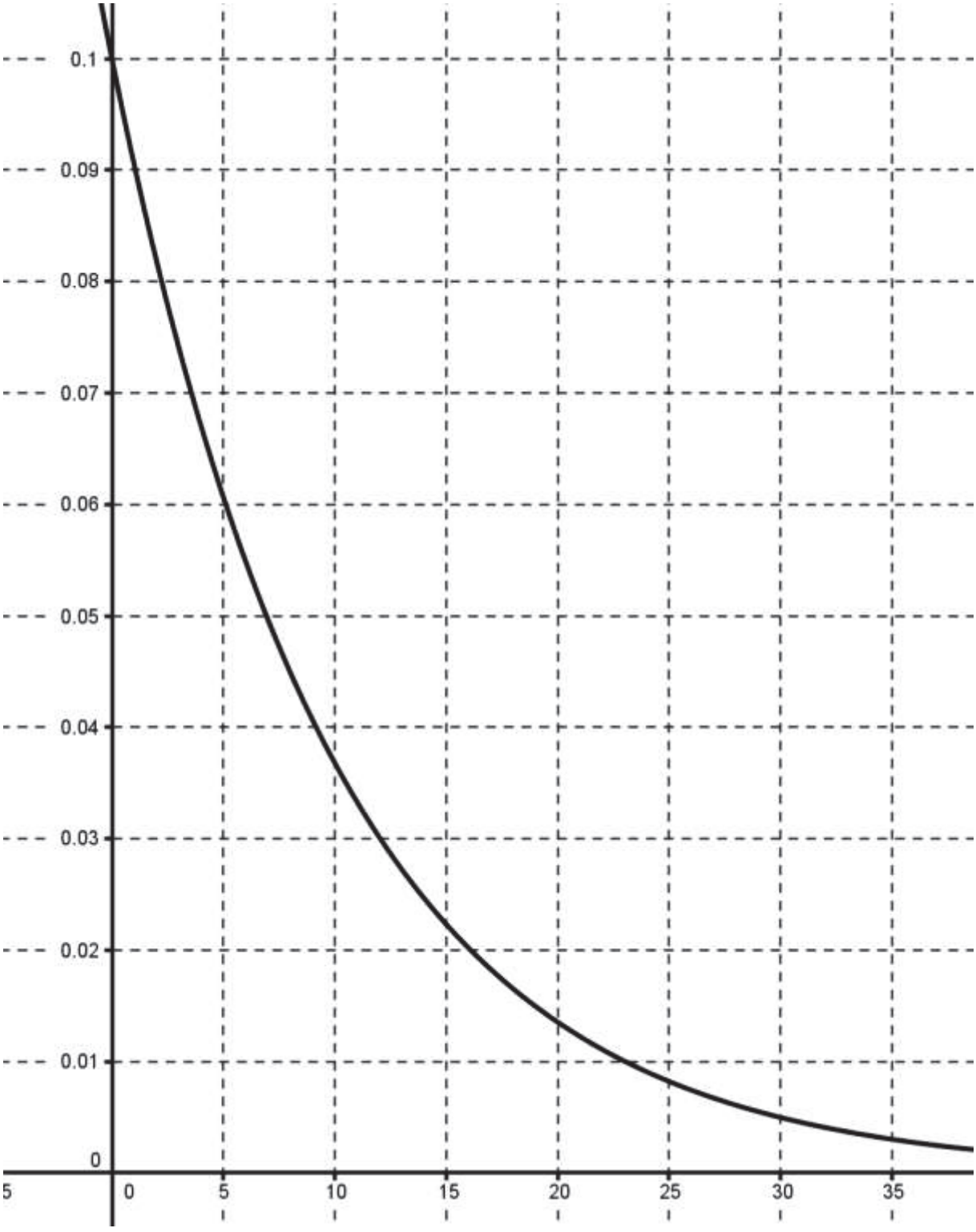
b) Au cours de l'année 2030 ?

Document distribué une fois l'histogramme construit sur *GeoGebra*

Histogramme, amplitude des classes 4 ans.

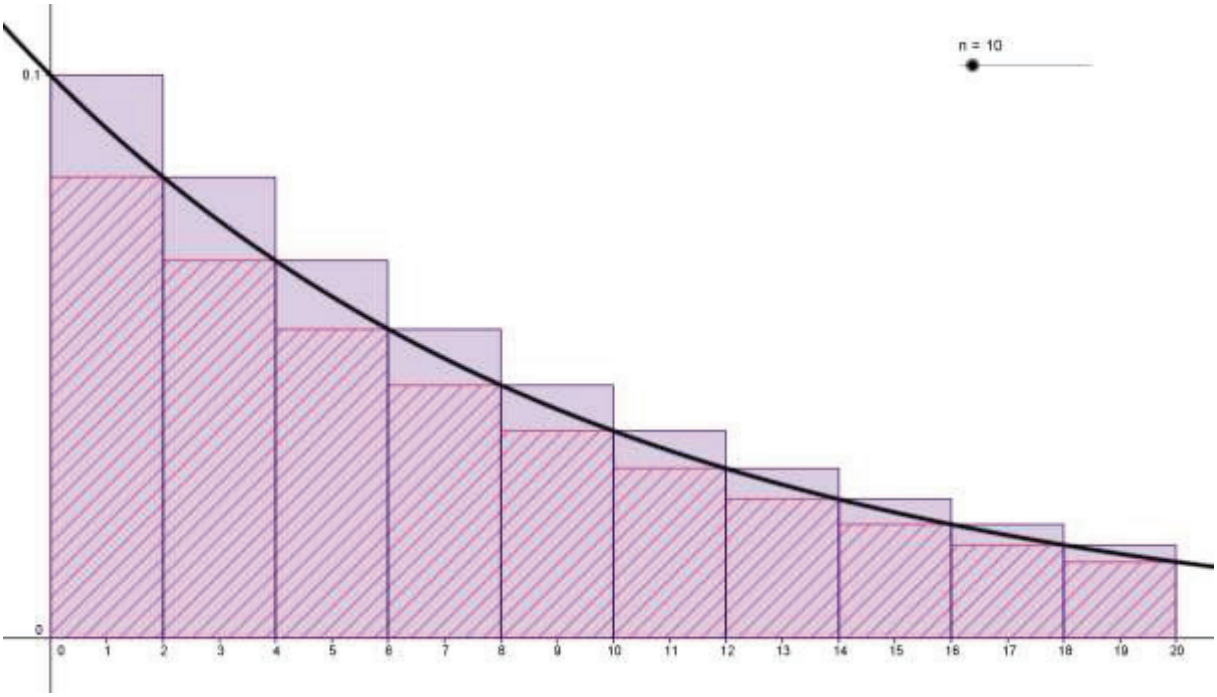


Document distribué pour la recherche de méthode pour calculer l'aire sous la courbe sur $[0\ ;\ 20]$



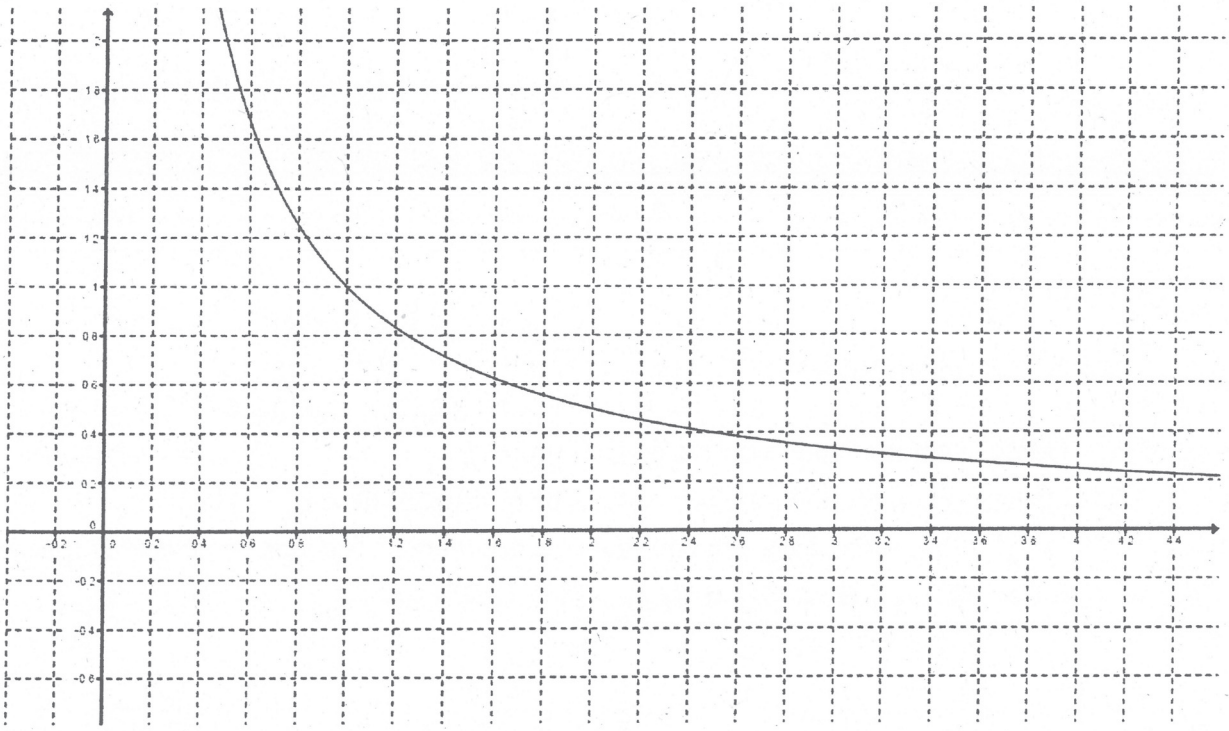
Document à distribuer pour illustrer la méthode des rectangles

Méthode des rectangles : $f : x \longmapsto 0{,}1 \times e^{-0{,}1x}$



G.5 Les exercices

Exercice 1 :
Utilisez la méthode des rectangles sur l'intervalle [1,2] avec une subdivision de largeur 0,2 pour encadrer l'aire sous la courbe représentative de la fonction inverse sur [1; 2].



Exercice 2 : En utilisant un calcul d'aire, calculer les intégrales suivantes (vous commencerez par faire un dessin) :

a) $\int_2^3 (x-2)\,dx$

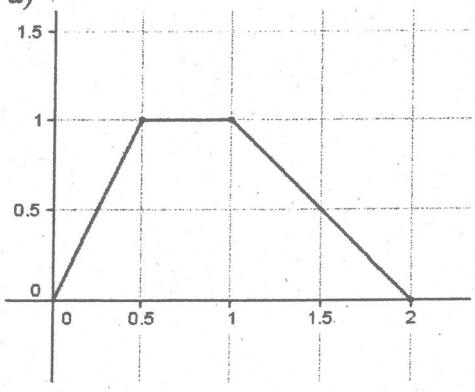
b) $\int_{-7}^{-4} 5\,dx$

c) $\int_{-5}^{3} |x|\,dx$

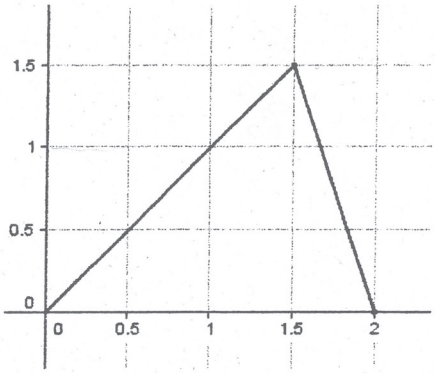
d) $\int_{-3}^{8} |x+2|\,dx$

Exercice 3 : Parmi les fonctions représentées graphiquement ci-dessous, déterminer celles qui définissent une densité de probabilité sur [0 ; 2] (pour les deux premières) ou sur [-1,5 ; 1,5] (pour les deux dernières).

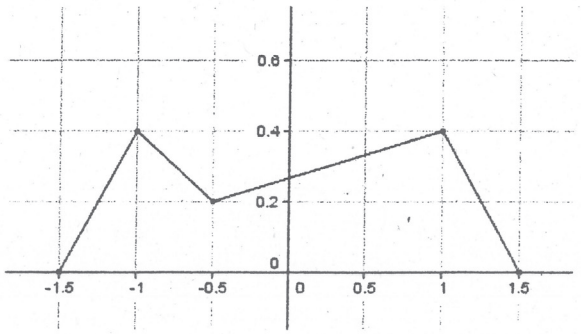
a)



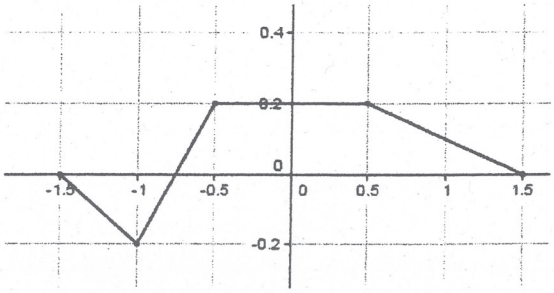
b)



c)



d)



Exercice 4 :

Soit f la fonction définie sur $[0;4]$ par $f(t) = \begin{cases} 0{,}2 \text{ sur } [0;2[\\ 0{,}2t - 0{,}2 \text{ sur } [2;3[. \\ -0{,}2t + 1 \text{ sur } [3;4] \end{cases}$

1) Tracer la courbe représentative de la fonction f.
2) Justifier que f est une fonction de densité de probabilité sur $[0\,;4]$.
3) Soit X une variable aléatoire continue qui admet pour fonction de densité la fonction f. Calculer $P(X \in [1\,;3])$ et $P_{X\in[1\,;4]}(X \in [2{,}5;4])$.

Exercice 5 : On considère la fonction carrée sur [0; 1].On subdivise l'intervalle $[0;1]$ en n intervalles de même longueur $\frac{1}{n}$. Sur chaque subdivision, on construit un « petit » rectangle r (hachuré sur le dessin) et un « grand » rectangle R. L'aire sous la parabole entre 0 et 1 est comprise entre la somme s_n des aires des « petits » rectangles et la somme S_n des aires des « grands » rectangles.

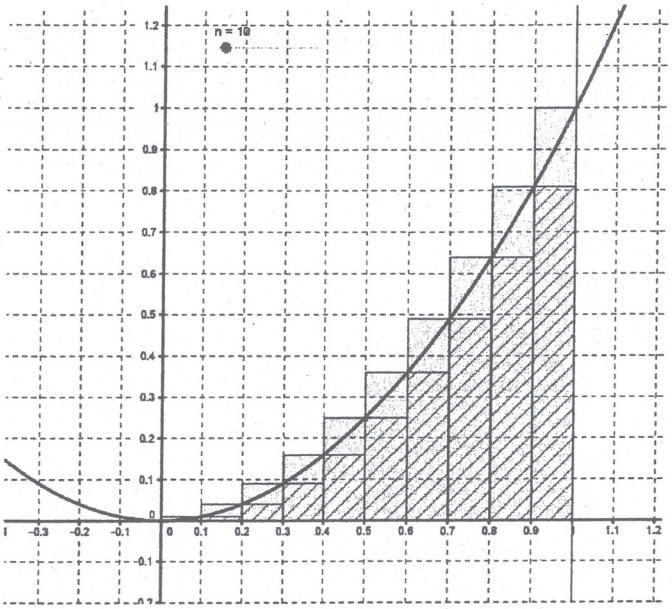


Figure dans le cas n = 10

1. On se place dans le cas où n = 10.
   a) Quelle est l'aire d'un « petit » rectangle r_k de base $\left[\frac{k}{10};\frac{k+1}{10}\right]$ (k *variant de* 0 à 9)? d'un « grand » rectangle R_k ?
   b) Calculer alors s_{10} et S_{10}. En déduire un encadrement de A.
2. Cas général : On subdivise l'intervalle $[0;1]$ en n intervalles de longueur $\frac{1}{n}$.
   a) Donner l'expression, en fonction de n, des sommes s_n et S_n.
   b) Déterminer les limites des suites (s_n) et (S_n). En déduire la valeur de A.
3. On considère l'algorithme ci-dessous :

Variables : a, b, h, U et V sont des variables réelles n est une variable entière Entrées : Lire a Lire b Lire n Traitement : Affecter 0 à U Affecter 0 à V Affecter $\frac{b-a}{n}$ à h	Pour k allant de ..... à..... faire Affecter $U + h \times (a + k \times h)^2$ à U Affecter $V + h \times (a + (k+1) \times h)^2$ à V Fin de pour Sortie : Afficher U Afficher V	a) Compléter cet algorithme. b) Qu'affiche l'algorithme si on entre $a = 0,\ b = 1\ et\ n = 5$? c) Programmer cet algorithme sur votre calculatrice. Qu'affiche le programme pour $a = 0,\ b = 1\ et\ n = 100$?

Exercice 6 :

Tous les samedis matins, un élève dont je tairai le nom, arrive au hasard entre 7h15 et 7h35 à l'arrêt de bus. A partir de 7h00, les bus passent exactement[1] toutes les 10 minutes. On appelle X la variable aléatoire égale à l'heure d'arrivée de cet élève à l'arrêt de bus.

1. Quelle est la loi de X ?
2. Quelle est la probabilité que cet élève attende le bus moins de 5 minutes ?
3. S'il rate le bus de 7h30, cet élève est en retard à son cours de mathématiques. Quelle est la probabilité que cet élève arrive en retard en mathématiques[2] ?

[1] Au pays des mathématiques, les bus passent à l'heure !

[2] Même au pays des mathématiques, les élèves sont en retard..... rien n'est parfait !

Exercice 7 : Soit X une variable aléatoire continue de fonction de densité de probabilité f sur l'intervalle I.
Démontrer les propriétés suivantes :

(i) Pour tout intervalle J inclus dans I, $P(X \in J) \geq 0$.

(ii) Pour tout $a \in I, P(X = a) = 0$.

(iii) Pour tout $a \in I, P(X \leq a) = P(X < a)$.

(iv) Pour tous $a, b, c \in I$, tels que $a \leq b \leq c$,

$$P(a \leq X \leq c) = P(a \leq X \leq b) + P(b < X \leq c).$$

(v) Pour tout $a \in I, P(X > a) = 1 - P(X \leq a)$.

(vi) Pour tous $a, b \in I$, tels que $a < b, P(a \leq X < b) = P(X < b) - P(X < a)$.

Exercice 8 : On considère la fonction $f(x) = e^x$ définie sur $[0;1]$ et C sa courbe dans un repère orthonormé.
On se propose de déterminer un encadrement de l'aire sous la courbe représentative de la fonction f sur [0 ;1] en utilisant la méthode des rectangles.
On subdivise l'intervalle $[0;1]$ en n intervalles de même longueur $\frac{1}{n}$. Sur chaque subdivision, on construit un « petit » rectangle r (hachuré sur le dessin) et un « grand » rectangle R.
L'aire que l'on cherche à calculer, est comprise entre la somme s_n des aires des « petits » rectangles et la somme S_n des aires des « grands » rectangles.

a) Montrer que, pour tout $n \geq 1$: $s_n = \frac{1}{n} \times \frac{e-1}{e^{\frac{1}{n}}-1}$ et $S_n = s_n \times e^{\frac{1}{n}}$

b) Déterminer les limites des suites (s_n) et (S_n). En déduire la valeur de A.

Ex 10: Déterminer une primitive de f sur l'intervalle I.

a) $f(x) = -5x + \pi$ $\quad I = \mathbb{R}$

b) $f(x) = \dfrac{3}{x^2}$ $\quad I = \mathbb{R}_+^*$

c) $f(x) = \dfrac{-5}{x}$ $\quad I =]0;+\infty[$

d) $f(x) = \dfrac{x}{\sqrt{x^2+1}}$ $\quad I = \mathbb{R}$

e) $f(x) = x e^{x^2}$ $\quad I = \mathbb{R}$

f) $f(x) = \dfrac{x}{x^2+4}$ $\quad I = \mathbb{R}$

g) $f(x) = e^{-x}$ $\quad I = \mathbb{R}$

h) $f(x) = \cos(3x+1)$ $\quad I = \mathbb{R}$

i) $f(x) = \dfrac{e^x}{e^x+2}$ $\quad I = \mathbb{R}$

j) $f(x) = \dfrac{4}{(x+1)^2}$ $\quad I = \mathbb{R}$

k) $f(x) = e^x(e^x+3)^2$ $\quad I = \mathbb{R}$

Ex 11: Déterminer les réels a et b pour que F soit une primitive de f sur I.

a) $F(x) = (ax+b)e^x$; $f(x) = 3xe^x$; $I = \mathbb{R}$

b) $F(x) = (a\ln x + b)x^2$; $f(x) = x\ln x$; $I =]0;+\infty[$

Ex 12: Soit $f : x \longmapsto \dfrac{5x^2+21x+22}{(x-1)(x+3)^2}$ définie sur $]1;+\infty[$.

1. Montrer qu'il existe 3 réels a, b et c tels que $\forall x \in]1;+\infty[$, $f(x) = \dfrac{a}{x-1} + \dfrac{b}{x+3} + \dfrac{c}{(x+3)^2}$.
2. En déduire une primitive de f sur $]1;+\infty[$ puis, la primitive de f sur $]1;+\infty[$ qui s'annule en 2.

Exercice 13 : Calculer les intégrales suivantes.

$$I_1 = \int_0^3 (x-4)\,dx \qquad I_6 = \int_1^3 \frac{du}{u+1}$$

$$I_2 = \int_{-2}^{-1} \left(x^2 - \frac{1}{x^2}\right) dx \qquad I_7 = \int_{-3}^{3} x e^2\,dx$$

$$I_3 = \int_0^2 (2x+1)(x^2+x+1)\,dx \qquad I_8 = \int_1^2 e^{2x}\,dt\,dx$$

$$I_4 = \int_1^2 \frac{dt}{t^5} \qquad I_9 = \int_0^{\pi/2} \sin 2t\,dt$$

$$I_5 = \int_{-\pi/6}^{3\pi/2} \cos t\,dt \qquad I_{10} = \int_{-7}^{+9} dt\,.$$

$$I_{11} = \int_1^2 \frac{dx}{3x+2} \qquad I_{12} = \int_{-\pi/6}^{\pi/4} \sin^2 t\,dt$$

$$I_{13} = \int_0^1 t e^{t^2-1}\,dt \qquad I_{14} = \int_e^{e^2} \frac{1}{t \ln t}\,dt$$

Exercice 14

a) Déterminer les réels a, b, c tels que :
pour tout $t \in \mathbb{R} \setminus \{\frac{1}{2}\}$, $\dfrac{t^2-1}{2t-1} = at + b + \dfrac{c}{2t-1}$

b) Calculer $\displaystyle\int_{-1/2}^{0} \frac{t^2-1}{2t-1}\,dt$

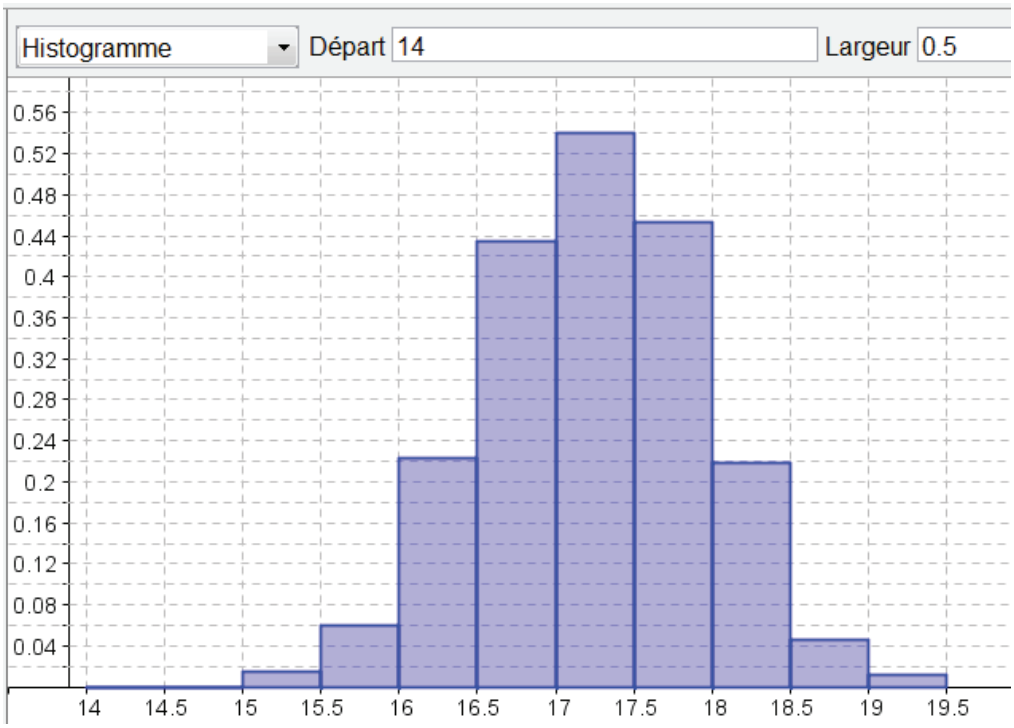
G.6 Problème des tailles

DM Mathématiques – Terminale S
Première rencontre avec la loi normale

Problème

Dans le cadre de ses recherches sur l'hérédité, le statisticien britannique Karl Pearson (1857-1936), a récolté des données et a établi un fichier des mesures de la taille du père et du fils (adulte) de 1078 familles à la fin du XIXè siècle. Ici, nous ne nous intéresserons qu'aux données sur la taille des pères, en **décimètres**.

Voici les données représentées sous forme d'histogramme de fréquences.



Etude de l'échantillon

1. Quelle est la fréquence des pères de famille mesurant entre 1,80 m et 1,90 m ? Justifier.

Modélisation

A l'aide de cet échantillon, on cherche à modéliser la taille des pères de famille à la fin du XIXè siècle. On appelle X la variable aléatoire « taille des pères de famille à la fin du XIXè siècle ».
En annexe, nous vous proposons trois fonctions et leur représentation graphique.

Remarque : Ces trois courbes sont symétriques par rapport à la droite d'équation $x = 17{,}2$.

2. Rappeler quelles sont les propriétés que doivent vérifier une fonction de densité de probabilité.

3. Vérifier que f_A (fonction donnée en annexe) est une fonction de densité de probabilité. On admettra que les fonctions f_B et f_C sont des fonctions de densité.

4. Quelle fonction de densité choisir, parmi les trois proposées, pour modéliser la variable aléatoire X (la taille des pères de famille en mètres) ? Justifier votre réponse. On appellera f cette fonction, à partir de maintenant.

5. A l'aide de la méthode des rectangles (avec 4 subdivisions), encadrer la probabilité (à 10^{-2} près) qu'un père de famille de la fin du XIXè siècle mesure entre 1,60 m et 1,70 m. Détailler votre méthode. Vous pouvez ajouter des commentaires et figures sur le graphique de la fonction choisie.

6. Voici deux algorithmes :

Algorithme 1 :

Variables : k est un nombre entier U est un nombre réel Initialisation : Affecter à U la valeur 0 Traitement : Pour k variant de 0 à 3 | Affecter à U la valeur U+ $\frac{1}{4}\ f(\frac{k}{4})$ Fin Pour Sortie : Afficher U

Algorithme 2

Variables : k est un nombre entier V est un nombre réel Initialisation : Affecter à V la valeur 0 Traitement : Pour k variant de 0 à 3 | Affecter à V la valeur V+ $\frac{1}{4}\ f(\frac{k+1}{4})$ Fin Pour Sortie : Afficher V

Que représente chacun des deux algorithmes ? Justifier.
Donner une valeur approchée à 10^{-2} près du résultat affiché par l'algorithme 1, puis par l'algorithme 2.

7. En justifiant, déterminer :

a) $P(X = 1{,}72)$;
b) La probabilité qu'un père ait une taille inférieure ou égale à 1,72 m ;
c) La probabilité qu'un père ait une taille strictement inférieure à 1,72 m ;
d) $P(X \geq 1{,}72)$;
e) $P(X > 1{,}72)$.

8. On note $I(b) = \int_{17,2}^{17,2+b} f(x)dx$ où b est un nombre réel positif ou nul.

On donne dans le tableau suivant, pour quelques valeurs de b, des valeurs approchées de $I(b)$ obtenues à l'aide de la calculatrice.

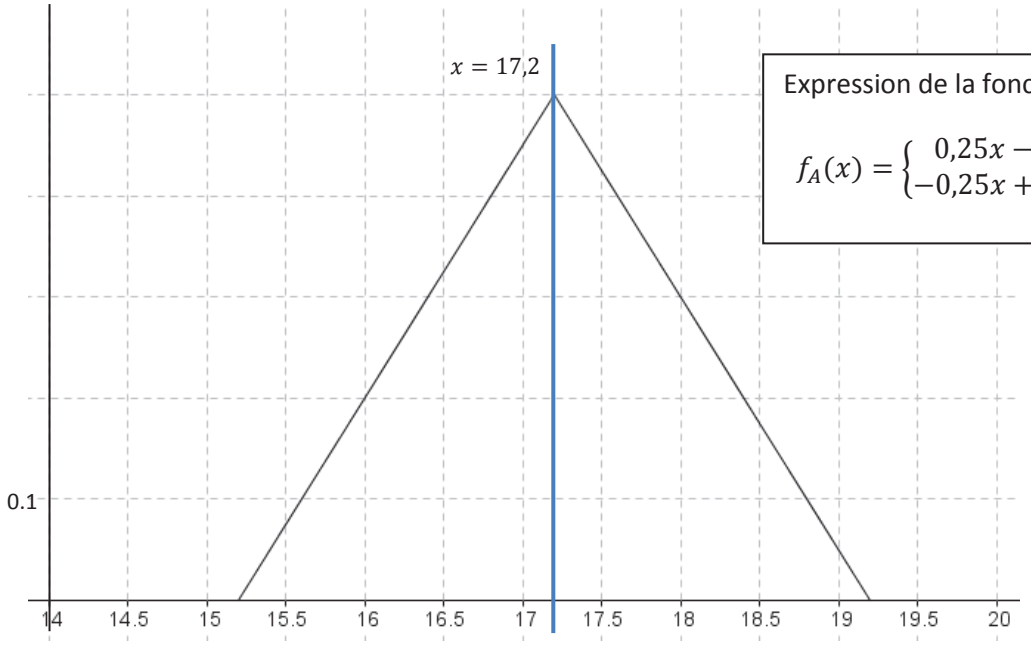
b	0,5	1	1,5	2	2,5
$I(b)$	0,2625	0,4234	0,4839	0,4979	0,4998

En déduire des valeurs approchées de :

a) La probabilité qu'un père mesure entre 1,67 m et 1,77 m.
b) La probabilité qu'un père mesure entre 1,72 m et 1,82 m.
c) La probabilité qu'un père mesure entre 1,47 m et 1,62 m.

Annexe

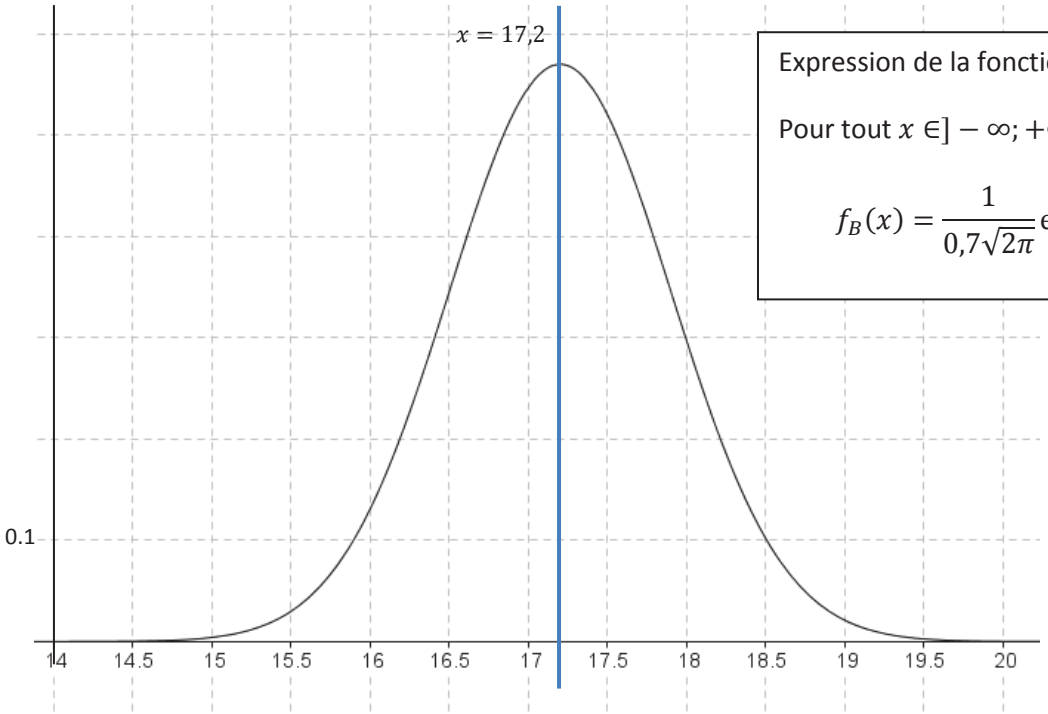
A.



Expression de la fonction :

$$f_A(x) = \begin{cases} 0{,}25x - 3{,}8 & si\ x \in [15{,}2; 17{,}2] \\ -0{,}25x + 4{,}8 & si\ x \in [17{,}2; 19{,}2] \end{cases}$$

B.



Expression de la fonction :

Pour tout $x \in]-\infty; +\infty[$,

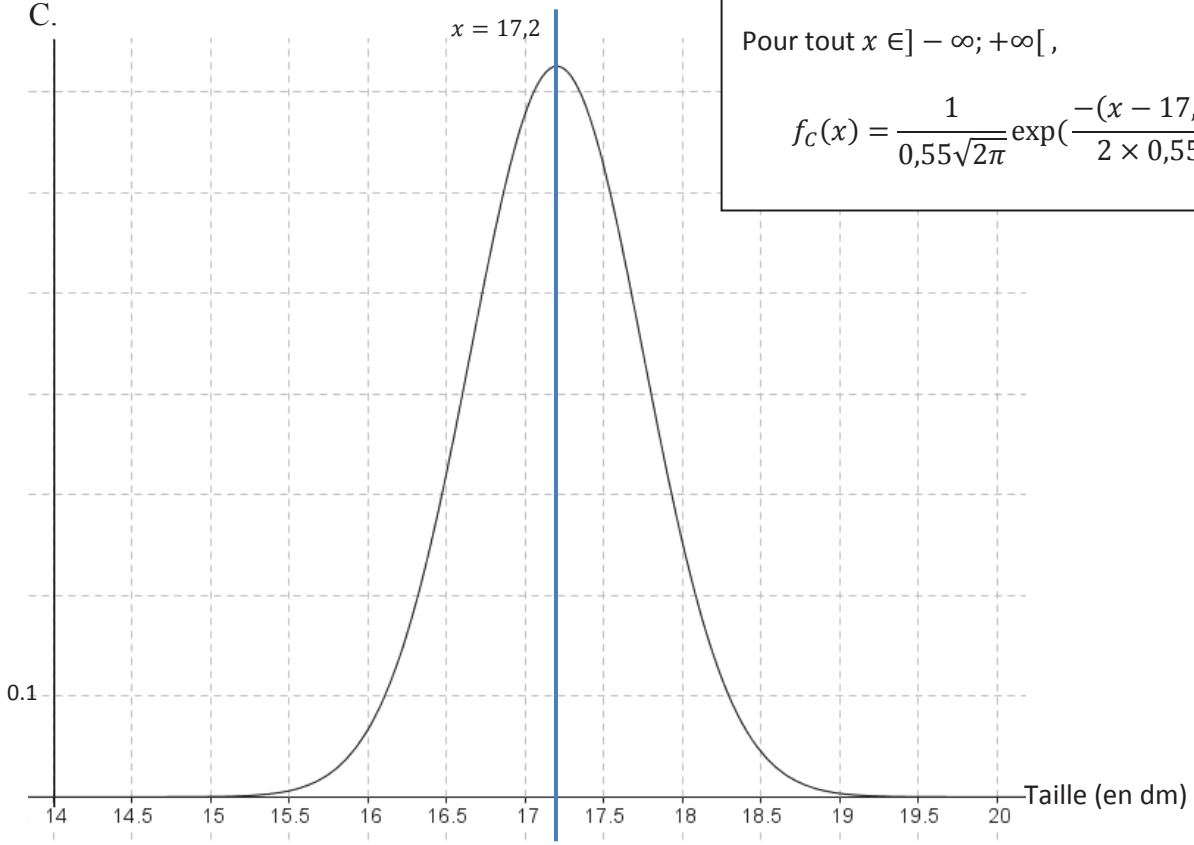
$$f_B(x) = \frac{1}{0{,}7\sqrt{2\pi}} \exp\left(\frac{-(x-17{,}2)^2}{2 \times 0{,}7^2}\right)$$

Expression de la fonction :

Pour tout $x \in]-\infty;+\infty[$,

$$f_C(x) = \frac{1}{0{,}55\sqrt{2\pi}} \exp\left(\frac{-(x-17{,}2)^2}{2 \times 0{,}55^2}\right)$$

C.



Annexe H

Planning de la séquence

Séance et date	**Contenu de la séance**	**Travail à faire pour la séance suivante**
Séance 1 (2h) Samedi 14 mars	Problème 1	Rédiger une synthèse sur le déroulement de la séance. + Terminer l'exercice sur le problème 1 (calcul de probabilités).
Séance 2 (1h) Mardi 17 mars	Synthèse du problème 1. Correction de l'exercice sur le problème 1. (Le reste de l'heure a été consacré à un autre chapitre).	
Séance 3 (1h) Mercr. 18 mars	Problème 2 (jusqu'à la construction de l'histogramme et la recherche d'une "courbe de tendance").	A partir du fichier *GeoGebra* avec l'histogramme (envoyé par l'enseignante), déterminer l'expression d'une fonction de densité possible.
Séance 4 (2h) Vendr. 20 mars	Fin du problème 2 (recherche de méthodes d'approximation d'aire sous une courbe). Cours : Chapitre XI - Calcul intégral et lois à densité. I. Aire sous la courbe / II. Intégrale d'une fonction continue, positive. a) Définition.	Feuille synthèse problèmes 1 et 2 à compléter.
Séance 5 (2h) Samedi 21 mars	Cours (suite) : II. a) Définition. Premiers exemples d'intégrales. Recherche des premiers exercices de la feuille + Correction de l'exercice 1.	Exercices 2, 3 et 4.
Séance 6 (1h) Mardi 24 mars	Correction des exercices 2 et 3. Cours (suite) : II. b) Premières propriétés / c) Théorème fondamental (visualisation avec *GeoGebra*).	Exercice 8 à faire sur feuille.
Séance 7 (1h) Mercr. 25 mars	Cours (suite) : II. b) Premières propriétés / c) Théorème fondamental (visualisation avec GeoGebra + démonstration du théorème). Exemple : Retour sur le problème 2.	Exercice 7 à faire sur feuille + Déterminer l'expression de F (dans l'exemple du cours).
Séance 8 (1h) Vendr. 27 mars	Correction exercices 4 et 8. Cours (suite) : II. c) Théorème fondamental (Exemple : Retour sur le problème 2) / III. Primitives. (Théorème et démonstration de l'existence)	
Séance 9 (1h) Vendr. 27 mars	Cours (suite) : III. Primitives. (Reprise de la démonstration de l'existence). Propriétés (Démonstration). Exercice : recherche de primitives au brouillon.	Exercice 5 (pour le cours en demi-groupe).
Séance 10 (1h) Mardi 31 mars	Cours (suite) : III. (Remarque) / Tableau des primitives Exercice 10 (voire exercices 11 et 12).	Exercices 10 et 12.

Séance et date	**Contenu de la séance**	**Travail à faire pour la séance suivante**
Séance 11 (1h) Mercr. 1er avril	Correction exercices 10 et 12. Cours (suite) : IV. Calcul intégral (théorème, démonstration). Exercice 13. Correction des premiers exemples.	Exercice 5 (pour le cours en demi-groupe). Exercice 13 (I_4, I_11, I_13, I_6, I_7 et I_8).
Séance 12 (1h) Vendr. 3 avril	Correction exercice 13 (I_4, I_11, I_13, I_6, I_7 et I_8). Cours (suite) : V. Lois à densité (reprise de la synthèse des problèmes 1 et 2, polycopiés loi uniforme, loi exponentielle).	
Séance 13 (2h) Samedi 4 avril	Cours (suite et fin) : V. Lois à densité (à partir de l'espérance de la loi exponentielle). Exercices 39 p. 386, 48 et 51 p. 387.	Exercices 59 p. 390 et 30 p. 385 (pour le cours en demi-groupe).

Annexe I

Les données relatives à la séquence

I.1 Transcription de la séance 1

Transcription de la séance du 14/03/15 (2h – classe entière)

31 élèves présents.

[00:00] *Sonnerie. Mise en place des élèves.*
[00:44] *Marie écrit au tableau : « Problème 1 ». Les élèves continuent de s'installer.* […]
[01:04] Marie : Bien. Cahier d'exercices s'il vous plait.
Marie fait le point sur les absents.
[01:44] Marie : Alors on va… Chut. On va d'abord être attentif…
Marie : Et donc, je vais vous donner l'énoncé d'un petit problème.
Arrivée d'une élève en retard.
Marie (à l'élève) : C'est bien parce que c'est un cours important. Je suis déjà assez embêtée qu'il en manque un certain nombre. […]
[02:37] Marie : Donc je vous distribue… Vous prenez votre cahier d'exercices, A11. D'accord ? Je vous distribue l'énoncé d'un petit problème. Je vous laisse le temps de le lire en… lecture silencieuse.
Marie : Silencieuse et individuelle. Vous voyez ce que je veux dire. Et je vous laisse 5 minutes pour vous lancer… Un petit truc… et au bout de 5 minutes, je répondrai à d'éventuelles questions. D'accord ? Sans répondre sur comment on fait, bien sûr. Si la réponse est « on fait quoi ? »… enfin si la question est « on fait quoi ? », il n'y aura pas de réponse. Allez ! Alors vous prenez bien en note tout ce qu'on va faire aujourd'hui…
[03:16] *Marie distribue l'énoncé à chaque élève en silence. Chuchotements d'élèves. Le problème est vidéoprojeté.*

Problème 1 :

Karine et Olivier décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Ils peuvent arriver à tout moment entre 7h et 8h. Que peut-on dire du temps d'attente du premier arrivé ?

[…]
Marie : Tout le monde a bien un petit papier qui s'appelle « Problème 1 » ?
Marie : Bien.
[…]
[04:24] Marie : Lecture silencieuse pour l'instant…
Quelques chuchotements. Les élèves se mettent progressivement au travail (lecture de l'énoncé).

Marie passe dans les rangs.
Certains élèves posent des questions à l'enseignante, en individuel [peu audible].

D'après mes notes et l'enregistrement (sauts de lignes quand l'enseignante change d'interlocuteurs) :

Marie (à un élève) : Qu'est-ce que ça t'inspire ?

L'élève B31 parle avec l'enseignante : elle semble avoir des idées. Marie lui dit de ne pas parler trop fort pour ne pas dévoiler aux autres ses idées tout de suite.

Élève A41 : Je comprends pas la question.
Marie : Tu comprends pas la question…
Marie : Qu'est-ce que tu peux dire du temps d'attente du premier arrivé ?
Élève A41 : Rien.
Marie : Mais des trucs basiques… pas…

Certains élèves réfléchissent à deux.

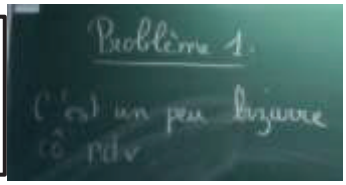
Marie : Tu vois pas l'intérêt de la question ?
Élève B52 : Non mais il arrivera jamais à 7h…
Marie : D'accord…
Élève B53 : Mais c'est pas possible…
B52 et B53 discutent ensemble. Ils utilisent les mots « probabilités », « comme on a fait hier », « 1/60 »…

[08:22] Marie (à la classe) : Chut chut chut. J'ai bien dit que c'était individuel pour l'instant…

Certains élèves réfléchissent toujours à deux.

[09:04] Marie (à la classe) : Alors…
Marie : D'abord… Est-ce que vous avez… des questions… sur le problème ? Est-ce que la situation est comprise par tout le monde ?
Marie : B22 ?
[09:20] Élève B22 : Je disais que c'est un peu bizarre comme rendez-vous…
Marie : Alors. C'est un peu bizarre comme rendez-vous. [*Marie commence à écrire au tableau].*
Élève B22 : Non, non mais écrivez pas ça…
Marie : Si si, on va écrire ça. « C'est un peu bizarre comme rendez-vous ». [*Marie écrit au tableau en même temps*]. Alors… est-ce que vous êtes d'accord ? Qu'est-ce que tu veux dire par « c'est un peu bizarre » ?

Problème 1 :
C'est un peu bizarre
cô rdv

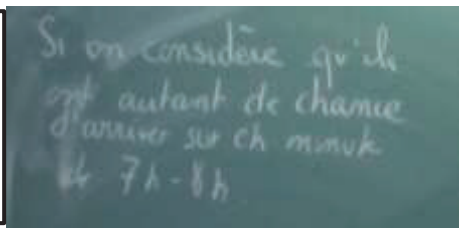


Élève B22 : On se donne pas rendez-vous entre 7h et 8h…
Plusieurs élèves parlent en même temps.
Élève B41 : Soit c'est pour manger, soit…
Marie : Attendez, attendez… Juste, vous prenez la parole dans l'ordre sinon on va jamais y arriver. Donc, B22, tu dis « on se donne pas rendez-vous entre 7 et 8h » et… tout de suite, B41…
Élève A52 : Objection !
Marie : Objection !
Marie : Alors, A52, qui lève la main. Objection ?
Élève A52 : Je réponds vraiment à ça ou… ?
Marie : Oui. Oui, vu que c'est sa question.
Élève A52 : Ah ok… Je sais pas moi… par exemple, on peut avoir un problème de transport, arriver à 7h30 comme à 7h, comme à 8h…
Élève B22 : Oui mais à l'origine on se donne… pardon…
Élève A52 : On peut pas se communiquer si on n'a pas de téléphone donc…
Marie : Tu… voudrais dire que c'est un problème de l'époque de Madame A. [elle-même] et… ?
[*Inaudible*].
Marie : B41, pourquoi tu disais… aussi tu réagissais à la… à le « c'est un peu bizarre comme rdv » ?
Marie : Dis.
Marie : Qu'est-ce que tu veux dire ? Toi, tu donnes toujours rendez-vous… entre… un truc large quoi…
Élève B41 : Pour moi quand il y a écrit entre 7h et 8h, il y a aucune chance que quelqu'un arrive à 7h. Aucune chance.
Marie : D'accord. Mais alors tu veux qu'on écrive quoi ? 7h01 ?
Élève B41 : Non, je sais pas. Entre… si c'est entre 7h et 8h, c'est pas avant 7h15 et limite… plus de 8h quoi.
Plusieurs élèves parlent en même temps.

Marie : Non, j'ai pas compris. Non alors… On est d'accord qu'après 8h, ils s'en… ils s'en vont, c'est fini. Le problème est mort, enfin voilà. Mais, j'ai pas compris ce que tu disais entre 7h et 7h15.
Élève B41 : Et bien, il faut…Vu que…
Marie : J'ai pas entendu… J'ai pas bien entendu…
Élève B41 : Vu qu'entre 7 et 8h, c'est grand… c'est une heure… si on arrive à 7h…
Élève A52 : Il y a le quart d'heure de bienséance.
Élève B41 : … il y a presque plein de chance d'attendre longtemps.
Marie : Il y a le quart d'heure de bienséance. Je ne le connaissais pas celui-là. Tu veux dire quand, par exemple, on t'invite à une soirée à partir de 19h… t'arrives pas avant 19h15… C'est ça le quart d'heure de bienséance ?
Élève A52 : Voilà.
Marie : D'accord. Ok.
Marie : Alors bon, peut-être, se donner rendez-vous entre 7h et 8h, c'est une plage horaire relativement grande… mais aussi pour que ce soit plus facilement traitable pour nous. On aurait pu se dire « ils se donnent rendez-vous entre 7h et 7h15 », est-ce que ça changeait ce qu'on va… est-ce que ça changeait le problème ?
Élève A52 : Non.
Marie : Non. Donc on va garder cette plage horaire grande même si… elle est peut-être un petit peu irréaliste. Alors par contre, revenons quand même sur le « c'est un peu bizarre ». Vous prenez des notes sur ce que vous êtes en train de dire ?
Marie : Alors.
Marie : Qui dit… parlait des problèmes… C'est A52 qui a dit « on peut avoir des problèmes de transport » etc. Prenons un créneau que vous connaissez… Prenons un créneau que vous connaissez beaucoup mieux, celui du samedi matin.
Marie : En gros, on part tous de chez nous à la même heure sachant qu'on a cours à 8h. On va pas commencer à dire sauf A41… Bon, sachant qu'on a cours à 8h10… mais malgré tout… Je sais pas moi j'arrive ici… […] Tout en partant toujours exactement à la même heure de chez moi, j'arrive ici entre 7h30 et 7h45. Tout à fait… aléatoirement.
Marie : D'accord ?
Marie : Donc parce que effectivement il y a les transports, que il y a des gens qui se promènent sur la voie, que je sais pas ce jour-là, il y en a un qu'est pas venu prendre son métro donc il y en a un de moins… enfin il peut se passer plein de choses donc disons que c'est tout à fait réaliste. Alors je suis d'accord avec vous se donner rdv entre 7h et 8h, c'est pas très sympa mais bon… ça change pas le problème si on avait dit entre 7h et 7h15 pour le traitement qu'on va en faire… pour la solution qu'on va apporter. C'est bon ?
Marie : Donc est-ce que. B52 ?
[13:14] Élève B52 : Non mais moi c'est pour après.
Marie : Oui. Oui oui bien sûr.
Élève B52 : Vu qu'il y a 60 minutes entre 7h et 8h, si on considère…
Marie : Voilà. C'est vrai.
Élève B52 : Non mais… si on considère qu'il y a autant de chance qu'ils arrivent à… chaque minute… entre 7h et 8h… donc ils ont une chance… en fait… Karine a 1 chance sur 60 d'arriver à… x heure… donc ils ont un soixante… un sur soixante au carré…
[…]
Marie : Alors attends, attends, attends. Donc si on considère… T'as dit quoi comme hypothèse de travail ? Qu'ils arrivent… au hasard c'est ça ?
Élève B52 : Autant de chance.
Marie : Autant de chance.
Élève B52 : Autant de chance d'arriver sur… sur chaque minute de l'intervalle.
Marie commence à écrire en même temps au tableau.
Marie : Alors est-ce qu'on est d'accord que ça c'est une hypothèse raisonnable et… Prenons Karine, quand elle donne… elle a autant de chance d'arriver entre 7h… enfin autant de chance de quoi… finis ta phrase pardon B52.

Élève B52 : D'arriver sur chaque minute de l'intervalle.
Marie écrit au tableau.

Si on considère qu'ils ont autant de chance d'arriver sur ch. minute de 7h-8h



Marie : Alors de l'intervalle 7… je l'écris comme ça, ça te va ?
Marie : Alors est-ce que… Oui alors, il y a déjà des réactions j'ai l'impression sur cette phrase. Moi je note pas tout, c'est vous qui prenez vos notes. Je note des choses pour qu'on fixe un tout petit peu parce que sinon j'oublie ce que vous dites surtout. B31 ?
Élève B31 : Moi je dirais que c'est… on parle pas en minute, là on parle à tout moment donc… je dirais que même les secondes, il y a déjà beaucoup plus de… Par exemple, si on arrive à 7h28min37s, c'est pas pareil. En plus, il y a les millièmes de seconde et…
Plusieurs élèves : Et les milliards de seconde.
Élève B31 : Et les milliardièmes de seconde après.
Élève A22 : C'est comme la puce hier.
Élève B31 : Oui, c'est infini en fait.
Marie : Ah, c'est comme la puce.

Comme la puce dm6

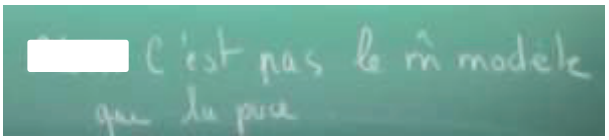
Élève B31 : Oui, c'est comme la puce.
Élève A52 : Sauf que…
Élève B22 : Sauf que c'est des minutes…
Élève B31 : Oui, mais c'est comme la puce.
Élève A22 : Quand t'arrives… quand t'arrives à 7h31 et que l'autre arrive à 7h32…
Marie : Chut, chut.
Élève A22 : Non à 7h31 virgule zéro cinq et 7h virgule trente-deux, t'es à la même heure.
Élève B22 : Oui, mais on est en maths.
Marie : Ah ! « Mais on est en maths ».
Marie écrit au tableau.

Mais on est en math

Marie : Non, non mais ça c'est intéressant. On est en maths. Alors, elle a raison B22. Alors, d'abord, pourquoi il y en a qui ont dit… Qui a dit « c'est comme la puce » ?
Élève ? : A22.
Marie : J'ai pas entendu parce que vous parlez un peu de façon désordonnée quand même. A22, pourquoi tu lui disais « c'est comme la puce » ?
Élève A22 : C'est… Il y a une possibilité infinie de… d'arriver… enfin…
Élève ? : Zéro.
Marie : Non, laisse-le finir sa phrase sinon c'est pas possible qu'on discute.
Élève A22 : Entre 8… entre 7h et 8h, on a une possibilité infinie d'arriver à une certaine heure.
Marie : Une infinité de moments tu veux dire ? d'instants ?
Élève A22 : C'est comme hier quand B52, il voulait marcher et…

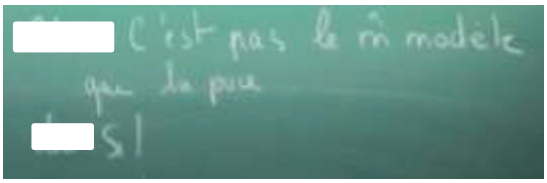
Marie : Oui, il voulait sauter… […] Comme quand ta puce elle sautait de un en un, d'accord ? Est-ce que tu vois ? Est-ce que tu es d'accord avec ça ?
Élève B52 : Oui, mais c'est pas le même modèle.
Certains élèves ne semblent pas d'accord.
Marie : Alors… Alors… Donc… B52… [*Marie écrit au tableau*].
[…]
Marie : Donc, c'est pas… C'est pas le même modèle que la puce. [*Marie écrit en même temps au tableau*]. Oui alors, « ce n'est pas » en français, d'accord ? C'est pas le même modèle que la puce. Pardon. [*Marie termine d'écrire la phrase au tableau*].

B52. C'est pas le même modèle que la puce



Marie : Il y a que moi qui prends des notes ? Non ? C'est pas le même modèle que la puce. Alors j'ai entendu… De toute façon je crois qu'il y a une partie de… B53 ?
Élève B53 : Bah si.
Marie : Ah bah si !
Marie complète le tableau.

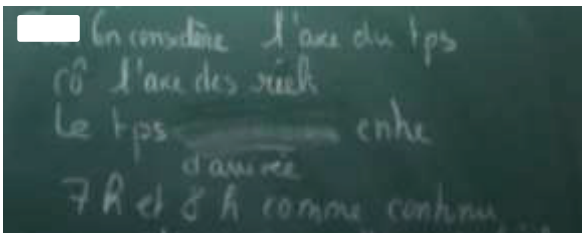
B52. C'est pas le même modèle que la puce
B53 : Si !



Élève B53 : Puisque… Si, parce que la puce elle avait… enfin elle pouvait varier sur tous tous tous les… les points.
Élève B53 : Enfin, les points petits… et là on peut s'arrêter…
Marie : Tous les points « petits » ? Il y avait des petits points et des gros points sur le machin ?
Marie : Qu'est-ce tu veux dire ?
Élève B53 : On commence par un, deux, trois… et les points : un virgule…
Élève ? : Intermédiaires.
Élève A52 : Les réels, quoi.
Marie : Sur tous les nombres réels.
Élève B53 : Oui voilà.
Marie : Voilà. D'accord. Alors… C41 ?
Élève C41 : On peut aussi considérer l'axe du temps comme l'axe des réels et ainsi considérer… le temps d'attente entre 7h et 8h comme un ensemble continu et non comme un ensemble discret.
Marie : Je n'aurais pas mieux dit.
Élève C41 : Euh… Non.
Marie : Je n'ai pas tout noté mais… Donc on peut considérer l'axe du temps comme l'axe des réels. [*Marie écrit au tableau en même temps*].
Marie : L'axe des réels… Et excuse-moi, tu…
Élève C41 : Et ainsi considérer le temps d'attente entre 7h et 8h comme un ensemble continu et non un ensemble discret.
Marie : Le temps d'attente entre… Alors, on était sur le temps… le moment d'arrivée mais c'est pareil enfin… parce que là, on était en train de parler du temps d'arrivée de l'un des… des personnages. Le temps d'attente ou d'arrivée.
Élève C41 : Ah oui.
Marie : Parce que entre 7 et 8h, c'est pas le temps d'attente, c'est le moment d'arrivée d'une personne.

Élève C41 : Ah.
Marie : Tu es d'accord ? Le temps d'arrivée entre 7h et 8h comme… continu tu as dit ? [*Marie écrit au tableau en même temps*].

C41 : On considère l'axe du tps
cô l'axe des réels
Le tps d'arrivée entre
7h et 8h comme continu



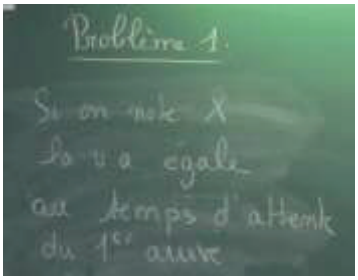
Marie : Il y a quelqu'un qui voulait rajouter quelque chose ? Alors attends parce que… […] Alors, B52 ?
Élève B52 : Ce que je voulais dire là ?
Marie : Oui, qu'est-ce que tu en penses de… de ce que dit C41, de ce que… de ce que dit B53 ?
Élève B52 : Non, mais parce que le modèle de la puce qu'on a vu hier, c'était un modèle où on raisonnait en terme de distance alors que là on raisonne en terme de temps.
Marie : Oui mais c'est… ce que te dit C41 c'est qu'elle… le temps on peut… c'est continu et que donc ça serait comme la puce qui se déplaçait entre 0 et 12.
Élève B52 : Non mais ce que je veux dire c'est que les… les probabilités qu'on a vu hier, c'était… On les a vu en fonction d'une distance sur des distances alors que là on parle en terme de temps. C'est pas la même chose. Enfin, je suis pas sûre qu'on puisse appliquer le même modèle.
Marie : D'accord. T'es pas sûr. Alors il y a quand même… Je voudrais juste revenir à… à la remarque de B22, je crois, « Mais on est en maths ». Qu'est-ce que tu voulais dire par cette phrase « Mais on est en maths » ?
Élève B22 : Il disait que quand on arrivait à un moment on disait pas qu'on arrivait à 7h30 et 28 minutes…
Marie : 28 secondes.
Élève B22 : 28 secondes mais… Donc ça, c'est dans la vraie vie et en maths, on peut arriver à 7h30…
Marie : Voilà.
Élève B22 : Et 28 secondes.
Marie : Voilà. Et… voilà, on va on va… Effectivement on va dissocier la… la réalité, la vraie vie. On va pas… On a des montres, c'est des appareils de mesure, il y a une certaine précision. Et donc on va peut-être considérer, comme le disait C41, que c'est continu. Ça l'est d'ailleurs, le temps est continu. Et c'est indépendamment de l'appareil de mesure qu'on a de ce temps, B52. Donc évidemment, elle a raison B22, quand on prend une situation concrète, d'accord, et que l'on passe en mathématiques, on est en maths, qu'est-ce que… qu'est-ce qu'on fait forcément de cette situation concrète ?
Élève B31 : On la simplifie.
Marie : B31 ?
Élève B31 : On… On l'arrange pour que ça…
Marie : On l'arrange ?
Élève B31 : Non mais on…
Marie : Oui, mais non mais vas-y, dis les mots qui… que tu veux. « On l'arrange » c'est une façon de le dire.
Élève B31 : On… En gros, on rajoute des choses pour… ça… le… pour pouvoir faire un modèle.
Marie écrit au tableau.
Marie : D'accord. Pour qu'on puisse le faire. Donc, entre réalité et maths, on arrange le problème pour qu'on puisse faire des maths. [*Marie écrit au tableau en même temps*]. Pour qu'on… enfin… qu'est-ce tu veux dire ? Pour que…

Entre réalité et math, on « arrange » le pb

Élève B31 : En gros… Je sais pas comment dire.

Marie : Bon, on arrange le problème comme… Bon alors, est-ce que tout le monde est d'accord pour qu'on fasse cette hypothèse, c'est-à-dire que… Donc là on est sur l'heure d'arrivée d'un personnage… […] Est-ce que tout le monde est d'accord pour qu'on considère, on fasse l'hypothèse, que l'instant d'arrivée d'un enfin des deux personnes, ce soit… comment on va dire alors ? En prenant ce qu'a… ce qu'a dit C41. C41 ?
[21:43] Élève C41 : P de X appartient à l'intervalle [0 ; 60].
Marie : Alors, qu'est-ce que tu… c'est qui ? Alors, P je présume que c'est une probabilité que tu veux mettre. X c'est quoi ?
Élève C41: X c'est la variable aléatoire qui représente le temps d'attente…
Marie : D'accord.
Élève C41 : … du premier arrivé.
Marie : D'accord. Alors, on va écrire ça. C22 ?
[22:05] Élève C22 : On peut dire qu'il y a le temps d'arrivée de Karine et le temps d'arrivée d'Olivier, et le temps d'attente, ce sera $T1 - T2$.
Marie : Alors. Donc. Pour résumé. Alors si on note X, mais alors quand tu me parles de probabilité du coup C41, qu'est-ce que tu dis de ce temps d'attente ? Enfin de la situation ?
Élève C41 : En vrai, je pense qu'il pourrait y avoir une chance sur un certain…
Marie : Non, quand tu commences à me dire, j'ai mis X une variable aléatoire… Je sais pas quoi, le temps d'attente ou… c'est ça que tu as dit ? Donc tu es en train de nous dire que la situation c'est quoi ? Enfin qu'on est dans quel cadre ? A52 ?
Élève A52 : Moi je dirais, si on prend la variable aléatoire qu'elle pourrait prendre les valeurs entre 0 et 60, les valeurs réelles entre 0 et 60. Et à partir de là…
Marie : Oui mais… D'accord. Donc on est dans l'aléatoire. On est d'accord ? Le temps d'attente, c'est aléatoire. C'est aléatoire parce que quoi ?
Élève A52 : On sait pas quand est-ce que le deuxième va arriver.
Élève C41 : C'est une fluctuation d'échantillonnage. [*Marie n'entend pas cette proposition parallèle*].
Marie : On sait pas quand est-ce que les deux… Les deux arrivent aléatoirement entre 7h…
Plusieurs élèves : Et 8h.
Marie : Et 8h. Alors, donc donc c'est… on va noter X, c'est ça que vous… Si on note X. Donc c'est quoi X ? X que tu as pris, C41 ?
Élève C41 : La variable aléatoire… Peut-être qu'on peut prendre le modèle de C22 alors ? Enfin si elle, elle a quelque chose…
Marie : On va mettre juste après. On va mettre les deux. Finis… finis ta phrase et après je vais noter ce qu'a dit C22. Donc si on note X la variable aléatoire… [*Marie écrit en même temps au tableau*].
Élève C41 : Qui est…
Marie : Alors égale.
Élève C41 : Egale… le temps d'attente du premier arrivé.
Marie : Alors au temps d'attente du premier arrivé. Au temps d'attente… [*Marie écrit en même temps au tableau*].

Problème 1 :
Si on note X
la v.a. égale
au temps d'attente
du 1er arrivé



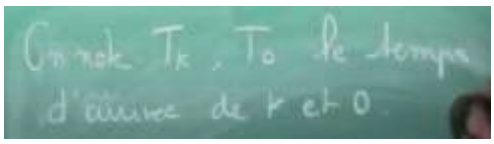
Élève C41 : Juste j'ai une question. On peut quand même considérer que le premier arrive à 7h pile et que le deuxième arrive à 7h pile ?
Marie : Pourquoi ?
Élève C41 : Est-ce qu'on peut considérer ça ?

Élève A52 : Oui.
Marie : Est-ce que ça peut arriver ?
Élève A52 : C'est possible.
Plusieurs élèves : Oui.
Marie : S'il n'y pas de temps d'attente, ça veut dire qu'il vaut combien ?
Plusieurs élèves : Zéro.
Marie : Zéro.
Élève B11 : Il y a pas de premier alors.
Marie : Comment ?
Élève B11 : Il y a pas de premier alors.
Marie : Ah d'accord. Enfin ils sont ex aequo. Premier ex aequo alors. Et on dira qu'on considère que l'ex aequo c'est possible. L'égalité, c'est possible. Alors est égale au temps d'attente du premier arrivé, tu m'as dit. [*Marie écrit au tableau en même temps*]. Et après, je sais plus, tu as dit autre chose.
Élève C41 : P de X appartient à l'intervalle [0 ; 60].
Marie : Alors, tu me dis « P de X », donc j'imagine c'est la probabilité de X de quoi ? [*Marie écrit en même temps au tableau*].

$p(X \in [0;60])$

Élève C41 : Enlevez la deuxième parenthèse… Appartient… Pardon. C'est comme hier…
Marie : Comme hier. Oui d'accord et alors, qu'est-ce que j'en fais de cette chose, moi ?
Élève C41 : C'est la probabilité.
Marie : Oui ?
Élève C41 : C'est la probabilité du temps d'attente. C'est la probabilité.
Marie : D'accord.
Élève C41 : C'est le chapitre.
Marie : Ah non. T'as vu un chapitre quelque part toi ?
Élève C41 : Non, mais je sais bien ce qu'on va faire.
Marie : D'accord. Alors bon là je comprends pas trop ce que tu veux nous dire mais peut-être que… C22. Je vous prends, je vous donne la parole après. C22, toi tu voulais qu'on note ta… je sais pas quoi.
Élève C22 : Je voulais qu'on note le temps…
[…]
Élève C22 : Je voulais qu'on note le temps d'attente en fonction du temps d'arrivée des deux…
Marie : D'accord, alors tu as noté des choses c'est ça ?
Élève C22 : J'ai mis T_K et T_O…
Marie : Alors T_K… T_O… [*Marie écrit au tableau en même temps*].
Élève C22 : Le temps d'arrivée des deux.
Marie : D'accord. Alors l'horaire… enfin le temps on va dire, tout le monde comprend. Le temps d'arrivée. Donc c'est ça que tu as noté toi ? D'arrivée de Karine et Olivier, que j'écris juste par les initiales.

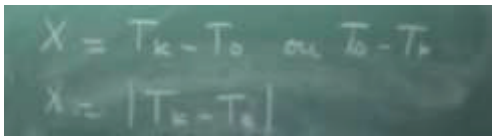
On note T_K, T_O le temps d'arrivée de $\mathcal{K}$ et O.



Élève C22 : Après, j'ai noté T_A le temps d'attente…
Marie : Oui alors du coup on va peut-être reprendre le… X parce que si on… on multiplie les notations, on va pas y arriver. Donc X du coup…
Élève C22 : C'est égal à $T_K - T_O$ ou $T_O - T_K$ suivant…

Marie : Alors, $T_K - T_O$ ou ... [*Marie écrit au tableau en même temps*]. Alors comme vous le voyez, nous sommes au brouillon. Bien sûr, jamais je n'oserais écrire… - On lève le doigt - ce genre de chose dans une… Alors, on est au brouillon du coup on a le droit de faire ce qu'on veut. B52 ?
Élève B52 : On a qu'à prendre la valeur absolue directement.
Marie : Oui. Donc en fait, valeur absolue… Est-ce que tout le monde est d'accord que le temps d'attente, ça serait… c'est égal à ça ? [*Marie montre le tableau*].

$$X = T_K - T_O \ ou \ T_O - T_K$$
$$X = |T_K - T_O|$$



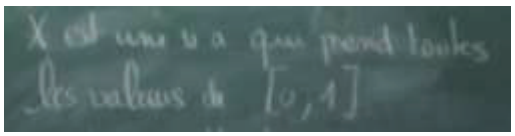
[26:55] Marie : Oui ? Alors, il y avait des mains levées là-bas tout à l'heure. Donc A22 ? A22.
Élève A22 : Moi j'aurais mis le temps en seconde et pas en minute. Parce qu'en minute, enfin c'est… on est sur 60 et ça va être moins facile pour le calculer alors que si on se met directement en seconde… ça va être sur 100. Enfin c'est… Enfin… Non sur 10, pardon. [*La classe s'agite*]. Enfin… Non mais c'est…
Marie : B31, je crois que tu veux répondre à A22.
Élève A22 : Non, ça fait 3600 enfin…
Rires.
Élève A22 : C'est 3600 mais…
Marie : Souffle-lui un petit peu plus discrètement, A21.
Rires.
Marie : Et ça sera plus facile de faire sur 3600 que sur 60 ?
Élève A22 : Oui parce que après on peut passer aux décimales plus facilement.
Marie : Bon, attends, je crois que B31 veut…
Élève B31 : En fait je pense que ça change rien que… puisque c'est comme hier donc… l'intervalle que tu prends, faut juste un intervalle… en gros on s'en fiche des valeurs.
Marie : C'est aussi que quand tu es en train de me parler de minutes et de secondes, tu restes sur une situation de mesure du temps et pas sur une situation continue du temps… en mathématiques, tu vois.
Élève A22 : Comme hier on avait fait… Dix puissance n et… enfin… pour… on avait dit que la valeur…
Marie : Oui mais elle… elle a pris [0 ; 60]. On n'a pas dit d'abord qu'on allait le [Inaudible]. Elle a… Si c'est entre 7h et 8h, on va pouvoir travailler entre quoi et quoi ?
Élève ? : 0 et 1.
Marie : Entre 0 et 1 ? D'accord ? C'est ce que te répond B31, peu importe en fait… Alors, il y avait A31.
[28:29] Élève A31 : Oui donc si par exemple on prend l'intervalle entre 0 et 1 donc on ne peut pas avoir la probabilité pour une valeur alors c'est égal à 0.
Marie : Alors si on prend entre 0 et 1…
Élève A31 : La probabilité pour qu'il soit…pour que le temps d'attente soit égal à 0,50 par exemple 530… 0,530…
Rires.
Marie écrit au tableau.

$$p(X = 0{,}530) = 0$$

Elève ? : C'est subtil…
Élève A31 : Il y a une infinité de… de chance, donc…
Marie : Il y a une infinie de chance… Ça c'est infini ça.
Élève A31 : C'est égal à 0 puisque…
Marie : Ah d'accord, c'est égal à zéro.
Élève ? : Oui puisque c'est l'infini.
Élève A31 : C'est ce qu'on a fait hier.

Marie : Oui. D'accord. Pourquoi ?
Élève A31 : Puisqu'on l'a démontré hier.
Marie : Ah ! C'est comme la puce.
Élève A31 : Oui.
Élève C41 : Oui.
Marie : D'accord. Ok ? A41 ?
[29:25] Élève A41 : Je comprends pas comment… le temps peut être égal à une probabilité.
Marie : On n'a pas écrit…
Élève A41 : X la valeur aléatoire…
Marie : On n'a pas écrit que le temps… enfin… que la probabilité que le… ah oui alors, je reviens sur ce qu'on a écrit précédemment. Est-ce que… peut-être ici, tu es d'accord que… Alors voilà, on note T_K le temps… non mais je vais revenir dans l'autre sens… le temps d'arrivée de… l'instant où arrive Karine d'accord et celui où arrive Olivier. Tu es d'accord que notre variable aléatoire, elle s'écrit comme ça ? C'est en… ces deux… T_K et T_O, d'accord ? C'est aléatoire puisqu'ils arrivent à tout moment entre 7h et 8h… donc le temps d'attente c'est aussi quelque chose qui est…
Élève A41 : Aléatoire.
Marie : Aléatoire. Donc on a bien une variable aléatoire. Pas déterminée.
Élève A41 : Oui.
Marie : D'accord ? C'est bon ?
Certains élèves discutent de cela.
Marie : Alors du coup… Revenons peut-être sur… Qu'est-ce que vous vouliez dire, là, par en fait, je pense c'est mal… Cette variable aléatoire-là, qui est égale au temps d'attente du premier arrivé, il y a plein d'élèves qui ont mis qu'une seule chose sur leur phrase… sur leur copie, avant qu'on discute.
Marie : Qu'est-ce que vous aviez mis ? Alors je sais plus qui j'ai vu moi. A52 ?
Élève A52 : C'est que le temps d'attente peut être égal à 0 ou encore 0 et 60 minutes…
Marie : Que le temps d'attente varie entre 0 et…
Élève A52 : et 60 minutes.
Marie : Et 1, on va le dire autrement. Donc. Quand on a une variable aléatoire, qu'est-ce qu'on regarde généralement ? La première chose qu'on regarde ?
Élève A52 : Toutes les valeurs possibles.
Marie : Voilà. L'ensemble des valeurs prises. Et donc, ça prend ses valeurs dans… Ça prend dans l'intervalle… ?
Élève A52 : Dans l'intervalle.
Marie : Toutes les valeurs de l'intervalle, d'accord ? Donc X une variable aléatoire qui prend toutes les valeurs de l'intervalle [0 ;1]. [*Marie écrit au tableau en même temps*].

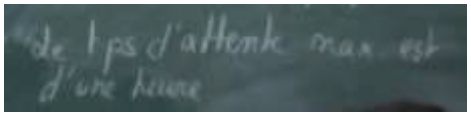
X est une v.a. qui prend toutes les valeurs de [0 ; 1].



Marie : On va se mettre entre 0 et 1, on peut prendre entre 0 et 60, entre 7 et 8… ça changera pas le… le problème. Ah non, entre 7 et 8, c'est stupide. Entre 0 et 1 ou 0 et 60, enfin ce que tu veux. Alors, alors est-ce qu'il y a d'autres choses que vous aviez de prime abord ? Alors, tout de même, il y a à peu près je sais pas 20 minutes, il y a plein d'élèves qui me disaient « il y a rien à dire »… Visiblement on avait quand même des choses… à dire. [32:20] Alors est-ce que vous avez d'autres choses parce que ma question… […] Que peut-on dire du temps d'attente du premier arrivé ? Est-ce que vous avez d'autres… est-ce que vous avez d'autres choses qui vous sont venues tout de suite ? C11 ?
Élève C11 : Le temps maximum [Inaudible].
Marie : Le temps maximum, il est… ?
Élève C11 : D'une heure.

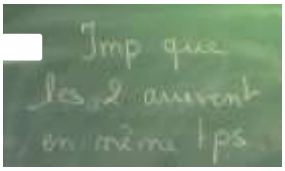
Marie : Le temps d'attente maximum, il est d'une heure. [*Marie écrit au tableau*]. Ce qui est une autre façon de dire que c'est entre 0... Le temps d'attente maximum est d'une heure.

Le tps d'attente max est d'une heure.



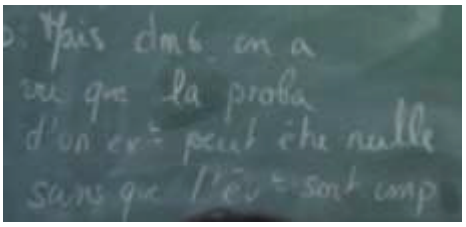
Marie : Alors. Est-ce que…
Marie : B43 ?
Élève B43 : Le temps minimum d'attente est différent de 0.
Marie : Alors, le temps minimum… Tu voulais qu'on ouvre ici alors ? [*Marie montre le crochet au niveau du zéro de l'intervalle* [0 ; 1] *écrit au tableau*]. Alors, est-ce que tu peux expliquer pourquoi tu veux pas qu'on prenne zéro ?
Élève B43 : Le point le plus proche du point de vue mathématique donc X peut prendre toutes les valeurs entre 0 et 1 il y en a une infinité.
Marie : Oui.
Élève B43 : Donc la possibilité qu'ils… de tomber deux fois sur la même… Donc par exemple, si le premier donc la valeur on va dire 0,5… la probabilité que le deuxième arrive à 0,5 pile est égale à 0.
Marie : D'accord. Dans la réalité ?
Élève B43 : Non, du point de vue mathématique.
Marie : Ah d'accord.
Élève B43 : La probabilité que les deux à…
Élève ? : Je suis pas d'accord.
Marie : C'est-à-dire que…
Élève B43 : Que les deux arrivent à 7h30, c'est pas possible.
Marie : D'accord. Alors, on va mettre : B43, donc tu… Impossible que les deux arrivent en même temps.
Élève B43 : Du point de vue mathématique…
Marie : Impossible que les deux arrivent en même temps. [*Marie écrit en même temps au tableau*].

B43 : Imp que les 2 arrivent en même tps



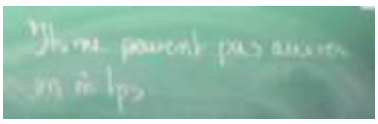
Marie : Alors, il y en a qui n'étaient pas d'accord avec B43. Qui est d'accord avec la proposition de B43 ? […] Qui est d'accord avec la proposition de B43 ? […]
Marie : Je sais plus là. Un, deux, trois, quatre, cinq, six… Qui n'est pas d'accord avec B43 ? Je me méfie, il y a des gens qui s'abstiennent. Il y en a qui ont pas levé la main ! […]
Marie : C31, qu'est-ce que tu en penses ? […] C31 ?
Élève C31 : En fait, je sais pas. Non, mais non.
Marie : Tu sais pas. Il faut choisir.
Élève C31 : En vrai, c'est pas possible.
Marie : Bon, on va laisser en suspens cette proposition. On y reviendra. D'accord ? A32 ?
Élève A32 : On a vu que hier si la probabilité d'un évènement était nulle, ça voulait pas dire que l'évènement il était impossible.
Élève A52 : Voilà, c'est ça.
Marie : Alors, A32. Mais DM6, on a vu que événement… [*Marie écrit en même temps au tableau*]. Non, qu'est-ce que j'ai dit ? Redis-moi ce que tu as dit, toi.
Élève A32 : La probabilité peut être nulle sans que l'événement soit impossible.
Marie : D'accord. Un événement peut être nul sans que l'événement soit impossible.

A32 : Mais dm6, on a vu que la proba d'un évt peut être nulle sans que l'évt soit imp



Élève B43 : Oui, je suis complètement d'accord.
Marie : B43 ? Oui. Pas complètement convaincu non plus.
Élève B43 : On a fait quand même noté P… oui, égal zéro donc... Hier.
Marie : Oui, c'est ce qu'elle dit.
Élève B43 : Oui.
Marie : P de… c'était 3 ? Arriver sur le… 3 ? C'est ce qu'elle dit.
Élève B43 : Oui mais la probabilité est infime.
Marie : Non, elle valait zéro. C'était pas… Infime, ça veut dire quoi ?
Marie : Très très très très petite, c'est ça ? C'est comme les petits rectangles dans l'histogramme ?… T'étais là hier ? D'accord. J'ai pas convaincu tout le monde hier visiblement. Je vois que B52 me regarde encore en se disant…
Élève B52 : En fait il y a rien à comprendre, c'est juste [Inaudible].
Marie : Alors, B14 ?
Élève B11 : En fait c'est pas que c'est impossible ou nul, c'est juste que c'est pas compris dans le problème.
Marie : Pas compris dans le problème ? Qu'est-ce que tu veux dire ?
Élève B11 : C'est hors-sujet.
Marie : C'est hors-sujet ? B43, tu te poses des questions hors-sujet ? […] B11 ?
Élève B11 : Parce que justement dans le sujet ils nous demandent le premier arrivé mais s'ils arrivent tous les deux en même temps il y a pas de premier donc c'est pas possible.
Marie : Ah oui, c'est vrai, toi tu veux pas… toi tu veux pas qu'ils arrivent… Ah oui, d'accord. T'es resté sur ton… Alors. B11.
Élève B11 : L'intervalle…
Marie : D'accord, d'accord je note toutes les propositions, on reviendra sur les propositions sur lesquelles on n'est pas… Donc B11…
Discussions dans la classe.
Marie : Donc ils ne peuvent pas arriver en même temps. On l'a écrit déjà ou pas ? Ils ne peuvent pas… [*Marie écrit au tableau*].
[…]

B11 : Ils ne peuvent pas arriver en même tps

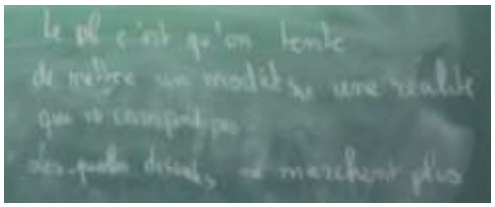


Marie : Il veut pas que les deux personnages puissent… arrivent… soient ex aequo.
Élève B11 : Que peut-on dire du temps d'attente du premier arrivé…
Marie : Donc j'aurais pas dû écrire premier ?
Élève B11 : Oui.
Marie : Et j'aurais dû écrire quoi alors ?
Discussions dans la classe.
Élève B11 : Je sais pas… Non mais… Quel est le premier temps d'attente ?
Marie : Quel est le temps d'attente ?
Élève B11 : Quel est le temps d'attente.
Marie : A52 ?

Élève A52 : En fait l'énoncé.
Marie : [...]
Élève A52 : Non mais je veux dire que l'énoncé il ne sous-entend en aucun cas le fait que les deux ne puissent pas arriver en même temps.
Marie : D'accord. Donc ce serait une hypothèse qu'il faudrait qu'on mette. On suppose que les deux personnes peuvent arriver...
Élève A52 : Oui.
Marie : ...en même temps. D'accord. Est-ce que c'est raisonnable de prendre cette hypothèse ou pas ?
Élève A52 : Oui. Mais en fait on peut pas vraiment dire qu'on peut pas arriver en même temps puisque l'énoncé ne le dit pas.
Marie : D'accord. Tu le fais... Tu le prends à l'envers de B11. C'est pas parce que... comme c'est pas dit, ça veut dire que c'est possible.
Élève A52 : Si on veut raisonner de la probabilité en mathématiques, on peut pas... interpréter... Comme l'exercice de logique qu'on a fait [...]. C'est pas parce que le père dit qu'on n'aura pas de bonbons quand on sera pas sage qu'on aura un bonbon quand on sera sage.
[...]
Élève A52 : Voilà, c'est dans le même esprit.
Marie : D'accord. B52 ?
Élève B52 : En fait, le problème je pense c'est de... de mettre un modèle sur des mathématiques mais alors que la réalité elle correspond pas du tout avec ce qu'on attend de...
Rires.
Élève ? : J'ai pas compris ce qu'il a dit.
Marie : Alors, il est en train de dire que ça ne sert à rien.
Rires.
Marie : Je résume. Non, alors vas-y redis-le B52. Le problème...
Élève B52 : Non le problème, c'est qu'on essaye de mettre... de fabriquer un modèle selon une réalité qui correspond pas avec nos probabilités.
Marie : Alors attends... Qu'on tente de mettre un modèle sur une situation... [*Marie écrit au tableau en même temps*].
Élève B52 : Oui, par rapport à une situation réelle qui correspond pas donc avec nos probabilités... qui ne correspond pas avec nos probabilités.
Marie : Et avec quelle probabilité ? Tu me parles de quoi comme probabilité ?
Élève B52 : Des nôtres, celles qu'on connait... habituellement.
Marie : C'est-à-dire ?
Élève B52 : Que si on trouve zéro, c'est qu'il y a un problème dans... quelque part. Si on trouve zéro, mais que c'est quand même possible, c'est qu'il y a un problème quelque part.
Marie : Oui, c'est que je t'ai pas convaincu.
Élève B52 : Bien non, non.
Marie : Si, si. Mais t'as le droit B52. C'est toujours ça le problème ?
Élève B52 : Oui, mais non. En fait je comprends pas pourquoi on essaye quand même de le faire alors que... on a compris que ça fait zéro, c'est bizarre, c'est un paradoxe, ok mais ça sert à quoi de...
Marie : Mais on n'a pas dit... Moi je vous laisse réagir. J'ai pas dit qu'on allait faire que ça ou que... Donc il y a ce problème, j'ai bien compris, du fait que... qu'ils peuvent pas arriver en même temps parce que ça pose un problème toujours avec ce P de X égal une valeur. D'accord. Peut-être que c'est le modèle qu'est tout faux alors ?
Élève B52 : Donc c'est...
Marie : Donc ta conclusion, c'est... Alors on n'a pas fait grand-chose sur le modèle pour le moment. On a juste dit « on va poser X le temps d'attente ». On sait que ça prend toutes les valeurs entre 0 et 1.
Élève B52 : Non, non mais ce que je veux dire c'est que peut-être les probabilités qu'on connait nous, elles sont pas forcément...
Marie : Ah d'accord. Et c'est quoi les probabilités que tu connais jusqu'à présent ?
Élève B52 : Non, mais... enfin.

Marie : Non, mais si. Elles s'appellent comment ce qu'on a fait jusqu'en première ? [...] Vous avez travaillé en probabilité comment jusqu'à présent ?
Élève ? : Discret.
Marie : Discret. Les probabilités discrètes qu'on connait, c'est ça que tu me dis ? Les probabilités qu'on connait jusqu'à présent donc ce qu'on a fait de la troisième à la... à la première et puis avec le début de l'année, nous, ce sont des probabilités discrètes. Ça veut dire vos variables... Ça veut dire quoi probabilités discrètes ? Qu'est-ce que vous entendez dans le mot ?
Élève B31 : Ça veut dire les valeurs sont séparées les unes des autres. Les valeurs que peut prendre...
Marie : Les valeurs que peut prendre la variable aléatoire... et c'est surtout qu'elles étaient en nombre...
Plusieurs élèves : Fini.
Marie : Fini. D'accord ? Donc les probabilités discrètes ne... marchent plus, c'est ça que t'as dit ? Ne correspondent pas. [*Marie écrit au tableau*].

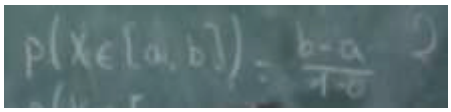
B52 : Le pb c'est qu'on tente
de mettre un modèle sur une réalité
qui ne correspond pas.
Les proba discrètes ne marchent plus



Marie : Oui ! Donc on va être obligé de changer des choses. C41 ?
[41:51] Élève C41 : Non mais je voulais juste dire qu'en fait, enfin je vais un peu reprendre ce que vient de dire B52 mais c'était avant qu'il ait fini...
[...]
Élève C41 : En fait, c'est juste que le modèle est différent de la réalité mais c'est enfin c'est pas faux. C'est pas parce que...
Marie : J'ai l'impression... J'ai pas compris tout à fait ça dans ce que disait B52. Mais j'ai peut-être mal compris B52. Il voulait dire que les probabilités qu'on sait faire, nous, ça va pas marcher pour régler ce problème. Est-ce que c'était ça ta pensée, B52 ? Alors toi ce que tu dis, c'est...
Élève C41 : Alors, je voulais dire que le modèle des probabilités continues ne... enfin n'est pas... ne correspond pas à la réalité mais c'est juste que... enfin...
Marie : Alors, si notre modèle ne va pas correspondre à la réalité, c'est qu'on va... c'est qu'on s'est trompé.
[42:38] Élève C41 : Non, parce que la probabilité de X quand X appartient à l'intervalle $[a\,;b]$, c'est $b-a$ sur... l'intervalle en entier.
Marie : Alors tu penses... ce que tu nous dis là, c'est que la probabilité que le temps d'attente soit entre a et b, c'est ça ?
Élève C41 : ...
Marie : Ce sera $b-a$ sur... ?
Élève C41 : $a\ b$.
Marie : Sur $a\ b$. C'est quoi $a\ b$?
Élève C41 : Sur l'intervalle $[a\,;b]$.
Marie : C'est quoi $a\ b$? l'intervalle $[a\,;b]$?
Élève C41 : L'intervalle a et b...
Rires.
Élève C41 : Non mais a... par exemple ici, a ça serait 0 et b ça serait 60.
Marie : On avait dit qu'on travaillait sur [0 ; 1] quelque part, finalement.
Élève C41 : Ou [0 ; 1], pardon. b égal 1.
Marie : Oui mais le $a\ b$ dont tu me parles là, c'est pas les mêmes que ceux que tu as mis en haut, C41 ?
Marie : Moi j'ai entendu ça. La probabilité que X soit entre a et b, c'est $b-a$ sur $a\ b$, c'est les mêmes lettres.
Élève C41 : $a\ b$ avec les crochets.

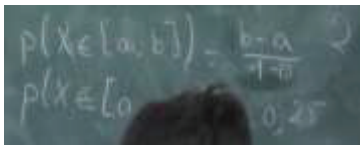
Marie : Non, ça c'est un segment, c'est c'est…
Élève C41 : Faut… faut dire que c'est pas le segment alors.
Rires.
Élève C41 : La distance a b.
Marie : Mais non, non…
Élève B31 : La distance c'est $b - a$. [*Marie n'entend pas cette intervention*].
Marie : … tu vois bien que non.
Élève C41 : La distance a b, je sais pas comment…
Marie : Non, moi non plus.
Élève C41 : a b peut-être…
Marie : Voilà. a b… Bon, est-ce que les autres comprennent ? Non mais vous avez le droit de comprendre, moi c'est… B31 ?
Élève B31 : En gros, elle veut dire que c'est $b - a$…
Rires.
Élève B31 : Sur…
Marie : En gros.
Élève B31 : Non, sur la… la… en gros, la valeur de toutes… l'addition de toutes les valeurs comprises dans l'intervalle et comme c'est infini, en gros c'est l'infini.
Plusieurs élèves ne semblent pas d'accord.
Marie : Non, c'est pas ça qu'elle a dit.
Rires.
Élève C41 : Non t'as raison, la puce c'était sur 12.
Marie : Ah ! Donc là c'est 1 alors ? Entre 0 et 1…
Élève C41 : Oui mais non. Si vous enlevez, si vous mettez le 1, on va dire que ce sera $b - a$, parce qu'on va enlever le 1.
Marie : Oui, là du coup… ça fait $b - a$, on est d'accord ? Oui. Ok. C'est ça que tu proposes ?
Élève C41 : Voilà. C'est pas parce que c'est égal à zéro que le modèle est faux, c'est juste qu'il faut le comprendre comme ça, c'est tout.
Rires.
Élève C41 : Non mais…
Marie : C'est fini. On n'en parle plus. Alors, donc ça c'est la proposition de C41. [*Marie écrit au tableau*].

$$C41 : p(X \in [a;b]) = \frac{b-a}{1-0}$$



Marie : Alors, vous êtes d'accord ? A52 ?
Élève A52 : On n'a pas compris.
Marie : Elle dit, finalement… Prenons un exemple numérique. Tu dis que la probabilité que le premier arrivé attende… alors on va prendre un truc simple entre 0 et 15 min, entre 0 et 0,25. Un quart d'heure.
Élève C41 : 0,15 oui.
Marie : Non, 0,25. Ça y est tu me fais douter. Un quart, ça fait 0,25. C'est bon, c'est bon, je l'ai.
Élève C41 : C'est pas en 60 ?
Marie : Non. Mais entre 0 et 1, peu importe. On a dit entre 0 et 1.
Élève C41 : Ah oui.
Marie : Parce que si y en a qui font entre 0 et 1 et d'autres entre 0 et 60, on va pas s'y retrouver. Donc en fait ce que tu nous dis, c'est que la probabilité que X appartienne à l'intervalle [0 ; 0,25], pour un quart d'heure, d'accord ? Donc le premier arrivé attend entre zéro et un quart d'heure, ce serait égal à donc la différence $b - a$ donc divisé par 1, ça fait ? 0,25. [*Marie écrit au tableau en même temps*].

$C41 : p(X \in [a;b]) = \frac{b-a}{1-0}$
$p(X \in [0;0{,}25]) = 0{,}25$

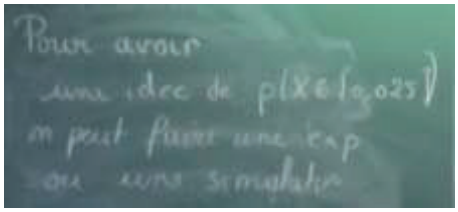


Élève C41 : Oui, enfin l'endroit où je voulais plus en venir c'est que si la probabilité que X appartienne à l'intervalle [0,25 ; 0,25] et bien c'est comme ça, c'est pas très grave. Enfin. Si la probabilité égale zéro, c'est juste que dans le modèle mathématique, c'est… on le traduit par un zéro mais qu'en vrai, ça se traduit par une chance sur… beaucoup.
Rires.
Élève C41 : Non, mais… Une chance sur l'infini.
Marie : Ok. Donc ça, ça… Pour $P(X = 0{,}25)$, ça ferait 0 ? Ce que tu proposes, ça serait cohérent avec ce qu'on a dit avant.
Élève C41 : Oui.
[46:27] Marie : Alors qu'est-ce que vous en pensez de la proposition de C41 ? A31, « ça sert à rien ».
Élève A31 : Non mais c'est que c'est indistant (sic). C'est écart nul. On prend l'intervalle entre 0 et 1, ça veut dire que l'écart… c'est le temps d'attente, mais je veux dire…
Maris : Oui.
Élève A31 : C'est égal à… le temps du premier arrivé et du temps du deuxième arrivé. Le temps… Et le fait de… Une diminution entre eux.
Marie : Oui, on prend la valeur absolue de la différence. Je comprends… Je suis désolée A31, je comprends vraiment pas ce que tu veux nous dire. Tu as compris, A32 ?
Élève A32 : En gros, il a dit que hier on travaillait sur des distances alors que là on travaille sur un écart…
Plusieurs élèves disent que ce n'est pas ça qu'il a dit.
Élève A32 : Si, si c'est ça.
Rires.
Marie : Bon, A31, A32 elle sait mieux que toi ce que tu veux dire.
Rires.
Marie : Enfin, on y reviendra.
Marie : Donc ce serait comme la puce ? Donc ce serait la formule pour calculer la probabilité que le temps d'attente soit entre je sais pas quoi et je sais pas quoi ? [47:38] Qu'est-ce que vous… Est-ce que vous pensez que c'est comme la puce, la probabilité que X soit entre 0 et 0,25 c'est la longueur de l'intervalle de temps divisé par l'intervalle de… total ? […]. C22 ?
Élève C22 : Non c'est comme si il y avait deux puces, non ?
Rires.
Marie : Non, c'est vrai qu'au départ, c'est comme si y avait deux puces et nous on s'intéresse au… à… aux écarts. Donc tu penses que […] en fait… Est-ce que vous êtes d'accord que le temps d'attente de Karine et Olivier c'est comme la puce ?
Plusieurs élèves répondent « oui ».
Marie : Oui ? Donc, comme dit C22, on a deux puces et donc là, cette fois-ci on s'intéresse à la valeur absolue de la différence. Et vous pensez que… Ce que dit C41, c'est que pour calculer la probabilité sur la valeur absolue de cette différence c'est comme calculer la probabilité pour… pour l'un d'entre eux…
Élève C41 : Ah non, c'est…
Marie : Ah je sais pas. Finalement, c'est ça que tu dis.
Élève C41 : Ah non, a c'est le premier et b le deuxième, je pense. Ou b le premier…
Marie : Non, ça c'est le temps d'attente, c'est pas le premier arrivé, le deuxième arrivé. C21 ?
Élève C21 : A ce moment-là, c'est comme si on calculait le temps d'arrivée du premier en fait sans…
Marie : Oui. C'est ce que je suis en train de dire. Elle est en train de nous dire que calculer le temps d'attente, enfin la probabilité du temps d'attente dans un intervalle, ça ce serait la même chose que de calculer la probabilité du moment d'arrivée dans… de l'un d'entre eux dans ce même intervalle. Est-ce que vous pensez que c'est ça ? Qu'en fait, ces deux variables aléatoires c'est les mêmes… ou pas ?

Marie : B52 ? Est-ce que vous avez entendu ce que je… ce que veut dire C41 ?
Élève B52 : Oui. Mais je suis pas d'accord avec ça.
Marie : […] Tu penses que c'est pas la même chose ? Nos deux… Notre temps d'attente et notre moment d'arrivée.
Élève B52 : Non, mais soit on pense que notre modèle, il est faux et que il faut simplement raisonner avec un autre modèle, soit il est vrai, on applique bêtement ce qu'on a vu et c'est tout…
Marie : Bêtement ?
Élève B52 : Non pas bêtement… Enfin on applique… mathématiquement le…
Marie : Est-ce qu'on peut… On est d'accord, c'est pas du discret ?
Élève B52 : Non.
Marie : Donc on peut pas appliquer notre… ce qu'on faisait. Alors, B31, tu veux réfuter enfin tu veux… ?
Élève B31 : Oui moi je pense que c'est pas vrai parce que quand on dit ça, ça veut dire que il y aurait… on en fixe un qui arrive forcément à la même heure alors que… comme…
Marie : C'est comme si on n'en avait considéré qu'un tu veux dire ?
Élève B31 : Oui.
Marie : En fixant l'un.
Élève B31 : En en fixant un et en fait c'est…
[50:04] Marie : Est-ce que vous pouvez pas, si vous êtes pas d'accord avec ça, est-ce que… vous pouvez pas… donner… Quand on est pas d'accord avec quelque chose, qu'est-ce qu'on peut fournir ?
Agitation dans la classe.
Marie : Un contre-exemple, oui. Alors ?
Marie : Bon, c'est peut-être un petit peu difficile. Bon, alors, il faudrait, il faudrait qu'on… Alors, B31 ?
Élève B31 : Par exemple, le premier, il arrive au bout de 15 minutes.
Marie : Donc il arrive… le premier arrive à 7h15.
Élève B31 : Il arrive à 7h15, si l'autre il arrive à 7h… 7h20 ?
Marie : Oui.
Élève B31 : Ça marche pas.
Discussions dans la classe.
Marie : Le truc marche pas ? Ce que tu dis marche pas ? Ou… ?
Élève B31 : La formule, elle marche pas je crois.
Marie : D'accord. Bon alors comment est-ce qu'on peut faire ? Qu'est-ce qu'on pourrait faire pour… réfuter le… le calcul de… de C41 ? On n'a pas entendu B22.
Élève B22 : Trouver un contre-exemple.
Marie : Oui, mais on a du mal visiblement. Donc qu'est-ce qu'on pourrait faire d'autre ? Qu'est-ce qu'on pourrait faire pour en savoir plus sur notre fameux temps d'attente, là ?
Élève B52 : Moi, j'ai pas compris pourquoi ça marchait pas.
Marie : Ah, parce que… J'ai pas dit que ça marchait pas ou que ça marchait, j'ai dit ce que disait C41, c'est que finalement calculer la probabilité que le temps d'attente soit dans l'intervalle [0 ; 0,25] - d'accord ?- ce serait… on ferait le même calcul pour le temps d'attente… le temps d'arrivée d'Olivier entre 0 et 0,25. C'est ce qu'elle dit. D'accord ? Et donc, comme disait C21 et C22, à leur sens, calculer la probabilité pour le temps T_K, dans un certain intervalle, c'est pas la même chose que de calculer… enfin c'est pas la même façon que de calculer la probabilité du temps d'attente qui est la différence… la valeur absolue de la différence entre T_K et T_O. […]
Marie : Alors, on va peut-être… A32 ?
Élève A32 : Du coup, j'ai pas compris pourquoi le contre-exemple de B31, il marchait pas.
Marie : C'est elle, elle dit « ça marche pas ».
Élève B31 : Non c'est pas ça. Parce que par exemple Karine, elle arrive entre 0 et 0,25 et après Olivier, il arrive juste 5 minutes après donc en gros, ça fait pas 0,25 ; le temps d'attente.
Élève A32 : Oui mais si j'ai compris. Pourquoi vous avez dit que ça marchait pas ?
Marie : Non, c'est elle qui a dit « ça marche pas », elle s'était arrêtée dans sa phrase.
Élève B31 : Ah non, pour moi, ça marche.

Marie : Ah, j'avais pas compris.
Élève B31 : Pour moi il marche.
Marie : Je sais pas, explique-lui, B31.
Élève A32 : Non mais j'ai compris. Elle a donné un contre-exemple qui prouvait que la formule de C41 était fausse et vous avez dit que le contre-exemple ne marchait pas.
Marie : Non, j'ai repris… Je n'ai pas compris ce qu'elle disait.
Élève A32 : Ah d'accord.
Marie : Comme elle a dit « ça marche pas », j'ai compris que ce qu'elle était en train de nous dire ne marchait pas. Et donc, ça te convainc son argument ? Parce que moi, j'ai rien compris.
[…]
Élève A52 : Mais pourquoi ça marche pas ?
Élève A32 : Réexplique ton…
Marie : Explique à tout le monde.
Élève B31 : Karine, elle arrive à… au bout de 15 min, donc elle arrive bien dans l'intervalle [0 ; 0,25]. Et si on dit que Olivier il arrive au bout de 20 min, ça veut dire que le temps d'attente il sera pas…
Marie : Il est de 5 minutes. Ah non le temps d'attente tu…
Élève B31 : Le temps d'attente quand on fait sur… Enfin en tout cas, il sera… la probabilité elle sera pas égale à 0,25.
Marie : B32 ?
Élève B32 : Mais moi je croyais que $P(X \in [a\,;b])$, c'était… X le temps d'attente.
Marie : Oui, c'est ça.
Élève B32 : Mais, dans ce cas-là… l'hypothèse que Karine arrive dans l'intervalle [0 ; 0,25] ne rentre pas dans la formule. Enfin, j'ai pas compris l'exemple de B31.
Marie : Moi non plus.
[…]
Marie : On va essayer d'avancer. […] [54:10] Qu'est-ce qu'on peut faire puisque on a une proposition […]. Que peut-on faire… pour avoir une idée de de cette valeur… de ça ? Qu'est-ce qu'on pourrait faire ?
Élève B22 : On fait l'expérience.
Marie : Une expérience
Élève A32 : Une simulation
Marie : Une simulation. Alors, pour avoir une idée de ça… [*Marie écrit au tableau en même temps*]. On va peut-être valider la proposition de votre camarade ou pas. On peut faire une expérience…
Marie : Alors, on peut faire une expérience… Qu'est-ce qu'il faudrait qu'on fasse ? B22 ?
Élève B22 : Un tirage de…
[0:55] Marie : Ça c'est une expérience ? D'accord. Oui. Alors, ou une simulation alors. Quelqu'un l'a dit.

Pour avoir
une idée de $P(X \in [0; 0{,}25])$
on peut faire une exp
ou une simulation

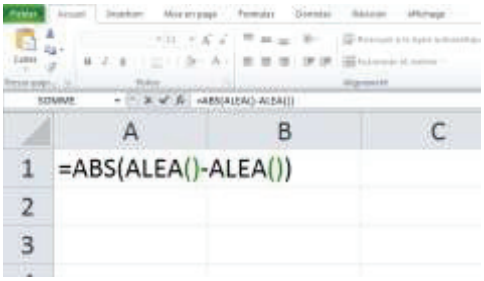


Marie : B53, qu'est-ce que tu proposes ?
[…]
Élève B53 : C'est pour faire un algorithme.
Marie : On va faire un algorithme.
Élève B53 : Pour faire une simulation.
Marie : Pour faire la simulation. D'accord. [55:10] Alors, je vous laisse 5 minutes. Est-ce que vous pouvez me proposer une simulation… du temps d'attente ?
[…]

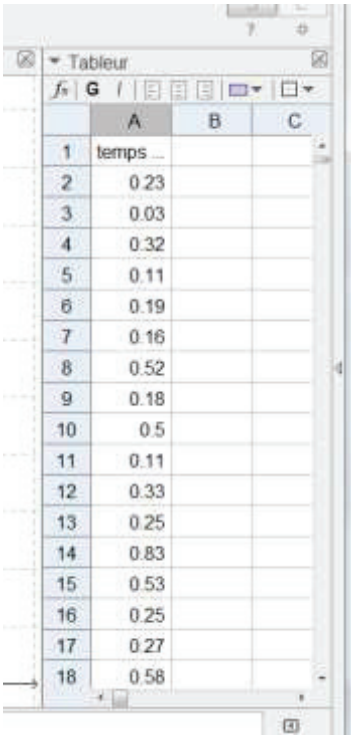
Marie : Mettez sur votre feuille. Qu'est-ce que vous proposez pour simuler…

Marie passe dans les rangs et discute avec certains élèves.
Certains élèves essaient d'implémenter un algorithme sur leur calculatrice.
[…]

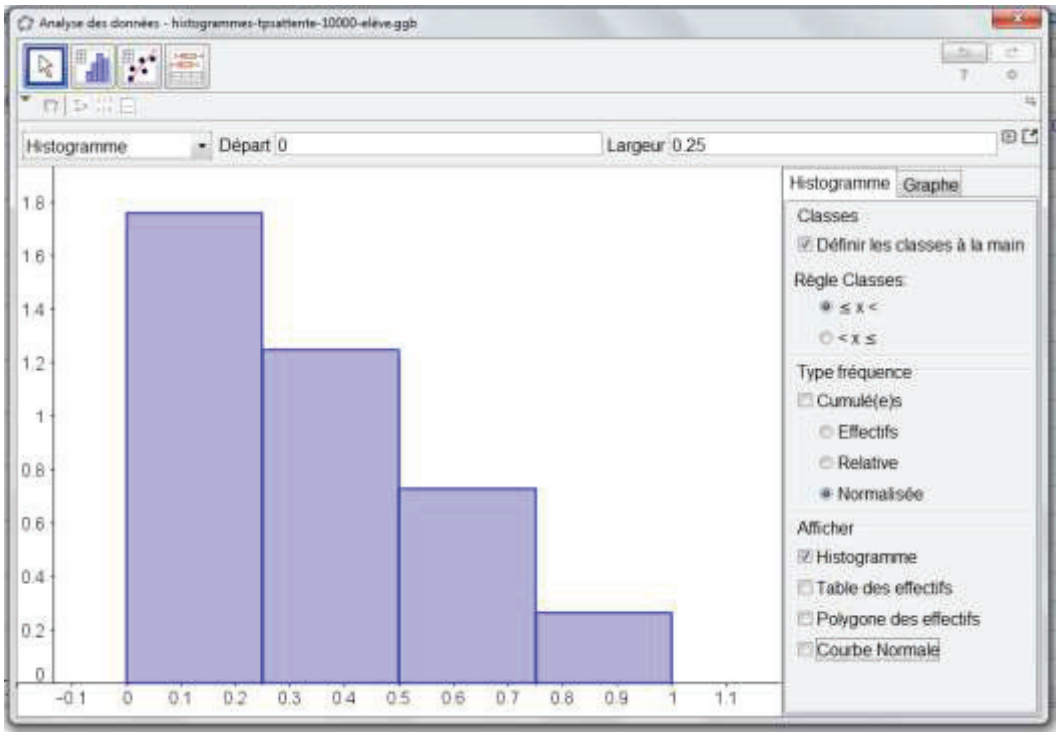
[1:02:31] Marie (à la classe) : Alors, on va reprendre. Je pense que tout le monde a eu des idées. Alors, qu'est-ce qu'on propose comme simulation ? Allez, allez, allez… Vous avez eu plein d'idées. Quand je suis passée dans les rangs, il y avait pleins de choses qui étaient en train d'être faîtes. Là, il y a plus rien. Allez. Qui prend la parole ? Qu'est-ce que vous voulez proposer ? […] Alors, B31 ?
Élève B31 : Ah non, en fait j'ai un autre contre-exemple pour…
Marie : Tu le gardes et tu nous le redonneras après. Qu'est-ce que vous nous proposez comme simulation ? Comment on s'y prend ?
[…]
Marie : B31 ?
Élève B31 : On peut faire des tirages sur Excel et on voit…
Marie : D'accord. Vous notez ? Tirage sur Excel. Alors tu vas me dire… tu vas me dire… [*Marie ouvre le logiciel en même temps*]. Tu vas me dire… Ça vient, ne vous inquiétez pas. Alors, qu'est-ce que je mets comme formule ?
Élève B31 : On met ALEA.
Marie : On met…
Élève B31 : ALEA 0-1.
Marie : [*En écrivant sur l'ordinateur en même temps*] ALEA… alors c'est celui-ci ALEA 0-1.
Élève B31 : 0-1… Donc on fait un grand nombre de tirages et après on compte le nombre de chiffres qui sortent entre…
Marie : Alors on va, d'accord. Alors est-ce que ça, ça simule l'expérience du temps d'attente du premier arrivé ?
Plusieurs élèves disent « non ».
Élève B31 : Non, puisque je ne suis pas d'accord avec ça mais...
Marie : Ah « non », donc tu… donc ça, c'est pas comme ça qu'il faudrait le simuler, le temps d'attente. Ça, ça simule quoi B53 ? Zéro… ?
Élève B53 : Quand est-ce qu'elle arrive.
Marie : Le… le temps d'arrivée de l'un des personnages. On a dit que X ça se calculait… Oui, justement, on va faire X, nous, parce que si on fait T_K, on va tomber là-dessus, ça c'est sûr. Donc comment est-ce qu'on met la formule là pour simuler le temps d'attente ? Vous l'avez écrit la formule tout à l'heure. C22 nous l'a donnée. […] Oui, $T_K - T_O$, en valeur absolue. Donc qu'est-ce qu'on met comme formule ? Il faut calculer la valeur absolue.
Élève ? : de ALEA machin…
Marie : [*Marie tape la formule à l'ordinateur en même temps*]. De ALEA « machin », c'est ça ?
Élève ? : Moins ALEA…
Marie : Moins ALEA pour… C'est raté je crois. Encore. Machin. Ça, ça fait quoi ? C'est bon ? Tout le monde est d'accord ou pas ?



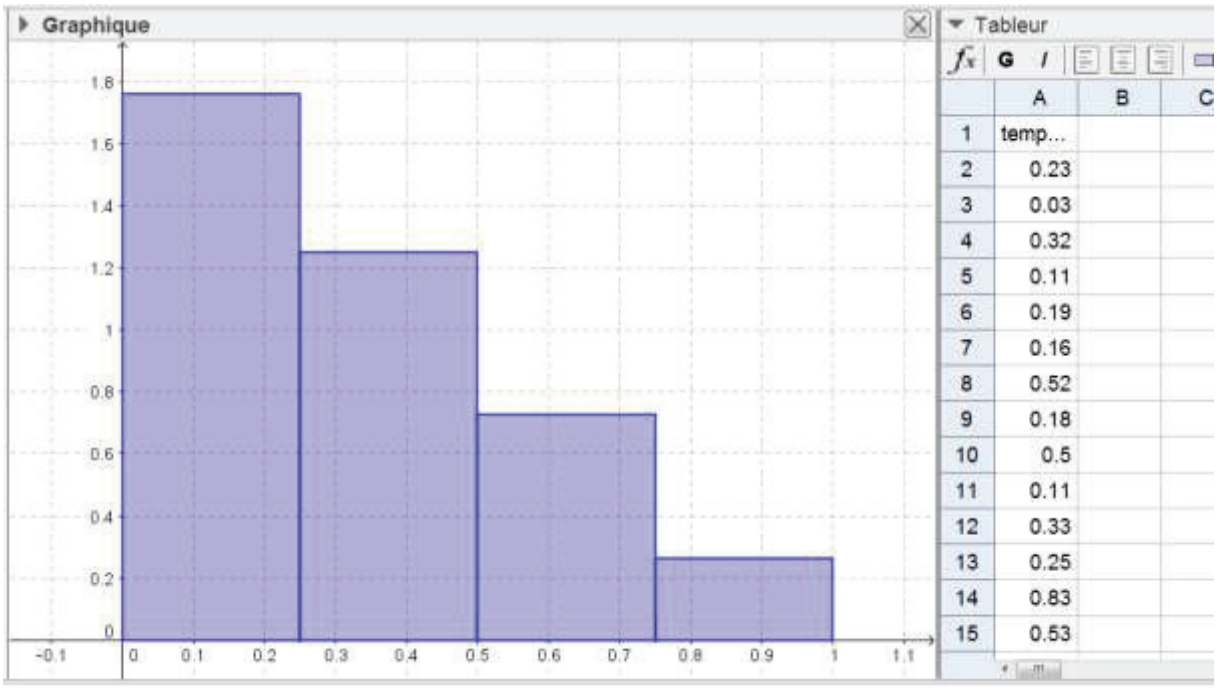
Une élève répond « oui ».
Marie : Oui ? Bon et après bien sûr on va simuler pleins… Bon là je le fais juste un petit peu. [*Marie déroule la formule*]. D'accord ? On va simuler pleins de fois. Un certain nombre de fois. [1:05:19] Et on va compter… tous les résultats entre 0 et 0,25, par exemple. D'accord. Alors, donc… On pourrait faire quoi avec toutes ces données ? Cette distribution de fréquences, là ?
Élève B32 : Un histogramme.
Marie : Oui, B32 ?
Élève B32 : Un histogramme.
Marie : Un histogramme. D'accord ? Ça sera plus simple que de compter avec Excel, parce qu'il faut le… c'est un peu plus compliqué. Alors, je vais pas prendre ces données-là parce que c'est long de les recharger sur le… [*Marie ouvre un autre fichier et prend des données qu'elles avaient simulées en amont de la séance*]. Vous avez noté la formule ?
Plusieurs élèves répondent « non ».
Marie : Alors. C'était valeur absolue enfin ABSOLU de ALEA moins ALEA. Alors là… [*La fenêtre de GeoGebra est ouverte*].



Marie : J'ai… j'ai fait ce que tu as dit B31, j'ai… C'est-à-dire, dans cette colonne-là, j'ai recopié le tableur… la simulation sur tableur pour 10 000. D'accord ? Je l'ai fait tourner… j'ai simulé 10 000 fois. C'est quand même déjà relativement important. Alors on va tracer un histogramme. Alors, GeoGebra le fait pour nous. […]. [*Marie manipule le logiciel*]. Alors, on va lui… on va définir les classes. J'ai déjà mis. J'avais mis ça.



Marie : Alors du coup je lui demande de faire un histogramme et du coup je choisis… les classes et il a compté dans les 10 000 données, il a compté combien il y avait de données entre 0 et 0,25, puis entre 0,25 et 0,5, puis entre 0,5 et 0,75 et 0,75 et 1. D'accord ? Donc il a compté pour nous. Il a dessiné directement… les rectangles. Ok ? Donc, on va aller le mettre sur… le graphique. [*L'histogramme apparaît dans la fenêtre graphique*]. Donc voilà ce que ça nous donne. [1:07:25] Qu'est-ce qu'on pourrait faire ?

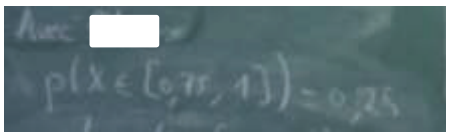


Marie : Alors, déjà rien qu'avec ça, peut-être que… on peut… rediscuter de la proposition de C41 ? B32 ?

Élève B32 : Si on prenait la… le modèle de C41, on pourrait aussi faire l'intervalle… $p(X \in [0{,}25\,;\,0{,}5])$ et…

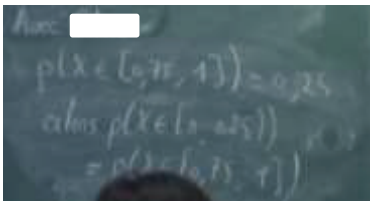
Marie : Oui alors on peut peut-être aller carrément… Tu vois ? Sur le dernier parce que c'est encore plus peut-être… Tu vois ce que je veux dire ?
Élève B32 : Bon alors $[0{,}75\ ;\ 1]$.
Marie : Oui, voilà. Donc si on utilisait la formule de C41, normalement avec sa formule, on trouverait quoi là ? [*Marie écrit au tableau en même temps*].
Plusieurs élèves répondent « 0,25 ».
Marie : 0,25. D'accord ? 0,25. Alors pardon. Avec le modèle de C41, donc avec C41 [*Marie écrit au tableau en même temps*]. Vous êtes d'accord que ça ferait ça, puisque c'est la même largueur d'intervalle ?

Avec C41 :
$p(X \in [0{,}75; 1]) = 0{,}25$



Marie : Donc, qu'est-ce qu'on en pense ?
[…]
Marie : Qu'est-ce qu'on en pense ? Ça semble pas… Alors, on aura que la probabilité que le temps d'attente soit entre 0 et 0,25 serait égale enfin est égale à la probabilité que X appartienne à l'intervalle entre 3/4h et 1h. Qu'est-ce qu'on en pense au vue de ce premier histogramme, de cette première simulation ? C41, qu'est-ce que tu en penses ?

Avec C41 :
$p(X \in [0{,}75; 1]) = 0{,}25$
alors $p(X \in [0; 0{,}25]) = p(X \in [0{,}75; 1])$



Élève C41 : C'est faux.
Marie : C'est faux. D'accord ? Ça n'a pas l'air juste en tout cas.
Marie : Donc, ne correspond pas aux résultats de la simulation. [*Marie écrit au tableau*].

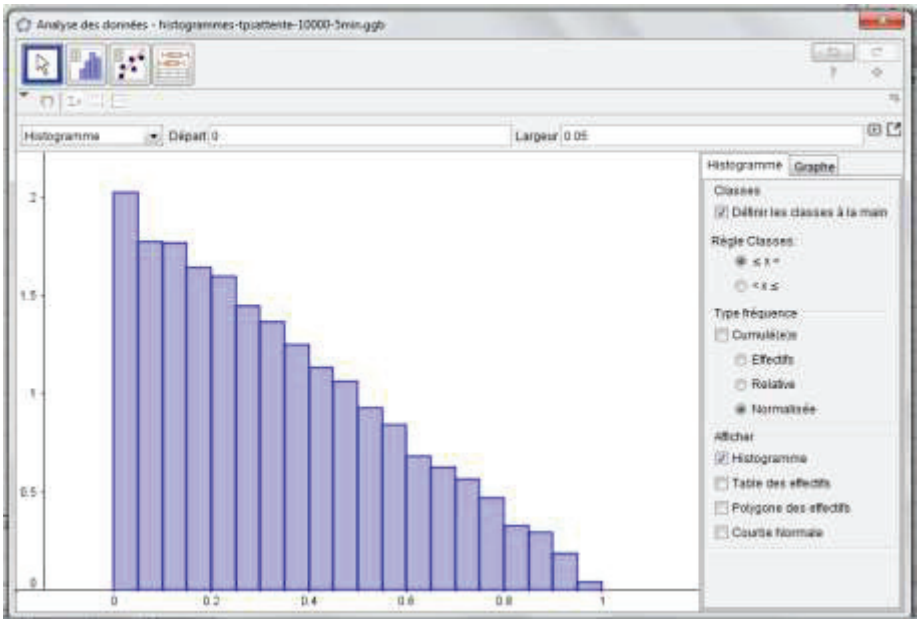
Ne correspond pas aux résultats de la simul

[1:09:28] Marie : Donc visiblement, c'est pas comme ça qu'on va calculer la probabilité que X appartienne à un intervalle $[a\ ;\ b]$. D'accord ? Notre question c'est toujours ça finalement. C'est réussir à calculer la probabilité que X appartienne à un intervalle. [*Marie écrit au tableau en même temps*]. Alors évidemment l'intervalle $[a\ ;\ b]$ étant inclus dans l'intervalle [0 ; 1] parce que sinon ça n'a pas d'intérêt.

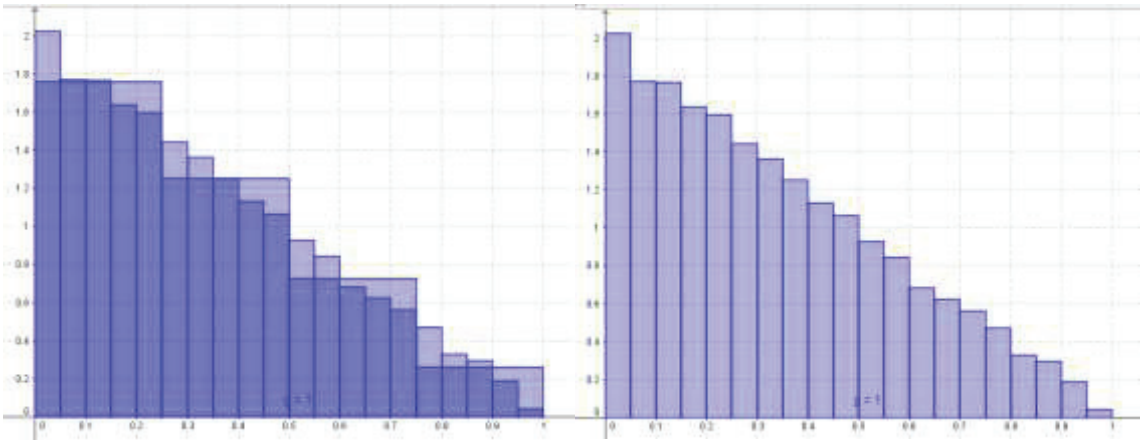
$Q : p(X \in [a; b])$ *où* $[a; b] \subset [0; 1]$

[1:09:52] Marie : Qu'est-ce que… puisqu'on a une jolie petite machine, qu'est-ce qu'on pourrait lui demander de faire ?
Marie : A21 ?
Élève A21 : De fragmenter encore.
Marie : De fragmenter ? Qu'est-ce que tu veux dire ?
Élève A21 : De prendre des intervalles plus petits.

Marie : Alors des intervalles… Tu veux dire quoi ? Si, si mais utilise le vocabulaire que tu connais quand même, A21. C'est un histogramme. J'imagine que les intervalles, c'est ça, là ? [*Marie montre la base d'un rectangle de l'histogramme vidéoprojeté*].
Élève A21 : Oui.
Marie : De prendre des amplitudes, des classes plus petites. On y va ? Alors. Donc du coup, du coup on va lui changer ça. Alors qu'est-ce que vous voulez comme largeur ?
Élève C41 : 0,1.
Élève B32 : 0,05.
Marie : Zéro virgule ?
Élève B32 : 0,05.
Marie : C'est quoi 0,05 ?
Élève B32 : C'est dans… l'intervalle.
Marie : Oui.
Élève B32 : Ah, c'est…
Marie : Mais… Ça correspond à un temps ou… ? Parce que là c'était 15 minutes.
Élève B32 : …
Marie : Tu veux qu'on fasse… Bon. 0,05. Moi je veux bien. [*Marie manipule le logiciel en même temps*]. C'est beau.



Élève B22 : 0,05 heure.
Marie : J'enlève l'autre. D'accord ? Pour garder seulement celui-ci.



Marie : Donc là, on a fragmenté comme disait A21. Donc, on a toujours notre série de données, 10 000 simulations du temps d'attente et on a compté, on lui a fait compter cette fois-ci dans des classes de largeur 0,05. Alors… Bon, il devient de plus en plus clair que la proposition de C41… Ça a pas l'air d'être quand même ça. [1:11:28] Est-ce que vous avez… des propositions ?

[1:11:44] Marie : Si je vous demandais de… avec ça, de… d'évaluer le… la probabilité que le temps d'attente soit entre 0,2 et 0,3. D'accord ? Imaginons. Comment vous feriez ?

Marie : Est-ce que c'est bien ça qu'on essaie d'évaluer ? La probabilité que le temps d'attente soit entre tant et tant ? Donc si je vous demandais ça… Si je vous posais cette question « évaluer le temps d'attente entre 0,2 et 0,3 », comment vous feriez avec cet… à l'aide de cet histogramme ?

Marie : […] A52 ?

Élève A52 : Entre 0,2 et 0,3 ?

Marie : Oui. Juste, comment tu ferais, pas ce que tu trouverais. Mais comment tu ferais ?

Élève A52 : On peut faire comme dans le DM, c'est-à-dire… on peut tracer une courbe de tendance. On va pouvoir calculer l'aire de ce triangle… et puis prendre une amplitude qui correspond à 0,2 ~ 0,3.

Marie : D'accord. Est-ce que C22 tu as entendu la proposition de A52 ?

Élève C22 : Un peu.

Marie : D'accord. Alors est-ce que tu pourrais nous la faire… à haute et claire voix ?

Élève A52 : On pourrait tracer la courbe de tendance.

Marie : Donc on pourrait essayer de lisser notre histogramme comme on l'a fait dans notre… dans le devoir ?

Élève A52 : Oui.

Marie : D'accord.

Élève A52 : Puis, calculer l'aire de ce triangle-là.

Marie : Triangle. Quel triangle ?

Élève A52 : Enfin, ce qui pourrait ressembler à un triangle.

Marie : Ah oui, tout tout le… J'avais pas compris. D'accord.

Élève A52 : C'est égal à 1.

Marie : Ah oui, alors, on n'a pas besoin de le calculer puisque… ça fait 1. Pourquoi ça ferait 1 ?

Élève A52 : Parce que… ça indique la fréquence.

Marie : D'accord.

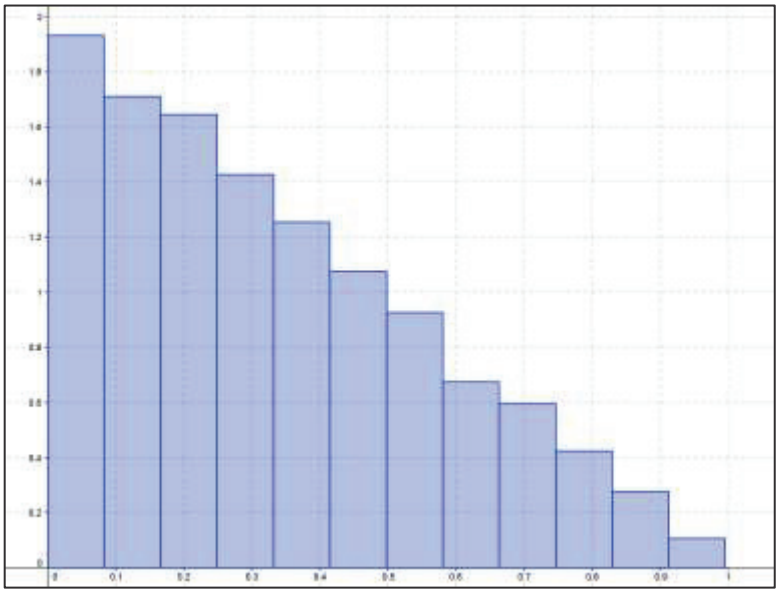
Élève A52 : Et donc après on peut prendre juste cette partie-là. Et puisqu'on a… la valeur, grâce à la courbe de tendance, qui correspond à l'ordonnée, on obtient donc une hauteur, faut la multiplier par l'amplitude de l'intervalle. Donc ça donne…

Marie : Est-ce que vous êtes d'accord avec la proposition de A52 ? Ceux qui levaient la main ? B31, tu voulais dire quoi ?

Élève A52 : Faut qu'on trace un trapèze, parce que…

Marie : Oui… Ça dépend de votre courbe de tendance, jeunes gens. Donc tu la vois comment cette courbe de tendance, toi ? Quand tu dis ça…

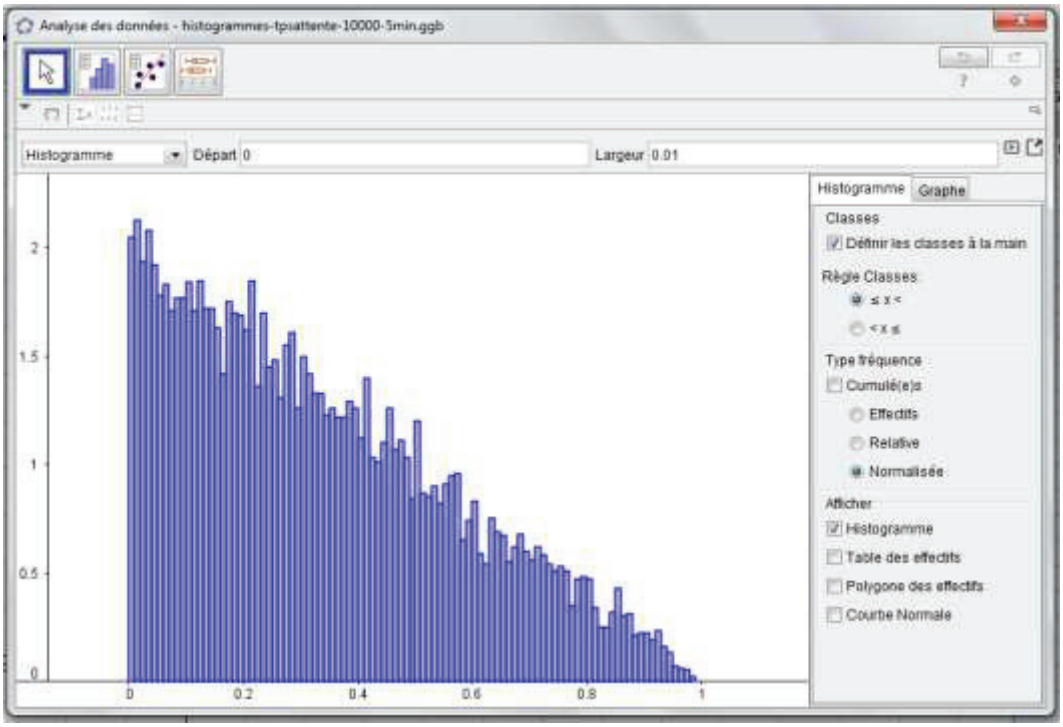
Élève A52 : Ça va être une affine.
Marie : Une droite ?
Élève A52 : Une droite.
Marie : Oui. B31, tu levais le doigt pour proposer de… de calculer… d'évaluer… C'était ça ? La même chose ? D'accord. A32 ?
Élève A32 : En fait, je comprends pas ce qu'il y a en ordonnées.
Marie : Alors. Tu comprends pas ce qu'il y a en ordonnées.
Élève A52 : C'est la densité.
Élève A32 : Ça va de zéro… Ah oui, c'est vrai.
[1:14:29] Marie : Ah oui, c'est vrai. Oui. D'accord ? Alors, on fait ce qu'a proposé A52 ? […].
Marie : Alors, vous notez ce qu'a proposé A52. A52, tu peux venir… tu veux venir tracer le… la courbe que tu nous proposes ?
Élève A52 : Je sais pas faire mais je vais essayer.
Marie : Dis… si tu fais… tu fais l'onglet droit…
Élève A52 : Non, je sais pas utiliser ça.
Marie : Attends… Si tu fais l'onglet droit et puis tu positionnes deux points pour tracer ta droite. Montre-moi à la main. Tu la vois comment alors cette droite ?
Élève A52 : Elle fait un truc comme ça. [*L'élève est au tableau, devant l'histogramme, et « trace » avec le doigt une droite le long du haut des rectangles de l'histogramme*].
Marie : D'accord. Ok. Alors… […]. Attendez, normalement, j'ai dû vous préparer un petit document. [*Marie cherche dans ses affaires*]. Tenez, là y a un histogramme pour des amplitudes de 5 minutes. Donc il est moins… les classes sont plus grandes que ce qu'a choisi… nous a demandé de prendre B22, d'accord ? [1:15:40] Alors là-dessus, tracez votre… Tiens. Là. Tracez… votre courbe. Vous avez noté ce qu'a dit A52 ? […]. Ce qu'il a proposé c'est quoi en fait ? On résume. Qu'est-ce qu'il a proposé ? B21 ?
Élève B21 : Tracer une courbe de tendance.
Marie : Donc, tracer une courbe de tendance. Donc ici, il pense c'est une fonction affine.
Élève B21 : On calcule l'aire.
Marie : Et de calculer l'aire.
Élève B21 : Et ensuite on calcule l'amplitude de ce qui était demandé.
Marie : Non, enfin… il s'est aperçu en le disant que l'aire sous la courbe ça fera forcément 1 parce que… c'est la somme des fréquences. Et donc il dit « et bien, faut calculer l'aire de la partie entre zéro… », je sais plus ce que j'ai dit. 0,2 et 0,3. Donc sur ce… histogramme, tracez votre courbe de tendance.
[1:16:28] *Distribution du polycopié ci-dessous dans la classe. Recherche individuelle ou par deux.*



[1:17:26] Marie : Alors, qu'est-ce que vous avez tracé ?
Marie : D'accord. Donc toi tu la vois comme ça. Tu l'as pas tracée ta courbe de… Alors, toi, tu aurais fait l'aire des deux rectangles. Toi, tu l'as pas tracé ta courbe de tendance ? Qu'est-ce qu'il a proposé A52 ? Vous avez pas compris ce qu'il a proposé, A52 ?
[Inaudible]
Marie : Oui, donc il a dit qu'il va essayer de tracer une droite un peu comme ça. Alors. Qui c'est qui vient tracer la… On la trace comme ça ?
Élève B22 : A main levée.
Marie : A main levée ? Peut-être pas à main levée… […]. B22 ?
Élève B22 : Y a pas une fonc… enfin un truc qui donne la ligne de tendance directement ?
Marie : Un truc qui donnerait la courbe… Dans le logiciel ? Si tu crois que je vais l'utiliser… Alors ? Peut-être, je ne sais pas, je suis pas au courant. Alors ? Donc, A52 il dit « on trace une courbe de tendance ». [*Marie écrit au tableau en même temps*].

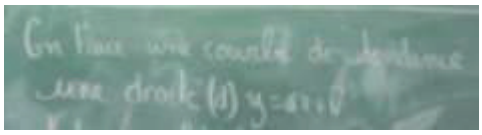
On trace une courbe de tendance

Marie : Vous l'avez tracée sur votre dessin ?
Plusieurs élèves : Oui.
Marie : Alors, à votre avis, là, vous êtes trente je sais pas combien aujourd'hui, vous avez tous la même ou pas, là ?
Plusieurs élèves : Non.
Marie : Non, parce que vous l'avez faite au jugé.
[…]
Marie : Non, là c'est parce que je trace au jugé une droite qui avait l'air… de passer par où quand même ? Enfin, vous me l'avez pas tracée là, la droite ? [*Marie montre une droite qui ne suit pas l'allure du haut des rectangles de l'histogramme*].
[…]
Marie : Non. On a essayé de… Si je réduis encore les classes, qu'est-ce qui va se passer ? Donc là, on avait pris quoi ? Tu avais dit 0,05. Oui, si on prend un histogramme qu'est encore plus… avec des amplitudes plus petites… Donc… [*Marie manipule le logiciel*]. Non, c'est trop moche. Non, ça m'arrange pas ça.



Marie : Ça va pas nous aider, je sens ça.
Bug du logiciel.
[...]
Marie : Alors, on ne va plus y toucher. Je ne vais pas essayer de modifier les classes. Alors, peut-être qu'on peut pas tous se retrouver avec une courbe de tendance différente quand même ? Donc ce serait bien qu'on se mette d'accord. Alors, on est quand même d'accord sur l'allure de cette courbe qu'on cherche. Est-ce qu'on est d'accord ? [...]. Qu'est-ce qu'on cherche comme genre de fonction ? B53 ?
Élève B53 : En fait, pour tracer faut qu'on essaye... de faire en sorte que... notre courbe quand on la trace qu'il y ait une partie qui soit en haut enfin...
Marie : Oui. A l'œil, je suis d'accord avec toi. On va essayer de de (sic) faire un truc...
Élève B53 : Une moyenne.
Marie : ...qui passe à peu près voilà au milieu des hauts des rectangles. Qu'il y en ait autant au-dessus qu'en-dessous de vide. Je suis d'accord mais t'es d'accord que même en faisant comme ça au mieux, chacun d'entre vous, on va se retrouver avec des droites tous différentes. Donc est-ce qu'on pourrait pas se... Comme disaient tes camarades, faudrait qu'on se mette d'accord sur cette droite quand même. J'ai... je demandais juste si vous êtes d'accord qu'on cherche une droite ?
Plusieurs élèves : Oui.
Élève B52 : Faut qu'elle passe par deux points. Qu'on définisse des points.
Marie : Alors. Donc, on cherche une courbe de tendance. Une droite. [*Marie écrit au tableau*]. Alors une droite, donc qu'est-ce qu'on cherche donc comme genre de choses ?
[Inaudible].
Marie : Oui. Donc c'est-à-dire ?
Élève ? : Une équation de droite.
Marie : Oui, donc c'est quoi ?
Plusieurs élèves : $ax + b$.
Marie : Non.
[...]
Marie : Non, non, non, une droite, c'est pas $ax + b$. B22 ?
Élève B22 : $y = ax + b$.
Marie : Merci. $y = ax + b$. [*Marie écrit au tableau*].

On trace une courbe de tendance une droite $y = ax + b$



[...]

Marie : D'accord. On cherche ça. Alors je vous laisse 5 minutes. [1:21:17] Tout le monde, trouvez-moi l'équation de cette droite. Trouvez-moi les conditions. On est d'accord qu'il faut trouver a et b.

[...]

[1:21:30] *Recherche individuelle.*

[1:22:06] Marie : Alors j'en trace une, moi, à l'écran un peu au jugé. Bon alors, je sais pas si je vais y arriver au jugé, moi. Bon alors je sais pas… C'est pas très facile. Bon, allez, je me lance. [*Marie manipule le logiciel en même temps. Photo du tableau non disponible*].

Élève ? : Pourquoi vous la faites pas passer par 2 et 1 ? Enfin… zéro, (0 ; 2)…

Marie : Attends… […] Bon, ça va pas être top. Allez, adjugé ?

Élève A22 : Madame, on peut la faire passer par des points.

Marie : On peut la faire passer par des points. Alors, est-ce qu'il y a des points dont on est sûr ?

Plusieurs élèves parlent en même temps.

Élève A22 : Par exemple (0 ; 2).

Marie : Si c'est des points où on n'en est pas sûrs… Pourquoi (0 ; 2) ?

Élève A22 : Parce que ça a l'air arbitrairement…

Marie : Ça a l'air… Non, non, non, moi je veux…

Élève A22 : Arbitrairement.

Marie : Non, si on veut trouver vraiment l'équation de la droite, il nous faut des vrais arg… Il nous faut des choses objectives.

Élève A22 : Ça le fait plutôt sur le dessin.

Marie : « Ça le fait plutôt ». Est-ce que ça, A22, c'est un argument « Ça le fait plutôt » ?

Légers rires.

Marie : Alors, est-ce qu'il y a… est-ce qu'il y a des points…

Élève A22 : En physique, ça serait passé.

Marie : Ne me tends pas de perche, A22. Alors, A21 ?

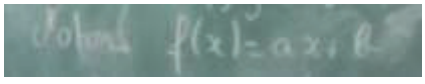
Élève A21 : (1 ; 0).

Marie : (1 ; 0). A21 nous dit « cette droite, elle passe forcément… ». Alors, on va peut-être lui donner un petit nom parce que… On va appeler $f(x)$ la fonction qu'on cherche, ça sera plus simple.

[...]

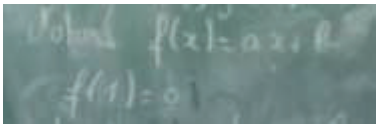
Marie écrit au tableau.

$\mathcal{N}otons\ f(x) = ax + b$



Marie : Qu'est-ce que… qui… A22, tu dis « en 1, ça vaut 0 ».

$\mathcal{N}otons\ f(x) = ax + b$
$f(0) = 1$

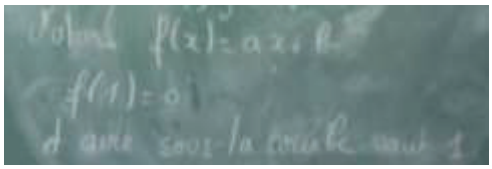


Élève A22 : Oui.

Marie : A22, il nous dit ça. Est-ce que vos… Pourquoi tu nous dis ça ?

Élève A22 : Parce qu'en 1, ça veut dire qu'il est 8h, donc on peut plus attendre.
Marie : D'accord.
Élève ? : Quoi ?
Marie : En 1, ça veut dire qu'il est 8h. Donc le temps d'attente…
Une autre élève : Ah oui.
Marie : Alors, on a une valeur. Sauf que… pour trouver a et b, il faut quoi ?
Élève A52 : Une deuxième.
Marie : Il faut deux renseignements, je dirais.
Élève A52 : Ou deux points.
Marie : […] Allez, il nous faut un autre renseignement pour retrouver a et b.
Élève ? : Soit il nous faut un autre point, soit il nous faut un nombre.
Marie : Alors, est-ce qu'on peut… ? Et pas « ça le fait », A22.
Marie : On n'a pas un autre renseignement ?
[…]
[1:24:41] Marie : C'est mort si on trouve pas un autre truc. On trouvera pas a et b. Bon je vais donner la parole à B31 parce qu'elle lève la main depuis 5 minutes.
Élève B31 : En fait, il faut… C'est le point (0 ; 2) parce que… on trace en fait le rectangle qui a une aire de 2 et on trace la diagonale du rectangle et ça…
Marie : Donc c'est quel renseignement que t'utilises sans le dire ?
Élève B31 : L'aire du rectangle.
Marie : Non. Dans l'autre sens, B31. Tu sais que l'aire sous la courbe doit valoir…
Élève B31 : 1.
Marie : Tu sais que l'aire sous la courbe doit valoir 1.
Élève A52 : Ah oui.
Marie : Donc on sait que l'aire… Alors ici, ça fait un triangle donc on sait aussi calculer l'aire d'un triangle rectangle… Donc on sait que l'aire, là, ça doit valoir 1. Alors, avec ces deux renseignements, tout le monde me trouve a et b, s'il vous plait.
Marie : [*en écrivant au tableau*]. Et l'aire sous la courbe vaut 1.

Notons $f(x) = ax + b$
$f(0) = 1$
et aire sous la courbe vaut 1



[1:25:46] Marie : Alors, allez-y. Tout le monde me trouve l'équation… enfin l'expression de ma fonction.
Marie : Allez, allez, tout le monde s'y colle. Pour de vrai cette fois-ci, on vous donnera pas la réponse, je vous préviens.
Marie (à un élève) : Du coup tu la traces en rouge sur le… document que… ton histogramme.
Marie (à la classe) : Quand vous l'avez votre équation de droite, vous la tracez en couleur sur votre histogramme.

Phase de recherche individuelle. Marie passe dans les rangs.

Marie : Moi je veux les calculs maintenant. Parce que B31 vous l'a donné mais moi je veux les calculs.

Marie discute avec les élèves dans les rangs.

[1:28:56] Marie : Est-ce que quelqu'un pourrait rappeler la formule de l'aire… […] De l'aire d'un triangle ? C31 ?
Élève C31 : Base fois hauteur sur 2.
Marie : Oui, base fois hauteur divisé par 2. Puis c'est un triangle rectangle alors…

[...]
Marie : Alors, on attend que tout le monde ait trouvé et en attendant, je vais donner la suite aux autres.
Marie distribue le polycopié avec les questions.

Exercice 1 : En utilisant le modèle mis en place, répondez aux questions suivantes.

a) Quelle est la probabilité que le premier arrivé attende exactement 40 minutes ?
b) Quelle est la probabilité que le premier arrivé attende plus de 40 minutes ?
c) Quelle est la probabilité que le premier arrivé attende moins de 40 minutes ?
d) Quelle est la probabilité que le premier arrivé attende entre 20 et 40 minutes ?
e) Quelle est la probabilité que Karine et Olivier arrivent en même temps ?
f) Finalement, ils ont décidé de ne pas attendre plus d'un quart d'heure. Quelle est la probabilité qu'ils se retrouvent ?

[...]
[1:29:41] Marie : Donc une fois que vous avez trouvé votre équation de droite… Attends. Vous commencez à répondre aux questions de l'exercice. D'accord ? Et je veux voir l'équation de droite sur toutes les copies, là.
Marie passe dans les rangs, vérifie si les élèves ont trouvé l'équation de la droite, et pour ceux qui ont terminé, elle leur distribue le polycopié avec les questions pour qu'ils y répondent, après avoir tracé en couleur la courbe.

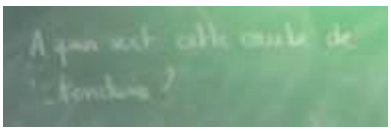
[1:31:44] Marie : Alors qui a pas eu l'énoncé du petit exercice à faire… à la suite, une fois qu'on a trouvé notre… notre courbe de tendance.
[...]
Marie distribue le polycopié pour les derniers élèves et vérifie leurs équations de droite.
Certains élèves ont des problèmes pour trouver l'équation au niveau de la méthode. Certains font des erreurs de calcul.

Marie (à l'élève B21) : Alors est-ce que tu as l'équation de droites, B21 ? [...] Tu l'as tracée ? Bien. Alors maintenant, comment grâce à cette courbe, tu vas répondre ?
Élève B21 : On peut pas faire par lecture graphique ? Comme au DM, on peut faire lecture graphique et ensuite mathématiquement ?
Marie : Non, fais directement avec les calculs parce que sinon… Utilise la courbe, qu'est-ce qu'il a dit, A52 ? On calcule l'aire…
Élève B21 : Que l'aire était égale à 1.
Marie : Non, on calcule… par exemple, quand on te demande la probabilité je sais pas que le temps d'attente… Vous allez en faire une autre peut-être sur votre cahier, ça sera plus pratique. Si je te demande de calculer maintenant la probabilité que X appartienne à l'intervalle [0 ; 0,25], comment tu vas faire ?
Élève B21 : Je vais additionner les…
Marie : Non, non. Pourquoi tu te sers pas de la courbe de tendance ?
Élève B21 : Ah, faut se servir de la courbe ?
Marie : Oui, ça sert à quoi qu'on l'ait fait cette courbe ? Parce que ces histogrammes, si je change l'amplitude, si je change les données, si je relance mon tableur, ça donnera un autre histogramme.
Élève B21 : Oui.
Marie : T'es d'accord ? Si j'utilise l'histogramme pour faire les calculs à chaque fois, j'obtiendrais des valeurs différentes. Donc c'est… c'est pas possible. D'accord ?
Élève B21 : Oui.
Marie : Cette courbe... Par contre, tous les histogrammes, ils ont tous la même tête. Donc c'est toujours cette même tête. Donc l'aire si tu choisis entre 0 et 0,25, tu vas dire quoi ? Que la probabilité, c'est quoi ? Qu'est-ce que tu aurais fait avec les histogrammes ?

Élève B21 : J'aurais calculé l'amplitude de chacun, je les aurais additionnées…
Marie : Ah bon, c'est comme ça que tu vas calculer la fréquence de chacun ?
Élève B21 : Et ensuite, j'aurais divisé par…
Marie : Non, c'est pas ça. Qu'est-ce qu'on a fait quand on a fait les histogrammes ?

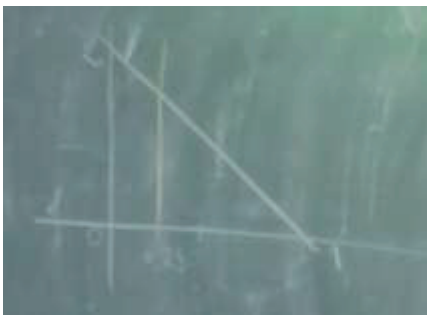
[1:34:56] Marie : Alors, juste, est-ce que tout le monde a enfin l'équation de droite quand même là ? Proprement ? Donc, B43, ça donne quoi ?
Élève B43 : $-2x + 2$.
Marie : Non.
Élève B43 : $y = -2x + 2$.
Marie : $y = -2x + 2$. On est d'accord ? Tout le monde l'a tracée sur son histogramme ? En couleur. [1:35:13] Ça nous sert à quoi alors du coup ce… d'avoir trouvé cette fonction ? Qu'est-ce qu'on en fait maintenant ? Non, on peut juste maintenant régler le problème de la question qu'on s'était posée, A52. En quoi, cette courbe, elle va nous aider ? Chut. J'aimerais bien que tout le monde soit attentif parce que là c'est important. Donc maintenant qu'on a cette espèce de… cette courbe de tendance. […] Vous êtes d'accord que si j'avais relancé le générateur aléatoire, j'aurais obtenu des histogrammes, d'accord ? Autant que je veux. Ils auraient eu presque la même…
Élève ? : Forme.
Un autre élève : Allure.
Marie : La même allure. La même forme, d'accord. Une espèce de chose comme ça décroissante. Mais si j'utilise les histogrammes pour… Est-ce que je peux utiliser mes histogrammes pour calculer mes probabilités ? En calculant l'aire des rectangles ? En sommant l'aire des rectangles, comme on a fait. B32 ?
Élève B32 : Non, parce que ça changerait à chaque fois.
Marie : Oui. Non, parce que ça change à chaque fois. D'accord ? A chaque échantillon, j'ai un nouvel histogramme, donc du coup, j'obtiens des fréquences observées qui seront différentes. Bon normalement elles seront proches, mais elles seront pas les mêmes. Alors, à quoi nous sert d'avoir trouvé cette droite ? Je suis pas sûre que cette phrase était très française au demeurant… A quoi ça sert alors cette droite ? Qu'est-ce qu'on va en faire, A52, maintenant, de ta droite, là ?
Élève A52 : Maintenant, on a un repère. Par exemple, …
Marie : Alors, on a une droite qui, on pense, modélise tous nos histogrammes en quelque sorte. Un… un bon…
Élève A52 : C'est le modèle théorique.
Marie : Voilà. D'accord. Alors qu'est-ce qu'on va en faire ? [*Marie écrit au tableau*].

A quoi sert cette courbe de tendance ?

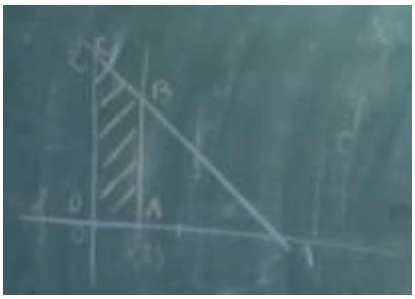


Élève A52 : Pourquoi ? Pour… Dans quel but ? Pour… Par exemple, pour déterminer…
Marie : Oui.
Élève A52 : L'attente de plus de 40 minutes, ou un truc comme ça…
Marie : Parle plus fort et articule s'il te plait, A52.
Élève ? : J'ai pas compris…
Marie : On était en train d'essayer de calculer la probabilité que X soit entre 0 et 0,25.
Élève ? : Ah oui.
Marie : Donc on aurait pu faire la somme des aires des petits rectangles de 0 à 0,25. Maintenant on fera quoi ?
Élève A52 : On trace la droite $x = 0{,}25$.
Marie : D'accord.
Élève A52 : Maintenant, on obtient un petit triangle.
Marie : Donc on va prendre [0 ; 0,25], on va considérer… [*Marie trace une courbe au tableau*].

Marie : Tu as pas… tu as dit qu'on obtient quoi ? Vas-y, excuse-moi, continue.
Élève A52 : On trace la droite $x = 0{,}25$.



Marie : Voilà.
Élève A52 : On obtient un triangle.
Marie : Un triangle ?
Élève ? : A droite.
Marie : Ah, d'accord. Ok.
Élève A52 : Là, c'est un rectangle donc on fait base fois hauteur… divisé par deux.
Marie : Je vois ton problème. Si, si, je vois. Dis-moi, comment, à l'aide du dessin, tu vas calculer la probabilité qu'on souhaitait évaluer ?
Élève A52 : Ah pardon, pardon, pardon. Oui, oui. C'est un trapèze.
Marie : Ah, d'accord. Est-ce que tout le monde est d'accord que du coup…
Élève A52 : Oui, oui.
Marie : … la probabilité que X soit entre 0 et 0,25, on va l'écrire comme ça, c'est égal à quoi ?
Élève A52 : Mais non mais après on peut utiliser celui du triangle parce qu'on peut faire 1 moins…
Marie : Oui bien sûr, il y a pleins de façons de faire, mais enfin en terminale S, l'aire d'un trapèze, peut-être que je peux compter dessus ?
Élève A52 : Oui.
Marie : Oui, d'accord. On va le rappeler. Mais répondez à la question, là. La probabilité que X soit entre 0 et 0,25, ça… ça va être quoi ? […]. A52 ? Tu vas calculer quoi alors ?
Marie : Comment est-ce que tu vas calculer cette probabilité, à l'aide de la courbe de tendance que tu… que tu as déterminée ?
Élève A52 : Avec les… avec… […] Je veux dire, on se base sur l'aire de…
Marie : L'aire ! Voilà, ça sera… c'est l'aire. C'est l'aire de quoi ?
Élève ? : L'aire du trapèze.
Marie : Bon, alors, on va lui donner des… On va peut-être le nommer ce trapèze, ça sera peut-être plus facile. [*Marie écrit au tableau*]. Donc, OABC pour vous… l'aire du trapèze OABC.



$p(X \in [0;0{,}25]) = aire\ du\ trapèze\ OABC$

Marie : Est-ce que tout le monde est d'accord ?
Silence. Une élève répond « oui ».

Marie : Non, tout le monde n'est pas d'accord… Si ? Et ça, ça va nous permettre de calculer la probabilité de tous nos intervalles ? Tout ce qu'on veut pouvoir calculer ? Oui ? Alors, allez-y. Alors, faut que je rappelle l'aire d'un trapèze peut-être ?
Certains élèves répondent « oui ».
Marie : [*en écrivant au tableau*]. On fera comme si c'était pas écrit. D'accord ?
Marie : Bon, je vous rappelle que ça, c'est un trapèze. Oui, enfin bon… C'est un quadrilatère qui possède deux côtés opposés parallèles. Ça, c'est la petite base, ça, c'est la grande base, ça, c'est la hauteur. Donc l'aire d'un trapèze, « petit base » plus « grande base » fois « hauteur », divisé par deux. C'est bon ?

Marie : Alors, sinon, y a d'autres façons de calculer des aires, A52, je crois que ce que t'étais en train de dire c'était le grand triangle moins le petit triangle, par exemple. On peut s'arranger si on… sait pas calculer l'aire d'un trapèze. C11, c'est bon ? Allez-y.

[1:40:25] *Phase de recherche individuelle ou par deux.*
Marie et le chercheur passent dans les rangs.
Plusieurs élèves ont des problèmes pour comprendre pourquoi on se sert de cette courbe et pas de l'histogramme. Marie réexplique individuellement à certains élèves.

[1:44:54] Marie (à la classe) : C'est peut-être plus simple de faire, sur votre cahier quadrillé, votre courbe de tendance… d'accord ? Une droite, celle-là. Et de… qu'est-ce que vous allez faire pour calculer toutes les proba que j'ai demandé ? A22 ?
Élève A22 : Trouver le point sur la… enfin, trouver le point sur la droite, calculer la distance jusque… de ce point… et la distance entre le point qu'on a trouvé et la… l'axe des abscisses et… faire… le calcul… de l'aire…
Marie : Oui, donc on va calculer l'aire, à chaque fois, de l'intervalle qui nous intéresse. Oui.

Marie aide un groupe de deux élèves, qui rencontrent un problème pour écrire 40 minutes en décimal. La première question pose des problèmes à certains. Marie leur conseille alors de passer directement à la deuxième question.

[1:50:00] *Sonnerie.*
Marie : Alors vous notez s'il vous plait pour mardi. Deux choses importantes. […] Pour mardi. Un, vous traiter l'exercice 1. […] Puis, vous allez prendre une feuille simple de cahier que vous allez me couper en deux, d'accord. Vous allez me prendre une feuille de classeur que vous allez couper en deux et sur cette petite partie, vous allez faire un résumé de ce que nous avons fait aujourd'hui. Ah alors là, c'est… Voilà. Et on met son nom sur la petite feuille.
Élève A52 : C'est noté ?
Marie : Ça c'est la question qui fâche. On met son nom et sur une demi-page de classeur, je ne veux pas que ça dépasse, vous me faites une synthèse de ce qu'on a fait aujourd'hui. Je vous rappelle que vous avez pris des notes. C'est bien compris ?
[…]
Marie : Tu prends une feuille. Oui ce qu'on a fait aujourd'hui, pendant deux heures, tu me fais un résumé de ce que tu en as compris, quel était l'objectif… Vous me faites une synthèse.

[...]
Marie : Tu mets ce que tu veux dans ta synthèse, c'est ta synthèse. C'est ce que toi tu as compris de ce qu'on a fait aujourd'hui. Les notes devraient vous aider.

I.2 Transcription de la séance 2

Transcription de la séance 2 du 17/03/15 (55 min – classe entière)

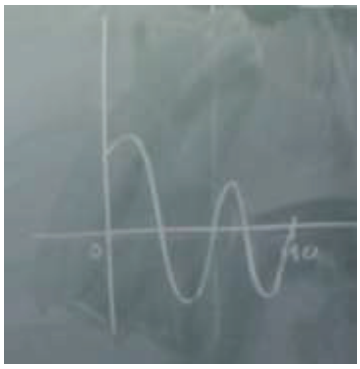
34 élèves présents.

[00:00] [>Marie]: Bien, B52 de ton côté, C41, du tien, veux-tu bien ramasser les synthèses ? En silence. A32.
[00:10] [>Marie]: Alors, je vais juste...
[00:25] [>Marie]: Les gens qui ramassent voulez-vous bien le faire rapidement, les autres donnez juste votre synthèse sans discuter. B22. [...] Oui c'est un brouillon, c'est pas grave B11. [...]
[1:52] [>Marie]: Alors, donc notre petit problème, on va écrire une synthèse. Vous prenez votre cahier d'exercices. Histoire tout de même de mettre au courant vos petits camarades absents samedi. Donc pour synthétiser ce qui a été fait en classe samedi. Alors donc le problème : nos deux amis se rencontrent entre 7h et 8h. Ils arrivent tous les deux à tout moment entre 7h et 8h et on s'est interrogé sur le temps d'attente du premier arrivé. C'est bon tout le monde l'a en tête ? Alors, ce que vous avez proposé successivement. [*Marie écrit au tableau*]. On a noté T indice O [T_O] le moment d'arrivée... Alors pourquoi O ? Ah si, il s'appelle Olivier. Le moment d'arrivée d'Olivier. Et idem le moment de K, le moment d'arrivée de Karine. D'arrivée de Karine. Alors ces deux éléments T_O et T_K, ce sont des variables aléatoires. Et qui prenaient leurs valeurs... Qui prenaient toutes les valeurs de... A32.
[3:35] [>Elève A32]: De l'intervalle [0; 1].
[>Marie]: Oui, on a tout de suite été sur [0; 1] en fait c'est [7; 8] mais qu'on peut ramener à [0; 1] sans aucune difficulté. Donc T_K et T_O sont des variables aléatoires qui prennent toutes les valeurs réelles de l'intervalle donc on va mettre [7; 8] mais que l'on peut ramener à l'intervalle... à une étude sur l'intervalle [0; 1] [*Marie écrit en même temps au tableau*]. Du coup, une variable aléatoire qui prend toutes les valeurs... toutes les valeurs réelles d'un intervalle donné, on...
[4:24] [>Marie]: On l'appelle comment ? C32.
[4:29] [>Elève C32]: Variable continue.
[>Marie]: Donc ce sont des variables aléatoires continues.
[4:39] [>Marie]: [*Marie écrit au tableau*]. Donc ce sont des variables aléatoires continues.
[4:49] [>Marie]: Et alors ce que vous avez très vite vu c'est que ces variables aléatoires c'était analogue à... B53 ?
[4:54] [>Elève B53]: A la puce du DM...
[>Marie]: Voilà, elles suivent le même modèle que la puce, on va mettre ça comme ça. Et qui suivent [Inaudible] le modèle de la puce. [*Marie écrit au tableau en même temps*] .
[5:16] [>Marie]: Est-ce que vous êtes d'accord avec ça ?
[5:21] [>Marie]: Oui ? Non ?
Plusieurs élèves répondent "oui".
[5:36] [>Marie]: Alors ensuite, c'était pas ça qui nous intéressait mais on l'a défini à partir de ces deux-là, c'était le temps d'attente du premier arrivé et ça on l'a noté X samedi. Donc on note [*Marie écrit au tableau*] X la variable aléatoire égale au temps d'attente... du premier arrivé. Alors celle-ci elle prend toutes les valeurs de quel intervalle ? C11 ?
[6:10] [>Elève C11]: [0; 1].
[>Marie]: Oui. Donc X prend toutes les valeurs dans... de l'intervalle [0; 1]. [*Marie écrit au tableau*]. Donc c'est aussi une variable aléatoire continue. Et donc là, C41 avait aussi proposé que ça suive aussi le même modèle que la puce. C'était ça son idée. Et... Du coup on a fait quoi pour valider l'hypothèse de C41 ? B11.
[6:40] [>Elève B11]: Modélisé.
[>Marie]: On a ?
[>Elève B11]: Modélisé.
[>Marie]: Alors, on a modélisé. Avant de modéliser on a...
[6:43] [>Elève B11]: Simulé.
[6:48] [>Marie]: On a simulé le temps d'arrivée avec le tableur d'accord. Enfin le temps d'attente excusez-moi avec le tableur. Alors ça, c'était pas très difficile à faire. Puis on a... On avait une série... Un échantillon de taille 10 000. On pourrait en avoir plein d'échantillons. D'ailleurs, j'ai oublié de vous donner samedi. Et à partir de ça effectivement, des échantillons... Alors on va les donner parce qu'il y en a qui étaient pas là donc ils ont pas vus. Tu pourrais distribuer rapidement sans parler à chacun de tes camarades, A12, toi ? [*Il est question du document Echantillons temps d'attente*].
[7:17] [>Elève A12]: Peut-être. [...]

[7:32] [>Marie]: Alors donc voilà, ça par exemple, parce que je vous ai montré à l'écran mais il y avait pas de... Je vous avais pas donné de poly. Alors on a modélisé après. Qu'est-ce que tu voulais dire par là ?
[7:39] [>Elève B11]: On a tracé une courbe de tendance.
[7:44] [>Marie]: Oui on a vu que ces histogrammes. Alors je vais en récupérer un. [...] Je vais en garder une pour moi aussi. Alors là par exemple, c'est plusieurs simulations d'échantillons de taille 10 000, soit réparties en classe de 15 min, soit ici à peu près en classes de 5 minutes, à chaque fois. D'accord ? Alors on a constaté que ces histogrammes ils avaient quoi en... J'ai à peu près mangé la réponse là. B21.
[8:17] [>Elève B21]: La même allure.
[>Marie]: Ils avaient tous la même allure. D'où l'idée de A52 d'utiliser une... de chercher une fonction qui alors... qui qui faisait quoi A52 en fait ?
[8:29] [>Elève A52]: Qui lissait [Inaudible].
[>Marie]: Qui ?
[8:34] [>Elève A52]: Qui lissait [Inaudible].
[>Marie]: Oui c'est ça. Plus on fait, alors là, je les ai pas imprimé mais on a regardé à un moment, si on fait des histogrammes de 1 minute et non pas de 5 minutes, ça devient en haut le haut des rectangles ils prennent presque l'allure d'une ici d'une droite, en tout cas d'un segment de droite. Alors pour décrire la loi de X... [*Marie écrit au tableau*]. La loi de probabilité de X, on a introduit une fonction F une fonction f pardon. On l'a noté comme ça je crois. On a introduit une fonction f. Alors en fait elle est définie sur quoi cette fonction ? A42.
[9:25] [>Elève A42]: [0; 1].
[>Marie]: Sur l'intervalle [0; 1]. Définit sur l'intervalle [0; 1]. C'est l'intervalle sur lequel X prend toutes ces valeurs. Donc sur l'intervalle [0; 1] [*Marie écrit au tableau*]. Alors on a mis un petit peu de temps samedi à me trouver l'équation de ma... enfin l'expression de ma fonction mais finalement nous avons vaincu et nous avons trouvé, B43 ?
[9:50] [>Elève B43]: $f(x)$ égal... c'était $-2x$...
[9:55] Plusieurs élèves: $+2$.
[>Marie]: $+2$.
[>Elève B43]: $-2x + 2$.
[>Marie]: Donc par $f(x)$ égal -2 oui $-2x + 2$ si tu veux. Je vais y arriver. $-2x + 2$. Alors cette fonction donc elle était telle que sa courbe... On met la phrase de A52, là ? Sa courbe lisse... Qu'est-ce que vous voulez qu'on mette ?
[10:21] [>Elève ?]: C'est bien.
[>Marie]: C'est bien quoi ?
[10:26] [>Elève ?]: La phrase de A52.
[>Marie]: La phrase de A52, que A52 va donc nous répéter.
[>Elève A52]: J'ai oublié.
[>Marie]: Qu' il a oublié.
[10:31] [>Elève B11]: La courbe lisse le haut...
[>Marie]: Le haut des rectangles. Telle que la courbe de f, on va mettre tout ça entre guillemet, lisse le haut des rectangles de l'histogramme. Est-ce que tout le monde comprend l'idée en tout cas à défaut de le dire de façon facile ?
[10:53] [>Elève A42]: Oui.
[>Marie]: Merci A42.
[10:58] [>Elève A42]: Je vous en prie.
[>Marie]: [*Marie finit d'écrire*]. Et alors? C'est là que ça semblait pas très clair après. Enfin je suis pas sûre que jusque-là c'est... Est-ce que jusque-là ça va ? C'est ce qu'on a fait samedi ?
Plusieurs élèves répondent "oui".
[11:12] [>Marie]: Et alors du coup... A quoi ça nous sert d'avoir trouvé cette jolie petite fonction ?
[>Marie]: A32.
[>Elève A32]: C'est pour calculer des probabilités quand X appartient à un certain intervalle.
[11 :27] [>Marie]: D'accord. Alors. [*Marie écrit au tableau*]. Ainsi pour calculer la probabilité que X appartienne à l'intervalle… alors prenne ses valeurs dans l'intervalle $[a; b]$, alors $[a; b]$ c'est un intervalle inclus dans [0; 1] . Alors on fait quoi ? Comment est-ce qu'on calcule cette probabilité, alors A32 ?
[11:52] [>Elève A32]: On calcule l'aire du triangle ou du trapèze enfin de la forme... on calcule l'aire de la forme géométrique qui va se trouver en-dessous de la courbe.

[12:00] [>Marie]: Oui c'est ça. On calcule l'aire on le dit comme ça l'aire sous la courbe, donc sous C_f, je l'écris en toutes lettres. Sous la courbe représentative de la fonction f. D'accord. Entre a et b. Alors.
[12:25] [>Elève A52]: Entre...
[>Marie]: Pardon ?
[>Elève A52]: Entre y et la courbe.
[12:31] [>Marie]: Oui alors on va faire un dessin car c'est ce que j'allais dire, un beau dessin vaut mieux qu'un long discours. Mais en fait vous allez voir qu'on va garder cette expression sans plus de précisions en se mettant d'accord de ce que... sur ce que ça veut dire, c'est-à-dire en fait ça veut dire… Alors on va repréciser des petits choses sur la fonction f mais ça veut dire entre la courbe, l'axe des abscisses, on va pas en-dessous, et entre deux droites verticales. On formalisera ça dans le cours mais je vais faire un dessin, ça sera aussi bien que d'écrire tout ça je pense. Oui, C41 ?
[13:01] [>Elève C41]: Sinon on peut l'appeler le rectangle a...
[13:06] [>Marie]: Mais c'est pas forcément un rectangle.
[>Elève C41]: Ah oui.
[13:11] [>Marie]: Je peux pas mettre une forme parce que en plus là on avait une fonction sympa d'accord? Donc c'est vrai que du coup quand... Après on va corriger les petites questions que... auxquelles vous aviez à répondre aujourd'hui mais il se trouvait qu'on obtenait soit des triangles, soit des trapèzes,...
[13:26] [>Elève C41]: La forme géométrique on pourrait l'appeler $a, f(a), f(b), b$.
[>Marie]: Oui si tu veux, c'est vrai. Mais on va juste faire un dessin tu veux bien ? Mais tu as raison on pourrait s'en sortir comme ça. Le tout c'est que tout le monde comprenne bien ce qu'on entend par aire sous la courbe et c'est comme ça qu'on va le désigner après. Donc je vais quand même prendre l'exemple de notre... [Inaudible]. [*Marie trace la courbe au tableau*]. Donc là ça valait 1, là ça valait 0. D'accord ? Donc ça c'est notre y... Notre courbe. Donc si on prend un intervalle $[a; b]$ dans notre... inclus dans $[0; 1]$, par exemple ça [*Marie montre un intervalle sur son graphique*], ce qu'on appelle l'aire sous la courbe, c'est la partie que je vais hachurer là. L'aire sous la courbe entre a et b c'est ça. Est-ce que c'est bon pour tout le monde ? C'est bien aire sous la courbe non ? Ca suffit ? Plusieurs élèves répondent "oui".
[14:19] [>Marie]: J'étais sûre que vous alliez être d'accord. Donc l'aire qui est ici, de ce domaine plan, d'accord, donc l'aire hachurée c'est égal à quoi ?
[14:39] [>Elève A52]: P de X.
[>Marie]: On lève le doigt, merci. Allez, allez, tout le monde peut donner la réponse là. Merci. A14.
[14:50] [>Marie]: Oui alors, je te révèle la règle du jeu A14, on lève le doigt quand on a la réponse.
[>Elève]: Petite base...
[14:55] [>Marie]: Non. Non. Excuse-moi ma question est pas claire. Non c'est vrai. Bien sûr tout le monde sait calculer l'aire de cette figure qui est un...
Plusieurs élèves répondent "trapèze".
[15:02] [>Marie]: Trapèze. Surtout depuis qu'on l'a rappelée. Oui. Non, je dis par rapport à notre problème à nous, notre problématique, cette aire elle mesure quoi pour... dans le cadre de notre problème, A14. c'était ça ma question. Je me suis mal exprimée. Non, tu vois pas ma question. Je la pose toujours mal je crois. A32.
[15:17] [>Elève A32]: La probabilité que X appartienne à l'intervalle $[a; b]$.
[15:22] [>Marie]: D'accord ? Donc c'est la probabilité que X appartienne à l'intervalle $[a; b]$. Alors soit dit en passant, on note ça soit de cette façon X appartient à l'intervalle $[a; b]$ mais on peut aussi noter la probabilité que X soit compris entre a et b de cette façon. Vous avez le droit aux deux... notations. Ok ?
[15:42] [>Marie]: Alors, revenons juste un petit peu sur du vocabulaire. Alors cette fonction f on va l'appeler la fonction de densité de probabilité de X.
[15:57] [>Marie]: Il y a A52 qui fait "Ouah". Ça c'est un mot quand même, faut m'expliquer un peu... Donc f s'appelle... [*Marie écrit au tableau*]. Alors... Il y a des questions ? Une rébellion ? Non, c'est bon ? Donc, f s'appelle la fonction de densité de probabilité de X. [*Marie écrit au tableau*].
[16:40] [>Marie]: On peut déjà noter que... Alors est-ce que vous pouvez m'expliquer un petit peu ce vocable ?
[16:47] [>Marie]: Ça a l'air affreux ça : fonction de densité... Pourquoi à votre avis ça s'appelle comme ça ou est-ce que c'est parce que ça fait bien de mettre un truc moche ?
Quatre/cinq élèves lèvent la main.
[17:07] [>Marie]: B31.
[17:12] [>Elève B31]: C'est parce que sur l'axe des...
[17:17] [>Marie]: Oui, je suis en train d'essayer de les remettre les histogrammes.
[>Elève B31]: C'est parce que sur l'axe des ordonnées, c'est la densité.

[17:22] [>Marie]: Oui c'est ça le lien... dans le devoir à la maison, on a vu que lorsqu'on représentait un histogramme sur l'axe des ordonnées c'est la densité qui était portée, de chaque classe, donc forcément votre fonction $f(x)$ vous imaginez, vous l'avez fait, vous avez un histogramme avec la courbe dessinée dessus enfin sur l'histogramme, et bien sur l'axe des ordonnées c'est bien la densité des histogrammes et donc c'est pour ça que ça s'appelle la fonction de densité de probabilité. Alors cette fonction f qu'est-ce qu'elle doit quand même un petit peu... Est-ce qu'elle est... Elle peut être quelconque ? Ou est-ce qu'il y a quand même des contraintes dessus ? Est-ce que toute fonction pourra être une fonction de densité de probabilité ou est-ce qu'il faut qu'elle possède tout de même quelques propriétés ?
Cinq/six élèves lèvent la main.
[>Marie]: Ne me regarde pas comme ça [Inaudible].
[18:20] [>Marie]: Est-ce que... Vous faites pas. Est-ce que si je me place sur l'intervalle [*Marie trace une courbe au tableau*] [0; 10] par exemple une fonction comme ça, ça pourra être une fonction de densité de probabilité.



Un élève répond "non", l'élève A52 répond "oui".
[18:35] [>Marie]: Pourquoi ? Tu sais pas mais là comme ça tu la vois pas. Elle est trop moche.
[18:43] [>Elève C22]: On peut pas trop calculer d'aire.
[>Marie]: C'est vrai, c'est exact. Ok. Alors... [*Marie trace alors une autre courbe au tableau*]. Là tu peux plus me dire que tu peux pas calculer... Ça c'est bien ça ?

L'élève A52 répond "non".
[18:58] [>Elève C22]: C'est dans les négatifs.
[>Marie]: Ah c'est dans les négatifs. Alors A52, qu'est-ce que tu voulais dire ?
[>Elève A52]: Pareil parce que en fait c'est par rapport à l'histogramme donc on peut pas avoir de rectangles en-dessous.
[19:03] [>Marie]: Oui donc f elle doit être ?
[19:08] [>Elève A52]: Positive.
[>Marie]: Positive sur l'intervalle. f... [*Marie écrit au tableau*].
[>Elève B52]: Et continue.
[>Marie]: Alors f possède les propriétés suivantes. [*Marie écrit au tableau*].

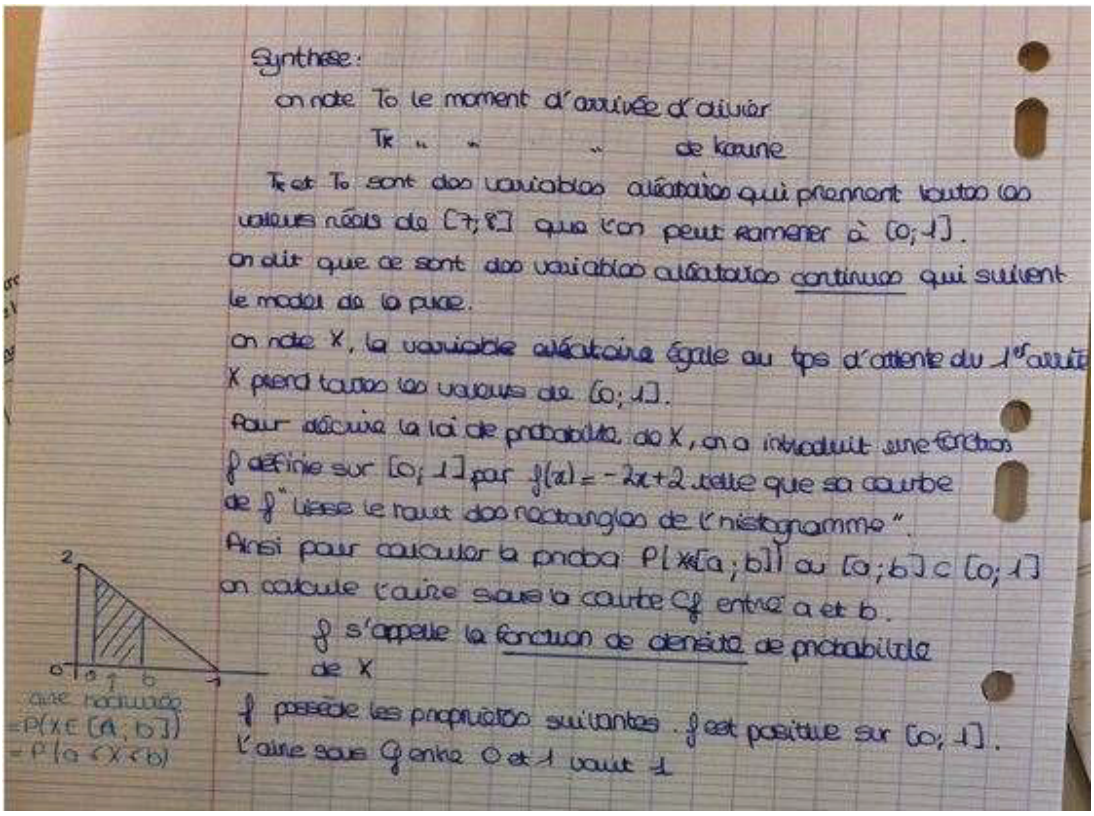
[>Marie]: f est positive donc ici c'est sur [0; 1], on généralisera un tout petit peu plus tard. Est-ce qu'il y a une autre propriété ? On les aura peut-être pas toutes là maintenant. Qu'est-ce... Qu'est-ce qui vous saute aux yeux, sans aucun jeu de mots. B52.
[19:38] [>Elève B52]: Elle est continue.
[19:43] [>Marie]: Alors pourquoi est-ce que tu veux qu'elle soit...
[>Elève A52]: Je suis pas d'accord.
[>Marie]: Ah. Alors A52 n'est pas d'accord. Donc toi B52 tu dis elle doit être continue et pourquoi A52 n'est pas d'accord.
[19:53] [>Elève A52]: Enfin peut-être qu'il a raison mais je me dis que finalement sur un histogramme il peut y avoir des valeurs qui sont...
[19:58] [>Marie]: Tu veux dire on pourrait on pourrait par exemple...
[>ElèveA52]: ... qui valent zéro.
[>Marie]: Si on avait...
[20:03] [>Elève A52]: Si c'est pas une expérience qui est complètement... où c'est pas équiprobable je veux dire, ça serait possible.
[20:13] [>Marie]: Parce que là tu penses que c'est de l'équiprobabilité ?
[>Elève A52]: Non. Enfin je veux dire si...
[20:18] [>Marie]: Tu veux dire la probabilité d'arriver entre... d'attendre...
[>Elève A52]: Non, je veux dire si jamais on a un rectangle avec une aire qui vaut zéro et bien du coup la condition [Inaudible].
[20:28] [>Marie]: D'accord donc du coup tu dessinerais, si tu avais un histogramme un peu comme ça, un truc plat après, encéphalogramme plat, tu ferais quoi comme genre de courbe alors pour le... ? [*Marie trace un histogramme au tableau*].

[20:38] [>Elève A52]: Je sais pas mais... Enfin je pense ça serait deux fonctions.
[20:43] [>Marie]: Alors une fonction définie par morceaux tu veux dire ? [*Marie trace au tableau*]. Alors là non j'exagère peut-être. Et puis après tu ferais un morceau comme ça ?

[20:54] [>Elève A52]: Oui j'aurais fait ça.
[>Marie]: Alors on va pas trancher ça maintenant mais effectivement les deux cas sont possibles. Mais dans le cadre du cours, nous, elles seront continues. Et alors par contre, une autre propriété qu'on a utilisé d'ailleurs ?

[21:09] [>Marie]: Qui est impérative. Allez, allez.
Quatre/cinq élèves lèvent la main.
[21:14] [>Marie]: Il y a bien un phénomène du côté du couloir. A22.
[21:24] [>Elève A22]: C'est avec X supérieur à 0.
[>Marie]: C'est-à-dire X supérieur à 0 ?
[>Elève A22]: On peut pas...
[>Marie]: Ah non. Ah non, c'est parce que là le temps d'attente... parce que là c'était un temps donc ça se mesure c'était de 0 à 1 mais je pourrais très bien imaginer quelque chose défini sur -10... Ca dépend des valeurs que prend ta variable aléatoire, ton expérience. B11 ?
[21:44] [>Elève B11]: L'aire en-dessous de la courbe vaut 1.
[>Marie]: Oui. L'aire sous la courbe vaut 1. [*Marie écrit au tableau*]. Donc entre 0 et 1. L'aire totale vaut 1. Et donc ça vous êtes d'accord on l'a utilisé pour obtenir notre équation et pourquoi ça doit valoir 1 ? A21.
[>Elève A21]: Parce que ça correspond à la fréquence. L'aire correspond...
[>Marie]: Moi je suis sûre que C41 a pas du tout entendu ce que tu as dit.
[>Elève A21]: L'aire correspond à la somme de toutes les probabilités...
[>Marie]: Non des fréquences oui enfin de ce qu'on avait fait.
[>Elève A21]: La somme des fréquences.
[>Marie]: Et donc on sait que ça fait...
[>Elève A21]: Un.
[>Marie]: Un. On l'a utilisé dans le devoir et on l'a réutilisé samedi. Alors, on va on va répondre un petit peu rapidement si vous le voulez bien. Je pense que les calculs vous ont pas trop posés de problèmes. J'ai perdu les questions moi. Mince. Voilà. Le petit exercice. On corrige. [22:51]

Synthèse notée par l'élève B53 :



I.3 Transcription de la séance 3

Transcription de la séance 3 du 18/03/15 (55 minutes – classe entière)

34 élèves présents.

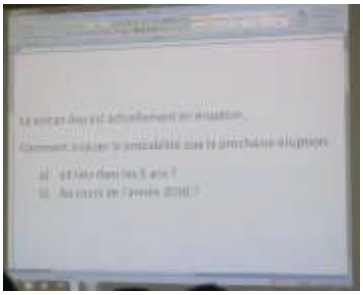
[00:00] *Mise en place des élèves.*
[00:19] Marie : Vous prenez vos cahiers d'exercices, s'il vous plait.
Marie écrit au tableau : « Problème 2 ».
Marie : Alors, comme pour le premier problème, vous prenez des notes. […]
Marie : Donc à la suite du problème 1, nous allons nous attaquer à un deuxième problème. […]
Un élève distribue le problème à chacun. […]
Marie : Alors. Je vous laisse prendre connaissance […] en silence du nouveau problème. […]
Marie : Alors. Qui lit le petit préambule ? A haute voix, pour toute la classe, s'il vous plait. A21. Merci.
Elève A21 : Le volcan Aso, situé sur l'île de Kyushu au Japon, est l'un des plus actifs au monde. On possède les statitiques (sic) de ses éruptions…
Rires.
[…]
Elève A21 : On possède les statistiques de ses éruptions régulièrement tenues depuis le XIIIè siècle. Nous avons, dans le document tableur, les années d'éruptions jusqu'au XIXè siècle. A partir du XXè siècle, les données, d'une autre nature, ne sont pas compatibles. On s'intéresse au temps écoulé, en années, entre deux éruptions successives.
Marie : Voilà. Alors, vous avez le tableau, d'accord ? Des relevés des éruptions du volcan… Alors je sais pas… Si vous savez comment on prononce le nom du volcan[1]… pour être honnête avec vous.
Plusieurs élèves : [Asso].
Marie : Ok.
Deux retardataires arrivent. […]
Marie : Alors. Je vous laisse regarder les données.
Marie : Et prendre note. Alors, il se trouve que le volcan Aso est actuellement en éruption. Là, au mois de mars 2015. On peut… regarder des vidéos sur internet.
Un élève : On va le faire ?
Une autre élève : Moi, j'y suis allée.
Marie : Tu y étais ? Où ça ?
L'élève : Au volcan.
Marie : Ah au volcan. Et il était en éruption quand tu y étais ?
L'élève : Oui, je crois.
Marie : Oui, tu crois.
Rires.
L'élève : Non, mais y a tout le temps des…
Marie : Ça doit se voir. Ah il y a tout le temps des petits nuages… ? D'accord. Ah ce qui expliquerait… Parce qu'on reparlera des données un peu plus tard, enfin de celles que nous avons sélectionnées. […] Voilà ce que je vous propose, puisque… […] puisque le volcan est en éruption en mars 2015 […].
Un troisième élève arrive en retard. […]
Marie : Donc, voilà. Comment évaluer que la prochaine éruption ait lieu dans les 5 ans ou au cours de l'année 2030 ?
Les questions sont vidéoprojetées.

[1] Certains élèves de la classe font japonais comme langue vivante.

Le volcan Aso est actuellement en éruption.

Comment évaluer la probabilité que la prochaine éruption:

a) ait lieu dans les 5 ans ?
b) Au cours de l'année 2030 ?



Marie : Alors je vous laisse quelques petites… instants. Chacun… pour me faire des propositions honnêtes…
Marie : Vous voyez les questions, ceux qui sont du côté… C31 et C42 ? […]
[04:49] Marie : Donc, c'est bon tout le monde voit le tableau ? Bien, je vous laisse 3 minutes, 4 minutes… Voilà. Comment évaluer la probabilité que notre volcan est en éruption… enfin soit en éruption dans les 5 années à venir ? ou au cours de l'année 2030 ? Enfin et au cours de l'année 2030… Et donc vous possédez… tout le monde a eu ou va bientôt avoir le document. Les retardataires, vous levez le doigt s'il vous plait.
[…]
Recherche individuelle. Silence.
Marie passe dans les rangs.

Quelques extraits, nourris par des notes prises pendant la séance.

Marie (à un élève) : Alors, la première éruption relevée, elle a lieu en 1929 et la suivante a été relevée en dix-neuf cent… en dix-neuf cent n'importe quoi en 1239. Donc il s'est passé 10 ans.

Un élève pose une question sur l'astérisque sur le document. Marie lui explique que c'est lié aux types d'éruptions.

Marie propose aux élèves de discuter éventuellement avec son voisin du problème posé.

Marie (à une élève) : Non, dans les 5 ans, c'est-à-dire au cours des 5 années prochaines, donc 1, 2, 3, 4, 5 mais pas plus.
L'élève : Ok.

Marie (à l'élève B52) : Alors. Toi tu as fait la moyenne. Oui, vas-y, dis au moins que tu as fait la moyenne. Note ton résultat. Et… En quoi… En quoi cette moyenne nous permet d'évaluer… pour répondre aux questions que je vous ai posées ?
Marie (aux élèves B52 et B53) : Mettez ce que vous avez fait, mettez le résultat que vous avez trouvé…

Marie (à la classe) : Ne perdez pas de vue les questions que j'ai posées.

Marie (aux élèves C21 et C22) : D'accord, tu veux… des données brutes, tu veux les mettre dans un tableau.
Elève C21 : Pour faire un histogramme.
Marie : Ah vous voulez faire un histogramme. Alors si vous voulez faire un histogramme, comment on fait ? C'est des données…
[Inaudible].

[…]

Marie (à un élève) : Alors X une variable aléatoire de I. C'est qui I ? Qu'est-ce que tu… « Posons le temps d'attente ». [Inaudible]. En fait X et I c'est pareil, t'es pas obligé de le mettre.

Marie (à l'élève A12) : Alors tu veux qu'on trace une courbe de tendance grâce à l'histogramme du temps d'attente j'imagine.
Elève A12 : Oui.
Marie : Cet histogramme du temps d'attente, là, à partir de ça, comment je le fais ?
Elève A12 : Je prends le temps d'attente. Je regarde… les fréquences à chaque… quelle va enfin… quelle valeur…
Marie : Tu connais les temps d'attente, du coup tu peux calculer la fréquence de chaque temps d'attente. D'accord. [*Inaudible*].
Marie : Si on fait ça comme ça c'est une variable continue ou bien discrète ?
Elève A12 : Discrète.
Marie : Ouais. Est-ce que tu… Est-ce que c'est discret le temps d'attente ?
Elève A12 : Non.
Marie : Non, c'est continu. Donc… Pour faire ton histogramme, il faut pas… Faut peut-être pas mettre le 1 an [*inaudible*], 2 ans [*inaudible*]. Tu vois ce que je veux dire ? Dans un histogramme, il faut que tes données elles soient réparties comment ?
Elève A12 : De manière continue.
Marie : Oui. Donc en classes. D'accord ? Donc faut que tu… tu me dises quel histogramme tu veux que je fasse si…
Elève A12 : Ok.

Marie (à la classe) : Alors, je vous laisse encore 1 ou 2 minutes. Je vois que tout le monde a proposé quand même ou presque…

Marie (aux élèves A41 et A42) : Pourquoi tu dis c'est difficile ? Qu'est-ce qui te gène ?
Elève A42 : En fait au début je voulais faire un histogramme exactement dans cet ordre-là pour voir si y avait un côté cyclique.
Marie : Oui d'accord.
Elève A42 : *[Inaudible].* Vu que ce que je dis c'est pas facile à évaluer.
Marie : Oui, alors. Parce que c'est des données brutes. Des données brutes en statistiques. *[Inaudible].*

L'élève B43 se demande s'il ne s'agit pas d'une suite, avec n l'année de l'éruption et u_n *la probabilité.*

Marie réexplique le problème à l'élève B44.

Les élèves B52 et B53 ont regroupés les données en classes, sous forme d'un tableau.

Plusieurs élèves sont en train de regrouper les données pour essayer de faire un histogramme.

Marie (aux élèves B21 et B22) : Alors qu'est-ce que vous proposez ?
Elève B22 : De faire un histogramme.
Marie : D'accord.
Elève B22 : Et après, on fait comme… l'aire…
Marie : On cherche une courbe.
Elève B22 : De tendance. Et il y aura pas de trapèze, je pense.
Marie : Pensez pas encore à la tête qu'elle a, la pauvre. On sait pas encore.
Ça se trouve c'est des carrés, c'est des triangles. On sait pas. D'accord. Ça se trouve elle est comme ça ta courbe de tendance. Pourquoi pas ?

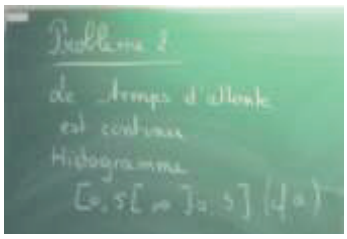
L'élève B31 a dessiné une courbe de tendance parabolique (sommet en haut) passant par (0 ; 0). *Elle me dit que, avec les données, ça a l'air d'être ça.*

[18:45] Marie : Alors. On va peut-être… Je crois que tout le monde a… a un peu avancé sur ses propositions. Alors… Qu'est-ce que vous proposez ? Avec ces données. Alors effectivement ce sont des données brutes, qui n'ont pas été triées. Alors, est-ce que tout le monde est de nouveau ensemble ? Je vous laisse 10 ou 15 secondes pour finir votre phrase. […] Alors, ça se sont des données brutes, données brutes en statistiques. Donc souvent ce passage-là, il est pris en charge, quand vous avez fait des statistiques en troisième, etc. Pas tout à fait ? En seconde. Mais, c'est vrai que le premier réflexe par exemple, c'est ce que C42 a eu comme réflexe et d'autres. C42, oui c'est toi. Quel est le premier réflexe que tu as eu face à ces données brutes ?
Elève C42 : C'est de les… c'est de les regrouper.
Marie : Oui, c'est de les… mettre dans un tableau de distribution de… de séries statistiques, pardon. C'est-à-dire de… pour chaque du… temps d'attente de compter le nombre d'occurrences, en tout cas l'effectif et de calculer la fréquence correspondante. Donc ça c'est un des premiers réflexes que beaucoup ont eu. Alors, après… avec ce tableau-là, par exemple… C42 voulait faire un histogramme. Est-ce qu'avec un tableau de ce type, on va faire un histogramme ? A52 ?
[20:12] Elève A52 : En fait, on pourrait mais faudrait d'abord, prendre un intervalle de…
Marie : Oui, parce que si on… si on fait le tableau dont parle C42, on a une série discrète, d'accord ? Parce qu'on a « une année » apparait, je dis n'importe quoi, 10 fois… « 2 ans » est apparu 3 fois… etc… Donc c'est pas un histogramme qu'on trace dans ces cas-là, c'est un diagramme…
Plusieurs élèves : En bâtons.
Marie : En bâtons. D'accord ? Or, le temps d'attente, à votre avis, est-ce que c'est du discret ou du continu ? Qu'est-ce qu'on va considérer ? B53 ?
Elève B53 : Continu.
Marie : Du continu. Pourquoi tu dis ça ?
Elève B53 : Parce que… enfin…
[…]
Elève B53 : Parce que c'est l'année, enfin…
Marie : C42, tu voulais dire aussi… Pourquoi c'est du continu ?
Elève C42 : Parce que ça peut arriver à tout moment durant l'année [*inaudible*].
Marie : Oui. Par exemple… Prenons… Entre 1272 et 1273, il y a eu un an d'attente. Quand on est… quand on relève qu'il y a eu un an d'attente entre ces deux éruptions, ça veut dire… Qu'est-ce que ça veut dire ? Ou qu'est-ce que ça peut vouloir dire ?
Marie : Est-ce que c'est un an pile ?
Plusieurs élèves : Non.
Marie : C'est ça ma question. B31 ?
Elève B31 : Non. C'est… ça peut être… c'est autour d'un an en fait.
Marie : C'est autour d'un an… Mais alors du coup, ça peut… ça peut vouloir dire quoi ?
Elève B31 : Entre 0 et 1 an… Ou entre…
Marie : Mieux que ça.
Elève B31 : … 1 an et 2 ans.
Marie : Entre 0 et… ?
Elève A52 : 2 ans.
Elève B31 : Et 2 ans.
Marie : 2 ans. Est-ce que tout le monde le comprend ?
Un élève : Si c'est en janvier…
Marie : Oui. Si il y en a une en décembre de la fin de l'année… précédente et que la suivante est en janvier, c'est pas la même année civile donc c'est noté « 1 an » en années mais y aura eu qu'un mois d'écart. D'accord ? Par contre, si l'autre a eu lieu en janvier de l'année 1272 et que la suivante a lieu en décembre de l'année 1273, il s'est écoulé en fait… 2 ans. D'accord ? Donc, déjà ça, c'est un problème avec ce type de données. D'accord ? Parce que le « 1 an », il représente quand même une durée de temps qu'est pas du tout… c'est pas date à date… C'est pas il y a une éruption le 5 septembre 1272 et l'autre a eu lieu exactement le 5 septembre 1273. Donc déjà… et puis le temps… On en a quand même déjà beaucoup parlé… C'est plutôt comment, le temps ? Comme phénomène ? B31 ?

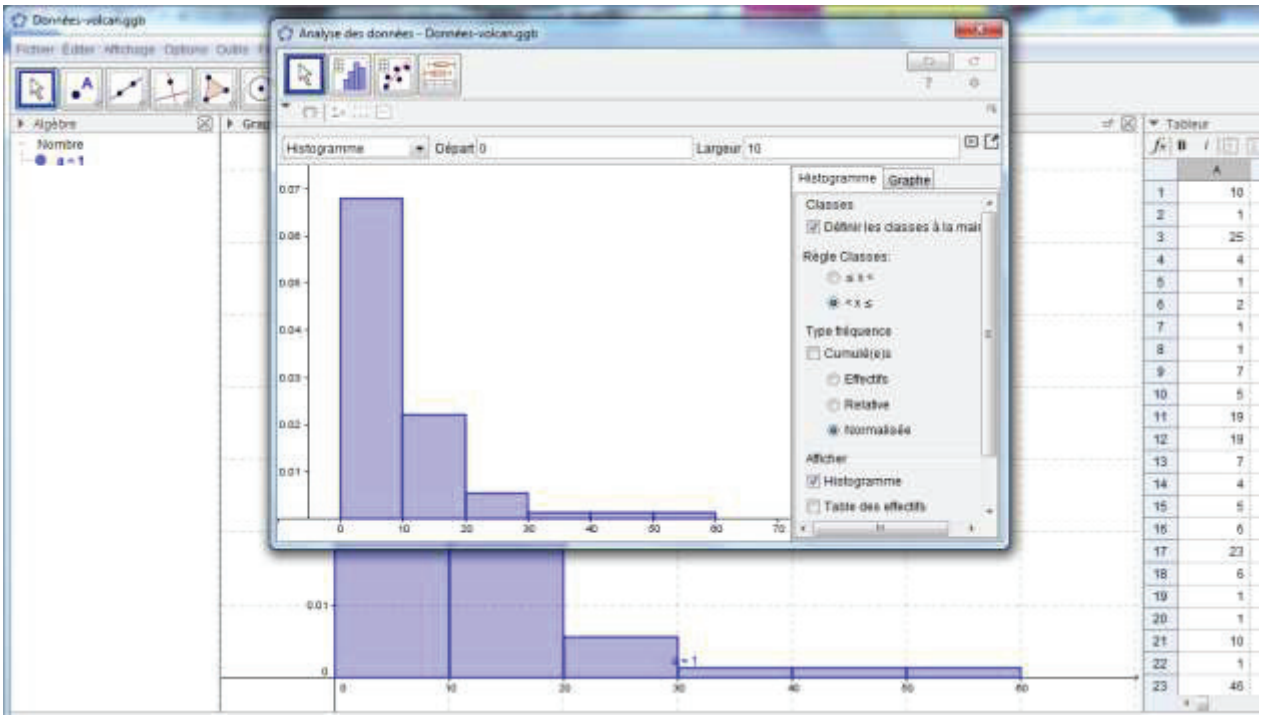
Elève B31 : C'est continu.
Marie : C'est continu. Donc effectivement au lieu de faire un tableau… données par données, on va les regrouper en classes. Alors, comment est-ce qu'on fait pour regrouper en classe ce genre de… Qu'est-ce que vous prenez ? Qu'est-ce que vous décidez ? B11 ?
Elève B11 : Moi j'ai regardé la question et il nous dit qu'il nous faut… y ait lieu dans les 5 ans, du coup, j'ai pris une classe de 5 ans à chaque fois.
Marie : Ah c'est pas bête. J'y avais pas pensé. On est d'accord que le temps d'attente, c'est continu. […] Donc, qu'est-ce qu'on a dit d'important jusqu'à présent ?
Marie écrit au tableau. Silence.
Marie : Rien. D'accord. Le temps d'attente donc est continu. Vous voulez faire un histogramme. J'écris que des mots, tant pis pour vous. […] Et donc B11, il veut regrouper dans des classes de [0 ; 5[, grâce à la première question. D'accord. Donc… Prenons des… des classes, on peut choisir ça. On peut choisir ça. [*Il s'agit des différents sens pour les crochets (cf tableau)*].

Problème 2 :

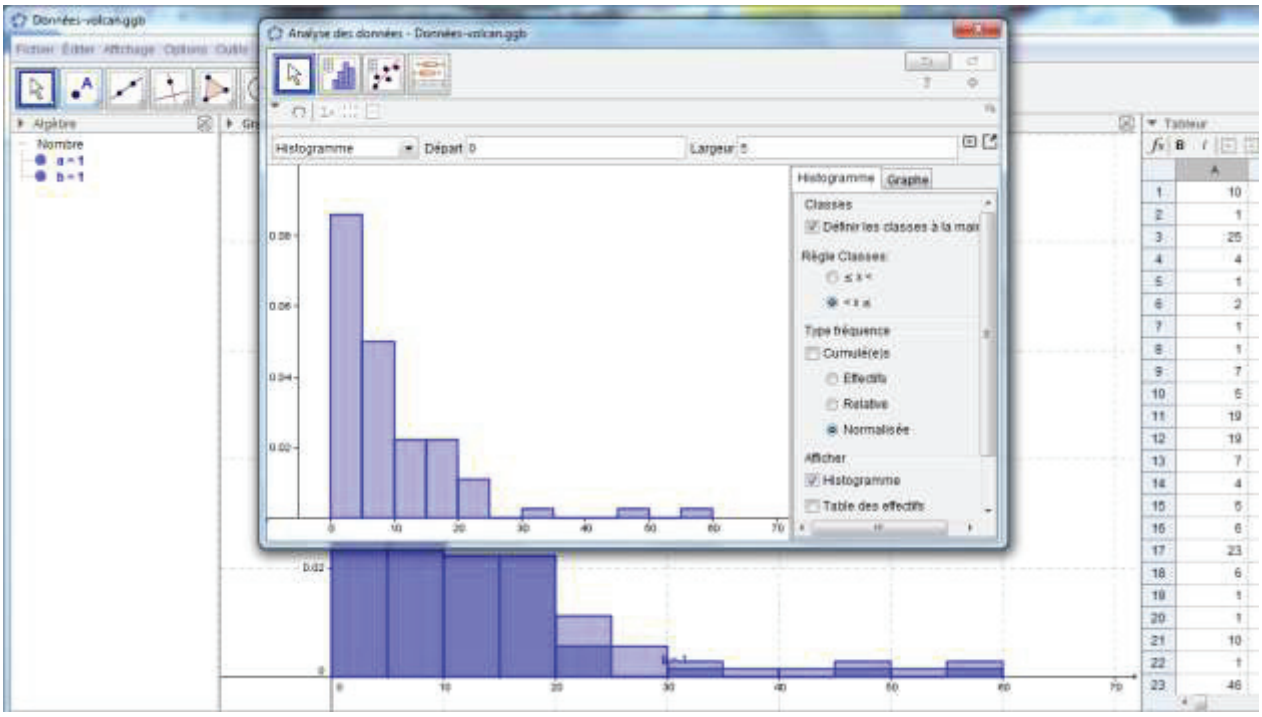
Le temps d'attente est continu
Histogramme
[0 ; 5[*ou*]0 ; 5] *(cf a)*



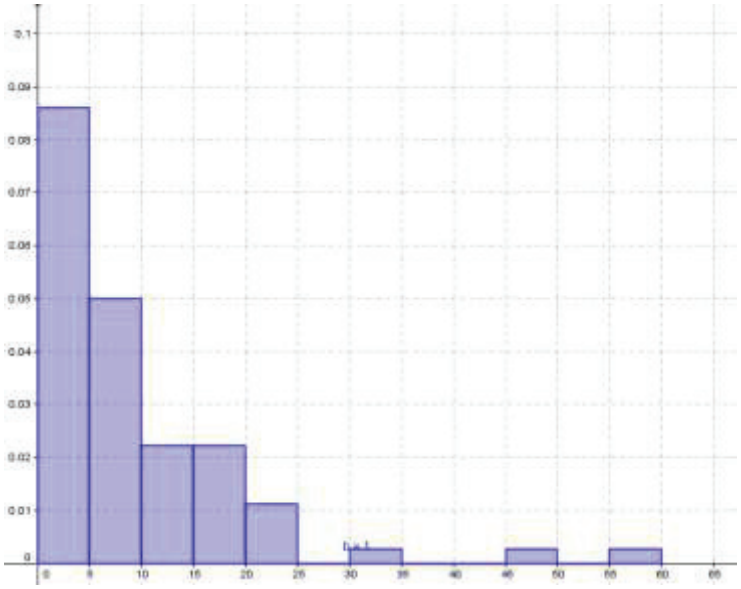
Elève B11 : Moi j'ai pris la deuxième.
Marie : Tu as pris celle-ci [*montrant l'intervalle]0 ;5]*].
Elève B11 : Ouais, parce que il y a pas de 0 en fait.
Marie : Il y a pas de données… Il y a pas de 0 donc tu l'as exclu d'office pour prendre ça. Donc confère guidé par la question petit a. Est-ce que ça vous semble raisonnable ce que propose B11 ?
Plusieurs élèves : Oui.
Marie : Oui. Alors, pour construire ce…
[24:12] Marie : Est-ce que quelqu'un a fait autre chose pour répondre au petit a ? Est-ce qu'il y a une autre proposition ? Est-ce que quelqu'un a trouvé une évaluation ? Là ?
Marie : Non, toujours pas ?
Marie : Non. D'accord.
[…]
[24:40] Marie : Alors, et bien… Allons-y alors parce que le pauvre… il y en a qui sont en train de compter, qui ont commencé, c'est ça ?
[…]
Marie : On va utiliser GeoGebra quand même… ça ira un peu plus… Ah, je vois le sourire sur le visage de B31 qui se dit « je veux pas les compter, je veux pas les compter »… Alors, j'ai mis les données dans la colonne A, d'accord ? Et donc, on va faire un bel histogramme comme veut B11, d'abord ? Alors… Attendez… Le voici.
Marie manipule le logiciel en même temps.
Marie : Alors, on veut… On veut donc les fréquences… Tu veux ça… [*En parlant du sens des crochets*]. Tu veux des classes de largeur 5. C'est ça que tu m'as dit ? Ok. Alors, voyons ce que cela donne.
Silence. Puis certains élèves ne semblent pas d'accord.



Un élève : C'est pas 10 là ?
Elève B32 : C'est largeur 10…
Elève C42 : C'est 10.
Marie : Qu'est-ce qui y a ? C'est pas bien ?
Elève B32 : C'est largeur 10…
Marie : Qu'est-ce que j'ai fait ?
Une élève : C'est pas 5…
Marie : Oui, non c'est pas 5. Vous avez raison.
Marie manipule le logiciel à nouveau.
Marie : Ah. Intéressant…
Marie : Qu'est-ce que ça fait ? Qu'est-ce qu'il s'est passé ? Alors, attendez, on va changer.
Marie : Je teste juste. Là, c'est largeur 1 ?
Marie : On va l'avoir. 2. Alors, qu'est-ce qu'il a ? 3. Il se rebelle. 5. Je l'ai eu !



Marie : Alors. Bon, il faut pas chercher. Non, non mais c'est parce que j'ai pas enlevé le premier. Bon, là c'est largeur 5 ? On s'explique pas vraiment ce qu'il vient de se passer précédemment, c'est le logiciel. Bon, c'est sympa. Alors, B11 ? Qu'est-ce qu'on en fait ?



[26:51] Elève B11 : On pourrait… Dans l'idéal, ça aurait été de voir une courbe… de tendance mais c'est…
Marie : Ah. Dans l'idéal, je sais pas si les autres ont entendu.
[…]
Elève B11 : J'ai dit dans l'idéal, on pourrait faire une courbe de tendance mais vu l'histogramme, ça va être compliqué…
Marie : Alors, dans l'idéal, on pourrait faire une courbe de tendance mais vu la tête de l'histogramme, ça va être compliqué ?
Elève B11 : Oui.
Marie : Pourquoi ?
Elève B11 : Ça fait des trous…

Marie : Oui il y a des trous… Forcément. On a des données…
Elève B11 : Oui.
Marie : On peut pas… B52 ?
Elève B52 : C'est ce que disait A52 en fait hier, il faut pas que la courbe de tendance elle soit continu.
Marie : Tu veux qu'on fasse une courbe de tendance qui soit pas continu ?
Elève B52 : Oui.
Elève A52 : Non, on a dit que c'était pas dans le programme.
Elève B52 : Ah…
Elève A52 : Ah non, on a le droit de faire une continue.
Marie : Il y avait des doigts levés par-là ou… ? A32 ?
Elève A32 : Je pense qu'il faudrait faire un histogramme avec des classes d'amplitudes différentes pour que ça puisse prendre toutes les valeurs… Enfin pour pas qu'il est de trous.
Marie : Ah. Alors, avec l'histogramme, on aimerait une courbe… Faire… Enfin on aimerait… On voudrait une courbe de tendance… mais il y a des trous. [*Marie écrit en même temps au tableau*].

Avec l'histogramme

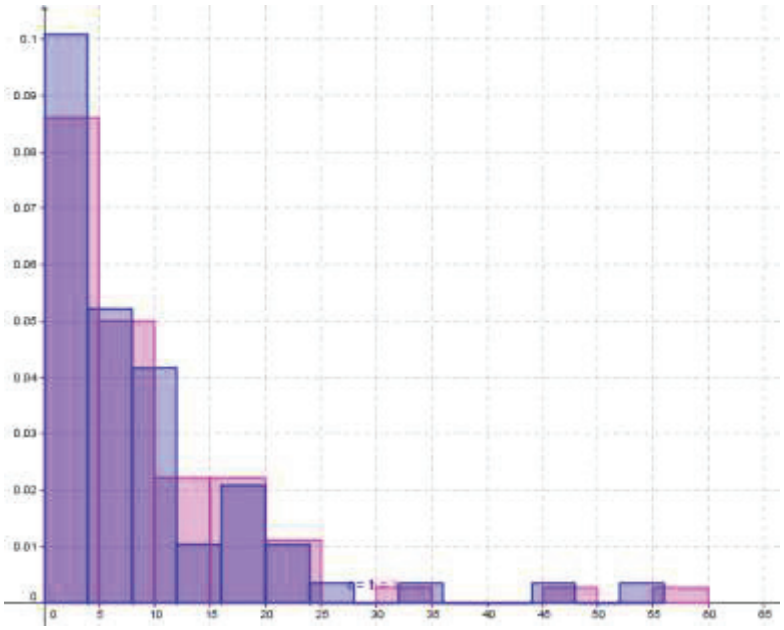
on voudrait une courbe de tendance mais il y a des trous.

Marie : Amplitudes différentes.

amplitude ≠ pr pas de trous

Marie : En fait, ça vous gêne que les rectangles soient pas collés quoi ?
Elève A52 : Non ça gêne pas… Ça gêne pas du tout.
Elève A32 : Non, mais c'est le fait que après…
Marie : Ah non non non, mais ce que je comprends. Vous me dites qu'il y a des trous… C'est pas ce que tu as dit ?
Elève A32 : Non mais c'est pas gênant. Enfin c'est juste que comme ça… on pourrait peut-être trouver une courbe de tendance plus facilement, en fait.
Marie : J'ai pas entendu.
Elève A32 : On pourra certainement trouver une courbe de tendance plus facilement.
Marie : Si quoi ?
Elève A32 : Si on change les… l'amplitude des classes. Ça fera peut-être un truc plus… plus harmonieux.
Marie : Ah. Je crois qu'il y a deux choses. J'avais compris que ce qui vous gênait c'est qu'il y ait des trous et que du coup, ça serait pas continu.
Elève A32 : Ah non.
Marie : Quand on modélise, on… on peut combler les… justement, essayer de combler les trous. Mais ça, c'est pas ça que tu voulais dire toi. Ce que tu voulais dire, c'est les… les… les rectangles qui sont dessinés pour l'instant ne semblent pas… faire un truc… voilà.
Elève A32 : Non, mais je pense que faire un… des classes de 5, c'est pas…
Marie : Je peux en faire des… je peux… Comme vous pouvez voir, s'il est de bonne humeur, je peux faire autre chose que 5.
Elève A32 : Mais on peut pas faire…
Marie : Alors, attendez, attendez… Alors je crois que A42 est pas d'accord avec toi en fait.

Elève A42 : En fait, les… les 3… les 3 rectangles là, c'est des éléments isolés. C'est… 1 éruption sur 73 éruptions, on peut peut-être les négliger éventuellement. On peut faire comme une courbe de tendance qui continue…
Marie : Alors du coup, pourquoi tu veux les négliger ?
Elève A42 : Bah.
Rires.
Marie : Oui d'abord, c'est un peu…
Elève A42 : Ce que je veux dire, c'est…
Marie : On a déjà pas beaucoup de données alors…
Elève A42 : Mais ce que je veux dire, c'est…
Marie : Mais tu as raison, ça pourrait être des valeurs aberrantes. Parfois on les enlève…
Elève A42 : C'est pas des valeurs aberrantes, c'est qu'il y en a qu'un, à chaque fois, dans chaque tranche d'âge… enfin non d'âge…
Marie : Oui, d'attente. Oui.
Elève A42 : Et, donc du coup… Enfin…
Marie : Donc, enfin, tu m'as montré quelque chose, toi. Montre-moi ta copie... Viens... Viens montrer ce que tu veux comme courbe.
L'élève A42 se lève pour montrer la courbe sur l'histogramme vidéoprojeté (courbe de type exponentiel décroissante).
Marie : D'accord. Donc, est-ce que… chercher un truc de ce genre, ça pourrait…
Elève A52 : Oui, je sais pas comment on pourrait faire.
Elève A42 : C'est une parabole.
Marie : Attendez, on va chercher. Un problème à la fois. Soulevons un problème à la fois.
Elève A52 : Oui, c'est clair.
Marie : Est-ce que, si je change les amplitudes… Du coup, merci A42.
Marie manipule à nouveau le logiciel, pour superposer à l'histogramme existant, un histogramme d'amplitude 4.
Marie : Alors, je vais essayer de le faire vite… Donc là, on a mis 5. Si on met… 4. Pardon… Voilà… A11, tu veux quelle couleur ? Rose. D'accord. Merci, A11. Donc ça c'est pour 4. Le rose c'est 5 et le bleu, c'est des amplitudes de 4.



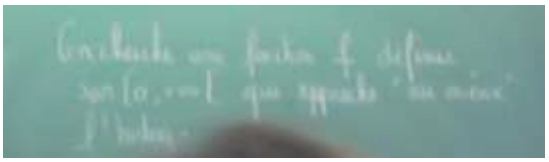
Une élève : On voit rien.

Marie : Comment ça on voit rien ? Alors, le bleu, j'ai mis des codes couleur exprès… Le bleu, c'est dans des… quand je répartis dans des classes de 4 ans et le rose quand je répartis dans des classes de 5 ans.
Une élève : 5 ans, c'est mieux.
Marie : D'accord. Mais… Est-ce que… est-ce que si je… Voilà, je vais y arriver. Excusez-moi.
Marie manipule le logiciel pour enlever l'histogramme d'amplitude 5.
Marie : A42, si je te mets l'histogramme de 4 ans, tu peux venir dessiner la courbe de tendance que tu voudrais.
Elève A42 : C'est la même.
Marie : Ah, c'est la même. Donc finalement les histogrammes… qu'est-ce qu'ils ont… ?
Elève A42 : Ils ont la même courbe.
[31:58] Marie : Enfin, ils ont la même allure on va dire… la même forme… Donc on peut peut-être se lancer. On en choisit un et on se lance pour… Comment on va faire pour trouver la courbe de tendance ?
[…]
Marie : Alors… On… Vous voulez lequel ?
Marie : Le 4 ans ?
Quelques élèves réfléchissent à voix haute.
Elève A52 : Ah non quoi que.
Marie : Quoi que. Je ne sais pas moi.
Marie : Alors, j'ai imprimé le 4 ans. Je repose la question, vous voulez lequel ?
Rires.
Elève A42 : 4 ans, c'est mieux que…
Marie : Non, non… C'est pas… Non, non, il y a pas de… C'est que je peux pas tout imprimer non plus donc… Donc on vous distribue l'histogramme qu'est au tableau.
Marie : Alors, donc le problème c'est quoi finalement là ? On en est où là tous ? Je vous laisse chacun 5 minutes.
Elève A42 : Faut trouver l'équation…
Marie : C'est quoi l'objectif ? Trouver…
Elève A42 : L'équation.
Elève A52 : Une courbe de tendance.
Plusieurs élèves : Une courbe de tendance.
Marie : L'équation, l'expression de… de votre courbe de tendance.
[33:18] Marie : Alors… ça s'appelle…
Un élève : Une parabole[2].
Marie : Non. La courbe de tendance, là, la fonction… Elle va nous servir à quoi cette courbe de tendance ? Dites-moi.
Plusieurs élèves : Trouver des probabilités.
Marie : En faisant quoi ?
Plusieurs élèves : Avec l'aire.
Marie : D'accord. Elle s'appelle comment cette fonction-là ? Qu'est-ce qu'on a dit hier ?
Elève A32 : Ah, courbe de densité.
Plusieurs élèves répondent en même temps.
Marie : La fonction de densité. D'accord.
Elève A42 : Densité de probabilité.
[33:37] Marie : Alors peut-être que… Alors, on n'a rien formalisé du tout mais… Ce temps d'attente-là entre deux éruptions… Donc on était d'accord que c'était un temps, que c'est du continu. Ça prend toutes les valeurs de quel intervalle quand même, pour qu'on sache sur quoi on travaille ?
Elève A52 : [0 ; 55].
Plusieurs élèves parlent en même temps pour proposer [0 ; 55] *ou* [0 ; 56].
Marie : Ah non, ça c'est les données. Ça, c'est les données qu'on a.

[2] L'élève veut dire « hyperbole ».

Elève C42 : [0 ; +∞[.
Marie : [0 ; +∞[? Est-ce que tout le monde est d'accord avec C42 ?
Plusieurs élèves répondent « oui ».
Elève A21 : Non.
Marie : Alors, « non », A21 ?
Elève A21 : Parce qu'il y a la fin du monde !
Marie : Non, il y a pas la fin du monde. On… Comme on la connait pas, tu pourras pas mettre… Tu veux qu'on mette en dernière bande « la fin du monde », mais comme on la connait pas.
Elève A21 : On peut pas mettre +∞.
Elève A42 : On met x.
Marie : Non.
Elève A52 : On peut pas mettre +∞.
Marie : Pourquoi on peut pas mettre +∞ ?
Elève A52 : Ça serait contre… contre…
Marie : Contre nature. [*Rires*]. B52 ?
Elève B52 : En fait, c'est comme si on se ramenait à un intervalle de [0 ; 1].
Marie : Ah non.
Marie : Qu'est-ce que…
Elève B52 : Nan mais comme hier, en fait.
Marie : Ah hier c'est pas pareil, on avait des intervalles fermés bornés, on a juste translater avec...
Marie : Alors. On cherche…
Marie écrit au tableau en même temps.
Marie : Une fonction, on l'appelle tous f … Je vous laisse 5 minutes. Une fonction f … Alors, on n'est pas tout à fait d'accord mais… définie sur [0 ;+∞[… qui approche, enfin comme d'habitude… Alors, je vous laisse 5 minutes pour me faire des propositions ? [35:13]

On cherche une fonction f définie sur $[0;+\infty[$ qui approche « au mieux » l'histog.



[…]
Les élèves travaillent individuellement. Marie passe dans les rangs.

Marie (à un élève) : Non, on en fait une.
[…]
Marie : Vas-y, propose. Moi j'ai dit on propose, alors on propose.

Marie (à la classe) : Oui, alors, si je vous ai donné un document écrit, c'est pour que vous puissiez faire au crayon de papier des trucs dessus. Je vous en redonnerai un autre après si vous voulez, plus propre, quand vous aurez fini de bidouiller.

Marie (à l'élève B23) : Alors tu veux un truc comme ça. D'accord. Et ça, ça aurait quel genre d'expression, B23 ?
Elève B23 : Une fonction inverse.
Marie : Alors, propose-moi un truc, là, un machin avec des nombres enfin… Donne-moi l'expression de ta fonction, quoi. Tu me dis c'est du genre fonction inverse, ok. Alors, c'est $1/x$, c'est $f(x) = 1/x$?
[…]
Marie : A ton avis ? Non ? Donc il faut que tu bidouilles un peu pour que tu… $1/x$, elle coupe pas l'axe des ordonnées.

Elève B23 : Ah ok.
Marie : Ah voilà. Donc… Débrouillez-vous.

Marie (à d'autres élèves) : Là, vous avez pas envie de dessiner un truc les filles ? Allez-y, prenez votre crayon papier, lâchez-vous.

Marie (à l'élève B32) : Alors, B32, c'est quel genre de truc ça ?
[Inaudible].
Marie : Moi je propose une fonction…

Marie (à un autre élève) : Alors ? Elle te plaît pas ? Tu veux quoi comme histogramme ? Quelle largeur ? Peut-être que je l'ai sous le coude, je sais pas. Bon celui-là te va.
[…]

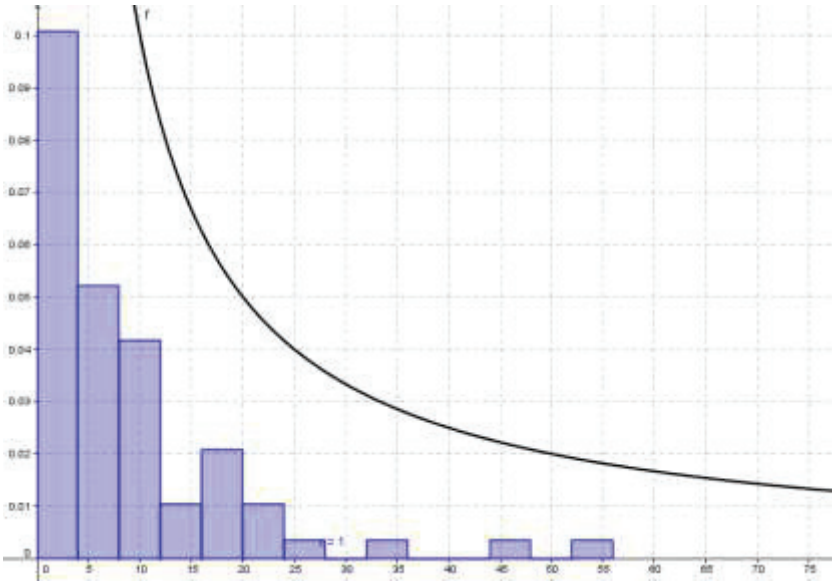
[37:55] Marie (à toute la classe) : Je veux voir une expression de fonction écrite sur toutes les feuilles quand même. Tout le monde est capable de me faire une proposition. En tout cas de forme, d'accord.

[…]
Marie : Ça n'a pas de sens. Qui c'est qui dit des bêtises ? Bravo, ça sert à rien. Encore que…
Elève A42 : C'est juste qu'il me demandait comment on calculait l'aire sous la courbe…
Marie : Ça c'est pas une réponse, ça. La réponse n'est pas une bonne réponse.
Elève A42 : Non, mais c'est…
Marie : Donc ça sert à rien.
Elève A42 : Oui, mais c'est… Il disait…
[…]

Marie (à l'élève A52) : Attends déjà, déjà si on trouvait la courbe.
Elève A52 : Oui mais d'accord…
Marie : On a dit, A52, un problème à la fois.
Elève A52 : Et après ?
Marie : Et après, on va chercher.
[…]
Elève A52 : Mais comment calculer l'aire sous une courbe ?
Marie : On verra ça après. On va s'y mettre. Tu nous fais pas confiance. Tu penses qu'on n'y arrivera pas ?
Elève A42 : T'as pas foi en nous.
Marie : Mais oui, moi j'ai foi. Je sais qu'on va y arriver tous ensemble.

Marie : Alors qu'est-ce que vous proposez là ?
[*Inaudible*]
Marie : Moi non plus je suis pas d'accord. Parce que celle-ci, c'est pas possible.
[*Inaudible*]

Marie : Bien. Allez. […]
[39:30] *Marie s'installe à son bureau pour utiliser GeoGebra.*

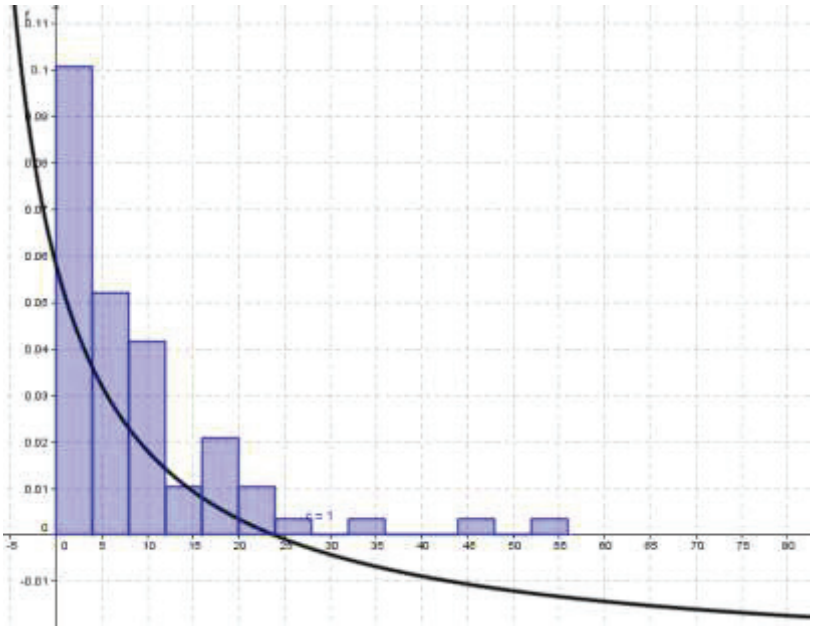


Marie : Alors.
Marie : B53 ?
Elève B53 : Il faut la…
Marie : C'est… T'as écrit ? Alors, première proposition.
Marie efface le tableau.
Marie : Alors, sur beaucoup de feuilles j'ai vu… En fait pour l'instant, je n'ai vu que ça. Proposition une, un truc du genre $1/x$.
Marie écrit au tableau en même temps qu'elle parle.

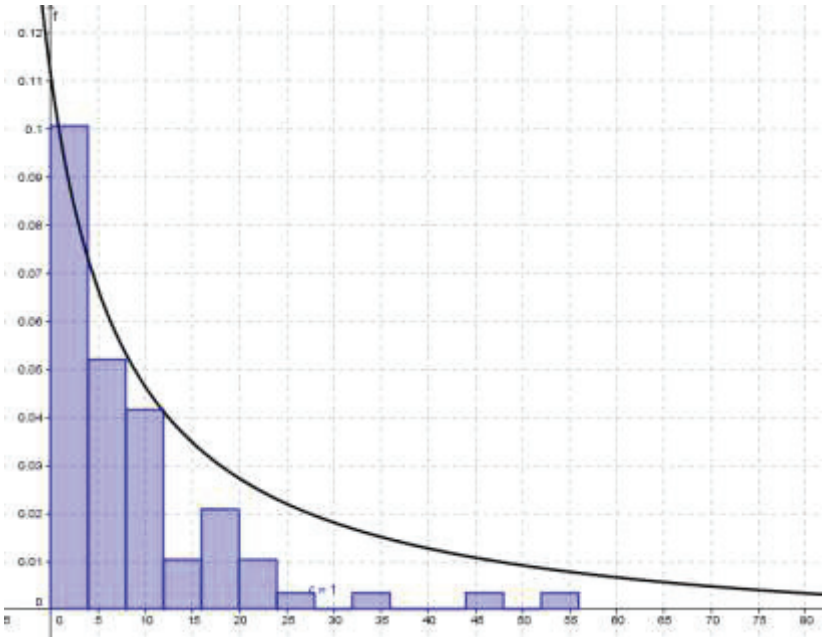
Prop 1

$$f(x) = \frac{1}{x}$$

Une élève ne semble pas d'accord.
Marie : Non, mais c'est une famille… de fonctions. Alors, c'est clair que $1/x$, ça va pas être ça.
Elève B53 : Non.
Marie : Donc tu veux qu'on la…
Elève B53 : Voilà.
Marie : Ca veut dire quoi qu'on la… ?
Elève B53 : On la décale vers le bas.
Marie : Tu veux que je la décale vers le bas.
Marie retourne sur le logiciel GeoGebra.
Marie : Ah oui, c'est sûr, c'est plus facile que sur le papier.
Rires.
Marie : Alors, je la décale… [*Marie déplace la courbe avec le curseur*]. [40:28] Alors oui mais est-ce que je peux la décaler… ? Regardez bien ce qui se passe partout parce que…

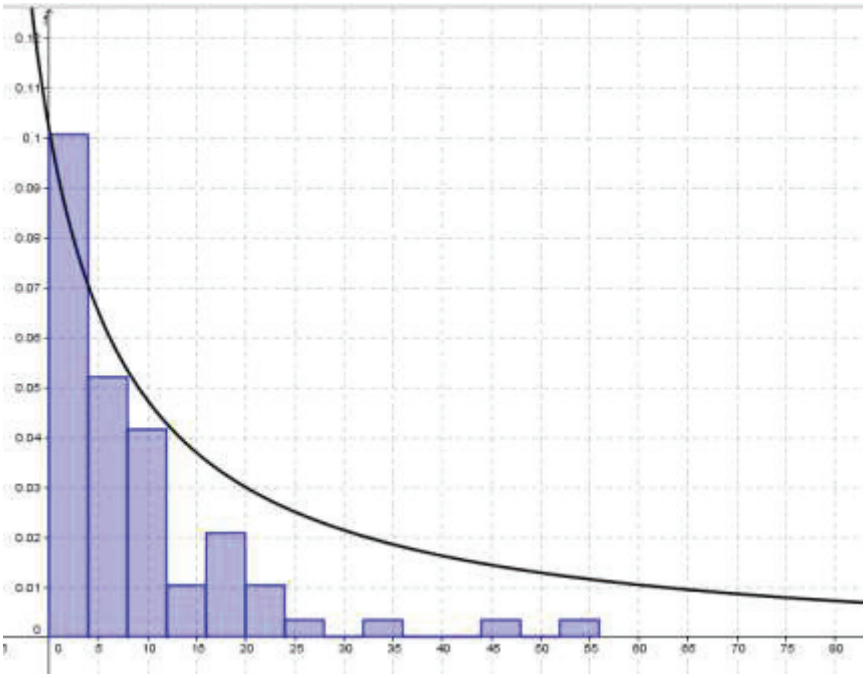


Elève A11 : Non, pas là, pas là.
Marie : Pourquoi je peux plus faire ça ?
Plusieurs élèves : Parce qu'elle est négative.
Marie : Parce qu'elle est négative. On a dit qu'on cherchait des fonctions ?
Plusieurs élèves : Positives.
Marie : Positives. On l'a dit hier, vous vous souvenez les contraintes ? Alors, donc, il faut que je la remonte.

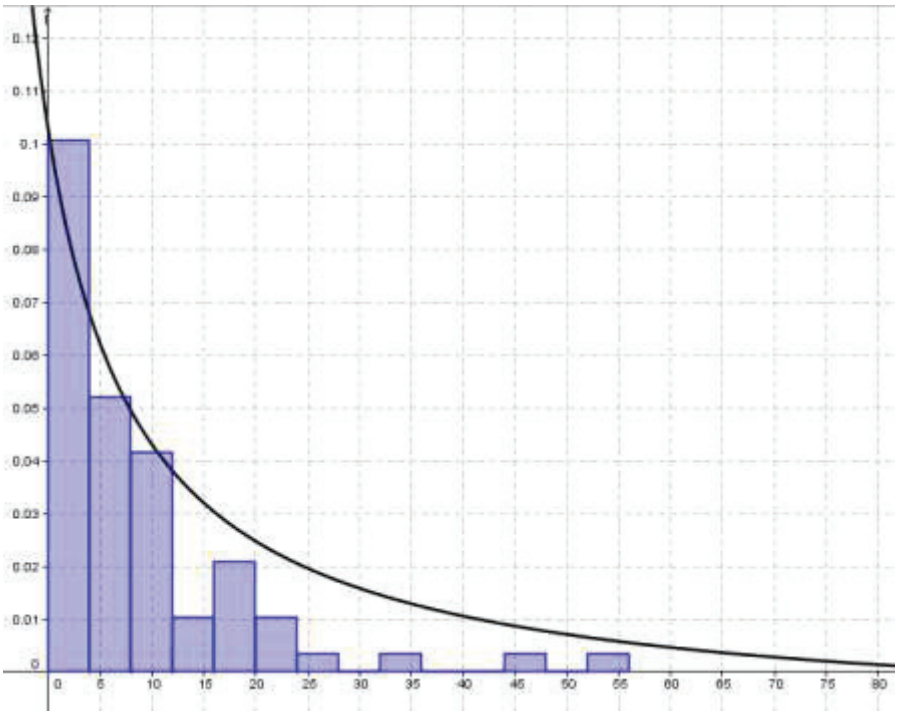


Elève A52 : Ah là, c'est bien.
Marie : Oui mais…
Elève B31 : Faut la décaler sur la gauche.
Marie : Faut la décaler vers la gauche.
Tous les élèves ne sont pas d'accord, certains trouvent que c'est bon, d'autres qu'il faut encore décaler.
Marie : Pas tous en même temps. Attendez. Je vais bouger un truc. Attendez, attendez, attendez. Criez pas.
Voilà.

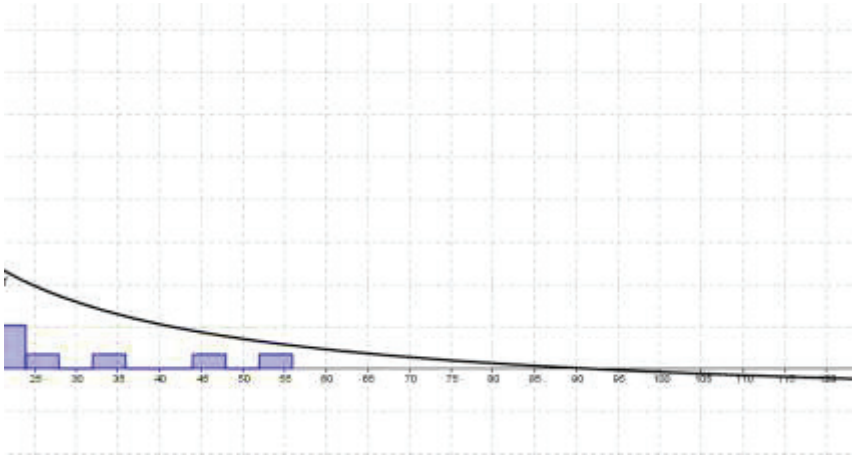
Marie : Alors ?
[...]



Marie : C'est bon là ? J'arrête ? Je continue ? Je fais quoi là ? Vous êtes contents ou pas tant ?
Elève C41 : Il faut un nombre entier. Ah il faut peut-être... Là, c'est bien.
Marie : Là c'est bien ? B53, c'est toi qu'avait proposé ça, tu me dis stop parce que là...
Marie bouge la courbe sur GeoGebra.
Marie : B53 ?
[...]
Marie : Alors, je la reprends. Donc on la laisse au-dessus de l'axe des abscisses quand même et tu veux que je la décale vers la... par... enfin vous voyez ce que je veux dire.
Elève B53. Où est-ce que le maximum...
Marie : Où est-ce que tu veux que je la décale, B53 ?
Elève B53 : A peu près là.
Marie : A peu près là ?
Elève B53 : Oui.



Tous les élèves ne semblent pas d'accord.
Marie : Oui, c'est vrai que c'est pas…
Elève B52 : Oui là c'est bien.
Elève B42 : Moi je dirais que faut la remonter.
Marie : Là, ça vous plait ?
Plusieurs élèves ont l'air d'accord.
Elève A32 : Oui. Moi ça me plait.
[…]
Marie (à l'attention d'une partie de la classe) : […] Est-ce qu'elle, elle vous plait ?
Plusieurs élèves de ce coin de classe ont l'air d'accord.
Marie : C52 ?
Elève C52 : Non, pas tout à fait.
Marie : Pas trop. Pourquoi ?
Elève C52 : Parce qu'elle va peut-être couper… elle va couper l'axe des abscisses si on va très vers la droite.
Marie : Oui, si on va très très très vers la droite tu veux dire.
Elève C52 : Ca va dépasser.
Elève A52 : Oui mais bon…
Marie déplace la courbe pour voir pour des abscisses plus grandes.

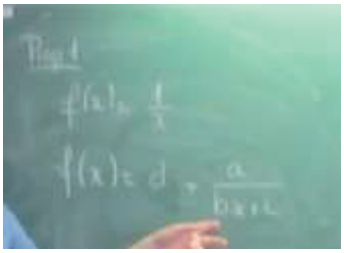


Elève A52 : Il a raison. On voit le métier.
Rires.
Marie : Bon, alors. C'est… Vous…
Marie : Les hyperboles… Enfin c'est pas un secret, ça s'appelle une hyperbole, on est d'accord ? En tout cas, il y a une branche qui nous intéresse.
[42:43] Marie : Il y a que le problème que ça coupe l'axe des abscisses qui vous gêne ?
Plusieurs élèves parlent en même temps.
Marie : Donc on en voudrait une qui coupe pas l'axe des abscisses, qui reste au-dessus… [42:53] C'est quoi la forme générique, là ?
Marie : La forme canonique pour une hyperbole ?
Une élève : Ah…
Marie : Ah. Vous vous souvenez pas ? C'est quoi ? Vous savez pour les suites quand vous voulez… étudier le… B31 ?
Elève B31 : C'est α…
Marie : α. Oui.
Elève B31 : α+β sur…
Marie : Oui alors α et β, on va pas pouvoir faire… On va mettre des lettres… Alors α sur ? Tu te souviens pas ?
Elève B31 : x + …
[43:30] Marie : On va mettre des lettres… $bx + c$ et d. Cette forme-là.
Marie écrit au tableau en même temps.

Prop 1

$$f(x) = \frac{1}{x}$$

$$f(x) = d + \frac{a}{bx + c}$$



Elève A42 : Ah oui.
[43:33] Marie : Bon… Il faut que la limite en +∞, il faut que ça fasse quoi de notre… fonction ?
Deux élèves répondent +∞, d'autres 0.
Marie : Non. Zéro. Si c'est +∞, ça va remonter. Donc si on veut que la limite de f en +∞, ça fasse 0, qu'est-ce qu'il faut qu'on impose ?
Plusieurs élèves parlent en même temps.
Marie : Donc ça, ça tend vers ? Ça, ça tend vers quoi lorsque x tend vers +∞ ? [*Marie montre le dénominateur*].

Plusieurs élèves répondent +∞.
Marie : En tout cas, vers un infini. Ça, ça tend vers ? [*En montrant le quotient*].
Une élève répond +∞, d'autres 0.
Marie : Zéro. Ça, ça tend vers un infini, son inverse vers zéro donc ça tendra vers ?
Plusieurs élèves répondent 0.
Elève A42 : Vers d.
Marie : Vers d. Donc il faut que d soit égal à quoi ?
Plusieurs élèves répondent 0.
Marie : A zéro.
Marie écrit au tableau.

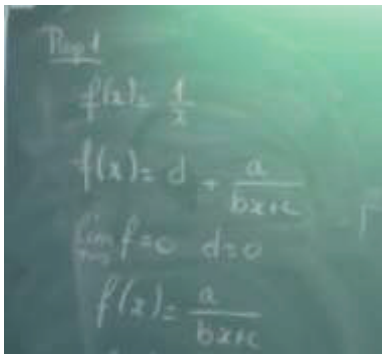
<u>*Prop 1*</u>

$$f(x) = \frac{1}{x}$$

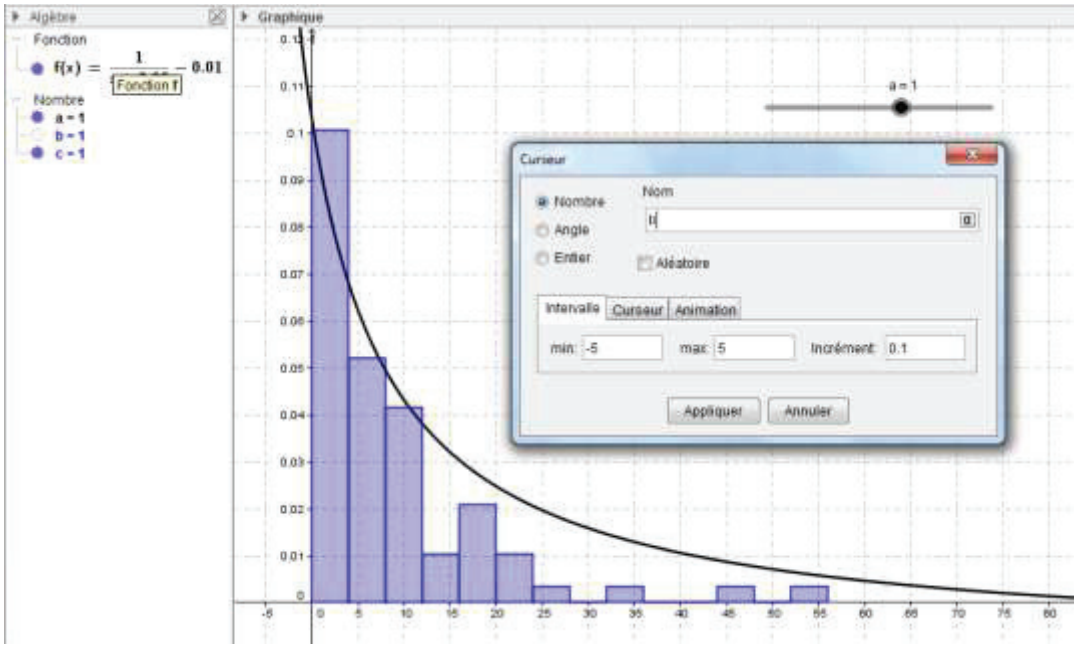
$$f(x) = d + \frac{a}{bx + c}$$

$$\lim_{+\infty} f = 0 \qquad d = 0$$

$$f(x) = \frac{a}{bx + c}$$

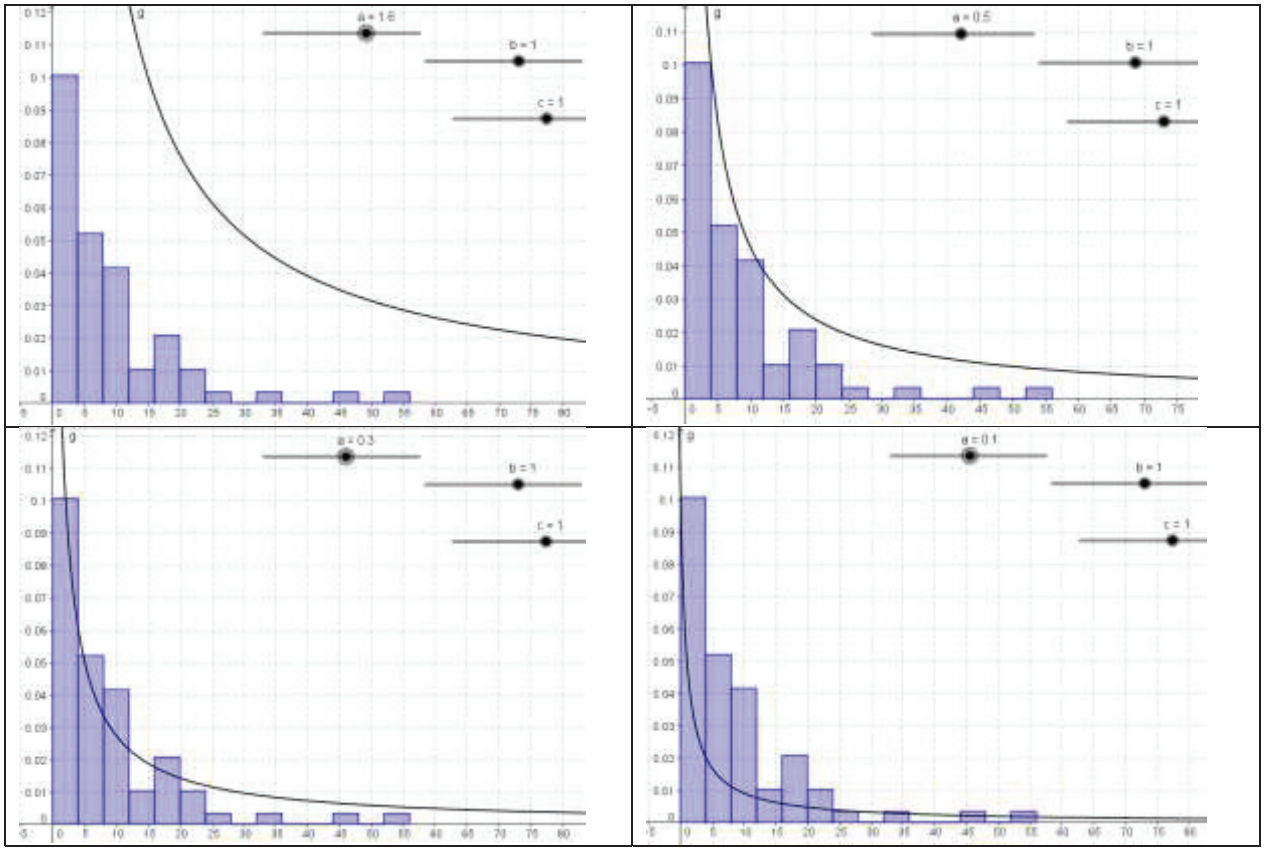


Marie : Donc il reste que ça. [*En montrant le quotient*]. Quelque chose de ce genre. Alors. [44:24] L'avantage du logiciel… […]
Marie manipule le logiciel pour insérer des curseurs.
Marie : Curseur. Alors, attendez, on va l'appeler a.
[…]

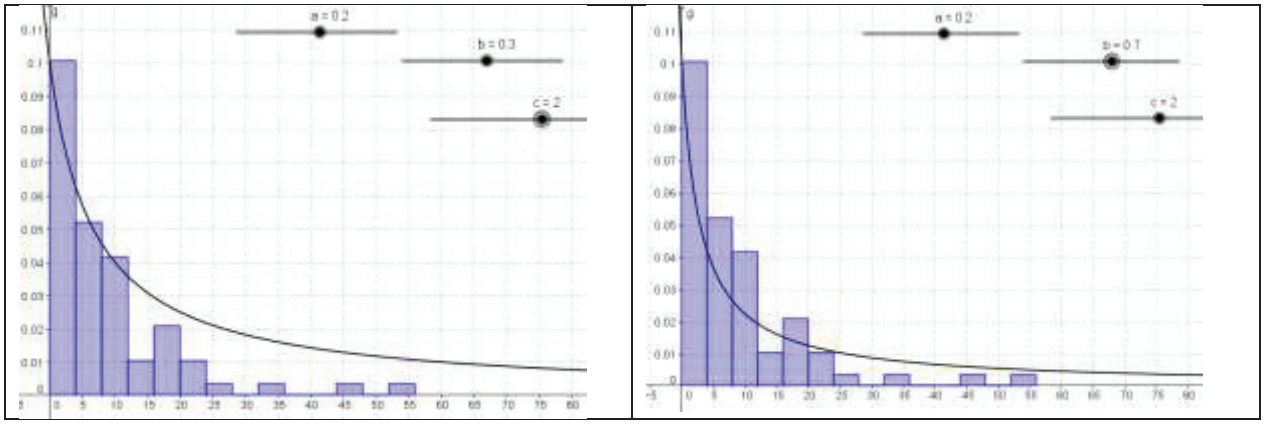


Marie : On y va. Je l'appelle $g(x)$ égale… J'ai mis quoi déjà, dans quel ordre ? J'ai mis a
Un élève dicte.
Marie : a sur $bx + c$.

Marie : On va supprimer celle-ci. Alors ? Qu'est-ce qui se passe quand je bouge a ? On va les garder positifs.
Marie fait varier le curseur a.
[...]
Marie : Ca bouge vite.

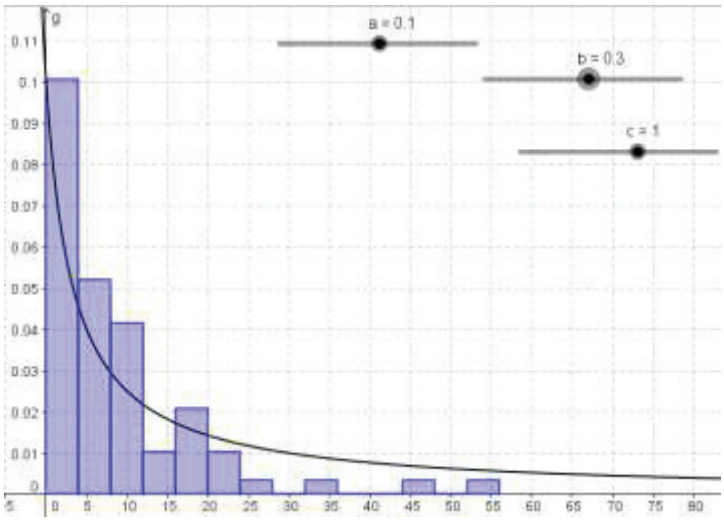


Marie fait varier les différents curseurs.



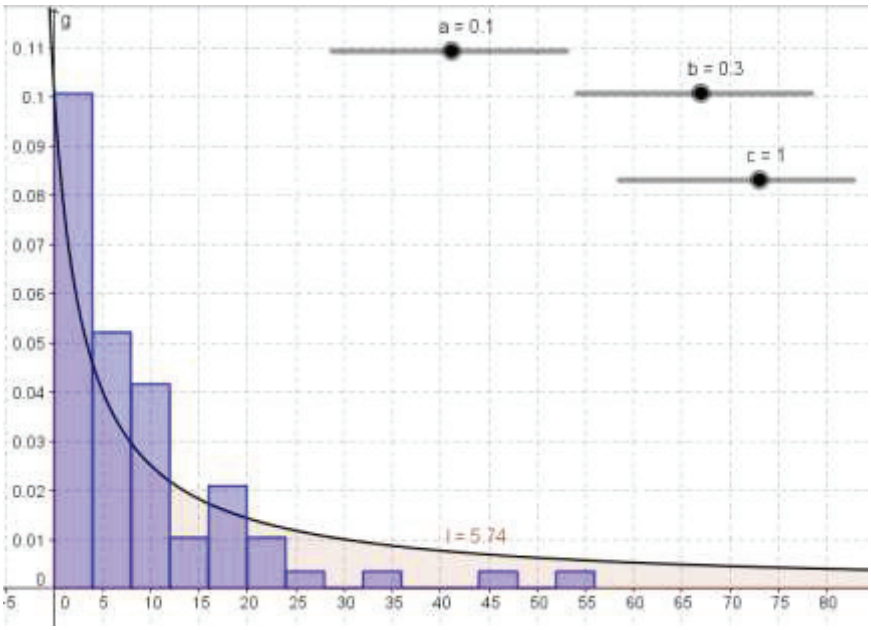
Pendant ce temps, certains élèves ont une discussion autour la fonction tracée.
Une élève : Il y a trop de trucs en haut.
Elève B31 : Non, puisque après c'est jusqu'à $+\infty$.
Elève A52 : Ah. Oui sauf que... il y aura jamais...

Elève C41 : C'est pas grave si elle touche jamais ?
Elève A42 : Il faut des temps d'attente de ouf (sic) avant.
Elève B31 : Oui c'est ça justement. Comme il y aura plein de trucs blancs, ça compense… le bleu.
Elève B52 : Il était peut-être en éruption.
Elève A42 : Oui, je suis d'accord, la fréquence elle peut être… [*Inaudible*].
Marie a terminé de manipuler le logiciel.

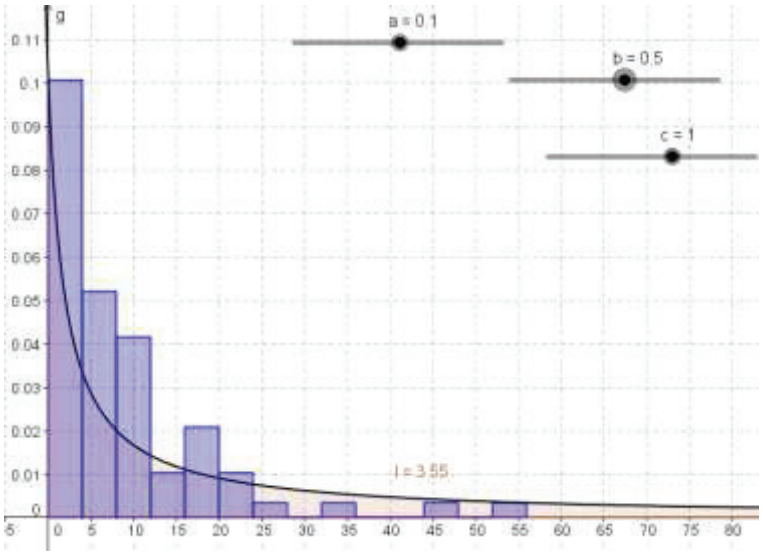


Marie : Alors, est-ce que nos hyperboles on y arrive ou pas ? Est-ce que… est-ce que ça vous satisfait ?
Elève A42 : Il y a beaucoup de trucs au-dessus de la courbe.
Elève A52 : Mais comme B31 a dit.
Marie : J'ai pas entendu ce qu'a dit B31.
Elève A52 : Ah c'est pas grave.
Marie : Qu'est-ce qu'elle a dit ?
[…]
Elève B31 : J'ai dit que c'était pas grave si y avait beaucoup de bleu parce que comme après, ça va jusqu'à l'infini il y aura plein… Donc si il y a du bleu, ça se compensera après.
Marie : D'accord. Le morceau que j'ai mangé ici…
Elève B31 : Oui.
Marie : […] Le morceau que j'ai mangé ici, tu veux dire c'est pas grave parce qu'on le retrouve… C'est ça ? Ça va compenser…
Elève A42 : Oui mais du coup, c'est pas au bon endroit.
[46:28] Marie : Rappelez-moi, il y avait une autre contrainte sur la courbe qu'on cherche quand même. C'est quoi ?
Elève B31 : L'aire c'est 1.
Marie : L'aire doit faire…
Plusieurs élèves répondent 1.
Marie : 1.
Elève A52 : Ah oui, oui. Du coup, ça marche pas si on met $+\infty$.
Elève C41 : Faut que ça touche.
Elève A52 : J'avais raison, alors.
Marie : Ah non, pourquoi tu dis ça ?
Rires.
Elève A52 : Parce que c'est une forme… parce que c'est une forme qui n'est pas finie.
Marie : Et qu'est-ce que… Alors, tu veux dire que parce que…
Elève A42 : On l'a vu, on l'a vu. L'aire elle grandit…. Ah oui, c'était un DM.
Marie : Oui. Il y en a un qui fait les DM.

Rires.
Marie : On a vu, A52, que c'était pas un problème. Qu'on pouvait avoir un domaine infini et une aire…
Elève A52 : C'est pas ça, c'est pas ça. C'est pas une forme parce qu'il n'y a pas de fermeture.
Marie : Oui mais non plus, dans les rectangles y avait pas de fin… On les continuait à l'infini. Moi j'en avais dessiné 4, mais la figure elle était infinie. Pense à ton périmètre du flocon de Von Koch, c'est pareil. Donc c'est pas ça le problème. Alors le logiciel est sympa, parce qu'il calcule les aires sous la courbe.
Marie entre dans la barre de saisie la commande avec intégrale (mais les élèves ne voient pas cette commande).
Marie : Alors, on va mettre un truc très très grand. D'accord ?
Elève A52 : Mais comment on peut calculer l'aire d'un truc comme ça ?
Marie : Un problème à la fois, A52. Moi je suis mono-tâche.



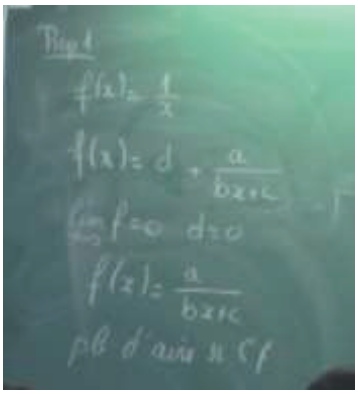
Marie : Ah, regardez, il est là. J'ai calculé l'aire entre 0 et… je sais pas quoi. 10 millions, 100 millions. Et ça fait 4… 5,74 déjà.
Marie : Donc ?
Elève B31 : Faut baisser encore.
Marie : On baisse encore.
Marie manipule à niveau le logiciel.
[…]
Elève B53 : En fait faut qu'on trouve l'équation pour que ça fasse 1.
Elève B52 : Ah oui, on pourrait résoudre une inéquation.
Petit bug avec le logiciel.
[…]
Marie : Tu veux que je la baisse.



Elève B52 : Oui mais là ça devient n'importe quoi.
Elève B31 : Là on est à combien là ?
Plusieurs élèves répondent 3,55.
Marie : 3,55 c'est pas 1.
Plusieurs élèves parlent en même temps.
Elève B52 : On se rapproche de 1. Elle tend vers 1.
Marie : Non mais attendez, elle s'approche de 1, mais moi j'ai mis jusqu'à 100 millions. Si je mets 10^{12}, qu'est-ce qui va se passer à votre avis ?
Plusieurs élèves répondent en même temps : « Ça sera encore plus », « Plus »,…
Marie : Ça sera trop grand. Ça va dépasser 1.
Marie complète le tableau.

Prop 1

$$f(x) = \frac{1}{x}$$
$$f(x) = d + \frac{a}{bx + c}$$
$$\lim_{+\infty} f = 0 \qquad d = 0$$
$$f(x) = \frac{a}{bx + c}$$
pb d'aire ss C_f



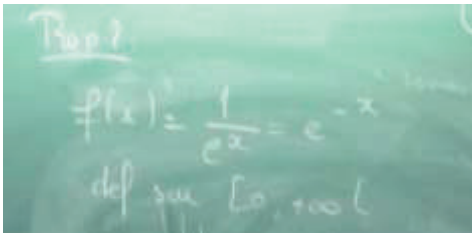
[48:52] Marie : Est-ce qu'il y avait d'autres propositions de fonctions ?
Une élève : Non, mais.
Marie : Tout le monde a pensé à une fonction inverse ?
Plusieurs élèves parlent en même temps.
Marie : Je sais pas, C42 dénonce C41.
Elève C41 : J'avais pensé à la fonction exponentielle. Parce qu'elle est toujours supérieure à 0 et… elle est positive. Enfin elle est toujours positive et continue. Mais quand je teste $1/e^x$ *[sur sa calculatrice]*, elle… elle… elle va de l'autre côté. Enfin de…
Marie : Elle va de l'autre côté ?
Marie commence à écrire au tableau.
Marie : Alors proposition 2, tu nous proposes $f(x) = \frac{1}{e^x}$.

Elève C41 : Oui.
Marie : Ça s'écrit comment ça aussi ?
Elève C41 : ...
Elève C42 : e^{-x}.
Marie : e^{-x}. Alors, qu'est-ce qu'elle... qu'est-ce qu'elle va de l'autre côté, là ?
Elève C42 : En fait, elle fait comme ça. Enfin, elle passe de l'autre côté des ordonnées.
Marie : On peut la définir seulement sur [0 ; +∞[.
Elève C42 : Ah oui.
Marie : On peut se restreindre à l'intervalle qui nous intéresse. On définit sur [0 ; +∞[.

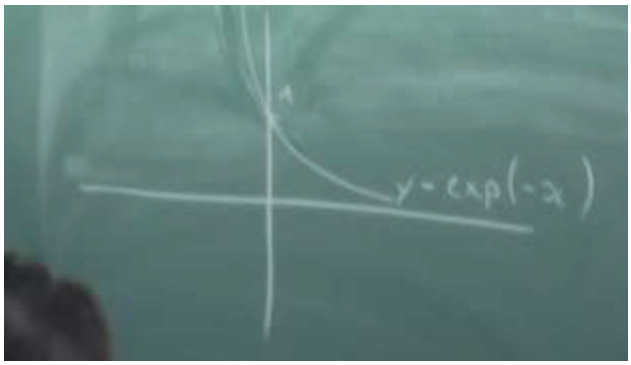
Prop 2

$$f(x) = \frac{1}{e^x} = e^{-x}$$

$$def\ sur\ [0; +\infty[$$



Marie : Alors ça donne quoi son...
Marie s'installe à l'ordinateur.
Marie : Si on prend sa... Alors attendez, je vais effacer tout ça sinon, on va plus rien voir... à un moment. [...]. Alors, $f(x)$ égal exp... ah non, pardon, de $-x$. [*Marie écrit en même temps dans la barre de saisie de GeoGebra*].
La courbe n'apparait pas sur l'écran car la fenêtre n'est pas adaptée.
Une élève : Elle est confondue ?
Marie : Non, non, c'est parce qu'elle...
Un élève : On aura le même problème, madame.
Marie : Qu'est-ce que vous... ?
Elève B31 : Faire comme avec...
Marie : Oui, on pourrait faire...
Marie manipule le logiciel pour faire apparaître l'aire sous la courbe.
Marie : Alors, là, c'est pas convaincant.
Marie : Qu'est-ce que vous savez sur l'exp$(-x)$. Vous connaissez l'allure de la courbe.
Marie : Ca fait quoi ?
Quelques élèves parlent entre eux.
Marie : Comment ça tout droit ? J'ai peur. B31 ?
L'élève B31 fait le tracé de la courbe dans l'air.
Marie : Oui c'est ça. Tu peux me dire au moins à peu près quelque chose où ça passe ou pas ? un point qu'on connait ou... ?
Elève B31 : En 0.
Marie : En 0, ça vaut ?
Elève B31 : 1.
Marie : 1. Donc une exponentielle, elle est comme ça. Enfin... elle est ratée celle-là.
Marie trace la courbe au tableau en même temps, puis efface une partie et la refait.



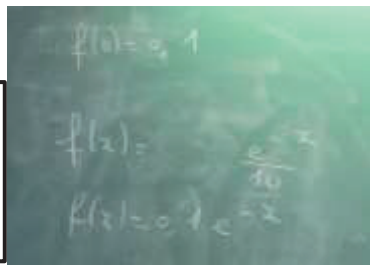
Marie : On est d'accord ? Bon alors, nous, on veut pas qu'elle passe en 1 ? Ça va être trop haut pour nous. C'est pour ça que là, elle est pas… [*en parlant du vidéoprojecteur*].
[51:28] Marie : Donc faut qu'on la fasse passer… Vous voulez qu'elle passe par où ?
Plusieurs élèves répondent en même temps.
Marie : 0… Vous voulez que $f(0)$ soit 0,1. Non ? Oui, A42 ?
Elève A42 : Mais après ça va passer sur l'axe des abscisses.
Marie : Sur quoi on peut jouer encore, alors ? Donc, déjà, on veut que $f(0)$ soit alors disons 0,1.
Marie écrit au tableau en même temps.

$$f(0) = 0{,}1$$

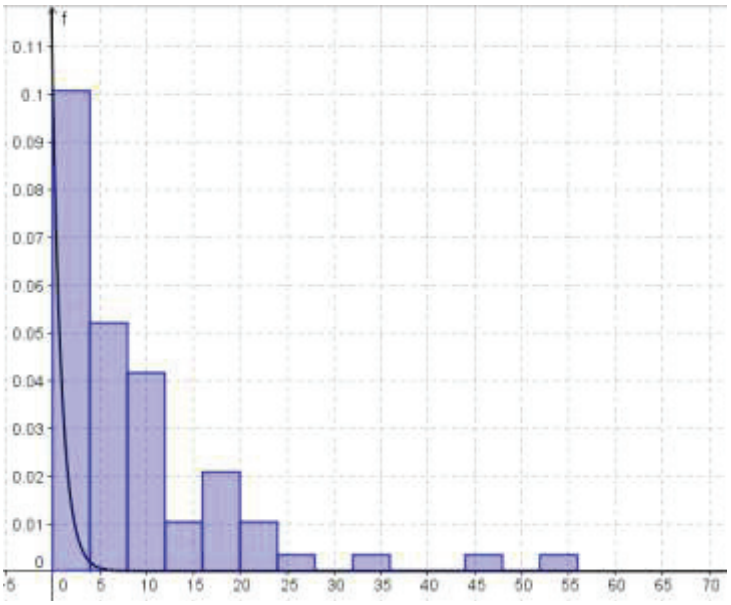
Puis Marie retourne sur l'ordinateur et manipule le logiciel.
Marie : Alors, allons-y.
[…]
Marie : Alors qu'est-ce que vous proposez ?
Marie bataille avec le logiciel.
[…]
Marie : Alors, sur quoi on peut jouer pour que ça aille mieux ?
Marie : Alors, déjà le… le $f(0)$ égal 0,1 comment vous faîtes ?
Elève B31 : On place un point.
Marie : Oui, mais sur l'expression, comment vous faîtes pour qu'en 0, ça fasse à 0,1 ?
Marie : B31 ?
Elève B31 : On fait moins 0,9.
Marie : Qu'est-ce qu'on fait moins 0,9.
Elève B31 : $y = e^{-x} - 0{,}9$.
Marie : D'accord. Alors, à ce moment-là, qu'est-ce qui va se passer pour la courbe ?
Elève B31 : Elle va descendre.
Marie : Elle va descendre, comme ceci. [*Marie montre la courbe de cette fonction sur le logiciel*]. D'accord. Alors, qu'est-ce que vous… Qu'est-ce qui vous plaisait pas non plus… ? […]
Un élève : La limite en $+\infty$, elle sera pas à 0. Elle sera en moins…
Marie : Oui. Alors, comment on change différemment ? B32 ?
Elève B32 : En divisant par 10.
Marie : Alors en… en divisant par 10.
Marie commence à écrire au tableau.
Marie : Donc, ton $f(x)$, ça fait $0{,}1 \times \mathrm{e}^{-x}$. Alors, diviser par 10, c'est multiplier par 0,1.
Elève B32 : Oui. [*Très discret*].
Rires.
Marie : Pardon. Ok.
Elève B32 : Oui.

Marie : En divisant par 10…
Marie réécrit au tableau.

$$f(0) = 0{,}1$$
$$f(x) = \frac{e^{-x}}{10}$$
$$f(x) = 0{,}1e^{-x}$$



Marie : Bon, est-ce que celle-ci $f(0)$ vaut 0,1 ? On va la tracer ? On regarde ce que ça donne ?
Marie écrit l'expression de la fonction dans la barre de saisie.
Un élève : Ça marche.
Marie : Ah ça je sais pas si ça marche.
La courbe apparait.



Marie : Bon alors, ça passe bien… Vous la voyez ou pas ?
Marie essaie de rendre plus visible la courbe.
Marie : Bon, alors, plus rien ne marche quand je sélectionne des trucs mais c'est pas grave on va dire que tout va bien. Elle passe bien par celui qu'on voulait, mais alors qu'est-ce que… Qu'est-ce qui va pas là ?
Marie : B32 ?
Elève B32 : Elle est… elle tend vers 0 dès…
Marie : Trop vite. Oui, c'est ça. Elle… elle chute… tragiquement. Alors, sur quoi on peut jouer ?
Marie : On connait que $exp(-x)$, nous ?
La sonnerie retentit.
Marie : Attendez, bougez-pas parce que… B31 ?
Elève B31 : *[Inaudible].* C'est un nombre inférieur à 1 fois x.
[55:10] Marie : Alors, pour vendredi, vous prenez votre cahier de texte […].
Marie : Alors, pour vendredi, je vais vous envoyer par mail les fichiers, d'accord ? Les histogrammes *GeoGebra*. Et donc, pour vendredi, vous essayez d'affiner notre fonction, là. D'accord ? Pour me faire une proposition. Alors, on se met d'accord, on choisit tous qu'elle passe par 0,1 et après le problème, est-ce que tout le monde a compris le problème de celle-ci ? C'est qu'elle descend…

842 *Plusieurs élèves répondent « Trop vite ».*
843 Marie : Trop vite. Donc faudrait essayer qu'elle descende en suivant mieux nos rectangles. On est
844 d'accord ? Alors, chez vous, vous avez tous *GeoGebra*, bien sûr. Je vous envoie les fichiers… dans la fin
845 d'après-midi. D'accord ? Pour pas que vous ayez besoin de faire les histogrammes, classer tout ça. Jouez
846 avec *GeoGebra*. Je vous rappelle qu'il y a des curseurs dans *GeoGebra*. Et faites-moi une proposition de
847 fonction qui vous parait convenable. D'accord ? Ok, c'est bon ?
848 […]
849 Marie : Vous me remettez au propre aussi vos notes dessus.
850 […]
851 Marie : Et donc, A52, après, il nous restera encore un problème à régler, on est bien d'accord… une fois
852 qu'on aura trouvé notre fonction. Donc on fera ça, vendredi. D'accord ?

Marie a envoyé un mail à ses élèves le jour même, avec le fichier GeoGebra de l'histogramme avec des amplitudes de 4 ans.

```
Pour chercher votre courbe de tendance, je vous envoie le fichier
géo-gébra (histogramme avec un pas de 4 années), je le  mets aussi
dans educ horus.
Pour calculer l'aire sous la courbe avec le logiciel, tapez dans la
ligne de saisie : I=Intégrale[le nom de la fonction, borne inf, borne
sup]. Mettez une borne sup très grande ( 10 000 000).
N'oubliez pas de répondre aux questions de départ : Evaluez
probabilité que la prochaine éruption ait lieu dans les 5 ans à venir,
au cours de l'année 2030?
```

I.4 Transcription de la séance 4

Transcription de la séance 4 du 20/03/15 (1h50 – classe entière)

34 élèves présents.

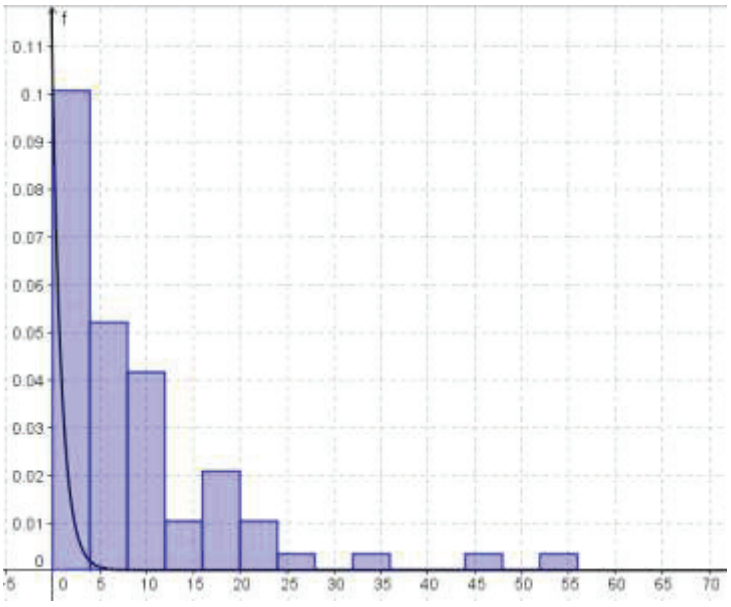
[00:00] *Sonnerie. Mise en place des élèves.*
[01:39] Marie : Bien. Chut. Tout le monde ressort son cahier d'exercices.
Marie : Chut. Alors je vais avoir besoin de votre attention. […]
Marie : Alors.
Marie : Alors quelqu'un peut-il… Alors, on n'y est pas tous encore. Alors, tout le monde a son cahier d'exercices devant lui, ouvert pour le problème 2 ? Alors, qu'est-ce que vous deviez faire pour aujourd'hui ?
Marie : On va attendre le silence complet pour démarrer.
Marie : A42.
Elève A42 : On devait conjecturer…
Marie : J'entends rien.
Elève A42 : On devait conjecturer l'équation de la courbe de tendance de l'histogramme et ainsi essayer de répondre aux questions.
Marie : D'accord. Est-ce que tout le monde est d'accord ?
Plusieurs élèves répondent « oui ».
Marie : Alors… Qu'est-ce que vous avez donc… On l'avait appelé f, j'imagine. Donc proposition. [*Marie écrit « Prop » au tableau*].
Marie : Alors, est-ce que vous avez des propositions à me faire ? Tout le monde devait me faire une proposition.
Silence
Marie : B23 alors, qu'est-ce que tu proposes ?
Elève B23 : Alors… [*Inaudible*].
Marie : Vous voyez pas l'écran ? Si ça va ? B23.
Elève B23 : C'est l'opposé de la... de la fonction exponentielle.
Marie : L'opposé de la fonction exponentielle. [*Marie écrit au tableau*].

$\underline{Prop}$:
$f(x) = -e^x$

Marie : Donc ça ?
Elève B23 : Non c'est devant x.
Marie : Ah, oui donc c'est pas…
Plusieurs élèves disent « l'inverse ». Marie modifie l'expression de la fonction f.

$\underline{Prop}$:
$f(x) = e^{-x}$

Marie : Oui, mais ça on avait déjà… Alors, tu veux que je la trace ?
Marie trace sur GeoGebra la courbe correspondante.



Marie : On l'avait déjà testée celle-ci. [...] B23, tu as pas cherché ?
Elève B23 : Non.
Marie : B11.
Elève B11 : J'ai trouvé avec les curseurs.
Marie : Donc tu as utilisé un curseur ?
Elève B11 : Oui.
Marie : On avait déjà dit quelque chose quand même, on avait imposé que $f(0)$ soit égal à...
Quelques élèves répondent « 0,1 ».
Marie : 0,1. On avait imposé $f(0) = 0{,}1$. [*Marie écrit au tableau en même temps*].

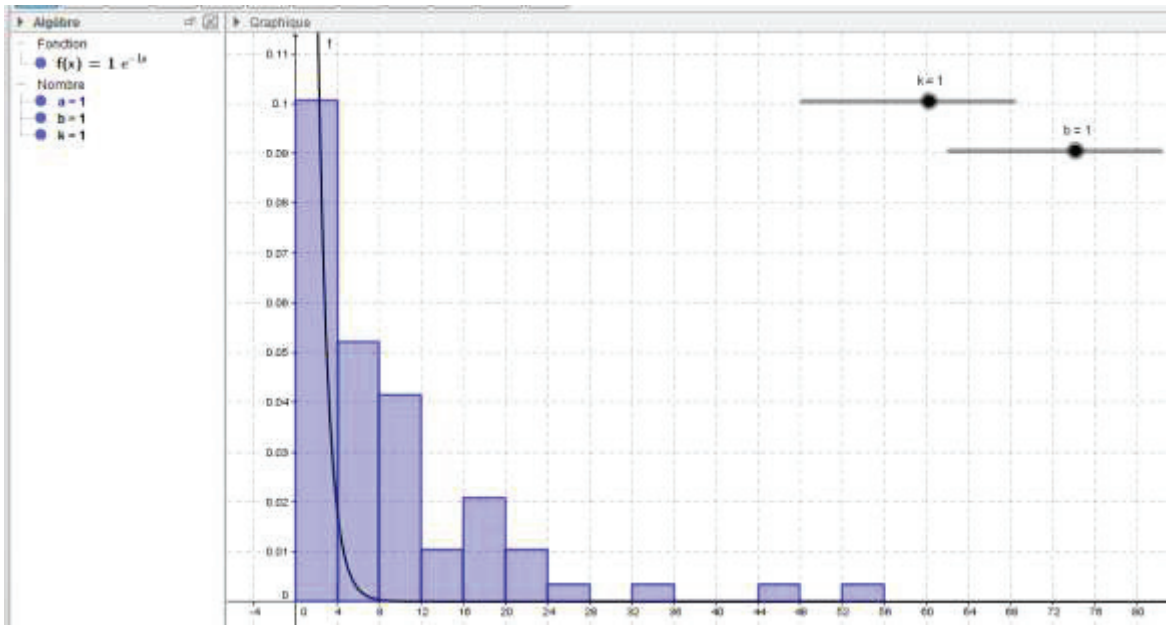
Prop :
$f(x) = e^{-x}$
$f(0) = 0{,}1$

Marie : Tu as pris deux curseurs ou... ?
Elève B11 : Oui j'en ai pris deux.
Marie : T'en as pris deux , d'accord. Donc du coup... Ok. Alors moi je l'ai appelé k. [*Marie crée les curseurs sur GeoGebra*]. Du coup l'expression que tu as... Attendez pardon. L'expression que tu as choisi c'était quoi ? On va appeler b le suivant. T'as cherché quoi ? Sur Geogebra, t'as choisi ?
Elève B11 : J'ai mis un curseur qui permettait de trouver le truc avant l'exponentielle.
Marie : Le truc avant l'exponentielle.
Elève B11 : Oui.
Marie : Alors du coup je vais l'appeler b. Donc tu cherches une constante fois une exponentielle...
Elève B11 : Moins une autre constante.
Marie : D'accord. kx avec les... deux curseurs... [*Maire écrit au tableau en même temps*].

Prop :
$f(x) = e^{-x}$
$f(0) = 0{,}1$
$f(x) = be^{-kx}$

Marie : Qui a choisi aussi de... d'utiliser les curseurs là, pour tester les fonctions ?
Cinq élèves lèvent la main.
Marie : Les autres ont fait leur travail ou ? B32, comment tu as fait toi ?
Elève B32 : J'ai testé avec...

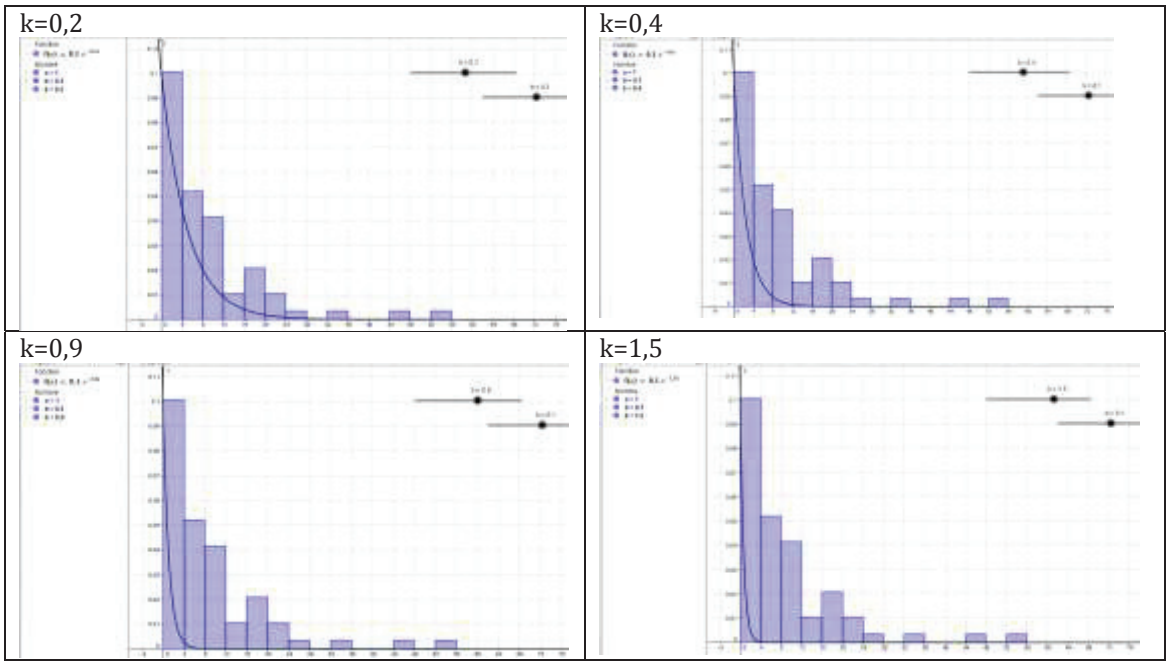
Marie : Pour me faire une proposition honnête.
Elève B32 : J'ai testé avec des valeurs.
Marie : A la main ?
Elève B32 : J'ai rentré…
Marie : Oui c'est ça que je veux dire. Tu as tapé, tu as ajusté… T'as pas pensé à utiliser les curseurs ?
Elève B32 : J'ai oublié.
Marie : D'accord. Alors notez-le quand-même. Là effectivement. Alors du coup on va faire ce que dit… Est-ce qu'il y a eu d'autres méthodes ? Alors attendez… Tant pis. Effacer. On va l'appeler $f(x)$. Qu'est-ce que j'ai mis au tableau s'il-te-plaît B11 ? Ah oui.
Elève B11 : be^{-kx}
Marie : be^{-kx}. Il y a des bavardages !
Marie : Ah, qu'est-ce que j'ai tapé ? Voilà ! Donc on reconnaît l'allure d'une exponentielle décroissante.



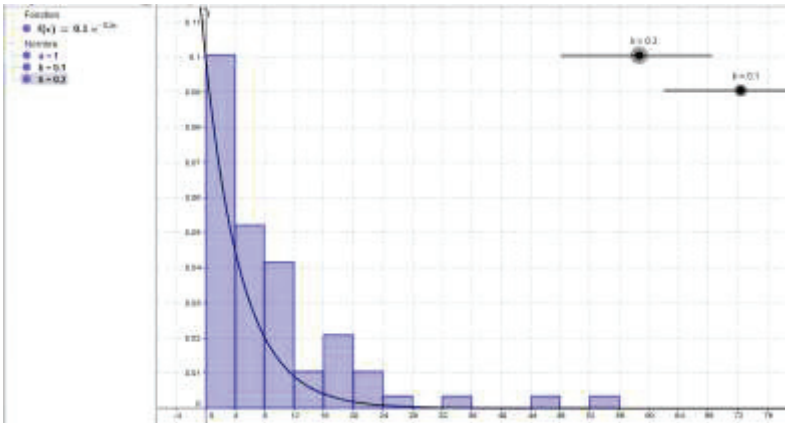
Marie : N'oubliez pas que… Donc $f(0)$ pour celle-ci elle vaut quoi si tu as pris ça, B11 ? Donc si on choisit ce type de fonction, quand tu calcules $f(0)$… [*Marie commence à écrire au tableau*].
Elève B11 : Ça fait b.
Marie : Ca fait b, et comme on voulait que ce soit 0,1.
Elève B11 : b c'est 0,1.

$$f(0) = b = 0{,}1$$

Marie : Est-ce qu'on avait besoin vraiment de ce curseur ?
Elève B11 : Non mais je l'ai mis au cas où.
Marie : Oui. Non mais, c'était juste… Alors peut-être que b je vais le mettre tout de suite à 0,1.
Un élève répond « oui ».
Marie : Ah moi je fais ce que vous voulez.
Plusieurs élèves répondent « oui ».
Marie : Si vous voulez pas je le fais pas. Alors évidemment j'ai encore fait une bêtise. Alors allons-y. On met b à 0,1. [*Marie met le curseur b sur 0,1*]. Voilà, donc on n'a plus qu'à faire bouger k alors ?
Elève B11 : Moi j'ai mis à 0,1.
Marie : Alors tu l'as pas mis tout de suite... Mais alors quand on le bouge comme ça [*Marie augmente les valeurs du curseur k, positif*], vous voyez ce que ça fait sur la courbe ? Est-ce que tout le monde voit ?

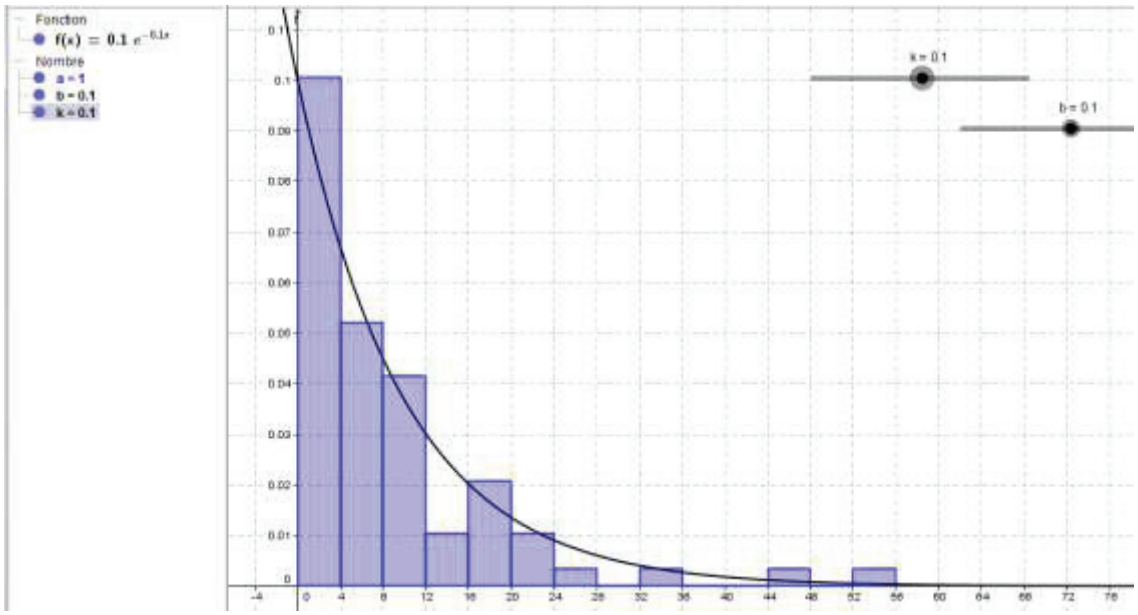


Elève A42 : Ca se rapproche de zéro.
Marie : Oui. Ça s'aplatit très très vite en fait on va dire. Alors du coup il faut le… le baisser. [*Marie diminue la valeur de k*]. Donc tu nous dis que tu l'as mis… Ça a l'air pas mal…



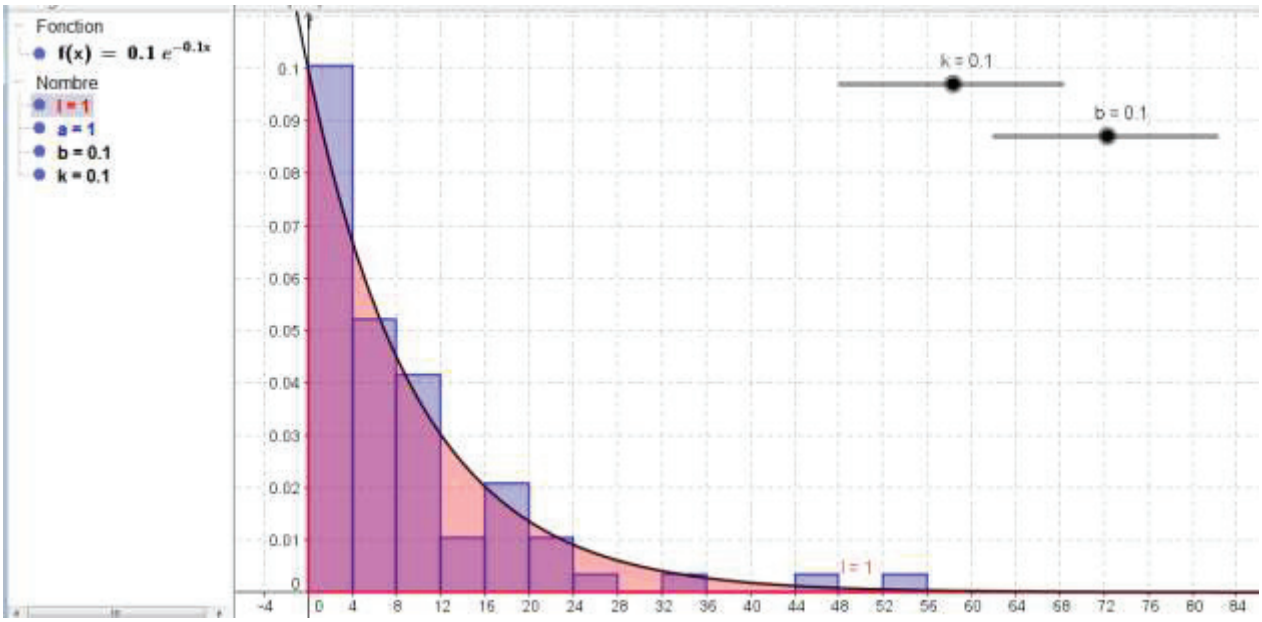
k=0,2

Elève A42 : On peut faire mieux.
Marie : On peut faire mieux. Bon.



k=0,1

Elève B11 : Moi je l'ai mis comme ça.
Marie : Alors pourquoi ? Parce que si je fais ça c'est pas mal non plus ? [*Marie remet la courbe avec k=0,2*].
Elève B31 : Mais il faut avoir l'aire sous la courbe.
Marie : Ah il faut avoir l'aire sous la courbe. Donc vous avez vérifié avec la deuxième contrainte ?
Elève B31 : Oui.
Elève B11 : Ah oui. J'ai pas fait comme ça.
Marie : Ah tu l'as pas fait… Alors pourquoi tu t'es arrêté à 0,1 ?
Elève B11 : Parce que je me suis dit que ça coupe à peu près tous les rectangles, donc…
Marie : Oui ça te paraissait visuellement bien. D'accord. OK. C'est un argument comme… Pour l'instant on en est là, donc voilà… Mais il y avait aussi, on pouvait vérifier, avec l'aire sous la courbe. D'accord ? Car c'est ce qui nous a permis de rejeter les inverses. D'accord ? Donc effectivement du coup je vous avais donné la commande par mail. Vous avez réussi tous à ouvrir le fichier pour faire marcher ça chez vous ?
Plusieurs élèves répondent « oui ».
Marie : Donc si on calcule l'intégrale… enfin l'aire sous la courbe pardon. [*Marie utilise la commande Integrale pour calculer l'aire sur le logiciel*]. Donc je l'ai appelé f. Alors entre zéro et on va prendre une borne très très grande. Alors il est là I. Vous le voyez ? Alors normalement, ça se colorise (sic). Non c'est l'opacité qu'il faut que je change.
[…]
Marie : Alors donc ça se colore quand je fais ça bien. Et donc voilà vous la voyez apparaître l'aire. Alors on a calculé jusqu'à je sais plus quoi. Et donc ça fait 1.



Une élève : C'est bien.
Marie : Oui c'est bien. Alors attention c'est pas une preuve que l'aire sous cette courbe de zéro à plus l'infini fait ?
Plusieurs élèves répondent « 1 ».
Marie : 1. On est bien d'accord ? Mais enfin pour ce qu'on a… notre travail c'est déjà pas mal. Est-ce qu'il y en a qui ont testé d'autres fonctions exponentielles décroissantes ? Non ? Tout le monde était d'accord pour dire que celle-ci c'était bien ? Qui a trouvé… qui proposait celle-là ?
Une dizaine d'élèves lèvent la main.
Marie : Et les autres alors ? C22, tu faisais quoi comme proposition pour aujourd'hui ?
Elève C22 : J'avais pas réussi… J'ai pas réussi à…
Marie : T'as pas réussi à faire quoi ?
Elève C22 : A faire que ça marche.
Marie : Tu pouvais faire comme B32, à l'arraché, saisir des nombres là et puis voir… T'as pas été assez courageuse ?
Elève C22 : Non.
Marie : T'as fait un deux et puis tu t'es arrêtée ? Et les curseurs, jeunes gens, on l'a utilisé déjà. C12, qu'est-ce que tu proposais toi pour aujourd'hui ?
Elève C12 : Ça.
Marie : Celle-là ?
Elève C12 : Oui.
Marie : Alors on reste sur celle-ci. Est-ce que tout le monde est d'accord ou pas ?
Marie écrit au tableau.

$$f(x) = 0{,}1e^{-0{,}1x}$$

Marie : Adjugé vendu. Oui ? Donc $f(0)$ vaut 0,1 et avec le logiciel, vous mettez avec… Vous n'oubliez pas que vous prenez des notes ? C'est quand-même pas… Et avec le logiciel il semblerait que l'aire sous la courbe entre zéro et plus l'infini fasse bien… ?
Plusieurs élèves répondent « 1 ».
Marie : 1. Est-ce que ça gène quelqu'un que l'aire sous cette courbe fasse 1 ? Je repensais à la remarque de A52 là. Tu es d'accord, A52 ? Même si le domaine est infini… On a déjà rencontré des domaines infinis qui avaient des aires.
Elève A42 : Finies.
Marie : Finies. D'accord ? Alors on le démontrera plus tard dans le cours en fait. Comme ça je convaincrai complètement B52 que cette aire vaut 1.

[10:33] Marie : Alors comment vous avez répondu du coup aux deux petites questions enfin aux questions qui étaient les questions de départ ? Qui étaient alors donc je vous rappelle que le volcan Aso est actuellement en éruption, alors je crois que c'est fini malheureusement aujourd'hui.
Elève A42 : Non mais c'était en janvier.
Marie : Pas de chance. Comment ?
Elève A42 : C'était en janvier l'éruption.
Marie : Non. Ça a commencé en décembre et c'était au début mars c'était toujours en éruption. Parce que une éruption, enfin ça dure sur un certain laps de temps quand même.
Elève A42 : Du coup on peut dire que c'était...
Marie : Oui on va dire qu'aujourd'hui ça l'est encore. Ça nous arrange.
Rires.
Marie : Donc on avait posé la question quelle est… d'évaluer la probabilité que la prochaine éruption ait lieu au cours des cinq ans à venir. Alors A21 ?
Elève A21 : On calcule l'aire sur l'intervalle [0; 5].
Marie : Oui alors et est-ce que tu peux me dire ce qu'on a…
[…]
Marie : Est-ce que tu peux me dire ce qu'on calcule en termes de notation ? On lui avait donné un nom à cette variable aléatoire, non ?
Plusieurs élèves répondent « X ».
Marie : X. Donc X est dans quoi ? [*Marie commence à écrire au tableau*].
Elève A21 : [0; 5].
Marie : Donc appartient à l'intervalle [0; 5]. [*Marie écrit au tableau en même temps*].

$$P(X \in [0;5]) =$$

Marie : Et donc ce sera… Redis ce que tu as dit.
Silence.
Marie : Redis ce que tu viens de me dire.
Elève A21 : J'ai encore rien dit.
Marie : Alors comment t'as fait ? Excuse-moi.
Grégoire : Non. Avec intégrale de f de 0 à 5.
Marie : Avec le logiciel ? D'accord. Alors. Ok.
Un élève : Moi j'ai fait sans le logiciel.
Marie : Alors je vois pas très bien ce que je fais aujourd'hui encore. […]. Oui, évaluer avec le logiciel entre 0 et 5. Et on trouve à peu près zéro…
Elève A21 : 0,39.
Marie : 39. [*Marie écrit au tableau*].

$$P(X \in [0;5]) \approx 0{,}39$$

Marie : Donc comment est-ce qu'on a trouvé ça là ?
Elève B31 : Avec la formule. Enfin avec la fonction.
Marie : Comment ?
Elève B31 : Avec la fonction.
Marie : Comment ?
Elève B31 : En rentrant l'intervalle. Enfin, pour trouver l'aire.
Marie : Avec le logiciel on a évalué l'aire. D'accord ? Pour avoir une valeur approchée.
[12:37] Marie : Et l'autre question c'était quoi ? Au cours de l'année 2030. Alors d'abord pourquoi j'avais posé cette question ?
[12:41] Marie : Est-ce que ça on pouvait… Vous auriez pu me faire des propositions là-dessus quand je vous ai donné le premier document ?
Plusieurs élèves répondent « non ».
Marie : Non, toujours pas ?
Elève B31 : Mais si.
Marie : D'accord on y reviendra. A21 ?

Elève A21 : Si, en étudiant le, enfin, en étudiant le nombre de fois où il est rentré en éruption d'une année sur l'autre.
Marie : Alors.
Elève A21 : Si on considère qu'il est en éruption… Non, dans les cinq, dans les cinq années prochaines.
Marie : Oui. C'est quelque chose… ça m'a étonné mardi que vous me le proposiez pas. C'est-à-dire que vous aviez le tableau des données brutes. On pouvait entre zéro et cinq compter le nombre d'occurrences pour obtenir la fréquence de…
Elève ? : C'est ce que je voulais dire au début.
Marie : De cet intervalle. D'accord ? Et personne me l'a proposé mardi mais on aurait pu le faire tout de suite.
[13:37] Marie : Alors après j'avais demandé que la prochaine éruption ait lieu au cours de l'année 2030. Alors, ça, qu'est-ce que ça donne ? Je suis sûre qu'il y en a encore plein qu'ont pas réfléchi chez eux ? B22 ?
Elève B22 : Ça veut dire que $P(X = 15)$. Puisqu'on est en 2015, donc 2030 ça va faire 15.
Marie : Bien vu. Quoique. Tout le monde a calculé ça ?
Certains élèves répondent « oui », d'autres « non ».
Marie : Ça, ça fait quoi ça ?
Elève ? : Le temps d'attente…
Elève B31 : Ça fait zéro.
Marie : Oui ça fait zéro.
Marie : C'est ça que je demande ?
Elève : Non.
Marie : J'ai demandé… N'oubliez pas que c'est continu. Et j'ai demandé au cours de l'année 2030.
Plusieurs élèves réagissent.
Marie : Alors ? Donc du 1er janvier au 31 décembre. Alors qu'est-ce que vous proposez de calculer ? Je vais juste écrire comme ça : au cours de 2030, c'est déjà écrit…

Au cours de 2030 :

Marie : Alors, A22 pardon ?
Elève A22 : Sur l'intervalle… On calcule P sur l'intervalle 15 inclus et 16 exclu.
Marie écrit au tableau.

Au cours de 2030 :
$p(X \in [16; 16[\,)$

Marie : Alors, est-ce que vous êtes d'accord ?
Plusieurs élèves disent « Non, 15 ».
Marie : 15. 16. [*Marie efface et réécrit*].

Au cours de 2030 :
$p(X \in [15; 16[\,)$

Certains élèves discutent entre eux.
Marie : Levez le doigt si vous êtes pas d'accord. Qui est… qui n'est pas d'accord, pardon, avec A22 ?
Elève A42 : Il y a une subtilité.
Marie : Il y a une subtilité. Alors on t'écoute, A42.
Elève A42 : Il faut considérer que l'éruption a eu lieu le 1er janvier de cette année, parce que sinon vu qu'elle a eu lieu en décembre de l'année dernière…
Marie : Alors on va considérer que j'ai posé la question mardi et qu'on était le 18 mars. Et que c'est le 18 mars que…
Elève A42 : Qu'elle a eu lieu l'éruption ?
Marie : Oui. On est obligé, tu as raison. De toute façon elle a démarré l'éruption en vrai sur l'année fin 2014.
Elève A42 : Oui c'est ça.

Marie : Donc il faut le début. On va considérer comme hypothèse que c'était le 18 mars 2015, le début. Alors du coup…
Elève A42 : Du coup, c'est pas ça.
Marie : Du coup, c'est pas tout à fait ça. C42 ?
Elève C42 : Entre 14,75 et 15,75 à peu près.
Marie : 14,75 et 15,75 ?
Elève C42 : Oui.
Marie : Alors dis-nous. Hypothèse. [*Marie commence à écrire au tableau*]. L'éruption… Non, on va mettre comme ça. Le volcan est en éruption le 18 mars 2015. Et on va partir de cette date. A42, tu as raison, bien sûr ça dépend de la date de début d'éruption. Alors, tu nous proposes C42, probabilité que X soit compris entre 14,75, c'est ça que tu as dit ? Et…
Elève C42 : 15,75.
Marie : Tu es d'accord ? Pour que ça fasse un an. Si tu as une borne, tu as l'autre.

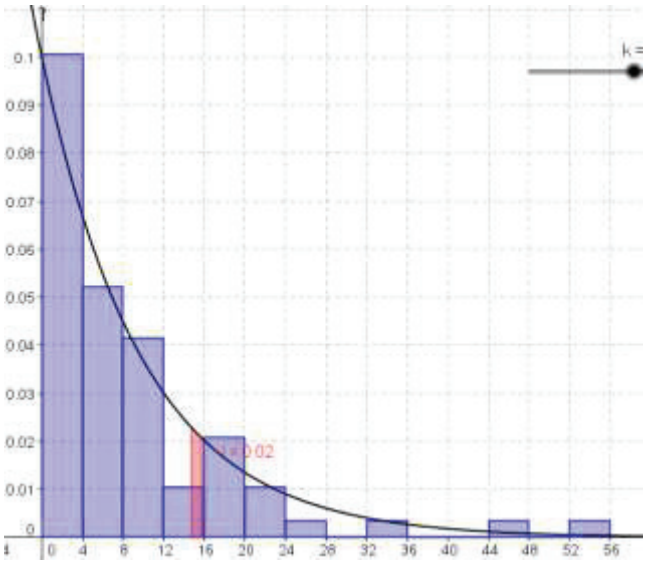
hyp : Le volcan est en éruption le 18 mars 2015
$p(14{,}75 \leq X \leq 15{,}75)$

Marie : Les autres, qu'est-ce que vous en pensez ? C'est quoi ces chiffres, là, bizarres que tu nous as sorti ? B52 ?
Elève B52 : En fait, j'ai toujours pas compris pourquoi faut qu'on soit en éruption le 18 mars.
Marie : Non, on va considérer que le début de l'éruption c'était le 18 mars 2015. D'accord ? Et que le prochain début d'éruption ça soit au cours de l'année 2030. C'est ça la question.
Elève A52 : Et peu importe.
Elève B52 : Oui peu importe.
Elève A52 : Puisque c'est…
Marie : Si tu veux compter 15 ans… Enfin si tu veux vraiment que ce soit au cours de l'année 2030, il faut que tu calcules la durée écoulée depuis l'éruption d'aujourd'hui.
Elève A42 : Mais de toute façon, on prend en compte que les années donc…
Elève B52 : Oui.
Marie : Non, justement c'est continu. C'est pas que des… Non, ça c'est les données qu'on avait. Les données qu'on avait on a bien dit que, par exemple, quand ils disaient 1 an, dans le tableau, ça pouvait vouloir dire entre 0 et 24 mois quand même, 0 exclu et 24 mois.
Elève A52 : D'accord.
Marie : D'accord ? C'est justement ça qu'est en train de traduire C42. Donc il faut bien qu'on compte, puisque c'est la durée entre 2 éruptions successives, faut bien qu'on compte le temps qui s'est passé entre l'éruption actuelle et le cours de l'année 2030. Donc on est en train de compter cette durée. Faut bien qu'on choisisse un point de départ dans l'année 2015. C'est ce que disait A42 en disant « on a qu'à considérer que ça commence le 1er janvier », alors effectivement ç'eût été plus simple. Parce que si on considère que c'est le 1er janvier, ça tombe très bien. Mais alors là c'est pas le 1er janvier, c'est le 18 mars. Alors pourquoi est-ce que tu proposes… Est-ce que vous comprenez les nombres là qu'il a donnés, C42 ? A52.
Elève A52 : Moi je dirais que par exemple on est en mars donc on prend un mois, un mois en année 2030 ça fait 2029 et en septembre. Du coup ça fait 14,75.
Marie : Oui. Est-ce que c'est ça que tu as fait ?
Elève C42 : Oui.
Marie : C'était à peu près, à vue d'œil ou tu as compté le nombre de jours ? 15 ans, si on rajoute 15 ans, on sera le 18 mars 2030. D'accord ? Donc il suffit de compter combien on a de jours du 1er janvier au 18 mars. D'accord ? En fait, il y en a 66 jours. Et c'est 66 jours si vous les ramenez en système décimal car une année c'est 365 jours.
Elève A52 : Ca dépend si c'est une année…
Marie : Alors moi j'avais compté que ça faisait 14,80 mais 14,75 ça va très bien et puis bien sûr après comme c'est sur 1 an, il n'y a pas de difficulté pour trouver l'autre borne. D'accord ? Pour compter de date à date. B52 ?
Elève B52 : J'ai toujours pas compris pourquoi 18 mars. Pourquoi on prend pas 1er janvier ?

Marie : Parce que, on a fait… Parce que ça m'arrange pas qu'on prenne le 1er janvier. Parce que moi je vous ai parlé de l'éruption le 18 mars. Mais d'ailleurs, c'est pas le 1er janvier l'éruption. En vrai.
Elève B52 : Vous avez dit elle commence en décembre.
Marie : 2014.
Elève B52 : Alors pourquoi on prend pas… Si l'éruption elle commence en décembre et elle va jusqu'en mars, on peut prendre n'importe quel jour entre décembre et mars.
Marie : Oui on peut prendre n'importe quel jour. Non on peut prendre le premier jour du mois de… c'est novembre en plus je crois 2014. Mais c'est… ça ferait le même… Ça donnerait le même principe, c'est-à-dire que c'est continu et que ça sera pas entre 14 et 15 ou 15 et 16. D'accord ? Il faudra compter le nombre de jours qui s'est écoulé réellement. Donc tu vas trouver des nombres avec… Tu vas trouver 14 virgule et des poussières et plus 1 évidemment, pour une année. Ok ? Non, t'as pas compris ?
Elève B52 : Mais. Bref.
Marie : Tu vas réfléchir, c'est ça ? Et je vais revenir te voir. Alors, est-ce que tu l'as calculé avec le… C42 ?
Elève C42 : Ça fait 0,02.
Elève A42 : Oui.
Marie : Oui, c'est ça. [*Marie écrit au tableau*].

$$p(14{,}75 \leq X \leq 15{,}75) \approx 0{,}02$$

Marie : Avec le logiciel, on change les bornes. [*Marie manipule le logiciel*]. Propriétés. Donc là on met… Alors, tu as pris 14,75 ? C42 ?
Elève C42 : Oui.
Marie : Tu l'as fait pour de vrai ou tu…
Elève C42 : En fait j'avais pris 15 et 16, mais…
Quelques élèves discutent entre eux.
[…]
La manipulation sur GeoGebra ne fonctionne pas.
Marie : Faut pas mettre de virgules sur GeoGebra quand on tape les décimaux. Oui faut mettre des points et j'ai mis… On va croire votre camarade. Ok.
Un élève : Vous avez g. Vous avez mis la fonction g.
Marie : Ah oui, merci. […]. Enfin, on va y arriver là. On est bon, là ? Yes. 0,02. Pardon. C'est bon ?



Elève A42 : Mais si on met 15 et 16, ça fait 0,02.
Marie : Oui parce que… Qu'est-ce que vous en pensez ? Pourquoi ?
Plusieurs élèves parlent en même temps.
Un élève : Parce que c'est négligeable.
Marie : Oui exactement parce que là la courbe elle devient très très fine et du coup la différence en terme de valeurs, elle est très faible. Donc on la sent pas. On la voit pas. Alors, maintenant… Oui, A52 ?
Elève A52 : Moi je comprends pas pourquoi… [*Inaudible*].

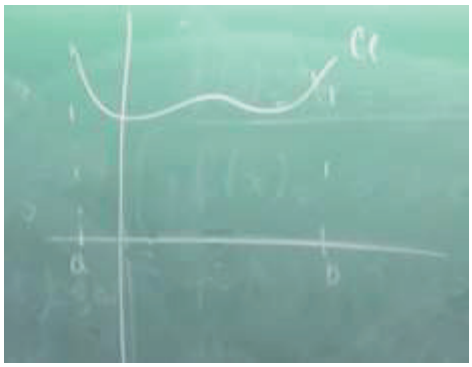
Marie : J'entends pas.
Elève A52 : Pourquoi la courbe doit passer par le point (0; 0,1) ?
Marie : Alors. Non, justement pourquoi ?
Elève ? Parce qu'on l'a dit.
Elève A52 : C'était intuitif.
Marie : Oui, on a décidé ensemble… on a dit « on va chercher une exponentielle décroissante », c'est C41 qui nous l'a proposé. D'accord ? On a tracé $exp(-x)$ mais celle-ci on sait qu'elle coupe en (0; 1). D'accord ? Au point de coordonnées (0; 1). Et on s'est dit ensemble mardi « mais non, il faut que ça coupe plus bas » et je vous ai demandé « voilà, vous voulez qu'on la fasse couper à peu près où ? » et on s'est mis d'accord. D'accord, encore une hypothèse. On a fait beaucoup de choix. 0,1. Rappelez-vous on a choisi de travailler avec un histogramme d'amplitude 4, on a choisi… voilà. Si je colle dessous d'autres histogrammes d'amplitudes 5 etc, on voit que la courbe, elle lisse bien les différents histogrammes aussi. Alors. Oui, C41 ?
Elève C41 : Moi j'ai bien compris les valeurs, mais je comprends pas pourquoi on fait un intervalle au cent…
Marie : A deux chiffres après la virgule ?
Elève C41 : Deux chiffres après la virgule. [*Rires*].
Marie : Parce qu'on a calculé le nombre de jours entre le 1er janvier et le…
Quelques élèves disent « 18 mars ».
Marie : Et le 18 mars, pardon. D'accord ? Ca fait à peu près 66 jours. Et ça faisait 66 jours quand tu ramènes sur 365, ça fait 0,75 année[1].
Elève C41 : Oui mais comme c'est un intervalle continu, on peut pas prendre un… Si c'est continu, il y a pas de valeurs enfin. Parce que il y a des valeurs entre les valeurs puisque c'est l'axe des réels… Non mais si enfin…
Marie : Non mais non mais si. Moi je veux bien mais je comprends pas ce que tu veux me dire.
Elève C41 : Comme c'est l'axe des réels, on peut pas prendre 75 pile.
Marie : On peut quand même… Là j'imagine… Ça tombait juste ton truc ? Non. Donc t'as fait un choix de décider de couper. Si on… C'était un décimal quand même ou pas ?
Elève C42 : [*Inaudible*].
Marie : Pourquoi ?
Elève C41 : Parce que c'est pas une puissance de 10.
[*Rires*].
Marie : D'accord. Donc on a fait un choix ici aussi de tronquer à deux enfin il a fait le choix de tronquer à deux décimales après la virgule. D'accord ? Mais sinon, tu calcules… C'était quoi ? C'était 66 j'ai dit ? Donc 14… 14 + 66 sur 365[2]. [*Marie écrit au tableau en même temps*].

$$14 + \frac{66}{365}$$

Marie : Ça c'est un nombre réel, c'est la valeur exacte.
Elève C41 : Ah d'accord.
Marie : D'accord ? Donc il a fait le choix de… d'arrondir. Ok ? Tu peux mettre ça avec le logiciel, tu peux le calculer. Il va te donner 0,02 aussi. Ça sera pas sensible sur l'aire, sur la valeur. D'accord ?
[24:58] Marie : Alors. Donc. Tout ça c'est bien beau mais A52, tu avais une interrogation mardi.
Elève A52 : Effectivement.
Marie : Effectivement. Peux-tu nous en faire part ?
Elève A52 : Il y en a eu tellement madame.
[*Rires*].
Marie : Donc, pour résumer. Nous avons une variable aléatoire. Prenons une variable aléatoire continue. Donc qui prend toutes ses valeurs… Ça veut dire quoi une variable aléatoire continue, A22 ?
Elève A22 : Qui prend toutes les valeurs… [*Inaudible*].
Marie : Vous ne pouvez pas parler à haute voix ?
Elève A22 : Une infinité de valeurs entre…
Marie : Non. Prend toutes les valeurs d'un intervalle donné. D'accord ? Et on trouve la courbe… la fonction qui nous permet de calculer quoi ? Elle s'appelle comment cette fonction f là qu'on a cherchée ?

[1] L'enseignante voulait dire : 1-66/365=0,75.
[2] Même remarque que précédemment.

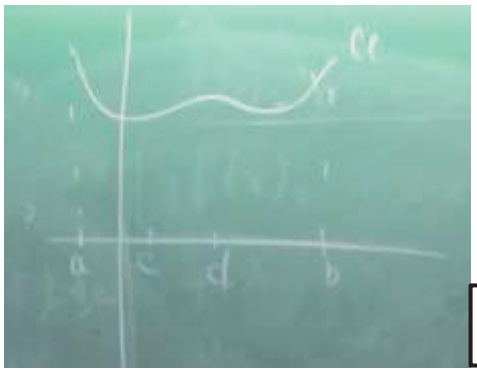
Elève A12 : Courbe représentative.
Elève B31 : De tendance. Courbe de tendance.
Elève B32 : Courbe de densité de probabilité.
Marie : On lève le doigt. C42.
Elève C42 : De densité.
Marie : De densité de probabilité. D'accord ? Donc sur l'intervalle $[a\,;b]$ donc imaginons ça. [*Marie trace une courbe au tableau en même temps*]. Et comment je fais grâce à cette fonction que j'ai trouvée… Cette fonction qu'est-ce qu'elle doit avoir comme contraintes, on a dit ? […] B32.



Elève B32 : $f(0) = 0{,}1$.
Marie : Non, non. Dans le cas général. On prend une variable aléatoire… [*Marie écrit au tableau*].
Elève B32 : Strictement positive.
Marie : Continue, voilà. Donc notre fonction f qui est définie sur l'intervalle $[a\,;b]$. Elle est donc continue. [*Marie écrit au tableau*]. D'autres choses ?
Elève B32 : Positive.
Marie : Oui, positive. D'accord. Et puis ? Qu'est-ce qu'il y avait ? B21.
Elève B21 : Sur l'intervalle $[a\,;b]$, l'aire elle vaut 1.
Marie : Donc l'aire sous C_f entre a et b vaut 1. Bon on a trouvé tout ça.

X v.a. continue
f def. $[a\,;b]$
cont., pos.
aire sous Cf e/ a et b = 1

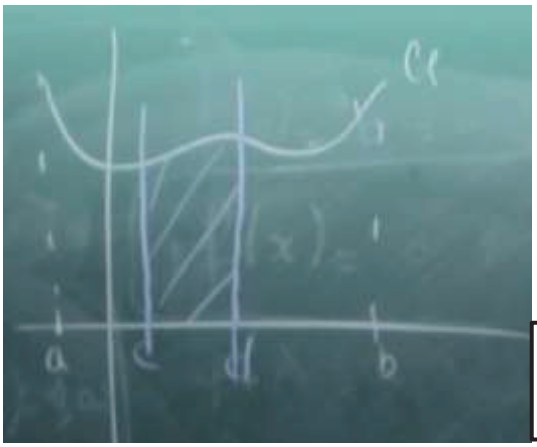
Marie : Comment fait-on pour calculer la probabilité que X soit dans l'intervalle je vais dire celui-là… Par exemple, voilà. [*Marie place les abscisses c et d sur le graphique*]. C'est un d ça. Comment je fais pour calculer, évaluer la probabilité que X soit compris entre c et d ? [35 :35]



$p(c \leq X \leq d)$

Marie : C22 ?
Elève C22 : On calcule… on calcule l'aire.
Marie : D'accord. Alors je vois qu'il faut faire des trucs comme ça visiblement.

Elève C22 : Placer les points c et d et après calculer l'aire entre c et d.
Marie : Donc c'est ça, c'est l'aire sous la courbe entre c et d, on dit comme ça. Aire sous C_f entre c et d. [*Marie écrit en même temps au tableau*].



$p(c \leq X \leq d)$ *aire sous* C_f *e/ c et d*

Marie : Alors. Quel est notre… problème ? En tout cas, on a eu à faire face à deux situations très différentes. C11 ?
Elève C11 : La courbe, elle est un peu arrondie.
Marie : La courbe, elle est un peu arrondie. Oui. Et ça t'aime pas ? Non. Alors, vous vous souvenez de la puce. C'est quoi la fonction de densité pour la puce ? Est-ce que vous êtes capable de me dire, pour la puce, c'était quoi f ? Oui, A32 ?
Elève A32 : C'était une constante.
Marie : Oui. Une constante. Tu te souviens de… ?
Elève A32 : C'était 1 sur…
Marie : C'était 1/12, oui. 1/12. [*Marie écrit au tableau*].

$$puce : f(x) = \frac{1}{12}$$

Marie : Donc ça c'était sympa, parce que là c'est pas arrondi, C11. Voilà, ça faisait des rectangles donc ça on savait faire. Pour… Olivier et Karine, c'est ça ?
Quelques élèves répondent « oui ».
Marie : C'était quoi notre fonction de densité ?
Plusieurs élèves répondent.
Marie : C'était $-2x + 2$. [*Marie écrit en même temps au* tableau].

$$O\ et\ \mathcal{K} : f(x) = -2x + 2$$

Marie : Donc c'était une droite, pas d'arrondi. Donc C11 était contente. Et donc on a calculé des aires de… ?
Plusieurs élèves répondent « trapèzes ».
Marie : Trapèzes, triangles, enfin bon des choses qu'on sait faire. Et le problème c'est… voilà, par exemple, celle-ci c'est une exponentielle décroissante, là, on n'a pas plus une figure simple. C'était ça ton interrogation, A52, sur lequel t'es revenu plusieurs fois mardi [la séance précédente]. Alors. Ça va être l'objet de votre travail là maintenant. J'aimerais que… Alors, on va prendre comme exemple, notre courbe de tendance, celle-ci. D'accord ? Et je vous demande de me proposer un… moyen, un algorithme pour calculer l'aire sous notre courbe. Alors on va se fixer un intervalle, le même pour tout le monde, sinon on va pas s'en sortir. On va faire entre 0 et 20, pour l'instant. D'accord ? Donc. Problème suivant, enfin… Problème suivant. [*Marie efface le tableau*]. Alors allez-y, proposez-moi. A22. Vous travaillez à deux si vous voulez, en chuchotant. Non, t'es pas à côté. Deux, en chuchotant. D'accord ? Deux, deux, deux, deux, deux, deux… B23, tu vas à côté de A52, s'il-te-plait ? Ok ? Je vous laisse… un petit moment. Je vais vous distribuer la courbe sur papier quand même. Une propre, belle, pour vous aider pour voir, si vous avez des choses à faire. D'accord ?
La classe s'agite légèrement.

Marie : A11, qu'est-ce que tu cherches là ? C'est quoi la question ? Qu'est-ce que tu vas chercher avec A12 ? A11, que vas-tu chercher avec A12 ? Chut.
Elève A11 : [Inaudible].
Marie : D'accord. Donc, on considère notre fonction. [*Marie écrit au tableau*].

$f(x) = 0{,}1e^{-0{,}1x}$
Algo pr calculer
l'aire ss C_f
entre 0 et 20

Marie : Entre 0 et 20, pour qu'on puisse échanger facilement. Je vous donne la courbe. [*Marie et un élève distribuent les photocopies*]. [39 :39]
[31:38] *Les élèves travaillent par deux.*
[...]

Marie : Si vous avez besoin de quelque chose, vous m'appelez.

Marie passe dans les rangs.
Voici quelques extraits.

Marie (à un élève) : Ah non, pas d'indications mathématiques.

Marie (à l'élève B53) : Une méthode, un algorithme. [...] Algorithme, ça veut pas dire, on prend sa calculatrice. Pose-la pour l'instant. [...] Un algorithme c'est pas forcément quelque chose qu'on programme. Une méthode, quelque chose qu'on va faire pour pouvoir raisonner, qui résout le problème. D'accord ?

[41:30] Elève B31 : On propose de… en fait on trouve l'équation de la tangente.
Marie : Donc ça on sait faire.
Elève B31 : Oui, oui. A chaque point. On commence à 20 et après on va jusqu'à 0. Et du coup, là, après, on a un truc qu'on peut calculer. On calcule cette aire-là, et ensuite on calcule celle-là et celle-là, et on fait la somme.
Marie : D'accord.

[42:08] Marie : Alors, qu'est-ce que vous proposez ?
Elève B12 : Alors en fait de faire ça [*l'élève montre les rectangles qu'elle a tracé sur sa feuille*].
Marie : Très bien. Oui. C'est l'idée la plus simple.
Elève B11 : Mais on sait pas comment la décrire.
Marie : Est-ce que vous pensez que vous allez obtenir une valeur exacte de toute façon comme ça ? [Inaudible].
Elève B11 : Non.
Marie : En tout cas, vous aurez une valeur approchée. Vous avez une méthode. Si vous faites… vous avez fait ça, donc vous avez fait un grand…
Elève B12 : Rectangle.
Marie : Un grand rectangle. Après si tu as rajouté celui-là, celui-là et… celui-là. Enfin… Faut peut-être faire les contours en couleur parce qu'on voit pas trop bien. Donc pourquoi vous avez fait des… 4 grands rectangles ?
Elève B11 : Parce que c'était plus simple.
Marie : Non, mais pourquoi 4 ?
Elève B12 : Parce qu'on suit la…
Marie : Parce que j'ai donné une graduation de 5 en 5. Donc pour améliorer votre valeur, qu'est-ce que vous pouvez imaginer de faire ?
Elève B11 : Réduire l'intervalle.
Marie : Réduire la taille des rectangles.
Elève B11 : Oui voilà.
Marie : Là, ils sont très grands donc c'est grossier. Mais imagine déjà que tu coupes là, ça sera mieux. Un rectangle qui sera là et un autre là. Oui, ça sera plus précis. Tu vois bien que la perte, elle va être moins

grande. […] Entre 0 et 20, vous pourriez faire quoi ? Oui, de 2 en 2, de 1 en 1… Voilà, plus vos rectangles sont… de bases petites… plus votre précision est juste. On va perdre moins de choses. D'accord ? Alors, dans un premier temps, si vous savez faire pour 4 rectangles, vous saurez faire pour 10, 20, 30 rectangles. Faîtes-le pour 4.
Elève B12 : 1, 2, 3, 4.
Marie : Oui ceux-là. Repasse-les mieux en couleur.
[…]
Marie : Donc c'est bien ça. Repassez plus en couleur, parce que j'ai pas bien compris quel était le dernier.
L'élève B12 montre à Marie le rectangle en question.
Marie : D'accord. Et bien, calculez les quatre aires que vous avez.
Elève B11 : Oui, d'accord.
Marie : Alors, la valeur que vous allez approcher, elle va être comment par rapport à l'aire qu'on cherche ? Plus grande. […]
Marie : Donc plus grande.
[…]
Elève B12 : On fait de 2 en 2 ?
Marie : Non. Calcule déjà avec ça. On a dit, si vous pouvez faire pour 4,… Non, là on considère que vous calculez l'aire sous les rectangles oranges [*l'élève B11 a repassé les rectangles en orange*]. C'est plus grand que celle qu'on cherche. Donc, qu'est-ce qui serait bien de trouver aussi ? Ça, ça c'est curviligne donc ça change pas le problème que tu avais de l'autre côté, tu vois ? C'est le même problème. Calculez-moi déjà ça. Je reviens. On est d'accord, tout ton truc orange, là. Ce que tu m'as dit, ce que tu m'as montré. Bon, allez-y.
[…]
Marie : Mais on est d'accord que l'aire que vous allez proposer est plus grand que l'aire qu'on cherche.
Elève B12 : D'accord.

[53:36] Marie (à la classe) : Dessinez-moi sur votre courbe ce que vous faites parce que…

[59:46] Marie : Alors, j'ai vu pleins pleins de choses. On va faire un petit bilan.
Marie vidéo-projette la courbe au tableau.
Marie : Alors, c'est votre courbe avec le… Alors. Du coup, globalement, ce que j'ai vu… […] Donc quand même l'idée générale qu'est au fond de tout ce que vous proposez, c'est de se débrouiller pour avoir quoi ?
Elève B31 : Des figures.
Marie : Levez le doigt, B31. Quoi que vous ayez proposé… […] finalement vous vous êtes tous débrouillés pour obtenir quoi ? Suivant un très grand principe en mathématiques.
Sept élèves lèvent la main.
Marie : Les autres n'ont pas conscience de ce que vous avez essayé de faire ?
Marie : Vous m'avez tous parlé de rectangles, de triangles, de trapèzes, de… voilà de ça. Donc, B42 ?
Elève B42 : On a cherché à avoir des formes géométriques faciles…
Marie : Des formes géométriques simples. Vous avez tous chercher, quelque soit la méthode dans laquelle vous vous êtes lancés à, sous cette aire qui est curviligne, à reconstituer des figures simples pour réussir à faire du calcul d'aire. Est-ce que vous êtes d'accord que tout le monde à essayer de faire ça ?
Des élèves répondent « oui ».
Marie : Alors après la méthode pour y arriver, elle est très diverse. Il y a des méthodes à base de tangentes. Il y a plusieurs groupes. Alors, disons qu'on va grossièrement le voir. Par exemple, si l'on se place à 10, il y en a qui trace la tangente à la courbe en ce point et du coup, on récupère l'aire d'un trapèze. Mais comme il y a de la perte, enfin, en tout cas, il y a des gros trous, vous recommencez sur cette figure-là en reprenant des tangentes. Alors, par contre, vous les prenez un petit peu au hasard vos tangentes, en des points qui sont pas très bien identifiés. A partir du moment où on cherche un algorithme, ce serait bien que ce soit répétitif… enfin que l'on puisse répéter de façon… Vous voyez ce que je veux dire ? Parce que si c'est un peu au hasard, on va pas s'en sortir quand même. A12 ?
Elève A12 : On… par exemple, la première tangente, on la prend en 20. Après on va décaler la deuxième tangente en un espace qu'on va définir.
Marie : D'accord.
Elève A12 : Puis, après, cet espace, il va se réduire au fur à mesure que l'on va se rapprocher de 0.
Marie : D'accord. Mais alors, donc là, quand on fait ça, on prend la tangente, là, vous vous avez décomposé en un trapèze et après c'était un peu compliqué parce que du coup, avec votre deuxième tangente, il fallait évaluer l'aire entre le morceau qui avait déjà été calculé et le nouveau et enlever la partie qui était commune.
Elève A12 : Maintenant c'est bon.

Marie ; Oui, maintenant c'est bon mais c'était un peu ça l'idée. Donc, il y a quand même, dans ce que tu dis, l'idée qu' on prend un point de départ, qu'on choisit un… je sais pas quoi… un…
Elève A12 : Un pas.
Marie : Un pas, voilà exactement. On choisit un pas et on va recommencer la méthode.
Elève A12 : Sauf que ce pas, il se réduit au fur à mesure.
Marie : Alors, toi, tu choisis que le pas se réduit au fur à mesure. D'accord. Alors, qui a fait des tangentes là un peu partout ? il y en a plein de groupes qui ont fait des tangentes. Vous avez fait avec les tangentes ? La méthode que vous avez sortie, c'était avec les tangentes ?
Elève B53 : …
Marie : Oui ?
Elève B53 : A peu près.
Marie : Vous avez trouvé quoi, B52 et …? [*Marie commence à écrire au tableau*]. Alors, soit avec les tangentes. Et vous avez trouvé quoi, B52 et B53 ?
Elève B53 : 0,8725.
Marie : Zéro virgule ?
Elève B52 : 8725.

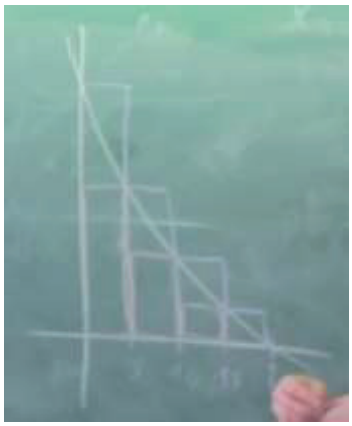
Tangentes :
$B52\text{-}B53 : A \approx 0{,}8725$

Marie : Il y en a d'autres qui ont abouti à des calculs… à des valeurs numériques, en utilisant des tangentes ? C42 ?
Elève C42 : Nous, on a 0,863.
Marie écrit au tableau.
Marie : C'est ça, C42 ?
Marie : C31.
[…]
Elève C31 : 0,865.

Tangentes :
$B52\text{-}B53 : A \approx 0{,}8725$
$B42\text{-}B41 : A \approx 0{,}863$
$B31\text{-}B32 : A \approx 0{,}865$

Marie : Est-ce que tout le monde a compris, avec les tangentes, ce qu'on fait vos camarades ? Il y en a plein qui ont commencé par… Pourquoi vous avez pensé aux tangentes, au fait ? Ça vient d'où cette idée ? Levez le doigt. C31, C32, pourquoi vous avez pensé à prendre les tangentes ? C31, C32, je vous parle. Je sais pas, pourquoi vous avez pris les tangentes ?
[…]
Elève C31 : On connait les équations…
Marie : Parce que tu sais trouver les équations de tangentes. B32 ?
Elève B32 : Parce que au voisinage de la courbe, la tangente… tend vers la courbe…
Marie : Alors, oui. Parce que c'est ce que vous avez vu l'année dernière, c'est-à-dire qu'au voisinage d'un point, la droite tangente, elle est assez proche de la courbe. Oui. Donc ça venait de ce que vous avez fait en première pour introduire le nombre dérivé, ce que vous vous êtes amusé à faire. Alors, il y a eu d'autres méthodes que les… que d'utiliser les tangentes. B11 et B12 ? Vous avez fait quoi, vous ?
Elève B11 : On a fait des rectangles.
Marie : Vous avez pris des rectangles.
Marie trace la courbe au tableau, ainsi que les rectangles en question.
Marie : Vos camarades ont découpé grossièrement… enfin grossièrement… Ils sont comment ? Ils sont comme ça. Ça, ça et ça.

Elève A52 : En fait, c'est l'histogramme.
Marie : C'est les mêmes… Oui, ils ont reconstitué un histogramme en-dessous (sic). A42 et A41, c'est les mêmes rectangles à peu près ?
Elève A41 : Non, non, non, en-dessous.
Elève A42 : Ils sont en-dessous et on rajoute des triangles.
Marie : Non, eux, c'est des rectangles aussi, j'ai bien vu des rectangles.
Elève B11 : Mais nous, on n'a pas pris les triangles.
Elève A42 : Oui, c'est des rectangles, mais c'est pas…
Marie : Alors, vous, ce sera… [*Marie trace de nouveaux rectangles sur le graphique précédent*]. Je mets en rose sur le même découpage parce que sinon… Donc vous, ce serait… Alors celui-ci, j'imagine.
Elève B42 : C'est ça.



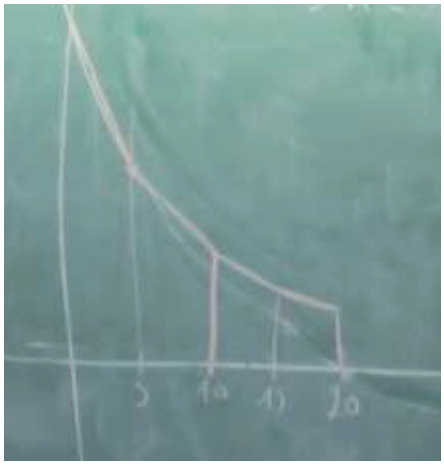
Marie : Ça va pas être très beau.
Elève B42 : Et on rajoute dessus…
Marie : C'est ça ? C'est ceux-là, vous ?
Marie : B22.
Elève B22 : Nous, on a fait des trapèzes. Mais…
Marie : Alors, attends, je vais faire un autre dessin parce que si c'est des trapèzes, on va plus rien voir. Vous avez calculé ?
Elève B11 : On était en train.
Marie : Ah, on était en train.
Elève A42 : Nous, on a une valeur.
Marie : Oui, mais tu m'as rajouté les… Tu as fait des traits. En fait, c'est la même chose que ce qu'elle va dire. C'est pour ça. Qu'est-ce que vous en pensez de l'aire qu'on cherche par rapport à ce qui est en rose et ce qui est en violet si je peux dire comme ça.
Plusieurs réponses semblent émerger « c'est plus grand », « c'est la moyenne ».
Marie : Qu'est-ce qu'on peut en dire comme ça ?

Une élève parle [inaudible].
Marie : L'aire en violet, elle est plus grande que l'aire qu'on cherche [*Marie écrit en même temps au tableau*] et l'aire rose ?
Plusieurs élèves répondent.
Marie : Elle sera plus petite.

$$A_R \leq A \leq A_V$$

Marie : Alors évidemment, si on prend 4 rectangles, là, c'est… Où est l'erreur qu'on commet ? Est-ce que vous êtes d'accord que l'aire qu'on cherche, elle est comprise entre l'aire rose et l'aire violette ? On commet une erreur, quand même importante. Elle est où ? Sur le dessin, vous la voyez ? B53 ?
Elève B53 : C'est que la courbe, elle passe pas exactement… Enfin, c'est assez imprécis vu que les rectangles...
Marie : Non, c'est pas imprécis.
Elève B53 : Les rectangles, ils sont assez épais enfin assez gros.
Marie : Oui, là, on a pris des gros rectangles.
Elève B53 : Du coup, si on diminue la taille des rectangles, ça serait plus…
Marie : Si on diminue la taille des rectangles, ça va mieux se passer. Alors, vous avez quand même pas répondu à ma question. Elle est où visuellement l'erreur ? Quand je prends un grand rectangle, là, comme ça, le grand violet qui est ici [*Marie montre le deuxième rectangle violet*], par rapport à l'aire que je veux calculer, elle est où l'erreur ? Elle mesure où… comment ? B41.
Elève B41 : Le triangle au-dessus.
Marie : Oui, alors c'est pas un triangle, c'est un triangle curviligne. Ça, ça va être l'erreur. [*Marie montre la figure en question*]. Ça, c'est celle qu'on commet en sous-estimant (sic) l'aire qu'on veut calculer. Alors, B53, tu dis, si on prend des triangles plus petits… Non, pardon, des rectangles plus petits, qu'est-ce qu'il va se passer ?
Plusieurs élèves discutent.
Marie : Non, non, que les rectangles.
Elève B32 : Oui, je pensais qu'ils rajoutaient après…
Marie : Non, non, seulement les rectangles. Je vais donner la parole à B22 pour ça après. Juste des rectangles comme ça. Et du coup, ce que tu disais, c'est que là, c'est grossier, c'est sûr avec 4 rectangles… Donc si on prenait des rectangles plus petits, on va en faire plus et qu'est-ce qui va se passer ? B31 ?
Elève B31 : Ça va être plus proche de la courbe.
Marie : Oui. Est-ce que vous voyez que l'erreur, elle va être moins grande ?
Elève A52 : Oui.
Marie [*en modifiant son graphique*] : Déjà si je coupe là, le prochain il va être là. D'accord ? Donc l'erreur que je vais commettre, elle va être moins grande que celle que j'avais au départ. Et pareil en-dessous. Je vous montrais quelque chose plus clair. Alors, est-ce que ça, ça va être quelque chose d'assez facile à réaliser ?
Marie : C'est des aires de rectangles, c'est base fois…
Plusieurs élèves répondent « hauteur ».
Marie : Hauteur. C'est quoi la base ? Alors, là, la base qu'on a choisi, par exemple, prenons celui-ci [*Marie montre le quatrième rectangle violet*], c'est 5, tout le monde est d'accord ? C'est grossier. C'est quoi la hauteur de ce rectangle-là ?
Plusieurs élèves répondent « c'est $f(15)$ ».
Marie : C'est $f(15)$, donc on connait f, donc on sait calculer. On est d'accord ? Donc il suffit d'être patient. Oui. Et vous avez des machines, après.
Elève A22 : Madame.
Marie : Oui, A22.
Elève A22 : Ça va être super long. Et ça va pas être très précis. Si on choisit de le faire…
Marie : Si on le fait avec beaucoup, beaucoup de rectangles, qu'est-ce qu'il va se passer ?
Elève B11 : Ça va être très précis.
Elève A22 : Ça va être très très long et la précision…
Rires.
Elève A22 : Non mais le temps qu'on passe en le faisant, c'est vrai que la précision, elle va s'améliorer, mais on passe beaucoup trop de temps.
Elève B11 : C'est un algorithme.
Marie : Est-ce que tu penses que c'est quelque chose que l'on fait à la main ?

Elève B11 : C'est un algorithme, en fait c'est bon.
Marie : Un programme, oui. Alors, B22, elle n'a pas fait des… alors, c'est comme ce que vous avez complété [*en s'adressant à A41 et A42*]. B22, c'est pas des rectangles.
Elève B22 : C'est des trapèzes.
Marie : C'est des trapèzes. [*Marie refait un nouveau graphique*]. Alors, non, pas ça quand même, faut pas que j'exagère. Alors, je les montre sur 5, 10, 15 pareil, c'est pas grave ?
Elève B22 : Non, c'est 2 et demi.
Marie : Oui, mais c'est pareil. On est d'accord. Parce que déjà que c'est pas beau au tableau, si j'en fait des tout petits, on ne va plus rien voir. Remarquez, j'aurais pu faire le dessin plus gros au départ. Alors, ils sont où ces fameux trapèzes dont tu parles ?
Elève B22 : En fait, c'est les mêmes…
Marie : D'accord, c'est les mêmes. Là je monte.
Elève B22 : Sauf que…
Marie : Et donc tu as rejoint… Mince, j'ai raté. Comme ça ?
Elève B22 : Oui, voilà.
Marie : C'est ça ? Et donc, bien sûr, pareil ici. Pareil là. Pareil là.



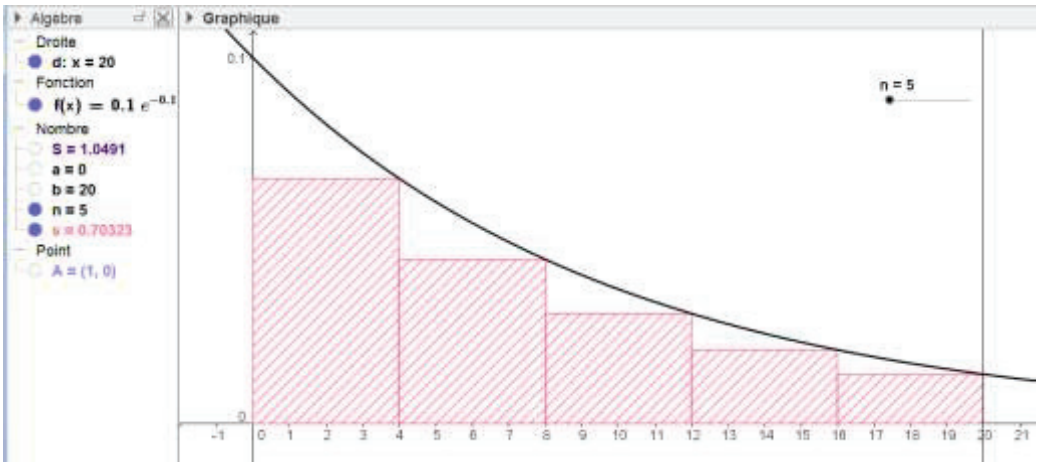
Marie : Oui, non mais après on fait comme les rectangles. On fait de plus en plus de trapèzes, pour que ce soit de mieux en mieux. D'accord ? C'est pareil, est-ce qu'on sait calculer l'aire de chacun de ces trapèzes ?
Plusieurs élèves répondent « oui ».
Marie : Oui, la base alors là c'est simple, on l'a choisi en fait et puis… Ah pardon, la largeur. La petite base, c'est ?
Plusieurs élèves répondent « $f(5)$ ».
Marie : Voilà. La grande base ?
Plusieurs élèves répondent « $f(0)$ ».
Marie : $f(0)$. Donc on sait aussi calculer l'aire. Et c'est pareil, plus on va prendre de trapèzes…
Elève B22 : Mais c'est plus précis que les rectangles.
Marie : Alors on voit là par exemple, effectivement, on voit sur le… On peut percevoir que sur les largeurs 5, l'erreur qu'on commet elle est moins grande ici que les rectangles que j'avais dessiné tout à l'heure. D'accord ? Ici, on obtient quel type de renseignement sur l'aire ? Enfin… La méthode des rectangles, là qu'on a… quand vous prenez des rectangles, ça vous donne un encadrement de l'aire. Ici, ça vous donnera quoi ?
Elève B11 : Une approximation.
Elève B32 : Une majoration.
Elève A22 : Une valeur arrondie.
Marie : Alors. Une majoration mais donc une seule valeur. Une valeur appro…
Elève A22 : Arrondie.
Marie : [*Rires*]. Valeur approchée qu'on va arrondir, qui sera ici par excès. Ça sera peut-être pas toujours le cas. Est-ce qu'il y a eu d'autres méthodes ? Donc vous avez construit des rectangles, des trapèzes et avec des tangentes, vous avez reconstitué des rectangles et des… Non, pas des rectangles. Des… triangles et des… des trapèzes. Globalement, c'est bon ? Alors, avec la méthode… avec les rectangles, là, il y en a qui ont obtenu… Ah non, vous avez pas fini vos calculs. Ils étaient grossiers de toute façon. A42 ?
Elève A42 : Non, nous c'est trapèze.
Marie : Oui, alors avec les trapèzes ?

Elève A42 : Zéro virgule…
Elève A41 : 8691.
Elève A42 : 869163…
Marie : Alors, attendez. [*Marie écrit au tableau en même temps*]. Zéro combien ?
Elève A41 : 8691.
Elève A42 : Après il y a beaucoup de chiffres.
Marie : Oui. Sûrement que eux aussi, ils ont dû me donner une valeur approchée.
Marie : A32.
Elève A32 : Nous, en arrondissant, ça fait 0,7.
Marie : Zéro virgule ?
Elève A32 : 7.
Marie : Ah oui. T'as fait des trapèzes ?
Elève A32 : Oui.
Marie écrit au tableau.
Elève A32 : Et avec l'algorithme…
Marie : Vous avez arrondi, oui.
Elève A31 : L'algorithme, ça donne 0,86467.
Elève A32 : 0,865, avec l'algorithme.
Marie : D'accord. B14.
Elève B14 : 0,874.
Marie : T'as fait des trapèzes ?
Elève B14 : Oui.
Marie : Ah oui, c'est vrai. 87 ? Bon on tourne à peu près.

Trapèze :
$\mathcal{A}_{42}$-$\mathcal{A}_{41}$: $A \approx 0{,}8691$
$\mathcal{A}_{32}$-$\mathcal{A}_{31}$: $A \approx 0{,}865$
$\mathcal{B}_{14}$-$\mathcal{B}_{13}$: $A \approx 0{,}874$

Elève A32 : On n'a pas la même valeur.
Marie : Oui c'est bizarre… Vous avez fait combien de trapèzes ?
Elève B14 : J'ai fait 0,1.
Marie : Ah oui. Vous avez fait combien de trapèzes, vous ?
Elève A32 : 7.
Marie : 10. 7 trapèzes. 10. Et toi ?
Inaudible.
Marie : Ok tu ne sais pas. Il y en a plus. Ok. Alors. Donc, toutes ces méthodes que vous proposez, je vais vous montrer ce que ça donne sur les rectangles.
Elève A32 : Madame.
Marie : Oui.
Elève A32 : Avec l'algorithme, il a fait des amplitudes de 0,01.
Marie : Là, pour ça ? Pour cette valeur-là ?
Elève A32 : Oui.
Marie écrit au tableau pour compléter.
Marie : Amplitude 0,01. Alors notez quand même que pour… - alors, on va laisser les tangentes parce que c'est un peu plus compliqué – ça c'est facile de faire un algorithme pour le programmer. Vous êtes d'accord ?
Plusieurs élèves répondent « oui ».
Marie : Soit… les tangentes c'est compliqué pour l'instant, d'accord, mais que ce soit les trapèzes ou les rectangles, ça va être facile à programmer.
Quelques élèves répondent « oui ».
Marie : Oui. Vous noterez ça pour demain ? On verra ça à la fin de l'heure. Alors, je vous montre ce qui se passe…
[…]
Marie manipule le logiciel GeoGebra.

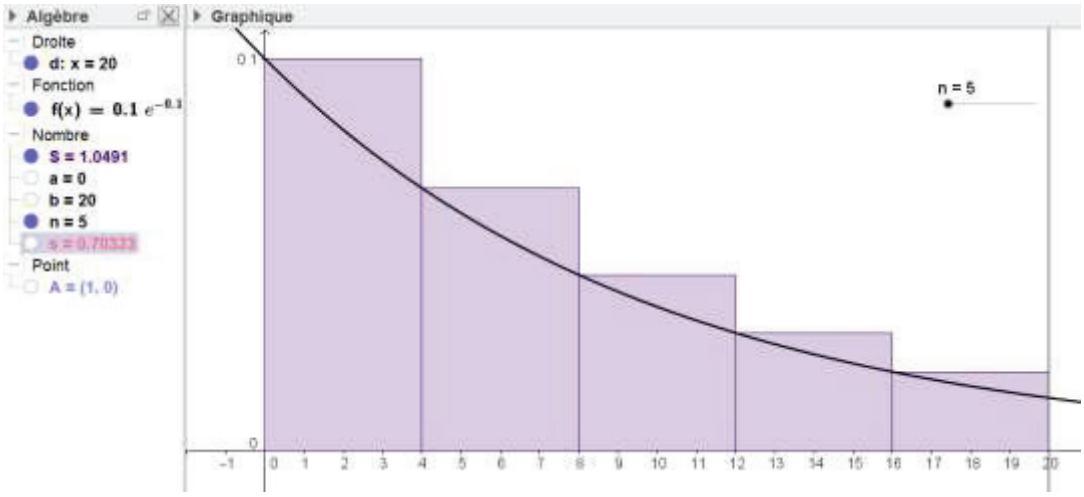
Marie : Alors, regardez la courbe, là. Donc les rectangles, si j'en fais… […] Si j'en fais 5… Regardez, ça c'est les petits rectangles, on va les appeler. […] Alors, au fait, les méthodes que vous avez proposez, elles existent toutes et elles sont plus ou moins rapides pour obtenir des valeurs approchées ou des encadrements de l'aire que l'on cherche. Déjà, c'est une bonne nouvelle. Alors, nous, celle qui est plus précisément à notre programme, c'est la méthode que l'on appelle la méthode des rectangles. Donc on monte… On choisit un pas comme tu disais, d'accord. En fait on choisit le nombre de subdivisions que l'on veut faire dans notre intervalle. Donc je l'ai appelé n, d'accord ? Donc là par exemple j'ai choisi de faire 5 bases de rectangles sur [0; 20]. Donc chacune d'amplitude 4. Et donc, on va monter des rectangles. Ceux qui sont en rose, on va les appeler entre nous, sinon c'est un petit peu lourd […], les petits rectangles. C'est eux qui sont montés sous la courbe, d'accord ? Donc là, vous avez la somme qui s'affiche ici [*Marie montre « s » dans la fenêtre Algèbre de GeoGebra*], enfin l'aire de tous ces petits rectangles, là ça fait 0,7 à peu près. D'accord ?



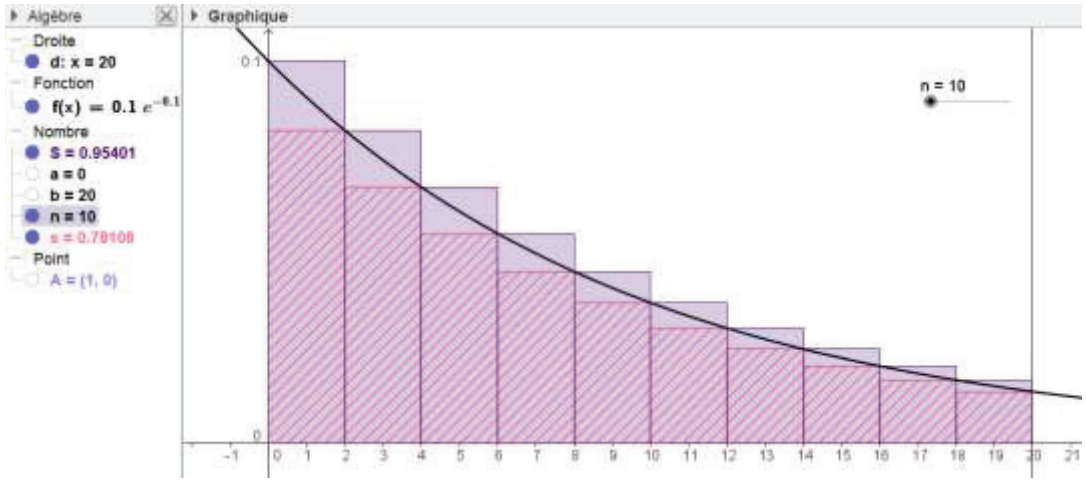
Marie : Et donc les autres… Je vais effacer celles-ci. Donc les petits rectangles, vous avez compris comment on les montait ?

Plusieurs élèves répondent « oui ».

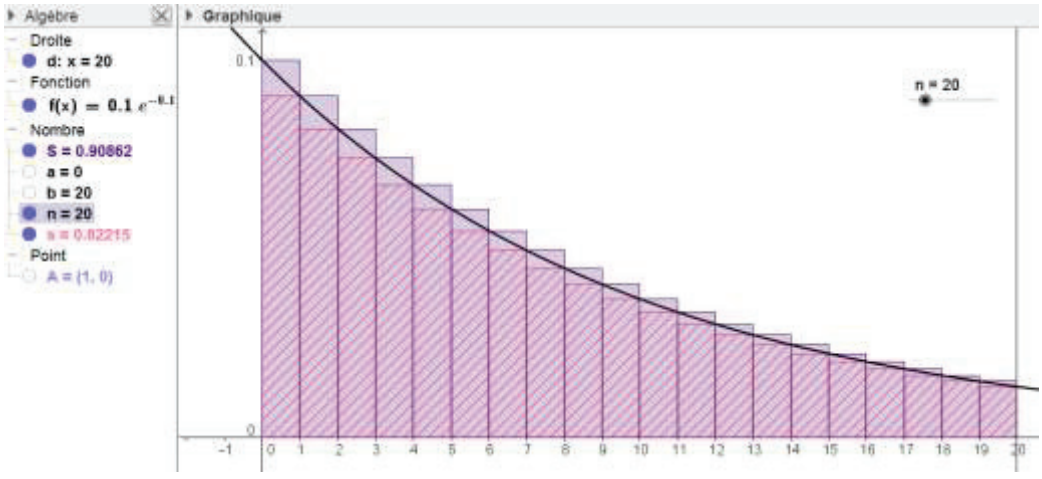
Marie : On va voir ça. Ça, c'est les grands. C'est les premiers qu'on a donné. […]

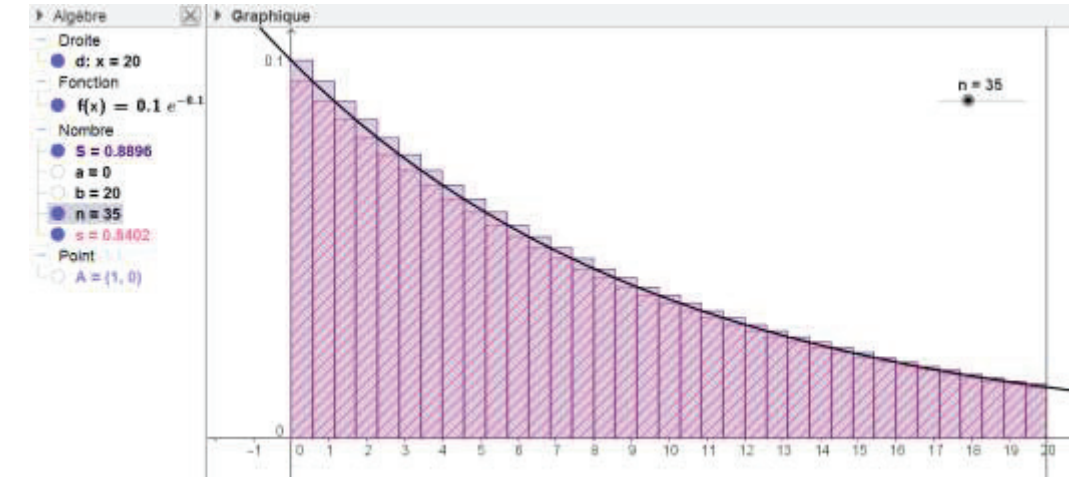


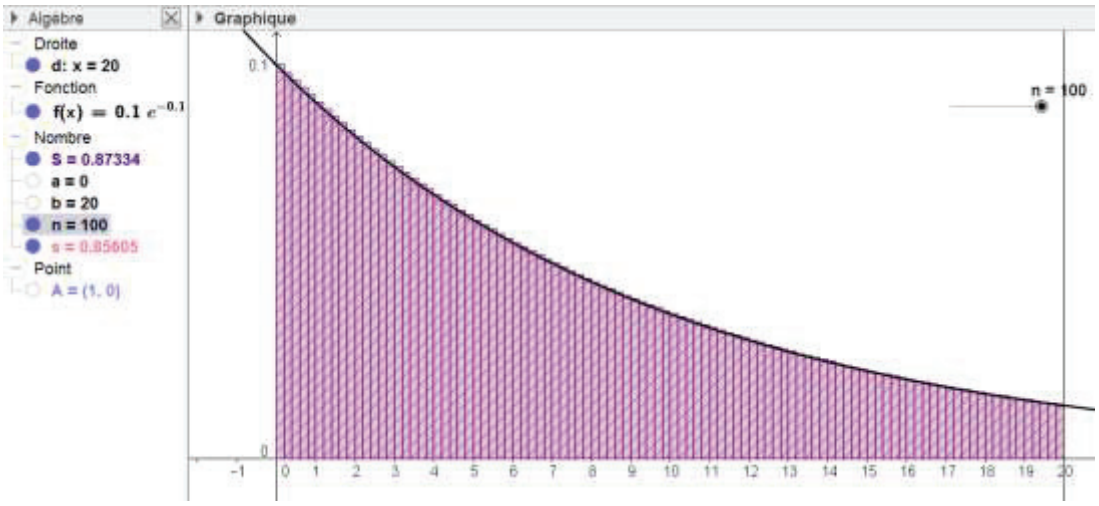
Marie : Donc ça c'est les grands rectangles, sur la même base bien sûr. Ok ? Et donc, là, l'aire vous voyez qu'elle vaut 1,05 à peu près. Alors, c'est ce que tu… ce que vous proposiez, B53, tout le monde, si on fait de plus en plus de rectangles… Alors, je vais mettre les deux en même temps. Donc, le hachurage rose, c'est les petits, ce qui est en violet, c'est les grands. Les grands ils montent de l'axe des abscisses, faîtes attention. Et donc si on augmente là par exemple… A22 qui n'avait pas l'air très convaincu… Donc déjà 10.



Marie : Vous voyez en même temps les valeurs qui bougent dans la page d'Algèbre, là ?
Les réponses sont mitigées.
Marie : Alors… Attendez. Quand n vaut 5, vous allez noter s'il vous plait, d'accord ? Pour la méthode des rectangles, quand n vaut 5, donc quand on fait 5 rectangles ça veut dire ici, on obtient une aire qui est comprise entre 0,70323 avec la précision que j'ai sur le logiciel et 1,0491. B11, tu peux me servir de script au tableau pour écrire les valeurs approchées ? Je te dicte.
L'élève B11 va au tableau.
Marie : Donc, quand on prend 5 rectangles…
Elève B11 : Je peux effacer ?
Marie : Oui, oui, efface tout. Quand on prend 5 rectangles, on obtient que notre aire est comprise alors entre… allez 0,70 ça nous suffira. Non, tu mets un encadrement s'il te plait. *A* compris entre 0,70 et 1,05.
[…]
Marie : Quand on prend n=10, on obtient que l'aire est comprise entre 0,78 et 0,96 (sic). Alors, regardez ce qui se passe visuellement quand je fais vraiment… Regardez.







Marie : Là quand même, l'erreur elle devient pas très grande. D'accord ? Et évidemment… Par exemple, pour n=100, vous voyez que les aires des petits rectangles, ceux qui sont sous la courbe, et la somme des aires des grands rectangles, au-dessus de la courbe, ça commence vraiment… enfin là, sensiblement on ne voit pas trop la différence. Bon il y a des problèmes de zooms mais quand même. Alors je te donne les valeurs pour n=100. On trouve que c'est compris entre 0,85 et 0,88. Entre 0,85 et 0,88. Et attends, je vous le donne pour 1000. Pour 1000, on trouve que l'aire est comprise entre… alors en borne inf 0,8638 et 0,8656.

$* n = 5$	$1{,}05 > A > 0{,}7$
$* n = 10$	$0,96 > A > 0{,}78$
$* n = 100$	$0{,}88 > A > 0{,}85$
$* n = 1000$	$0{,}8638 > A > 0{,}8656$

Elève A52 : C'est l'aire de quoi ?
Marie : Comment ?
Elève A52 : C'est l'aire de quoi ?
Marie : La borne inf, c'est l'aire des petits rectangles sous la courbe, la somme des aires des petits rectangles roses. Attends, je les diminue parce qu'on voit plus rien. D'accord, la borne inf c'est la somme des aires des petits rectangles, ceux qui sont construits sous la courbe, et la borne sup c'est celle… c'est la somme des aires des grands rectangles. Et tu vois bien que si j'affiche les deux et je fais croître n, donc si je réduis les amplitudes, mon encadrement devient… de mieux en mieux en terme… enfin pour approcher l'aire et ça se voit visuellement puisque l'erreur… Tu vois l'erreur elle est là, l'amplitude elle est là, c'est la

petite différence qu'il y a en haut et elles sont de plus en plus petites. Donc en fait, cette méthode des rectangles, elle va nous permettre… Qu'est-ce qui va se passer, à votre avis, si je fais tendre n vers +∞ ?
Plusieurs élèves répondent en même temps.
Marie : Ca va tendre vers l'aire parce que… Qu'est-ce qu'on a fait quand on a fait ça ? Qu'est-ce qu'on va obtenir ? A42. A42, qu'est-ce qui se passe ? Je vais l'appeler a_n, l'aire qui est en-dessous…
Elève A52 : Ca tend vers l'aire.
Marie : Voilà. On va avoir notre aire, notre A, qui est comprise entre… on va écrire ça comme ça, ça c'est la somme des aires des petits rectangles, ça c'est la somme des aires des grands rectangles et quand n tend vers +∞, qu'est-ce qui se passe ?
Plusieurs élèves parlent en même temps.
Marie : Qu'est-ce qu'on va obtenir pour la limite de a_n visiblement ? En tout cas, qu'est-ce qui se passe quand ces deux…
Un élève dit « ils sont égaux » ; un autre « ils tendent vers A ».
Marie : Alors, ils tendent vers le même nombre effectivement. C'est celui-ci qu'on va appeler… qu'on va donner comme valeur à notre aire. D'accord ? Ça marche aussi avec la méthode des trapèzes.
Elève A42 : Donc ça veut dire que si on fait des suites, ça marche.
Marie : Tout ceci ne vous dit pas encore comment exactement on va calculer l'aire sous nos courbes, d'accord. Mais en tout cas, ça montre que cette aire elle existe et qu'on va pouvoir en avoir des valeurs approchées. Et donc nous on va voir comment on peut calculer des valeurs exactes. Avec tout ça, on va pouvoir calculer nos probabilités maintenant, quand on saura faire ça. D'accord ? Quand on saura calculer l'aire sous une courbe d'une fonction continue positive sur un intervalle donné et bien on pourra calculer nos probabilités. Vous avez pris des notes ? Vous avez vu ce qui se passait pour les rectangles ? Vous avez compris ? Oui ? La méthode des rectangles, on va se ré-entrainer. Tiens, tu vas donner ça à tout le monde, pour que au moins il y ait un document avec des rectangles propres.
Un élève distribue un polycopié.
Marie : C'est sur votre courbe. Tout le monde a bien vu que ce que ce que j'appelle les petits rectangles, c'est ceux qui sont sous la courbe, ils sont hachurés sur votre dessin. Et les grands rectangles, ils partent bien de l'axe des abscisses et ils montent tout en haut, ils sont un peu grisés sur le dessin.
[83:30] Marie : Alors, vous allez prendre votre cours, s'il vous plait.
Les élèves sortent les affaires dont ils ont besoin.

I.5 Documents complémentaires

I.5.1 Notes d'élèves sur le problème de la rencontre et le problème du volcan Aso

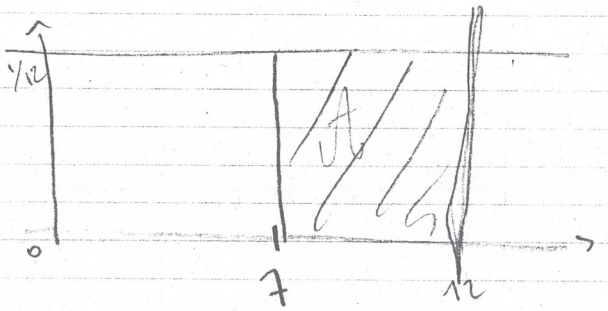
a) Notes de l'élève C41

4. Alea( ) renvoie un réel entre 0 et 1.

~~12 Alea( ) renvoie un réel~~ entre 0 et 12.

Ce phénomène s'appelle la fluctuation d'échantillonnage.

La somme des aires des rectangles de l'histogramme vaut 1.



Donc $p_1 = p(X \in [7;12])$ est l'aire sous la courbe de tendance entre 7 et 12. $p_1 = 5 \times \frac{1}{12} = \frac{5}{12}$

$p_2 = p(X = 3) = 0$ car on obtient un segment.

Problème 1 :

Karine et Olivier décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Ils peuvent arriver à tout moment entre 7h et 8h. Que peut-on dire du temps d'attente du premier arrivé ?

X varie entre 0 et 60 minutes. On considère l'axe du temps comme l'axe des réels
Le temps entre d'arrivé entre 7h et 8h est continu.
$p(X \in [0;60])$, Si on note X la v.a. égale au temps d'attente du 1er arrivé.
⇒ Lorsqu'on prend une situation concrète et que l'on étudie en maths, il faut modifier le problème pour l'adapter à un modèle mathématique.

On note T_K, T_0 le temps d'arrivée de K et O. $X = |T_K - T_0|$

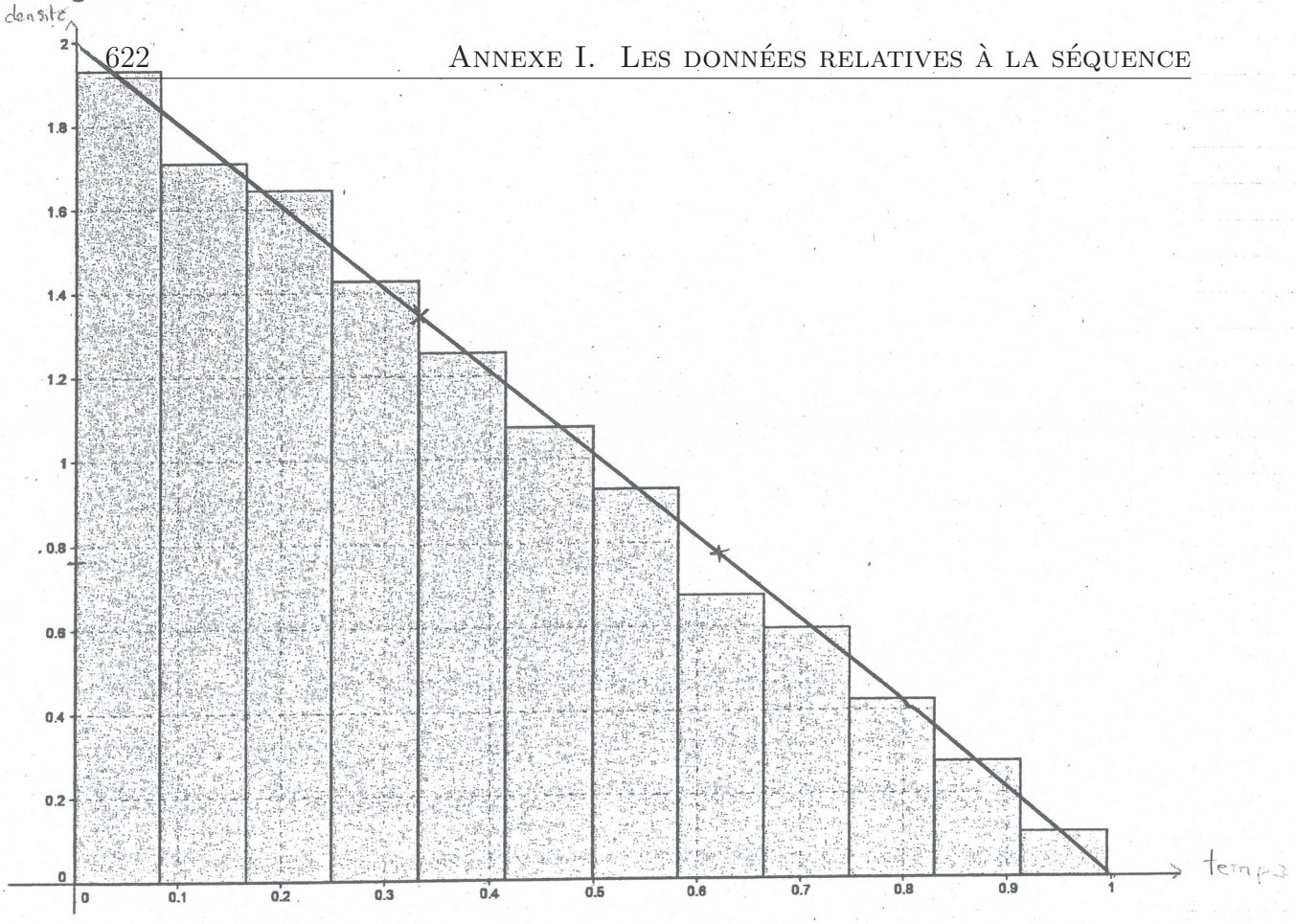
X est une variable aléatoire qui prend toutes les valeurs de [0,60].

déclarer ⇒ valeur (Alea() - Alea())

Traçons une courbe de tendance pour calculer l'aire de la partie entre 0,2 et 0,3.
Cette droite doit passer par 2 points. tq $y = ax + b$.
tq $f(1) = 0$ et d'aire sous la courbe vaut 1. donc $b = 1$
$A_{(triangle)} = \frac{b \times h}{2} = 1 \iff \frac{1 \times h}{2} = 1 \iff h = 2 \iff f(0) = 2$

Donc $y = -2x + 2$

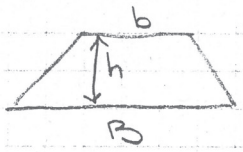
Histogramme d'un échantillon de taille 10 000, classe d'amplitude 5 minutes.



Exercice 1 : En utilisant le modèle mis en place, répondez aux questions suivantes.

a) Quelle est la probabilité que le premier arrivé attende exactement 40 minutes ? 0,63
b) Quelle est la probabilité que le premier arrivé attende plus de 40 minutes ? 0,63
c) Quelle est la probabilité que le premier arrivé attende moins de 40 minutes ? 0,63
d) Quelle est la probabilité que le premier arrivé attende entre 20 et 40 minutes ? 0,33 et 0,63
e) Quelle est la probabilité que Karine et Olivier arrivent en même temps ?
f) Finalement, ils ont décidé de ne pas attendre plus d'un quart d'heure. Quelle est la probabilité qu'ils se retrouvent ?

Rappels :



$A_{trapèze} = \frac{(b+B)\times h}{2}$

a) $A = \frac{(0+0)\times\frac{2}{3}}{2}$

b) $A \geqslant \frac{\frac{1}{3}+f(\frac{2}{3})}{2} = \frac{1}{9}$

c) $A \leqslant \frac{(0,77+2)\times 0,63}{2} \simeq 0,87$

d) $A = \frac{(1,37+0,77)\times 0,3}{2} = 0,32$

e) Comme la différence d'attente entre les deux amis est de 0, alors la hauteur est de 0 donc $A = \frac{(x+y)\times 0}{2} = 0$

f) $p\left(X \leq \frac{1}{4}\right) = \frac{(2+f(\frac{1}{4}))\times\frac{1}{4}}{2} = \frac{7}{16}$

Donc $p\left(X \in \left[\frac{2}{3};\frac{2}{3}\right]\right) = 0$

$p\left(\left(X \in \left[\frac{2}{3};1\right]\right)\right) = \frac{1}{9}$

$p\left(X \in \left[0;\frac{2}{3}\right]\right) = 1 - \frac{1}{9} = \frac{8}{9}$

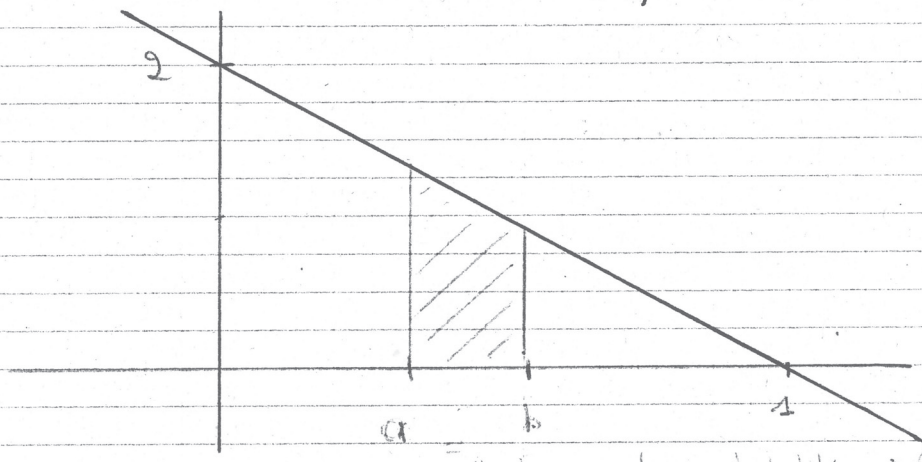
$p\left(\frac{1}{3} \leq X \leq \frac{2}{3}\right) = \frac{\left(f(\frac{1}{3})+f(\frac{2}{3})\right)\times\frac{1}{3}}{2} = \frac{1}{3}$

Synthèse : On a noté T_O : le moment d'arrivée d'Olivier
T_K : de Karine

T_K et T_O sont des variables aléatoires qui prennent toutes les valeurs réelles de $[7;8]$ que l'on peut ramener à $[0;1]$
On dit que ce sont des variables aléatoires continues qui suivent le modèle de la puce

On note X la variable aléatoire égale au temps d'attente du 1er arrivé. X prend toutes les valeurs de $[0;1]$.

Pour décrire la loi de probabilité, on a introduit une fonction f définie sur $[0;1]$ par $f(x) = -2x + 2$ telle que la courbe de f "lisse le haut des" rectangles de l'histogramme.
Ainsi pour calculer la probabilité $p(X \in [a;b])$ où $[a;b] \in [0;1]$.
On calcule l'aire sous la courbe $\mathcal{C}_f$ entre a et b.



aire hachurée ⇒ la probabilité $p(X \in [a;b]) = p(a \leqslant X \leqslant b)$

f s'appelle <u>la fonction de densité</u> de probabilité de X.

f possède les propriétés suivantes : — f est positive sur $[0;1]$.
— f est continue sur $[0;1]$.
— l'aire sous la courbe $\mathcal{C}_f$ entre 0 et 1 vaut 1.

Problème 2 :

Le volcan Aso, situé sur l'île de Kyushu au Japon, est l'un des plus actifs au monde. On possède les statistiques de ses éruptions, régulièrement tenues depuis le XIIIè siècle. Nous avons, dans le document tableur, les années d'éruptions jusqu'au XIXè siècle (à partir du XXè siècle, les données, d'une autre nature, ne sont pas compatibles(*)).On s'intéresse au temps écoulé, en années, entre deux éruptions successives.

Années des 73 éruptions du volcan Aso du XIIIe au XIXe siècle

Donnée n°	Année	Temps d'attente	Donnée n°	Année	Temps d'attente
1	1229		34	1562	4
2	1239	10	35	1563	1
3	1240	1	36	1564	1
4	1265	25	37	1576	12
5	1269	4	38	1582	6
6	1270	1	39	1583	1
7	1272	2	40	1584	1
8	1273	1	41	1587	3
9	1274	1	42	1598	11
10	1281	7	43	1611	13
11	1286	5	44	1612	1
12	1305	19	45	1613	1
13	1324	19	46	1620	7
14	1331	7	47	1631	11
15	1335	4	48	1637	6
16	1340	5	49	1649	12
17	1346	6	50	1668	19
18	1369	23	51	1675	7
19	1375	6	52	1683	8
20	1376	1	53	1691	8
21	1377	1	54	1708	17
22	1387	10	55	1709	1
23	1388	1	56	1765	56
24	1434	46	57	1772	7
25	1438	4	58	1780	8
26	1473	35	59	1804	24
27	1485	12	60	1806	2
28	1505	20	61	1814	8
29	1506	1	62	1815	1
30	1522	16	63	1826	11
31	1533	11	64	1827	1
32	1542	9	65	1828	1
33	1558	16	66	1829	1
			67	1830	1
			68	1854	24
			69	1872	18
			70	1874	2
			71	1884	10
			72	1894	10
			73	1897	3

Le Volcan Aso est actuellement en éruption.

Comment évaluer la proba que la prochaine éruption a) ait lieu dans les 5 ans?

b) Au cours de l'année 2030?

- probabilités discrètes
- f[illegible] d'attente

b) - ajouter cette fréquence k fois à 1897. avec $k \in \mathbb{N}$

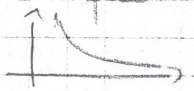
a) - voir si cette moyenne est de 5 années

faire la probabilité de chaque temps d'attente ~~cette~~ entre 1229 et 1897

temps d'attente continu mais données du doc discrètes

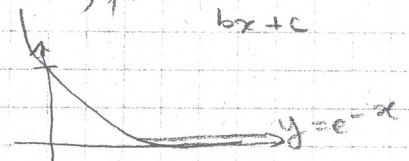
Histogramme $[0;5[$ ou $]0;5]$ (cf a)

courbe de tendance plus ou moins identique qqsoit la tranche d'âge



On cherche une fonction f définie sur $[0;+\infty[$ qui approche "au mieux" l'histogramme

$f(x) = d + \frac{a}{bx+c}$ avec $\lim\limits_{x \to +\infty} f(x) = 0$ et $d=0$; $f(x) = \frac{a}{bx+c}$



Proposition 1 $f(x) = 1/x$

Proposition 2 $f(x) = e^{-x}$

avec $f(0) = 0,1 \Rightarrow f(x) = e^{-x} - 0,9$

$\Leftrightarrow f(x) = e^{-x} \times 0,1$

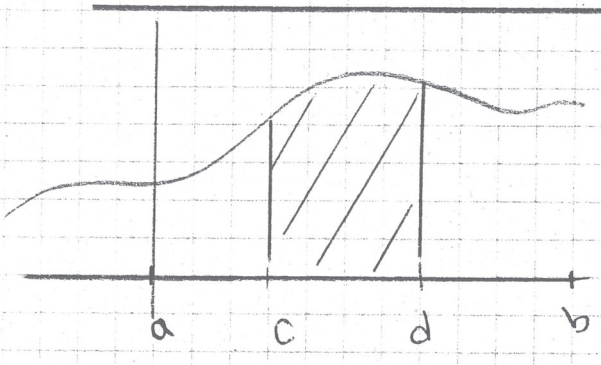
$\Rightarrow b e^{\lambda x}$ avec $\lambda = -0,1$ et $b = 0,1$

a) $p(X \in [0;5]) \approx 0,39$

b) $p(X \in [15;16[)$

Hypothèse: Le volcan est en éruption le 18 mars 2015 $p(14,75 \leq X \leq 15,75) = 0,02$

$14 + \frac{86}{365}$ $16 - \frac{33}{365}$



X v.a. continue

f déf. $[a;b]$ continue, positive, aire sous la courbe C_f entre a et b = 1

$p(c \leq X \leq d)$ aire sous la courbe C_f entre c et d

Algorithme

Variable

$f(x) = 0,1e^{-0,1x}$

$f'(x) = 0,1 \times (-0,1) \times e^{-0,1x} = -0,01 \times e^{-0,1x}$

$y = f'(a)(x-a) + f(a)$

$y = (-0,01 \times e^{-0,1a})(x-a) + 0,1e^{-0,1a}$

0 -> A

Pour (X ; 1

Algorithme

Variables

0 → A

0 → Y

Pour (X; 0 à 19)

$\frac{(f(x)+f(x+1))}{2} \rightarrow Y$

A + Y → A'

Fin pour

Affiches A'

Variables

0 → A''

Pour (X, 1 20)

$(2 \times 0,1 e^{-x0,1} + 0,01 e^{-x0,1}) \div 2 + A'' \rightarrow A''$

Fin

Afficher A''

$A \in [A'; A'']$

$\left(f'(11)(11-11) + \frac{f(11)}{2} + f'(11)(10-11) + f(11)\right) \times 1$

$= \frac{f(11) - f'(11) + f(11)}{2} = \frac{2f(11) - f'(11)}{2}$

$f(x) = 0,1e^{-x0,1}$ et $f'(x) = -0,01e^{-0,1x}$

Méthode rectangles

n : nombre de rectangles

pour n = 5 $0,7723 \leq A \leq 1,0491$

pour n = 10 $0,78 \leq A \leq 0,96$

pour n = 100 $0,85 \leq A \leq 0,88$

pour n = 1000 $0,8638 \leq A \leq 0,8656$

Feuille 1 Exercice 1 Calcul intégral

Notons $\mathcal{A}_5$ la somme des aires des petits rectangles

$A_5 = 0,2 \times f(1,2) + 0,2 \times f(1,4) + 0,2 \times f(1,6) + 0,2 \times (1,8) + 0,2 \times f(2)$

$= 0,2 \left(\frac{1}{1,2} + \frac{1}{1,4} + \frac{1}{1,6} + \frac{1}{1,8} + \frac{1}{2} \right) \simeq 0,64$

Notons $\mathcal{S}_5$ la somme des aires des grands rectangles :

$S_5 = 0,2 \times f(1,0) + (0,2) \times f(1,2) + 0,2 f(1,4) + 0,2 f(1,6) + 0,2 f(1,8)$

$= 0,2 \left(\frac{1}{1} + \frac{1}{1,2} + \frac{1}{1,4} + \frac{1}{1,6} + \frac{1}{1,8} \right) \simeq 0,75$

$A \in [0,64; 0,75]$

b) Notes de l'élève A12

Problème 1 :

Karine et Olivier décident de se retrouver au café de l'Hôtel de Ville entre 7h et 8h. Ils peuvent arriver à tout moment entre 7h et 8h. Que peut-on dire du temps d'attente du premier arrivé ?

"C'est un peu bizarre comme RDV" Anne-Laure
Si on considère qu'il y a autant de chance d'arriver sur chaque minute de 7^H à 8^H.
Comme le puck du DM6.
"On considère l'axe du temps comme l'axe des réels le temps d'arrivée entre 7^H et 8^H comme continu" Flore

* On note X la variable aléatoire égale au temps d'attente du 1er arrivé.
* On note T_K et T_O le temps d'arrivée de Karine et Olivier.
$X = |T_K - T_O|$ X est une variable aléatoire qui prend toutes les valeurs de $[0;1]$.

Brice : impossible que les 2 arrivent en même temps
Albane : Mais DM6, on a vu que la probabilité d'un événement peut-être nulle sans que l'événement soit impossible.

Abs(Alea() - Alea()) sur excel.

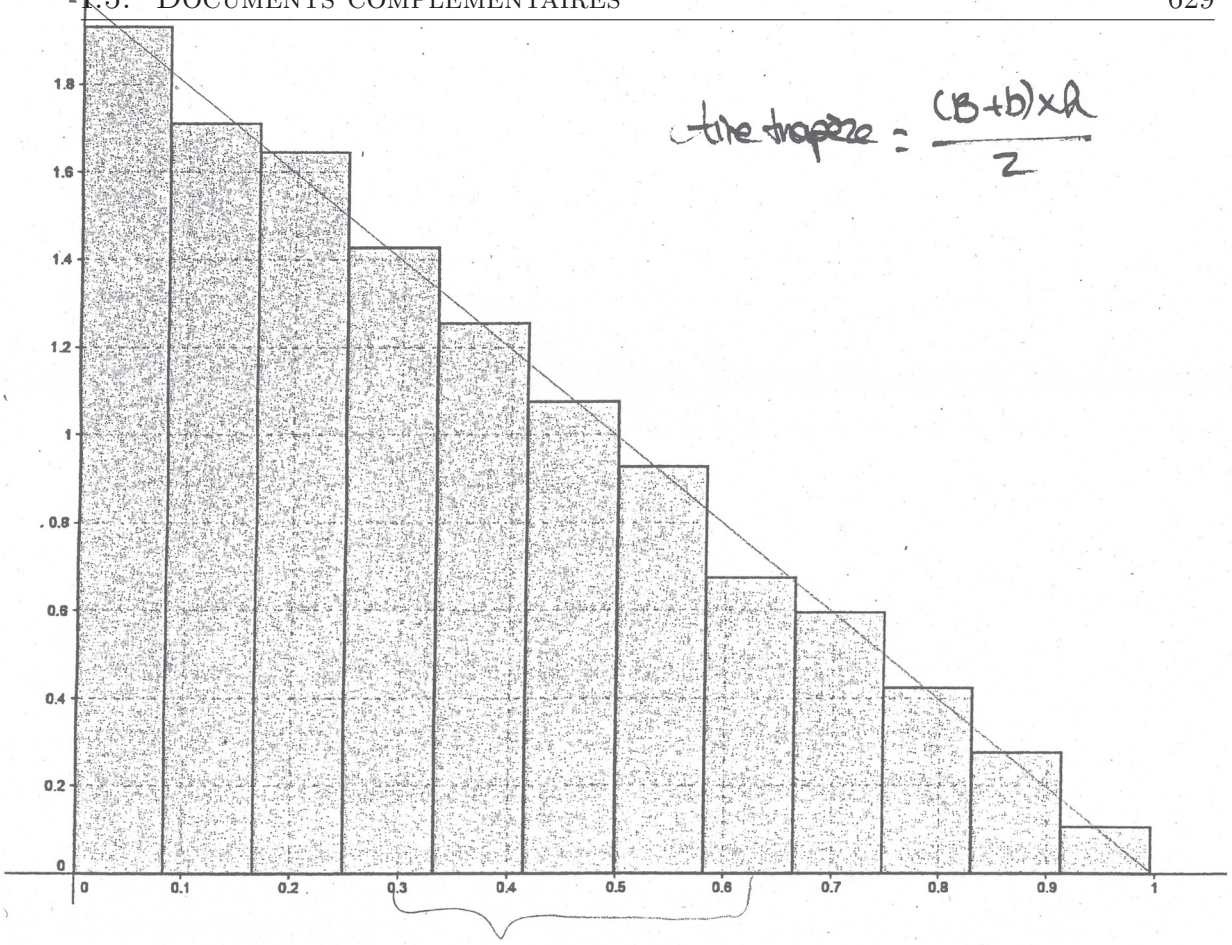
(Comment trouver $p(X \in [a;b])$ ou $[a;b] \subset [0;1]$
Tracer courbe de tendance et donc calculer la probabilité
une droite $y = ax + b$
$f(1) = 0$ et on sait que l'aire sous la courbe vaut 1.
donc $f(0) = 2$.
$\Rightarrow y = -2x + 2$.

aire triangle $= \frac{b \times h}{2}$

Exercice 1 : En utilisant le modèle mis en place, répondez aux questions suivantes.

a) Quelle est la probabilité que le premier arrivé attende exactement 40 minutes ?
b) Quelle est la probabilité que le premier arrivé attende plus de 40 minutes ?
c) Quelle est la probabilité que le premier arrivé attende moins de 40 minutes ?
d) Quelle est la probabilité que le premier arrivé attende entre 20 et 40 minutes ?
e) Quelle est la probabilité que Karine et Olivier arrivent en même temps ?
f) Finalement, ils ont décidé de ne pas attendre plus d'un quart d'heure. Quelle est la probabilité qu'ils se retrouvent ?

Histogramme d'un échantillon de taille 10 000, classe d'amplitude 5 minutes.



$\mathcal{A}$ trapèze $= \frac{(B+b)\times h}{2}$

a) $p(X=40) = 0$

b) $p(X \in [40;60]) = p(X \in [2/3; 1])$

$\mathcal{A} = \frac{b \times h}{2}$ base $= 1 - 2/3 = 1/3$.

hauteur $= f(2/3) = \frac{2}{3}$

$\mathcal{A} = \frac{(1/3)\times(2/3)}{2} = \frac{1}{9}$

c) $1 - 1/9 = \frac{8}{9}$.

d) 40 min $\Leftrightarrow$ 2/3 20 min $\Leftrightarrow$ 1/3.

$f(2/3) = 2/3$ $f(1/3) = 4/3$.

$\mathcal{A} = \frac{[(2/3)+(4/3)]\times(1/3)}{2} = \frac{2\times 1/3}{2} = \frac{1}{3}$.

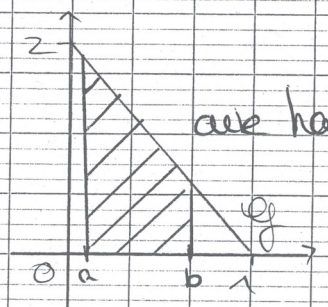
e) temps d'attente $= 0$ $p = 0$.

f) Evt $[0; 15]$. 15 min $\Leftrightarrow$ 1/4 $f(1/4) = 3/2$

$\mathcal{A}_{triangle} = \frac{(1/4)\times(3/2)}{2} = \frac{3/8}{2} = \frac{3}{16}$

$\frac{7}{16}$

Synthèse: On a noté T_O le temps d'arrivée d'Olivier et T_K le moment d'arrivée de Kawtar. T_K et T_O sont des v.a qui prenent toutes les valeurs de $[7;8]$ (que l'on peut ramener à $[0;1]$. On dit que ce sont des variables aléatoires continues qui suivent le même modèle de la pièce.
On note X la v.a égale au temps d'attente du premier arrivé. X peut toutes les valeurs de l'intervalle $[0;1]$. Pour définir la loi de probabilité de X, on a introduit une fonction f définie sur $[0;1]$ par $f(x) = -2x+2$ telle que la courbe de f "lise le haut des rectangles de l'histogramme".
Ainsi, pour calculer $p(X \in [a;b])$ où $[a;b] \in [0;1]$ on calcule l'aire sous la courbe $\mathcal{C}_f$ entre a et b.



avec hachurée égale à $p(X \in [a;b]) = p(a \leq X \leq b)$

f s'appelle la fonction de densité de probabilités de X.
f est positive sur $[0;1]$. L'aire sous $\mathcal{C}_f$ entre 0 et 1 vaut 1

Problème 2:
On fait un tableau de fréquence par des classes de temps d'attente d'amplitude $\neq$ $[1;4]$ - $[5;7]$; $[8;11]$; $[12;19]$; $[20;25]$; $[26;0]$

Le temps d'attente est continu. Lorsque qu'on parle d'une durée d'attente, la variable peut être comprise entre $[0;2]$
On fait un histogramme avec des classes d'amplitude 5.
On voudrait une courbe de tendance mais il y a des trous.
Malgrés les trous on peut trouver une courbe de tendance.

On cherche une fonction f définie sur $]0;+\infty[$ qui approche "au mieux" l'histogramme.

Prop 1 : $\frac{1}{x}$

$f(x) = d + \frac{a}{bx+c}$.

$\lim\limits_{+\infty} f = 0 \quad d = 0$ donc $f(x) = \frac{a}{bx+c}$

Prop 2 : $f(x) = \frac{1}{e^x} = e^{-x}$ définie sur $]0;+\infty[$

On veut que $f(0) = 0,1$.

$f(x) = 0,1\, e^{-\frac{x}{2}}$

Proposition : $f(0) = 0,1$

$f(x) = be^{-kx} \quad f(0) = b = 0,1$

$\Rightarrow f(x) = 0,1\, e^{-0,1x}$

$p(X \in [0;5]) \approx 0,39$.

Un cours de 2030 :

Hypothèse : le volcan a eu éruption le 18 Mai 2013.

$p(X \in [4,75; 15,75]) = 0,02$

Trouver un algorithme pour calculer l'aire sous la courbe de f entre 0 et 20

$S \approx 0,69678$

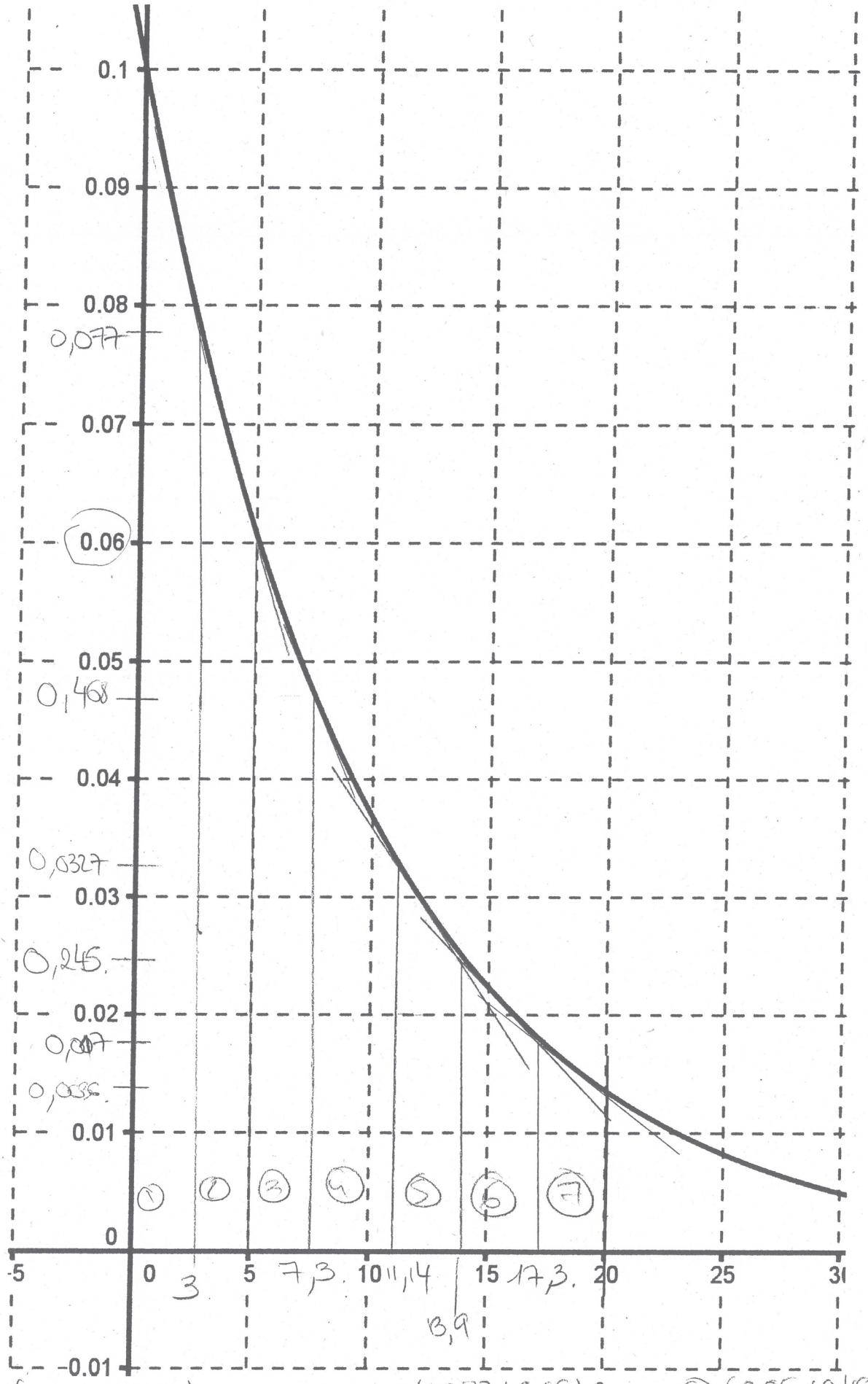
① $\frac{(0,468 + 0,0327) \times 4,1}{2} = 0,0031$

② $\frac{(0,0327 + 0,245)}{2} \times 7,5 = 0,07$

③ $\frac{0,245 + 0,017}{2} \times 3,4 = 0,07055$.

④ $\frac{0,017 + 0,0036}{2} \times 2,7 = 0,02751$

doc 1



(B+D)×h/2

① $\frac{(0,1+0,077)\times 3}{2} = 0,1655$ ② $\frac{(0,077+0,06)\times 2}{2} = 0,137$ ③ $\frac{(0,06+0,468)\times 7,3}{2}$ = 0,122282

I.5.2 Méthodes de calcul d'aire

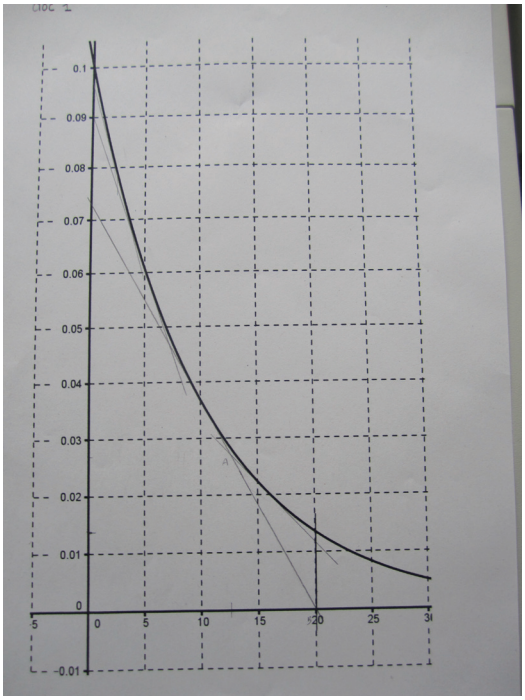


Figure I.1 – Méthode avec les tangentes rencontrée chez plusieurs élèves

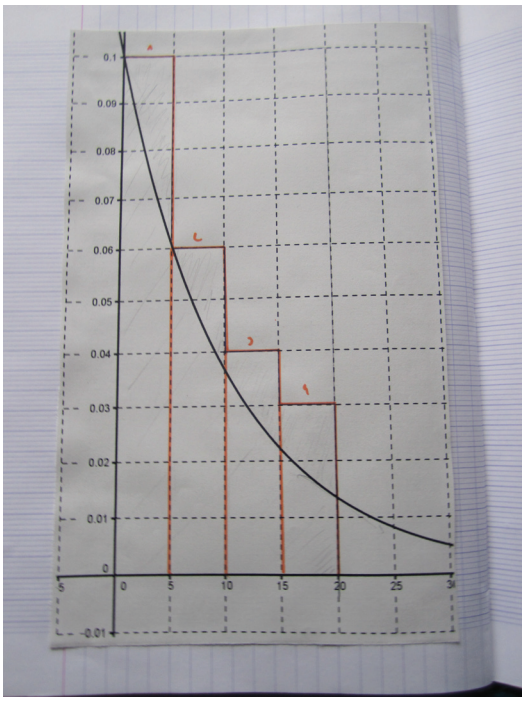


Figure I.2 – Méthode des rectangles (supérieurs) rencontrée chez deux élèves

I.6 Synthèse donnée à compléter par les élèves

Synthèse des problèmes 1 et 2 :

En classe de première, nous avons étudié des variables aléatoires qui prennent un nombre fini de valeurs : on les appelle des variables aléatoires discrètes.
Dans le dm6 et les problèmes 1 et 2, nous avons rencontré des variables aléatoires qui prennent toutes les valeurs réelles d'un intervalle (l'intervalle $[0; 12]$ pour la puce, l'intervalle $[0; 1]$ pour le temps d'attente du premier arrivé et $[0; +\infty[$ pour le volcan). Ces variables aléatoires sont dîtes continues.

Pour décrire la loi d'une variable aléatoire continue X à valeurs dans l'intervalle I, on lui associe une fonction f définie, positive et continue sur I telle que l'aire sous la courbe C_f sur I soit égale à 1.
Cette fonction est appelée fonction de densité de probabilité associée à la v.a. X.
On dit que la v.a. continue X a pour densité de probabilité la fonction f si, et seulement si, pour tout intervalle $[a; b]$ inclus dans I, la probabilité que X prenne les valeurs de $[a; b]$ soit égale à l'aire sous C_f entre a et b.

Par exemple :

- Dans le dm6, la v.a. égale à la distance parcourue par la puce jusqu'à son arrêt, a pour fonction de densité de probabilité la fonction f définie sur $[0; 12]$ par $f(x) = \frac{1}{12}$.
- Pour le problème 1, la v.a. égale au temps d'attente du premier arrivé, a pour fonction de densité de probabilité la fonction f définie sur $[0; 1]$ par $f(x) = 2 - 2x$.
- Pour le problème 2, la v.a. égale au temps écoulé entre deux éruptions successives a pour fonction de probabilité la fonction f définie sur $[0; +\infty[$ par

Pour calculer ces probabilités, on a besoin de savoir calculer l'aire sous la courbe d'une fonction positive et continue sur un intervalle.

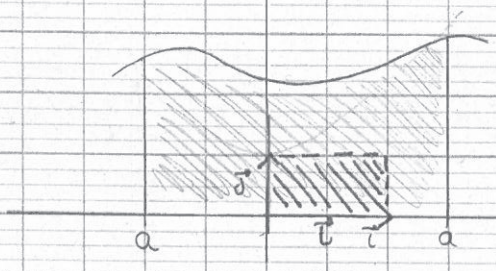
- Dans le cas, du dm6 et du problème 1, les aires à calculer sont celles de figures polygonales élémentaires (triangle, rectangle, trapèze) et on sait calculer leur aire.
- Dans le cas du volcan,

I.7 Le cours

Chap 11 : Calcul intégral et loi à densité

I) Aire sous la courbe

On considère une fonction f définie sur $[a;b]$, continue et positive sur $[a;b]$. On note $\mathcal{C}_f$ sa courbe représentative dans un repère orthogonal du plan $(O;\vec{i};\vec{j})$



On note $\mathcal{D}$ le domaine plan compris entre les droites d'éq $x=a$ et $x=b$ et $\mathcal{C}_f$ et l'axe des abscisses (hachuré)

$$\mathcal{D} = \{M(x;y);\ a \leq x \leq b \text{ et } 0 \leq y \leq f(x)\}.$$

L'unité d'aire est l'aire du rectangle construit sous les vecteurs de base.

On veut calculer l'aire en u.a du domaine $\mathcal{D}$. On dit que l'on calcule l'aire sous la courbe $\mathcal{C}_f$.

Cf ex 5

On admet que la méthode des rectangles permet de construire deux suites s_n (somme des aires des petits rectangles) et S_n (somme des aires des grands rectangles) telles que

$s_n \leq \mathcal{A}_n \leq S_n$ et $\lim\limits_{n\to+\infty} s_n = \lim\limits_{n\to+\infty} S_n = a \in \mathbb{R}$

On pose, par définition $a = \mathcal{A}_\mathcal{D}$.

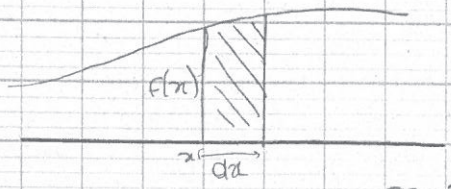
II) Intégrale d'une fonction continue positive.

a) Définition

Déf : Soit f une fct, déf, continue et positive sur $[a;b]$. On appelle intégrale de f entre a et b, la mesure de l'aire en u.a sous la courbe représentative de f ($\mathcal{C}_f$)

On note $\int_a^b f(x)\,dx$.

Rmq : on lit $\int_a^b f(x)dx$ se lit "intégrale de a à b de f(x)dx"
le symbole est un $\int$ "s" notation due à Leibniz (1646-1716)



f(x) représente l'aire d'un rectangle de base dx et de hauteur f(x).

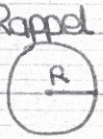
. on écrit indifféremment $\int_a^b f(x)dx = \int_a^b f(t)dx$..
x est appelé variable muette.

Exemple

① Calculer $\int_{-2}^{3} (x+2)\,dx$

Posons $f(x) = x+2$.
f est positive et continue sur $[-2; 3]$.
$\int_{-2}^{3} (x+2)dx = \mathcal{A}_D$ (u.a)
$\rightarrow \int_{-2}^{3} (x+2)dx = \frac{5\times 5}{2} = 12,5$

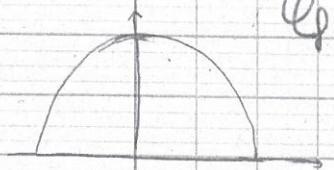
Rappel :



P. $2\pi R$
$\mathcal{A} = \pi R^2$

② Calculer $\int_{-1}^{1} \sqrt{1-x^2}\,dx$

Soit $f : x \mapsto \sqrt{1-x^2}$ $\quad D_f = [-1; 1]$



$\mathcal{C}_f$ a pour éq $y = \sqrt{1-x^2}$ $\quad y \geq 0$.

$\Leftrightarrow y^2 = 1-x^2$

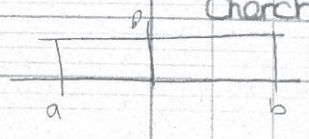
$\Leftrightarrow x^2 + y^2 = 1$ éq du cercle trigo.

$\int_{-1}^{1} \sqrt{1-x^2}\,dx = \frac{\pi \times 1^2}{2} = \frac{\pi}{2}$.

③ Soit k un réel positif

On considère la fonction f définie sur $[a; b]$ par $f(x) = k$.
Pour quelle valeur de k la fonction f est-elle une fonction densité de probabilité.

f est positive sur $[a; b]$ car $k \geq 0$. f est continue sur $[a; b]$.
Cherchons k tel que l'aire sous $\mathcal{C}_f$ entre a et b fasse 1



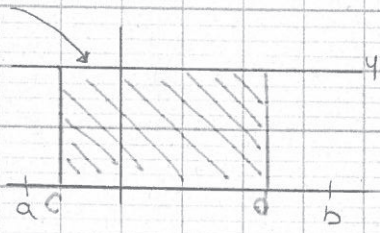
$k \times (b-a) = 1$
$\Leftrightarrow k = \frac{1}{b-a}$

exemple à connaître

$\forall x \in [a;b]$, $f(x) = \frac{1}{b-a}$.

Si une variable aléatoire continue X a pour fonction de densité de probabilité la fonction f définie sur $[a;b]$ par $f(x) = \frac{1}{b-a}$ on dit que X suit une loi uniforme sur $[a;b]$ on note $X \sim \mathcal{U}_{[a;b]}$. ($\mathcal{C}_f$ exemple de la puce)

aire hachurée $= P(X \in [c;d])$



Pour tout $[c;d]$ inclus $[a;b]$.

$P(X \in [c;d]) = P(c \leq X \leq d)$.

$= (d-c) \times \frac{1}{b-a}$

$P(X \in [c;d]) = \frac{d-c}{b-a}$

b) premières propriétés

Prop : Soit f une fonction continue et positive sur $[a;b]$

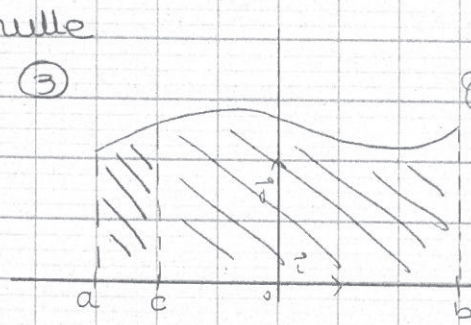
1) $\int_a^b f(x)dx \geq 0$
2) $\int_a^a f(x)dx = 0$
3) Soit c un réel de $[a;b]$ on a :

$\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx$ (Relation de Chasles)

démo ① $\int_a^b f(x)dx$ mesure une aire donc $\int_a^b f(x)dx \geq 0$.

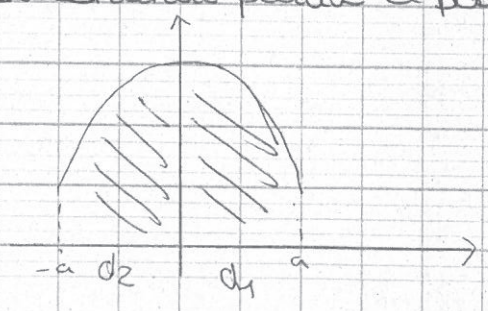
② le domaine sous $\mathcal{C}_f$ est réduit à un segment donc l'aire est nulle

③



$\int_a^c f(x)dx + \int_c^b f(x)dx$ est l'aire sous $\mathcal{C}_f$ entre a et b (additivité des aires).

Application : Soit f une fonction continue positive et paire sur $[-a;+a]$



Application (suite)

$\int_{-a}^{a} f(x)dx = \int_{-a}^{0} f(x)dx + \int_{0}^{a} f(x)dx$

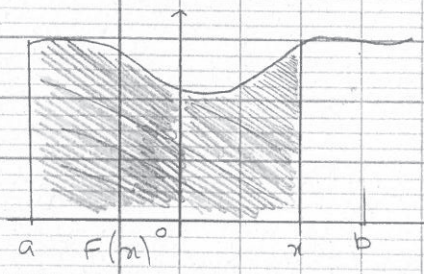
$\int_{a}^{a} f(x)dx = \mathcal{A}d_1 + \mathcal{A}d_2$

D_1 et D_2 ont la même aire par symétrie

$\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx$

c) Théorème fondamental

Soit f une fonction continue positive $[a;b]$ on considère la fonction F définie sur $[a;b]$ par $F(x) = \int_{a}^{x} f(t)dt$



$\mathcal{C}_f$ f continue et ⊕ sur $[a;b]$

F définie sur $[a;b]$ par $F(x) = \int_{a}^{b} f(t)dt$

(aire sous $\mathcal{C}_f$ entre a et x)

$[a;b] \to \mathbb{R}$

$x \to F(x)$

Considérons les pts $M(x, F(x))$ ou (x; aire ss $\mathcal{C}_f$ entre a et x)

Quelques exemples nous permettent de conjecturer que f est la dérivée de F

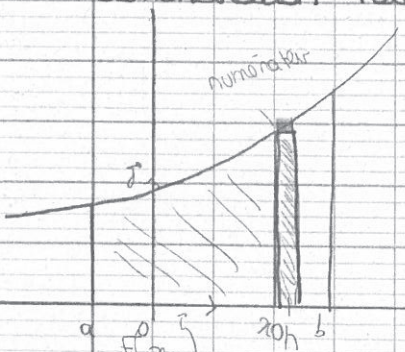
$f(x) = x^2$ $F(x) = \frac{1}{3}x^2$

$f(x) = \frac{1}{x}$ $F(x) = \ln(x)$

$f(x) = e^x$ $F(x) = e^x$

Th: Soit f une fonction continue et positive sur $[a;b]$. La fonction F définie sur $[a;b]$ par $F(x) = \int_{a}^{b} f(t)dx$ est dérivable sur $[a;b]$ et pour tout $x \in [a;b]$ $F'(x) = f(x)$

démonstration : Pour une fonction f continue, positive et croissante sur $[a;b]$

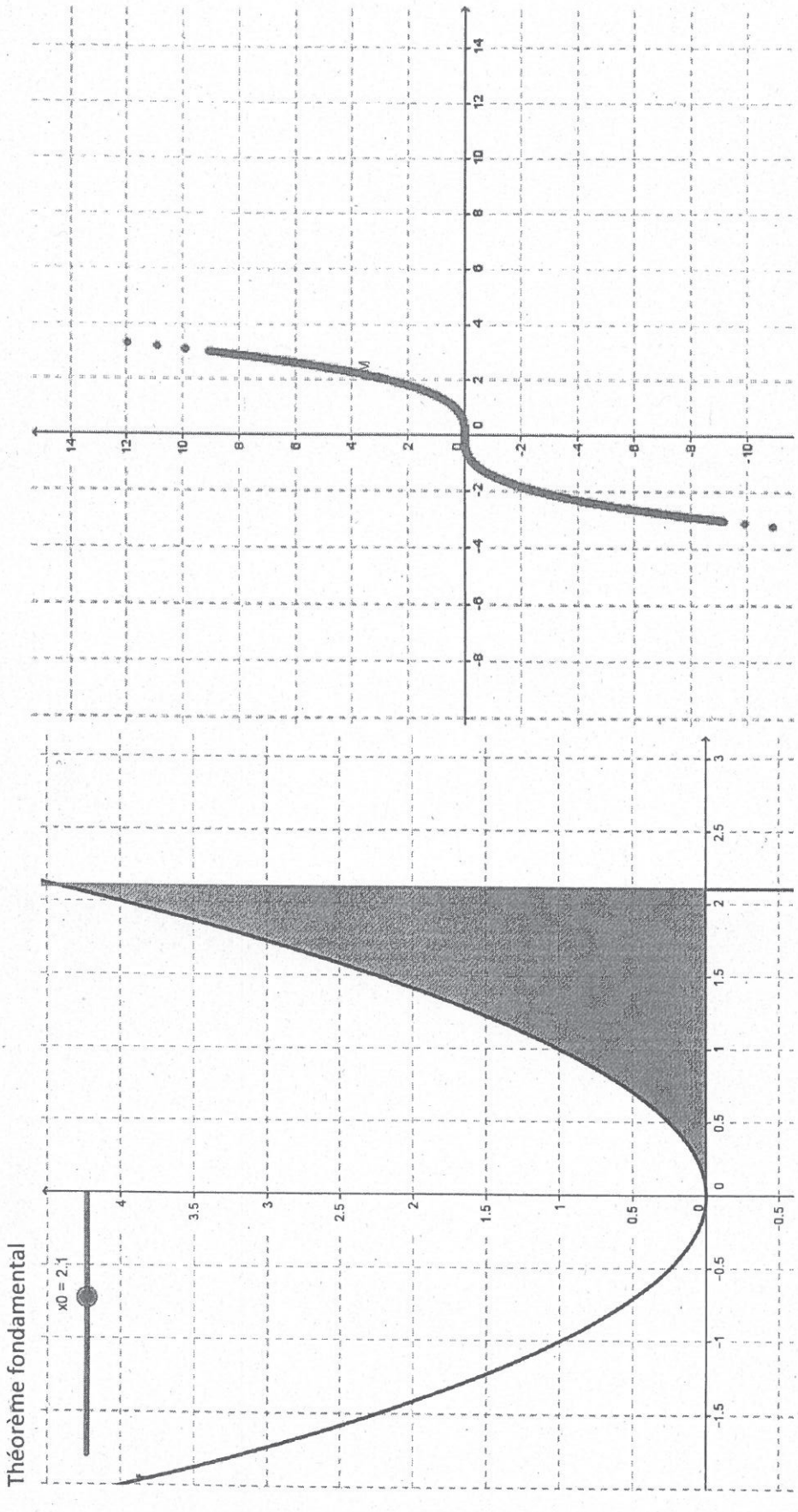


Considérons le taux d'accroissement de F en $x_0 \in [a;b]$

$\frac{F(x_0+h) - F(x_0)}{h}$

1er cas : $h > 0$

$h \times f(x_0) \leq F(x_0+h) - F(x_0) \leq h \times f(x_0+h)$.

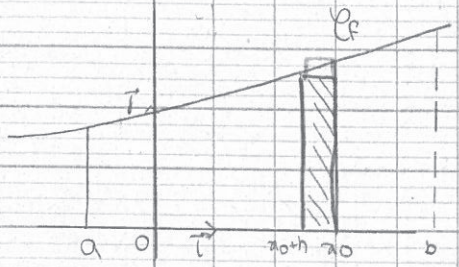

Théorème fondamental
x0 = 2,1

chap 11
②

même chose pour f

$\Rightarrow f(x_0) \leqslant \dfrac{F(x_0+h)-F(x_0)}{h} \leqslant f(x_0+h)$

2ème cas : $h<0$ $\quad \Delta\ h<0$



$-h \times f(x_0+h) \leqslant F(x_0)-F(x_0+h) \leqslant -h \times f(x_0)$

$\Rightarrow f(x_0+h) \leqslant \dfrac{F(x_0)-F(x_0+h)}{-h} \leqslant f(x_0) \quad (-h>0)$

$\Rightarrow f(x_0+h) \leqslant \dfrac{F(x_0+h)-F(x)}{h} \leqslant f(x_0)$

Or $\lim\limits_{h\to 0} f(x_0) = f(x_0)$

$\lim\limits_{h\to 0} f(x_0+h) = f(x_0)$ car f est continue en x_0.

D'après le théorème des gendarmes, $\lim\limits_{h\to 0} \dfrac{F(x_0+h)-F(x_0)}{h} = F(x)$.

Par définition F est dérivable en x_0 et $F'(x_0) = f(x_0)$

Conclusion F est dérivable sur $[a;b]$ et $\forall x \in [a;b]$

$F'(x) = f(x)$

Retour aux urnes : X v.a égale au temps des 2 éruptions

X va continue dont la fonction de densité de proba est

f définie sur $[0;+\infty[$ par $f(t) = 0,1e^{-0,1t}$

On a essayer de calculer $P(0 \leqslant X \leqslant 20) = \int_0^{20} f(t)dt$

$p(0 \leqslant X \leqslant 20) = F(20)$

où $F(x) = \int_0^{20} 0,1e^{-0,1t}\,dx$.

Cherchons une fonction définie sur $[0;+\infty[$ dont la dérivée est f.

$F(x) = -e^{-0,1x} \qquad f(x) = 0,1e^{-0,1x}$

Pb $F<0$ $\quad F'(x) = -(-0,1)e^{-0,1x} = 0,1e^{-0,1x}$

$F(x) = -e^{-0,1x} + k \quad k \in \mathbb{R}$

or $F(0) = 0$; $k = 1$

$\forall t \geq 0 \quad F(t) = -e^{-0,1t} + 1$

$P(0 \leqslant X \leqslant 20) = F(20) = -e^{-0,1\times 20} + 1 = -e^{-2} + 1$

III) Primitives

Dans ce §, les fonctions considérées ne sont plus nécessairement positives

Déf : Soit f une fonction définie sur un intervalle I. On appelle primitive de f sur I, toute fonction F définie et dérivable sur I telle que, pour tout $x \in I$ $F'(x) = f(x)$.

Rq : le Th fondamental nous dit que f est continue sur I et positive sur $[a;b]$ alors F définie sur $[a;b]$ par $F(x) = \int_a^b f(t)dt$ est une primitive de f sur $[a;b]$

Th : Toute fonction continue sur I admet des primitives sur I

ROC démo dans le cas où f est continue sur $[a;b]$.

On admet qu'alors F admet un minimum m sur $[a;b]$.

• Si $m \geq 0$ alors f est positive sur $[a;b]$

Comme est y est continue alors $x \mapsto \int_a^x f(t)dt$ est une primitive de f sur $[a;b]$

• $m < 0$. Considérons g définie sur $[a;b]$ par $g(x) = f(x) - m$.

g est continue et positive sur $[a;b]$ donc $G(x) = \int_a^x g(t)dt = \int_a^x (f(t) - m)dt$ est une primitive de g sur $[a;b]$.

Considérons $F(x) = G(x) + mx$ pour $x \in [a;b]$

F est dérivable sur $[a;b]$ comme somme de fonctions dérivables sur $[a;b]$. $\forall x \in [a;b)$

$F'(x) = G'(x) + m$

$\Leftrightarrow F'(x) = g(x) + m = f(x) - m + m$

$\Leftrightarrow F'(x) = f(x)$.

On a une primitive de f sur $[a;b]$

Prop 1 : Soit f une fonction continue sur I et F une primitive de f sur I, l'ensemble des primitives de f sur

est l'ensemble des fonctions de la forme $F+k$, $k \in \mathbb{R}$.

démo : * Soit $k \in \mathbb{R}$

$$\forall x \in I,\ (F+k)'x = F'(x) + 0 = f(x)$$

Les fonctions de la forme $F+k$ où $k \in \mathbb{R}$ sont des primitives de f.

* Soit G une primitive de f sur I.

$\forall x \in I \quad G'(x) = f(x)$

$\forall x \in I \quad F'(x) = f(x)$ (par hyp)

$\Rightarrow \forall x \in I \quad G'(x) - F'(x) = 0$

$\Rightarrow \forall x \in I \quad (G-F)'(x) = 0$

$\Rightarrow \forall x \in I \quad (G-F)x = k,\ k \in \mathbb{R}$

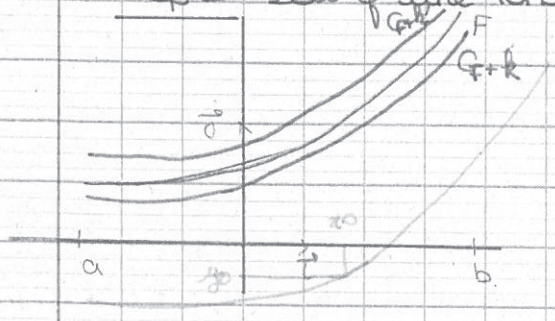
$\Rightarrow \forall x \in I \quad G(x) = F(x) + k,\ k \in \mathbb{R}$.

Rmq : On dit que sur un intervalle, deux primitives d'une même fonction diffèrent d'une constante.

Prop 2 : Soit f une fonction continue sur un intervalle I

Soit $x_0 \in I$, $y_0 \in \mathbb{R}$

Il existe une unique primitive qui prend la valeur y_0 en x_0.



Rmq : Soit f une fonction continue et positive sur $[a;b]$ la fonction F définie sur $[a;b]$ par $F(x) = \int_a^x f(t)dt$ est la primitive de f sur $[a;b]$ qui s'annule en a.

IV. Calcul intégral

Th : Soit f une fonction continue et positive sur $[a;b]$. F une primitive de f sur $[a;b]$. On a que $\int_a^b f(t)dt = F(b) - F(a)$

Ex : Calculons $I = \int_1^2 (x^2 - 1)\,dx$

Primitives des fonctions de référence

f	F	F est une primitive de f sur tout intervalle de
$f(x) = a$ où a est un réel	$F = ax + k$	$\mathbb{R}$
$f(x) = x$	$F = \frac{x^2}{2}$	$\mathbb{R}$
$f(x) = x^2$	$F = \frac{x^3}{3}$	$\mathbb{R}$
$f(x) = x^n$	$F = \frac{x^{n+1}}{n+1}$	• $\mathbb{R}$ si $n \in \mathbb{N}^*$ • $]-\infty; 0[$ ou $]0; +\infty[$ si $n \in \mathbb{Z}, n \leq -2$
$f(x) = \frac{1}{x^2}$	$F = -\frac{1}{x}$	$]-\infty; 0[$ ou $]0; +\infty[$
$f(x) = \frac{1}{x}$	$F = \ln(x)$	$]0; +\infty[$
$f(x) = e^x$	$F = e^x$	$\mathbb{R}$
$f(x) = \frac{1}{\sqrt{x}}$	$F = 2\sqrt{x}$	$]0; +\infty[$
$f(x) = \cos(x)$	$F = \sin(x)$	$\mathbb{R}$
$f(x) = \sin(x)$	$F = -\cos(x)$	$\mathbb{R}$

Formules de dérivation : $u$et $v$désignent deux fonctions définies et dérivables sur un intervalle I.

f est de la forme	F est de la forme	Conditions
$k \times u'$ où k est un réel	$k \times u$	
$u' + v'$	$u + v$	
$u' \times e^u$	(e^u)	
$u' \times u^n$ $n \in \mathbb{Z} - \{0; 1\}$	$\frac{1}{n+1} u^{n+1}$	u non nulle si $n \in \mathbb{Z}, n \leq -2$
$\frac{u'}{u^2}$	$-\frac{1}{u}$	u non nulle sur I.
$\frac{u'}{\sqrt{u}}$	$2\sqrt{u}$	u strictement positive sur I.
$\frac{u'}{u}$	$\ln(u)$.	u strictement positive sur I.

$$\int_1^2 (x^2-1)\,dx = \left[\frac{1}{3}x^3 - x\right]_1^2 = \frac{1}{3}2^2 - 2 - \frac{1}{3}1^3 + 1 =$$
$$= \frac{8}{3} - 2 - \frac{1}{3} + 1 = \frac{7}{3} - 1 = \frac{4}{3}$$

démo: f est continue et positive sur $[a;b]$ donc la fonction G définie sur $[a;b]$ par $G(x) = \int_a^x f(t)dt$ est une primitive de f sur $[a;b]$. F est une primitive de f sur $[a;b]$ donc il existe un réel k tel que, pour tout $x \in [a;b]$ $F(x) = G(x) + k$

$F(b) - F(a) = (G(b)+k) - (G(a)+k)$

$F(b) - F(a) = G(b) - G(a)$

$F(b) - F(a) = \int_a^b f(t)dt - \int_a^a f(t)dt$

$F(b) - F(a) = \int_a^b f(t)dt$ CQFD

Ce théorème est indépendant de la primitive choisie.

F(b) - F(a)
G(B) + k - G(A) - k
x(10 - x)

V. Loi à densité

cas du cadre

Remarque : Si $I = [a;+\infty[$. $\int_a^{+\infty} f(t)dt$.

Pour définir $\int_a^{+\infty} f(t)dt$ on a $\lim_{X \to \infty} \int_a^X f(t)dt$.

Si cette limite est un réel L on posera que $\int_a^{+\infty} f(t)dt$.

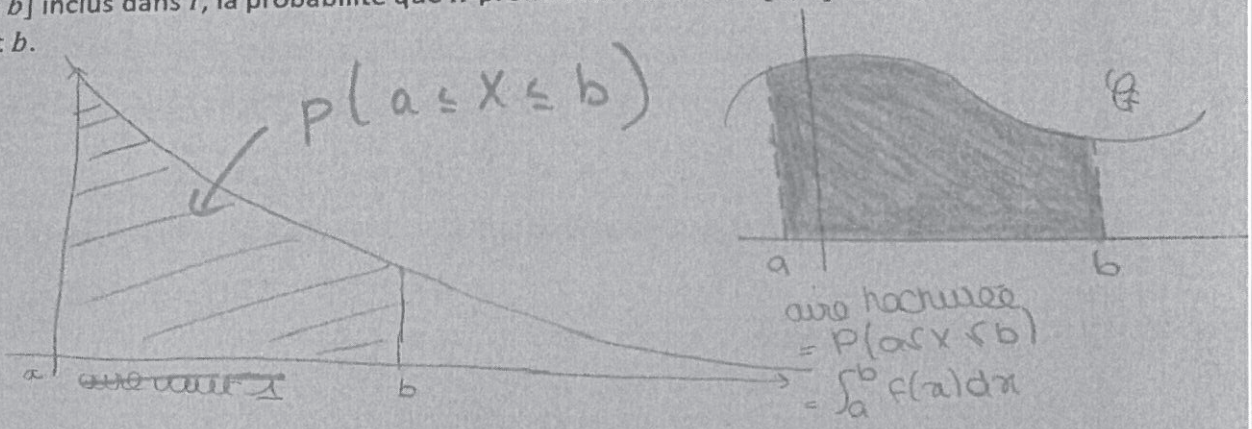
De la même façon, si $I =]-\infty; a]$ alors $\lim_{X \to -\infty} \int_X^a f(t)dt = L \in \mathbb{R}$

alors on pose $\int_{-\infty}^a f(t)dt = L$.

Synthèse des problèmes 1 et 2 :

En classe de première, nous avons étudié des variables aléatoires qui prennent un nombre fini de valeurs : on les appelle des variables aléatoires discrètes.
Dans le dm6 et les problèmes 1 et 2, nous avons rencontré des variables aléatoires qui prennent toutes les valeurs réelles d'un intervalle (l'intervalle $[0; 12]$ pour la puce, l'intervalle $[0; 1]$ pour le temps d'attente du premier arrivé et $[0; +\infty[$ pour le volcan). Ces variables aléatoires sont dîtes continues.

Pour décrire la loi d'une variable aléatoire continue X à valeurs dans l'intervalle I, on lui associe une fonction f définie, positive et continue sur I telle que l'aire sous la courbe C_f sur I soit égale à 1.
Cette fonction est appelée fonction de densité de probabilité associée à la v.a. X.
On dit que la v.a. continue X a pour densité de probabilité la fonction f si, et seulement si, pour tout intervalle $[a; b]$ inclus dans I, la probabilité que X prenne les valeurs de $[a; b]$ soit égale à l'aire sous C_f entre a et b.



Par exemple :

- Dans le dm6, la v.a. égale à la distance parcourue par la puce jusqu'à son arrêt, a pour fonction de densité de probabilité la fonction f définie sur $[0; 12]$ par $f(x) = \frac{1}{12}$.
- Pour le problème 1, la v.a. égale au temps d'attente du premier arrivé, a pour fonction de densité de probabilité la fonction f définie sur $[0; 1]$ par $f(x) = 2 - 2x$.
- Pour le problème 2, la v.a. égale au temps écoulé entre deux éruptions successives a pour fonction de probabilité la fonction f définie sur $[0; +\infty[$ par

Pour calculer ces probabilités, on a besoin de savoir calculer l'aire sous la courbe d'une fonction positive et continue sur un intervalle.

- Dans le cas, du dm6 et du problème 1, les aires à calculer sont celles de figures polygonales élémentaires (triangle, rectangle, trapèze) et on sait calculer leur aire.
- Dans le cas du volcan, l'aire à calculer était l'aire de trapèzes collés puis on sommait toutes ces aires. Plus l'amplitude de la base était petite plus la précision était grande.

Propriétés : Soit X une variable aléatoire continue sur un intervalle I.

(i) Pour tout intervalle J inclus dans I, $P(X \in J) \geq 0$.

(ii) Pour tout $a \in I, P(X = a) = 0$.

(iii) Pour tout $a \in I, P(X \leq a) = P(X < a)$.

(iv) Pour tous $a, b, c \in I$, tels que $a \leq b \leq c$,

$$P(a \leq X \leq c) = P(a \leq X \leq b) + P(b < X \leq c).$$

(v)

(vi) Pour tout $a \in I, P(X > a) = 1 - P(X \leq a)$.

(vii) Pour tous $a, b \in I$, tels que $a < b, P(a \leq X < b) = P(X < b) - P(X < a)$.

Démonstration : voir exercice 7

Exemple 1 : La production quotidienne d'un produit en tonnes est une variable aléatoire continue dont la fonction de densité de probabilité est la fonction f définie sur $[0; 10]$ par $f(x) = 0{,}006(10x - x^2)$.

1. Vérifier que f est une fonction de densité de probabilité.

• f est continue sur $[0;10]$ Cf f fonction polynome

• $\int_0^{10} 0{,}006(10x - x^2) = \left[0{,}006\left(5x^2 - \frac{x^3}{3}\right)\right]_0^{10}$

$= 0{,}006\left(5 \times 10^2 - \frac{10^3}{3}\right) = 1$

• $\forall x \in [0;10]$ $f(x) = 0{,}006\,x(10-x)$

sur $[0;10]$; $0{,}006\,x \geq 0$

x	0		10
$10-x$	$+$	0	$-$
f	0	$+$	0

Sur $[0;10]$; $f \geq 0$

2. Calculer $p(X \leq 7)$. En déduire la probabilité de l'événement « la production quotidienne dépasse 7 tonnes ».

$P(X \leq 7) = \int_0^7 0{,}006(10x - x^2)\,dx$

$P(X \leq 7) = \frac{98}{125} = 0{,}784$.

Calculons $P(X > 7)$

$P(X > 7) = 1 - P(X \leq 7) = 0{,}216$

Des modèles à connaître

a) ***Loi uniforme*** (Cf la puce)

Définition 1 : Soit a et b deux réels tels que $a < b$. On dit qu'une variable aléatoire X suit une loi uniforme sur l'intervalle $[a; b]$ si, et seulement si, sa fonction de densité de probabilité est la fonction définie sur $[a; b]$ par $f(x) = \frac{1}{b-a}$. On note $X \sim U_{[a;b]}$.

Nous avons déjà vérifié que f est bien une densité de probabilité.

Propriété 1 : Pour tout intervalle $[c; d]$ inclus dans $[a; b]$, $p(X \in [c; d]) = \frac{d-c}{b-a}$.

Démonstration : Faite en application dans le cours.

Définition 2 : Soit $X \sim U_{[a;b]}$. On appelle espérance de X et on note $E(X)$, le réel défini par $E(X) = \int_a^b tf(t)dt$.

Propriété 2 : Si $X \sim U_{[a;b]}$ alors $E(X) = \frac{a+b}{2}$.

Démonstration :

$E(X) = \int_a^b \frac{t}{b-a}\,dt = \left[\frac{1}{b-a} \times \frac{1}{2}t^2\right]_a^b = \frac{1}{b-a} \times \frac{1}{2}b^2 - \frac{1}{b-a} \times \frac{1}{2}a^2$

$= \frac{1}{2(b-a)}(b^2 - a^2) = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}$

b) Loi exponentielle

Propriété 1 : Soit λ un réel strictement positif. La fonction f définie sur $[0; +\infty[$ par $f(x) = \lambda e^{-\lambda x}$ est une fonction de densité de probabilité.

Démonstration : $f(x) = \lambda e^{-\lambda x}$ $x \geq 0$ f est continue sur $\mathbb{R}^+$

. $\lambda > 0$, $\forall x \geq 0$ $e^{-\lambda x} > 0$ donc f est positive sur $\mathbb{R}^+$.

. $\int_0^X f(x)dx = \int_0^X \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x}\right]_0^X$

$= -e^{-\lambda X} - (-e^{-\lambda \times 0}) = -e^{-\lambda X} + 1$

$\lim\limits_{X \to +\infty} -\lambda X = -\infty$ $(\lambda > 0)$ } par composition $\lim\limits_{x \to +\infty} e^{-\lambda X} = 0$

$\lim\limits_{x \to -\infty} e^x = 0$

donc $\lim\limits_{X \to +\infty} -e^{-\lambda X} + 1 = 1$

ainsi $\int_0^{+\infty} \lambda e^{-\lambda x} dx = 1$

Définition 1 : On dit qu'une variable aléatoire X suit une loi exponentielle de paramètre λ sur si, et seulement si, sa fonction de densité de probabilité est la fonction définie sur $[0; +\infty[$ par $[0; +\infty[$ $f(x) = \lambda e^{-\lambda x}$. On note $X \sim \mathcal{E}(\lambda)$.

Propriété 2 : Soit $X \sim \mathcal{E}(\lambda)$. Pour tout réel a, b tels que $0 \leq a \leq b$:

a) $p(X \leq b) = 1 - e^{-\lambda b}$

b) $p(X > b) = e^{-\lambda b}$

c) $p(a < X \leq b) = e^{-\lambda a} - e^{-\lambda b}$

Démonstration :

a) $P(X \leq b) = \int_0^b \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x}\right]_0^b = -e^{-\lambda b} - (e^{-\lambda \times 0}) = 1 - e^{-\lambda b}$

b) $P(X > b) = 1 - P(X \leq b)$

$P(X > b) = 1 - (1 - e^{-\lambda b}) = e^{-\lambda b}$

c) $P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = (1 - e^{-\lambda b}) - (1 - e^{-\lambda a})$

$P(a \leq X \leq b) = e^{-\lambda a} - e^{-\lambda b}$.

Définition 2 : Soit $X \sim \mathcal{E}(\lambda)$. On appelle espérance de X et on note $E(X)$, le réel défini par $E(X) = \int_0^{+\infty} t\lambda e^{-\lambda t} dt$.

Propriété 3 : Si $X \sim \mathcal{E}(\lambda)$ alors $E(X) = \frac{1}{\lambda}$.

ROC

Démonstration :

Lemme : la fonction $f : t \mapsto \lambda t e^{-\lambda t}$ admet pour primitive sur $[0; +\infty[$ la fonction $F : t \mapsto \left(-t - \frac{1}{\lambda}\right) e^{-\lambda t}$

$\left\{ \begin{array}{l} x \mapsto -x - \frac{1}{\lambda} \text{ est dérivable sur } [0; +\infty[\\ x \mapsto e^{-\lambda x} \text{ est dérivable sur } [a; +\infty[\end{array} \right\}$ F est dérivable sur $[0; +\infty[$

$\forall x \in [0; +\infty[$

$F'(x) = -e^{-\lambda x} + \left(-x - \frac{1}{\lambda}\right)(-\lambda e^{-\lambda x})$ $F'(x) = -e^{-\lambda x} + x\lambda e^{-\lambda x} + e^{-\lambda x}$

$F(x) = x\lambda e^{-\lambda x}$ donc $F'(x) = f(x)$

démo : Soit $X \geq 0$

$I(x) = \int_0^X \lambda x e^{-\lambda x} dx = \left[\left(-x - \frac{1}{\lambda}\right) e^{-\lambda x}\right]_0^X = \left(\left(-X - \frac{1}{\lambda}\right) e^{-\lambda X}\right) + \frac{1}{\lambda} = -Xe^{-\lambda X} - \frac{1}{\lambda} e^{-\lambda X} + \frac{1}{\lambda}$

$\lim\limits_{X \to +\infty} -\frac{1}{\lambda} e^{-\lambda X} = 0$

$\lim\limits_{X \to +\infty} -Xe^{-\lambda X} = 0$ (C.O)

} $\lim\limits_{X \to +\infty} I(x) = \frac{1}{\lambda}$ donc $E(x) = \frac{1}{\lambda}$

Exemple 2 : La durée de vie d'une ampoule, en heures, est une variable aléatoire continue qui suit une loi exponentielle de paramètre 0,002.

a) Calculer la probabilité pour qu'une ampoule ne tombe pas en panne avant 100 heures d'utilisation.

$P(X \geq 100) = e^{-0,002 \times 100} = e^{-0,2} = 0,82$

b) Calculer la probabilité pour qu'une telle ampoule fonctionne au bout de 600h sachant qu'elle fonctionne encore au bout de 500h.

$P_{(X \geq 500)}(X \geq 600) = \dfrac{P(X \geq 500) \cap P(X \geq 600)}{P(X \geq 500)}$

$= \dfrac{P(X \geq 600)}{P(X \geq 500)} = \dfrac{e^{-1,2}}{e^{-1}} = e^{-0,2} = 0,82$

Propriété 4 (Durée de vie sans veillissement) :
Si $X \sim \mathcal{E}(\lambda)$ alors pour tous réels positifs t et h, on a : $p_{(X \geq t)}(X \geq t + h) = p(X \geq h)$.

ROC Démonstration :

$P_{(X \geq t)}(X \geq t+h) = \dfrac{P((X \geq t) \cap (X \geq t+h))}{P(X \geq t)} = \dfrac{P(X \geq t+h)}{P(X \geq t)}$

$P_{(X \geq t)}(X \geq t+h) = \dfrac{e^{-\lambda(t+h)}}{e^{-\lambda t}} = e^{-\lambda t - \lambda h + \lambda t} = e^{-\lambda h}$

$P_{(X \geq t)}(X \geq t+h) = P(X \geq h)$

Annexe J

Compléments

J.1 Questionnaire

Lycée : **Juin 2015**
Classe :
Nom/Prénom : Durée : 1 h

Ce questionnaire n'est pas une évaluation. Il sera analysé dans le cadre d'une recherche. Vous pouvez donc laisser (sans crainte) toute trace de réflexion (même si vous n'êtes pas sûr, même si vous pensez ne pas bien justifier,...). Ne perdez pas trop de temps sur une question, écrivez seulement vos idées si c'est trop long et passez à la suite.

Merci pour votre participation.

Pour les trois premières questions :
Soit X une variable aléatoire continue à valeurs dans un intervalle I de $\mathbb{R}$. On note f la fonction de densité de probabilité associée à X.

Question 1 Cette fonction de densité de probabilité f vérifie-t-elle des conditions particulières ? Si oui, lesquelles ?

Si la question 1 vous pose problème, levez la main pour être débloqué et cochez cette case ☐

Question 2 Pouvez-vous expliquer, avec vos mots, pourquoi la fonction de densité f doit vérifier les conditions que vous avez cité à la question 1 ?

Question 3 A quoi sert la fonction de densité f ? Pourquoi s'intéresse-t-on à l'aire sous la courbe représentative de cette fonction ?

Question 4 Dans chacun des cas suivants, dire si la fonction définit une fonction de densité de probabilité sur l'intervalle donné ou non. **Justifier chaque réponse.** Toute forme de justification est possible.

(a) Soit g la fonction définie sur $[0;1]$ par $g(x) = 4x - 1$.

☐ Oui ☐ Non

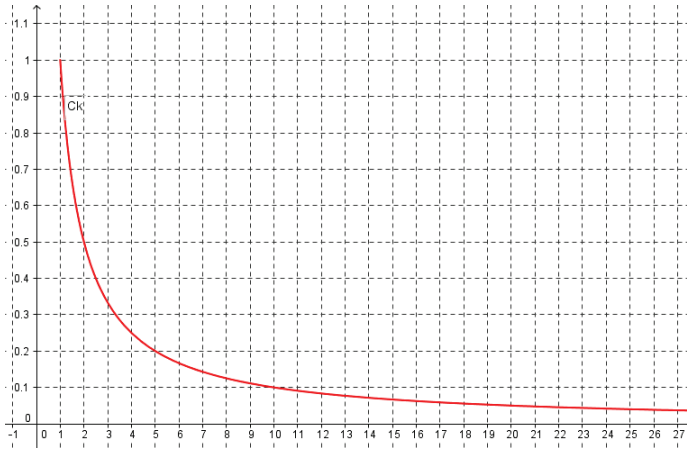
Justification :

(b) Soit h la fonction définie sur $[0;+\infty[$ par $h(x) = 3e^{-3x}$.

☐ Oui ☐ Non

Justification :

(c) Soit k la fonction définie sur $[1;+\infty[$ par $k(x) = \dfrac{1}{x}$, dont la courbe C_k est représentée ci-dessous.



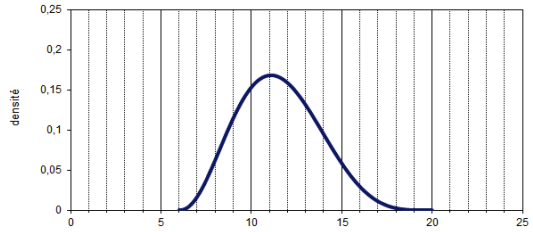
☐ Oui
☐ Non

Justification :

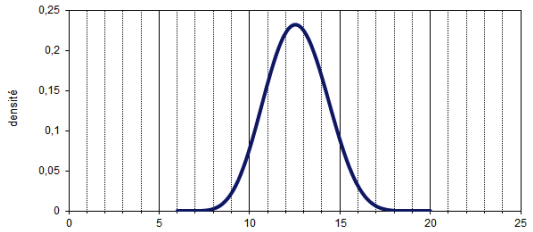
Question 5 On veut modéliser l'heure de départ de M. Martin, le matin. Il pense qu'il a de fortes chances de partir entre 8h08 et 8h15. En tout cas, il ne partira jamais avant 8h06 ni après 8h20.

Parmi les trois fonctions de densité proposées, laquelle choisiriez-vous ? Justifiez.

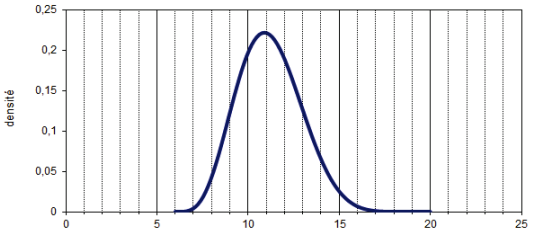
(a)



(b)



(c)



Justification :

Question 6 Peut-on trouver une valeur exacte ou approchée de l'intégrale suivante : $I = \displaystyle\int_{-0,5}^{1} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \,\mathrm{d}x$? Expliquez la (ou les) méthode(s) que vous utiliseriez. Vous pouvez proposer une solution qui utilise la calculatrice, mais, bien sûr, sans utiliser la touche de la calculatrice qui permet de calculer directement des intégrales !

Question 7 Soit f une fonction définie et continue sur $[a;b]$. Complétez l'expression suivante par la lettre manquante pour définir la fonction F :
$F(...) = \displaystyle\int_{a}^{u} f(t)\,\mathrm{d}t$

Question 8 Soit X une variable aléatoire continue à valeurs dans $[a;b]$ et f sa fonction de densité. Complétez l'expression suivante par les lettres manquantes.
$P(... \leqslant X \leqslant ...) = \displaystyle\int_{a}^{u} f(t)\,\mathrm{d}t$

Question 9 (a) Connaissez-vous une fonction continue sur un intervalle qui n'admet pas de primitives ? Si oui, donnez un exemple. Si non, justifiez.

(b) Connaissez-vous une fonction continue sur un intervalle qui admet une primitive dont vous ne connaissez pas explicitement une expression ? Si oui, donnez un exemple. Si non, justifiez.

Question 10 Soit X une variable aléatoire qui a pour fonction de densité f sur l'intervalle $[0;10]$. Soit a et b deux nombres réels appartenant à $[0;10]$ tels que $a < b$.

On pose, pour tout $x \in [0;10]$, $g(x) = P(X \leqslant x)$.

(a) Ecrire $g(x)$ à l'aide d'une intégrale.

(b) Quelle relation existe-t-il entre la fonction f et la fonction g ?

(c) **En déduire** l'égalité suivante : $P(a \leqslant X \leqslant b) = P(X \leqslant b) - P(X \leqslant a)$.

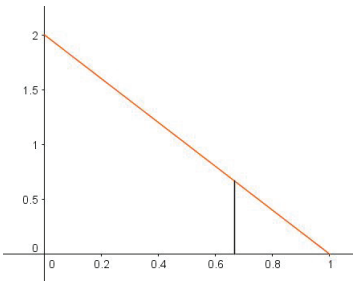
J.2 Analyse de tâches de l'Exercice 1

Il s'agit d'un exercice d'application qui fait partie intégrante du problème, il permet notamment de vérifier la compréhension de l'utilisation de la courbe de tendance dans la détermination de probabilités, qui vient d'être construite.

Nous allons présenter l'analyse de tâches des 4 premières questions, les deux suivantes reprennent les précédentes.

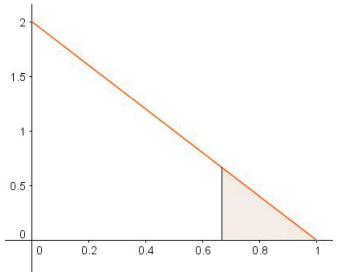
a) Quelle est la probabilité que le premier arrivé attende exactement 40 minutes ?

Tout d'abord, il faut passer du paradigme PaD-P1 au paradigme PaD-P2, en interprétant la question par le fait que l'on cherche $p(X = \frac{40}{60})$ ou encore $p(X = \frac{2}{3})$. Il faut pour cela que la conversion minutes/heures, connaissance ancienne, soit disponible. Il faut ensuite mettre en jeu la connaissance en construction, c'est-à-dire que cette probabilité est égale à une aire sous la courbe de densité. Comme dans tout l'exercice, la représentation graphique de la courbe est une aide indispensable, nous semble-t-il. Il s'agit de l'aire du segment ci-dessous, qui vaut 0 (connaissance ancienne, supposée disponible). Nous avons ici un changement de domaine. Cependant ici le registre graphique peut suffire pour répondre.



b) Quelle est la probabilité que le premier arrivé attende plus de 40 minutes ?

Il faut à nouveau interpréter la question dans le paradigme PaD-P2. On cherche donc $p(X \geqslant \frac{40}{60})$ ou encore $p(X \geqslant \frac{2}{3})$. X étant à valeurs dans $[0;1]$, on a alors $p(X \geqslant \frac{2}{3}) = p(\frac{2}{3} \leqslant X \leqslant 1)$. Il s'agit d'un changement d'écriture dans le registre symbolique probabiliste. Cette probabilité correspondant à l'aire sous la courbe ci-dessous, c'est-à-dire à l'aire du triangle de base $\frac{1}{3}$ et de hauteur $f(\frac{2}{3})$. On passe ici au domaine du calcul intégral, dans le niveau A1. Il faut alors déterminer $f(\frac{2}{3})$ en utilisant l'expression de f et ensuite utilisée la formule de l'aire d'un triangle rectangle, connaissances anciennes supposées disponibles. Après calculs, on obtient $\frac{1}{9}$.

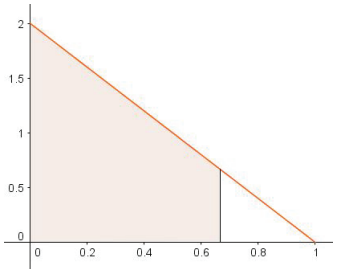


c) Quelle est la probabilité que le premier arrivé attende moins de 40 minutes ?

Il faut à nouveau interpréter la question dans le paradigme PaD-P2. On cherche donc $p(X \leqslant \frac{40}{60})$ ou encore $p(X \leqslant \frac{2}{3})$. X étant à valeurs dans $[0;1]$, on a alors $p(X \leqslant \frac{2}{3}) = p(0 \leqslant X \leqslant \frac{2}{3})$. Il s'agit d'un changement d'écriture dans le registre symbolique probabiliste. Ensuite deux méthodes sont possibles : soit la probabilité est interprétée comme l'aire du trapèze ci-dessous, soit on utilise la propriété $p(X \leqslant a) = 1 - p(X > a)$. Nous ne détaillons ici que la seconde démarche car la première sera présentée dans l'analyse de la question suivante. La propriété $p(X \leqslant a) = 1 - p(X > a)$ est une connaissance ancienne dans le cas des probabilités discrètes finies peut se retrouver rapidement pour le cas continu, en faisant le lien entre les probabilités et les aires sous la courbe et le fait que l'aire totale sous la courbe vaut 1. Pour cette question, on peut écrire :

$$\begin{aligned} p\left(X \leqslant \frac{2}{3}\right) &= 1 - p\left(X > \frac{2}{3}\right) \\ &= 1 - \left(p\left(X \geqslant \frac{2}{3}\right) - p\left(X = \frac{2}{3}\right)\right) \\ &= 1 - \left(\frac{1}{9} - 0\right) \\ &= \frac{8}{9}. \end{aligned}$$

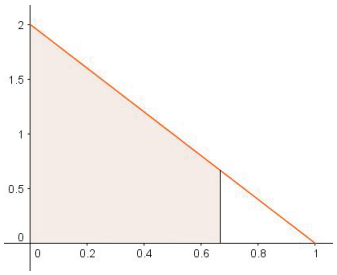
Donc en utilisant les questions précédentes, $p(X \leqslant \frac{2}{3}) = \frac{8}{9}$.



d) Quelle est la probabilité que le premier arrivé attende entre 20 et 40 minutes ?

Il faut à nouveau interpréter la question dans le paradigme PaD-P2. On cherche donc $p(\frac{20}{60} \leqslant X \leqslant \frac{40}{60})$ ou encore $p(\frac{1}{3} \leqslant X \leqslant \frac{2}{3})$. Il s'agit par un changement

de registre dans les probabilités à densité (passage du symbolique probabiliste au graphique), accompagné ensuite d'un changement de sous-domaine (des probabilités à densité au calcul intégral, niveau A1), d'interpréter la probabilité comme l'aire du trapèze ci-dessous, avec pour bases $f(\frac{1}{3})$ et $f(\frac{2}{3})$ et hauteur $\frac{1}{3}$.
On obtient alors $p(\frac{1}{3} \leqslant X \leqslant \frac{2}{3}) = \frac{1}{3}$, après calculs. La formule de l'aire d'un trapèze est un connaissance ancienne supposée disponible.



La représentation graphique joue clairement un rôle important. Il permet notamment le changement de domaine entre les probabilités et l'analyse. Tous les calculs d'aire sont de niveau A1, ce qui explique la faisabilité de ces calculs bien que le calcul intégral n'ait encore été étudié en classe. A chaque question, il y a un changement de paradigme de P1 à P2, par notamment un changement de registre (passage du langage naturel au registre symbolique probabiliste) puis un second changement de registre en passant du registre symbolique probabiliste au registre graphique, qui par un changement de domaine permet d'aboutir. On pourrait dire que l'on passe des probabilités au domaine de la géométrie mais comme nous l'avons précisé dans le début de cette thèse, nous considérons le calcul d'aire élémentaire dans le sous-domaine du calcul intégral, ici dans le niveau A1 – sans pour autant renier qu'il fasse bien entendu partie de la géométrie.