



HAL
open science

Modèles praxéologiques dans la transition secondaire-supérieur : le cas des probabilités en filière biologie

Camille Doukhan

► **To cite this version:**

Camille Doukhan. Modèles praxéologiques dans la transition secondaire-supérieur : le cas des probabilités en filière biologie. Education. Université de Bretagne occidentale - Brest, 2021. Français. ⟨NNT : 2021BRES0094⟩. ⟨tel-03632311⟩

HAL Id: tel-03632311

<https://theses.hal.science/tel-03632311v1>

Submitted on 6 Apr 2022

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.



HAL Authorization

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

ÉCOLE DOCTORALE N° 603
Éducation, Langages, Interactions, Cognition, Clinique
Spécialité : *Didactique des mathématiques*

Par

Camille DOUKHAN

**Modèles praxéologiques dans la transition secondaire-supérieur :
le cas des probabilités en filière biologie**

TOME 1 : TEXTE

Thèse présentée et soutenue à Rennes, le 9 décembre 2021

Unité de recherche : Centre de recherche sur l'éducation, les apprentissages et la didactique

Rapporteurs avant soutenance :

Alejandro GONZALEZ-MARTIN Professeur, Université de Montréal
Corinne HAHN Professeure (HDR), ESCP Europe

Composition du Jury :

Présidente :	Patricia MARZIN-JANVIER	Professeure (HDR), Université de Bretagne Occidentale
Examineurs :	Alejandro GONZALEZ-MARTIN	Professeur, Université de Montréal
	Corinne HAHN	Professeure (HDR), ESCP Europe
	Laurent VIVIER	Maître de conférences (HDR), Université de Paris

Dir. de thèse :	Ghislaine GUEUDET	Professeure (HDR), Université de Bretagne Occidentale
Co-dir. de thèse :	Fabrice VANDEBROUCK	Professeur (HDR), Université de Paris

Ce projet a bénéficié du soutien financier de la Région Bretagne

Remerciements

Dans l'objectif d'être concise je ne pourrais remercier individuellement toutes les personnes qui ont oeuvré de près ou de loin à la production de ce travail, pour cela je m'en excuse.

En premier lieu je souhaite remercier Ghislaine Gueudet et Fabrice Vandebrouck qui m'ont encadrée dans la préparation de cette thèse et fait découvrir le monde de la recherche en didactique. Merci à Ghislaine qui me suit depuis le master, a encadré mon mémoire et m'a lancée dans la thèse en m'accordant sa confiance. Je la remercie pour sa réactivité sans faille, sa disponibilité et ses encouragements. Merci à Fabrice qui a accepté de participer à ce projet et de me suivre durant ces trois années, je le remercie pour ses conseils et ses remarques toujours avisées. Merci à eux deux de m'avoir guidée dans cette passionnante traversée que fut la préparation de cette thèse.

Je tiens à remercier Alejandro González Martín et Corinne Hahn d'avoir accepté de rapporter cette thèse. Je remercie également Patricia Marzin-Janvier et Laurent Vivier d'avoir accepté de faire partie de mon jury. A travers leurs rapports et les riches échanges que nous avons eus, j'espère que cette recherche se poursuivra et s'étoffera.

Ce travail n'aurait pas été possible sans le soutien financier de la région Bretagne qui m'a permis de me consacrer entièrement et sereinement à l'élaboration de ma thèse. Je remercie à nouveau Ghislaine Gueudet pour avoir monté le dossier qui a su me permettre d'obtenir ce financement.

Je remercie chaleureusement les membres du CREAD qui m'ont accueillie au cours de ces trois années. En particulier merci à Jean-Noël Blocher, Jean-Marie Boilevin et Sophie Joffredo-Lebrun qui m'ont conseillée ou accompagnée à un moment ou à un autre dans ce travail. J'ai eu la chance de partager mon bureau à l'INSPE avec de supers doctorant(e)s (certaines devenues docteurs), un grand merci à Nolwenn, Valérie, Muriel et Antoine pour m'avoir notamment aiguillée sur le fonctionnement du doctorat.

Les financements octroyés par le laboratoire du CREAD et l'école doctorale ELICC m'ont permis de faire connaître mon travail et de participer à des rencontres scientifiques riches et captivantes. Au cours de celles-ci j'ai eu le plaisir d'échanger avec des chercheuses et chercheurs que je remercie pour leurs remarques, questions et autres encouragements. Merci au GDR DEMIPS pour l'accueil qu'il réserve aux jeunes chercheuses et chercheurs et l'opportunité enrichissante des rencontres scientifiques qu'il propose. Merci particulièrement à Hussein Sabra pour les échanges que nous avons pu avoir à propos de ce travail et pour ses conseils. Ces événements m'ont également permis de rencontrer les jeunes

chercheuses et chercheurs de l'ARDM et ainsi de partager nos expériences autour de nos recherches respectives. Certains d'entre eux sont devenus des ami(e)s et j'adresse ici un clin d'oeil particulier à Elann.

Cette recherche étant expérimentale et pour tout ce que cela a pu apporter à ce travail, je souhaite remercier les enseignantes et enseignants qui m'ont accueillie dans leur classe et les élèves, étudiantes et étudiants qui ont accepté de participer à cette étude. Je remercie également Marie-Pierre Lebaud qui m'a initiée à BRAISE et m'a ainsi permis d'y apporter ma contribution.

Si j'ai pu terminer la rédaction du manuscrit et préparer cette soutenance sereinement c'est en grande partie grâce à l'équipe de Strasbourg qui m'a si bien accueillie et a su adapter mon emploi du temps pour me libérer du temps les derniers mois ; pour cela un grand merci à toutes et tous et particulièrement à Charlotte Derouet.

Ce travail de thèse clôture plus de vingt années de scolarité, primaire, secondaire et supérieur, au cours desquelles j'ai eu le plaisir de connaître des enseignant(e)s bienveillant(e)s et inspirant(e)s. Sans aucun doute ces rencontres m'ont amenée jusqu'ici, je leur adresse toute ma reconnaissance et un merci tout particulier à Mme Jestin, Mr Bastard et Mme Adam.

Pour le soutien qu'elles et ils m'ont apporté toutes ces années, pour les vacances passées ensemble et tous les autres bons moments partagés un immense merci à mes fantastiques ami(e)s brestois(e)s, rennais(e)s et à celles et ceux parti(e)s trouver le soleil. Merci tout spécialement à Julie et Martin d'être présents à mes côtés depuis tant d'années.

Enfin, j'adresse ces derniers remerciements à ma famille et tout particulièrement à Juliette et à mes parents. Merci pour vos patientes relectures, vos attentions et votre confiance. Merci pour vos encouragements qui m'ont portée durant ces trois années.

Table des matières

1	Revue de littérature	13
1.1	A propos des difficultés rencontrées par les étudiants à la transition secondaire-supérieur	13
1.2	A propos de la modélisation en probabilités	17
1.3	A propos de la spécificité des étudiants non-spécialistes	21
1.4	A propos des particularités liées à l'enseignement et à l'apprentis- sage des probabilités et notamment des probabilités conditionnelles	25
1.5	A propos de l'enseignement et de l'apprentissage des variables aléa- toires continues	30
2	Cadre théorique et questions de recherche	35
2.1	Registres de représentations sémiotiques	36
2.2	Éléments issus de la Théorie Anthropologique du Didactique (TAD)	37
2.3	Éléments de la Théorie de l'Activité (TA)	38
2.4	Modèle Praxéologique de Référence Étendu et adapté à la Transi- tion secondaire-supérieur : MPR-ET	40
2.5	Questions problématisées	42
3	Méthodologie générale de la recherche	45
3.1	Présentation des terrains	45
3.2	Outils pour l'analyse de la transition institutionnelle	46
3.3	Outils d'analyse de l'activité et des productions des élèves et des étudiants	46
4	Enseignement des probabilités au secondaire : analyse du savoir à en- seigner	49
4.1	Les différentes approches des probabilités et leur enseignement . .	49
4.1.1	Méthodologie de l'étude historico-épistémologique	50
4.1.2	Éclairages historiques	51

4.1.3	Pluralité des approches dans l'enseignement des probabilités	54
4.2	Programmes officiels	56
4.2.1	Méthodologie de l'analyse des textes officiels	56
4.2.2	Éléments généraux issus des textes officiels concernant l'en- seignement des probabilités	58
4.2.3	Conditionnement et indépendance	59
4.2.4	Lois à densité	60
4.2.5	Modélisation dans les textes officiels	62
4.3	Analyse d'un manuel et de sujets de baccalauréat	63
4.3.1	Méthodologie de l'étude de manuel et de sujets de baccalau- réat	64
4.3.2	Probabilités conditionnelles	65
4.3.3	Lois de probabilités à densité	85
5	Enseignement et apprentissage des probabilités à la transition secondaire-supérieur	105
5.1	Observation de séances de classe : recueil de données et outils d'analyse	105
5.2	Enseignement des probabilités au secondaire : analyse du savoir enseigné	108
5.2.1	Analyse du savoir enseigné pour le thème des probabilités conditionnelles	109
5.2.2	Analyse du savoir enseigné pour le thème des lois continues	130
5.2.3	Le cas particulier de la modélisation	190
5.2.4	Éléments caractéristiques indépendants des contenus	193
5.3	Enseignement des probabilités en première année d'Université : le cas des étudiants de biologie	194
5.3.1	Analyse du savoir enseigné pour le thème des probabilités conditionnelles	195
5.3.2	Analyse du savoir enseigné pour le thème des lois continues	221
5.3.3	Éléments caractéristiques en termes de déroulement	252
6	Probabilités à la transition secondaire-supérieur : analyses relatives aux élèves et étudiants	255
6.1	Test sur les probabilités conditionnelles	255
6.1.1	Présentation du test	257
6.1.2	Analyse a priori	257

6.1.3	Méthodologie de l'analyse des productions d'élèves et d'étudiants	264
6.1.4	Analyse a posteriori	265
6.2	Questionnaire en ligne à destination des étudiants de biologie . . .	275
6.2.1	Présentation du questionnaire	275
6.2.2	Analyse des réponses au questionnaire	277
6.3	Entretiens avec des étudiants de première année de biologie	283
6.3.1	Présentation de la grille d'entretien	284
6.3.2	Analyse des entretiens	287
6.4	Examen du module de probabilités en première année de biologie .	291
6.4.1	Méthodologie du recueil et de l'analyse des copies	293
6.4.2	Analyse a priori du sujet d'examen	294
6.4.3	Analyse de copies d'étudiants	297
7	Conclusion et perspectives	301
7.1	Réponses aux questions initiales	301
7.2	Apport théorique pour l'étude de la transition et proposition pour le cas des probabilités pour les non-spécialistes	307
7.2.1	Méthodologie de construction du MPR-ET	308
7.2.2	MPR-ET pour les probabilités conditionnelles et les lois continues, le cas des étudiants de biologie	310
7.3	Limites de cette étude	314
7.4	Perspectives	315
7.4.1	Travaux complémentaires en cours	315
7.4.2	Travaux envisagés pour la suite	317
	Références	319

Introduction

Les étudiants qui arrivent à l'Université pour suivre un cursus de biologie ne s'attendent pas toujours à trouver un enseignement de mathématiques aussi conséquent que celui qui est dispensé en première année d'Université. Des auteurs comme Heublein (2014) et Larsen, Kornbeck, Kristensen, Larsen, et Sommersel (2013) ont d'ailleurs constaté que les difficultés que ces étudiants non-spécialistes rencontrent en mathématiques sont une des causes d'abandon de leurs études. En outre, certains étudiants de biologie choisissent cette filière pour s'éloigner d'une matière qu'ils n'apprécient pas, ou pour laquelle ils rencontraient d'ores et déjà des difficultés dans le secondaire. Il s'agit d'un constat que nous avons pu faire en enseignant nous-même auprès de ces étudiants. Aussi, et suite à des discussions avec d'autres membres de l'équipe enseignante de mathématiques de l'Université qui observaient également de nombreuses difficultés chez ces étudiants de biologie, il nous sembla pertinent de nous intéresser, dans cette thèse, aux questions de transition secondaire-supérieur concernant les étudiants non-spécialistes au début de l'Université.

Après une comparaison des contenus de mathématiques entre la classe de terminale, série scientifique, et les modules de mathématiques enseignés aux étudiants de première année de biologie, nous avons choisi de nous intéresser aux probabilités, et plus précisément aux probabilités conditionnelles et aux lois continues, thèmes communs aux deux curricula.

L'objectif de cette thèse est, premièrement, d'analyser finement les attendus à l'Université pour ces deux thèmes de probabilités et de les comparer aux contenus enseignés au lycée. Deuxièmement, de faire l'état des lieux des difficultés rencontrées par ces étudiants, pour ces enseignements de probabilités, dans le but d'y remédier.

Plan de l'étude

Cette thèse est composée de sept chapitres qui s'articulent comme suit.

Afin de situer notre travail au sein de la recherche en didactique pour l'enseignement supérieur, nous présentons, dans un premier chapitre, une revue de littérature qui aborde les différents angles de cette étude : la transition secondaire-supérieur, la modélisation en probabilités, la spécificité des étudiants

non-spécialistes, le cas particulier des probabilités conditionnelles, et enfin celui des lois continues.

Le deuxième chapitre expose le cadre théorique que nous avons choisi pour répondre aux problématiques de cette thèse. Une première partie de ce chapitre est dédiée aux registres de représentations sémiotiques de Duval (1993). Dans une deuxième partie nous présentons des éléments de la Théorie Anthropologique du Didactique (Chevallard, 2003) qui nous permettront de mener des analyses du point de vue institutionnel. Afin d'étudier avec attention l'activité de l'élève et de l'étudiant, dans une troisième partie, nous complétons cette théorie avec des notions de la Théorie de l'Activité adaptées à la didactique des mathématiques (Vandebrouck, 2018). Dans une quatrième partie, nous exposons un outil théorique que nous avons construit à partir du concept de Modèle Praxéologique de Référence issu de la Théorie Anthropologique du Didactique. Cet outil, appelé MPR Étendu par la théorie de l'activité adapté pour la Transition (noté MPR-ET), nous permettra d'explicitier de manière synthétique les attentes institutionnelles envers les étudiants à la fin du cours de probabilités de première année de biologie en croisant le point de vue de l'institution avec celui des étudiants. Dans une dernière partie nous énonçons nos questions de recherche au regard du cadre théorique précédemment développé.

Le troisième chapitre est bref, il présente la méthodologie générale retenue pour cette étude. Nous reviendrons dans chaque chapitre sur sa méthodologie spécifique.

Dans la perspective de concevoir un MPR-ET, nous proposons dans un quatrième chapitre, un éclairage historique et épistémologique des courants probabilistes et de leurs impacts sur l'enseignement des probabilités. Le second objectif de ce quatrième chapitre est d'analyser le savoir à enseigner en classe de terminale scientifique au lycée, pour les deux thèmes de probabilités retenus.

Le cinquième chapitre est consacré à l'analyse du savoir enseigné en classe de terminale scientifique et en première année d'Université en filière de biologie. Pour ce faire, nous prenons appui sur des documents recueillis en classe et sur de nombreuses séances observées, que nous analysons au regard des résultats du quatrième chapitre. Ce chapitre est l'occasion de comparer les institutions en termes d'attendus, de déroulements des séances, de praxéologies enseignées, etc.

Afin de compléter ces résultats, et dans l'objectif de construire un MPR-ET étendu adapté aux attentes et aux besoins des étudiants biologistes débutants à l'Université, nous nous intéressons, dans un sixième chapitre, au point de vue des étudiants et à leurs apprentissages. Dans une première partie nous présentons et analysons un test que nous avons soumis à des élèves et à des étudiants. Nous recueillons ensuite le point de vue des étudiants sur l'enseignement qu'ils reçoivent à l'Université pour ces contenus de probabilités, au moyen d'un questionnaire en ligne que nous approfondissons par des entretiens. Dans une dernière partie

nous analysons des copies d'examens de ces mêmes étudiants avant de tirer des conclusions en vue du dernier chapitre.

L'analyse de l'existant, du point de vue des institutions, de l'activité de l'élève et de l'étudiant, nous permet de présenter dans un septième chapitre la construction de l'outil MPR-ET. Ce dernier chapitre fait le bilan général du travail exposé tout au long des six précédents chapitres. Nous présentons finalement des perspectives pour de futures recherches.

Chapitre 1

Revue de littérature

L'étude des probabilités à la transition secondaire-supérieur est un sujet qui a été assez peu abordé dans la recherche en didactique des mathématiques. Nous avons fait le choix de spécifier ce travail aux étudiants d'Université de filières utilisatrices de mathématiques, autrement dit les étudiants non-spécialistes. On trouve dans la littérature de nombreux travaux qui se sont intéressés aux questionnements relatifs à la transition secondaire-supérieur, à la modélisation, aux probabilités ou aux étudiants non-spécialistes mais il n'existe, à notre connaissance, pas de travaux mêlant ces différents aspects.

Pour commencer nous présentons des travaux de recherche portant sur les difficultés rencontrées par les étudiants lors de la transition secondaire-supérieur de façon générale. Nous présenterons ensuite des travaux concernant la modélisation en probabilités, puis des travaux centrés sur les étudiants non-spécialistes et enfin, plus précisément, des travaux sur l'enseignement et l'apprentissage des probabilités conditionnelles.

Il y a certains travaux dont nous n'allons pas parler dans cette partie car ils ont davantage leur place comme soutien à la construction du modèle praxéologique de référence. Ces travaux seront plus détaillés dans la partie 4.1, page 49.

1.1 A propos des difficultés rencontrées par les étudiants à la transition secondaire-supérieur

La transition secondaire-supérieur est un thème largement étudié par la recherche en didactique des mathématiques, en France, comme à l'international, depuis plusieurs années maintenant. De nombreux auteurs se sont intéressés aux difficultés en mathématiques, rencontrées par les étudiants à l'entrée dans le supérieur (Gueudet & Thomas, 2019; Artigue, 2004; Robert, 1998), notamment en analyse (Vivier & Durand-Guerrier, 2016) et en algèbre linéaire (Dorier, 2000a), et des synthèses de ces travaux ont déjà été réalisées (Gueudet, 2008; Gueudet & Vandebrouck, 2019; Biza, Giraldo, Hochmuth, Khakbaz, & Rasmussen, 2016). Nous n'entendons pas ici lister exhaustivement ces résultats, mais à partir de ces

références nous présentons certains éléments concernant les changements lors de la transition secondaire-supérieur et les difficultés des étudiants qui peuvent en résulter, susceptibles d'éclairer nos analyses futures.

Comme l'a souligné Gueudet (2008) l'institution Université est un nouveau monde, avec sa propre langue et ses nouvelles règles. Dans ce nouveau monde, les étudiants rencontrent de nombreux obstacles et difficultés dans l'apprentissage des mathématiques. Comme l'ont décrit Gueudet et Thomas (2019), les difficultés rencontrées par les étudiants lors du passage de l'enseignement secondaire à l'enseignement supérieur peuvent avoir des causes assez générales (changement de ville, quitter ses parents, etc.) et des causes plus spécifiques aux contenus mathématiques. C'est le second point qui nous intéresse ici.

Des difficultés liées aux contenus

Une première source de difficultés identifiée dans de nombreux travaux est celle des contenus. En effet, d'après certains auteurs les difficultés rencontrées par les étudiants proviennent des contenus qui sont plus difficiles et des notions plus abstraites à l'Université qu'au lycée. Les connaissances seraient construites, dans le supérieur, suivant une "*pensée mathématique avancée*" (concept introduit par Tall (1991)). Cette pensée n'a pas été assez développée chez les élèves dans le secondaire et serait une cause des difficultés qu'ils rencontrent.

Dans le cas qui nous intéresse, des probabilités pour les non-spécialistes, il semblerait que les contenus ne deviennent pas particulièrement plus abstraits ni avancés à l'Université. Il convient donc de s'intéresser à d'autres facteurs de difficultés chez les étudiants débutants à l'Université.

Une transition entre deux cultures différentes aux fonctionnements bien distincts

Comme le souligne Artigue (2004), les difficultés que présente le passage du lycée au supérieur sont notamment dues à la transition entre deux cultures mathématiques différentes. Le passage du lycée à l'Université entraîne de nombreux changements. Ces changements sont d'abord d'ordre structurels. A l'Université le groupe est d'une taille plus importante, les enseignants de mathématiques sont différents d'un cours à l'autre, parfois l'enseignant qui fait le cours magistral n'est pas le même que celui qui encadre la séance de travaux dirigés.

Les attentes de l'institution Université sont différentes de celles du lycée, par exemple, en ce qui concerne l'autonomie attendue chez les étudiants. Les attentes sont aussi différentes en ce qui concerne la mise en fonctionnement des connaissances. Si au lycée les exercices mettent en jeu des connaissances qui sont explicitées et qui doivent être immédiatement appliquées, à l'Université en revanche, les connaissances doivent être plus disponibles et sont mises en

fonctionnement de façon plus complexe (Gueudet & Vandebrouck, 2019).

Les niveaux de mises en fonctionnement (mobilisables, disponibles), décrits par Robert (1998), permettent d'analyser les types de tâches selon l'activité de l'élève et notamment d'étudier la complexité de la tâche. Les outils développés en théorie de l'activité à la suite du travail de Robert (1998) sont complémentaires d'une approche anthropologique exclusive pour l'analyse que nous présentons dans la suite (voir cadre théorique présenté partie 2, page 35).

Des pratiques d'enseignement différentes

Dans le secondaire, les connaissances mathématiques se construisent assez souvent par imitation d'une procédure et répétition d'un certain nombre d'exercices types (Artigue, 2004). L'enseignant du lycée a pour habitude de faire appel à la mémoire du groupe classe pour mobiliser certaines connaissances déjà étudiées ou débloquer une situation. Ces moyens d'actions disparaissent à l'Université. En effet, dans l'institution Université, le groupe classe est différent (beaucoup plus grand ou scindé en petits groupes pour les séances de travaux dirigés) et l'enseignant qui intervient n'est pas forcément le même que celui qui a donné le cours. A l'Université la diversité des exercices est plus grande, les notions défilent plus rapidement et les techniques (voir définition partie 2.2, page 37) n'ont plus le temps d'être routinisées à force de répétitions (Gueudet, 2008). On parle d'accélération du temps didactique.

Ainsi les pratiques d'enseignement dans l'institution Université sont différentes de celles présentes dans l'institution lycée. Ces nouvelles pratiques entraînent une complexité supplémentaire lors de la transition pour les étudiants.

La transition du point de vue de l'approche anthropologique

Si l'on se situe dans une approche anthropologique (voir TAD partie 2.2, page 37), Gueudet et Thomas (2019) soulignent que l'enseignement dans le secondaire se concentre principalement sur le bloc pratico-technique. Au sein de l'institution lycée, l'activité mathématique est rarement théorique. L'enseignement des mathématiques se focalise sur les aspects pratiques de cette activité. Il est habituel que pour un type de tâches donné, une technique unique soit proposée (Gueudet, 2008).

Dans le supérieur l'enseignement est davantage focalisé sur l'organisation théorique des contenus mathématiques, la présentation de preuves et de théorèmes. L'enseignement des mathématiques à l'Université est plus axé sur le bloc technologico-théorique que l'enseignement au lycée (Gueudet & Thomas, 2019). Ceci est d'autant plus problématique qu'à l'Université les enseignants s'attendent à ce que les étudiants maîtrisent pleinement certains contenus dont ils n'ont en

fait appris que les aspects techniques au lycée (Gueudet & Thomas, 2019).

Dans l'étude qui suit nous comparons les praxéologies du secondaire et du supérieur pour le cas qui nous intéresse, des probabilités enseignées aux non-spécialistes, ces analyses sont à retrouver chapitre 4 et chapitre 5.

De nouvelles attentes liées aux pratiques expertes des mathématiciens

Comme nous l'avons déjà décrit précédemment, les difficultés que rencontrent les étudiants à leur entrée à l'Université, proviennent notamment des nouvelles attentes de l'institution Université. Robert (1998) explique que ces attentes découlent également des pratiques "expertes" du mathématicien. En effet, la position du mathématicien dans l'institution Université lui confère sa place d'expert.

Ces pratiques expertes sont considérées comme nécessaires au développement des modes de pensée et de raisonnement mathématiques attendus dans le supérieur : connaissances réparties dans plusieurs domaines, autonomie des choix de raisonnement, disponibilités et organisation des connaissances.

A propos des outils numériques et des dispositifs de remédiation

De nombreux outils technologiques à usages pédagogiques existent pour l'Université, qu'il s'agisse de logiciels spécifiques, on peut citer WIMS, ou d'environnements numériques de travail. La crise sanitaire a aussi mis en avant l'usage des dispositifs de communication distante. Cependant Gueudet (2017) observe que les outils numériques sont très peu présents dans les ressources utilisées par les enseignants et proposées aux étudiants à l'Université. L'auteure explique pourtant que parmi ces outils numériques, on peut trouver des solutions de remédiation à la complexité que représente la transition secondaire-supérieur. Lors de nos observations d'enseignement nous serons attentifs à l'utilisation qui est faite des outils numériques par les enseignants et par les étudiants.

Pour résumer

Le passage de l'enseignement secondaire à l'enseignement supérieur est un changement de cultures et d'institutions dans lesquelles les pratiques d'enseignements sont différentes. Les difficultés rencontrées par les étudiants sont les conséquences, entre autre, de nouvelles pratiques attendues. Pour des mêmes contenus, les mises en fonctionnement des connaissances sont différentes dans la culture Université. Ces nouvelles attentes sont parfois liées aux pratiques expertes des mathématiciens.

Dans la partie suivante, nous allons désormais nous intéresser aux travaux de recherche en didactique des mathématiques qui concernent la modélisation.

1.2 A propos de la modélisation en probabilités

La littérature en recherche en didactique des mathématiques regorge d'études sur la modélisation en mathématiques. Nous présentons dans cette partie des résultats issus de travaux qui portent sur la modélisation, de façon générale, et des résultats qui portent plus particulièrement sur les probabilités. Nous n'exposerons ici que les éléments qui nous permettront d'enrichir, par la suite, notre recherche et nos analyses.

L'une des spécificités du raisonnement probabiliste est le lien très fort qui l'unit aux phénomènes réels. La modélisation de phénomènes réels et aléatoires est un processus de construction de modèles.

Processus de modélisation, modèle pseudo-concret et modèles mathématiques

Avant toute chose il nous semble nécessaire de présenter les définitions des termes "*modélisation*" et "*modèle*" que l'on trouve dans les travaux de recherche en didactique des mathématiques et que nous retenons pour la suite de ce travail.

Selon Soury-Lavergne et Bessot (2012) la **modélisation** est "*un processus de résolution de problème qui consiste à construire un modèle à partir d'une situation réelle*".

Henry (2010) définit la notion de **modèle** comme "*une représentation simplifiée et idéalisée de la réalité*".

Les représentations (langage mathématique et symbolisme) qui permettent des descriptions sur lesquelles des propriétés générales (ou algorithmiques) peuvent fonctionner sont appelées par Chaput, Girard, et Henry (2011) **modèles mathématiques**. Ces modèles permettent de décrire une situation et de faire fonctionner des propriétés mathématiques générales.

De nombreux chercheurs se sont appliqués à décrire les étapes du processus de modélisation. Chevallard (1989) propose une description en trois étapes de ce processus, en articulant un système -mathématique ou non- et un modèle mathématique. Blomhøj et Jensen (2003) détaillent le processus de modélisation en sous-compétences. Cette description correspond mieux à ce que nous retiendrons, pour notre étude, des étapes de modélisation :

- (1) formulation de la tâche
- (2) production d'une représentation mathématique (modèle)

- (3) traduction des objets et des relations en mathématiques
- (4) utilisation de méthodes mathématiques pour obtenir un résultat mathématique
- (5) interprétation des résultats
- (6) validation du modèle

Certains auteurs spécifient cette description aux contenus de probabilités qui nous intéressent. Selon Henry (2005b) les étapes du processus de modélisation dans le cadre des probabilités, sont les suivantes :

- (1) observation, description, puis simplification afin d'obtenir un **modèle pseudo-concret**
- (2) **mathématisation** : formalisation du modèle pseudo-concret et représentation de ce modèle dans un système symbolique adapté au calcul probabiliste
- (3) traduction des résultats mathématiques en termes de modèle pseudo-concret et **validation** du modèle

Le *modèle pseudo-concret* mentionné dans la description ci-dessus, est défini par Henry (2005b) comme :

(...) une représentation simplifiée et idéalisée de la réalité qui peut être présentée dans un vocabulaire courant renvoyant à des objets réels. Ces objets possèdent, dans le modèle, des caractéristiques idéales et abstraites.

L'enseignement de la modélisation dans le cadre des probabilités

De nombreux chercheurs s'accordent à dire que les probabilités devraient être enseignées comme un moyen de modéliser les phénomènes réels plutôt que comme une simple mesure abstraite. Les chercheurs Bakker, Hahn, Kazak, et Pratt (2018) précisent qu'il est essentiel, pour l'enseignement des probabilités, de combiner une perspective centrée sur les données avec une perspective de modélisation. Burrill et Biehler (2011) expliquent que les étudiants auxquels on enseigne les probabilités avec une approche trop formelle, apprennent les formalismes sans comprendre les phénomènes qui sont décrits par ces notions mathématiques. Selon ces auteurs, la probabilité ne doit pas être enseignée sans données. Batanero, Henry, et Parzysz (2005) ajoutent que la modélisation de situations concrètes est une étape obligatoire dans le raisonnement scientifique.

L'enseignement des probabilités combine plusieurs approches, l'approche classique (ou laplacienne), l'approche fréquentiste et l'approche subjectiviste (ou bayésienne) (plus de détails dans la partie 4.1 page 49). Ce mélange peut causer des difficultés dans l'apprentissage des élèves. Chaput et al. (2011) préconisent d'opter pour la perspective de la modélisation dans l'enseignement des probabilités. Celle-ci permet de mieux agencer l'approche classique et l'approche fréquentiste.

Bien qu'il n'y ait pas de règle fixe pour modéliser un problème, suivant l'approche retenue, les situations probabilistes peuvent être interprétées de façons différentes et les modèles qui en résultent seront eux aussi différents (Eichler & Vogel, 2014). Girard (2005) constate que la confusion entre modèle et réalité, dans l'enseignement des probabilités, est à l'origine de nombreuses difficultés chez les étudiants.

Dans le document d'accompagnement des programmes des années 2000, la place attribuée à la modélisation dans l'enseignement des probabilités est très claire :

Le programme de probabilités des classes de première et terminale S concerne la modélisation d'expériences de référence (Ministère de l'Éducation Nationale de la Jeunesse et de la Vie Associative (MENJVA) & DGESCO, 2001).

Dans ce document, il y a une définition de l'activité de modélisation pour les thèmes de probabilités qui nous intéresse :

Modéliser une expérience aléatoire, c'est lui associer une loi de probabilités. L'objectif est que les élèves comprennent à l'aide d'exemples que **modéliser**, c'est ici choisir une loi de probabilités (Ministère de l'Éducation Nationale de la Jeunesse et de la Vie Associative (MENJVA) & DGESCO, 2001).

Dans le programme de mathématiques de la classe de première et de terminale scientifique des années 2010/2011, l'enseignement par la modélisation est mentionné pour les thèmes de probabilités au programme. On peut lire dans le document d'accompagnement (Ministère de l'Éducation Nationale de la Jeunesse et de la Vie Associative (MENJVA) & DGESCO, 2013) qui liste et détaille les compétences attendues en mathématiques au lycée :

Modéliser : traduire en langage mathématique une situation réelle (à l'aide d'équations, de suites, de fonctions, de configurations géométriques, de graphes, de lois de probabilités, d'outils statistiques ...). Utiliser, comprendre, élaborer une simulation numérique ou géométrique prenant appui sur la modélisation et utilisant un logiciel. Valider ou invalider un modèle.

Nous présenterons dans le chapitre 4, partie 4.2 (page 56), une analyse détaillée des programmes de terminale scientifique pour les thèmes de probabilités qui nous intéressent.

Selon Prodromou (2014), l'approche par la modélisation choisie dans les programmes renforce l'utilisation de modèles formalisés dans un système symbolique. Ces modèles, développés pour représenter des situations réelles, sont présentés aux élèves en langage symbolique. Cela implique de leur proposer uniquement des situations où ces modèles fonctionnent, ce qui pourrait, dans ce cas de figure, atténuer l'activité de modélisation chez les élèves.

Des difficultés liées au contexte

Dans le domaine des probabilités l'interprétation contextuelle est importante.

Les difficultés varient suivant le contexte des exercices (Huerta, 2014). Selon Martignon et Wassner (2002) les tâches qui manquent d'authenticité, c'est-à-dire qui sont présentées dans un contexte artificiel, sont une cause de difficulté supplémentaire. Selon ces auteurs, le contexte influe largement sur la motivation et l'intérêt de l'étudiant à s'intéresser aux problèmes.

Selon Chaput et al. (2011) il est nécessaire qu'une phase de modélisation soit présente dans les applications de probabilités. Ils expliquent que, bien que les exercices soient présentés dans un emballage concret, leurs contextes sont suffisamment proches de situations déjà rencontrées pour que cela permette aux étudiants de lier les tâches prescrites aux lois de probabilités qu'ils connaissent.

Après s'être intéressés aux tests soumis à l'entrée à l'Université en Andalousie, Batanero, Lopez-Martin, Arteaga, et Gea (2018) constatent qu'une part trop importante des problèmes de probabilités, proposés dans ces tests, ne sont pas liés à des situations de la vie réelle. Certains de ces problèmes demandent un raisonnement plus algébrique que probabiliste.

Les contextes des exercices sont, en général, assez artificiels, notamment dans les manuels scolaires. Henry (2005b) souligne à propos des programmes de 1991 :

Bien que la notion d'expérience aléatoire relève d'un habillage pseudo-concret dans les manuels, elle concerne une véritable observation du réel dans le programme.

Le programme de mathématiques de la classe de terminale scientifique (Ministère de l'Éducation Nationale (MEN), 2011) suggère l'enseignement des probabilités discrètes, pour les thèmes "*conditionnement et indépendance*", au travers de situations concrètes. On retrouve des suggestions analogues concernant les contenus de lois à densités :

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.

Eichler et Vogel (2014) ont établi une classification des situations problèmes : situations problématiques virtuelles / situations problématiques virtuelles du monde réel / situations problématiques du monde réel. Nous utiliserons cette classification dans la suite afin d'étudier les contextes des exercices (voir chapitre 4, analyse de manuels partie 4.3 page 63).

Nous terminons cette partie avec une recommandation de Chiel, McManus, et Shaw (2010). Selon ces auteurs il est nécessaire de repenser les mathématiques

en tant qu'outil de modélisation pour l'étude de problèmes extra-mathématiques. Il serait donc intéressant de proposer des cours où les étudiants participeraient à la construction de modèles mathématiques dans le cadre des probabilités.

Pour résumer

Il est important qu'une partie du processus de modélisation soit à la charge de l'élève. En effet, ceci peut entraîner un meilleur engagement, de la part de l'élève, dans la tâche. Il faut cependant chercher à ce que les élèves distinguent bien le modèle de la réalité.

Plusieurs modèles peuvent être présentés pour une même réalité. La modélisation doit faire partie de l'activité de l'élève et les contextes doivent être le moins artificiels possible pour permettre cette activité. Ceci est d'autant plus vrai en probabilités pour lesquelles la perspective de la modélisation est à intégrer pleinement dans leurs enseignements.

1.3 A propos de la spécificité des étudiants non-spécialistes

Nous nous intéressons dans cette partie à l'enseignement des mathématiques auprès d'étudiants non-spécialistes. Ces étudiants ont des besoins et des attentes spécifiques. Dans le cadre de notre étude il est nécessaire de prendre le temps de s'intéresser plus précisément à ce public et à ses particularités.

Nous avons vu partie 1.1, page 13, que les étudiants rencontrent des difficultés en entrant à l'Université et les étudiants non-spécialistes n'échappent pas à la règle. Ce qui les différencie des étudiants "spécialistes" c'est d'être dans des filières utilisatrices de mathématiques autres que celle de mathématiques (physique, économie, chimie, biologie, etc). Ces étudiants reçoivent un enseignement de mathématiques sans qu'ils aient choisi d'étudier cette matière. Et alors qu'ils sont entrés à l'Université pour étudier la physique, la biologie ou l'économie, Heublein (2014) et Larsen et al. (2013) expliquent que les difficultés qu'ils rencontrent en mathématiques sont une des causes d'abandon de leurs études.

De nombreux auteurs se sont intéressés aux difficultés propres à ces étudiants non-spécialistes.

Un enseignement de mathématiques déconnecté des attentes et des besoins des étudiants non-spécialistes

Premièrement, il semblerait que beaucoup de ces étudiants considèrent les mathématiques comme trop abstraites, ce qui les empêche de mobiliser les outils mathématiques qui leur sont enseignés. Gonzalez-Martin, Guedet, Barquero, et Romo-Vázquez (2021) expliquent que cela proviendrait du manque de lien entre les contenus de mathématiques enseignés et les autres disciplines. L'utilisation des outils mathématiques dans ces matières (physique, chimie, biologie), n'est pas assez motivée et trop implicite.

La formation scientifique, dans tous les domaines, nécessite des connaissances mathématiques et donc un enseignement de ces concepts, y compris à un niveau universitaire (Batanero, Chernoff, Engel, Lee, & Sanchez, 2016). Comme le soulignent Chiel et al. (2010), il est important de construire une base solide en mathématiques pour préparer les étudiants non-spécialistes à une recherche interdisciplinaire.

Cependant, les étudiants considèrent souvent leurs enseignements de mathématiques trop éloignés de leur future activité professionnelle. Gonzalez-Martin et al. (2021) ajoutent à ce sujet que, de façon générale, les programmes de mathématiques sont trop déconnectés des attentes professionnelles de ces étudiants non-spécialistes.

Viirman et Nardi (2018) insistent, les compétences et la confiance des étudiants non-spécialistes en mathématiques doivent être améliorées. Il est possible d'améliorer la motivation de ces étudiants non-spécialistes, pour les mathématiques, en leur montrant qu'elles sont utiles pour résoudre des problèmes dans d'autres disciplines scientifiques et qu'elles leur seront utiles dans leurs futures pratiques professionnelles.

La spécificité de la biologie

La biologie est une science essentiellement descriptive. Le développement récent de nouveaux outils de modélisation mathématique a eu un impact fort en la rendant plus mathématique, c'est-à-dire en lui donnant un rôle prédictif et décisionnel (Duran & Marshall, 2019).

Les relations entre la biologie et les mathématiques sont très importantes et pourtant Chiel et al. (2010) constatent que les étudiants de biologie ont un fort sentiment d'antipathie envers les matières quantitatives et notamment les mathématiques. Pourtant, et afin de répondre aux exigences de leur futur métier, les étudiants de biologie auront besoin d'avoir des connaissances mathématiques et surtout des compétences quant à l'utilisation d'outils et de procédés mathématiques Duran et Marshall (2019).

Comme l'expliquent Chiel et al. (2010) les enseignants de biologie et les enseignants de mathématiques ont une manière très distincte de penser leurs enseignements. De façon générale, le processus de formation des étudiants de mathématiques et de biologie est très différent. Pour simplifier grossièrement, on peut dire que la formation en mathématiques se base sur l'abstraction, le formalisme, l'intuition et la rigueur ; tandis que la biologie repose davantage sur la mémorisation, la terminologie, la caractérisation et l'expérimentation. Ainsi, un enseignant de mathématiques, parfois également chercheur, qui a suivi une formation uniquement de mathématiques risque d'être en décalage lorsqu'il enseigne à ces étudiants de biologie. Ces propos rejoignent le paragraphe sur les attentes du mathématicien expert.

Chiel et al. (2010) nous expliquent que ce qui attire les étudiants vers la biologie c'est finalement ce qui repousse les étudiants qui s'intéressent aux mathématiques. Le clivage entre ces différents groupes d'étudiants est généralement instauré dès le lycée, mais les auteurs insistent sur ce point, "*c'est un obstacle fondamental qui doit être surmonté*".

L'intérêt de la modélisation chez les étudiants non-spécialistes

La modélisation pourrait permettre de remédier aux difficultés que de nombreux étudiants en biologie rencontrent pour maîtriser les mathématiques et la modélisation. En effet, l'activité de modélisation permettrait de faire le lien entre les mathématiques et les autres disciplines. Henry (2010) explique que des exemples concrets pourraient aider les étudiants à surmonter certains obstacles.

Chiel et al. (2010) présentent des résultats très significatifs sur l'impact d'activités de modélisation mathématique de systèmes biologiques chez les étudiants de biologie. Il semblerait qu'après une première activité de modélisation, les étudiants se sentent plus à-même de construire et d'utiliser à nouveau des modèles mathématiques. Les auteurs constatent que les étudiants habituellement les plus en difficultés en mathématiques, sont ceux qui réussissent le mieux face à des problèmes ouverts issus de situations réelles. Les auteurs concluent en expliquant qu'en travaillant à partir de situations concrètes et en permettant aux étudiants de faire des expériences, il est possible de surmonter les obstacles précédemment mentionnés.

Selon Gonzalez-Martin et al. (2021) l'activité de modélisation permet une meilleure perception de l'utilité des mathématiques par les étudiants non-spécialistes en général. D'après ce que nous avons dit précédemment, l'activité de modélisation rend donc possible l'amélioration des compétences en mathématiques. De plus, l'implication des étudiants de biologie dans des activités de modélisation est un facteur de motivation (Viirman & Nardi, 2018).

Cependant il n'est pas toujours facile de faire cohabiter le raisonnement mathématique avec le raisonnement biologiste. En effet, Viirman et Nardi (2018) détaillent dans leur étude une activité de modélisation organisée par des enseignants de mathématiques avec des étudiants de biologie. Pour l'enseignant il est nécessaire d'utiliser des relations mathématiques pour construire un modèle tandis que les étudiants souhaitent commencer par un recueil des données empiriques. Il y a ici un conflit entre le point de vue du mathématicien et celui du chercheur en biologie. Dans leur étude, l'enseignant de mathématiques se place dans la version mathématique du problème et propose d'utiliser des outils mathématiques pour, notamment, vérifier la validité du modèle alors que les étudiants interprètent les résultats de façon empirique.

Ainsi l'alliance du raisonnement mathématique et du raisonnement biologiste est délicate. Mal amenée elle pourrait entraver l'activité de modélisation mathématique nécessaire afin d'analyser et de comprendre des phénomènes issus de la biologie.

Même si, comme nous venons de le voir, l'activité de modélisation peut être un levier pour aider les étudiants qui rencontrent des difficultés en mathématiques, l'activité de modélisation est une source de complications chez certains étudiants (voir partie 1.2, page 17). Comme Gonzalez-Martin et al. (2021) le précisent, les étudiants non-spécialistes ont du mal à transférer leurs connaissances mathématiques dans des problèmes issus d'autres disciplines, telle que la biologie. Comme nous venons de le décrire, ces étudiants savent utiliser des modèles mathématiques qu'ils ont déjà vus, mais ils ont des difficultés à argumenter la validité du modèle ainsi qu'à trouver des raisons pour lesquelles ce modèle a un sens dans le cas particulier de l'exercice.

Enfin, concernant plus particulièrement le cas qui nous intéresse des probabilités, soulignons que la modélisation probabiliste est largement utilisée par les chercheurs en biologie. Ces chercheurs, qui travaillent dans des domaines assez divers (biostatistique, biochimie, écologie, biologie moléculaire, comportement animal, etc) se servent de modèles probabilistes, d'outils de statistiques descriptives et inférentielles (Duran & Marshall, 2019). On peut par exemple citer les modèles stochastiques en immunologie (Nankep, 2018) ou encore les modèles statistiques en épidémiologie (e.g. modèles SIR (Weiss, 2013)).

Pour résumer

Les difficultés qu'éprouvent les étudiants non-spécialistes en mathématiques sont dues en partie au manque de liens entre l'enseignement des mathématiques qui est proposé à l'Université et les futures pratiques professionnelles de ces étudiants. En ce qui concerne les étudiants de biologie, ils éprouvent généralement peu d'intérêt pour les mathématiques et rencontrent de nombreuses difficultés dans cette discipline. Il est nécessaire de montrer à ces étudiants que les mathématiques leur sont utiles dans d'autres domaines scientifiques.

L'activité de modélisation pourrait être un levier pour éviter ces difficultés mais il est nécessaire d'être attentif aux nombreux obstacles que cela peut provoquer.

1.4 A propos des particularités liées à l'enseignement et à l'apprentissage des probabilités et notamment des probabilités conditionnelles

Les étudiants, et notamment les étudiants non-spécialistes, ont besoin de connaissances et de compétences en probabilités. Ils devront être capable de prendre des décisions adaptées à des situations où il y a du hasard. Selon Batanero et al. (2016), il est nécessaire d'apprendre à dépasser la pensée déterministe, c'est-à-dire, accepter l'existence d'une chance fondamentale dans la nature. Pour cela, il faut avoir des connaissances en probabilités. D'après Martignon et Wassner (2002), de façon générale, les étudiants considèrent les probabilités comme plus intéressantes que les autres domaines des mathématiques mais aussi comme plus difficiles.

Mais alors quelles sont les spécificités des probabilités, notamment en ce qui concerne les difficultés rencontrées par les étudiants ? Nous essayons de répondre à cette question en reprenant des travaux qui se sont intéressés à ce sujet, tout d'abord en ce qui concerne les probabilités en général puis plus spécifiquement pour les probabilités conditionnelles.

Premiers éléments concernant les difficultés en probabilités

Le champ des probabilités est très vaste et les difficultés que rencontrent les étudiants dépendent des contenus. Dans ce travail nous nous intéressons aux thèmes des probabilités conditionnelles dans le champ des probabilités discrètes ainsi qu'aux variables aléatoires continues, dans le champ des probabilités continues. Cependant il existe des difficultés communes à ces deux thèmes, propres au domaine des probabilités.

Carranza et Kuzniak (2006) indiquent que les exercices présents dans les manuels de lycée en France, sont souvent très techniques et que les questions présentent beaucoup d'indices afin que l'élève fasse les calculs demandés.

Batanero et al. (2018) expliquent que les étudiants éprouvent des difficultés lorsqu'ils rencontrent plusieurs types de probabilités à calculer dans un même exercice (une probabilité d'intersection et une probabilité conditionnelle par exemple). Les étudiants rencontrent également des difficultés dans les preuves de propriétés abstraites.

Une difficulté pour beaucoup d'étudiants, spécifique au domaine des probabilités, est la relation, parfois trompeuse, entre l'intuition et le développement conceptuel des notions de probabilités. Afin d'y remédier, Batanero et al. (2005) préconisent de renforcer la compréhension intuitive des concepts avant l'enseignement formel des probabilités.

Un enseignement des probabilités très formel

Un constat qui ressort des lectures que nous avons faites est que les difficultés rencontrées par les étudiants seraient liées à un enseignement trop formel des probabilités. Les problèmes de probabilités issus de situations réelles sont concrets et pourtant l'enseignement des probabilités, des concepts associés et les instruments présentés aux étudiants sont dans un langage très formel (Martignon & Wassner, 2002). Comme le soulignent Amrani et Zaki (2015), un enseignement des probabilités basé sur la théorie des ensembles ou sur des techniques de dénombrement, empêche la bonne compréhension, par les étudiants, des phénomènes associés (les auteurs donnent comme exemples de ces phénomènes mal compris : faire à tort l'hypothèse d'une équiprobabilité ou d'une indépendance d'évènements).

Les étudiants éprouvent des difficultés lorsqu'ils rencontrent des exercices qui portent sur les aspects théoriques des concepts de probabilités enseignés (Amrani & Zaki, 2015). Et inversement, Martignon et Wassner (2002) constatent que, face à des problèmes dans un contexte de la vie réelle, les étudiants ont des difficultés à mobiliser les concepts abstraits qu'ils connaissent et qui se prêtent à la situation.

Selon Amrani et Zaki (2015), un enseignement basé sur une approche plus intuitive des probabilités permettrait à la fois une meilleure appréhension de la modélisation probabiliste par les étudiants et améliorerait la production d'exemples par ces étudiants.

Le cas particulier des probabilités conditionnelles

Nous nous intéressons dans cette partie aux travaux et aux résultats qui concernent plus particulièrement les probabilités conditionnelles.

Carrasco et Elicer (2017) se sont intéressés à des scénarios de prise de décisions dans le cadre d'enseignements de probabilités au lycée. Les probabilités conditionnelles servent justement lorsqu'il faut s'adapter à de nouvelles conditions. Les auteurs soulignent que les décisions des élèves se font souvent suivant des considérations heuristiques ou des expériences personnelles. Les élèves s'appuient sur des arguments non mathématiques pour faire leurs choix et expliquent ne pas voir la nécessité de faire un calcul de probabilités pour répondre. C'est là un des paradoxes des probabilités, face à un problème décisionnel, les élèves considèrent l'expérience (personnelle ou heuristique) comme plus puissante que le calcul de probabilités.

Totohasina (1992) recommande d'utiliser la notation indicielle $P_A(B)$ pour une première introduction de cette notion et de lire cette expression de la façon suivante : "*probabilité sachant A de (l'évènement) B*". Selon lui la notation $P(B|A)$ pourrait laisser penser - à tort - qu'il s'agit d'une fonction P appliquée à deux variables B et A . Henry (2005a) explique que la notation $B|A$ fait penser à un "évènement" conditionnel, terme qui pourrait avoir un sens mais qui ajoute une complexité, selon lui, inutile.

Difficultés spécifiques aux probabilités conditionnelles

Selon plusieurs auteurs (Huerta, 2014; Diaz & De la Fuente, 2007), les étudiants ont des difficultés à interpréter les quantités numériques présentes dans l'énoncé. Ils confondent, par exemple, les probabilités d'intersection d'évènements avec les probabilités conditionnelles. Diaz et De la Fuente (2007) explique qu'il y a souvent une ambiguïté verbale dans l'expression de la probabilité conditionnelle, ce qui rend difficile pour l'étudiant de distinguer la probabilité conditionnelle de la probabilité de l'intersection.

De la même façon, Diaz et De la Fuente (2007) indique que les étudiants ont des difficultés pour inverser les évènements A et B dans l'expression $P_B(A)$ car ils confondent souvent le rôle de l'évènement qui conditionne et de l'évènement qui est conditionné. L'auteur ajoute que ces difficultés proviennent de la définition même de la probabilité conditionnelle, définition qui n'a pas été bien comprise mais aussi d'une idée fautive que l'on retrouve chez certains élèves : $P_B(A)$ est une relation temporelle.

Martignon et Wassner (2002) expliquent que pour les théorèmes les plus importants des règles mnémotechniques sont parfois enseignées comme les "règles du chemin" (pour les arbres de probabilités) et la "règle de la probabilité totale". Le

problème est que ces règles sont rapidement oubliées.

Pour Parzysz (2011), la conception "cardinaliste" de la probabilité conditionnelle est une cause de difficulté. Il est nécessaire de passer des effectifs aux fréquences et de choisir des situations dans lesquelles il n'est pas possible d'avoir recours aux effectifs.

Batanero et al. (2005) expliquent que les difficultés que rencontrent les étudiants en probabilités proviennent du grand nombre de paradoxes dans ce domaine. Beaucoup d'entre eux reposent sur des malentendus et de la contre-intuition, y compris en ce qui concerne les probabilités conditionnelles.

Un exemple remarquable est celui de la notion d'indépendance d'évènements. Comme l'expliquent Batanero et al. (2005), historiquement l'indépendance de deux évènements était naturelle et intuitive, c'est comme ça que le concept a émergé : *deux évènements sont indépendants s'il n'y a aucune raison de penser que l'un d'eux pouvait influencer l'autre*. Ensuite la notion d'indépendance a été formalisée et il est devenu difficile de relier l'idée intuitive d'indépendance et sa définition formelle. Ce paradoxe se retrouve dans l'enseignement des probabilités et de la notion d'indépendance de deux évènements, pour laquelle certains étudiants ont des difficultés car ils raisonnent en utilisant l'intuition d'indépendance.

Registres de représentations et arbres

En probabilités, les différents registres de représentation de Duval (1993) (voir détails partie 2.1 page 36) sont très présents dans l'enseignement secondaire, à la fois comme illustrations et comme outils de résolution de problèmes. Les registres suivants sont associés à la résolution de problèmes : langagier, symbolique, arbres et tableaux. Les représentations permettent de traduire des situations concrètes en pseudo-concrètes et ont une influence sur la bonne compréhension des problèmes par les étudiants (Martignon & Wassner, 2002).

Cependant, comme le souligne Parzysz (2011), l'utilisation de plusieurs registres de représentations (tableaux, graphiques, arbres, boîtes etc) ne doit pas empêcher de construire du sens pour chacun d'eux, ni d'apprendre à les articuler, au risque de voir disparaître le bénéfice de leur utilisation.

Les registres de représentations, et notamment les arbres de probabilités, ont une place importante dans l'enseignement des probabilités conditionnelles. A ce propos, Diaz et De la Fuente (2007) explique qu'il est nécessaire d'enseigner les probabilités conditionnelles en utilisant des représentations, comme des arbres de probabilités. L'intérêt des arbres de probabilités est qu'ils sont lisibles pour les élèves et qu'ils permettent de faire apparaître une grande quantité d'informations.

Eichler (2007) qui s'est intéressé à des élèves de lycée en Allemagne, explique que les élèves savent utiliser l'arbre de probabilités dans des situations

1.4. A propos des particularités liées à l'enseignement et à l'apprentissage des probabilités et notamment des probabilités conditionnelles

appropriées. Il observe également que l'enseignant utilise l'arbre pour faciliter la compréhension de l'indépendance d'évènements et du théorème de Bayes. Totohasina (1992) qui s'est intéressé à des étudiants de début d'Université sur des problèmes de Bayes, observe que les étudiants ont des difficultés pour résoudre de tels problèmes lorsqu'ils ont fait le choix d'une représentation graphique de type tableau à double entrée plutôt que celui d'un arbre. L'auteur observe des confusions entre probabilités conditionnelles et probabilités d'intersection chez ces étudiants, il explique que "*le tableau à double entrée apparaît ici comme un piège de l'intuition*" alors qu'à sa place un "*arbre de probabilités joue un rôle à la fois descriptif et heuristique*".

L'utilisation des arbres de probabilités peut également être une cause de difficultés chez les étudiants. En effet, comme l'explique Parzysz (2011), l'énoncé est généralement en langage naturel, il faut produire un modèle "pseudo-concret" et il est nécessaire d'effectuer une conversion correcte de cet énoncé avant de pouvoir utiliser l'arbre.

Concernant le format des données numériques dans les énoncés

Une autre source de difficultés, dans le cas des probabilités, serait celle du format des données. En effet, plusieurs auteurs s'accordent à dire que celui-ci impacterait les étudiants, et leurs réussites dans la résolution des exercices et problèmes. Cependant ces auteurs ne sont pas tous d'accord sur le format à privilégier : pourcentages, fréquences relatives ou fréquences naturelles.

Huerta (2014) constate, dans son étude, que les élèves ont de meilleurs résultats lorsque les données numériques de l'énoncé sont en pourcentage ou en fréquences.

Batanero et al. (2018) qui ont analysé les résultats des étudiants aux tests d'entrée à l'Université en Andalousie, observent que les données en pourcentage ainsi que le mélange de formats de données au sein d'un même exercice posent des difficultés aux étudiants.

Selon Diaz et De la Fuente (2007) les calculs de probabilités conditionnelles sont plus faciles pour les étudiants lorsque les données chiffrées des exercices sont en fréquences naturelles.

Pour faciliter la compréhension de l'indépendance et du théorème de Bayes, Eichler (2007) préconise l'utilisation de l'arbre de probabilités avec des données en fréquences naturelles, plutôt que l'arbre traditionnel (avec des fréquences relatives). Il a également constaté dans son étude que la compréhension des élèves est entravée par le choix de l'enseignant d'utiliser un tableau de probabilités dans le cas où les données sont en pourcentage.

Nous concluons cette sous-partie par une recommandation de Diaz et De la Fuente (2007) qui explique qu'il serait souhaitable d'enseigner les probabilités conditionnelles conjointement aux statistiques en présentant notamment des applications à des situations réelles. En effet, l'arbre de probabilités est commun aux statistiques et aux probabilités et peut donc constituer un point d'articulation entre ces deux domaines (Parzysz, 2011).

Pour résumer

Les difficultés que les étudiants du supérieur rencontrent en probabilités sont dues, en partie, à un enseignement trop formel de ces notions de probabilités. Les probabilités peuvent être difficiles car elles sont parfois contre-intuitives.

L'utilisation d'un arbre de probabilités associée à des données en fréquence permet une meilleure compréhension chez l'élève des phénomènes aléatoires.

1.5 A propos de l'enseignement et de l'apprentissage des variables aléatoires continues

Dans cette partie nous nous intéressons plus particulièrement au cas des probabilités continues et notamment des lois continues. Les lois continues que l'on retrouve dans l'enseignement en classe de terminale scientifique sont la loi uniforme, la loi exponentielle et la loi normale.

Les lois continues dans l'enseignement secondaire

Les lois continues font partie du programme de lycée depuis 2002 en France, à cette époque seule la loi uniforme et la loi exponentielle sont au programme (voir partie 4.2 page 56 pour une analyse des programmes). Le programme de terminale opte pour une définition analytique via la densité de probabilités et Henry (2003) s'interroge sur la pertinence de l'introduction des lois continues en classe de terminale scientifique au lycée. Il considère comme "audacieux" d'introduire ces notions à ce niveau scolaire et s'interroge sur la définition qui est proposée de la notion de densité, et sur comment donner un sens probabiliste à cette notion auprès des élèves de lycée.

A ce sujet, Girard (2003) insiste, l'introduction de la notion de probabilités continues, comme elle est faite dans les programmes de lycée, est porteuse d'un obstacle épistémologique, car celle-ci ne permet pas de voir le caractère "approximatif" de ce champ des mathématiques.

L'introduction des lois continues en terminale offre une source d'exercices variés qui sont, d'après Henry (2003), pour les élèves, plus intéressants que les problèmes traditionnels de combinatoire (jeux de hasard par exemple). Girard (2003) constate néanmoins que les élèves de lycée rencontrent des difficultés lorsqu'il y a par exemple des données en pourcentage, un vocabulaire trop ensembliste, et plus largement vis-à-vis du langage probabiliste utilisé.

Modélisation dans le cas des lois continues

Selon Henry (2003) le point de vue de la modélisation est incontournable dans le cas des lois continues. En effet, ces lois sont des outils mathématiques indispensables permettant de calculer des probabilités lorsque l'on se place dans un modèle adéquat.

Comme nous l'avons exposé dans la sous-partie 1.2 page 17, l'enseignement des probabilités doit être enrichi par l'analyse de données issues d'expériences. Henry (2003) insiste, la modélisation par des lois continues est très importante. L'enseignement dans le secondaire des lois continues, associé à la simulation informatique de situations simples rend incontournable la notion de modèle.

Concernant le cas particulier de la loi exponentielle qui est désignée dans le programme de terminale par "loi de durée de vie sans vieillissement", Henry (2003) explique que certains exemples, comme celui de la radioactivité, permettent une bonne illustration de cette loi. En revanche il précise que la discrétisation et le passage par la loi géométrique, qui permettent de comprendre pourquoi la densité de probabilités intervenant dans ce modèle est de type exponentielle, sont hors d'atteinte en terminale.

Selon Batanero, Tauber, et Sanchez (2004), l'activité de modélisation, dans le cas des lois continues, occasionne des difficultés spécifiques. Par exemple, les étudiants rencontrent des difficultés à traiter les données empiriques et le modèle mathématique simultanément, car il faut à la fois analyser les caractéristiques de l'ensemble des données et les caractéristiques du modèle, tout en faisant attention à bien distinguer le modèle des données. Comme le souligne Henry (2003), il est important de distinguer la statistique, qui décrit la réalité, des probabilités, qui sont une modélisation de la réalité.

Batanero et al. (2004) relèvent d'autres difficultés, en lien avec la modélisation probabiliste et rencontrées par les étudiants dans le cas particulier d'exercices sur la loi normale. Les étudiants ont des difficultés à comprendre l'utilité et l'intérêt des modèles théoriques pour décrire des données empiriques. Ils ont des difficultés à distinguer les données empiriques des modèles mathématiques. Pour changer cela il faudrait faire davantage d'activités reliant les données au modèle mathématique. Enfin, les auteurs observent des difficultés à reconnaître quand une variable discrète peut ou ne peut pas être modélisée par une loi normale.

Wilensky (1997) s'est intéressé aux comportements d'étudiants face à des problèmes de modélisation faisant appel à la loi normale. Il constate que même si les étudiants savent résoudre les différents problèmes ils sont cependant dans l'incapacité d'expliquer pourquoi ils ont choisi la distribution normale plutôt qu'une autre. L'auteur observe un sentiment de confusion et d'indécision chez ces étudiants lorsqu'ils sont face à des possibilités différentes pour résoudre un problème, il définit cela comme "l'anxiété épistémologique".

Le cas particulier de la loi normale

Dans leur recherche Batanero et al. (2004) se sont intéressés plus particulièrement au cas de la loi normale et présentent des résultats très intéressants que nous exposons dans ce paragraphe. Il est notamment envisageable d'élargir ces résultats aux autres lois continues, nous y reviendrons.

La loi normale est importante pour les étudiants à l'Université, ils doivent la connaître et savoir l'utiliser. La loi normale permet notamment de modéliser des phénomènes biologiques. Cependant, comme le soulignent Batanero et al. (2004), derrière la loi normale il y a des notions complexes qui requièrent l'intégration de beaucoup d'idées et concepts statistiques.

Dans leur étude, Batanero et al. (2004) constatent que les étudiants utilisent beaucoup les représentations et graphiques dans leurs argumentations dans le cas de la loi normale. Les auteurs relèvent aussi de nombreuses difficultés rencontrées par les étudiants lorsqu'il s'agit de représentations graphiques. Par exemple les étudiants ont des difficultés à interpréter certains histogrammes de fréquences, graphiques statistiques et résumés statistiques. Les étudiants ont des difficultés à interpréter les probabilités comme des aires sous la courbe et à faire des calculs dans les cas où un changement d'intervalle est nécessaire (par exemple : modifier les bornes d'un intervalle).

Batanero et al. (2004) soulignent que la bonne compréhension et manipulation des représentations graphiques et des aires sous la courbe - pour le cas de la loi normale - est à la fois nécessaire pour résoudre les problèmes de loi normale et un outil didactique indispensable pour que les étudiants comprennent convenablement le calcul des probabilités.

Il est plutôt vraisemblable que ce constat ainsi que certaines des difficultés relevées par Batanero et al. (2004) ne soient pas spécifiques à la loi normale, et nous y serons attentifs dans la suite de ce travail.

Recommandations concernant l'enseignement des lois continues

Pour terminer cette section concernant les lois continues nous présentons plusieurs recommandations pour l'enseignement des lois continues et notamment de la loi normale, issues du travail de Batanero et al. (2004). Ces recommandations qui peuvent s'étendre aux lois continues autres que la loi normale, nous semblent particulièrement pertinentes dans le cadre de notre travail, et nous veillerons à en tenir compte.

Selon les auteurs, il est nécessaire de maîtriser les notions de base des probabilités continues (probabilité élémentaire, courbe de densité, symétrie, histogramme etc) avant de commencer l'étude de la loi normale.

Batanero et al. (2004) recommandent de travailler sur la compréhension intuitive des notions liées à la loi normale. Pour cela, il peut être intéressant de mettre en place des activités afin de permettre aux étudiants de lier les données réelles aux modèles, et ainsi de faciliter l'apprentissage des notions basiques sur la loi normale.

Selon ces auteurs, il est nécessaire de faire travailler les étudiants sur des données réelles, que ce soit pour la loi normale ou pour les autres lois continues. Travailler à partir de données pourrait notamment aider à améliorer l'intuition chez les étudiants, à leur apprendre les étapes du travail de modélisation et à les aider à faire la distinction entre modèle et réalité.

Enfin, Batanero et al. (2004) insistent sur le fait qu'il est important de travailler sur et avec les représentations graphiques pour aider les étudiants à bien comprendre et assimiler les concepts en lien avec la loi normale.

Pour résumer

Il est très important de mettre en évidence, dans l'enseignement des lois continues, le lien entre probabilités et modélisation. En effet, un enseignement formel de ces concepts peut entraver la compréhension des phénomènes associés.

Les étudiants rencontrent des difficultés spécifiques aux lois continues. Il s'agit à la fois de difficultés dans l'interprétation graphique de la probabilité, mais aussi pour mettre en place des modèles et les décrire.

Chapitre 2

Cadre théorique et questions de recherche

Nous présentons dans cette partie notre cadre théorique. Il s'agit de notions et de concepts, issus de différentes théories de didactique, qui nous permettront de répondre aux problématiques de cette étude. Ce cadre théorique nous servira à formuler nos questions de recherche, à construire une méthodologie, à mener à bien les analyses et à présenter nos résultats.

Nous allons nous appuyer sur des concepts issus de l'approche psychologico-cognitive de Duval (1993), de la Théorie Anthropologique du Didactique (TAD) introduite par Chevallard (1998) que nous compléterons avec des éléments de la Théorie de l'Activité (TA) adaptés à la didactique des mathématiques par Robert et Rogalski (Vandebrouck, 2008). Nous avons retenu pour notre travail la TAD car nous pensons qu'une perspective institutionnelle est susceptible d'éclairer les questions de transition institutionnelle. Nous l'avons complété par la Théorie de l'Activité car, seule, la TAD est insuffisante pour notre projet qui est d'analyser la transition d'un point de vue institutionnel mais aussi du point de vue de l'élève et de l'étudiant. La complémentarité de ces deux cadres nous permettra de voir comment la complexité de la tâche impacte l'organisation praxéologique et de répondre aux problématiques soulevées.

Nous allons d'abord présenter les registres de représentations sémiotiques de Duval (1993), il s'agit ici d'une approche cognitive. Ensuite nous proposons d'introduire différents concepts issus de la TAD qui vont nous être utiles pour la suite. Dans une seconde partie nous exposerons des éléments de la TA susceptibles de nous aider dans l'analyse de l'activité des élèves. Nous poursuivrons en présentant le concept de Modèle Praxéologique de Référence (MPR) et comment nous l'adaptions à notre travail. Enfin, nous terminerons par la formulation de nos questions de recherche au regard des apports qui auront été faits.

2.1 Registres de représentations sémiotiques

Dans cette partie nous présentons la notion de registre de représentation sémiotique introduite par Duval (1993). Étant donné que les concepts mathématiques sont abstraits et imperceptibles, l'usage de représentations est indispensable. Lors de la résolution d'un exercice c'est par les représentations sémiotiques que l'élève manipule les objets mathématiques. Toutefois il est important de savoir distinguer l'objet mathématique de la représentation que l'on en fait.

Les représentations sémiotiques permettent de rendre communicables les représentations mentales que l'on peut avoir d'un objet mathématique. Ces représentations ont également un rôle important dans l'activité cognitive de l'élève qui s'en sert. Pour qu'un système sémiotique (par exemple : langage symbolique, langage naturel, figures géométriques, etc) soit un registre de représentation sémiotique tel que défini par (Duval, 1993) il doit rendre possible les trois activités cognitives essentielles suivantes :

- La production ou la reconnaissance d'une représentation selon les règles de formation propres au registre dans lequel la représentation est faite.
- Le traitement d'une représentation, tout en restant dans un même registre et donc en se pliant aux règles qui y sont données.
- La conversion d'une représentation en une autre tout en changeant de registre. Cette transformation permet souvent un apport de sens.

(Parzysz, 2011) explique que les arbres de probabilités ont des règles de traitement et de conversion précises et possèdent ainsi les caractéristiques d'un registre de représentation . Dans sa thèse Nechache (2016), a mis en évidence que les tableaux de probabilités sont également des registres de représentations sémiotiques dans lesquels des traitements mathématiques sont permis.

Duval (1993) souligne qu'il est indispensable de pouvoir utiliser plusieurs registres sémiotiques de représentation pour un même objet mathématique, tout comme il est nécessaire d'être capable de choisir un registre plutôt qu'un autre.

Voici quelques exemples de registres que nous rencontrerons dans ce travail :

- registre symbolique probabiliste ou de l'intégrale (formules mathématiques),
- registre de la langue naturelle,
- registres de représentations : tableau de probabilités, arbre de probabilités (appelés aussi arbres pondérés), diagramme de Venn, graphique de fonction, aire sous la courbe, etc.

De nombreux registres sont mobilisés dans l'enseignement des probabilités, notamment pour ce qui concerne les probabilités conditionnelles (arbres, tableaux). Nous nous intéresserons aux attentes institutionnelles en termes d'emploi de registres et de conversion de registres, au lycée puis à l'Université.

2.2 Eléments issus de la Théorie Anthropologique du Didactique (TAD)

Nous présentons, dans cette sous-partie, certains des concepts fondamentaux issus de la Théorie Anthropologique du Didactique qui nous seront utiles dans ce travail, il s'agit des notions d'*institution* et de *praxéologie*.

Nous compléterons ces définitions avec des notions issues du cadre T4TEL développé par Chaachoua (2018) qui s'inscrit lui-même dans la Théorie Anthropologique du Didactique. Ce cadre est une extension du modèle praxéologique et permet d'enrichir la description des organisations mathématiques.

Nous commençons cette sous-partie avec un concept très important en TAD, la notion d'*institution*. Chevallard en donne la définition suivante :

Une institution I est un dispositif social "total", qui peut certes n'avoir qu'une extension très réduite dans l'espace social (il existe des "micro-institutions"), mais qui permet - et impose - à ses *sujets*, c'est-à-dire aux personnes x qui viennent y occuper les différentes *positions* p offertes dans I , la mise en jeu de *manières de faire et de penser propres* (Chevallard, 2003).

Nous sommes donc "*sujets*" des institutions qui nous entourent, une classe, ou un établissement scolaire peuvent être considérés comme des institutions.

Le postulat de base de la TAD est que les savoirs sont façonnés par les institutions dans lesquelles ils vivent. L'analyse des savoirs mathématiques doit aller de pair avec l'étude des pratiques institutionnelles. De cette façon nous considérons la transition secondaire-supérieur comme une transition institutionnelle : dans l'institution Université les mathématiques enseignées aux étudiants de mathématiques ne sont pas les mêmes que les mathématiques enseignées aux étudiants de biologie et ne sont pas les mêmes que les mathématiques enseignées dans l'institution lycée.

Dans la TAD, les savoirs sont considérés comme des productions humaines et l'activité mathématique fait partie de l'ensemble des activités humaines. Toute activité humaine peut être modélisée par ce que l'on appelle une "*praxéologie*". La praxéologie est relative à la personne qui la met en oeuvre ou à l'institution dans laquelle elle vit. Une praxéologie (ou organisation mathématique) est un quadruplet $(T, \tau, \theta, \Theta)$, composé des éléments suivants :

- * Le **type de tâches**, noté T , est toujours décrit par un verbe d'action. La notion de tâche qui est utilisée ici fait référence à une action au sens large, par exemple : "multiplier deux matrices entre-elles" ou "composer une chanson" sont ainsi des types de tâches.
- * Soit le type de tâches T à effectuer. Il existe au moins une manière d'accomplir T , c'est ce qu'on appelle la **technique**, notée τ . Le type de tâches et la

technique forment le bloc de la *praxis* ou *bloc pratico-technique* noté $[T/\tau]$ que l'on désigne couramment par le terme *savoir-faire*.

- * La **technologie**, notée θ , est le discours qui justifie la technique τ . La technologie θ a deux fonctions principales. Elle justifie que la technique donne bien ce qui est annoncé et permet d'accomplir T , c'est une fonction de *justification*. La technologie permet également d'expliquer et de rendre compréhensible la technique, c'est une fonction d'*explication*.
- * Il existe un discours, noté Θ , dont le rôle est de justifier la technologie θ , ce discours est appelé **théorie**. La technologie et la théorie forment le bloc du *logos* ou *bloc technologico-théorique* noté $[\theta/\Theta]$ que l'on nomme aussi *savoir*.

A partir d'un type de tâches T , on trouve une *technique* τ , une *technologie* θ , et une *théorie* Θ . L'ensemble forme un quadruplet qu'on appelle **praxéologie relative au type de tâches T** , ou encore *praxéologie ponctuelle*, notée $[T/\tau/\theta/\Theta]$.

Afin de mieux décrire les organisations mathématiques, on introduit la notion de **sous-type de tâches**, notée T' issue du cadre T4TEL. Un sous-type de tâches T' , du type de tâches T , est à la fois un sous-ensemble de T et un type de tâches.

Par exemple : le type de tâches "calculer la probabilité de l'intersection de deux événements indépendants" est un sous-type de tâches de "calculer la probabilité de l'intersection de deux événements" qui est lui-même un sous-type de tâches de "calculer la probabilité d'un événement".

Chaachoua (2018) propose de décrire la technique par un ensemble de types de tâches. Ces types de tâches sont alors appelés **ingrédients de la technique**.

Par exemple : pour le type de tâches "calculer la probabilité de l'intersection de deux événements indépendants", les ingrédients d'une technique peuvent être : "calculer les probabilités des événements un à un" et "multiplier les valeurs de ces probabilités".

Ainsi, chaque type de tâches qui est également ingrédient de technique, est réalisé par une ou plusieurs techniques qui s'expriment elles aussi par un ensemble de types de tâches. Afin de ne pas se retrouver face à une description infinie il sera à notre charge de choisir le niveau de granularité pertinent où il sera bon de s'arrêter pour décrire ces techniques.

2.3 Eléments de la Théorie de l'Activité (TA)

Il nous semble pertinent d'utiliser la Théorie de l'Activité afin d'examiner de plus près la complexité de l'activité de l'élève (et de l'étudiant) pour une tâche donnée et de prendre en compte les spécificités des étudiants débutants à l'Université, qu'il s'agisse de leurs ressources, leurs histoires, le contexte de leur

enseignement ou de leurs connaissances antérieures (habitudes, contrats, etc.).

Nous nous plaçons en Théorie de l'Activité, telle qu'elle a été adaptée en didactique des mathématiques (Robert, 1998 ; Vandebrouck, 2008). Les définitions principales de ce cadre sont rappelées dans la suite de cette partie et nous commençons par cette description de Rogalski (2008) :

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 cote de la situation" et "du cote du sujet". (Rogalski, 2008)

Ainsi, l'étude de l'activité est toujours définie du point de vue du *sujet*, ici il s'agit de l'élève ou de l'étudiant. Le sujet a des intentions, des compétences et des responsabilités.

Pour analyser l'activité de l'élève il faut tenir compte du contexte en jeu, ici les mathématiques et le cadre scolaire, ce que nous appelons la *situation* (il ne s'agit pas de la situation au sens de la TSD définie par Brousseau). Le sujet agit dans la situation qui lui apporte des contraintes mais aussi des ressources.

Il faut bien distinguer la *tâche* de l'*activité*. La tâche est l'objet de l'activité, c'est en fait ce que le sujet doit faire. La tâche provoque des *adaptations* à la charge de l'élève.

La tâche correspond au but que le sujet doit atteindre sous certaines conditions. (Leontiev, 1984)

Tandis que l'activité correspond aux moyens d'actions et aux interactions de l'élève avec son environnement.

L'activité est ce que développe le sujet lors de la réalisation de la tâche (actions, interactions visibles mais aussi inférences, hypothèses, décisions, non actions, manières de gérer le temps, état personnel...). (Vandebrouck, 2008)

Les activités mathématiques sont des mises en fonctionnement de connaissances mathématiques que l'on cherche à qualifier. Les connaissances sont-elles anciennes ou bien nouvelles ? Les niveaux de mises en fonctionnement sont-ils disponibles ou mobilisables ? S'agit-il d'applications directes des connaissances ou d'applications avec des adaptations ?

L'emploi de la Théorie de l'Activité dans ce travail va nous permettre par exemple de souligner que la complexité de lier une tâche proposée à un type de tâches est une caractéristique de la transition secondaire-supérieur.

Afin de préciser du mieux possible l'activité mathématique des élèves et des étudiants nous utilisons le concept de *sous-activités mathématiques* défini par Robert et Vandebrouck (2014). Ces sous-activités sont de trois types :

- sous-activités de reconnaissance d'outils ou d'objets mathématiques à mettre en fonctionnement. Par exemple reconnaître la situation ou la tâche (objet), et reconnaître la méthode ou le théorème à utiliser (outil).
- sous-activités d'organisation du raisonnement global. Par exemple organiser les raisonnements entre eux ou les étapes, leur ordre, etc.
- sous-activité de traitement interne. Il s'agit par exemple des calculs à mener et des changements de registres.

Une tâche est alors définie comme complexe lorsqu'elle combine plusieurs de ces sous-activités (Robert & Vandebrouck, 2014). L'emploi de la Théorie de l'Activité en complément de la Théorie Anthropologique du Didactique nous permet donc d'étudier comment la complexité de la tâche impacte l'organisation praxéologique et en particulier les techniques à mettre en oeuvre.

Désormais, et pour le reste de ce travail, nous utilisons ici la notion de tâche telle que décrite dans la Théorie de l'Activité, c'est-à-dire en référence à l'objet de l'activité et à sa description.

2.4 Modèle Praxéologique de Référence Étendu et adapté à la Transition secondaire-supérieur : MPR-ET

Modèle Praxéologique de Référence dans la TAD

Le concept de Modèle Praxéologique de Référence (MPR) a été développé dans le cadre de la TAD. Il s'agit d'un modèle visant l'analyse critique des complexes praxéologiques mis en oeuvre dans l'institution étudiée. Dans notre cas, les institutions étudiées sont la dernière année de lycée et la première année d'Université en filière de biologie, en France.

Une organisation mathématique (OM) à enseigner constitue un modèle praxéologique du curriculum mathématique obtenu à partir d'une analyse des programmes et des manuels. L'identification de ces OM à enseigner passe donc par la caractérisation des types de tâches institutionnels et peut être vue comme une "reconstruction" du chercheur. [Le chercheur] peut procéder à un autre découpage que celui de l'institution voire de le compléter; il construit alors un modèle praxéologique de référence (MPR) regroupant les praxéologies à enseigner, enseignées mais également enseignables. (Chaachoua, Ferraton, & Desmoulins, 2017) (page 302)

Cette citation nous donne une idée générale de ce sur quoi peut s'appuyer la construction d'un MPR : analyse praxéologique des curricula (savoirs à enseigner), des manuels, des séances en classe (savoirs enseignés).

Comme il est mentionné par les auteurs ci-dessus, le MPR est une construction du chercheur qui le conçoit selon ses questions de recherche. Dans notre cas nous souhaitons, préalablement à l'analyse des programmes et des manuels, commencer par une analyse historique des savoirs. En effet, comme le souligne Dorier (2000b) :

L'étude du savoir mathématique 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.

Un éclairage historique et épistémologique des savoirs en jeu est donc susceptible d'aider à situer les difficultés et les obstacles épistémologiques qui leur sont inhérents. Il peut être intéressant de comprendre à quels problèmes historiques la notion répond ou de connaître l'état de la science au moment de l'apparition de ces notions.

Modèle Praxéologique de Référence Étendu - par la Théorie de l'Activité - et adapté à la Transition secondaire-supérieur

De façon générale, un MPR se construit à partir d'analyses praxéologiques (curricula, manuels, etc) en appui sur l'étude historique et épistémologique des savoirs. Cependant nous souhaitons ici construire un MPR adapté à l'institution Université mais aussi adapté aux étudiants qui sortent du lycée. Aussi, nous compléterons ce premier travail d'analyse des praxéologies par des analyses du point de vue de la Théorie de l'Activité, afin de prendre en compte les acteurs qui sont ici les élèves et étudiants.

Les modalités d'analyses seront, par exemple, de recenser les variations de type de tâches, les adaptations et les sous-activités à la charge de l'élève (ou de l'étudiant). Nous appellerons le modèle ainsi obtenu : **MPR Étendu**, sous-entendu *étendu par la Théorie de l'Activité*. Nous aurions pu, avec l'analyse épistémologique et l'analyse des supports offerts par l'institution, constituer un MPR Étendu de l'Université et un MPR Étendu du lycée, puis les comparer. Mais étant donné que l'objectif final est de faire des propositions en vue de construire un enseignement innovant qui réduirait les difficultés des étudiants à la transition, nous sommes centrés sur l'Université où il nous semble plus simple de faire ce type de proposition.

Le modèle [praxéologique de référence] rend ainsi possible l'analyse de ce qui a cours dans différentes instances d'un système d'enseignement (comme les manuels ou le cours d'un enseignant), permet de rendre compte de la variété des OM à enseigner, de repérer et donc de pallier les manques éventuels. (Chaachoua et al., 2017) (page 302)

C'est dans le sens décrit à la toute fin de cette citation que nous nous plaçons pour la construction de notre MPR. En effet, en complément de la double approche TAD-TA, nous avons pour objectif de produire un MPR étendu pour l'Université qui serait lui-même un outil pour accompagner la transition : proposer des pistes pour l'enseignement des contenus, pallier les manques, servir d'appui pour la construction d'une ingénierie, etc.

Nous appellerons le modèle ainsi obtenu : **MPR Étendu par la théorie de l'activité et adapté pour la Transition secondaire-supérieur**, noté désormais **MPR-ET**.

Dans cette thèse nous n'allons donc pas, comme cela se fait habituellement, construire un MPR préalable comme outil pour l'analyse de séances, mais nous allons analyser et synthétiser nos résultats dans un MPR étendu pour la transition (MPR-ET).

Le MPR-ET ainsi obtenu ne vise pas l'exhaustivité des organisations mathématiques à enseigner ou enseignées dans l'institution car nous nous appuyons sur les éléments dont nous disposons suite aux observations et aux données recueillis auprès des étudiants.

Le MPR-ET que nous allons construire va permettre de synthétiser les réponses aux questions suivantes :

- Quelles organisations mathématiques présentes au lycée sont reprises et adaptées à l'Université ?
- Quelles techniques, nouvelles ou non, vise-t-on pour les étudiants ?
- Quelles sont les continuités ou ruptures entre institutions en terme de types de tâches et d'adaptations à la charge des étudiants ?
- Comment décrire les attentes institutionnelles envers les étudiants à la fin du cours de probabilités de première année de biologie ?

La construction de notre MPR-ET apparaîtra ainsi comme une synthèse de notre étude des continuités et ruptures lors de la transition lycée-Université. Ce modèle devrait nous permettre de donner à voir de manière synthétique les attentes institutionnelles envers les étudiants à la fin du cours de probabilités de première année de biologie en croisant le point de vue de l'institution avec celui des étudiants.

2.5 Questions problématisées

La revue de travaux et les éléments théoriques introduits ci-dessus nous permettent de présenter maintenant les questions de recherche auxquelles nous allons tenter de répondre, en lien avec les objectifs présentés dans l'introduction

page 9.

Dans le but de concevoir un Modèle Praxéologique de Référence Étendu et adapté à la Transition secondaire-supérieur nous nous questionnons sur les enjeux didactiques relatifs aux différentes approches probabilistes et sur les attendus d'un point de vue institutionnel dans le secondaire. C'est l'objet de notre première question de recherche **QR.1**.

QR.1 : Quel est le savoir à enseigner en probabilités dans le secondaire en classe de terminale scientifique (selon une perspective épistémologique et institutionnelle) ?

QR.1.1 Comment historiquement sont apparus les différents courants et approches probabilistes et quels sont les impacts de ces approches sur l'enseignement des probabilités ?

QR.1.2 Quels sont les attendus et recommandations institutionnelles dans le secondaire en classe de terminale scientifique, pour le thème des probabilités ?

QR.1.3 Comment décrire le savoir à enseigner pour le thème de probabilités, d'après les manuels et sujets de baccalauréat ?

Comme nous l'avons développé précédemment, notre problématique concerne la transition secondaire-supérieur, en probabilités, pour les étudiants de biologie. C'est le sujet de notre seconde question de recherche **QR.2**. Dans cette question nous nous intéressons au savoir enseigné à l'Université vis-à-vis du lycée.

QR.2 : Quelles sont les caractéristiques de la transition secondaire-supérieur en probabilités pour les étudiants non-spécialistes ? (le cas d'étudiants de première année de biologie)

QR.2.1 Quelles sont les caractéristiques de la transition secondaire-supérieur en termes de praxéologies ?

QR.2.2 Quelles sont les caractéristiques de la transition secondaire-supérieur en termes d'activité des étudiants et de pratiques des enseignants ?

QR.2.3 Quelles sont les praxéologies de modélisation et quelles sont les adaptations des connaissances mathématiques ainsi que les sous-activités à la charge des élèves et des étudiants lors d'activités de modélisation ? (Comparaison entre le lycée et l'Université)

Les résultats présentés dans la revue de travaux (chapitre 1 page 13), ont mis en évidence la spécificité des étudiants non-spécialistes et notamment de biologie. C'est en appui sur ces résultats que nous présentons notre troisième question de recherche **QR.3**.

QR.3 : Quelles sont les spécificités des étudiants non-spécialistes dans l'apprentissage des probabilités ? (le cas d'étudiants de première année de biologie)

QR.3.1 : Quelles sont les difficultés rencontrées par les étudiants non-spécialistes dans ces apprentissages ?

QR.3.2 : Quels sont les attentes et les besoins des étudiants non-spécialistes concernant les probabilités ?

Le chapitre 4 page 49 est consacré à la première question. La seconde question de recherche fera l'objet du chapitre 5 page 105 et la troisième question sera abordée dans le chapitre 6 page 255. Enfin, l'ensemble des éléments de réponses à ces questions seront synthétisés dans le chapitre 7 page 301.

Nous venons d'énoncer nos questions de recherche. Dans le chapitre suivant nous décrivons les terrains sur lesquels nous avons mené notre étude ainsi que la méthodologie globale adoptée. Nous détaillerons dans chaque chapitre ensuite la méthodologie spécifique de ce chapitre.

Chapitre 3

Méthodologie générale de la recherche

Nous décrivons, dans cette partie, la méthodologie globale associée à chaque question de recherche précédemment énoncée. Une synthèse des données recueillies se trouve dans le tableau 3.1. Nous faisons ici le choix de ne pas décrire dans le détail la méthodologie suivie mais plutôt de présenter les terrains et les différentes analyses que nous allons mener dans cette thèse. Nous reviendrons cependant, au début de chaque partie, sur la méthodologie spécifique de cette partie.

Nous avons choisi d'étudier plus particulièrement les probabilités conditionnelles et les lois continues qui sont deux thèmes de probabilités à la fois au programme de mathématiques de la classe de terminale scientifique et enseignés en première année de biologie dans l'Université où nous avons l'opportunité d'effectuer nos observations.

Dans une étude précédente (Doukhan & Gueudet, 2019), nous avons mis en évidence par des analyses de manuels scolaires qu'un même contenu de probabilités (variables aléatoires discrètes) peut conduire à des praxéologies très différentes au lycée et à l'Université. En appliquant des méthodes similaires ici (analyse des ressources collectées lors d'observations en classe, telles que les films des séances, les photocopiés de cours et les manuels scolaires), nous souhaitons ainsi comparer les praxéologies du secondaire avec celles de l'Université pour les deux thèmes de probabilités retenus.

3.1 Présentation des terrains

Nos recherches se déroulent en France, dans un lycée en milieu rural et dans une Université de taille moyenne. Des cours de mathématiques sont proposés aux étudiants en biologie dès la première année d'Université, dont un cours entièrement consacré aux probabilités. Ce cours se compose de 14 heures de cours magistraux avec environ 300 étudiants et 14 heures de travaux dirigés, qui sont des séances consacrées à des exercices, en groupes de 30 étudiants.

Dans le secondaire nous nous sommes intéressés à la classe de terminale série scientifique dont proviennent majoritairement les étudiants de licence de biologie.

Nous avons choisi d'étudier la transition secondaire-supérieur à travers les deux seuls thèmes de probabilités, communs aux deux curricula, les probabilités conditionnelles et les lois continues.

3.2 Outils pour l'analyse de la transition institutionnelle

Afin de replacer le sujet d'étude dans un contexte historique et épistémologique et dans l'objectif de construire un MPR, nous avons commencé par une étude historique des concepts et des approches probabilistes (chapitre 4 page 49). Dans un second temps nous avons analysé les documents officiels qui régissent les programmes dans le secondaire, mais aussi des sujets de baccalauréat et un manuel destiné à la classe de terminale scientifique (chapitre 4 page 49).

Le cours de probabilités en biologie à l'Université est assez général, il reprend des notions vues au lycée (probabilités conditionnelles et indépendance, variables aléatoires discrètes, variables aléatoires continues) mais il contient également des chapitres consacrés à de nouveaux concepts (le modèle probabiliste et l'espace probabilisé, l'indépendance des variables aléatoires, les théorèmes limites et leurs applications, l'approximation des lois, etc).

Pour cette recherche, nous nous sommes particulièrement intéressés à ce cours d'Université et avons assisté aux cours magistraux et aux séances de travaux dirigés concernant les deux thèmes de probabilités précédemment cités. Les séances observées (2 cours magistraux et 2 séances de TD) sont analysées dans le chapitre 5.3 page 194. Nous avons également observé une classe de terminale série scientifique pour ces mêmes enseignements de probabilités, ces séances (5 pour les probabilités conditionnelles et 7 pour les lois continues) sont analysées dans le chapitre 5.2 page 108.

3.3 Outils d'analyse de l'activité et des productions des élèves et des étudiants

Dans un troisième temps nous avons souhaité poursuivre ce travail auprès des élèves et des étudiants et notamment accéder à leurs écrits. Pour ce faire nous avons élaboré un test et nous l'avons soumis aux élèves de la classe de terminale scientifique dans laquelle nous avons fait nos observations et à des étudiants en première année de biologie dans l'Université où se déroule nos observations. L'analyse de ce test et des productions des étudiants et élèves est à retrouver dans

3.3. Outils d'analyse de l'activité et des productions des élèves et des étudiants

le chapitre 6, partie 6.1 page 255.

Nous avons également élaboré un questionnaire en ligne à destination des étudiants de première année de biologie, qui leur a été soumis durant le module de probabilités décrit ci-dessus (voir partie 6.2 page 275). Ce questionnaire nous a permis, entre autres, de solliciter certains d'entre eux pour des entretiens plus approfondis (voir partie 6.3 page 283). Enfin, nous avons pu recueillir et analyser des copies d'examens de ces mêmes étudiants. Ces analyses sont détaillées dans la partie 6.4 page 291).

Date	Données recueillies
Octobre - Novembre 2018	Observations de séances en terminale scientifique sur les probabilités conditionnelles
Décembre 2018	Test soumis aux élèves de terminale scientifique et aux étudiants de L1 de biologie
Février - Mars 2019	Observations de séances sur les probabilités conditionnelles et les lois continues en L1 de biologie
Mars 2019	Lancement du questionnaire en ligne à destination des étudiants de L1 de biologie
Mars - Avril - Mai 2019	Observations de séances en terminale scientifique sur les lois continues
Avril - Mai 2019	Entretiens avec 4 étudiants de L1 de biologie
Juin 2019	Recueil des copies d'examens de L1 de biologie

FIGURE 3.1 – Méthodologie générale - données recueillies

Chapitre 4

Enseignement des probabilités au secondaire : analyse du savoir à enseigner

L'objectif de ce chapitre est double. Premièrement, et afin de concevoir à la toute fin de ce travail un MPR-ET, nous apportons un éclairage historique et épistémologique des courants probabilistes et de leurs impacts sur l'enseignement des probabilités (question de recherche QR.1.1). Le second objectif de ce chapitre est d'analyser le savoir à enseigner, pour les deux thèmes de probabilités qui nous intéressent, en classe de terminale scientifique au lycée (questions de recherche QR.1.2 et QR.1.3). Pour ce faire nous avons choisi de nous intéresser dans une seconde partie aux documents officiels qui exposent les programmes, aux sujets de baccalauréat qui constituent une référence pour les attendus en classe de terminale et aux manuels qui, dans le secondaire, servent de référence institutionnelle auprès des enseignants.

4.1 Les différentes approches des probabilités et leur enseignement

Dans le but de disposer d'un recul historique sur les contenus enseignés au secondaire il nous semble nécessaire de commencer par un éclairage historico-épistémologique des notions en jeu. Dans un premier temps nous présentons les différents courants probabilistes et les définitions de probabilités qui en émanent. Nous exposons, dans une seconde partie, des résultats, issus de divers travaux, qui décrivent les impacts de ces approches sur l'enseignement des probabilités.

Avant toute chose nous commençons par présenter la méthodologie de ces analyses.

4.1.1 Méthodologie de l'étude historico-épistémologique

Comme nous l'avons déjà mentionné dans la partie de théorie *Modèle Praxéologique de Référence Étendu et adapté à la Transition secondaire-supérieur* (partie 2.4 page 40), il est important de procéder préalablement à une étude historique et épistémologique des savoirs afin de mener à bien des analyses didactiques. Pour rappel, Dorier (2000b) explique que :

L'étude du savoir mathématique 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.

Notre étude historique s'intéresse à mettre en avant l'émergence des différents courants et approches probabilistes. Dans un second temps nous souhaitons présenter l'impact de ces courants sur l'enseignement des probabilités. Pour cela nous nous sommes appuyés sur les travaux en didactique des mathématiques de Henry (2003, 2010), Batanero et al. (2005), Carranza et Kuzniak (2006), Eichler (2007), Chaput et al. (2011), Parzys (2011), Eichler et Vogel (2014) et la thèse de Derouet (2016) qui présentent des éléments historiques sur les probabilités et l'enseignement de ces contenus. Pour les questions qui concernent l'enseignement par la modélisation nous avons complété ces lectures avec l'article de Prodromou (2014).

Les questions qui guident notre travail sont les suivantes :

- Comment sont apparus historiquement les différents courants et approches probabilistes ?
- Quelles définitions ont émergé de ces différentes approches ?
- Quels sont les impacts de ces approches sur l'enseignement des probabilités ?
- Comment l'enseignement des probabilités a-t-il évolué et suivant quelle(s) approche(s) les probabilités sont-elles enseignées aujourd'hui ?

Nous exposons dans cette partie des résultats identifiés par les travaux sus-mentionnés qui nous semblent répondre aux questions ci-dessus. Cette partie a pour objectif de donner quelques éclairages historiques et en aucun cas nous ne cherchons à décrire avec exhaustivité l'histoire des probabilités à la manière d'un historien.

Nous commençons par présenter, dans la sous-partie suivante, des éléments historiques concernant l'émergence des différents courants probabilistes et par la même occasion nous présentons les définitions de la notion de probabilités qui en résultent.

4.1.2 Éclairages historiques

Il existe deux grands courants philosophiques lorsque l'on s'intéresse aux probabilités : objectiviste et subjectiviste. Mais selon Henry (2003) ces deux courants appartiennent au même paradigme, celui de la description de la réalité et non de sa modélisation. Ainsi il propose de considérer un paradigme supplémentaire, celui de la modélisation. Nous reviendrons sur ce dernier dans la partie 4.1.3 page 54.

En complément des courants objectiviste et subjectiviste il y a différentes approches des probabilités, autrement dit différentes manières de définir et de concevoir la notion de probabilité. Les approches **classique** et **fréquentiste** sont définies suivant le courant objectiviste tandis que l'approche **bayésienne** est définie suivant le courant subjectiviste. Enfin, l'approche **formelle** définit la probabilité suivant l'axiomatique de Kolmogorov. Selon l'approche choisie (classique, fréquentiste, bayésienne, formelle), l'interprétation des situations est différente (Eichler, 2007). Dans cette première sous-partie nous reprenons chronologiquement l'émergence de ces approches.

« *Alors nous dirons que le résultat est un effet du hasard et nous appellerons effets de la fortune tous les effets du hasard qui figurent parmi les choses que l'on peut choisir et qui sont accomplis par des êtres capables de choix.* » (Aristote traduit par Sentis (2005)).

Aristote s'intéressait déjà à l'apparition d'événements qui pourraient être intentionnels, mais qui ne le sont pas, c'est-à-dire aux événements issus d'expériences aléatoires. C'est au XVI^e siècle, avec les jeux de hasard, que les probabilités discrètes commencent à émerger et deviennent des objets de recherche et de questionnements plus avancés.

C'est à Fermat et Pascal que l'on attribue l'invention du calcul des probabilités. Au travers de leur correspondance portant sur le *Problème des partis* (1654), Pascal propose une solution à ce problème en faisant appel aux combinaisons. Le terme de probabilités n'existe pas encore mais, à propos de la répartition du hasard dans les jeux, Pascal explique :

« [...] ainsi, joignant la rigueur des démonstrations de la science à l'incertitude du hasard, et conciliant ces choses en apparence contradictoires, elle peut, tirant son nom des deux, s'arroger à bon droit ce titre stupéfiant : la Géométrie du hasard. »

A cette époque il s'agit surtout d'estimer les gains moyens possibles à toutes sortes de jeux. En 1657, Huygens donne pour la première fois la définition de la notion d'espérance :

« [l'espérance est] le juste prix auquel en joueurs accepterait de céder sa place dans une partie. » (*Du calcul dans les jeux de hasard*).

Au début du XVIII^e un tournant s'opère, Montmort, De Moivre et Bernoulli s'intéressent désormais aux relations, dans les jeux de hasard, entre rapports des chances de gagner et gains espérés.

Dans *Ars Conjectandi* (1713), Jacques Bernoulli est le premier à faire le lien entre probabilité et fréquence de réalisation d'un événement. Il s'intéresse à la détermination de la probabilité par la méthode expérimentale et cherche à répondre à des problèmes dans lesquels on ne peut pas déterminer la probabilité *a priori* d'un événement. Pour cela il propose d'observer un grand nombre de répétitions d'expériences toutes semblables puis de prendre la fréquence observée comme probabilité. Bernoulli expose pour la première fois la loi faible des grands nombres.

Les probabilités sont alors définies de façon fréquentiste, on parle aussi de définition « *a posteriori* ». La probabilité qu'un événement A se réalise est la fréquence limite d'apparition de cet événement, observée expérimentalement durant un nombre très important d'expériences aléatoires.

Approche fréquentiste : l'observation de la stabilisation des fréquences d'un événement permet l'estimation expérimentale d'une probabilité.

En 1733, Buffon introduit le calcul de probabilités dans un contexte géométrique. Il est le premier à avoir recours aux probabilités continues et plus particulièrement à la loi uniforme. Il s'agit du premier exemple de probabilités à densité.

Dans *The Doctrine of Chances* (1756), De Moivre propose d'améliorer le théorème de Bernoulli. C'est la première fois qu'apparaît la loi normale qui est donnée ici comme limite d'une loi binomiale.

Durant cette période le concept de probabilités possède deux faces : l'une est subjective, portée sur le degré de croyances, l'autre est objective, portée sur une fréquence stable.

Les probabilités conditionnelles apparaissent dans les écrits de Bayes, publiés à titre posthume dans *An essay towards solving a problem in doctrine of chance* en 1763. Il souhaite y « analyser les causes supposées inconnues à partir d'événements observés ». A cette occasion, apparaît pour la première fois, la célèbre formule suivante :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B \text{ si } A)$$

Une nouvelle approche du concept de probabilités découle de ces écrits. Il s'agit de l'approche bayésienne dans laquelle on tient compte du point de vue de l'observateur.

Approche bayésienne : la probabilité correspond au degré de crédibilité attribuable à une hypothèse dont on ne peut établir si elle est vraie ou fausse. La probabilité est alors une mesure du degré de certitude (ou de croyance) d'un événement. Le degré de certitude correspond au niveau des connaissances sur l'événement. Ces connaissances dépendent de la quantité d'informations disponibles sur l'évènement qui sont représentées sous la forme d'une probabilité conditionnelle. (Carranza & Kuzniak, 2006)

Selon Eichler et Vogel (2014) le point de vue subjectiviste sur les probabilités est étroitement lié à la formule de Bayes qui permet de revoir l'estimation *a priori* de la probabilité en traitant de nouvelles informations et d'estimer une nouvelle probabilité *a posteriori*.

De la théorie des erreurs émerge les premières lois densité et la notion de fonction de densité. C'est Simpson (1710-1761) qui en travaillant sur la loi uniforme discrète introduit pour la première fois la notion de fonction de densité à l'aide de la fonction génératrice (Derouet, 2016).

En 1825, dans le premier principe de *l'Essai philosophique sur les probabilités*, Laplace donne une nouvelle définition des probabilités. La définition Laplacienne, que l'on appelle aussi définition « *a priori* » ou définition classique d'une probabilité est exprimée comme suit :

« Le premier de ces principes est la définition même de la probabilité qui, comme on l'a vu, est le rapport du nombre des cas favorables à celui de tous les cas possibles. »

Approche classique : la probabilité de l'évènement A est obtenue par le rapport suivant :

$$\mathbb{P}(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

Il s'agit de la formule de Laplace.

A la suite des travaux de Borel sur la théorie de la mesure et des travaux de Lebesgues sur la théorie de l'intégration, Kolmogorov introduit au XX^e siècle les fondements de la théorie moderne des probabilités. Cette nouvelle théorie propose une axiomatique du calcul des probabilités.

Approche formelle : axiomatique de Kolmogorov dans laquelle est défini un espace probabilisé (Ω, \mathcal{A}, P) .

4.1.3 Pluralité des approches dans l'enseignement des probabilités

En 1780 le premier programme officiel de mathématiques contenant des éléments sur les probabilités est proposé au lycée par Condorcet. C'est en 1788 qu'elles sont finalement enseignées pour la première fois au lycée en France.

Approche classique

L'enseignement des probabilités s'appuie exclusivement sur l'approche classique jusqu'en 1980. Cette définition a des limites comme de supposer l'équiprobabilité des cas possibles. L'enseignement des probabilités est, à cette époque, principalement basé sur des problèmes de combinatoire et de dénombrement, source de nombreuses difficultés chez les élèves Henry (2010). Enseignée seule, l'approche classique réduit les probabilités à des problèmes de comptage d'événements et masque les multiples applications des probabilités aux situations de la vie réelle.

Approche fréquentiste

C'est seulement à partir de 1990 qu'un nouveau choix épistémologique est fait pour l'introduction des probabilités au lycée. L'approche fréquentiste de Bernoulli est introduite dans les programmes. Cette arrivée est la conséquence de l'évolution de l'enseignement des mathématiques vers des activités plus expérimentales, mais aussi du développement de l'analyse de données dans les sciences en général (Chaput et al., 2011).

Comme le souligne Parzysz (2011), l'approche fréquentiste permet de mieux rendre compte des phénomènes réels et permet une meilleure compréhension des phénomènes aléatoires. L'inconvénient de cette approche est qu'elle ne laisse place qu'aux expériences reproductibles où les épreuves sont indépendantes.

En revanche, un avantage de cette approche est qu'elle permettrait de donner une définition plus compréhensible de la notion de probabilités conditionnelle. En effet, selon Parzysz (2011) la probabilité conditionnelle pourrait être introduite en traitant d'abord des problèmes de fréquences, puis en expliquant le passage de la probabilité conditionnelle à un quotient de deux probabilités.

Malgré tout, enseignée seule, l'approche fréquentiste génère des difficultés chez les élèves, comme la confusion entre l'observation de la réalité (fréquence stabilisée) et la connaissance théorique (probabilité). Selon Chaput et al. (2011) la définition d'une probabilité comme fréquence relative stabilisée pose de ce fait des problèmes épistémologiques. Cette confusion entre modèle et réalité rend le processus de modélisation difficile à appréhender pour les étudiants (Batanero et al., 2005).

Approche par la modélisation

Le point de vue de la modélisation apparaît dans les programmes de lycée dans les années 2000. Ces nouveaux programmes suggèrent l'enseignement des probabilités par la modélisation.

Selon Prodromou (2014) l'approche par la modélisation permet de relier les approches classique et fréquentiste, mais cette nouvelle approche permet surtout de trancher le débat (philosophique) entre subjectivistes et objectivistes par le choix d'un modèle probabiliste qui doit être le plus adéquat possible (Henry, 2010).

Dans les documents d'accompagnement des programmes des années 2000 on peut lire « *une loi de probabilités est un objet mathématique ayant les mêmes propriétés qu'une distribution de fréquences* ». Henry (2003) estime que le terme d'« *objet mathématique* » employé ici signifie que la probabilité ne doit pas être considérée comme une grandeur de la réalité mais qu'elle fait partie des notions qui modélisent la réalité. Ce point de vue est, toujours d'après Henry (2003), en rupture avec celui de l'ancien programme, qui proposait une définition fréquentiste de la notion de probabilité, où il s'agissait d'une description de la réalité et non de modélisation.

Approche bayésienne

Carranza et Kuzniak (2006) observent que dans les programmes du lycée des années 2000 les probabilités sont introduites à partir de la stabilisation des fréquences ou à partir de situations d'équiprobabilité. Aucune référence n'est faite à l'approche bayésienne ni au degré de certitude. Pourtant, on trouve très peu d'exercices dans les manuels relevant de l'approche fréquentiste alors qu'il y a de nombreux exercices autour de la formule de Bayes. En fait, selon ces auteurs, il semblerait que le modèle bayésien n'apparaisse pas explicitement et qu'il se réduise uniquement à la formule de Bayes. Finalement, l'approche subjectiviste est quasiment inexistante dans l'enseignement des probabilités (Chaput et al., 2011).

Le modèle Bayésien s'appuie sur une mesure mathématique du degré de certitude qui permet une sorte de « mise-à-jour » de la mesure de probabilités lors de l'arrivée de nouvelles informations. La formule de Bayes joue un rôle essentiel car elle permet de réévaluer la crédibilité de chacune des hypothèses possibles en s'appuyant sur les données déjà obtenues. Pourtant on ne trouve aucune référence au degré de certitude dans l'enseignement secondaire, celui-ci semble réservé à l'enseignement supérieur. (Carranza & Kuzniak, 2006).

Pour conclure cette partie, Henry (2010) explique que « *les différents points de vue de la probabilité sont complémentaires, le fait de réduire l'enseignement à une seule approche peut expliquer certaines difficultés d'apprentissages chez les élèves* ».

Selon lui pour comprendre le lien entre valeur issue d'un calcul *a priori* (si une condition d'équiprobabilité est possible) et estimation *a posteriori* d'une valeur par l'observation de fréquences expérimentales, une très bonne compréhension de la loi des grands nombres est nécessaire. Or les élèves de lycée ne connaissent pas la loi des grands nombres, qui ne fait pas partie des contenus des programmes de lycée.

Pour résumer

Dans cette partie nous avons présenté comment historiquement et épistémologiquement les différentes approches des probabilités sont apparues dans l'enseignement des mathématiques. Les travaux présentés précédemment ont identifié une pluralité d'approches, celle-ci est intrinsèque à la notion de probabilité. C'est le contexte des exercices qui détermine le sens à utiliser.

D'après les auteurs précédemment mentionnés il est important que les différentes approches des probabilités soient enseignées ensemble et surtout articulées entre elles. Cela permettrait aux élèves de donner du sens à l'usage des probabilités dans le monde réel mais aussi de réduire les difficultés des élèves dans ces apprentissages. Dans l'analyse des programmes officiels, du manuel et des sujets de baccalauréat que nous présentons nous sommes attentifs à ces aspects.

4.2 Programmes officiels

Nous poursuivons ce travail en analysant des textes officiels régissant les programmes du lycée pour les thèmes qui nous intéressent, à savoir : les probabilités conditionnelles et les lois continues. Dans une section suivante (section 4.3) nous proposons une analyse d'un manuel de lycée et de sujets de baccalauréat, en regard des résultats de cette partie.

4.2.1 Méthodologie de l'analyse des textes officiels

Dans cette sous-partie nous présentons les textes officiels que nous avons retenus pour ce travail ainsi que la méthodologie d'analyse de ces derniers.

Nous faisons référence, tout au long de ce travail, aux programmes de 2011 bien que, suite à la réforme du baccalauréat de 2019/2020, de nouveaux programmes pour la classe de première et de terminale soient depuis mis en place au lycée en France.

Les questions qui guident notre travail sont les suivantes :

- Quels sont les attendus institutionnels pour les deux thèmes de probabilités qui nous intéressent ?
- Quelles sont les recommandations des textes officiels concernant la modélisation probabiliste ?

Le document principal qui guide nos analyses (Ministère de l'Education Nationale (MEN), 2011) présente le programme, en vigueur jusqu'en 2020, de l'enseignement spécifique et de spécialité de mathématiques en classe terminale de la série scientifique.

Ce document est découpé en plusieurs parties : Recommandations générales, Analyse, Géométrie, Probabilités et Statistiques, Algorithmiques et enfin Notations et raisonnement mathématiques. Pour les parties consacrées à l'analyse, la géométrie, les probabilités et les statistiques, le programme présente, sous forme d'un tableau, les **contenus**, les **capacités attendues** et des **commentaires**. Voici quelques explications que l'on peut lire au début de ce document :

Le programme fixe les objectifs à atteindre en termes de capacités. Il est conçu pour favoriser une acquisition progressive des notions et leur pérennisation. Son plan n'indique pas la progression à suivre.

Les capacités attendues indiquent un niveau minimal de maîtrise des contenus en fin de cycle terminal. La formation ne s'y limite pas.

Plusieurs démonstrations, ayant valeur de modèle, sont repérées par le symbole [...]. Certaines sont exigibles et correspondent à des capacités attendues.

Les commentaires notés [...], distinguent des thèmes pouvant se prêter à des ouvertures interdisciplinaires, en concertation avec les professeurs d'autres disciplines scientifiques. (Ministère de l'Education Nationale (MEN), 2011)

Nous complétons l'analyse du programme officiel de la classe de terminale scientifique avec celle des documents suivants d'accompagnement des programmes, à destination des enseignants et proposés par le Ministère de l'Education Nationale :

- un document, *Ressources pour le lycée général et technologique - Mathématiques*, présentant les compétences en mathématiques qui doivent être développées par la formation mathématique au lycée (Ministère de l'Education Nationale de la Jeunesse et de la Vie Associative (MENJVA) & DGESCO, 2013).
- un document, *Ressources en probabilités et statistique pour la classe terminale générale et technologique*, proposant des supports et des compléments notamment pour les nouveaux contenus de probabilités qui viennent enrichir les programmes à partir de 2011 (Ministère de l'Education Nationale de la Jeunesse et de la Vie Associative (MENJVA) & DGESCO, 2012).

En ce qui concerne la méthodologie d'analyse de ces textes, nous proposons d'analyser ces différents documents en repérant ce qui concerne les probabilités

conditionnelles, les lois continues ou la modélisation probabiliste. Nous nous attachons à repérer les capacités attendues évoquées dans ces textes. Nous tentons ici d'anticiper la comparaison qui sera faite entre les textes officiels et les manuels.

Ces documents nous informent sur le point de vue de l'institution concernant les thèmes qui nous intéressent et les façons de mettre en évidence la modélisation probabiliste.

Nous présentons dans une première sous-partie des éléments généraux, issus des textes officiels, concernant les probabilités. Nous approfondissons dans une seconde sous-partie, ce qui concerne le thème des probabilités conditionnelles. Nous nous intéressons ensuite aux éléments du programme concernant les lois continues. Enfin, dans une dernière sous-partie nous présentons ce que les documents officiels mentionnent concernant la modélisation dans le cadre des probabilités pour les classes de première et terminale, série scientifique.

4.2.2 Éléments généraux issus des textes officiels concernant l'enseignement des probabilités

Dans le programme officiel de la classe de terminale scientifique, la partie qui nous intéresse est celle consacrée aux probabilités et aux statistiques. Pour ces deux domaines des mathématiques, une liste de contenus et les capacités attendues correspondantes sont présentées. Aucune partie n'est consacrée à la modélisation dans ce document, pourtant elle a une place importante dans les différents textes et programmes officiels, nous y reviendrons dans la partie 4.2.5, page 62.

La partie "*Probabilités et Statistiques*" du programme débute par une courte introduction qui motive les notions à enseigner :

On approfondit le travail en probabilités et statistique mené les années précédentes.

Afin de traiter les champs de problèmes associés aux données continues, on introduit les lois de probabilités à densité. Le programme en propose quelques exemples et, en particulier, la loi normale qui permet notamment d'initier les élèves à la statistique inférentielle par la détermination d'un intervalle de confiance pour une proportion à un niveau de confiance de 95 %.

Cette partie se prête particulièrement à l'étude de problèmes issus d'autres disciplines. Le recours aux représentations graphiques et aux simulations est indispensable. (Ministère de l'Éducation Nationale (MEN), 2011)

Dans cette partie, les **contenus** (comme ils sont nommés dans le document) sont regroupés en quatre thèmes :

- Conditionnement et indépendance
- Notion de loi à densité à partir d'exemples

- Intervalle de fluctuation
- Estimation

On peut lire en guise d'introduction au document *Ressources en probabilités et statistique pour la classe terminale générale et technologique* ce qui suit :

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. Ces notions 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.

Ce paragraphe nous donne donc déjà des informations sur les problématiques qui guident notre travail, oui les contenus enseignés sont assez similaires entre la classe de terminale scientifique et le supérieur, cependant "*le point de vue adopté est assez différent*". Il s'agira donc dans la suite d'en savoir davantage sur cette distinction.

4.2.3 Conditionnement et indépendance

Les **contenus** présentés dans le programme officiel de la classe de terminale scientifique, pour le thème "*conditionnement et indépendance*", sont les suivants :

- * Conditionnement par un évènement de probabilités non nulle. Notation $P_A(B)$
- * Indépendance de deux évènements

Les **capacités attendues** officiellement, chez les élèves de terminale série scientifique, pour le thème "*conditionnement et indépendance*", sont, d'après le Ministère de l'Education Nationale (MEN) (2011) :

- ♣ **Construire un arbre pondéré en lien avec une situation donnée** (notée par la suite **C-ConsArb**)
Il est détaillé à côté : "*on représente une situation à l'aide d'un arbre pondéré ou d'un tableau de probabilités. On énonce et on justifie les règles de construction et d'utilisation des arbres pondérés.*"
- ♣ **Exploiter la lecture d'un arbre pondéré pour déterminer des probabilités** (notée **C-ExplArb**)
Il est précisé : "*un arbre pondéré correctement construit constitue une preuve.*"
- ♣ **Calculer la probabilité d'un évènement connaissant ses probabilités conditionnelles relatives à une partition de l'univers** (notée **C-CalcFPT**)
Il est précisé : "*le vocabulaire lié à la formule des probabilités totales n'est pas un attendu du programme, c'est seulement la mise en oeuvre de cette formule qui doit être maîtrisée.*"
- ♣ **Démontrer que si deux évènements sont indépendants alors il en est de même pour leurs complémentaires** (notée **C-DemInd**)
Il s'agit d'une démonstration exigible.

On remarquera qu'aucune capacité, autre que la démonstration de l'indépendance des complémentaires, n'est attendue pour le contenu *indépendance de deux évènements*, d'après le texte du programme.

A la suite du tableau présentant les contenus, capacités attendues et les commentaires se trouve cette recommandation officielle : "*Cette partie du programme se prête particulièrement à l'étude de situations concrètes*". Ainsi nous nous attendons à observer, dans l'analyse du manuel présentée dans la partie 4.3, page 63, de nombreux exercices et exemples faisant référence à des situations issues de la vie réelle.

4.2.4 Lois à densité

Les **contenus** au programme de la classe de terminale série scientifique, pour le thème "*notion de loi à densité à partir d'exemples*", sont les suivants :

- * Loi à densité sur un intervalle
- * Loi uniforme sur $[a, b]$
- * Espérance d'une variable aléatoire suivant une loi uniforme
- * Lois exponentielles
- * Espérance d'une variable aléatoire suivant une loi exponentielle
- * Loi normale centrée réduite $\mathcal{N}(0, 1)$
- * Théorème de Moivre-Laplace
- * Loi normale $\mathcal{N}(\mu, \sigma^2)$ d'espérance μ et d'écart-type σ^2

La loi normale et le théorème de Moivre-Laplace sont les nouveautés du programme de 2011 vis-à-vis de l'ancien programme datant de 2001. A ce sujet, se trouve dans le document *Ressources en probabilités et statistique pour la classe terminale générale et technologique* (Ministère de l'Éducation Nationale de la Jeunesse et de la Vie Associative (MENJVA) & DGESCO, 2012) l'explication suivante :

La loi normale est introduite en terminale S comme loi-limite d'une suite de variables aléatoires grâce au théorème de Moivre-Laplace. Bien qu'admis, ce théorème se visualise facilement grâce à des animations avec un logiciel de géométrie dynamique ou sur tableur.

Les **capacités attendues** chez les élèves de terminale série scientifique, pour le thème "*notion de loi à densité à partir d'exemples*", sont, d'après le programme officiel Ministère de l'Éducation Nationale (MEN) (2011) :

♣ **Connaître la fonction de densité de la loi uniforme sur $[a, b]$** (notée par la suite **C-ConUni**)

Il est précisé : "*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 $E(X) = \int_a^b tf(t)dt$. On note que cette définition constitue un prolongement dans le cadre continu de l'espérance d'une variable aléatoire discrète."

♣ **Calculer une probabilité dans le cadre d'une loi exponentielle** (notée **C-CalcExp**)

Il est précisé : "la variable aléatoire T suivant une loi exponentielle vérifie la propriété de durée sans vieillissement", cette propriété est indiquée comme étant une démonstration exigible, elle correspond donc également à une capacité attendue des programmes.

♣ **Démontrer que l'espérance d'une variable aléatoire suivant une loi exponentielle de paramètre λ est $\frac{1}{\lambda}$** (notée **C-DemEspExp**)

Il est précisé : "l'espérance est définie comme la limite quand x tend vers $+\infty$ de $\int_0^x tf(t)dt$, où f est la fonction de densité de la loi exponentielle considérée". Ce qui attire notre attention, au vu des résultats présentés dans la revue de travaux concernant notamment l'intérêt de l'étude situations concrètes en probabilités, c'est surtout la remarque qui suit : "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."

♣ **Connaître la fonction de densité de la loi normale $\mathcal{N}(0, 1)$ et sa représentation graphique** (notée **C-ConLN**)

Il est indiqué qu'on utilisera le théorème de Moivre-Laplace pour introduire la loi normale $\mathcal{N}(0, 1)$, il est détaillé comment retrouver la valeur de l'espérance d'une telle loi en sommant la valeur de la limite quand x tend vers $-\infty$ de $\int_x^0 tf(t)dt$ et la valeur de la limite quand x tend vers $+\infty$ de $\int_0^x tf(t)dt$. Il est admis en revanche que la variance, définie par $E((X - E(X))^2)$, vaut 1.

♣ **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 $\mathcal{N}(0, 1)$** (notée **C-DemUa**)

Il s'agit d'une démonstration exigible.

♣ **Connaître les valeurs approchées de $u_{0,05} = 1,96$ et $u_{0,01} = 2,58$** (notée **C-ConUa**).

♣ **Utiliser une calculatrice ou un tableur pour calculer une probabilité dans le cadre d'une loi normale $\mathcal{N}(\mu, \sigma^2)$** (notée **C-UtilCalcu**)

Il est rappelé comment centrer-réduire une telle variable aléatoire : "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."

♣ **Connaître une valeur approchée de la probabilité des événements $\{X \in [\mu - \sigma, \mu + \sigma]\}$, $\{X \in [\mu - 2\sigma, \mu + 2\sigma]\}$ et $\{X \in [\mu - 3\sigma, \mu + 3\sigma]\}$, lorsque X suit une loi normale $\mathcal{N}(\mu, \sigma^2)$** (notée **C-ConVal**)

Il est précisé : "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 remarquera qu'aucune capacité n'est attendue d'après le texte du programme, pour les lois à densité quelconque.

A la suite du tableau présentant les capacités attendues et les commentaires associés se trouve cette recommandation générale : "On illustre ces nouvelles notions par des exemples issus des autres disciplines." On peut penser à la loi exponentielle qui modélise des phénomènes de durée de vie sans vieillissement pour la physique.

4.2.5 Modélisation dans les textes officiels

Dans le programme de mathématiques de la classe de terminale (Ministère de l'Éducation Nationale (MEN), 2011), il n'y a ni de partie dédiée entièrement à l'activité de modélisation, ni mention du terme "modélisation" pour la partie de probabilités et statistiques. On trouve tout de même, dans les différents documents officiels, des recommandations et des indications relatives à l'activité de modélisation. C'est l'objet de cette sous-partie.

L'importance de la modélisation dans l'enseignement des probabilités apparaît dans le programme officiel de mathématiques pour la classe de première, pour les thèmes de probabilités au programme (à savoir : variable aléatoire discrète et loi de probabilité, modèle de la répétition d'expériences identiques et indépendantes, épreuve de Bernoulli et loi binomiale) :

La notion de loi de probabilités d'une variable aléatoire permet de **modéliser** des situations aléatoires, d'en proposer un traitement probabiliste et de justifier certains faits observés expérimentalement en classe de seconde. L'utilisation des arbres pondérés est développée pour **modéliser** la répétition d'expériences identiques et indépendantes. Elle est restreinte à ce cadre afin d'éviter toute confusion avec des situations relevant des probabilités conditionnelles. (Ministère de l'Éducation Nationale (MEN), 2010).

Dans le programme officiel de mathématiques pour la classe de terminale (Ministère de l'Éducation Nationale (MEN), 2011), les seules traces que nous avons trouvées, de l'activité de modélisation pour les thèmes de probabilités qui nous concernent, pour la partie "*Probabilités et Statistique*", sont les suivantes :

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.

Même si une situation concrète peut très bien ne pas offrir d'activité de modélisation, nous n'avons rien trouvé de plus évocateur dans ce texte officiel. En revanche, dans le document d'accompagnement des programmes (Ministère de l'Éducation Nationale de la Jeunesse et de la Vie Associative (MENJVA) &

DGESCO, 2013) qui liste et détaille les compétences à développer en mathématiques au cours du lycée, se trouve un paragraphe qui décrit la compétence "*modéliser*". La définition donnée par l'institution pour cette compétence est la suivante :

Modéliser : traduire en langage mathématique une situation réelle (à l'aide d'équations, de suites, de fonctions, de configurations géométriques, de graphes, de lois de probabilité, d'outils statistiques ...). Utiliser, comprendre, élaborer une simulation numérique ou géométrique prenant appui sur la modélisation et utilisant un logiciel. Valider ou invalider un modèle.

Par ailleurs, dans le document d'accompagnement du Ministère de l'Education Nationale de la Jeunesse et de la Vie Associative (MENJVA) et DGESCO (2012) qui présente les ressources en probabilités et statistique pour la classe de terminale, se trouvent de nombreux exemples concernant les lois à densité, faisant référence à une activité de modélisation.

Nous serons donc attentifs aux activités de modélisation présentées dans les manuels (voir partie 4.3) et proposées par l'enseignante lors des séances observées en classe (voir partie 5.2.3, page 190).

Pour résumer

Le programme actuel de mathématiques pour la classe de terminale scientifique présente l'enseignement des probabilités à travers l'approche classique et l'approche fréquentiste. En effet, le théorème de Moivre-Laplace, bien que sa démonstration soit admise, fait le lien entre fréquence et probabilité. Selon les résultats présentés dans la première partie 4.1.3, page 54, combiner ces approches est une bonne chose du point de vue de l'apprentissage.

Par ailleurs, nous venons de répertorier les compétences et capacités attendues dans les programmes officiels pour les deux thèmes de probabilités qui nous intéressent. Les textes des programmes font également référence à la modélisation, à des situations concrètes ou à des problèmes issus de d'autres disciplines. Nous allons compléter ces résultats par l'analyse suivante d'un manuel destiné à la classe de terminale scientifique.

4.3 Analyse d'un manuel et de sujets de baccalauréat

Nous présentons dans cette partie l'analyse de deux chapitres d'un manuel de mathématiques pour la classe de terminale, série scientifique. Nous complétons cette analyse avec un travail sur les sujets de baccalauréat de terminale scientifique des années 2018, 2019 et 2020 issus de différents centres. L'étude de ce manuel complétée par l'étude de sujets de baccalauréat devrait nous donner une

idée plus précise des tâches, des adaptations de connaissances qu'elles induisent de la part des élèves et des praxéologies proposées et attendues par l'institution.

Dans un premier temps, nous commençons par présenter les modalités d'analyse que nous avons suivies.

4.3.1 Méthodologie de l'étude de manuel et de sujets de baccalauréat

Considérant que dans le secondaire, au delà de l'institution et des programmes, les manuels reflètent les pratiques enseignantes, nous avons souhaité compléter l'analyse précédente des programmes officiels par l'analyse d'un manuel de mathématiques pour la classe de terminale scientifique. En effet, des enquêtes ont montré qu'en France les enseignants de mathématiques au secondaire utilisent très largement le manuel (Rapport IGAENR-IGEN (Séré & Bassy, 2010)).

Le manuel que nous avons retenu pour cette étude est le manuel de terminale scientifique, collection Indice aux éditions Bordas (Poncy et al., 2012). Il s'agit du manuel utilisé par l'enseignante de terminale scientifique de la classe dans laquelle nous avons fait des observations. Ces observations seront présentées dans le chapitre suivant (section 5.2, page 108).

Nous nous intéressons au chapitre sur les probabilités conditionnelles et au chapitre sur les lois continues de ce manuel.

Nous avons conscience que l'étude d'un seul manuel peut sembler insuffisante. Mais dans l'objectif de l'élaboration de notre MPR-ET et afin d'obtenir un ensemble de praxéologies représentatives de ce qui est attendu à la fin du lycée, plutôt que d'étudier plusieurs manuels de terminale scientifique, nous avons préféré compléter cette analyse par celle de sujets de baccalauréat et leurs corrigés fournis par l'APMEP, présentant des exercices sur les thèmes de probabilités conditionnelles ou des lois continues.

Les questions qui guident notre travail sont les suivantes :

- Retrouve-t-on les capacités attendues du programme dans le manuel ? Si oui, comment sont-elles mises en application ?
- Quelles sont les types de tâches proposés par le manuel et présents dans les sujets de baccalauréat, pour les deux thèmes de probabilités qui nous intéressent ?
- Pour les types de tâches relevés, quelles sont les techniques présentées dans le manuel et les sujets de baccalauréat ?
- Quelles adaptations sont à la charge de l'élève selon les exercices et exemples ?
- Quelle est la place de la modélisation dans ce manuel ?

Pour l'analyse du manuel nous procédons en plusieurs étapes. Tout d'abord nous proposons une analyse structurelle des deux chapitres. Nous présentons ensuite les registres de représentations (Duval, 1993) que nous avons relevés dans chacun des chapitres, la façon dont ils sont explicités ou non et si des changements de registres sont demandés. Nous nous sommes également intéressés aux contextes des exemples et exercices présents au sein des chapitres étudiés. Enfin nous présentons une analyse praxéologique des chapitres concernés en nous appuyant sur la méthodologie présentée par Chaachoua (2014), il s'agit de faire l'inventaire des types de tâches et techniques en répondant aux questions : que doit faire l'élève et comment doit-il le faire ?

Nous complétons cette analyse praxéologique avec l'analyse de sujets de baccalauréat qui donnent à voir les praxéologies attendues par l'institution. L'analyse des corrigés provenant de l'APMEP permettra de renseigner à la fois les techniques attendues pour les types de tâches complémentaires que nous aurons recensés mais également les adaptations de connaissances à la charge des élèves.

Dans la première sous-partie nous présentons l'analyse du chapitre "*Probabilités conditionnelles*" de ce manuel complétée par l'analyse de sujets de baccalauréat pour ce thème. Puis dans la seconde sous-partie nous nous intéressons au chapitre "*Lois de probabilités à densité*" et, de même, nous complétons cette analyse avec une étude d'exercices issus de sujets de baccalauréat sur le thème des lois continues.

4.3.2 Probabilités conditionnelles

Dans le manuel étudié, le chapitre de probabilités conditionnelles est le dixième chapitre du manuel. A la première page de ce chapitre, un petit encadré présente les "*notions du chapitre*" qui sont :

- * probabilités conditionnelles
- * arbres pondérés et probabilités totales
- * indépendance de deux évènements

Un second encadré motive par un exemple les probabilités conditionnelles et leur utilisation dans une situation de la vie réelle :

Un évènement, comme une catastrophe naturelle, peut être la conséquence de plusieurs causes. La notion de probabilités conditionnelle est un outil qui permet de mesurer l'influence de chacune de ces causes et de chiffrer le risque qu'elles ont de contribuer à une catastrophe. Cette notion est à l'origine d'une formule mathématique utilisée dans la gestion des risques. (Poncy et al., 2012)

Les *capacités attendues* du programme officiel n'apparaissent pas explicitement dans ce chapitre, elles sont ici présentées sous la forme de *savoir-faire* (notés "**s-f**" dans la liste qui suit). Voici les savoir-faire que l'on trouve dans le manuel concernant le chapitre de probabilités conditionnelles :

- ♣ s-f 1 : calculer une probabilité conditionnelle
- ♣ s-f 2 : calculer la probabilité d'une intersection
- ♣ s-f 3 : exploiter la lecture d'un arbre pondéré (on retrouve ici C-ExplArb)
- ♣ s-f 4 : construire et utiliser un arbre (on retrouve ici C-ConsArb)
- ♣ s-f 5 : étudier l'indépendance de deux évènements
- ♣ s-f 6 : utiliser l'indépendance de deux évènements

Nous considérons que ces savoir-faire sont présentés sous forme de types de tâches. Ils sont explicités dans les pages dédiées au cours, il s'agit, pour chacun des savoir-faire, d'un exercice à plusieurs questions, avec des solutions commentées et des éléments de méthode. Nous y reviendrons dans l'analyse praxéologique.

La correspondance entre capacités attendues dans les programmes officiels et savoir-faire présents dans le manuel n'est pas totale. Dans la figure 4.1 nous présentons un schéma qui fixe les relations entre capacités attendues par les programmes officiels et savoir-faire présents dans le manuel.

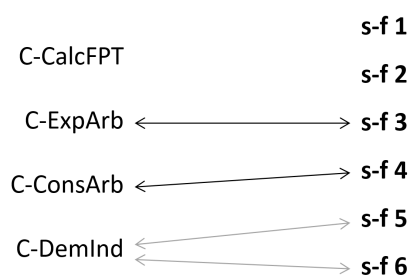


FIGURE 4.1 – correspondance entre *capacités attendues* et *savoir-faire*

Les flèches noires représentent une correspondance parfaite (s-f3 et C-ExplArb, s-f4 et C-ConsArb), tandis que les flèches grises représentent une correspondance partielle (la capacité apparaît dans le savoir-faire mais le savoir-faire du manuel, qui est sous la forme d'un exercice à plusieurs questions, propose plus que cela). Si la capacité **C-CalcFPT** n'est liée à aucun des savoir-faire du manuel c'est parce qu'elle ne fait pas l'objet d'un savoir-faire à part entière. En revanche, cette capacité se retrouve au centre de nombreux exercices et fera l'objet d'une praxéologie que nous présenterons plus loin.

Structure du chapitre

Nous présentons, dans ce qui suit, la structure détaillée du chapitre. Le chapitre commence par trois exercices permettant de "*réactiver les savoirs*".

- Le premier exercice est un vrai-faux qui permet de travailler les données en pourcentages, les données en fréquences naturelles et le passage de l'une à l'autre

- Le second exercice présente une situation d'équiprobabilité avec des calculs de probabilités où il est demandé de rappeler les définitions des événements \bar{A} , $A \cup B$ et $A \cap B$
- Le troisième exercice présente un arbre pondéré dans le cas de la répétition d'un tirage avec remise. Rappelons que l'arbre pondéré a déjà été vu en classe de première, filière scientifique, dans le cadre de la loi binomiale et de la répétition d'expériences identiques et indépendantes

Deux pages proposent ensuite des activités introductives. Le mot "*activité*" dans le sens des manuels n'est pas le même que celui de la Théorie de l'Activité. Les objectifs annoncés de ces activités introductives sont les suivants :

- ↔ introduire la notion de probabilités conditionnelle
- ↔ découvrir les arbres pondérés
- ↔ découvrir la formule des probabilités totales
- ↔ introduire la notion d'évènements indépendants

Les six pages suivantes sont une alternance de pages dédiées au cours et de pages présentant les différents "*savoir-faire*". Associée à chacune des pages de cours, une page fait face en présentant les "*savoir-faire*" liés au contenu de cours présentés. Le plan du cours et des "*savoir-faire*" du chapitre de ce manuel est le suivant :

1. Probabilités conditionnelles
 - Probabilité conditionnelle de B sachant A
 - Utilisation de tableaux de probabilités
 - ♣ s-f 1 : calculer une probabilité conditionnelle
 - ♣ s-f 2 : calculer la probabilité d'une intersection
2. Arbres pondérés et probabilités totales
 - Probabilité conditionnelle et arbre pondéré
 - Formule des probabilités totales
 - ♣ s-f 3 : exploiter la lecture d'un arbre pondéré
 - ♣ s-f 4 : construire et utiliser un arbre
3. Indépendance de deux évènements
 - Deux évènements indépendants
 - Indépendance et évènements contraires
 - ♣ s-f 5 : étudier l'indépendance de deux évènements
 - ♣ s-f 6 : utiliser l'indépendance de deux évènements

Le chapitre commence par l'explication des notations \mathbb{P} et Ω , s'ensuit la définition de la probabilité conditionnelle de A sachant B puis un premier exemple. Une indication précise que $P_A(B)$ se note parfois $P(A|B)$, ce qui est très intéressant car, comme nous allons le voir dans le chapitre 5, la première notation est celle utilisée par l'enseignante de terminale scientifique et la seconde est celle utilisée par l'enseignant de première année de biologie.

Une propriété suit cette définition, il s'agit du calcul de $P_B(A)$ dans le cas d'une situation d'équiprobabilité, c'est-à-dire en utilisant le premier principe de Laplace du rapport du nombre de cas favorables à celui de tous les cas possibles. Un second onglet présente deux propriétés $P(A \cap B) = P(A) \times P_A(B)$ et $P(A \cap B) = P(B) \times P_B(A)$ et leur démonstration. La seconde sous-partie est consacrée à l'utilisation d'un tableau à doubles entrées pour le cas des probabilités conditionnelles.

La seconde partie du chapitre est consacrée à la modélisation par des arbres de probabilités. Leurs utilisations ainsi que leurs règles de construction y sont détaillées. Une seconde sous-partie est ensuite consacrée à la formule des probabilités totales. La définition de partition de l'univers n'est pas explicitement au programme de terminale mais le terme est quand même employé et défini ici. La formule des probabilités totales est présentée dans le cas d'un nombre fini n d'événements, puis dans le cas plus élémentaire de deux événements.

La dernière partie de ce chapitre concerne les événements indépendants. On y trouve la définition de deux événements indépendants suivie d'une proposition listant quelques propriétés. Un exemple utilisant un tableau de probabilités est donné. La seconde sous-partie caractérise l'indépendance des événements complémentaires lorsque les événements sont indépendants.

A la suite de cela vingt pages proposent des exercices et des activités. Le mot "*activité*" dans le sens des manuels n'est pas le même que celui de la Théorie de l'Activité. Parmi ces pages on peut trouver (les termes en italiques sont issus du manuel) :

- 84 exercices dont des exercices résolus, des vrai-faux, des exercices chronométrés, des exercices *pour démarrer*, des exercices *pour s'entraîner*, des exercices *pour faire le point*, des exercices *pour aller plus loin*, d'autres pour se préparer à l'oral du baccalauréat et enfin des exercices de *prises d'initiatives*. Parmi ces 84 exercices, quatre portent sur des problèmes d'algorithmique.
- Cinq pages qui présentent des extraits de sujets de bac. Ces extraits sont présentés avec les thèmes du cours concernés, pour certains avec des éléments de correction, des *réflexes à avoir*, des rappels ou encore des indications quant au *savoir-faire* à utiliser.
- Trois pages dédiées à l'*accompagnement personnalisé*. Il s'agit de cinq exercices et d'une activité d'approfondissement concernant la formule de Bayes. Les exercices font référence aux *savoir-faire* du cours ou sont présentés avec une méthode à suivre.
- Une page consacrée à un TP d'algorithmique.

Dans la suite de notre analyse nous avons fait le choix d'exclure les exercices et activités exclusivement algorithmiques. Notre analyse ne portera donc plus que sur dix-neuf pages.

Registres de représentations

Nous nous intéressons, dans ce qui suit, aux registres de représentations (Duval, 1993) présents dans le chapitre de probabilités conditionnelles de ce manuel. Nous commençons par les recenser. Les résultats que nous exposons sont regroupés dans le tableau 4.2.

	introduction + activités introductives (4 pages)	cours + savoir-faire (6 pages)	exercices (19 pages)	TOTAL
tableaux	2	4	7	13
arbres de probabilités	4	4	16	20
illustrations diverses	2	0	11	13

FIGURE 4.2 – répartition des registres de représentations dans le manuel Indice pour le chapitre *probabilités conditionnelles*

Dans ce chapitre il y a treize tableaux à doubles entrées et vingt arbres de probabilités, ces représentations sont une façon de représenter la situation décrite et de déterminer les probabilités manquantes. De façon générale, dans les manuels du secondaire on trouve des illustrations diverses, il s'agit de photos ou de dessins n'ayant pas d'intérêt mathématique mais uniquement un rôle d'illustration. On compte treize illustrations diverses dans ce chapitre. Nous présentons des exemples dans la figure qui suit 4.3.

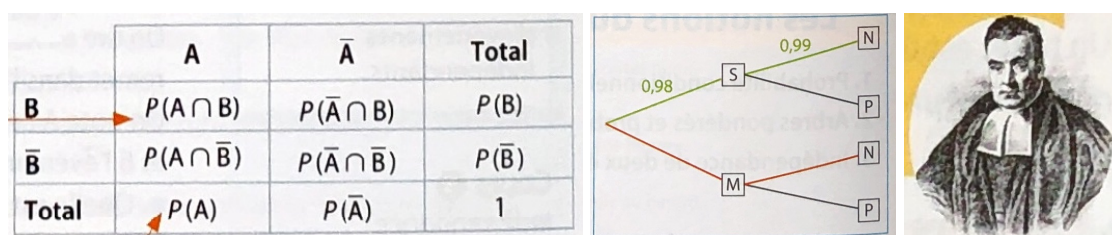


FIGURE 4.3 – exemples d'un tableau de probabilités (1) d'un arbre de probabilités (2) et d'une illustration (3) – extraits du manuel Indice pour le chapitre *probabilités conditionnelles*

Concernant les arbres et les tableaux de probabilités nous cherchons à savoir s'ils sont toujours explicités et s'il y a des changements de registres à la charge des élèves. En prenant le soin de regarder les exercices résolus du chapitre, nous n'avons pas trouvé d'exercice dans lequel des changements de registres (arbre \leftrightarrow tableau) étaient attendus ou demandés. En revanche, dans les éléments de corrections de ces exercices, il est très souvent présenté un arbre de probabilités, y compris lorsque celui-ci n'est pas explicitement demandé dans l'exercice.

Contextes des exercices

Nous nous sommes ensuite intéressés aux contextes des exercices et des exemples du chapitre. En effet, d'après les résultats évoqués dans la revue de travaux (chapitre 1 page 13), nous savons que le contexte impacte la motivation de l'élève à s'intéresser aux problèmes. Il semblerait que les exercices présentés dans un contexte artificiel, c'est-à-dire qui manquent d'authenticité, sont une cause de difficulté supplémentaire pour les élèves.

Nous avons donc choisi de distinguer trois catégories pour les contextes des exemples et exercices. En nous appuyant sur la classification de Eichler et Vogel (2014) qui ont établi une classification des situations problèmes, voici les trois catégories que nous proposons :

- exercice dans un contexte théorique mathématique.
- exercice relevant d'une situation virtuelle : il s'agit ici de situations présentées avec un contexte fortement réduit.
- exercice relevant d'une situation réelle : il s'agit de situations plus authentiques qui fournissent un ancrage narratif et parfois visent à reproduire les problèmes réels d'une société.

Nous avons donc recensé les contextes des exemples et exercices du chapitres. Les résultats quantitatifs sont regroupés dans le tableau 4.4.

	introduction + activités introductives (4 pages)	cours + savoir-faire (6 pages)	exercices (19 pages)	TOTAL	%
théorique	0	3	39	42	38.9%
virtuel	3	0	6	9	8.3%
réel	5	4	48	57	52.8%

FIGURE 4.4 – contexte des exercices et exemples dans le manuel Indice pour le chapitre *probabilités conditionnelles*

Près de 40% des exercices et exemples présentés dans ce chapitre sont dans un contexte mathématique théorique et environ 8% des exercices sont des situations problématiques virtuelles. Dans la figure 4.5 nous en présentons des exemples.

Plus de la moitié des exercices font référence à des situations relevant de la vie réelle d'après notre analyse. Ce sont des exercices qui correspondent à des situations concrètes, certains font référence à des situations réelles visant à s'intéresser à un problème de société comme des élections par exemple. Ces résultats sont en accords avec les recommandations du programme officiel (Ministère de l'Education Nationale (MEN), 2011), on peut citer : "*cette partie du programme*

<p>7 A et B désignent deux événements de l'ensemble des issues d'une expérience aléatoire tels que : $P(A) = 0,3$, $P_A(B) = 0,2$ et $P(B) = 0,54$. Déterminer $P(A \cap B)$ puis $P_B(A)$.</p> <p>8 C et D désignent deux événements de l'ensemble des issues d'une expérience aléatoire tels que : $P(C) = 0,4$, $P_C(D) = 0,25$ et $P_D(C) = 0,2$.</p> <ol style="list-style-type: none"> Déterminer $P(C \cap D)$. Exprimer $P(D)$ en fonction de $P(C \cap D)$ et $P_D(C)$. En déduire la valeur de $P(D)$. 	<p>51 On lance un dé cubique bien équilibré dont les faces portent respectivement 1, 2, 3, 4, 5 et 6 points. On considère les événements suivants :</p> <ul style="list-style-type: none"> - A : « le nombre obtenu est pair » ; - B : « le nombre obtenu est un multiple de trois » ; - C : « le nombre obtenu est inférieur ou égal à trois ». <ol style="list-style-type: none"> Les événements A et B sont-ils indépendants ? Répondre à la même question pour les événements A et C, puis pour les événements B et C. <p style="text-align: right;">→ Pour vous aider : Savoir-faire 6, p. 301</p>
---	---

FIGURE 4.5 – exemples d'un contexte théorique (à g.) et d'une situation virtuelle (à d.) – extraits du manuel Indice pour le chapitre *probabilités conditionnelles*

se prête particulièrement à l'étude de situations concrètes". Dans la figure 4.6 nous présentons des exemples de telles situations d'exercices.

Trois candidats A, B et C se présentent à une élection. Ils obtiennent respectivement la moitié, les trois dixièmes et le cinquième des suffrages. D'autre part, on sait que 50 % des électeurs de A, 30 % des électeurs de B et 40 % des électeurs de C sont des hommes.
 On interroge au hasard une personne s'étant prononcé pour l'un des trois candidats.

- Décrire l'expérience aléatoire à l'aide d'un arbre pondéré.
- En déduire la probabilité d'interroger un homme ayant voté pour le candidat C.
- On interroge au hasard une personne s'étant prononcée pour l'un des trois candidats. Déterminer la probabilité que ce soit une femme.

101 Quatre vacanciers Ahmed, Boris, Coralie et Damien tirent à la courte-paille pour savoir qui fera la vaisselle.
 L'un d'eux, Damien, présente aux trois autres quatre allumettes dont une brûlée. Ils tirent à tour de rôle et celui qui tire l'allumette brûlée fera la vaisselle, et bien sûr Damien prend l'allumette laissée par les autres.
 Coralie proteste : « Je tire toujours la troisième depuis des semaines après Ahmed et Boris, et ça tombe sur moi ! Désormais je veux tirer la première ».
 Qu'en pensez-vous ?

FIGURE 4.6 – exemples de situations réelles - extraits du manuel Indice pour le chapitre de probabilités conditionnelles

Analyse praxéologique du chapitre

Nous allons désormais présenter l'analyse praxéologique du chapitre. Cette analyse répond aux questions : que doit faire l'élève et comment doit-il le faire ?

Afin de lister avec le plus d'exhaustivité possible un ensemble de praxéologies représentatives de ce qui est attendu à la fin du lycée, nous complétons l'analyse de ce chapitre de manuel avec des sujets du baccalauréat et leurs corrigés fournis

par l'APMEP. Nous avons en particulier regardé les sujets de 2018 (Pondichéry, Liban), 2019 (Antilles-Guyane, Amérique du Nord, Liban, Métropole) et 2020 (Polynésie, Métropole).

Dans un premier temps nous proposons de faire l'inventaire des types de tâches. Afin de repérer les types de tâches importants pour l'institution nous allons regarder les exemples et activités du cours ainsi que les exercices résolus. Le mot "*activité*" dans le sens des manuels n'est pas le même que celui de la Théorie de l'Activité. Dans notre cas les activités du cours qui présentent des éléments d'explications et de méthodes sont les "*savoir-faire*" présentés plus haut, il s'agit d'exercices corrigés et commentés. Les exercices résolus du manuel sont quant à eux de plusieurs sortes : ceux dont une correction est détaillée et commentée et ceux pour qui des éléments de solution sont proposés à la fin du manuel. Nous complétons ces résultats par l'analyse des sujets de baccalauréat mentionnés ci-dessus.

Nous identifions les techniques en nous appuyant sur ces exercices résolus et sur les corrigés fournis par l'APMEP. Les éléments de technologie sont pour leur part identifiés avec les commentaires des auteurs, la partie de cours, ainsi qu'avec l'analyse du manuel du professeur.

La liste des principaux types de tâches que nous avons relevés, à la fois dans le chapitre du manuel Indice et dans les sujets de baccalauréat précédemment mentionnés, est la suivante :

- **T-Assoc** : associer les valeurs de l'énoncé à des probabilités d'évènements
- **T-CalcBar** : calculer, à partir de la probabilité d'un évènement, la valeur de la probabilité de son évènement contraire
- **T-CompleArb** : compléter un arbre de probabilités
- **T-CompleTab** : compléter un tableau de probabilités
- **T-ModéArb** : modéliser une situation probabiliste décrite en langage naturel par un arbre de probabilités
- **T-ModéTab** : modéliser une situation probabiliste décrite en langage naturel par un tableau
- **T-LirArb** : associer les valeurs numériques présentes sur les branches de l'arbre à des valeurs de probabilités
- **T-CalcProbaCond** : calculer une probabilité conditionnelle
- **T-CalcInter** : calculer la probabilité d'une intersection d'évènements
- **T-CalcSimple** : calculer la probabilité d'un évènement simple
- **T-EtudInd** : étudier l'indépendance de deux évènements
- **T-CourInd** : répondre à des questions de cours en lien avec l'indépendance de deux évènements
- **T-Interp** : interpréter le résultat

- **T-Suite** : établir une relation du type $r_{n+1} = f(r_n)$

Notre analyse des sujets de baccalauréat nous conduit à compléter cette liste par les trois types de tâches complémentaires suivants qui ont la particularité de faire appel à des raisonnements algébriques (équations ou inéquations) :

- à partir d'un énoncé en langage naturel, minorer la probabilité $P(A)$ sachant que $P(D) = P(A \cap D) + P(B \cap D)$ est majorée par une valeur donnée
- à partir d'un énoncé en langage naturel, à partir d'une nouvelle valeur de $P_E(U)$ et sans connaître $P(E)$, calculer $P(U)$
- déterminer x défini par $P_{\overline{M}}(E) = x$ en résolvant l'équation $x = P(E)$

Nous n'en tiendrons pas compte dans l'élaboration de notre MPR-ET compte tenu du faible nombre d'énoncés dans lesquels on les rencontre.

Le type de tâches **T-CourInd** fait référence au contenu du programme *Indépendance de deux évènements*. Dans le cas de deux évènements, ce thème est largement abordé dans les exercices du manuel sous forme de démonstrations ou de questions de cours, ce qui coïncide avec la (seule) capacité attendue dans le programme pour ce contenu. Par exemple, on trouve dans le manuel les variations (terme défini ci-dessous) suivantes : "*établir la formule de $P(A \cup B)$ pour deux évènements indépendants*", "*rappeler la formule de $P_B(A)$ pour deux évènements indépendants*" ou encore "*justifier que $P(A \cup B) + P(\overline{A}) \times P(\overline{B}) = 1$ lorsque A et B sont indépendants*".

Nous appelons **variations** du type de tâches T , des modifications dans l'énoncé des exercices qui influencent les techniques à utiliser et qui provoquent des adaptations de la part de l'élève. Par exemple, pour un même type de tâche, il peut s'agir du format des données ou de si l'énoncé est en langage naturel ou non.

Nous avons relevé de nombreuses variations pour les principaux types de tâches listés ci-dessus. Par exemple, suivant que l'énoncé de l'exercice ou celui de la question est en langage naturel, une étape de modélisation et/ou d'interprétation aura lieu lors de la résolution de la tâche. De même, dans un exercice dont l'énoncé est en langage naturel, les évènements sont souvent identifiés et nommés préalablement mais parfois cette étape est à la charge de l'élève. Suivant le format des données numériques (pourcentages, fréquences naturelles ou fréquences relatives) l'activité des élèves ne sera pas la même, des adaptations auront lieu.

Nous reprenons donc ci-dessous la liste des types de tâches exposée précédemment avec, cette fois, l'ajout des différentes variations que nous avons pu relever. Cette liste sera par la suite représentée sous la forme synthétique d'un schéma.

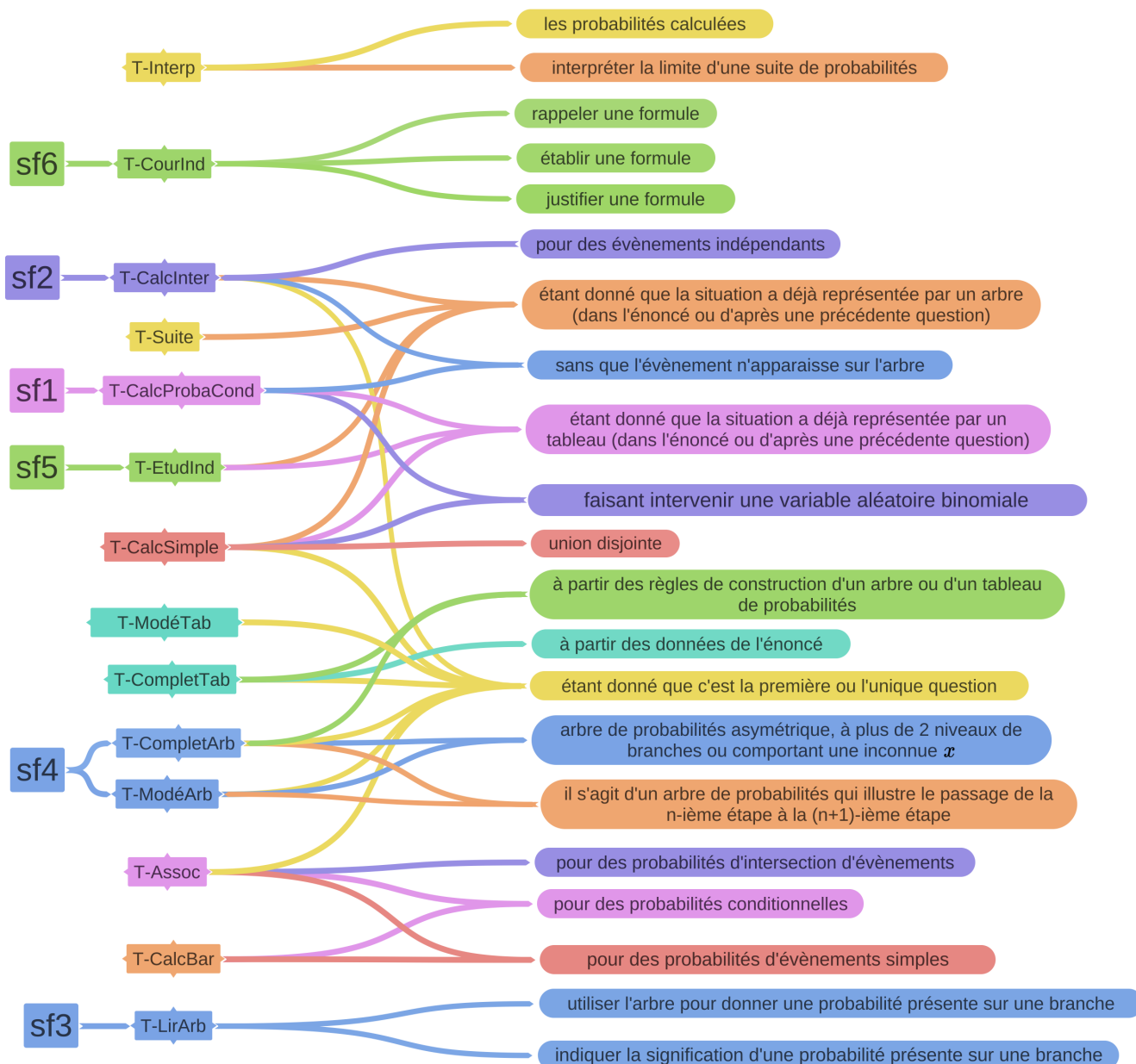
- **T-Assoc** : associer les valeurs de l'énoncé à des probabilités d'évènements
 - associer les valeurs de l'énoncé à des probabilités d'évènements simples
 - associer les valeurs de l'énoncé à des probabilités d'intersection d'évènements

- associer les valeurs de l'énoncé à des probabilités conditionnelles
- à partir d'un énoncé en langage naturel
- étant donné que c'est la première question de l'exercice
- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- sans que les évènements aient été explicitement nommés dans le texte
- les données sont en pourcentages ou fréquences relatives
- T-CalcBar : calculer, à partir de la probabilité d'un évènement, la probabilité de son évènement contraire
 - à partir d'un énoncé en langage naturel
 - calculer, à partir de la probabilité d'un évènement simple, la probabilité de son évènement contraire
 - calculer, à partir d'une probabilité conditionnelle, la probabilité de son évènement contraire
 - étant donné que les évènements ont été décrits et nommés dans le texte
 - sans que les évènements aient été explicitement nommés dans le texte
 - les données sont en pourcentages ou fréquences relatives
- T-CompletArb : compléter un arbre de probabilités
 - à partir d'un énoncé en langage naturel
 - compléter un arbre de probabilités obtenu après la n-ième étape d'une situation probabiliste où $r_n = P(R_n)$
 - compléter un arbre de probabilités asymétrique
 - étant donné que les évènements ont été décrits et nommés dans le texte
 - sans que les évènements aient été explicitement nommés dans le texte
 - à partir des données de l'énoncé
 - à partir des règles de construction de l'arbre de probabilités
 - étant donné que c'est la première question de l'exercice
 - les données sont en fréquences naturelles ou fréquences relatives
- T-CompletTab : compléter un tableau
 - à partir d'un énoncé en langage naturel
 - étant donné que c'est la première question de l'exercice
 - étant donné que les évènements ont été décrits/nommés dans le texte
 - sans que les évènements aient été explicitement nommés dans le texte
 - à partir des données de l'énoncé
 - à partir des règles de construction du tableau de probabilités
 - les données sont en fréquences naturelles ou fréquences relatives
- T-ModéArb : modéliser une situation probabiliste décrite en langage naturel par un arbre de probabilités
 - à partir d'un énoncé en langage naturel
 - l'arbre sera asymétrique ou aura plus de deux niveaux de branches

- l'arbre illustre le passage de la n -ième étape à la $(n+1)$ -ième étape, pour n quelconque
- l'arbre ne pourra pas être complet car il manque des données
- étant donné que les évènements ont été décrits et nommés dans le texte
- sans que les évènements aient été explicitement nommés dans le texte
- étant donné que c'est la première question de l'exercice
- les données sont en pourcentages, en fréquences naturelles ou fréquences relatives
- T-ModéTab : modéliser une situation probabiliste décrite en langage naturel par un un tableau de probabilités
 - à partir d'un énoncé en langage naturel
 - étant donné que les évènements ont été décrits et nommés dans le texte
 - étant donné que c'est la première question de l'exercice
 - les données sont en pourcentages ou fréquences relatives
- T-LirArb : associer les valeurs numériques présentes sur les branches de l'arbre à des valeurs de probabilités
 - étant donné qu'il est demandé de donner/préciser la valeur de la probabilité d'un évènement présent sur une branche de l'arbre
 - étant donné qu'il est demandé d'indiquer la signification d'une certaine valeur numérique qui correspond à une probabilité située sur une branche de l'arbre
- T-CalcProbaCond : calculer une probabilité conditionnelle
 - à partir d'un énoncé en langage naturel
 - étant donné que les évènements ont été décrits et nommés dans le texte
 - sans que les évènements aient été explicitement nommés dans le texte
 - sans que l'évènement apparaisse sur l'arbre
 - étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées
 - étant donné que la situation a déjà été représentée par un arbre ou un tableau (dans l'énoncé ou d'après une précédente question)
 - faisant intervenir une variable aléatoire binomiale
 - les données sont en pourcentages, en fréquences naturelles ou fréquences relatives
- T-CalcInter : calculer la probabilité d'une intersection d'évènements
 - à partir d'un énoncé en langage naturel
 - étant donné que les évènements sont indépendants
 - étant donné que les évènements ont été décrits et nommés dans le texte
 - sans que les évènements aient été explicitement nommés dans le texte
 - étant donné que la situation a déjà été représentée par un arbre (dans l'énoncé ou d'après une précédente question) mais sans forcément que l'évènement apparaisse sur l'arbre
 - étant donné que c'est la première ou la seule question de l'exercice

- les données sont en pourcentages, en fréquences naturelles ou fréquences relatives
- T-CalcSimple : calculer la probabilité d'un évènement simple
 - à partir d'un énoncé en langage naturel
 - étant donné que les évènements ont été décrits et nommés dans le texte
 - sans que les évènements aient été explicitement nommés dans le texte
 - étant donné que la situation a déjà été représentée par un arbre (dans l'énoncé ou d'après une précédente question)
 - étant donné que la situation a déjà été représentée par un tableau (dans l'énoncé ou d'après une précédente question)
 - étant donné qu'il s'agit d'une union d'évènements disjoints
 - étant donné que c'est la première ou la seule question de l'exercice
 - faisant intervenir une variable aléatoire binomiale
 - à partir d'une probabilité d'intersection inconnue notée "x"
 - les données sont en pourcentages, en fréquences naturelles ou fréquences relatives
- T-EtudInd : étudier l'indépendance de deux évènements
 - à partir d'un énoncé en langage naturel
 - étant donné que les évènements ont été décrits et nommés dans le texte
 - sans que les évènements aient été explicitement nommés dans le texte
 - étant donné que la situation a déjà été représentée par un tableau (dans l'énoncé ou d'après une précédente question)
 - étant donné que la situation a déjà été représentée par un arbre (dans l'énoncé ou d'après une précédente question)
 - les données sont en fréquences naturelles ou fréquences relatives
- T-CourInd : répondre à des questions de cours en lien avec l'indépendance de deux évènements
 - établir une formule
 - rappeler une formule du cours
 - justifier une formule
- T-Interp : interpréter le résultat
 - préciser à quoi correspondent les probabilités calculées
 - interpréter la limite de $r_n = P(R_n)$
 - étant donné que les évènements ont été décrits et nommés dans le texte
 - sans que les évènements aient été explicitement nommés dans le texte
- T-Suite : établir une relation du type $r_{n+1} = f(r_n)$
 - étant donné que la situation a déjà été représentée par un arbre (dans l'énoncé ou d'après une précédente question)

Dans l'objectif de réaliser un MPR-ET et afin d'exposer plus clairement l'ensemble des résultats ci-dessus nous proposons une représentation des liens entre savoir-faire, types de tâches et variations de types de tâches dans la figure 4.7.



sf 1	calculer une probabilité conditionnelle
sf 2	calculer la probabilité d'une intersection
sf 3	exploiter la lecture d'un arbre pondéré
sf 4	construire et utiliser un arbre
sf 5	étudier l'indépendance de deux évènements
sf 6	utiliser l'indépendance de deux évènements

FIGURE 4.7 – représentation des liens entre savoir-faire, types de tâches et variations de types de tâches – probabilités conditionnelles

Pour permettre de garder le plus de lisibilité possible dans la figure, nous avons fait le choix de ne pas représenter les variations qui sont communes à la quasi-totalité des types de tâches. Il s'agit pour ce chapitre des variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les évènements ont déjà été décrits/nommés dans le texte
- sans que les évènements aient été explicitement décrits/nommés dans le texte
- les données sont en pourcentages, fréquences naturelles ou fréquences relatives

Nous considérons que les variations - telles que nous les avons précédemment définies - portent exclusivement sur les formes et formulations d'énoncés des exercices tels qu'ils sont présentés dans le manuel et dans les sujets de baccalauréat. A l'aide des corrigés à notre disposition nous avons alors relevé, pour chacun des exercices et des types de tâches, les adaptations - *a priori* - à la charge de l'élève (noté \rightarrow) lors de la résolution de ces tâches. Nous nous plaçons donc du point de vue de l'élève et de son activité (voir partie 2.3 page 38). Suite à cela nous faisons l'hypothèse que chacune des variations de types de tâches précédemment mentionnées provoque des adaptations dans l'activité de l'élève. Nous faisons ici le choix de ne pas présenter l'ensemble de ces adaptations afin de ne pas surcharger cette partie.

Voici quelques adaptations provoquées par des variations de types de tâches, que nous avons relevées :

- T-Assoc : associer les valeurs de l'énoncé à des probabilités d'évènements, à partir d'un énoncé en langage naturel
 - \rightarrow l'élève doit prendre en charge le changement de registre
- T-CalcBar : calculer, à partir de la probabilité d'un évènement, la valeur de la probabilité de son évènement contraire, à partir d'un énoncé en langage naturel
 - \rightarrow l'élève doit prendre en charge le changement de registre
- T-CompletArb : compléter un arbre asymétrique de probabilités
 - \rightarrow l'élève doit adapter ses connaissances pour compléter un arbre qui est asymétrique
- T-CompletTab : compléter un tableau de probabilités lorsque les évènements n'ont pas été identifié ni nommé dans le texte
 - \rightarrow l'énoncé laisse à la charge de l'élève une étape du raisonnement
- T-Modé : modéliser une situation probabiliste décrite en langage naturel par un arbre de probabilités ou un tableau, étant donné que c'est la première question de l'exercice
 - \rightarrow l'énoncé laisse alors à la charge de l'élève des étapes du raisonnement
- T-LirArb : associer les valeurs numériques présentes sur les branches de l'arbre à des valeurs de probabilités

- T-CalcProbaCond : calculer une probabilité conditionnelle faisant apparaître X , une variable aléatoire de loi binomiale
 \rightsquigarrow l'élève doit mettre en jeu des connaissances antérieures/d'un autre chapitre
- T-CalcInter : calculer la probabilité d'une intersection d'évènements étant donné que la situation a déjà été représentée par un arbre à une précédente question
 \rightsquigarrow la réponse à la question nécessite l'utilisation d'un résultat établi dans les questions précédentes
- T-CalcSimple : calculer la probabilité d'un évènement simple à partir d'une situation d'équiprobabilité
 \rightsquigarrow l'élève doit mettre en jeu plusieurs connaissances (définition classique de Laplace)
- T-EtudInd : étudier l'indépendance de deux évènements
 \rightsquigarrow l'élève doit utiliser une connaissance et reconnaître les modalités de son application
- T-CourInd : répondre à des questions de cours en lien avec l'indépendance de deux évènements
- T-Interp : interpréter le résultat
 \rightsquigarrow la réponse à la question nécessite l'utilisation d'un résultat établi dans les questions précédentes
- T-Suite : établir une relation du type $r_{n+1} = f(r_n)$ à partir d'un arbre de probabilités
 \rightsquigarrow la réponse à une question nécessite l'utilisation d'un résultat établi dans les questions précédentes et l'élève doit mettre en jeu des connaissances antérieures/d'un autre chapitre

Pour chacun des principaux types de tâches nous détaillons les ingrédients de la technique (τ) qui représentent ici les actions successives pour effectuer la tâche. Les actions successives sont séparées par un point-virgule. Pour cela nous nous appuyons exclusivement sur les corrigés à notre disposition dans les manuels et sur les corrigés fournis par l'APMEP pour les sujets de baccalauréat. Parfois, plusieurs techniques sont proposées pour un même type de tâche, dans ce cas-là nous détaillons ces différentes techniques, notées $\tau_1, \tau_2 \dots$

Nous présentons également les éléments de technologies (θ), ceux-ci proviennent des commentaires des auteurs (manuel ou corrigés APMEP). De ce fait, certains types de tâches n'ont ni ingrédient de technique ni technologie proposés, on les appelle **types de tâches élémentaires**.

- **T-Assoc** : associer les valeurs de l'énoncé aux probabilités : $P(A)$, $P_B(A)$, $P(A \cap B)$ étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

$\tau : x$

θ : x

Il s'agit d'un type de tâches élémentaire sauf dans le cas où les données de l'énoncé sont en fréquences naturelles. Dans ce cas-là un ingrédient de la technique est celui du type de tâches **T-CalcSimple-FN** (voir plus loin).

- **T-CalcBar** : calculer, à partir de la probabilité d'un évènement, la valeur de la probabilité de son évènement contraire, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ : utiliser une formule du type $P(\bar{A}) = 1 - P(A)$

θ : $P(A) = 1 - P(\bar{A})$

- **T-CompletArb** : compléter un arbre de probabilités, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ : **T-Assoc** ; placer sur les branches de l'arbre les probabilités déjà connues ; déterminer les probabilités manquantes en utilisant la règle des noeuds

θ : la somme des probabilités des branches issues d'un même noeud est égale à 1, ce qui permet de compléter l'arbre ; $P(\bar{A}) = 1 - P(A)$

- **T-CompletTab** : compléter un tableau de probabilités, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ : **T-Assoc** ; placer dans le tableau les probabilités déjà connues ; déterminer les probabilités manquantes en utilisant des formules du type $P(B \cap \bar{A}) = P(B) - P(B \cap A)$

θ : $P(B) = P(B \cap \bar{A}) + P(B \cap A)$

Dans le cas où les données de l'énoncé sont en fréquences naturelles, un ingrédient supplémentaire de la technique est celui associé au type de tâches **T-CalcSimple-FN** (voir plus loin).

- **T-ModéTab** : modéliser une situation probabiliste décrite en langage naturel par un tableau de probabilités, à partir d'un énoncé en langage naturel, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ : placer les évènements simples dans la première colonne et la première ligne du tableau ; placer dans le tableau les valeurs des probabilités déjà connues ; déterminer les valeurs manquantes en utilisant des formules du type $P(B \cap \bar{A}) = P(B) - P(B \cap A)$

θ : x

Dans le cas où les données numériques de l'énoncé sont en fréquences naturelles, la représentation par un tableau se fait en gardant les fréquences naturelles plutôt que d'introduire dès cette étape les probabilités des évènements.

- **T-ModéArb** : modéliser une situation probabiliste décrite en langage naturel par un arbre de probabilités

τ : identifier les évènements en jeu et les nommer ; **T-Assoc** ; tracer l'arbre ; placer les évènements sur l'arbre et les valeurs des probabilités déjà calculées ; compléter l'arbre à l'aide des trois règles de construction de l'arbre (page de cours du manuel)

θ : pour construire un arbre pondéré on interprète et on traduit les informations de l'énoncé en distinguant bien les probabilités conditionnelles. Les évènements A , B et C figurent au premier niveau de l'arbre avec sur chaque branche la probabilité correspondante, au second niveau de l'arbre figurent les probabilités conditionnelles.

- **T-LirArb** : associer les valeurs numériques présentes sur les branches de l'arbre à des valeurs de probabilités

τ : indiquer la signification des valeurs présentes sur les branches de l'arbre partiellement complété

θ : on inscrit sur la branche allant de A vers B la probabilité conditionnelle $P_A(B)$

- **T-CalcProbaCond1** : calculer une probabilité conditionnelle $P_B(A)$

τ_1 : utiliser la définition d'une probabilité conditionnelle $P_B(A) = \frac{P(A \cap B)}{P(B)}$; remplacer les deux termes par les valeurs préalablement calculées

τ_2 : utiliser la définition d'une probabilité conditionnelle et le tableau de probabilités qui donne immédiatement $P(A \cap B)$ et $P(B)$

θ_1 : pour calculer $P_B(A) = \frac{P(A \cap B)}{P(B)}$ il faut au préalable avoir déterminé $P(A)$ et $P(A \cap B)$

θ_2 : x

Ici deux techniques sont possibles, notées τ_1 et τ_2 , avec θ_1 et θ_2 leurs éléments de technologies associés. Dans le cas où les données présentes dans le tableau sont en fréquences naturelles, un ingrédient supplémentaire de la technique est celui associé au type de tâches **T-CalcSimple-FN** (voir plus loin).

- **T-CalcProbaCond2** : calculer une probabilité conditionnelle, à partir d'un énoncé en langage naturel, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ : interpréter l'énoncé pour identifier les évènements en jeu et la probabilité recherchée ; utiliser la définition d'une probabilité conditionnelle $P_B(A) = \frac{P(A \cap B)}{P(B)}$; remplacer les deux termes par les valeurs préalablement calculées

θ : on veut calculer la probabilité de A sachant B , c'est-à-dire $P_B(A) = \frac{P(A \cap B)}{P(B)}$

- **T-CalcInter1** : calculer la probabilité d'une intersection d'évènements $P(A \cap B)$

τ_1 : utiliser la définition d'une probabilité conditionnelle $P(A \cap B) = P_B(A) \times P(B)$; remplacer les deux termes par les valeurs préalablement calculées

τ_2 : utiliser l'arbre de probabilités et la règle du chemin

θ_1 : pour calculer la probabilité d'une intersection, penser aux probabilités conditionnelles $P(A \cap B) = P_B(A) \times P(B)$

θ_2 : la probabilité d'un chemin est le produit des probabilités des branches composant ce chemin, ce qui permet d'écrire $P(A \cap B) = P_B(A) \times P(B)$

Ici deux techniques sont possibles, notées τ_1 et τ_2 , avec θ_1 et θ_2 leurs éléments de technologies associés.

- **T-CalcInter1bis** : calculer la probabilité d'une intersection d'évènements $P(A \cap B)$ lorsque les évènements (contraires) sont indépendants

τ : utiliser l'indépendance des évènements (contraires) pour calculer la probabilité de l'intersection comme le produit des probabilités des évènements

θ : lorsque A et B sont indépendants, \bar{A} et B le sont aussi, \bar{B} et A le sont aussi et on a $P(A \cap B) = P(A) \times P(B)$

- **T-CalcInter2** : calculer la probabilité d'une intersection, à partir d'un énoncé en langage naturel, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ_1 : interpréter l'énoncé pour identifier l'évènement recherché; utiliser la définition d'une probabilité conditionnelle $P(A \cap B) = P_B(A) \times P(B)$; remplacer les deux termes par les valeurs préalablement calculées

τ_2 : interpréter l'énoncé pour identifier l'évènement recherché; utiliser l'arbre de probabilités et la règle du chemin

θ_1 : $P(A \cap B) = P_B(A) \times P(B)$

θ_2 : $P(A \cap B)$ est la valeur obtenue au bout du chemin "vert" sur l'arbre pondéré

Ici deux techniques sont possibles, notées τ_1 et τ_2 , avec θ_1 et θ_2 leurs éléments de technologies associés.

- **T-CalcInter2bis** : calculer la probabilité d'une intersection d'évènements indépendants, à partir d'un énoncé en langage naturel, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ : interpréter l'énoncé pour identifier l'évènement recherché; utiliser l'indépendance des deux évènements pour calculer la probabilité de leur intersection

θ : dans un énoncé, lorsqu'il y a le mot "indépendant" il faut utiliser les égalités du cours

- **T-CalcSimple1** : calculer $P(A)$

τ : utiliser l'arbre et la formule des probabilités totales

θ : x

- **T-CalcSimple2** : calculer la probabilité d'un évènement, à partir d'un énoncé en langage naturel, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ : interpréter l'énoncé pour identifier l'évènement recherché; utiliser l'arbre et la formule des probabilités totales

θ : les trajets conduisant à la réalisation de l'évènement F sont les trajets aboutissant à un évènement écrit en rouge, sur l'arbre pondéré. La formule des probabilités totales : $P(F) = P(A \cap F) + P(B \cap F) + P(C \cap F) = P(A) \times P_A(F) + P(B) \times P_B(F) + P(C) \times P_C(F)$

- **T-CalcSimpleFN** : calculer la probabilité d'un évènement, à partir d'un énoncé en langage naturel, étant donné que les données sont en fréquences naturelles

τ : identifier qu'il s'agit d'une situation d'équiprobabilité et calculer la probabilité des évènements comme le quotient du nombre de cas favorables sur le nombre total de cas

θ : x

- **T-CalcSimple3** : calculer la probabilité d'une union d'évènements disjoints, à partir d'un énoncé en langage naturel, étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ : interpréter l'énoncé pour identifier l'évènement recherché, exprimer la probabilité de l'union comme somme de probabilités (évènements disjoints); utiliser les règles inhérentes à l'arbre de probabilités afin de calculer la probabilité des deux évènements

θ : Soit $A = R_1 \cup R_2$, où R_1 et R_2 sont deux évènements incompatibles, alors $P(A) = P(R_1) + P(R_2)$

- **T-EtudInd1** : étudier l'indépendance de deux évènements étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé

τ_1 : comparer le produit des probabilités des évènements avec la probabilité de l'intersection des évènements

τ_{1bis} : comparer le produit des probabilités des évènements avec la probabilité de l'intersection des évènements, en s'aidant d'un tableau de probabilités

τ_2 : comparer $P_A(B)$ avec $P(B)$

θ_1 : pour justifier l'indépendance de deux évènements on vérifie que $P(A \cap B) = P(A) \times P(B)$;

θ_2 : pour justifier l'indépendance de deux évènements on vérifie que $P_A(B) = P(B)$ ou $P_B(A) = P(A)$

Ici trois techniques sont possibles, dont deux très similaires notées τ_1 et $\tau_1 bis$.

- **T-EtudInd2** : étudier l'indépendance de deux évènements étant donné que les évènements ont déjà été décrits et nommés dans l'énoncé et que les évènements contraires sont indépendants

τ : citer une propriété du cours : si deux évènements A et B sont indépendants alors il en est de même pour \overline{A} et B , pour les évènements A et \overline{B} ainsi que pour les évènements \overline{A} et \overline{B}

θ : x

Lorsque l'un de ces types de tâches représente l'unique question ou la première question de l'exercice, et si l'énoncé est en langage naturel, un ingrédient supplémentaire de la technique est **T-Assoc** voire parfois **T-ModéArb**.

- **T-CourInd** : répondre à des questions de cours en lien avec l'indépendance de deux évènements, comme rappeler la formule de $P_B(A)$ pour deux évènements indépendants

τ : citer une propriété du cours : si deux évènements A et B sont indépendants alors $P_B(A) = P(A)$

θ : x

- **T-Interp** : après avoir calculé des probabilités du type $P(A \cap B)$ ou $P_A(B)$ interpréter le résultat

τ : utiliser l'énoncé de l'exercice pour traduire le résultat numérique en langage naturel ; faire un changement de registres.

θ : x

- **T-Suite** : établir une relation du type $r_{n+1} = f(r_n)$

τ : **T-ModéArb** illustrant le passage de la n -ième étape à la $(n+1)$ -ième étape ; utiliser l'arbre et la formule des probabilités totales pour exprimer r_{n+1} en fonction de r_n

θ : x

Il s'agit ici d'un usage très spécifique de l'arbre, il permet de donner une idée de ce qui se passe lors du passage de n à $n + 1$.

Synthèse

Les analyses présentées dans cette partie correspondent aux attendus de la fin du lycée pour le thème des probabilités conditionnelles, du point de vue des manuels et des sujets de baccalauréat qui influencent les pratiques enseignantes. Les recommandations du programme quant à l'étude de *situations concrètes* sont prises en compte. Les analyses que nous venons de présenter permettent également d'éclaircir les attendus pour le thème de l'indépendance de deux événements : il n'y a pas juste la capacité C-DemInd (cf p.57) mais bien de multiples types de tâches associés à ces contenus.

Nous constatons que de nombreux types de tâches et variations font référence à une activité de modélisation probabiliste, nous définirons ce terme plus précisément à la suite des observations menées en classe de terminale scientifique (partie 5.2.3 page 190). Il s'agit généralement, à partir d'un énoncé en langage naturel, de commencer par identifier les événements en jeu, les nommer et déterminer leurs probabilités. Selon la situation, l'utilisation d'un arbre ou d'un tableau de probabilités peut être pertinente ou non. Il s'agit alors de changements de registres liés à l'activité de modélisation (passage du registre de la langue naturelle au registre de représentation *arbre* ou *tableau*).

Nous avons observé que dans le cas où un arbre (ou un tableau) pourrait être une aide ou une étape - ingrédient de la technique - il n'est pas toujours suggéré par l'énoncé d'en faire un, sauf bien entendu lorsqu'il est demandé dans une première question de représenter la situation par un arbre ou un tableau de probabilités. Ainsi les changements de registres liés à une activité de modélisation ne sont pas toujours explicités et sont alors à la charge de l'élève.

Dans la suite nous nous intéressons au chapitre du manuel consacré aux lois à densité, nous proposons une méthodologie d'analyse très similaire à celle que nous avons appliquée pour ce chapitre sur les probabilités conditionnelles.

4.3.3 Lois de probabilités à densité

Le chapitre de lois continues, appelé dans ce manuel **lois de probabilités à densité**, est le 11^{ème} chapitre du manuel Indice (Poncy et al., 2012). L'organisation de ce chapitre est très similaire à ce que nous avons pu présenter précédemment. Au début de celui-ci un encadré présente les "*notions du chapitre*" qui sont ici :

- * lois de probabilités à densité et loi uniforme
- * loi exponentielle
- * loi normale centrée réduite
- * propriété de la loi normale centrée réduite
- * lois normales

De la même façon que dans le chapitre précédent, un paragraphe introductif motive les nouvelles notions par quelques exemples :

La loi binomiale est une loi discrète, c'est-à-dire qu'une variable suivant une telle loi ne prend qu'un nombre fini de valeurs. Lorsqu'une variable aléatoire prend toutes les valeurs d'un intervalle de \mathbb{R} , on dit que l'on a une loi à densité. Par exemple, le poids ou la taille des enfants par tranches d'âges, que l'on retrouve dans les carnets de santé, peuvent être modélisés par une loi normale, qui est une loi à densité. (Poncy et al., 2012).

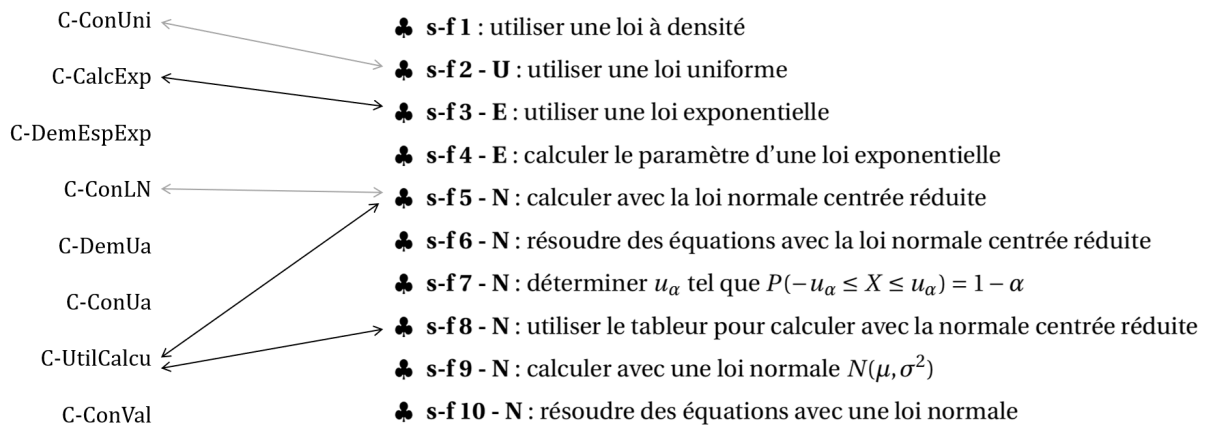
Les "capacités attendues" du programme officiel sont, dans ce manuel, des "savoir-faire" (notés "s-f"). Voici la liste des savoir-faire proposés dans ce chapitre sur les lois de probabilités à densité :

- ♣ s-f 1 : utiliser une loi à densité
- ♣ s-f 2 - U : utiliser une loi uniforme
- ♣ s-f 3 - E : utiliser une loi exponentielle
- ♣ s-f 4 - E : calculer le paramètre d'une loi exponentielle
- ♣ s-f 5 - N : calculer avec la loi normale centrée réduite
- ♣ s-f 6 - N : résoudre des équations avec la loi normale centrée réduite
- ♣ s-f 7 - N : déterminer u_α tel que $P(-u_\alpha \leq X \leq u_\alpha) = 1 - \alpha$
- ♣ s-f 8 - N : utiliser le tableur pour calculer avec la loi normale centrée réduite
- ♣ s-f 9 - N : calculer avec une loi normale $N(\mu, \sigma^2)$
- ♣ s-f 10 - N : résoudre des équations avec une loi normale

Ces savoir-faire sont présentés dans les pages dédiées au cours. Nous les considérons comme des types de tâches. Chaque savoir-faire est présenté sous la forme d'un exercice à plusieurs questions, avec des solutions commentées et des éléments de méthode.

La correspondance entre capacités attendues officiellement et savoir-faire présents dans le manuel n'est pas totale. Dans la figure 4.8 nous présentons un schéma qui fixe les relations entre ces deux familles.

Comme précédemment, dans la figure 4.8 les flèches noires représentent une correspondance parfaite, tandis que les flèches grises représentent une correspondance partielle (la capacité apparaît dans le savoir-faire mais l'exercice correspondant au savoir-faire ne présente pas que cette capacité). Si certaines capacités officielles ne sont reliées à aucun savoir-faire du manuel c'est parce qu'elles ne sont décrites dans aucun de ces savoir-faire. Les capacités **C-DemEspExp** et **C-DemUa** sont uniquement décrites dans la partie de cours et présentées comme des *démonstrations exigibles*. Les capacités **C-ConUa** et **C-ConVal** apparaissent également dans les parties de cours sous la forme de propriétés, ce sont des capacités que l'on retrouvera à mettre en oeuvre dans des exercices.

FIGURE 4.8 – correspondance entre *capacités attendues* et *savoir-faire*

Structure du chapitre

Nous présentons dans ce qui suit la structure détaillée du chapitre qui est très similaire à celle du chapitre précédent.

Le chapitre commence par trois exercices permettant de "réactiver les savoirs". Le premier exercice est un vrai-faux qui concerne la loi discrète binomiale. Le second exercice est un QCM sur la fonction exponentielle, qui associe équations, fractions, primitives et limites pour cette fonction particulière. Enfin, le troisième exercice permet de travailler sur le calcul d'aires et les intégrales de fonctions polynomiale.

Les deux pages suivantes proposent des activités introductives aux nouvelles notions de ce chapitre. Les objectifs annoncés de ces activités sont les suivants :

- ↔ introduire la notion de loi à densité
- ↔ introduire la notion de loi uniforme sur un intervalle
- ↔ introduire la loi exponentielle
- ↔ introduire la loi normale centrée réduite

Les dix pages suivantes sont, comme pour le chapitre précédent, une alternance de pages dédiées au cours et de pages présentant les différents "savoir-faire". Associée à chacune des pages de cours, une page fait face en présentant les "savoir-faire" liés au contenu de cours présentés. Le plan de cette partie du chapitre est le suivant :

1. Loïs de probabilités à densité et loi uniforme
 - Variable aléatoire à densité
 - Loi uniforme sur $[a, b]$
 - ♣ s-f 1 : utiliser une loi à densité
 - ♣ s-f 2 : utiliser une loi uniforme

2. Loi exponentielle
 - ♣ s-f 3 : utiliser une loi exponentielle
 - ♣ s-f 4 : calculer le paramètre d'une loi exponentielle
3. Loi normale centrée réduite
 - Théorème de Moivre-Laplace
 - Loi normale centrée réduite
 - Calculs de probabilités
 - ♣ s-f 5 : calculer avec la loi normale centrée réduite
 - ♣ s-f 6 : résoudre des équations avec la loi normale centrée réduite
4. Propriétés de la loi normale centrée réduite
 - Règles de calcul
 - Valeurs remarquables liées à la loi normale centrée réduite
 - ♣ s-f 7 : déterminer u_α tel que $P(-u_\alpha \leq X \leq u_\alpha) = 1 - \alpha$
 - ♣ s-f 8 : utiliser le tableur pour calculer avec la loi normale centrée réduite
5. Lois normales
 - Loi normale $N(\mu, \sigma^2)$
 - Calculs de probabilités
 - Les intervalles "un, deux, trois sigma"
 - ♣ s-f 9 : calculer avec une loi normale
 - ♣ s-f 10 : résoudre des équations avec une loi normale

Ce chapitre se poursuit par vingt pages d'exercices dont l'organisation est très similaire à celle du chapitre précédent. Parmi ces pages il y a (en italiques les dénominations du manuel) :

- 104 exercices dont des exercices résolus, des vrai-faux, des exercices chronométrés, des exercices *pour démarrer*, des exercices *pour s'entraîner*, des exercices *pour aller plus loin*, d'autres pour se préparer à l'oral du baccalauréat et enfin des exercices *de prises d'initiatives*. Parmi ces 104 exercices trois portent sur des problèmes d'algorithmique.
- Quatre pages qui présentent des extraits de sujets de bac. Ces extraits sont présentés avec les thèmes du cours concernés, pour certains avec des éléments de correction, des *réflexes à avoir*, des rappels ou encore des indications quant au *savoir-faire* à utiliser.
- Trois pages dédiées à l'*accompagnement personnalisé*. Il s'agit de cinq exercices et d'une activité d'approfondissement concernant la méthode de Monte-Carlo. Les exercices font référence aux *savoir-faire* du cours ou sont présentés avec une méthode à suivre.
- Une page consacrée à un TP d'algorithmique.

Dans la suite de notre analyse nous avons fait le choix d'exclure les exercices et activités exclusivement algorithmiques. Notre analyse ne portera donc plus que sur dix-neuf pages.

Registres de représentations

Nous nous intéressons, dans ce qui suit, aux registres de représentations présents dans le chapitre de lois à densité. Nous avons commencé par recenser ces registres. Les résultats que nous exposons sont regroupés dans le tableau 4.9.

	introduction + activités introductives (4 pages)	cours + savoir-faire (10 pages)	exercices (19 pages)	TOTAL
graphiques de fonctions	4	19	11	34
illustrations diverses	3	1	13	17
tableaux de valeurs de probabilités	0	0	3	3
représentations géométriques	0	0	2	2

FIGURE 4.9 – répartition des registres de représentations dans le manuel Indice pour le chapitre *lois à densité*

Etant donné que le chapitre est consacré aux variables aléatoires continues les représentations graphiques de fonctions sont essentiellement celles de fonctions de densité. Ainsi, les représentations graphiques de fonctions représentent plus de la moitié des registres de représentations recensés au sein de ce chapitre, mais dans le détail, seulement 10% des exercices (11 pour 101) proposent un travail à partir d'un graphique de fonctions. Cela ne correspond donc pas vraiment aux recommandations des textes officiels qui précisent, pour ce chapitre en particulier : "le recours aux représentations graphiques et aux simulations est indispensable".

On recense d'autres registres de représentations, comme des tableaux de probabilités et des schémas représentant des situations géométriques. De nouveau, on retrouve de nombreuses illustrations diverses, on compte dix-sept illustrations non-mathématiques dans ce chapitre. Nous présentons des exemples dans la figure 4.10.

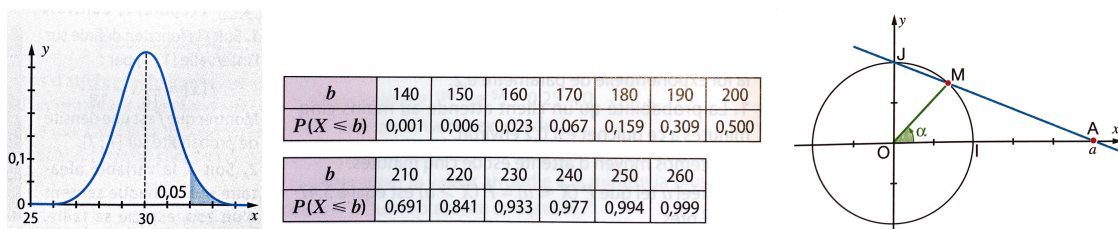


FIGURE 4.10 – graphique de fonction (1), tableau de probabilités (2) et représentation géométrique (3) – extraits du manuel Indice

Contextes des exercices

Comme nous l'avons évoqué dans la revue de travaux, des chercheurs ont mis en évidence que les contextes des exercices influent sur l'activité de l'élève et sur sa motivation (voir partie 1 page 13). De plus, nous rappelons les recommandations du programme (Ministère de l'Éducation Nationale (MEN), 2011) concernant les thèmes de probabilités qui nous concernent : "*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. [...]*". C'est pourquoi nous accordons une attention particulière aux contextes dans lesquels sont présentés les exercices et exemples du chapitre de ce manuel.

Nous présentons dans ce qui suit des éléments relatifs aux contextes des exercices de ce chapitre. Nous reprenons les trois catégories présentées dans l'analyse du chapitre précédent, à savoir :

- exercice dans un contexte théorique mathématique.
- exercice relevant d'une situation virtuelle : il s'agit ici de situations présentées avec un contexte fortement réduit comme relevant de la physique ou de la géométrie.
- exercice relevant d'une situation réelle : il s'agit de situations plus authentiques qui fournissent un ancrage narratif et parfois visent à reproduire les problèmes réels d'une société.

Après avoir recensé les contextes des exemples et exercices du chapitre nous présentons les résultats quantitatifs dans le tableau 4.11.

	introduction + activités introductives (4 pages)	cours + savoir-faire (10 pages)	exercices (19 pages)	TOTAL	%
théorique	5	7	70	82	62.6%
virtuel	1	1	2	4	3%
réel	1	4	40	45	34.4%

FIGURE 4.11 – contexte des exercices et exemples pour le chapitre *lois à densité*

Ainsi, près de deux tiers des exercices et exemples présentés dans ce chapitre sont dans un contexte mathématique théorique, ce qui peut sembler surprenant au vu des recommandations du programme officiel. Dans la figure 4.12 nous présentons deux exemples d'exercices dans un contexte très différent.

On trouve trace de quelques exercices dans un contexte que nous appelons ici *virtuel*, il s'agit de situations relevant de problèmes de géométrie ou d'autres disciplines comme la physique avec l'exemple de l'étude de la désintégration nucléaire.

93 Une densité de probabilité curieuse
Soit a un réel strictement positif et f la fonction définie sur $[-a; a]$ par $f(x) = e^{-|x|}$.

- Déterminer a pour que f soit une fonction de densité de probabilité.
- Soit X la variable aléatoire continue dont la fonction de densité de probabilité est la fonction f déterminée à la question 1.
 - Calculer $P(X > -0,5)$.
 - Soit g la fonction définie par $g(x) = xf(x)$.
Montrer que, pour tout réel x appartenant à l'intervalle $[-a; a]$, on a $g(-x) = -g(x)$.
 - En déduire par des considérations géométriques que l'espérance mathématique de X est nulle.

65 Un commerçant vend des moteurs électriques dont la durée de vie en années est une variable aléatoire qui suit la loi exponentielle de paramètre 0,2

- Si le moteur est garanti un an, quelle proportion de ses clients le commerçant devra-t-il dépanner avant la fin de la garantie ?
- Quelle durée de garantie maximale (en mois) doit être choisie pour que le commerçant ait à dépanner moins de 10 % de ses clients durant cette garantie ?

FIGURE 4.12 – exemples d'un contexte théorique (à g.) et d'une situation réelle (à d.) – extraits du manuel Indice pour le chapitre *lois à densité*

Cela correspond aux recommandations du programme officiel qui précise : "On illustre ces nouvelles notions par des exemples issus des autres disciplines". Voici un exemple d'un tel exercice dans la figure 4.13.

La variable aléatoire X qui, à chaque atome radioactif d'iode 131, associe sa durée de vie en jours, suit une loi exponentielle. On sait que la probabilité que cette durée de vie soit inférieure à deux jours est égale à $0,160$ à 10^{-3} près.

- Calculer, à 10^{-3} près, le paramètre λ de cette loi exponentielle.
- La demi-vie d'une substance radioactive est le temps t au bout duquel la moitié des atomes initiaux sont désintégrés.
Calculer, à $0,1$ près, la demi-vie de l'iode 131.

FIGURE 4.13 – exemples de situation virtuelle – extrait du manuel Indice pour le chapitre *lois à densité*

Analyse praxéologique du chapitre

Nous allons désormais présenter l'analyse praxéologique du chapitre. Cette analyse prend la même forme qu'au chapitre précédent.

Afin de lister avec le plus d'exhaustivité possible un ensemble de praxéologies représentatives de ce qui est attendu à la fin du lycée, nous complétons l'analyse de ce chapitre de manuel avec des sujets du baccalauréat et leurs corrigés fournis par l'APMEP. Nous avons en particulier regardé les sujets de 2019 (Centres étrangers, Antilles-Guyane, Polynésie, Métropole et Nouvelle-Calédonie) et 2020 (Polynésie).

Pour repérer les types de tâches importants pour le thème des lois continues, nous avons regardé, au sein de ce chapitre de manuel, les exemples et activités du cours ainsi que les exercices résolus. Nous avons complété ces résultats par l'analyse des sujets de baccalauréat mentionnés ci-dessus. Nous identifions les

techniques en nous appuyant sur ces exercices résolus et sur les corrigés fournis par l'APMEP. Nous présentons ensuite les technologies selon les commentaires des auteurs, la partie de cours, et l'analyse du manuel du professeur.

Voici la liste des principaux types de tâches que nous avons relevés, à la fois dans le chapitre du manuel Indice et dans les sujets du baccalauréat précédemment mentionnés :

- **T-Veriff** : vérifier qu'une fonction donnée est bien une densité de probabilités
- **T-DonF** : définir/donner la fonction de densité d'une variable aléatoire continue
- **T-DonLoi** : définir/donner la loi d'une variable aléatoire continue
- **T-CalcSimple** : calculer la probabilité d'un évènement simple du type $\{X < a\}$, $\{X > b\}$ ou $\{a < X < b\}$
- **T-Hach** : hachurer une certaine probabilité à partir de la représentation graphique de la fonction de densité
- **T-ExprEx** : exprimer $P(X > a)$ en fonction de λ le paramètre d'une loi exponentielle
- **T-CalcProbaCond** : calculer une probabilité conditionnelle du type $P_{X < a}(X > b)$
- **T-CalcEsp** : calculer l'espérance mathématique d'une variable aléatoire continue
- **T-Det(λ)** : déterminer le paramètre λ d'une loi exponentielle
- **T-Det(a)** : déterminer le réel a tel que $P(X < a) = \text{nombre donné}$
- **T-Det(u_α)** : déterminer u_α pour α un nombre donné
- **T-DetInterv** : déterminer un intervalle I centré en l'espérance de X tel que $P(X \in I) = \text{nombre donné}$
- **T-Det(σ)** : déterminer l'écart-type d'une variable aléatoire X de loi normale
- **T-Det($\sigma + \mu$)** : déterminer la moyenne et l'écart-type d'une variable aléatoire X suivant une loi normale
- **T-Interp** : interpréter le résultat
- **T-CalcMoy** : calculer la moyenne de $P(X < a)$ et $P(Y < b)$ lorsque X suit une loi uniforme donnée et Y suit une loi normale donnée

Dans la liste ci-dessus nous distinguons "calculer" de "déterminer" suivant que la technique se fasse par un calcul ou non.

Nous utilisons ici aussi la notion de **variations** du type de tâches T . Il s'agit de modifications dans l'énoncé des exercices qui influencent les techniques à utiliser ou qui provoquent des adaptations de la part de l'élève.

Nous avons relevé des variations concernant les principaux types de tâches listés ci-dessus. Par exemple pour un même type de tâche, il peut s'agir d'une densité quelconque, d'une loi uniforme, exponentielle ou normale. L'énoncé peut être en langage naturel ou non, et induire ainsi une activité de modélisation. Les évènements ne sont pas toujours identifiés et nommés préalablement, dans ce cas cette étape est à la charge de l'élève. Selon les variations présentes l'activité des élèves ne sera pas la même, des adaptations auront lieu. Nous exposons ci-dessous la liste des principaux types de tâches avec les variations que nous avons relevées. Comme pour le chapitre précédent sur les probabilités conditionnelles, un schéma, à retrouver figure 4.14 page 97, synthétise cette longue liste.

- **T-VerifF** : vérifier qu'une fonction donnée est bien une densité de probabilités
 - étant donné que la variable aléatoire est identifiée dans l'énoncé et que sa densité a été donnée dans l'énoncé ou établie dans une question précédente
 - à partir de la représentation graphique de la fonction de densité
- **T-DonF** : définir/donner la fonction de densité d'une variable aléatoire continue
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi uniforme
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi exponentielle
- **T-DonLoi** : définir/donner la loi d'une variable aléatoire continue
 - définie comme $\frac{X-\mu_X}{\sigma_X}$ où X suit une loi normale $\mathcal{N}(\mu, \sigma^2)$ donnée
- **T-CalcSimple** : calculer la probabilité d'un évènement simple du type $\{X < a\}$, $\{X > b\}$ ou $\{a < X < b\}$
 - étant donné que la variable aléatoire est identifiée dans l'énoncé et que sa densité a été donnée dans l'énoncé ou établie dans une question précédente
 - étant donné que l'énoncé est en langage naturel
 - étant donné que c'est la première question
 - étant donné que la densité n'a pas été donnée dans l'énoncé
 - à partir de la représentation graphique de la fonction de densité
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi uniforme
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi exponentielle
 - à partir d'une valeur de l'espérance, étant donné qu'il s'agit d'une loi exponentielle de paramètre λ inconnu
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite et en utilisant un tableur (obligatoirement)
 - étant donné qu'il s'agit d'une loi normale de paramètres μ et σ donnés
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale $\mathcal{N}(\mu; \sigma^2)$ de paramètres μ et σ NON donnés
 - il s'agit en fait d'un intervalle à 1, 2 ou 3 σ (loi normale $\mathcal{N}(\mu; \sigma^2)$)

- calculer $P(X > a)$ étant donné qu'on connaît $P(X < a - \mu)$ dans le cas où X suit une loi normale où μ son espérance est donnée
- calculer $P(X > a)$ ou $P(a < X < b)$ étant donné qu'on connaît $P(X < a)$ et/ou $P(X > b)$
- **T-Hach** : hachurer une certaine probabilité à partir de la représentation graphique de la fonction de densité
- **T-ExprEx** : exprimer $P(X > a)$ en fonction de λ le paramètre d'une loi exponentielle
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi exponentielle
 - à partir d'une valeur de l'espérance, étant donné qu'il s'agit d'une loi exponentielle de paramètre λ inconnu
- **T-CalcProbaCond** : calculer une probabilité conditionnelle du type $P_{X < a}(X > b)$
 - à partir d'un énoncé en langage naturel
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale de paramètres μ et σ donnés
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi exponentielle
 - calculer : $P_{X > a}(X > b)$, étant donné qu'il s'agit, d'après l'énoncé, d'une loi exponentielle et en utilisant la propriété du vieillissement
- **T-CalcEsp** : calculer l'espérance mathématique d'une variable aléatoire continue
 - à partir d'un énoncé en langage naturel (ex : "calculer le prix moyen")
 - étant donné que la variable aléatoire est définie et que sa densité a été donnée dans l'énoncé ou établie dans une question précédente
 - étant donné que c'est la première question
 - étant donné que la densité n'a pas été donnée dans l'énoncé
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi uniforme
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale de paramètres μ et σ donnés
 - à partir de considérations géométriques
 - à partir de la représentation graphique de la fonction de densité
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi exponentielle
 - à partir de la valeur d'une probabilité du type $P(a < X < b)$ ou $P(X < a)$ (décrite en langage naturel ou non), étant donné qu'il s'agit d'une loi exponentielle de paramètre λ inconnu
- **T-Det(λ)** : déterminer le paramètre λ d'une loi exponentielle
 - à partir de la valeur d'une probabilité du type $P(a < X < b)$ ou $P(X < a)$ (décrite en langage naturel ou non), étant donné qu'il s'agit d'une loi exponentielle de paramètre λ inconnu

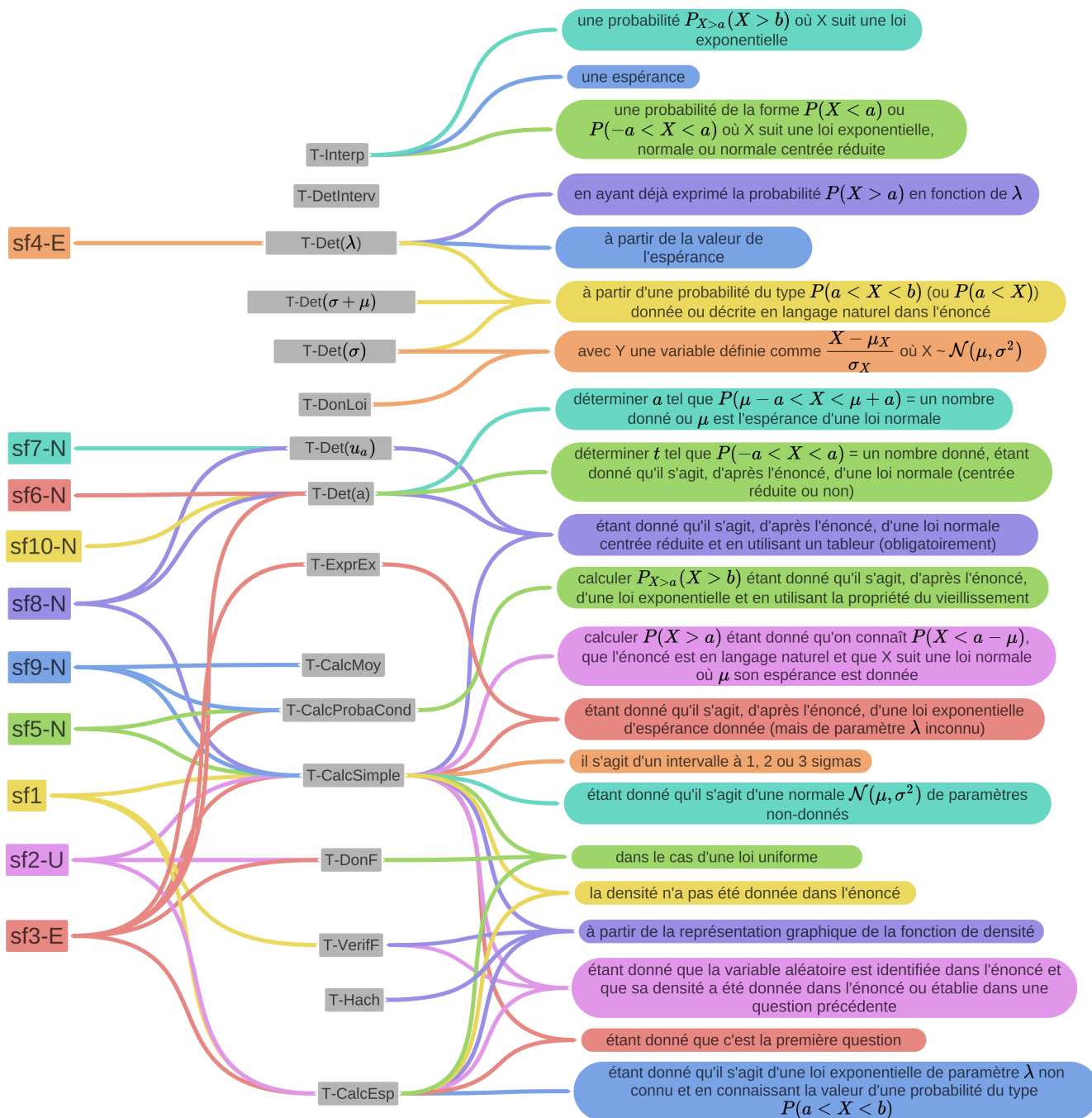
- à partir d'une valeur de l'espérance, étant donné qu'il s'agit d'une loi exponentielle de paramètre λ inconnu
- **T-Det(a)** : déterminer le réel a tel que $P(X < a) =$ un nombre donné
 - à partir d'un énoncé en langage naturel (ex : "déterminer la température t telle que 10% des journées aient une température supérieure à t ")
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite et en utilisant un tableur obligatoirement
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale de paramètres μ et σ donnés
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi exponentielle
 - déterminer le réel b tel que $P(X > b) =$ un nombre donné
 - déterminer a tel que $P(-a < X < a) =$ un nombre donné ; étant donné qu'il s'agit d'une loi normale (centrée réduite ou non)
 - déterminer a tel que $P(\mu - a < X < \mu + a) =$ un nombre donné ; étant donné qu'il s'agit d'une loi normale de paramètres μ et σ donnés
- **T-Det(u_α)** : déterminer u_α pour α un nombre donné
 - à partir d'un énoncé en langage naturel
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite et en utilisant un tableur obligatoirement
- **T-DetInterv** : déterminer un intervalle I centré en l'espérance de X tel que $P(X \in I) =$ un nombre donné
 - étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale de paramètres μ et σ donnés
- **T-Det(σ)** : déterminer l'écart-type d'une variable aléatoire X de loi normale
 - où Z définie comme $\frac{X-\mu_X}{\sigma_X}$ suit une loi normale centrée réduite
 - sans qu'une VA Z centrée réduite associée à X soit définie/donnée comme $\frac{X-\mu_X}{\sigma_X}$
 - à partir d'une probabilité du type $P(a < X < b)$ donnée ou décrite en langage naturel dans l'énoncé
- **T-Det($\sigma + \mu$)** : déterminer la moyenne et l'écart-type d'une variable aléatoire X suivant une loi normale
 - à partir de probabilités du type $P(a < X)$ données ou décrites en langage naturel dans l'énoncé
- **T-Interp** : interpréter le résultat
 - dans le cas d'une espérance
 - d'une probabilité " $P_{X>a}(X > b)$ " où X suit une loi exponentielle
 - d'une probabilité " $P(X < a)$ " où X suit une loi exponentielle
 - d'une probabilité " $P(X < a)$ " où X suit la loi normale centrée réduite

- d'une probabilité " $P(-a < X < a)$ " où X suit la loi normale centrée réduite
- d'une probabilité " $P(X < a)$ " où X suit une loi normale de paramètres μ et σ donnés
- d'une probabilité $P(X \in I)$ où X suit une loi normale de paramètres μ et σ donnés
- **T-CalcMoy** : calculer la moyenne de $P(X < a)$ et $P(Y < b)$ lorsque X suit une loi uniforme donnée et Y suit une loi normale donnée
 - à partir d'un énoncé en langage naturel

Dans l'objectif de réaliser un MPR-ET et afin d'exposer plus clairement l'ensemble des résultats ci-dessus, nous proposons (figure 4.14) une représentation des liens entre **savoir-faire**, **types de tâches** et **variations** de types de tâches pour le thème des lois à densité.

De la même façon que pour les probabilités conditionnelles et afin de garder le plus de lisibilité possible sur cette figure, nous avons fait le choix de ne pas représenter les variations de types de tâches qui concernent quasiment tous les types de tâches. Il s'agit des variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné qu'il s'agit d'une loi normale centrée réduite
- étant donné qu'il s'agit d'une loi normale de paramètres μ et σ donnés
- étant donné qu'il s'agit d'une loi exponentielle



sf 1	utiliser une loi à densité
sf 2 - U	utiliser une loi uniforme
sf 3 - E	utiliser une loi exponentielle
sf 4 - E	calculer le paramètre d'une loi exponentielle
sf 5 - N	calculer avec la loi normale centrée réduite
sf 6 - N	résoudre des équations avec la loi normale centrée réduite
sf 7 - N	déterminer u_α tel que $P(-u_\alpha \leq X \leq u_\alpha) = 1 - \alpha$
sf 8 - N	utiliser le tableur pour calculer avec la normale centrée réduite
sf 9 - N	calculer avec une loi normale $N(\mu, \sigma^2)$
sf 10 - N	résoudre des équations avec une loi normale

FIGURE 4.14 – représentation des liens entre savoir-faire, types de tâches et variations de types de tâches – lois à densité

Nous considérons que les variations - telles que nous les avons précédemment définies - provoquent des adaptations dans l'activité de l'élève (utiliser une connaissance antérieure, introduire un intermédiaire, mettre en jeu plusieurs connaissances, utiliser un résultat établi dans les questions précédentes, etc). Pour ne pas surcharger cette partie nous ne détaillons pas les adaptations pour ce chapitre, nous laissons au lecteur le plaisir d'imaginer les adaptations qui peuvent découler des variations sus-mentionnées.

Pour les principales variations de types de tâches nous détaillons ci-dessous les ingrédients de la technique (τ) qui représentent ici les actions successives pour effectuer la tâche. Pour cela nous nous appuyons sur les corrigés à notre disposition dans les manuels et sur les corrigés de sujets de baccalauréat fournis par l'APMEP. Lorsque plusieurs techniques sont proposées nous le détaillons.

Nous présentons également les éléments de technologies (θ), ceux-ci proviennent des commentaires des auteurs (manuel et corrigés issus de l'APMEP). Certains types de tâches n'ont ni ingrédient de technique ni technologie proposés, ce sont des types de tâches élémentaires. Enfin, certains types de tâches présentent parfois plusieurs techniques, dans ce cas-là nous détaillons ces différentes techniques, notées $\tau_1, \tau_2 \dots$

- **T-Veriff** : vérifier qu'une fonction polynomiale donnée est bien une densité de probabilités

τ : justifier que la fonction est continue et positive sur l'intervalle donné en utilisant les propriétés des fonctions polynomiales ; vérifier par un calcul que l'intégrale de cette fonction sur l'intervalle donné vaut 1

θ : pour vérifier que f est une densité sur I il faut montrer que f est continue, positive et d'intégrale 1 sur I .

- **T-DonF** : donner la fonction de densité d'une variable aléatoire uniforme sur $[a, b]$

τ : x

θ : x

Il s'agit d'une question de cours, c'est un type de tâches élémentaire.

- **T-DonLoi** : donner la loi d'une variable aléatoire continue définie comme $\frac{X-\mu_X}{\sigma_X}$ où X suit une loi normale

τ : x

θ : x

Il s'agit également d'une question de cours, c'est un type de tâches élémentaire.

- **T-CalcSimple1** : calculer la probabilité d'un évènement simple du type $\{X < a\}$, $\{X > b\}$ ou $\{a < X < b\}$ pour une densité quelconque

τ_1 : identifier les bornes a et b puis calculer l'intégrale de la fonction de densité sur cet intervalle en utilisant des techniques relevant de l'analyse.

τ_2 : identifier les bornes a et b puis calculer l'intégrale à partir d'une représentation graphique de la densité et de techniques de géométrie.

- θ_2 : d'après la formule de l'aire d'un triangle, d'après la formule de l'aire d'un rectangle.
- **T-CalcSimple2** : calculer la probabilité d'un évènement simple du type $\{X < a\}$, $\{X > b\}$ ou $\{a < X < b\}$ pour une loi uniforme ou exponentielle
 - τ : identifier les bornes a et b et utiliser la formule du cours $P(a < X < b) = \frac{b-a}{d-c}$ pour une loi uniforme sur $[c, d]$ et $P(a < X < b) = e^{-\lambda a} - e^{-\lambda b}$ pour une loi exponentielle de paramètre λ
 - θ : x
 - **T-CalcSimple3** : calculer la probabilité d'un évènement simple du type $\{a < X < b\}$ pour une loi normale centrée réduite
 - τ_1 : identifier les bornes a et b et utiliser la fonction adéquate de la calculatrice
 - τ_2 : avec un tableur, faire la différence de $P(X < b)$ et $P(X < a)$: **T-CalcSimple4**
 - θ_1 : pour ces questions on doit utiliser la calculatrice
 - θ_2 : pour calculer $P(a < X < b)$ on écrit : $P(a < X < b) = P(X < b) - P(X \leq a)$, les fonctions sont différentes sur la calculatrice et sur le tableur
 - **T-CalcSimple4** : calculer la probabilité d'un évènement du type $\{X < a\}$ ou $\{X > b\}$ pour une loi normale centrée réduite
 - τ_1 : attribuer la valeur 0 à une des bornes et identifier l'autre borne, **T-CalcSimple3** et ajouter cette valeur à 0.5 (ou retrancher cette valeur à 0.5) pour obtenir la probabilité souhaitée
 - τ_2 : avec un tableur, utiliser directement la fonction adéquate
 - θ_1 : pour les calculs du type $P(X < a)$ ou $P(X > b)$ on peut s'aider d'un croquis
 - θ_2 : les fonctions sont différentes sur la calculatrice et sur le tableur
 - **T-CalcSimple5** : calculer la probabilité d'un évènement simple du type $\{a < X < b\}$ pour une loi normale
 - τ_1 : identifier les bornes a et b et utiliser la fonction adéquate de la calculatrice
 - θ_1 : on calcule à l'aide de la calculatrice ; il faut faire attention à l'ordre des bornes
 - **T-CalcSimple6** : calculer la probabilité d'un évènement du type $\{X < a\}$ ou $\{X > b\}$ pour une loi normale
 - τ_1 : attribuer la valeur de l'espérance à une des bornes et identifier l'autre borne, **T-CalcSimple5** puis ajouter cette valeur à 0.5 (ou retrancher cette valeur à 0.5) pour obtenir la probabilité souhaitée
 - θ_1 : on peut utiliser le fait que $P(a < X)$ est peu différent de $P(a < X < 10^{99})$
 - **T-Hach** : hachurer une certaine probabilité à partir de la représentation graphique de la fonction de densité
 - τ : x

θ : x

Il n'y a aucun ingrédient de technique décrit dans le corrigé. C'est un type de tâches élémentaire.

- **T-ExprEx** : exprimer $P(X > a)$ en fonction de λ le paramètre d'une loi exponentielle

τ : x

θ : x

C'est une formule du cours $P(X > a) = e^{-\lambda a}$. Il s'agit donc ici d'un type de tâches élémentaire.

- **T-Det(λ)1** : déterminer le paramètre λ d'une loi exponentielle à partir de la valeur de son espérance

τ : x

θ : x

C'est à nouveau une formule du cours $E(X) = \frac{1}{\lambda}$. Il s'agit donc ici d'un type de tâches élémentaire.

- **T-Det(λ)2** : déterminer le paramètre λ d'une loi exponentielle à partir d'une probabilité du type $P(X < a)$ (exprimée en langage naturel ou non)

τ : identifier a , exprimer $P(X < a)$ en fonction de $P(X > a)$ puis **T-ExprEx**; résoudre l'équation algébrique $e^{-\lambda a} = \text{nombre}$, d'inconnue λ

θ : pour calculer le paramètre d'une loi exponentielle, on résout une équation

- **T-CalcProbaCond1** : calculer une probabilité conditionnelle du type $P_{X < a}(X > b)$

τ : utiliser la définition d'une probabilité conditionnelle, simplifier l'expression de $P(\{X < a\} \cap \{X > b\})$, calculer $P(b < X < a)$ et $P(X < a)$ avec **T-CalcSimple** adéquat

θ : l'intersection des événements " $T \leq a$ " et " $T \geq b$ " est l'évènement " $b \leq T \leq a$ "

- **T-CalcProbaCond2** : calculer une probabilité conditionnelle du type $P_{X > a}(X > b)$ dans le cas d'une loi exponentielle

τ : utiliser la propriété du cours $P_{X > a}(X > a + T) = P(X > T)$

θ : x

- **T-CalcEsp1** : donner l'espérance mathématique d'une variable aléatoire continue de loi uniforme, exponentielle ou normale à partir d'un énoncé en langage naturel

τ : x

θ : x

Ce sont des formules du cours. Par exemple pour la loi uniforme sur $[a, b]$ l'espérance est donnée par $E(X) = \frac{b+a}{2}$. Pour la loi exponentielle l'espérance est donnée par $E(X) = \frac{1}{\lambda}$ où λ est le paramètre de la loi. Pour la loi normale c'est immédiat. Il s'agit donc ici d'un type de tâches élémentaire.

- **T-CalcEsp2** : calculer l'espérance mathématique d'une variable aléatoire continue de densité quelconque
 τ : utiliser la définition de l'espérance, calculer l'intégrale à l'aide de techniques issues de l'analyse.
 θ : x
- **T-CalcEsp3** : donner l'espérance mathématique d'une variable aléatoire de loi normale à partir de sa représentation graphique
 τ : identifier l'axe de symétrie de la courbe, lire l'abscisse correspondante
 θ : la courbe représentative de la densité correspondant à la loi normale a pour axe de symétrie la droite d'équation $x = E(X)$
- **T-Det(a)1** : déterminer le réel a tel que $P(X < a) =$ nombre donné, pour X une variable aléatoire de loi normale (centrée réduite ou non)
 τ_1 : avec une calculatrice, utiliser directement la fonction adéquate
 τ_2 : avec un tableur, utiliser directement la fonction adéquate
 θ_1 : pour simplifier les calculs on introduit la fonction Φ définie sur \mathbb{R} par $\Phi(t) = P(X < t)$
 θ_2 : x
- **T-Det(a)2** : déterminer le réel a tel que $P(X > a)$ ou $P(a < X < 0) =$ nombre donné, pour X une variable aléatoire de loi normale (centrée réduite ou non)
 τ : exprimer $P(X > a)$ ou $P(a < X < 0)$ en fonction de $P(X < a)$ puis **T-Det(a)1**.
 θ : pour résoudre une équation où intervient une variable aléatoire de loi normale (centrée réduite), se ramener à $P(X < a)$
- **T-Det(a)2bis** : déterminer le réel a tel que $P(-a < X < a) =$ nombre donné, pour X une variable aléatoire de loi normale (centrée réduite ou non)
 τ : chercher a tel que $P(-a < X < a) =$ nombre, équivaut à chercher a tel que $2\Phi(a) - 1 =$ nombre, ce qui donne $\Phi(a) = \frac{\text{nombre donné} + 1}{2}$ puis **T-Det(a)1** donne la valeur de a .
 θ : pour résoudre une équation où intervient une variable aléatoire de loi normale (centrée réduite), se ramener à $P(X < a)$
- **T-Det(a)3** : déterminer le réel a tel que $P(X > a) =$ nombre donné, pour X une variable aléatoire de loi exponentielle
 τ : exprimer $P(X > a)$ en fonction de a et λ (formule du cours $P(X > a) = e^{-\lambda a}$) puis résoudre l'équation algébrique $e^{-\lambda a} =$ nombre, d'inconnue a
 θ : par croissance de la fonction logarithme
- **T-Det(u_α)** : déterminer u_α pour α un nombre donné
 τ : se ramener à un problème du type "chercher a tel que $P(-a < X < a) =$ nombre"; **T-Det(a)2**
 θ : " $P(-u_\alpha < X < u_\alpha) =$ nombre", équivaut à " $2\Phi(u_\alpha) - 1 =$ nombre"
- **T-DetInterv1** : déterminer le réel a tel que $P(\mu - a < X < \mu + a) =$ nombre donné, où X suit une loi normale $\mathcal{N}(\mu, \sigma^2)$

- τ : centrer et réduire l'expression avec $Y = \frac{X-\mu}{\sigma}$, alors Y suit une loi normale centrée réduite; le problème revient à "trouver b tel que $P(-b < Y < b) = \text{nombre donné}$ " où $b = -\frac{a}{\sigma}$. Ce qui correspond à **T-Det(a)2bis**. On en déduit a .
- θ : lorsque X suit une loi normale il est nécessaire de se ramener à une loi normale centrée réduite
- **T-DetInterv2** : déterminer un intervalle I centré en l'espérance de X de loi normale $\mathcal{N}(\mu, \sigma^2)$ tel que $P(X \in I) = \text{nombre donné}$

τ : identifier I comme un intervalle de la forme $[\mu - a; \mu + a]$; le problème devient **T-DetInterv1**, on en déduit I .

θ : un intervalle I centré en μ est de la forme $[\mu - a; \mu + a]$ où a est un réel positif
 - **T-Det(σ)** : déterminer l'écart-type d'une variable aléatoire X de loi normale $\mathcal{N}(\mu, \sigma^2)$

τ : identifier une probabilité de la forme $P(a < X < b)$ pour laquelle on connaît une valeur; centrer et réduire l'expression avec $Y = \frac{X-\mu}{\sigma}$, alors Y suit une loi normale centrée réduite; le problème devient "trouver σ tel que $P(-\frac{b}{\sigma} < Y < \frac{a}{\sigma}) = \text{nombre avec } b \text{ donné}$ "; **T-Det(a)2bis**

θ : x
 - **T-Det($\sigma + \mu$)** : déterminer la moyenne et l'écart-type d'une variable aléatoire X suivant une loi normale

τ : si elles ne sont pas déjà données, identifier deux probabilités de la forme $P(X > a)$ et $P(X < b)$ dont on connaît une valeur pour chacune; centrer et réduire les deux expressions avec $Y = \frac{X-\mu}{\sigma}$, alors Y suit une loi normale centrée réduite; le problème devient "trouver σ et μ tels que $\Phi(\frac{b-\mu}{\sigma}) = \text{nombre donné}$ et $\Phi(\frac{a-\mu}{\sigma}) = 1 - \text{nombre donné}$ ". **T-Det(a)1** permet d'obtenir une valeur pour $\frac{b-\mu}{\sigma}$ et $\frac{a-\mu}{\sigma}$, il reste à résoudre ce système.

θ : x
 - **T-Interp** : interpréter le résultat

τ : utiliser l'énoncé de l'exercice pour traduire le résultats numérique en langage naturel; faire un changement de registres.

θ : x
 - **T-CalcMoy** : calculer la moyenne de $P(X < a)$ et $P(Y < b)$ lorsque X suit une loi uniforme donnée et Y suit une loi normale donnée

τ : calculer $P(X < a)$ et $P(Y < b)$, c'est-à-dire **T-CalcSimple2** et **T-CalcSimple6**; faire la moyenne des deux nombres obtenus.

θ : x

Synthèse

Les analyses présentées dans cette partie concernent les attendus de la fin du lycée pour le thème des lois continues. Nous avons bien retrouvé des exercices et exemples issus des autres disciplines, comme le souhaite le programme de mathématiques de la classe de terminale scientifique. En revanche, les exercices sont nombreux à être présentés dans un contexte théorique et les situations concrètes - pourtant celles recommandées par les textes des programmes - plutôt rares.

Les registres de représentations de type graphiques de fonctions sont présents mais finalement que dans un exercice sur dix. On pourrait imaginer qu'il y ait plus souvent ce genre de représentations au vu des recommandations du programme officiel.

Les techniques proposées dans les corrigées et savoir-faire, reposent le plus souvent sur l'utilisation de la calculatrice (calculs avec la loi normale, déterminer des intervalles, etc.), même si des étapes préalables sont souvent nécessaires et donc des adaptations sont à la charge de l'élève (reconnaissance des modalités à mettre en oeuvre, utilisation de résultats issus de questions précédentes, transformation de l'expression). Pour la loi exponentielle, les techniques reposent essentiellement sur l'utilisation directe de formules et ne font pas intervenir de techniques d'intégration, ces résultats ont déjà été mis en avant par Derouet (2016) dans sa thèse.

Pour résumer

Dans cette partie 4.3 nous avons présenté l'analyse des chapitres de manuel selon des critères qui nous semblent pertinents pour l'élaboration de notre futur MPR-ET : structure du chapitre, registres de représentations, contextes des exercices et analyse praxéologique. Tout au long de cette analyse nous avons veillé à faire référence aux textes et aux recommandations des programmes officiels présentés dans la deuxième partie de ce chapitre.

Comme nous l'avons écrit à la fin de la première partie, à propos des paradigmes et approches des probabilités, c'est le contexte des exercices qui détermine le sens, et donc l'approche, à utiliser. Aussi, suivant que les données des exercices sont en pourcentages ou en fréquences naturelles, l'approche fréquentiste ou classique est sous-entendue. Concernant les probabilités conditionnelles, comme le soulignaient certains auteurs dans la première partie, l'approche bayésienne n'apparaît que sous la forme du théorème de Bayes. Ici, ce théorème qui n'est pas au programme de terminale scientifique, n'est évoqué que dans une activité d'approfondissement du manuel.

Nous avons ainsi analysé la structuration du savoir dans le manuel et dans les sujets de baccalauréat. La modalité de nos analyses nous permet d'ores-et-déjà d'anticiper des résultats relatifs à l'activité des élèves. Nous nous appuierons sur ces résultats pour analyser, dans le prochain chapitre, les données empiriques recueillies durant des séances en classe de terminale scientifique et en première année de biologie à l'Université.

Chapitre 5

Enseignement et apprentissage des probabilités à la transition secondaire-supérieur

L'objectif de ce chapitre est de caractériser la transition secondaire-supérieur en probabilités, en termes de praxéologies enseignées, d'activité des étudiants et de pratiques des enseignants (question de recherche n°2).

Pour ce faire, nous présentons l'analyse du savoir enseigné en classe de terminale scientifique et en première année d'Université en filière de biologie, pour les thèmes de probabilités qui nous intéressent. Ces analyses, qui s'appuient sur des documents recueillis en classe et sur de nombreuses séances observées, sont faites au regard des résultats du chapitre 4 page 49.

Nous commençons par exposer, dans une première partie, les outils et terrains d'analyse. Une seconde partie est consacrée à l'analyse du savoir enseigné dans le secondaire. Nous présentons enfin, dans une troisième partie, l'analyse du savoir enseigné en première année de biologie à l'Université.

5.1 Observation de séances de classe : recueil de données et outils d'analyse

Nous avons réalisé des observations dans le secondaire et à l'Université, en France, de séances portant sur les probabilités conditionnelles et des séances sur les lois continues. Ces observations ont eu lieu au cours de l'année scolaire 2018-2019.

Dans le secondaire nos observations se déroulent dans une classe de terminale série scientifique, au sein d'un lycée en milieu rural. Nous avons observé deux séquences dans cette classe, une première à l'automne 2018 pour le chapitre des probabilités conditionnelles et une seconde, plus tard dans l'année scolaire, pour le chapitre de lois continues.

Dans le tableau 5.1 (respectivement 5.2) nous décrivons les séances qui ont eu lieu pour le thème des probabilités conditionnelles (respectivement pour les lois continues), celles auxquelles nous avons assisté et celles que nous avons filmées.

Date	Durée	Thème principal	Présence	Film	Synopsis
08/10/2018	1h	Séance 1 : cours (partie I : Probabilités conditionnelles)	non	non	non
15/10/2018	2h	Séance 2 : cours (partie II : Arbres pondérés) + exercices	oui	non	oui
16/10/2018	2h	TP	x	x	x
05/11/2018	2h	Séance 3 : cours (partie III : Indépendance de deux évènements)	non	non	non
19/11/2018	1h	Séance 4 : extraits de sujets de baccalauréat	oui	oui	oui
26/11/2018	1/4h	Séance 5 : extraits de sujets de baccalauréat	oui	oui	oui

FIGURE 5.1 – Observations en classe de terminale scientifique –
Probabilités conditionnelles

A l'Université nous avons assisté aux séances d'un module de probabilités à destination des étudiants en première année de filière biologie. Ce module de probabilités, d'une durée de 28h réparties sur sept semaines, se compose de sept séances de cours magistraux (CM) alternées avec sept séances de travaux dirigés (TD). Dans le cadre de ce travail, nous avons assisté au CM n°2 et au TD n°2 concernant les probabilités conditionnelles et au CM n°4 et au TD n°4 concernant les lois continues. Dans le tableau 5.3 nous décrivons les séances auxquelles nous avons assisté.

Nous n'avons pas pu filmer toutes les séances (notamment à l'Université) mais à partir des notes prises au cours des séances, ou des films le cas échéant, nous avons produit des synopsis. Un synopsis de séance est un tableau qui présente un découpage en épisodes de la séance. Pour chacun des épisodes nous précisons dans le synopsis s'il est détaillé ou non dans la suite. Généralement les épisodes qui ne sont pas détaillés sont intégralement consacrés à des moments de cours (lecture du photocopié de cours dans le secondaire ou diapositives présentant exclusivement des théorèmes ou définitions à l'Université) ou portent sur des contenus autres que ceux qui nous préoccupent (nombres complexes, suites, etc).

Notre étude repose donc sur ces synopsis de séances, et sur les transcriptions lorsque nous avons pu en faire. Nos analyses sont illustrées par des extraits de

5.1. Observation de séances de classe : recueil de données et outils d'analyse

Date	Durée	Thème principal	Présence	Film	Synopsis
12/03/2019	2h	Séance 1 : cours (partie I : Lois de probabilités à densité) + exercices	oui	oui	oui
21/03/2019	2h	Séance 2 : cours (partie II : Loi uniforme) + exercices	oui	oui	oui
04/04/2019	2h	Séance 3 : cours (partie III : Loi normale centrée réduite) + exercices	oui	oui	oui
23/04/2019	2h	Séance 4 : cours (partie IV : Lois normales) + exercices	oui	oui	oui
29/04/2019	2h	Séance 5 : cours (fin partie IV) + exercices	oui	oui	oui
06/05/2019	2h	Séance 6 : cours (partie V : Lois exponentielles) + exercices	oui	oui	oui
13/05/2019	1/2h	Séance 7 : extraits de sujets de baccalauréat	non	non	non

FIGURE 5.2 – Observations en classe de terminale scientifique –
Lois continues

Date	Durée	Thème principal	Présence	Film	Synopsis
01/02/2019	2h	Cours magistral (CM2) sur les probabilités conditionnelles	oui	non	oui
05/02/2019	2h	Séance de travaux dirigés (TD2) sur les probabilités conditionnelles	oui	non	oui
15/02/2019	2h	Cours magistral (CM4) sur les variables aléatoires continues	oui	non	oui
07/03/2019	2h	Séance de travaux (TD4) sur les variables aléatoires continues	oui	non	oui

FIGURE 5.3 – Observations en L1 de biologie – module de probabilités

transcripts, ou de vidéo et s'appuient sur les résultats du chapitre 4 page 49.

Les questions qui guident notre travail sont les suivantes :

- Quels sont les types de tâches et techniques proposés aux élèves du secondaire et aux étudiants de l'Université ?
- Pour chaque type de tâches relevé, quelles sont les variations à la charge des élèves et étudiants? Comment ces variations sont-elles gérées dans la

classe? Quelles sont les variations systématiquement prises en charge par l'enseignant(e)? L'enseignant(e) élimine-t-il(elle) certaines variations?

- Quel est le rôle de l'enseignant(e) et quelles sont les aides procédurales qui modifient l'activité des élèves sur les types de tâches et variations relevés?
- Les sous-activités de reconnaissance, d'organisation et de traitement sont-elles prises en charge par les élèves ou par l'enseignant(e) et de quelle manière?

5.2 Enseignement des probabilités au secondaire : analyse du savoir enseigné

Dans cette partie, nous présentons l'analyse du savoir enseigné en classe de terminale scientifique pour les thèmes de probabilités qui nous intéressent, à savoir, les probabilités conditionnelles et les lois continues. Notre analyse s'appuie sur des données empiriques que nous avons recueillies en classe de terminale scientifique dans un établissement scolaire français au cours de l'année scolaire 2018-2019. Dans cet établissement, une enseignante de mathématiques a accepté de nous accueillir dans sa classe pour observer les séances concernant les deux thèmes précédemment cités.

Cette enseignante, que l'on peut considérer comme une experte du point de vue didactique (elle a participé à de nombreuses activités de l'IREM, elle est régulièrement tutrice de stagiaires), organise ces séances de la façon suivante :

- Une première période appelée « Mathoral » durant laquelle un élève présente au tableau, devant le reste de la classe, la correction d'un exercice qu'il a préparé chez lui. Cette présentation donne lieu à une note, les élèves sont évalués selon des critères bien précis qui leur sont donnés en amont.
- Une seconde période, elle aussi caractéristique de cette enseignante, appelée « Questions Rapides », l'enseignante pose oralement des questions aux élèves. Ces questions sont généralement en lien avec la suite de la séance, l'enseignante attend des élèves qu'ils y répondent rapidement car il s'agit souvent de questions proches du cours.
- Enfin, une partie associant cours et exercices. En ce qui concerne le cours, l'enseignante a pour habitude de distribuer un polycopié à trous qui guide le déroulement du chapitre. Dans ce polycopié on trouve les éléments classiques d'une leçon : définitions, théorèmes et propositions, ainsi que des exemples.

Le déroulement d'une séance type consiste généralement en la succession de ces trois périodes, ordonnées ou non. Dans la suite de ce document nous appellerons Alix l'enseignante.

Dans la partie suivante nous présentons l'analyse du savoir enseigné pour le thème des probabilités conditionnelles dans cette classe de terminale scientifique. Puis nous procédons de la même façon, dans une seconde partie, pour les lois continues. Dans une troisième et dernière partie nous nous intéressons au savoir enseigné dans le cadre d'activité de modélisation probabiliste. Enfin, nous présentons dans une quatrième et dernière parties des éléments qui nous semblent caractéristiques de cet enseignement des probabilités dans le secondaire.

5.2.1 Analyse du savoir enseigné pour le thème des probabilités conditionnelles

Dans cette partie nous présentons l'analyse du savoir enseigné pour le thème des probabilités conditionnelles.

Alix, l'enseignante de mathématiques de la classe, choisit de consacrer environ huit heures à ce chapitre. Ces heures sont réparties en six séances dont une séance de travaux pratiques d'algorithmique de 2h en salle informatique. Dans la figure 5.1 sont exposés, l'enchaînement de ces six séances ainsi que le thème principal abordé. Nous écartons volontairement cette séance de TP de la suite de nos analyses car l'algorithmique ne fait pas partie du programme du cours de probabilités de première année d'Université en biologie. Nous avons assisté à trois séances et deux d'entre elles ont pu être filmées. Lors de la dernière séance nous avons pu proposer aux élèves un test (voir énoncé du test et analyse des productions d'élèves partie 6.1 page 255).

A partir des films et des notes prises durant les séances nous avons produit les synopsis des séances auxquelles nous avons assisté. Nous avons retranscrit les épisodes des séances filmées identifiés comme intéressants d'après les synopsis. Les épisodes "intéressants" sont ceux dans lesquels le thème des probabilités conditionnelles apparaît, ils correspondent aux lignes grisées dans les synopsis. Nous présentons, pour chaque séance, l'analyse de ces épisodes avant de tirer des conclusions dans une partie de synthèse.

Les résultats du chapitre 4 page 49 guident notre analyse des séances en classe de terminale, au lycée. Il s'agit de répertorier les types de tâches et techniques proposés aux élèves au cours des séances puis, pour chaque type de tâches relevé, d'identifier les variations qui sont à la charge des élèves, celles qui sont systématiquement prises en charge par l'enseignante et comment elles sont gérées dans la classe. Nous nous demandons également si l'enseignante élimine certaines variations, quel est son rôle et quelles sont les aides procédurales qui modifient l'activité des élèves sur les types de tâches et variations relevés.

Séance n° 1

La séance n°1 du chapitre sur les probabilités conditionnelles à laquelle nous n'avons pas assisté, a été consacrée à la première partie du polycopié de cours distribué aux élèves : « I. Probabilités conditionnelles ». Le polycopié est à retrouver en annexe A.1.

Séance n° 2

Thème principal de l'épisode	Commentaires
« Mathoral » sur les suites	X
Travail en autonomie sur les exercices 10, 26 et 26 du manuel	Détails dans ce qui suit
Cours : « II. Arbre pondéré » : règles de construction + exemple	Détails dans ce qui suit
Travail en autonomie sur les exercices 4, 5 et 12 du manuel	Détails dans ce qui suit

FIGURE 5.4 – Synopsis de la séance n°2 – *Probabilités conditionnelles*

La séance n°2 à laquelle nous avons assisté mais que nous n'avons pas pu filmer, dont le synopsis se trouve figure 5.4, commence par un « Mathoral » avec un exercice sur les suites.

La séance se poursuit par le chapitre sur les probabilités conditionnelles et l'enseignante choisit de laisser les élèves travailler en autonomie durant 40 minutes sur des exercices corrigés du manuel *Indice* (voir énoncés des exercices dans la figure 5.5). Deux de ces exercices sont dans un contexte théorique mathématique, le troisième représente une situation de la vie réelle. L'enseignante circule dans les rangs et aide les élèves qui ont des difficultés. Il n'y a pas de correction ni de trace écrite au tableau étant donné qu'il s'agit d'exercices corrigés du manuel.

Les types de tâches (et leurs variations principales) présents dans ces trois exercices sont les suivants :

- **T-CompleTab** (à partir des règles de construction d'un tableau de probabilités)
- **T-CalcSimple** (à partir d'un tableau de probabilités)
- **T-CalcInter** (à partir d'un tableau de probabilités)
- **T-CalcProbaCond** (étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées/sont données dans l'énoncé)
- **T-CalcSimple** (faisant intervenir une variable aléatoire binomiale)
- **T-CalcProbaCond** (faisant intervenir une variable aléatoire binomiale)

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

10 1. A, B, C et D sont des événements de l'ensemble des issues d'une expérience aléatoire.
Recopier et compléter le tableau de probabilités ci-dessous :

	A	B	C	Total
D	0,2		0,1	
\bar{D}		0,15		
Total	0,45		0,3	

2. Déterminer les valeurs de $P(B \cap D)$, $P(B)$ et $P(\bar{D})$.
3. Calculer $P_{\bar{D}}(D)$ et $P_{\bar{D}}(B)$.

29 La variable aléatoire X suit la loi binomiale de paramètres $n = 50$ et $p = 0,3$. Les résultats seront donnés à 10^{-3} près.

1. Calculer la probabilité des événements « $X \leq 20$ » et « $X \geq 10$ ».
2. Calculer $P_{(X \geq 10)}(X \leq 20)$.

26 TRAVAIL EN AUTONOMIE

Énoncé
Parmi les élèves de Terminale ES d'un lycée, 21 % des élèves suivent l'option « LV3 », 15 % l'option « Arts » et 30 % suivent au moins une de ces deux options.
On choisit au hasard un des élèves de Terminale ES.

1. Calculer la probabilité qu'il suive les deux options.
2. Sachant que l'élève suit l'option « Arts », calculer la probabilité qu'il suive aussi choisi l'option « LV3 ».

Solution
On commence par nommer les différents événements.
Soit L l'événement « l'élève choisi suit l'option LV3 » et A l'événement « l'élève choisi suit l'option Arts ».
D'après les données de l'énoncé, on peut écrire que :
 $P(L) = 0,21$; $P(A) = 0,15$ et $P(A \cup L) = 0,3$.

1. On doit déterminer $P(A \cap L)$.
On sait que $P(A \cup L) = P(A) + P(L) - P(A \cap L)$, soit :
 $0,3 = 0,15 + 0,21 - P(A \cap L)$ d'où $P(A \cap L) = 0,36 - 0,3 = 0,06$.
La probabilité que l'élève choisi suive les deux options est 0,06 soit 6 %.

2. On doit déterminer $P_A(L)$.
On a $P_A(L) = \frac{P(A \cap L)}{P(A)} = \frac{0,06}{0,15} = \frac{2}{5} = 0,4$.
Sachant que l'élève suit l'option « Arts », la probabilité qu'il suive aussi l'option « LV3 » est 0,4 soit 40 %.

FIGURE 5.5 – Séance n°2 – Extraits du manuel Indice
Énoncés des exercices 10, 29 et 26

- **T-CalcInter** (à partir d'un énoncé en langage naturel, sans que les événements aient été explicitement nommés dans l'énoncé)
- **T-CalcProbaCond** (à partir d'un énoncé en langage naturel, étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées/sont données dans l'énoncé)

La séance se poursuit avec la suite du cours. Le polycopié à trous distribué lors de la première séance est projeté au tableau. La séance est consacrée à la seconde partie de ce polycopié qui est : « II. Arbre pondéré. » Cette partie permet à l'enseignante d'énoncer les règles de construction d'un arbre de probabilités et de faire le lien avec les formules vues dans la première partie du cours lors de la première séance. Il s'agit d'une aide à visée constructive. Les règles de construction de l'arbre écrites dans le polycopié sont les mêmes que celles du manuel, à savoir :

- Règle n°1 : la somme des probabilités des branches issues d'un même noeud est égale à 1.
- Règle n°2 : le produit des probabilités indiquées sur les branches d'un chemin est égal à la probabilité de l'intersection des événements rencontrés le long de ce chemin.
- Règle n°3 : formule des probabilités totales : la probabilité d'un événement est égale à la somme des probabilités de chacun des chemins qui y conduisent.

A ces trois règles, l'enseignante ajoute oralement une règle « zéro » : *Les événements indiqués au bout des branches issues d'un même noeud doivent former*

une partition de l'évènement indiqué sur le noeud : c'est-à-dire qu'ils doivent être deux à deux disjoints et leur réunion doit être égale à l'évènement indiqué sur ce noeud.

L'enseignante indique oralement aux élèves deux rôles importants pour l'arbre de probabilités : « *L'arbre permet de calculer des probabilités et de retrouver de façon pratique les formules vues dans le premier paragraphe.* » Ces formules sont d'ailleurs rappelées, écrites au tableau et les élèves doivent les écrire dans le photocopié. L'arbre de probabilités a un rôle central dans le cours, au point d'écrire du savoir procédural : à quoi ça sert, comment s'en servir, comment le relier au reste du cours, etc.

L'exemple présent à la page 3 de ce photocopié est ensuite traité. Il s'agit d'un exercice relevant d'une situation réelle, extrait du sujet de baccalauréat Antilles-Guyane 2010, dont l'énoncé est ci-dessous.

Lors d'une épidémie chez des bovins, on s'est aperçu que si la maladie est diagnostiquée suffisamment tôt chez un animal, on peut le guérir, sinon la maladie est mortelle. Un test est mis au point et essayé sur un échantillon d'animaux dont 2 % sont porteurs de la maladie. On obtient les résultats suivants :

- si un animal est porteur de la maladie, le test est positif dans 85 % des cas,
- si un animal est sain, le test est négatif dans 95 % des cas.

On choisit de prendre ces fréquences observées comme probabilités pour toute la population et d'utiliser le test pour un dépistage préventif de la maladie. Soient respectivement les événements M et T «être porteur de la maladie» et «avoir un test positif.»

1. Un animal est choisi au hasard, quelle est la probabilité que son test soit positif ?
2. Si le test du bovin est positif, quelle est la probabilité qu'il soit malade ?

La première question de cet exemple correspond au type de tâches **T- CalcSimple** : calculer la probabilité d'un évènement simple. Les variations présentes ici sont les suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que c'est la première question de l'exercice
- les données sont en pourcentages

Pour cette question il y a plusieurs adaptations à la charge des élèves. En effet, c'est la première question de l'exercice et l'énoncé est en langage naturel. L'enseignante prend en charge l'organisation au sens des sous-activités mathématiques (reconnaissance, organisation, traitement). Cela inclut une reconnaissance de la méthode à suivre (faire un arbre), elle dit aux élèves : « *Il faut construire l'arbre.*

» Le type de tâches à la charge des élèves devient alors **T-ModéArb** : modéliser une situation probabiliste décrite en langage naturel par un arbre de probabilités. L'enseignante poursuit et prend en charge l'activité d'organisation en indiquant : « *Il faut faire un choix pour les branches.* »

L'enseignante représente ensuite l'arbre au tableau et les élèves interviennent oralement afin de le compléter, cela renforce l'aiguillage des élèves vers l'arbre dans ce type de tâches **T-CalcSimple**. Un premier élève indique qu'il faut faire deux branches. Un autre élève propose d'appeler les branches « *malade, pas malade.* » Des élèves indiquent les valeurs à écrire sur les branches correspondant aux événements. L'activité de traitement final correspondant à l'adaptation suivante "*conversion fréquences/pourcentages*", en lien avec la variation "*les données sont en pourcentages*", semble être prise en charge facilement par les élèves.

Une fois l'arbre complété et commenté par l'enseignante (étant donné que la séance n'a pas été filmée nous n'avons pas plus de détails), des élèves prennent la parole tour à tour pour répondre à la première question. L'enseignante écrit le détail du raisonnement et du calcul au tableau. Un élève propose à un moment d'écrire $P(\overline{M} \cap T)$ comme $P(\overline{M}) \times P(T)$. L'enseignante, afin d'expliquer que cette propriété « *ne marche pas tout le temps* », prend un exemple où la probabilité de l'intersection n'est pas égale au produit des probabilités. Le résultat numérique obtenu est 0,066 ; un élève prend en charge le traitement de la dernière étape en interprétant ce résultat de la manière suivante : « *La probabilité que le test soit positif est 0,066* », un autre élève de la classe souligne : « *Soit 6,6 %.* »

Concernant la seconde question de l'exemple : « *Si le test du bovin est positif, quelle est la probabilité qu'il soit malade?* », il s'agit du type de tâches **T-CalcProbaCond**. Les variations sont les suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les événements ont déjà été décrits et nommés dans le texte
- étant donné que la situation a déjà été représentée par un arbre
- sans que l'évènement apparaisse sur l'arbre
- étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées
- les données sont en pourcentages

L'adaptation principale à la charge des élèves est ici de reconnaître qu'il s'agit du type de tâches calculer une probabilité conditionnelle. Il y a la reconnaissance de ce que l'on demande de calculer avant d'appliquer les techniques du cours. Plusieurs propositions sont faites par des élèves. Un élève propose de calculer $P_T(M)$, un autre $P_{\overline{M}}(T)$, un autre encore $P(M \cap T)$. C'est l'enseignante qui prend en charge la suite de la correction de cette question (étant donné que la séance n'a pas été filmée nous n'avons pas plus de détails).

La séance se termine par un travail en autonomie sur des exercices du manuel dont les énoncés sont dans la figure 5.6. L'enseignante circule dans la classe pour aider les élèves en difficulté.

4 A et B désignent deux événements de l'ensemble des issues d'une expérience aléatoire tels que :
 $P(A) = 0,75$, $P(B) = 0,8$ et $P(A \cap B) = 0,3$.
 Déterminer $P_A(B)$ et $P_B(A)$.

5 C et D désignent deux événements de l'ensemble des issues d'une expérience aléatoire tels que :
 $P(C) = 0,25$, $P(D) = 0,4$ et $P(C \cap D) = 0,1$.
 Déterminer la probabilité conditionnelle de l'événement C sachant que l'événement D est réalisé et déterminer $P_C(D)$.

12 On considère deux événements A et B associés à une expérience aléatoire. L'arbre pondéré ci-contre respecte les conventions habituelles et apporte des informations sur cette expérience aléatoire.

1. Indiquer la signification des nombres 0,6, 0,75 et 0,3.
 2. Compléter l'arbre pondéré et préciser les valeurs de :
 $P(\bar{A})$, $P_A(\bar{B})$ et $P_{\bar{A}}(\bar{B})$.

FIGURE 5.6 – Séance n°2 – Extraits du manuel *Indice*
Énoncés des exercices 4, 5 et 12

Les types de tâches (et leurs variations principales) présents dans ces trois exercices sont les suivants :

- **T-CalcProbaCond** (étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées/sont données dans l'énoncé)
- **T-LirArb** (étant donné qu'il est demandé d'indiquer la signification d'une certaine valeur numérique qui correspond à une probabilité située sur une branche de l'arbre)
- **T-LirArb** (étant donné qu'il est demandé de donner/préciser la valeur de la probabilité d'un événement présent sur une branche de l'arbre)
- **T-CompletArb** (à partir des règles de construction d'un arbre de probabilités)

Séance n°3

Nous n'avons pas pu assister à la séance n°3 du chapitre sur les probabilités conditionnelles. Celle-ci a été consacrée à la troisième et dernière partie du polycopié : « III. Indépendance de deux événements ». Le polycopié est à retrouver en annexe A.1.

Séance n°4

La séance n°4 du chapitre, dont le synopsis se trouve figure 5.7, commence par un « Mathoral » avec un exercice sur les nombres complexes. La séance se poursuit par un moment de « Questions Rapides », ces questions sont posées oralement par l'enseignante, toujours sur les nombres complexes.

L'enseignante pose ensuite aux élèves quelques questions pour les aider à mobiliser des éléments du cours sur les probabilités conditionnelles avant de

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

Thème principal de l'épisode	Commentaires
« Mathoral » sur les nombres complexes	X
« Questions Rapides » sur les probabilités conditionnelles	Détails dans ce qui suit
Travail en autonomie sur des extraits de sujets de baccalauréat issu du manuel	Détails dans ce qui suit

FIGURE 5.7 – Synopsis de la séance n°4 – *Probabilités conditionnelles*

les laisser travailler en autonomie sur des exercices : « *Vous avez des formules à connaître pour faire un exercice de bac sur les probabilités.* »

La première question est la suivante : « *si A et B sont deux événements lors d'une expérience aléatoire : exprimer la probabilité de A union B.* » L'enseignante passe dans les rangs, laisse quelques minutes aux élèves pour chercher la réponse puis elle interroge un élève qui répond oralement : « *$P(A)$ plus $P(B)$ moins $P(A \text{ et } B)$.* » L'enseignante demande à la classe quel est le nom de cette formule. Sans réponse de la part des élèves, elle modifie légèrement sa question et leur demande quel est le nom mnémotechnique de cette formule. Plusieurs élèves répondent alors en même temps qu'il s'agit de la formule des patates, l'enseignante ajoute : « *Oui la formule des patates, et le vrai nom c'est la formule de Poincaré.* »

La seconde question formulée par l'enseignante est la suivante : « *Comment tester que deux événements A et B sont indépendants ?* » ; l'enseignante procède de la même façon, elle laisse quelques minutes de réflexion aux élèves puis en interroge un. Cet élève répond : « *$P(A \text{ et } B)$ égale $P(A)$ fois $P(B)$.* »

L'enseignante pose une troisième et dernière question aux élèves : « *Récitez la formule des probabilités totales, qui me donnerait $P(B)$?* » L'enseignante passe dans les rangs et interroge une première élève qui dit ne pas savoir. L'enseignante interroge une autre élève qui répond oralement : « *$P(A \text{ inter } B)$... » , l'élève semble hésiter et l'enseignante l'encourage à poursuivre, l'élève fini par dire « *et $P(A \text{ barre et } B)$.* »*

La suite de la séance est consacrée à un travail en autonomie sur des extraits de sujets de baccalauréat. Les élèves se mettent en îlots de trois ou quatre, ils travaillent en autonomie et en groupe. L'enseignante se déplace dans la salle et envoie au fur et à mesure des élèves corriger au tableau les différentes questions. Le premier exercice traité correspond au sujet de bac n° D du manuel Indice dont l'énoncé est dans la figure 5.8.

La première question de cet exercice correspond au type de tâches **T-CompletArb**. C'est un type de tâches qui est en réalité un sous-type de tâches du type de tâches **T-ModéArb**. **T-ModéArb** a déjà été rencontré par les élèves dans

Sujet D

Capacités mises en œuvre

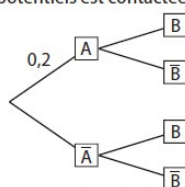
- Compléter un arbre de probabilités
- Calculer et utiliser des probabilités conditionnelles
- Tester l'indépendance de deux événements

Une revue est proposée en deux versions : papier ou électronique. Il est possible de s'abonner à une seule des deux versions ou de s'abonner simultanément aux deux versions. Un centre d'appel est chargé de démarcher les personnes figurant sur une liste de lecteurs potentiels. On admet que :

- lorsqu'un lecteur potentiel est contacté par le centre d'appel, la probabilité qu'il s'abonne à la version papier est égale à 0,2 ;
- s'il s'abonne à la version papier, la probabilité qu'il s'abonne aussi à la version électronique est égale à 0,4 ;
- s'il ne s'abonne pas à la version papier, la probabilité qu'il s'abonne à la version électronique est égale à 0,1. Une personne figurant sur la liste de lecteurs potentiels est contactée par le centre d'appel. On note :

- A : l'événement « la personne s'abonne à la version papier » ;
- B : l'événement « la personne s'abonne à la version électronique ».

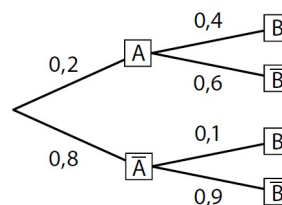
1. a. Reproduire et compléter l'arbre ci-contre.



- b. Donner la probabilité de \bar{B} sachant A et celle de \bar{B} sachant \bar{A} .
2. a. Calculer la probabilité que la personne contactée s'abonne à la version papier et à la version électronique.
- b. Justifier que la probabilité de l'événement B est égale à 0,16.
- c. Les événements A et B sont-ils indépendants ?
3. On suppose que la personne contactée s'est abonnée à la version électronique. Quelle est alors la probabilité qu'elle soit aussi abonnée à la version papier ?

Sujet D

1. a.



- b. $P_A(\bar{B}) = 0,6$ et $P_{\bar{A}}(\bar{B}) = 0,9$.
2. a. $P(A \cap B) = 0,2 \times 0,4 = 0,08$.
- b. $P(B) = P(A \cap B) + P(\bar{A} \cap B) = 0,08 + 0,08 = 0,16$.
- c. A et B ne sont pas indépendants :
 $P(A) \times P(B) = 0,2 \times 0,16 = 0,032 \neq P(A \cap B)$.
3. $P_B(A) = \frac{0,08}{0,16} = \frac{1}{2}$.

FIGURE 5.8 – Séance n°4 – *Extraits du manuel Indice*
 Énoncé du sujet D (à g.) – Corrigé issu du manuel du professeur (à d.)

l'exemple du cours lors de la séance n°2. Ils devaient alors reconnaître dans cet exemple qu'il fallait faire un arbre. Ici il n'y a pas l'étape de **T-ModéArb**, c'est directement **T-CompletArb**. Les variations présentes ici sont les suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les événements ont déjà été décrits et nommés dans le texte
- les données sont en fréquences relatives

L'élève qui est envoyée au tableau pour corriger cette première question reproduit et complète l'arbre mais ne donne aucune explication oralement. Ce que l'élève écrit au tableau se trouve dans la figure 5.9.

Les variations semblent avoir été prises en charge par l'élève. En effet, les variations de **T-CompletArb** sont ici les mêmes que celles présentes dans l'exemple traité lors de la séance n°2. Ce dernier avait été traité en classe sous la forme d'interactions entre l'enseignante et les élèves, ces variations ont donc déjà été proposées et traitées par les élèves. Cette question ne semble pas avoir posé de difficulté au reste de la classe. L'enseignante précise à l'oral : « *Faites attention*

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

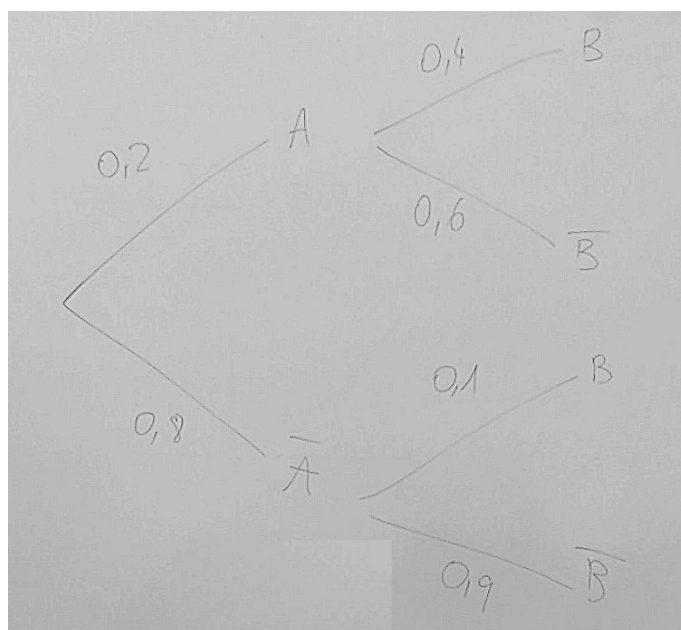


FIGURE 5.9 – Réponse à la première question du sujet n°D

aux mots dans l'énoncé. Qu'est-ce qu'ils nous demandent exactement?. » Une élève répond : « Reproduire et compléter. » L'enseignante insiste : « Donc on attend que toutes les branches soient reproduites et complétées. »

La seconde question (1.b) de cet exercice correspond au type de tâches **T-LirArb** qui est lui-aussi un type de tâches déjà rencontré (exercice 12 du manuel lors de la séance n°2). Ici les événements sont donnés et nommés dans la question, la variation présente est la suivante : *étant donné qu'il est demandé de donner/préciser la valeur de la probabilité d'un événement présent sur une branche de l'arbre*. Il s'agit là aussi d'une variation déjà proposée et traitée par les élèves lors de la séance n°2.

L'enseignante prend rapidement en charge la reconnaissance de la tâche et de la variation. En effet, elle indique oralement à l'ensemble des élèves : « C'est donner ceci cela et non pas calculer. Donc ne vous inquiétez pas s'il n'y a pas de calcul, c'est juste une lecture dans l'arbre. Ce n'est pas étonnant car on a dit "donner". » Tout en circulant dans la classe, l'enseignante s'adresse à un élève après avoir regardé les notes de celui-ci : « Ca tu le lis dans l'arbre, il n'y a pas de calcul, c'est une probabilité conditionnelle. Celle-ci se lit dans un arbre. » Une élève est envoyée au tableau et corrige cette deuxième question sans donner oralement d'explication supplémentaire.

La troisième question de cet exercice (2.a) correspond au type de tâches **T-CalcInter** qui a déjà été rencontré dans les exercices de la séance n°2. Les variations présentes ici sont les suivantes :

- à partir d'un énoncé en langage naturel

- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que la situation a déjà été représentée par un arbre
- les données sont en fréquences relatives

Le type de tâches **T-CalcInter**, étant donné que la situation a déjà été représentée par un arbre, a déjà été rencontré dans l'exemple du cours lors de la séance n°2 en tant qu'ingrédient de technique du type de tâches **T-CalcProbaCond**. Bien que dans ce dernier, l'enseignante avait pris en charge l'organisation au sens des sous-activités mathématiques, ici l'élève qui est envoyé au tableau prend en charge à son tour la reconnaissance, l'organisation et le traitement de la tâche. L'élève au tableau n'a pas de difficulté et écrit ce qui suit :

$$\begin{aligned}P(A \cap B) &= P(A) \times P_A(B) \\ &= 0,2 \times 0,4 \\ &= 0,08\end{aligned}$$

La quatrième question (2.b) de cet exercice correspond au type de tâches **T-CalcSimple** qui a déjà été rencontré dans les exercices de la séance n°2. Les variations présentes ici sont les suivantes :

- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que la situation a déjà été représentée par un arbre
- les données sont en fréquences relatives

L'enseignante envoie au tableau un élève pour corriger la quatrième question. Après avoir regardé des cahiers d'élèves, l'enseignante s'adresse à la classe, : « *je le vois pas assez dans vos cahiers, d'après la formule des probabilités totales* », il s'agit d'une aide à visée constructive.

Cette tâche a déjà été traitée par les élèves et notamment avec l'enseignante lors de la première question de l'exemple du cours. L'élève qui est envoyée au tableau corrige la question et écrit au tableau ce qui suit :

$$\begin{aligned}P(B) &= P(A \cap B) + P(\bar{A} \cap B) \\ &= 0,08 + P(\bar{A}) \times P_{\bar{A}}(B) \quad \text{d'après la formule des probas totales} \\ &= 0,08 + 0,1 \times 0,8 \\ &= 0,16\end{aligned}$$

La cinquième question (2.c) correspond au type de tâches suivant **T-EtudInd**. Les variations présentes ici sont les suivantes :

- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que la situation a déjà été représentée par un arbre
- les données sont en fréquences relatives

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

Etant donné que nous n'avons pas pu assister à la séance n°3 qui était sur le thème de l'indépendance de deux évènements nous ne pouvons pas apporter d'éléments chronologiques sur les tâches qui ont déjà été traitées ou non.

En passant près d'un élève et après avoir regardé son cahier, l'enseignante lui dit : « *Tu confonds évènements indépendants et évènements incompatibles. Et ici ils ne sont surtout pas indépendants. Indépendant c'est la leçon qu'on vient de faire.* » L'élève confond ici deux définitions :

« les évènements A et B sont incompatibles $\Leftrightarrow A \cap B = \emptyset$ »

et

« les évènements A et B sont indépendants $\Leftrightarrow P(A \cap B) = P(A) \times P(B)$. »

La caméra étant placée près d'un îlot de quatre élèves (que nous appellerons E1, E2, E3 et E4) nous avons pu recueillir des extraits de leurs échanges. Un élève (E2) s'adresse à un autre élève (E1) de son îlot :

E2 : « *Comment tu justifies que deux évènements ne sont pas indépendants ?* »

E1 : « *Tu fais A et B...* »

E2 : « *Ah oui, A et B pas égale à...* »

E3 : « *Après, il faut calculer avant* »

Une nouvelle élève est envoyée au tableau par l'enseignante pour corriger cette question. L'élève écrit sa réponse puis retourne s'asseoir, il n'y a pas d'interaction orale ni avec la classe ni avec l'enseignante.

La sixième et dernière question (3) de l'exercice : "*On suppose que la personne contactée s'est abonnée à la version électronique. Quelle est alors la probabilité qu'elle soit aussi abonnée à la version papier ?*", correspond au type de tâches **T-CalcProbaCond** qui a déjà été rencontré dans les exercices de la séance n°2. Les variations présentes ici sont les suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que la situation a déjà été représentée par un arbre (dans une précédente question)
- sans que l'évènement apparaisse sur l'arbre
- étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées
- les données sont en fréquences relatives

Un élève (E) s'adresse à l'enseignante (A) qui passe à côté de lui :

E : « *En fait la question 3 elle embrouille. Enfin elle est pas si compliquée, car la proba on l'a déjà !* »

A : « Ah ! On l'a déjà ? Tu es sûr ? »

L'élève (E) hésite, bafouille (inaudible) puis reprend : « Ça serait plutôt $P(A \text{ sachant } B)$? »

A : « Il y a une conditionnelle, oui c'est $P(A \text{ sachant } B)$. Et on l'a où dans l'arbre ? »

E : « Non on l'a pas. »

A : « Ça c'est classique, la dernière question c'est une conditionnelle mais pas lisible dans l'arbre, donc il faut la calculer. »

Face à ce type de tâche, du point de vue de l'élève, il y a une adaptation qui est l'évènement n'apparaît pas sur l'arbre. La sous-activité de reconnaissance attendue (quel calcul je dois faire ?) est à la charge de l'élève. Cette sous-activité est ici prise en charge par l'enseignante qui donne une aide à visée constructive.

Un second élève (P) du même îlot questionne l'enseignante (A) à propos toujours de cette question :

P : « Mais la question, le texte peut prêter à confusion, le "aussi" on peut le comprendre comme "et" ou "en plus" ? Ça peut être l'intersection ? »

L'enseignante accompagne la reconnaissance du bon type de tâches et répond à l'élève : « Non parce que c'est quand même une structure particulière de phrase. C'est même deux phrases et j'ai quelque chose qui est certain. On me dit que... » Il s'agit ici d'aides procédurales que l'enseignante adresse à l'élève. Une des adaptations ici est de reconnaître qu'il s'agit de calculer une probabilité conditionnelle. Une autre adaptation est que l'évènement n'apparaît pas sur l'arbre donc il y a une activité de reconnaissance à la charge de l'élève (quel calcul je dois faire).

Un troisième élève de l'îlot s'adresse à l'enseignante : « Ce genre de question on ne peut pas y répondre si on n'a pas fait le reste... », l'enseignante lui répond : « Mais si parce que si déjà tu traduis l'énoncé, tu dis $P(A \text{ sachant } B)$, tu as des points ! »

Un élève est envoyé au tableau pour corriger cette dernière question. Il n'y aura pas de remarque de l'enseignante ni d'interaction orale de cet élève avec la classe concernant la réponse écrite au tableau. L'élève écrit :

$$P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{0,08}{0,16} = 0,5$$

Les élèves commencent ensuite à travailler sur le sujet de baccalauréat du manuel n°A dont l'énoncé se trouve dans la figure 5.10.

L'enseignante indique aux élèves qu'il n'y aura pas assez de temps pour corriger cet exercice en détail. Etant donné qu'il s'agit d'un QCM seules les réponses seront données. Les types de tâches (et leurs variations principales) présents dans cet exercice sont les suivants :

Sujet A QCM

Capacités mises en œuvre

- Calculer et utiliser des probabilités conditionnelles
- Utiliser la formule des probabilités totales

Pour chacune des questions ci-dessous, une seule des trois propositions est exacte.

Un lecteur d'une bibliothèque est passionné de romans policiers et de biographies. Cette bibliothèque lui propose 150 romans policiers et 50 biographies. 40 % des écrivains de romans policiers sont français et 70 % des écrivains de biographies sont français.

Le lecteur choisit un livre au hasard parmi les 200 ouvrages.

- La probabilité qu'il choisisse un roman policier est :
a. 0,4 b. 0,75 c. $\frac{1}{150}$
- S'il choisit un roman policier, la probabilité que l'auteur soit français est :
a. 0,3 b. 0,8 c. 0,4
- La probabilité qu'il choisisse un roman policier français est :
a. 1,15 b. 0,4 c. 0,3
- La probabilité qu'il choisisse un livre d'un écrivain français est :
a. 0,9 b. 0,7 c. 0,475
- La probabilité qu'il ait choisi un roman policier sachant que l'écrivain est français est :
a. $\frac{4}{150}$ b. $\frac{12}{19}$ c. 0,3

Sujet A

On considère les événements :

- F : « le livre choisi est français » ;
- R : « le livre choisi est un roman policier ».

- $\frac{150}{200} = 0,75$: réponse **b.**
- $P_R(F) = 0,4$: réponse **c.**
- $P(R \cap F) = 0,4 \times 0,75 = 0,3$: réponse **c.**
- $P(F) = 0,4 \times 0,75 + 0,7 \times 0,25 = 0,475$: réponse **c.**
- $P_F(R) = \frac{0,3}{0,475} = \frac{12}{19}$: réponse **b.**

FIGURE 5.10 – Séance n°4 – Extraits du manuel *Indice*
Énoncé du sujet A (à g.) – Corrigé issu du manuel du professeur (à d.)

- **T-CalcSimple** (à partir d'un énoncé en langage naturel)
- **T-Assoc** (à partir d'un énoncé en langage naturel)
- **T-CalcInter** (à partir d'un énoncé en langage naturel)
- **T-CalcSimple** (à partir d'un énoncé en langage naturel)
- **T-CalcProbaCond** (à partir d'un énoncé en langage naturel et étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées)

Il y a une adaptation commune à chacune des questions, car les événements dont on cherche la probabilité sont décrits en langage naturel. Pour les deux premières questions il y a une activité de reconnaissance du type de tâches car il n'y a pas de calcul à faire ici, il faut uniquement associer des valeurs numériques de l'énoncé à des probabilités d'événements.

L'élève E1 de l'îlot que nous observons depuis le début, s'adresse aux élèves de son îlot (E2 et E4).

E1 : « Est-ce qu'il y a que les auteurs anglais ou français dans les questions ? »

E2 : « Oui, c'est soit français, soit anglais. »

E4 : « c'est toujours la même structure les énoncés. »

Un peu plus tard E1 interroge à nouveau les élèves de son îlot (ici E3).

E1 : « Est-ce qu'on connaît $P(P \text{ ou } F)$? Pour la question 3, est-ce qu'on connaît policier ou français ? Je crois pas ? »

E3 : « Utilise la formule $P(B \text{ sachant } A)$ égale $P(A \text{ inter } B)$ sur $P(A)$, tu la modifies pour avoir $P(A \text{ inter } B)$. T'as compris ? »

E1 : « Arrête d'utiliser des arbres, dis "et" ou "ou", c'est "et" ou "ou" ? »

E3 se lève et écrit la formule dans le cahier de E1 : « C'est ta formule. »

E1 : « Ah mais en fait je la connais ta formule. »

Ici l'élève E1 souhaite utiliser la formule $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ mais il se rend compte qu'il n'y a pas d'information permettant de calculer la probabilité de l'union. Une élève (E3) lui conseille alors d'utiliser plutôt la définition d'une probabilité conditionnelle et de transformer l'expression.

Séance n° 5

Nous ne proposons pas de synopsis pour la séance n°5 car nous n'avons assisté qu'à un seul et unique épisode d'une vingtaine de minutes, il s'agit d'un « Mathoral » sur les probabilités conditionnelles.

Un élève présente au tableau, devant le reste de la classe, la correction d'un exercice qu'il a préparé chez lui. Il s'agit du sujet de bac n°C dont l'énoncé est présenté figure 5.11.

Cette présentation orale donne lieu à une note, les critères de notations sont connus de l'élève et l'enseignante essaie d'intervenir le moins possible. L'exercice présenté ce jour là est un extrait de sujet de baccalauréat concernant les probabilités conditionnelles issu du manuel Indice.

La première question de cet exercice et une question de cours qui correspond au type de tâches **T-CourInd** et même à la capacité officiellement attendue des programmes C-DemInd. Il n'y a pas d'activité de reconnaissance étant donné qu'il s'agit d'une question de cours. En effet, il est inscrit "Restitution organisée de connaissances" qui correspond à un attendu des programmes connu par les élèves.

L'élève qui est au tableau prend la parole : « On doit démontrer que si $P(A \text{ inter } B)$ est égale à $P(A)$ fois $P(B)$ alors les événements \bar{A} et B le sont également. » L'enseignante propose de passer cette première question étant donné que la démonstration est dans la leçon et dans le manuel Indice. L'élève semble ne pas avoir entendu et commence la démonstration au tableau. L'enseignante l'interrompt mais l'élève indique qu'il veut terminer la démonstration entamée, il explique : « On sait que $P(\bar{A} \cap B)$ c'est égale à $P(B)$ fois $P_B(\bar{A})$. Comme $P_B(A)$ plus $P_B(\bar{A})$ est égale à 1 on a du coup (...) Comme A et B sont indépendants $P_B(A)$ est

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

Sujet C **ROC**

Capacités mises en œuvre

- Démontrer une propriété du cours
- Utiliser l'indépendance de deux événements
- Utiliser une loi binomiale

1. Restitution organisée de connaissances

Prérequis : Deux événements A et B sont indépendants pour la probabilité P si et seulement si $P(A \cap B) = P(A) \times P(B)$.

Soit A et B deux événements associés à une expérience aléatoire. Démontrer que, si A et B sont indépendants pour la probabilité P , alors les événements \bar{A} et B le sont également.

2. Application.

Chaque matin de classe, Stéphane peut être victime de deux événements indépendants :

- R : « il n'entend pas son réveil sonner » ;
- S : « son scooter, mal entretenu, tombe en panne ».

Il a observé que chaque jour de classe, la probabilité de R est égale 0,1 et que celle de S est égale à 0,05.

Lorsque qu'au moins l'un des deux événements se produit, Stéphane est en retard au lycée sinon il est à l'heure.

a. Calculer la probabilité qu'un jour de classe donné, Stéphane entende son réveil sonner et que son scooter tombe en panne.

b. Calculer la probabilité que Stéphane soit à l'heure au lycée un jour de classe donné.

c. Au cours d'une semaine, Stéphane se rend cinq fois au lycée. On admet que le fait qu'il entende son réveil sonner un jour de classe donné n'influe pas sur le fait qu'il l'entende ou non les jours suivants. Quelle est la probabilité que Stéphane entende le réveil au moins quatre fois au cours d'une semaine ?

Arrondir le résultat à la quatrième décimale.

Sujet C

1. Voir démonstration, manuel de l'élève p. 300.

2. a. $P(\bar{R} \cap S) = 0,9 \times 0,05 = 0,045$.

b. $P(\bar{R} \cap \bar{S}) = 0,9 \times 0,95 = 0,855$.

c. X , la variable aléatoire correspondant au nombre de fois où Stéphane entend son réveil, suit une loi binomiale de paramètres $n = 5$ et $p = 0,9$.

La probabilité que Stéphane entende le réveil au moins quatre fois est : $P(X = 4) + P(X = 5) = 0,91854$.

FIGURE 5.11 – Séance n°5 – Extraits du manuel Indice

Énoncé du sujet C (à g.) – Corrigé issu du manuel du professeur (à d.)

égale à $P(A)$ (...). » L'élève termine : « On arrive à $P(B \cap \bar{A})$ qui est égale à $P(B)$ fois $P(\bar{A})$ donc ça prouve l'indépendance de l'évènement \bar{A} et B »

$$\begin{aligned}
 P(B \cap \bar{A}) &= P(B) \times P(\bar{A}) \\
 P(A) + P(\bar{A}) &= 1 \\
 P(\bar{A}) &= 1 - P(A) \\
 P(A) &= P(A) \\
 P(B \cap \bar{A}) &= P(B) \times (1 - P(A)) \\
 P(B \cap \bar{A}) &= P(B) \times P(\bar{A})
 \end{aligned}$$

« On doit démontrer que si $P(A \text{ inter } B)$ est égale à $P(A)$ fois $P(B)$ alors les évènements \bar{A} et B le sont également

On sait du coup que $P(\bar{A} \text{ inter } B)$ c'est égal à $P(B)$ fois $P(\bar{A})$, comme $P_B(A)$ plus $P_B(\bar{A})$ est égal à 1 on a du coup (...)

Comme A et B sont indépendants $P_B(A)$ est égale à $P(A)$ (...)

On arrive à $P(B \cap \bar{A})$ qui est égal à $P(B)$ fois $P(\bar{A})$ donc ça prouve l'indépendance de l'évènement \bar{A} et B »

FIGURE 5.12 – Photo du tableau (à g.) et transcription du discours de l'élève (à d.) – Première question du sujet n°C

La seconde question de cet exercice correspond au type de tâches **T-CalcInter**

avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les évènements sont indépendants
- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que c'est la première question de l'exercice
- les données sont en fréquences relatives

Etant donné que nous n'avons pas pu assister à la séance n°3 qui était sur le thème de l'indépendance de deux évènements nous ne pouvons pas apporter d'éléments chronologiques sur les tâches qui ont déjà été traitées ou non pour ce thème.

L'élève écrit au tableau les données numériques de l'énoncé en même temps qu'il lit l'énoncé : « *Question 2, on sait que chaque matin de classe Stéphane peut être victime de deux évènements indépendants : R : son réveil n'a pas sonné, S : son scooter mal entretenu tombe en panne. On sait que la probabilité de R est de 0,1 , et que P(S) c'est égale à 0,05.* »

Il y a une première adaptation car l'évènement n'est pas donné et l'énoncé est en langage naturel, c'est une adaptation déjà rencontrée et traitée plusieurs fois. Les élèves ont intériorisé la méthode générale. En effet, l'élève au tableau reformule ici la question et prend l'initiative du changement de registre : « *On doit calculer la probabilité qu'un jour de classe donné Stéphane entende son réveil sonner et que son scooter tombe en panne. Donc on cherche $P(\overline{R} \cap S)$.* »

L'élève a pris en charge l'activité de reconnaissance du fait que les évènements sont indépendants, en effet, il explique : « *Comme on a démontré avant que si A et B sont indépendants alors \overline{A} et B sont indépendants, là on nous dit au tout départ que les deux évènements sont indépendants, donc on peut directement dire que $P(\overline{R} \cap S)$ c'est $P(\overline{R})$ fois $P(S)$.* »

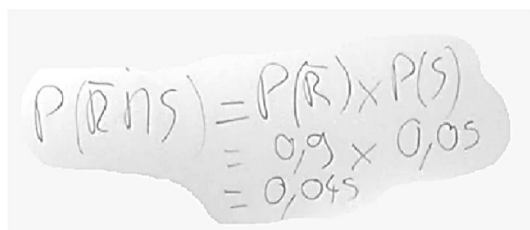
L'enseignante souhaite qu'il rajoute un élément de technologie pour justifier son discours, en effet, elle précise : « *Là tu es au tableau mais vous devez sinon écrire « car R et S sont indépendants donc " \overline{R} " et "S" aussi »* », c'est une aide à visée constructive.

Un autre élève de la classe (E2) s'interroge et pose une question à celui qui est au tableau : « *On ne dit pas que R et S sont indépendants, si ?* », l'élève au tableau lui répond : « *Si c'est écrit "chaque matin Stéphane peut être victime de deux évènements indépendants".* » L'élève E2 explique alors : « *J'ai fait avec un arbre et ça marche aussi* », sous-entendu qu'il ne s'est pas servi de l'indépendance pour répondre à la question. L'arbre ne se prête pas vraiment à cet exercice étant donné qu'il n'y a pas de probabilité conditionnelle donnée dans l'énoncé. La sous-activité de reconnaissance de la méthode à choisir pour faire ce calcul est à la charge de l'élève, même si la démonstration de la première question permet d'orienter ce

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

choix.

L'élève au tableau termine de répondre à cette question : « $P(\bar{R})$ c'est 0,9 donc ça nous donne 0,045. Donc la probabilité qu'un matin Stéphane entende son réveil et que son scooter tombe en panne est de 0,045. »


$$\begin{aligned} P(\bar{R} \cap S) &= P(\bar{R}) \times P(S) \\ &= 0,9 \times 0,05 \\ &= 0,045 \end{aligned}$$

« On cherche $P(\bar{R} \text{ inter } S)$, comme on a démontré avant que si A et B sont indépendants alors \bar{A} et B sont indépendants. Là on nous dit au tout départ que les deux événements sont indépendants, donc on peut directement dire que $P(\bar{R} \text{ inter } S)$ c'est $P(\bar{R})$ et $P(S)$

$P(\bar{R})$ c'est 0,9 donc ça nous donne 0,045. Donc la probabilité qu'un matin Stéphane entende son réveil et que son scooter tombe en panne est de 0,045 »

FIGURE 5.13 – Photo du tableau (à g.) et transcription du discours de l'élève (à d.) – Deuxième question du sujet n°C

La troisième (2.b) question de cet exercice est très similaire à la question précédente, elle correspond de nouveau au type de tâches **T-CalcInter** avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les événements sont indépendants
- sans que l'évènement n'ait été décrit ni nommé dans le texte
- les données sont en fréquences relatives

L'élève (E) lit la troisième question :

E : « Question b, on nous demande de calculer la proba que Stéphane soit à l'heure au lycée un jour de classe donné. »

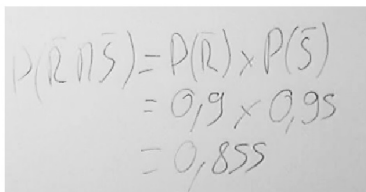
Il prend en charge l'adaptation d'interpréter la question qui est en langage naturel, et l'adaptation du fait que l'évènement n'ait pas encore été décrit ni nommé dans le texte. En effet, il explique :

E1 : « Cette fois-ci on cherche $P(\bar{R} \cap \bar{S})$. »

L'élève procède de la même manière qu'à la question précédente pour le calcul :

E : « Déjà les deux événements sont indépendants, pour les mêmes raisons que la question précédente. Donc là ils sont tous les deux aussi indépendants. (...) Donc la probabilité qu'un matin de classe Stéphane soit à l'heure au lycée est de 0,855. »

L'organisation, au sens des sous-activités mathématiques, est prise en charge ici par l'élève qui n'a qu'à suivre ce qu'il vient de faire dans la question précédente.


$$\begin{aligned} P(R|S) &= P(R) \times P(S) \\ &= 0,9 \times 0,95 \\ &= 0,855 \end{aligned}$$

« Question b, on nous demande de calculer la proba que Stéphane soit à l'heure au lycée un jour de classe donné, cette fois ci on cherche $P(\bar{R} \text{ inter } \bar{S})$,

déjà les deux événements sont indépendants, pour les mêmes raisons que la question précédente. Donc là ils sont tous les deux aussi indépendants (...) donc la probabilité qu'un matin de classe Stéphane soit à l'heure au lycée est de 0,855 »

FIGURE 5.14 – Photo du tableau (à g.) et transcription du discours de l'élève (à d.) – Troisième question du sujet n°C

La quatrième et dernière question (2.c) de cet exercice correspond au type de tâches suivant **T-CalcSimple** faisant intervenir une variable aléatoire binomiale : *Calculer la probabilité d'un évènement du type $P(X < a)$ où X est une variable aléatoire discrète suivant une loi binomiale étant donné que l'énoncé est en langage naturel.* Ce type de tâche, relatif au chapitre sur les variables aléatoires discrètes, a été traité dans l'exercice 29 lors de la séance n°2 mais il a aussi et surtout été traité en classe de première.

L'élève explique : « L'expérience consiste à répéter de façon identique et indépendante l'épreuve que Stéphane entend son réveil pendant 5 jours, chaque épreuve est une épreuve de Bernoulli dans laquelle le succès S : "Stéphane n'entend pas son réveil un matin" », il poursuit : « Donc la variable aléatoire X elle compte le nombre de matins où Stéphane n'entend pas son réveil durant les 5 jours. Le nombre de succès suit donc la loi binomiale de paramètres $n=5$ et $p=0,1$. On a la formule avec $P(X = k) = \dots$ »

L'élève revient ensuite au sujet de la question : « Comme on nous demande la probabilité d'entendre au moins 4 fois son réveil, donc on va prendre l'inverse, on va chercher la probabilité d'entendre pas son réveil 5 ou 4 fois, donc on va faire $P(X = 0) + P(X = 1)$. »

Un autre élève l'interrompt : « Pourquoi tu fais pour 0 et 1 ? », l'élève au tableau lui répond : « Parce qu'on cherche au moins 4 fois donc on prend l'inverse, le contraire. Au lieu de calculer $P(X = 2)$, $= 3$, $= 4$, le nombre de succès, on fait l'inverse. » Il insiste : « C'est plus simple, il y a moins de calculs à faire. »

L'enseignante complète cette explication et s'adresse à la classe : « Vous n'avez peut-être pas pris tous la même loi binomiale, ni la même variable aléatoire. Toi tu comptes le nombre de jour où.. ? », l'élève au tableau lui répond : « Où Stéphane n'entend pas son réveil. » L'enseignante s'adresse alors à E2 : « Peut-être que tu as choisi la variable aléatoire qui compte le nombre de jours où Stéphane entend son réveil ? Donc à ce moment-là quels sont les paramètres de X ? » L'élève au tableau répond : « $n = 5$ et $p = 0.9$. » L'enseignante conclut : « Exactement, et dans ce cas-là vous avez calculé $P(X = 4)$ et $P(X = 5)$ et normalement vous trouvez le même

résultat. »

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$
$$P(X \leq 1) = P(X=0) + P(X=1)$$
$$= 0,59049 + 0,38805$$
$$= 0,97854$$

« L'expérience consiste à répéter de façon identique et indépendante l'épreuve que Stéphane entend son réveil pendant 5 jours, chaque épreuve est une épreuve de Bernoulli dans laquelle le succès S : "Stéphane n'entend pas son réveil un matin",

donc la variable aléatoire X elle compte le nombre de matins où Stéphane n'entend pas son réveil durant les 5 jours. Le nombre de succès suit donc la loi binomiale de paramètres $n=5$ et $p=0,1$. On a la formule avec $P(X \text{ égal } k) = \dots$

Comme on nous demande la probabilité d'entendre au moins 4 fois son réveil, donc on va prendre l'inverse, on va chercher la probabilité d'entendre pas son réveil 5 ou 4 fois, donc on va faire $P(X \text{ égal } 0)$ plus $P(X \text{ égal } 1)$ »

FIGURE 5.15 – Photo du tableau (à g.) et transcription du discours de l'élève (à d.) – Quatrième question du sujet n°C

A la fin de cet épisode qui a eu lieu lors de la cinquième séance, nous avons proposé un test aux élèves de la classe, l'analyse de ce test et des productions d'élèves sont à retrouver au chapitre 6, partie 6.1 page 255.

Synthèse

Dans cette partie nous présentons une synthèse des résultats issus des analyses précédentes pour le thème des probabilités conditionnelles en classe de terminale scientifique au lycée.

Types de tâches et variations observées

Au cours de nos observations nous avons relevé, au sein des exemples et exercices traités en classe, les types de tâches et variations principales suivantes :

- **T-Assoc** (à partir d'un énoncé en langage naturel ; sans que les événements aient été explicitement nommés dans le texte)
- **T-LirArb** (étant donné qu'il est demandé d'indiquer la signification d'une certaine valeur numérique qui correspond à une probabilité située sur une branche de l'arbre ; étant donné qu'il est demandé de donner/préciser la valeur de la probabilité d'un événement présent sur une branche de l'arbre)
- **T-CompletTab** (à partir des règles de construction d'un tableau de probabilités)
- **T-CompletArb** (à partir d'un énoncé en langage naturel ; à partir des règles de construction d'un arbre)
- **T-CalcSimple** (à partir d'un tableau de probabilités ; à partir d'un énoncé en langage naturel ; faisant intervenir une variable aléatoire binomiale ; à partir

d'un énoncé en langage naturel et étant donné que c'est la première question de l'exercice ; sans que les événements aient été explicitement nommés dans le texte)

- **T-ModéArb** : en tant qu'ingrédient de technique pour T-CalcSimple (à partir d'un énoncé en langage naturel)
- **T-CalcInter** (à partir d'un arbre de probabilités ; à partir d'un énoncé en langage naturel et sans que les événements aient été explicitement nommés dans l'énoncé ; sans que les événements aient été explicitement nommés dans le texte ; étant donné que les événements sont indépendants)
- **T-CalcProbaCond** (faisant intervenir une variable aléatoire binomiale ; à partir d'un énoncé en langage naturel ; à partir d'un arbre de probabilités ; étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées/sont données dans l'énoncé ; sans que les événements aient été explicitement nommés dans le texte)
- **T-CourInd**
- **T-EtudInd**

En comparant avec la grande variété de praxéologies présentées dans le chapitre 4 page 49, nous n'avons pas observé les types de tâches suivants :

- T-CalcBar : calculer, à partir de la probabilité d'un événement, la valeur de la probabilité de son événement contraire
- T-ModéTab : modéliser une situation probabiliste décrite en langage naturel par un tableau
- T-Interp : interpréter le résultat
- T-Suite : établir une relation du type $r_{n+1} = f(r_n)$

Nous n'avons pas relevé les variations de types de tâches suivantes :

- Les données sont en fréquences naturelles
- L'arbre de probabilités est asymétrique ou incomplet
- L'arbre de probabilités illustre le passage de la $n^{\text{ième}}$ à la $(n + 1)^{\text{ième}}$ étape

Ces résultats sont à retrouver dans le tableau 5.16 qui suit.

Nous rappelons aux lecteurs que nous ne cherchons pas l'exhaustivité et que, n'ayant pas pu assister à toutes les séances de cette séquence, les résultats exposés ci-dessus ne reflètent pas entièrement la diversité des types de tâches et variations rencontrés par les élèves de cette classe de terminale scientifique.

	T-Assoc	T-CalcBar	T-CompleTab	T-CompleArb	T-ModéArb	T-ModéTab	T-LirArb	T-CalcProbaCond	T-CalcInter	T-CalcSimple	T-EtudInd	T-CourInd	T-Interp	T-Suite
A partir d'un énoncé en langage naturel	x	-	-	x	x	-		x	x	x	-			
Etant donné que c'est la première question de l'exercice	x		x	-	x	-			x	x		x		
Etant donné que les évènements ont déjà été décrits et nommés dans le texte	-	-	x	x	x	-		x	x	x	x		-	
Sans que les évènements aient été explicitement nommés dans le texte	x	-	-	-	-			x	x	x	-		-	
Les données sont en pourcentages	x	-			x	-		x	x	x	-			
Les données sont en fréquences relatives	-	-	x	x	-	-		x	x	x	x			
Les données sont en fréquences naturelles			-	-	-			-	-	-	-			
A partir des règles de construction d'un tableau de probabilités			x											
A partir des règles de construction d'un arbre de probabilités				x										
L'arbre est (ou sera) asymétrique				-	-									
L'arbre ne pourra être complet car certaines données ne sont pas connues					-									
L'arbre illustre le passage de la n-ième étape à la (n+1)-ième étape				-	-									
A partir des données de l'énoncé			-	x										
Donner/préciser la valeur de la probabilité d'un évènement qui est présent sur une branche de l'arbre							x							
Indiquer la signification (proba de quel évènement ?) d'une valeur située sur une branche de l'arbre							x							
Probabilités d'évènements simples	-	-												
Probabilités d'intersection d'évènements	x													
Probabilités conditionnelles	-	-												
Etant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées								x						
Sans que l'évènement n'apparaisse sur l'arbre								x						
Etant donné que la situation a déjà été représentée par un arbre ou un tableau (énoncé ou précédente question)								x	x	x	x			
Faisant intervenir une variable aléatoire binomial								x		x				
Etant donné que les évènements sont indépendants									x					
Etant donné qu'il s'agit d'une union d'évènements disjoints										-				

FIGURE 5.16 – "x" : variations de type de tâches relevées. "x" : variations présentes sous forme d'ingrédients de technique. "-" : variations absentes

Registres de représentations

Nous soulignons la place primordiale de l'arbre de probabilités dans ce cours, de la même façon que nous l'avions relevée dans l'analyse de manuel et des sujets de baccalauréat (chapitre 4). L'arbre de probabilités intervient à la fois comme aide procédurale dans l'activité de traitement ($P(A \cap B)$ apparaît dans l'arbre au bout d'une branche, ce qui permet de se rappeler comment calculer cette probabilité), comme intermédiaire dans l'activité de modélisation (à partir d'un énoncé en langage naturel l'arbre permet de traduire les données avant de commencer à répondre aux questions) et comme élément de la technique lors du calcul d'une probabilité. Nous reviendrons sur le rôle médiateur de l'arbre dans l'activité de modélisation dans la partie 5.2.3 page 190.

Éléments saillants pour cet enseignement

Compte-tenu des résultats du chapitre 4 page 49, nous relevons une grande variété de tâches proposées aux élèves mais aussi une certaine répétition. En effet, de nombreux types de tâches sont traités plusieurs fois au cours de cette séquence, par exemple une première fois en classe puis une seconde fois en autonomie, ou une première fois sur un exercice corrigé puis une seconde fois sur un sujet de type baccalauréat.

Chacune des variations proposées provoque dans l'activité de l'élève une (ou des) adaptation(s), par exemple : utiliser un intermédiaire, mettre en jeu plusieurs connaissances, utiliser une question précédente, choisir la méthode, etc. Ces adaptations sont généralement prises en charge par l'enseignante lors de la première rencontre. Du moins, l'enseignante accompagne ces variations, généralement dans les activités d'organisation et de reconnaissance. La reconnaissance de la méthode est très souvent à la charge de l'enseignante, notamment lorsqu'il s'agit d'une première rencontre, pour les élèves, avec le type de tâches en question.

Par ailleurs, nous relevons que certains éléments importants (aide ou indication par exemple) sont dits par l'enseignante à des groupes mais pas à toute la classe. En effet, les élèves travaillent beaucoup en autonomie et en groupe, ce qui favorise ce genre d'échanges (par exemple lors de la séance n°4).

5.2.2 Analyse du savoir enseigné pour le thème des lois continues

Dans cette partie nous présentons l'analyse du savoir enseigné en classe de terminale, pour le thème des lois continues. L'enseignante de mathématiques choisit de consacrer environ 12h à ce chapitre réparties en sept séances. Dans la figure 5.2 on retrouve l'enchaînement de ces six séances ainsi que le thème

principal abordé. Nous avons assisté aux six premières séances qui ont toutes été filmées.

A partir des notes prises durant les séances et des films nous avons produit les synopsis de ces séances. Nous avons également retranscrit les épisodes des séances filmées identifiés comme intéressants d'après les synopsis. Les épisodes "intéressants" sont ceux dans lesquels soit on voit apparaître un nouveau type de tâches (ou une nouvelle variation) soit un élève prend en charge la tâche (c'est notamment le cas des épisodes "Mathoral").

Pour ne pas trop alourdir cette partie nous présentons, pour chaque séance, le synopsis puis l'analyse des épisodes correspondant aux lignes grisées dans le synopsis. Ces analyses seront illustrées par des extraits de transcripts avant de tirer des conclusions sur ces enseignements dans une dernière partie de synthèse.

De la même façon que dans la partie précédente, nous prenons appui sur les résultats du chapitre 4 page 49. Il s'agit, de répertorier les types de tâches et techniques proposés aux élèves au cours des séances puis, pour chaque type de tâches relevé, d'identifier les variations qui sont à la charge des élèves ou prises en charge par l'enseignante. Pour chacune des tâches, nous nous intéressons aux sous-activité de reconnaissance, d'organisation et de traitement et nous nous demandons quelles sont les aides procédurales et constructives qui modifient l'activité des élèves.

Nous attirons l'attention du lecteur sur le fait que les élèves n'ont pas encore vu les primitives lorsqu'ils commencent le chapitre sur les lois continues. Lors des premières séances les intégrales sont donc nécessairement calculées soit à la calculatrice soit déterminées géométriquement, cette remarque est valable pour les séances n°1 et n°2.

Le plan du cours pour ce chapitre « Lois à densité » est le suivant :

- I. Lois de probabilités à densité
 - 1. Rappels
 - 2. Variables aléatoires continues, exemples
 - 3. Fonction de densité
 - 4. Espérance
- II. Loi Uniforme
 - 1. Définitions et propriétés
 - 2. Espérance
- III. Loi normale centrée réduite
 - 1. Définitions et propriétés
 - 2. Espérance mathématique

IV. Loïs normales

1. Définition
2. Déterminer l'espérance ou l'écart-type d'une loi normale
3. Intervalles à 1, 2 ou 3 sigmas

V. Loïs exponentielles

1. Définitions et propriétés
2. Espérance mathématique
3. Durée de vie sans vieillissement

Séance n° 1

n°	Thème principal de l'épisode	Commentaires
1	« Questions Rapides » sur le thème des fonctions	X
2	« Mathoral » exercice sur les complexes	X
3	Début du cours : « I. Loïs de probabilités à densité » : rappels et analogie avec les variables aléatoires discrètes puis activité introductive	Détails dans ce qui suit
4	Cours : « I.3. Fonction de densité » : exemple	Détails dans ce qui suit
5	Cours : « I.4. Espérance »	X

FIGURE 5.17 – Synopsis de la séance n°1 – *Loïs continues*

Détails à propos de l'épisode n°3 :

L'enseignante écrit le cours au tableau pour cette première partie : « I. Loïs de probabilités à densité. » Pour les loïs normales et exponentielles, qui feront l'objet des parties suivantes, le cours sera sous la forme d'un polycopié à trous, comme pour le chapitre des probabilités conditionnelles.

Le cours commence par deux exemples sur les variables aléatoires discrètes (voir en annexe B.3 des extraits du tableau). Dans le premier exemple la variable aléatoire prend un nombre fini de valeurs, dans le second elle prend ses valeurs dans \mathbb{N} . Ces deux exemples permettent à l'enseignante d'introduire la notion de variables aléatoires continues, elle explique aux élèves : « Cette année on va s'intéresser à des variables aléatoires qui ne seront plus seulement discrètes, on va aller vers des variables aléatoires continues. C'est-à-dire qu'elles ne vont plus prendre des valeurs en nombre fini, ni même des valeurs infinies mais bien séparées comme dans l'ensemble des nombres entiers, elles vont prendre des valeurs qui sont des

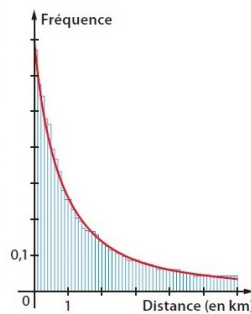
nombre réels dans des intervalles. »

Le cours se poursuit par une activité introductive issue du manuel de la classe dont l'énoncé se trouve dans la figure 5.18.

1 L'éco-point

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 $P(0 \leq X < 1)$?
 - d. Pour tout décimal t appartenant à $\{0; 0,1; 0,2; \dots; 5,8; 5,9\}$, 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 $\int_0^t f(x) dx$. Estimer sa valeur.
- b. Soit t un nombre réel appartenant à $[0; 6]$. Exprimer $P(0 \leq X < t)$ à l'aide d'une intégrale.

Activité 1 L'éco-point

Cette activité introduit la notion de loi continue sur un exemple très simple permettant de visualiser la courbe d'une densité de probabilité et le calcul de probabilités du type $P(0 \leq X \leq t)$ en utilisant l'aire « sous la courbe ».

La définition en cours devrait ainsi être plus facile.

1. a. L'aire est 0,077 et la base est 0,1 donc la hauteur est 0,77.
 - b. La somme des aires des 60 rectangles de l'histogramme vaut 1.
 - c. Sur le graphique la probabilité $P(0 \leq X < 1)$ est représentée par la somme des aires des 10 premiers rectangles.
 - d. Pour tout décimal t appartenant à $\{0; 0,1; 0,2; \dots; 5,8; 5,9\}$ la somme des aires des rectangles dont la base est sur $[0; t]$ représente la probabilité $P(0 \leq X \leq t)$.
2. a. $\int_0^t f(x) dx$ est l'aire exprimée en unité d'aire du domaine limité par la courbe représentative de la fonction f , les axes de coordonnées et la droite d'équation $x = t$. On peut estimer sa valeur à $P(0 \leq X < t)$, c'est-à-dire à 1.
 - b. Pour tout réel t appartenant à $[0; 6]$:

$$P(0 \leq X < t) = \int_0^t f(x) dx.$$

FIGURE 5.18 – Séance n°1 – Extraits du manuel Indice
Énoncé de l'activité (à g.) – Corrigé issu du manuel du professeur (à d.)

L'enseignante laisse les élèves travailler en autonomie sur ce problème, elle passe dans les rangs et interroge les élèves au fur et à mesure. L'enseignante explique : « Voilà le premier exemple de variable aléatoire continue qui prend ses valeurs dans un intervalle réel. »

Les élèves ont du mal à commencer l'exercice, il s'agit de nouvelles tâches pour eux. L'enseignante apporte des aides procédurales : « Vous avez l'histogramme à droite en bleu, l'axe des abscisses est gradué de kilomètres en kilomètres, qu'est-ce qui est égale à la fréquence d'une classe ? Si je regarde une barre de cet histogramme la fréquence est donnée par la hauteur. » L'enseignante fait relire la première question par un élève puis elle précise : « C'est l'aire qui donne la proportion de personnes qui habitent dans un intervalle de distance donnée, c'est ça un histogramme, c'est la différence avec un diagramme en barres ou en bâtons. »

Un élève demande à l'enseignante : « En fait il faut trouver quelque chose qui nous donne l'aire de 7,7% ? », l'enseignante acquiesce. Un élève est envoyé au ta-

bleau pour répondre à cette première question, il écrit :

$$\begin{aligned}P \times L &= a \\a &= 0,077 \text{ et } L = 0,1 \\P &= \frac{0,077}{0,1} = 0,77\end{aligned}$$

L'enseignante lui demande de donner des explications oralement, il dit : « *On sait que 7,7% c'est l'aire du rectangle, on connaît la largeur qui est de 0,1. Il faut diviser l'aire du rectangle par la largeur pour avoir la hauteur.* »

La seconde question pose des difficultés aux élèves, l'enseignante les aide en leur donnant des indications : « *L'aire d'un rectangle c'est la fréquence des habitants qui habitent à cette distance, du coup la somme des aires ? Que vaut la somme des aires comme les aires ce sont des fréquences ?* »

Un élève va corriger cette seconde question au tableau, il écrit « 100% » et dit : « *Parce que la somme de toutes les aires des rectangles c'est 100% parce que c'est la totalité de l'expérience* », l'enseignante lui demande des précisions : « *oui, c'est la somme des fréquences, et en nombre décimal 100 % c'est égal à quoi ?* », l'élève répond « 1 » et complète le tableau :

$$100\% = 1$$

L'enseignante laisse les élèves chercher la troisième question (1.c), tout en passant dans les rangs elle donne des indications : « *La probabilité que X soit compris entre 0 et 1, comment ça se traduit sur le graphique ?* », elle interroge une élève qui répond : « *C'est représenté par les aires des 10 premiers rectangles compris entre 0 et 1* », l'enseignante ajoute : « *Il manque un mot bien précis, que fait-on avec les aires des 10 premiers rectangles ?* », l'élève répond : « *La somme.* »

L'enseignante prend en charge la résolution au tableau de cette question : « *Si je prends tous les rectangles je devrais obtenir 1 et si je prends seulement la somme des aires des 10 premiers rectangles je devrais obtenir quelque chose qui approche la probabilité que X soit compris entre 0 et 1* » et elle écrit :

$$(c) P(0 \leq x < 1) \rightarrow \text{somme des aires des 10 premiers rectangles}$$

L'enseignante donne des aides procédurales pour la quatrième question (1.d) : « *On prend tous les rectangles qui correspondent à l'intervalle $[0, t]$, on prend la somme des aires et on vous demande de décrire à quoi ça correspond, en termes de probabilités.* » Une élève va corriger cette question au tableau et écrit :

$$(d) P(0 \leq x < t) \rightarrow \text{somme des aires des rectangles sur } [0, t]$$

L'élève explique : « *Sur le même principe, la somme des aires, enfin des rectangles, sur $[0, t]$ c'est l'ensemble des probabilités entre 0 et t* », l'enseignante la reprend : « *C'est LA probabilité que X soit entre 0 et t, la probabilité va être*

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

représentée graphiquement par la somme des aires des rectangles correspondants. »

L'enseignante apporte des éléments de synthèse concernant les premières questions et fait remarquer aux élèves le lien entre cette activité et la méthode des rectangles qu'ils ont vue dans le chapitre sur le calcul d'intégrales : « *On a découpé notre intervalle $[0, 6[$ en rectangles. Si j'étais en terminale ça me rappellerait quelque chose cette histoire de découper en rectangles un intervalle... Découper un intervalle en petits intervalles, et sur chaque intervalle je considère un rectangle et l'aire de ce rectangle. C'est un moyen de se ramener aux variables aléatoires discrètes, je vais prendre des tranches d'intervalles. Je ne prends plus des intervalles de 0,1 mais d'amplitude 0,01. En rajoutant des décimales je vais pouvoir augmenter le nombre de rectangles, autrement dit des rectangles de base de plus en plus resserrée. Et je vais pouvoir faire tendre la base de ces rectangles vers 0, et finalement la somme des aires de ces rectangles va pouvoir se rapprocher, en multipliant le nombre de rectangles, de l'aire sous la courbe rouge. Les premières questions c'est l'algorithme des rectangles comme on le fait dans le chapitre d'intégration. »*

Les élèves commencent à chercher la question suivante (2.a) et l'enseignante prend rapidement en charge l'activité de reconnaissance de la tâche : « *C'est quoi la définition d'une intégrale, c'est ce qu'ils vous demandent, on commence par l'aire.* » L'enseignante interroge ensuite plusieurs élèves jusqu'à avoir la réponse qu'elle attendait, réponse issue du cours sur l'intégration « *L'aire du domaine sous la courbe de f entre 0 et 6.* »

Un élève va écrire au tableau la correction de la dernière question (2.b) :

$$\int_0^t f(x) dx$$

L'enseignante termine avec des remarques générales sur cette activité : « *Pour une variable aléatoire continue on va être amené à étudier des aires que l'on peut approcher par une somme d'aires de rectangles et qu'on peut tenter de calculer par des intégrales.* » L'enseignante fait des remarques concernant le lien entre intégrale et variables aléatoires à densité : « *Déjà pour calculer ce genre d'intégrale il va falloir connaître la fonction f , quelle est la fonction f pour une variable aléatoire donnée? Quelle est la fonction f dont la courbe va correspondre à l'histogramme? Elle vous sera donnée en général en terminale.* »

Conclusion :

Au delà d'un premier exemple de variable aléatoire continue, cette activité est l'occasion de faire découvrir aux élèves le lien entre calcul d'aire, représentation graphique et probabilité. Ce lien sera omniprésent tout le long du chapitre, à la fois dans le cours et dans le discours de l'enseignante. Les élèves ont déjà vu en classe le chapitre sur l'intégration, ils connaissent donc déjà le lien entre intégrale et calcul d'aire "sous la courbe", c'est aussi l'occasion de revoir la méthode des rectangles

pour le calcul d'intégrales. Cette activité permet à l'enseignante d'introduire la notion de loi continue et de fonction de densité.

Détails à propos de l'épisode n° 4 :

L'exemple suivant relève d'une situation de la vie réelle et sert de fil conducteur pour la suite du cours.

Une entreprise fabrique des disques durs. On considère l'expérience aléatoire qui consiste à choisir au hasard un disque dur. On définit la variable aléatoire X qui à chaque disque dur associe sa durée de vie en heure. Cette variable aléatoire peut prendre toutes les valeurs réelles de l'intervalle $[0; +\infty[$. X est une variable aléatoire continue.

Après avoir donné la définition d'une densité de probabilités et celle de la probabilité d'un événement $[a, b]$ dans le cas d'une variable aléatoire continue, l'enseignante illustre ces nouvelles définitions sur l'exemple précédent. Elle commence par écrire au tableau :

$$P(5000 \leq X \leq 20000) =$$

Et elle explique : « Dans l'exemple précédent, la probabilité que X soit compris entre 5000 et 20000 correspond à ce que la durée de vie du disque dur soit comprise entre 5000 et 20000 heures, on va faire un petit dessin ». Pour représenter la fonction de densité de cette durée de vie dont la loi n'est pas explicitement donnée, l'enseignante dessine une courbe de fonction exponentielle décroissante mais elle ne dit pas aux élèves qu'il s'agit d'une fonction exponentielle ou d'une loi exponentielle.

A partir de la représentation graphique de la fonction de densité elle explique l'interprétation géométrique de cette probabilité : « L'axe des abscisses doit correspondre aux valeurs prises par la variable aléatoire que vous considérez, ici ça va correspondre aux nombres d'heures. Donc on place 5000 et 20000, puis la droite d'équation $x = 5000$ et $x = 20000$. On hachure l'aire du domaine sous la courbe de f sur l'intervalle $[5000; 20000]$, ça vous donne la probabilité qu'on a écrite juste au-dessus » et elle complète le tableau comme on peut le voir dans la figure 5.19.

Elle conclut quant à la fonction f : « c'est la fonction qui par son intégrale vous donne les probabilités que X soit dans un intervalle donné. »

Conclusion :

L'enseignante accompagne le passage "calcul de probabilité" vers "calcul d'intégrale" et fait travailler les élèves sur le passage de l'un vers l'autre. Cette activité de traitement sera par la suite à la charge des élèves (reconnaissance des bornes, hachurer sur le graphique la probabilité recherchée, etc.). Il s'agit de la première rencontre avec le type de tâches **T-CalcSimple** pour les élèves, bien que le calcul n'a pas été terminé.

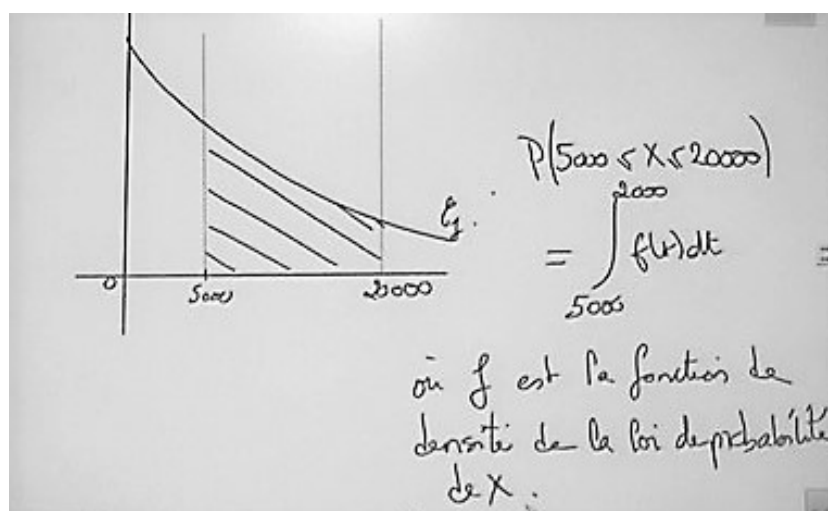


FIGURE 5.19 – Extraits du tableau lors de la séance n°1 « I.3. Fonction de densité »

Séance n°2

n°	Thème principal de l'épisode	Commentaires
1	« Questions Rapides » sur les premières définitions du chapitre de lois continues : T-Vériff et T-CalcSimple	Détails dans ce qui suit
2	« Mathoral » ex2 du manuel sur les lois continues : T-CalcSimple	Détails dans ce qui suit
3	« Mathoral » ex37 du manuel sur les lois continues : T-Vériff , T-CalcSimple et T-CalcEsp	Détails dans ce qui suit
4	Cours : « I.4. Espérance » interprétation de l'espérance puis un exemple : T-Vériff , T-CalcSimple et T-CalcEsp + T-Interp	Détails dans ce qui suit
5	Cours : « II. Loi Uniforme » activité algorithmique introductive, définitions, propriétés et espérance	X
6	Cours : « II. Loi Uniforme » premier exemple : T-DonF , T-CalcSimple + T-Interp et T-CalcEsp + T-Interp	Détails dans ce qui suit

FIGURE 5.20 – Synopsis de la séance n°2 – Lois continues

Détails à propos de l'épisode n°1 :

La séance commence par un moment de « Questions Rapides » sur le cours de la fois précédente. L'enseignante trace une allure de courbe et écrit au tableau :

Soit f une fonction de densité définie sur $[3; 10]$

Elle demande aux élèves de rappeler les propriétés d'une fonction de densité.

Il s'agit ici d'une tâche proche du type de tâches **T-Vérif** : vérifier qu'une fonction donnée est bien une densité de probabilité. Un premier élève répond : « *Il faut qu'elle soit continue et positive* », l'enseignante lui dit qu'il manque quelque chose et un autre élève complète : « *Il faut que l'aire sur l'intervalle $[3; 10]$ soit égale à 1.* »

L'enseignante demande à cet élève de traduire cette aire en terme d'intégrale, ce à quoi l'élève répond : « *L'intégrale de 3 à 10 de $f(x)dx$* », l'enseignante écrit au tableau :

$$1 = \int_3^{10} f(x) dx$$

L'enseignante demande ensuite aux élèves : « *Si on considère désormais X une variable aléatoire continue qui aurait pour densité cette fonction, comment peut-on alors traduire ça en termes de probabilités pour la variable aléatoire X ?* », l'enseignante fait travailler les élèves sur le passage "calcul d'intégrale" vers "calcul de probabilité".

Un élève répond : « *$P(X \text{ appartient à } [3; 10])$* », il prend en charge l'activité de traitement de la tâche (traduire une intégrale en termes de probabilité) et l'enseignante écrit au tableau :

$$= P(X \in [3, 10]) = P(3 \leq X \leq 10)$$

La question suivante est écrite au tableau :

Sachant que $P(X < 5) = 0,6$ calculer les probabilités suivantes :
 $P(X = 5)$, $P(X \leq 5)$, $P(X > 5)$

Cette question correspond au type de tâches **T-CalcSimple** avec les variations suivantes :

- étant donné que la densité n'a pas été donnée dans l'énoncé
- calculer $P(X > a)$ étant donné qu'on connaît $P(X < a)$

Les élèves ont un peu de temps pour chercher. C'est la première fois qu'ils rencontrent ce type de tâches pour lequel on attend ici un résultat numérique. L'enseignante prend en charge la reconnaissance de la méthode, elle dit aux élèves : « *Si vous êtes paumés vous traduisez tout en termes d'intégrales, on peut s'aider avec des aires.* »

Une élève volontaire passe au tableau, l'enseignante lui demande de commencer par hachurer sur le graphe de la fonction f la probabilité $P(X < 5) = 0,6$ ce qu'elle fait correctement, puis elle explique comment répondre à la première question : « *$P(X = 5)$, l'intégrale va de 5 à 5 donc ça fait 0* » et écrit au tableau :

$$P(X = 5) = 0$$

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

Pour la deuxième question, l'élève explique : « Pour $X \leq 5$ ça change pas parce qu'on additionne 0,6 et 0 » et elle écrit au tableau :

$$P(X \leq 5) = 0,6$$

Pour la troisième question, l'élève explique : « Ensuite pour $X > 5$, sachant que l'intégrale au complet fait 1 et que l'on connaît tout ce qui est inférieur, ça fait $1 - 0,6$ ça fait 0,4 », l'enseignante demande : « Comment ça s'appelle ? Les deux événements comment sont-ils ? », ce à quoi l'élève répond : « Contraires » puis elle écrit au tableau :

$$P(X > 5) = 0,4$$

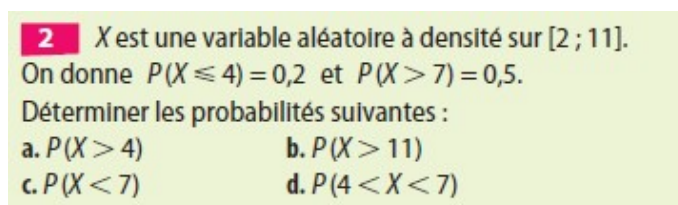
Conclusion :

Les élèves rencontrent **T-Veriff** pour la première fois. L'enseignante fait travailler les élèves sur le passage "aire" vers "calcul d'intégrale" puis sur le passage "calcul d'intégrale" vers "calcul de probabilité".

Dans l'exemple proposé les élèves rencontrent également pour la première fois **T-CalcSimple** avec un véritable calcul, sans fonction de densité, juste en utilisant les propriétés d'une fonction de densité. L'enseignante insiste sur la représentation graphique et la première étape du traitement, à savoir, reconnaître les bornes et hachurer sur le graphique la probabilité recherchée.

Détails à propos de l'épisode n°2 :

Un élève présente au tableau sa correction d'un exercice du manuel dans le cadre d'un « Mathoral ». L'énoncé de cet exercice se trouve dans la figure 5.21.



2 X est une variable aléatoire à densité sur $[2 ; 11]$.
On donne $P(X \leq 4) = 0,2$ et $P(X > 7) = 0,5$.
Déterminer les probabilités suivantes :

a. $P(X > 4)$	b. $P(X > 11)$
c. $P(X < 7)$	d. $P(4 < X < 7)$

FIGURE 5.21 – Séance n°2 – Énoncé de l'exercice 2
Extrait du manuel *Indice*

Cet exercice ressemble très fortement à ce qui vient d'être fait lors des questions rapides. L'élève commence par tracer la courbe de la fonction f et représente les probabilités recherchées. Il le dit lui-même, il s'inspire de ce qui vient d'être fait. En effet, l'enseignante a beaucoup insisté avec « le dessin » dans la procédure de résolution de l'exercice précédent.

L'élève commence par écrire au tableau les données de l'énoncé, puis répond oralement à la première question (a.) : « Pour $P(X > 4)$ on sait que $P(X < 4)$

c'est 0,2 donc j'ai fait 1 - 0,2, ça fait 0,8 » L'enseignante attend une justification particulière pour ce type d'évènements et demande à l'élève : « *Les deux évènements ils sont comment l'un par rapport à l'autre ?* », l'élève répond : « *Inverses* », l'enseignante le reprend : « *On ne dit pas inverses mais contraires* ».

Pour la seconde question (b.), l'élève explique : « *P(X > 11) j'ai dit que c'était l'évènement impossible donc ça fait 0* » et il écrit au tableau :

$$P(X > 11) = 0$$

Pour la troisième question (c.) , l'élève explique : « *P(X < 7) c'est pareil que le premier c'est l'évènement contraire* » et il écrit au tableau :

$$P(X < 7) = 0,5$$

Enfin, pour la dernière question (d.), il explique : « *On sait que P(X > 4) c'est 0,8 et P(X < 7) c'est 0,5 donc ça nous donne 0,3. J'ai fait 0,8 moins 0,5* » il écrit au tableau :

$$P(4 < X < 7) = 0,3$$

Conclusion :

Il s'agit d'une nouvelle rencontre pour les élèves avec **T-CalcSimple**, cela semble désormais automatique pour eux de commencer un exercice de ce type avec une représentation graphique. Lors du calcul d'une probabilité quand celle-ci découle de propriétés telles que $P(\bar{A}) = 1 - P(A)$, l'enseignante attend des éléments technologiques dans une réponse à ce type de tâche, comme par exemple : "*car les évènements sont contraires*".

Détails à propos de l'épisode n°3 :

Un élève vient au tableau pour un « Mathoral », il va présenter sa correction de l'exercice dont l'énoncé se trouve dans la figure 5.22.

37 1. a. Représenter dans un repère orthonormé d'unité 4 cm la fonction f définie sur $[0 ; 1]$ par :

$$f(x) = x + \frac{1}{2}.$$

b. Montrer que f est une fonction de densité de probabilité sur $[0 ; 1]$.

2. X est la variable aléatoire continue sur $[0 ; 1]$ dont la loi a pour densité de probabilité la fonction f définie à la question 1. Calculer :

a. $P(X < 0,25)$	b. $P\left(X > \frac{1}{3}\right)$
c. $P(0,1 < X < 0,7)$	d. $E(X)$

FIGURE 5.22 – Séance n°2 – Énoncé de l'exercice 37
Extrait du manuel *Indice*

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

L'élève commence par représenter graphiquement la fonction de densité qui est donnée dans l'énoncé.

La seconde question (1.b) correspond au type de tâches **T-Vériff** avec la variation suivante : à partir de la représentation graphique de la fonction de densité. La reconnaissance de la méthode et la sous-activité d'organisation pour ce type de tâches ont déjà été prises en charge au début de la séance lors des questions rapides, reste seulement ici le traitement qui est réellement à la charge de cet élève. La variation à *partir de la représentation graphique de la fonction de densité* provoque une adaptation à la charge de l'élève, il doit ici utiliser des techniques de géométrie.

L'élève présente sa correction : « *Pour montrer que f est une fonction de densité il faut montrer que la fonction est continue et positive et que l'intégrale est continue sur $[0, 1]$* », l'élève s'embrouille un peu et l'enseignante (A) lui fait relire ses notes. Il reprend :

E : « *C'est ce que j'ai écrit... l'intégrale de f est égale à 1. Donc f est continue et positive sur $[0, 1]$ car c'est une fonction croissante et affine. De plus l'intégrale de 0 à 1 est égale à 1, j'ai pris la primitive... mais c'était compliqué.* »

A : « *Surtout qu'on n'a pas vu ça encore pour le moment ! Je vous ai fait sentir qu'il y avait un lien entre le calcul d'intégrale et ce que l'on appelle une primitive mais on n'a pas encore fait le chapitre.* »

L'élève a voulu calculer l'intégrale de f sur son domaine de définition à l'aide d'une primitive, or le chapitre sur les primitives n'a pas encore été vu en classe, même si la notion a été un peu abordée dans le chapitre sur l'intégration. Nous précisons ici que l'élève en question est un redoublant, il a donc déjà eu connaissance du chapitre sur les primitives l'an passé. L'enseignante attend ici des élèves qu'ils calculent l'aire recherchée de façon géométrique. L'élève (E) au tableau est bloqué et l'enseignante (A) lui apporte une aide procédurale :

A : « *Calculer l'aire, quelle est cette forme ?* »

E : « *Un trapèze* »

A : « *Tu nous rappelles la formule rapidement* »

L'élève écrit finalement au tableau la formule qui donne l'aire d'un trapèze puis il applique cette formule à l'exercice et finit par obtenir une aire égale à 1.

Les trois questions suivantes correspondent au type de tâches **T-CalcSimple** avec les variations suivantes :

- étant donné que la variable aléatoire est identifiée et que sa densité a été donnée dans l'énoncé
- à partir de la représentation graphique de la fonction de densité

Ce type de tâches a déjà été proposé et traité par les élèves, mais ici la variation à partir de la représentation graphique de la fonction de densité provoque une adaptation dans l'activité de l'élève, il doit à nouveau utiliser des techniques de géométrie. Cet élève explique qu'il a essayé de faire avec une primitive mais sans succès. L'enseignante lui demande donc d'utiliser la formule de l'aire d'un trapèze qui est écrite au tableau. Elle précise pour l'ensemble de la classe : « De toute façon dans cet exercice, ce ne sont que des aires de trapèzes ». L'élève donne les résultats numériques oralement (il s'aide d'une calculatrice pour finir les calculs).

La dernière question (2.d) correspond au type de tâches **T-CalcEsp**, étant donné que la variable aléatoire est identifiée et que sa densité a été donnée dans l'énoncé, à partir de la représentation graphique de la fonction de densité.

C'est la première fois que ce type de tâches est traité par les élèves. L'enseignante prend en charge l'organisation et demande à l'élève au tableau de rappeler la formule de l'espérance : « Tu écris la formule en toutes lettres au tableau ça c'est nouveau, elle est définie comment ? Pour cette question je vous avais dit de faire avec la calculatrice » L'élève explique avoir fait le calcul avec sa calculatrice en utilisant la fonctionnalité qui donne la valeur d'une intégrale (ce qu'avait recommandé de faire l'enseignante) et il écrit au tableau :

$$E(X) = \int_0^1 xf(x)dx = 7/12 = 0,583$$

Pour finir l'enseignante reprend l'élève sur sa justification de la positivité, pour rappel il avait dit qu'elle était croissante et affine.

Conclusion :

Il s'agit d'une nouvelle rencontre, pour les élèves, avec les types de tâches **T-VerifF** et **T-CalcSimple** mais cette fois-ci à partir de la représentation graphique d'une fonction. L'enseignante prend en charge cette variation qui a vraisemblablement posé des difficultés aux élèves dans l'activité de traitement car pour calculer l'intégrale il faut utiliser des propriétés géométriques.

C'est la première rencontre avec **T-CalcEsp**. La fonction de densité est suffisamment simple pour permettre le calcul de l'espérance sans calculatrice (à l'aide d'un calcul à la main), mais comme le chapitre sur les primitives n'a pas encore été vu, seule la technique instrumentée est faisable.

Détails à propos de l'épisode n° 4 :

La séance continue par un moment de cours. L'enseignante explique oralement à la classe comment interpréter l'espérance : « L'espérance s'interprète comme la valeur moyenne probable de X si je répète l'expérience un suffisamment grand nombre de fois. » Le cours se poursuit par un exemple dont voici l'énoncé :

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

Une entreprise produit des dalles, soit X la variable aléatoire égale à la production en tonnes, X prend ses valeurs dans l'intervalle $[0; 20]$, la densité de sa loi de probabilités est définie sur $[0; 20]$ par $f(x) = 0,015x - 0,00075x^2$

1. Vérifions que f est une fonction de densité.
2. Calculons une valeur approchée de la probabilité de $P(E)$ où E est l'évènement « la production soit supérieure à 12 tonnes. »
3. Calculer l'espérance de X et l'interpréter.

La première question correspond au type de tâches **T-VérifF** qui a déjà été traité par les élèves dans le cas des Questions Rapides et qui vient d'être vu dans l'exercice précédent avec une fonction dont la représentation graphique était connue. L'enseignante prend en charge très rapidement la reconnaissance de la méthode à suivre, l'organisation et le traitement de cette question, elle demande aux élèves : « *Il s'agit de quel type de fonction? Comment on peut justifier qu'elle est continue et positive? Comment on fait pour avoir son signe? Quel est l'axe de symétrie de cette parabole?* » S'ensuit toute l'étude de la fonction f , l'enseignante fait s'entraîner les élèves à tracer sur leur calculatrice le graphe d'une telle fonction. Étant donné qu'ils n'ont pas encore vu les primitives l'enseignante leur demande de vérifier, au moyen de leur calculatrice, que l'intégrale de f sur $[0; 20]$ vaut bien 1.

La deuxième question correspond à un type de tâches déjà rencontré : **T-CalcSimple** avec les variations suivantes :

- étant donné que la variable aléatoire est identifiée et que sa densité a été donnée dans l'énoncé
- à partir d'un énoncé en langage naturel

L'énoncé est en langage naturel et cette adaptation est à la charge des élèves, comme c'est la première fois, elle est ici prise en charge par l'enseignante : « *Face à ce genre d'énoncé vous devez rapidement être capable de traduire en termes de probabilités* » puis elle écrit au tableau :

$$P(E) = P(X \in [12, 20]) = P(12 \leq X \leq 20)$$

L'enseignante insiste sur le passage "calcul de probabilité" vers "calcul d'intégrale" : « *Vous devez savoir le traduire en termes d'intégrale* », elle complète le tableau en écrivant :

$$\int_{12}^{20} f(t)dt$$

Finalement la tâche laissée aux élèves n'est plus qu'un calcul d'intégrale dont l'expression de la fonction est donnée dans l'énoncé. Les élèves déterminent le résultat numérique à l'aide de la calculatrice (utilisation et rappel sur l'algorithme des rectangles). Les élèves savent faire ce calcul à la calculatrice de deux façons, avec la fonction de la calculatrice dédiée et avec l'algorithme des rectangles qu'ils

ont déjà programmé en classe.

La troisième question correspond au type de tâches déjà rencontré **T-CalcEsp** suivi du type de tâches **T-Interp** proposé pour la première fois, avec les variations suivantes :

- étant donné que la variable aléatoire est identifiée et que sa densité a été donnée dans l'énoncé
- interpréter le résultat dans le cas d'une espérance

L'enseignante écrit directement le calcul et explique qu'il faut ici utiliser la calculatrice. L'interprétation, qui est une application directe du début du cours, est laissée ici aux élèves et l'un d'entre eux propose : « *La production de dalles par l'entreprise sera en moyenne de 10 tonnes si on fait ça un grand nombre de fois* ». L'enseignante reformule cette réponse et prend ainsi en charge la résolution de cette question : « *En reproduisant l'expérience un grand nombre de fois la production moyenne **probable** sera de 10 tonnes* »

Conclusion :

Il s'agit d'une première rencontre, pour les élèves, avec la variation "*l'énoncé est en langage naturel*", qui est prise en charge par l'enseignante. Elle explique qu'il faut d'abord traduire l'énoncé en terme de probabilité, puis elle détaille le passage du calcul de probabilités vers le calcul d'intégrale.

Les calculs d'intégrales sont tous faits à la calculatrice par les élèves (**T-Vériff**, **T-CalcSimple** et **T-CalcEsp**).

Il s'agit également de la première rencontre avec **T-Interp**, ici interpréter la valeur de l'espérance, qui est une application immédiate du petit moment de cours qui précède l'exemple. Cette tâche est laissée à la charge des élèves.

Détails à propos de l'épisode n° 6 :

La séance se poursuit avec une activité algorithmique introduisant ainsi la loi uniforme puis avec un exemple de temps d'attente dont l'énoncé se trouve ci-dessous. L'enseignante explique aux élèves : « *C'est une généralisation d'une situation d'équiprobabilité : je choisis un nombre entier au hasard, la variable aléatoire égale à ce nombre entier est une variable aléatoire discrète. Et si maintenant je choisis un nombre réel au hasard dans un certain intervalle la variable aléatoire égale à ce nombre elle est continue, on va dire qu'elle suit la loi uniforme.* »

Exemple : Soit X la variable aléatoire égale au nombre de minutes après 14h au bout duquel un technicien arrive, ce nombre étant aléatoire dans $[0; 60]$ X va suivre la loi uniforme $[0; 60]$. Soit f la fonction de densité de la loi de probabilités de X : f est une fonction constante sur $[0; 60]$.

L'enseignante représente la fonction f dans un repère puis explique aux élèves comment retrouver l'expression de la fonction de densité d'une loi uniforme au

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

moyen d'arguments géométriques (il s'agit du type de tâches **T-DonF** dans le cas d'une loi uniforme) :

$$\text{il faut que } \int_0^{60} f(x) dx = 1$$

$$\text{or } \int_0^{60} f(x) dx = (60 - 0)f(0) \quad (\text{aire d'un rectangle}) \\ = 60 \times f(0)$$

$$\text{donc } f(0) = \frac{1}{60} \quad \text{et pour tout } x \in [0; 60] \quad f(x) = \frac{1}{60}$$

L'exemple se poursuit avec un calcul de probabilité (voir extraits du tableau figure 5.23) et l'enseignante interprète le résultat de la façon suivante : « C'est la probabilité que le technicien intervienne entre 14h15 et 14h40 ». Il s'agit du type de tâches **T-CalcSimple** + **T-Interp** dans le cas d'une loi uniforme.

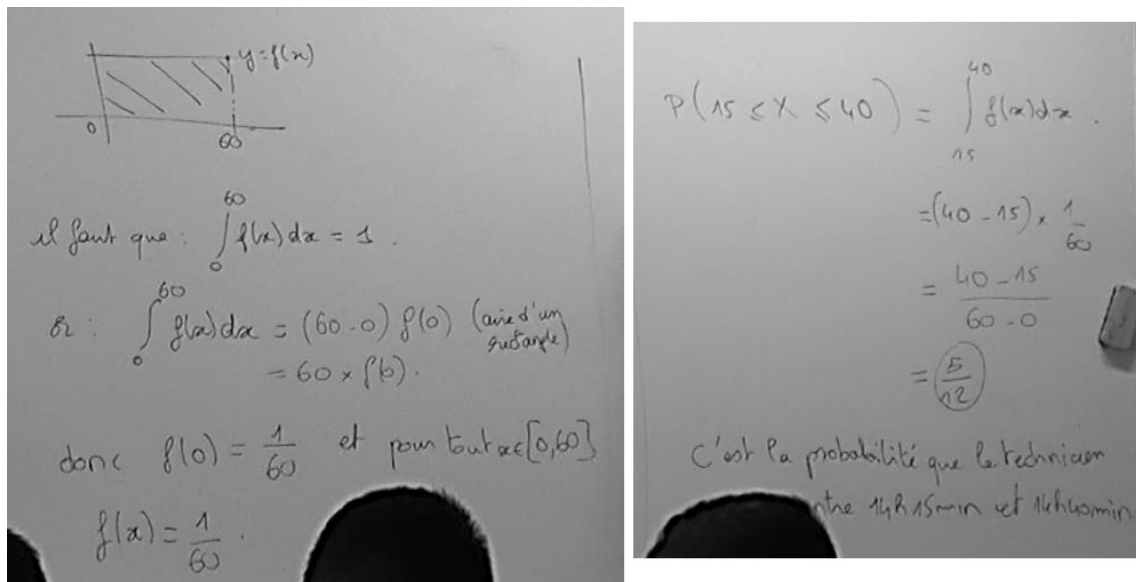


FIGURE 5.23 – Extraits du tableau lors de la séance n°2 « II. Loi Uniforme »

La séance se termine par une partie sur l'espérance, l'enseignante énonce la formule de l'espérance pour une variable aléatoire de loi uniforme puis démontre le résultat au tableau. L'enseignante conclut en calculant l'espérance pour l'exemple précédent (il s'agit du type de tâches **T-CalcEsp** + **T-Interp** dans le cas d'une loi uniforme) : « 0 + 60 divisé par 2, donc 30, ça veut dire qu'en répétant l'expérience un grand nombre de fois, sur un grand nombre d'appels, je peux espérer l'arrivée du technicien en moyenne à 14h30 ».

Conclusion :

C'est la première rencontre avec le type de tâches **T-DonF** : donner la fonction de densité d'une variable aléatoire continue, ici dans le cas d'une loi uniforme.

Les types de tâches **T-CalcSimple** et **T-CalcEsp** sont proposés pour la première

fois dans le cas d'une loi uniforme. Étant donné qu'il s'agit de la première rencontre, qui plus est dans le cadre d'un exemple du cours, toutes les variations et sous-activités de reconnaissance et d'organisation sont prises en charge par l'enseignante.

Séance n°3

n°	Thème principal de l'épisode	Commentaires
1	« Questions Rapides » sur la loi uniforme	Détails dans ce qui suit
2	Travail en autonomie sur les exercices 8 et 53 : loi uniforme (résolution géométrique) T-CalcInter , T-CalcProbaCond et T-Det([a;b])	Détails dans ce qui suit
3	Cours : « III. Loi normale centrée réduite » : définition et premières propriétés (symétrie, valeur de $f(0)$)	Voir polycopié de cours en annexe B.1
4	Premier exemple de T-CalcSimple et T-Det(a) pour la loi normale centrée réduite (utilisation d'une calculatrice)	Détails dans ce qui suit
5	Suite du cours : définition de u_α et détails de la démonstration associée exigible au bac. Cas particuliers $u_{0,05}$ et $u_{0,01}$	Voir polycopié de cours en annexe B.1
6	Premier exemple de T-Det(u_α) (loi normale centrée réduite)	Détails dans ce qui suit

FIGURE 5.24 – Synopsis de la séance n°3 – *Lois continues*

Détails à propos de l'épisode n°1 :

La séance commence par un moment de « Questions Rapides » sur le cours de la fois précédente. Le tableau est déjà pré-rempli comme on peut le voir dans la figure 5.25.

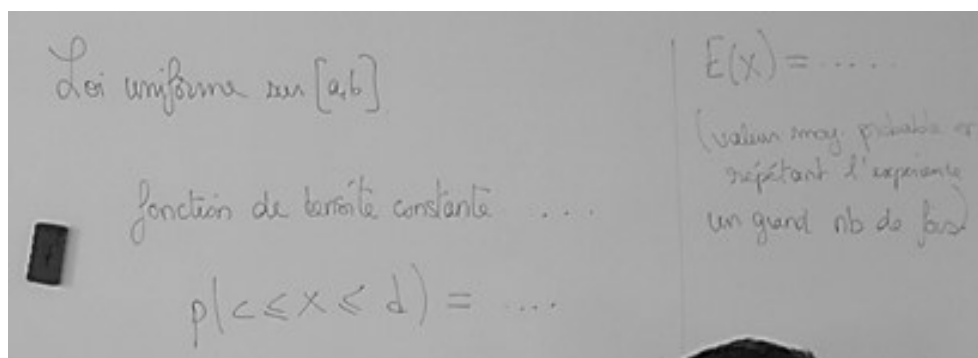


FIGURE 5.25 – Extrait du tableau lors de la séance n°3 « Questions Rapides »

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

La première question est la même que lors de la séance précédente, rappeler les propriétés d'une fonction de densité, des élèves répondent et l'enseignante complète.

L'enseignante fait ensuite quelques rappels à propos de la loi uniforme, elle complète le tableau et dessine la représentation graphique de la fonction de densité d'une loi uniforme sur $[a; b]$. La seconde question posée aux élèves est de rappeler la formule qui donne l'espérance pour une loi uniforme. Un élève répond en donnant la formule du cours et l'enseignante ajoute « *Vous devez être capables de l'interpréter avec une phrase du type, dans le contexte de la situation, c'est la valeur moyenne probable que je peux espérer obtenir en répétant l'expérience un grand nombre de fois* ».

Conclusion :

Tout comme lors de la séance précédente, l'enseignante insiste sur la définition d'une fonction de densité et sur l'interprétation de l'espérance.

Détails à propos de l'épisode n°2 :

La séance se poursuit par un travail en autonomie sur les exercices 8 et 53 du manuel Indice (voir énoncés figure 5.26). Ces exercices sont dans un contexte théorique mathématique. L'enseignante circule dans les rangs et aide les élèves qui ont des difficultés. Il n'y a pas de correction ni de trace écrite au tableau étant donné qu'il s'agit d'exercices corrigés du manuel.

8 La variable aléatoire X suit la loi uniforme sur $[2; 20]$.
On note A l'événement « $X > 5$ » et B l'événement « $X < 12$ ».

1. Déterminer $P(A)$ et $P(A \cap B)$.
2. En déduire $P_A(B)$.

53 a et b sont deux nombres réels.
La variable aléatoire X suit la loi uniforme sur $[a; b]$.
On sait que $P(X < 3) = 0,2$ et $P(X < 4) = 0,4$.
Déterminer les réels a et b .

FIGURE 5.26 – Séance n°3 – *Extraits du manuel Indice*
Énoncés des exercices 8 et 53

Les types de tâches (et leurs variations principales) présents dans l'exercice 8 sont les suivants :

- **T-CalcSimple** (loi uniforme, étant donné que la densité n'a pas été donnée dans l'énoncé)
- **T-CalcInter** (loi uniforme, étant donné que la densité n'a pas été donnée dans l'énoncé)
- **T-CalcProbaCond** (loi uniforme, étant donné que la densité n'a pas été donnée dans l'énoncé)

Le type de tâches **T-CalcSimple** a déjà été traité par les élèves avec les mêmes variations, la formule de la densité d'une loi uniforme et la formule de la probabilité $P(c < X < d)$ viennent d'être rappelées et sont toujours inscrites au tableau, cette première question ne devrait donc pas poser de difficulté aux élèves. Pour **T-CalcInter** il y a une adaptation à la charge des élèves qui est d'identifier l'évènement $A \cap B$ car il n'a pas été explicitement donné dans l'énoncé. Puis pour **T-CalcProbaCond** il faut mobiliser des connaissances issues d'un ancien chapitre (il faut se servir de la définition d'une probabilité conditionnelle).

En passant près d'une élève, l'enseignante lui dit : « l'évènement A il était donné dans l'énoncé mais l'évènement B il faut le décrire avant, et puis comment tu trouves cette probabilité? Si tout le reste est correct ça ne se voit pas, tu peux avoir 0 point éventuellement. »

Quelques instants plus tard l'enseignante s'adresse à la classe : « Je vois toute sorte de choses : 13/18, 14/18/ 15/18, mettez-vous d'accord. Vous êtes obligés soit de mettre un calcul soit de mettre un dessin ». Au bout de quelques minutes l'enseignante se dirige vers le tableau et pointe la formule rappelée au début de la séance lors des « Questions Rapides » :

$$P(c \leq X \leq d) = \frac{d - c}{b - a}$$

Elle s'adresse alors à la classe : « Je vois le même problème sur pleins de cahiers. Quand vous calculez ce genre de probabilités avec une loi uniforme il faut le détail du calcul, faut pas juste mettre la réponse, même si elle est facile à trouver mentalement. Il faut l'opération ou à la limite le dessin avec les dimensions du rectangle. » L'enseignante (A) se fait interpeller par une élève (question de l'élève inaudible) :

A : « Tu ne comptes pas le nombre d'entiers qu'il y a entre 2 et 20, tu comptes la distance, tu mesure la différence. C'est la longueur de l'intervalle. »

E : ...

A : « Justement on est sur une loi continue ici on n'est pas sur une loi discrète. »

En passant devant un autre élève :

A : « J'ai demandé du détail, il est où le détail ? »

E : « J'ai fait un schéma »

A : « Quelle est la différence entre 5 et 20? Pourquoi tu as mis 14? De toute façon je veux le détail, je veux voir écrit : 20-5 sur 20-2 »

Un nouvel élève pose une question (inaudible) à l'enseignante qui lui répond : « Alors tu sais que pour une loi continue, la probabilité que X soit compris strictement entre ... c'est la même que ... ce qui compte c'est l'aire, c'est la surface du rectangle ». En passant devant un autre élève l'enseignante lui fait remarquer qu'il a trouvé un résultat plus grand que 1 : « Tu aurais dû sonner l'alarme, une probabilité plus

grande que 1 ? ». Il n'y a pas de correction au tableau pour cet exercice.

Le type de tâches de l'exercice 53 est le suivant : **Déterminer les bornes a et b d'une loi uniforme sur $[a; b]$ en connaissant deux valeurs de probabilités.** Ce type de tâche, que l'on nommera **T-Det([a;b])**, est une nouveauté car il n'apparaît pas dans les résultats du chapitre 4 page 49.

L'enseignante prend en charge la reconnaissance de la méthode à utiliser ici, elle dit aux élèves : « *Pour le 53 géométriquement ça marche bien, sinon on pose un système* ».

L'enseignante prend également en charge l'organisation et le traitement de la résolution, elle dessine au tableau la fonction de densité de loi uniforme puis explique aux élèves : « *Pour une résolution éventuellement géométrique et astucieuse, là vous devez être capable de gérer le système jusqu'au bout et de trouver $a = 2$ et $b = 7$. Alors on peut s'amuser aussi sans système avec un dessin, on nous dit que l'aire, en termes de probabilités, la proba que X soit entre a et 4 elle doit être égale à $0,4$ et celle entre 3 et a ça doit être égale à $0,2$. Donc conclusion ces deux rectangles (elle pointe du doigt le schéma) ont mêmes aires : $0,2$ et s'ils ont mêmes aires et mêmes hauteurs " $b - a$ ", alors je n'ai pas besoin de connaître " $b - a$ ", car ils ont mêmes bases. Donc il y a la même distance entre 3 et 4 et entre a et 3 ; Donc a c'est 2 . Par contre faut le rédiger en français, un petit dessin, trois phrases bien élégantes et le tour est joué. Et en fait y a plein de problèmes mathématiques qu'on peut résoudre de cette façon* ».

L'enseignante, au travers de cet exercice, incite les élèves à utiliser des techniques de résolution géométriques pour des exercices faisant intervenir la loi uniforme.

Conclusion :

Il s'agit de la première rencontre avec **T-CalcInter**, **T-CalcProbaCond** et **T-Det([a;b])** dans le cadre d'une loi uniforme.

Des échanges transcrits ici nous relevons qu'il subsiste des confusions chez certains élèves entre variables aléatoires discrètes et continues, certains élèves "comptent" les entiers de l'intervalle $[2; 20]$ par exemple.

L'enseignante insiste fortement sur la résolution géométrique dans le cadre d'exercices relevant de la loi uniforme mais elle attend également des élèves une justification et un calcul détaillé.

Détails à propos de l'épisode n°3 :

La séance se poursuit par une nouvelle partie de cours, **III. Loi normale centrée réduite**. Un polycopié de cours à trous (à retrouver en annexe B.1) est distribué aux élèves, il est également projeté au tableau et servira de trame au déroulement du cours.

A propos de l'exemple suivant qui se trouve en bas de la page n°1 du polycopié de cours :

- (a) Calculer $P(-5 \leq X \leq 5)$
- (b) Calculer $P(X \leq 0,6)$
- (c) En déduire $P(X \geq 0,6)$ et $P(X \leq -0,6)$
- (d) Déterminer $u \in \mathbb{R}$ tel que $P(X \leq u) = 0,75$

L'enseignante explique : « On n'a pas encore vu les primitives mais vous avez déjà ressenti et remarqué que pour calculer une aire sous une courbe de façon exacte quand on ne dispose pas de formule de géométrie on va pouvoir utiliser ce qu'on appelle des primitives. Ici une telle fonction on ne peut pas la trouver, on va vous demander en terminale essentiellement, de calculer ce genre de probabilités avec la calculatrice. Il y a des tables de probabilités qui existent. »

L'enseignante ouvre une application sur son ordinateur qui représente une calculatrice graphique de terminale. Trois menus permettent de faire des calculs avec la loi normale :

1. normalFdp
2. normal FRép
3. FracNormale

Les questions (a), (b) et (c) relèvent du type de tâches **T-CalcSimple** : calculer la probabilité d'un événement simple du type $\{X < a\}$, $\{X > b\}$ ou $\{a < X < b\}$, étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite.

L'enseignante explique la procédure à suivre sur la calculatrice : « Vous avez 3 choix possibles qui indiquent le mot normal, le premier ne vous servira pas, on va regarder ensemble le choix 2 », puis se pose la question du choix pour les bornes a et b : « "lower", c'est la borne inférieure, dans la leçon je vous demande la probabilité que X soit entre -5 et 5 donc ici c'est -5. "upper" c'est la borne supérieure c'est 5. Alors ensuite la donnée suivante qu'on vous demande, que je pourrai expliquer un petit peu plus tard, c'est un "mu", comme "moyenne" et c'est 0 pour la loi normale centrée réduite. La donnée suivante c'est "sigma" c'est l'écart-type. Ensuite il faut valider. ». C'est la première rencontre avec ce type de tâches pour les élèves, ici l'enseignante prend tout en charge : la reconnaissance de la méthode l'organisation et le traitement de la tâche.

Le résultat numérique s'affiche et l'enseignante insiste à nouveau sur l'utilisation de la calculatrice car il faudra pour les élèves être à l'aise avec cette technique : « Est-ce que c'est bon pour tout le monde l'utilisation de la calculatrice ? Au bac ça va tourner beaucoup autour de ça pour la loi normale ».

L'enseignante fait ensuite une remarque sur l'allure de la courbe en utilisant le résultat trouvé qui est très proche de 1 : « On a trouvé 0,99 c'est presque 1, mais 1

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

c'est la probabilité sur tout l'ensemble des réels. Est-ce que quelqu'un a une remarque à faire? Est-ce que ça vous étonne ou est-ce que c'est normal que l'on trouve déjà presque 1 alors qu'on cherche la probabilité d'être entre -5 et 5? Regardez l'allure de la courbe.» Un élève répond : « Le gros de la courbe est déjà passé » et l'enseignante conclut : « Oui exactement, c'est une courbe qui va s'écraser et elle va s'écraser très vite sur l'axe des abscisses ».

Pour le deuxième exemple (b), l'enseignante explique aux élèves : « Si on cherche la probabilité d'être plus petit que 0,6 on a un problème car je n'ai pas de valeur inférieure; c'est là qu'on va utiliser la propriété fondamentale de symétrie de la courbe en cloche ». Elle fait un dessin au tableau : « On fait la cloche et ensuite on place l'axe des abscisses et surtout l'axe des ordonnées : 0 et là on a 1 sur racine de 2 pi; on place aussi le 0,6. On nous demande de calculer ... c'est-à-dire l'aire sous la courbe de f sur l'intervalle moins l'infini 0,6. Deux méthodes ici mais nous on va utiliser la symétrie et la remarque fondamentale du début de la leçon. Si je hachure l'aire sous la courbe en vert ici combien vaut-elle? », un élève répond : « 0,5 ». L'enseignante poursuit : « Donc en fait l'aire que je cherche à obtenir c'est 0,5 plus cette bande et sur cette bande j'ai une valeur inférieure « 0 » et une valeur supérieure. Je vais donc pouvoir utiliser la fonctionnalité de ma calculatrice ». L'enseignante écrit au tableau la décomposition de $P(X < 0,6)$:

$$P(X \leq 0,6) = P(X < 0) + P(0 \leq X \leq 0,6)$$

Elle justifie ce découpage au moyen d'éléments technologiques : « Cette égalité est vraie soit par la relation de Chasles si je le pense en termes d'intégrales, soit je le pense en termes d'événements qui sont disjoints, qui ne peuvent pas se produire en même temps, et donc la probabilité de la réunion c'est la somme des probabilités », l'enseignante termine en écrivant :

$$\begin{aligned} P(X \leq 0,6) &= P(X < 0) + P(0 \leq X \leq 0,6) \\ &= 0,5 + P(0 \leq X \leq 0,6) \text{ NormalFrep}(0; 0,6; 0 : 1) \\ &\simeq 0,725 \end{aligned}$$

Un élève demande à l'enseignante : « Est-ce qu'on peut l'écrire de moins l'infini à 0,6? », ce à quoi elle répond : « C'est tout à fait correct, j'évite de l'écrire comme ça car je ne l'ai pas défini proprement dans une leçon ». Suite au résultat numérique obtenu, l'enseignante conclut comme précédemment quant à l'allure de la courbe et à la dispersion des valeurs : « Les valeurs centrales sont très représentées et les valeurs extrêmes sont sous-représentées, il y a très peu de très très grands et très peu de très très petits ».

Le troisième exemple (c) relève du type de tâches **T-CalcSimple** calculer $P(0,6 < X)$ étant donné que l'on connaît $P(X < 0,6)$ et a déjà été rencontré dans le cadre de la loi uniforme. L'enseignante rappelle, au moyen de la représentation graphique, la relation entre ces deux événements contraires.

Calculer $P(X < -0,6)$ relève du même type de tâches mais pour la variation : calculer $P(X < -a)$ étant donné qu'on connaît $P(a < X)$. C'est une variation

propre à la loi normale qui fait intervenir la propriété de symétrie de sa densité. L'enseignante détaille l'explication géométrique (placer les bornes, hachurer les aires sur le graphe de la fonction) de la relation $P(X < -a) = P(X > a)$ et conclut : « *Donc vous voyez que même si on utilise beaucoup la calculatrice et même quand on n'a pas de valeur exacte on utilise énormément les propriétés de cette courbe. L'aire totale qui vaut 1 et la symétrie par rapport à l'axe des ordonnées* ».

Le quatrième exemple (d) relève du type de tâches **T-Det(a)** : déterminer le réel a tel que $P(X < a) =$ un nombre donné, étant donné qu'il s'agit d'une loi normale centrée réduite. A nouveau une fonctionnalité de la calculatrice donne immédiatement le résultat (cette fois-ci le choix n°3). Voici les explications de l'enseignante : « *On vous donne la probabilité, le résultat, 0,75 et on vous demande pour quel réel « u » on a la probabilité X inférieur ou égal à « u » qui soit égale à 0,75. En gros on cherche où placer une droite parallèle à l'axe des ordonnées pour que l'aire sous la courbe à gauche de la droite ça soit 75 % de l'aire totale. Tout à l'heure on avait 0,72 donc ça va être une valeur un tout petit peu plus grande ici* ».

Elle affiche la calculatrice numérique au tableau : « *La calculatrice vous la donne, c'est le troisième choix FracNormal. C'est un genre de fonction réciproque, ça marche avec une seule borne, la borne supérieure. On vous demande "area", c'est l'aire, on rentre 0,75 et voyez on a une valeur proche. Je ne suis pas étonnée de trouver un petit peu plus que 0,6.* » L'enseignante écrit au tableau le raisonnement et le résultat et indique que l'on peut le vérifier (en faisant « *la même manip que tout à l'heure* »). Elle demande aux élèves de le faire eux-mêmes sur leur propre calculatrice.

Un élève demande : « *Est-ce qu'on pourrait retrouver « u » algébriquement ?* », ce à quoi l'enseignante répond : « *Tu ne peux pas calculer explicitement cette intégrale comme j'ai dit en introduction donc on va voir les valeurs approchées les plus fines possibles par différentes méthodes.* »

Conclusion :

C'est la première rencontre avec **T-CalcSimple** et **T-Det(a)** pour une loi normale centrée réduite, l'enseignante prend tout en charge : la reconnaissance de la méthode, l'organisation et le traitement de la tâche. L'enseignante accompagne toujours la technique instrumentée (calculatrice) d'explications géométriques, à partir du graphe de la densité. Elle insiste fortement sur l'utilisation de la calculatrice car les élèves devront être à l'aise avec cette technique.

Il est intéressant de relever que l'enseignante indique clairement que tous les exercices seront à faire avec la calculatrice mais qu'elle parle aussi de tables de probabilités alors que celles-ci ne sont pas du tout utilisées dans le secondaire mais plutôt dans le supérieur où l'usage de la calculatrice n'est plus autorisé. C'est une forme d'accompagnement de la transition.

Même si la calculatrice permet de faire beaucoup de choses, une expression du type $P(X < a)$ ne peut pas se calculer immédiatement et il faut transformer l'ex-

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

pression. $P(X > a)$ se calcule avec l'évènement contraire mais est justifié géométriquement, tout comme $P(X < -a) = P(X > a)$ qui est détaillé et justifié sur le graphe de f .

Détails à propos de l'épisode n°5 :

L'exemple page 3 du polycopié de cours, relève du type de tâches **T-Det**(u_α) : déterminer u_α pour α un nombre donné, étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite et en utilisant une calculatrice obligatoirement.

Déterminer avec une calculatrice une valeur approchée du réel positif $u_{0,08}$ tel que $P(-u_{0,08} \leq X \leq u_{0,08}) = 0,92$

L'enseignante commence par dire aux élèves : « Alors comment on fait maintenant pour toutes les autres valeurs qu'on ne doit pas apprendre par coeur ? (elle fait allusion à $u_{0,05}$ et $u_{0,01}$). Ici je laisse 8% en tout, et je veux avoir 92% au milieu. Comment je vais faire, déjà je traduis l'énoncé je cherche la probabilité que X soit entre $-u_{0,08}$ et $u_{0,08}$, je cherche à que ce soit égal à $1 - 0,08$ donc $0,92$. Donc comment je fais ça ? Par quoi je commence ? », un élève répond : « Par un dessin ».

L'enseignante explique aux élèves comment faire le dessin d'une courbe de loi normale centrée réduite, puis comment placer les droites d'équation $x = u_{0,08}$ et $x = -u_{0,08}$. Elle hachure l'aire correspondant à $0,92$ et détaille, en utilisant la symétrie, à quoi correspondent chacun des morceaux sous la courbe. Le dessin se trouve dans la figure 5.27.

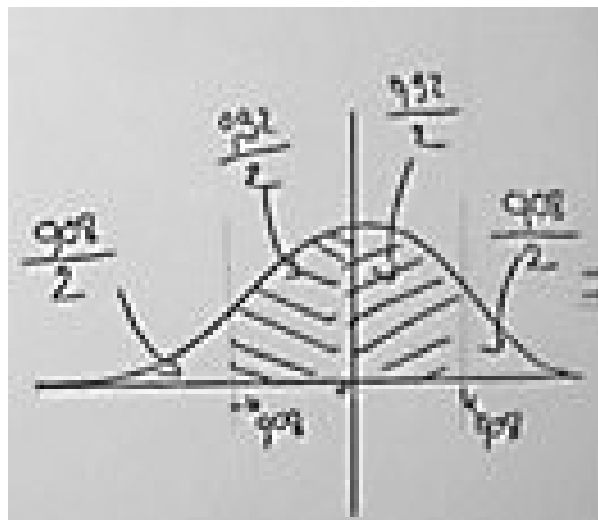


FIGURE 5.27 – Dessin de l'enseignante pour l'exemple du cours – Séance n°3

Une fois le dessin fait, l'enseignante explique aux élèves comment poursuivre la résolution : « *Alors comment on fait ? L'aire est donnée et on demande la borne. Je vous ai dit tout à l'heure FracNormale fait ça, mais FracNormale elle va nous donner une borne supérieure qui réalise l'aire sur l'intervalle de moins l'infini à cette borne (elle montre le borne $u_{0,08}$). Donc Fracnormale je ne peux pas l'utiliser avec 0,92, il faut que j'ajoute l'aire ici (elle montre l'intervalle $] - \infty; -u_{0,08}[$) donc 0,04, donc voilà, soit vous le voyez comme ça : 0,92 plus alpha divisé par 2. Ou alors comme d'habitude, sur cet intervalle $] - \infty; 0[$ on a 0,5 (elle montre l'intervalle en question) plus 0,92 divisé par 2. Vous suivez le raisonnement que vous voulez tant que vous avez un dessin et une écriture à côté qui est claire. »*

Elle écrit au tableau :

$$\text{Alors } P(0 \leq X \leq u_{0,08}) = \frac{0,92}{2} = 0,46 \text{ (par symétrie)}$$

Elle insiste sur la justification "par symétrie" puis continue à détailler la méthode : « *Donc, ce qu'il faut que j'exprime c'est la probabilité que X soit inférieur ou égal, ça je veux savoir combien ça fait parce que si je sais combien ça fait je pourrai à ce moment-là utiliser la calculatrice* » et elle complète le tableau :

$$\begin{aligned} \text{Donc } P(X \leq u_{0,08}) &= P(X \leq 0) + P(0 \leq X \leq u_{0,08}) \\ &= 0,5 + 0,46 \\ &= 0,96 \end{aligned}$$

Elle termine en expliquant comment conclure à l'aide de la calculatrice : « *Avec la fonction réciproque FracNormale, l'aire je peux lui donner, c'est 0,96. J'obtiens alors $u_{0,08} \simeq 1,75$. On peut vérifier ce résultat avec la calculatrice, si je demande la probabilité d'être entre -1,75 et +1,75 elle doit me donner 92 %.* »

La séance se termine avec la partie sur l'espérance mathématique, l'enseignante explique : « *C'est pour ça que dans la calculatrice pour μ on met 0, μ comme moyenne, et on met σ qui vaut 1. C'est pour ça qu'on l'appelle la loi centrée sur 0 et réduite* ».

Conclusion :

Première rencontre avec **T-Det**(u_α), l'enseignante prend ici en charge la reconnaissance de la méthode (commencer par un dessin, transformer l'expression, utiliser la calculatrice), l'organisation et le traitement. Elle passe beaucoup de temps sur la représentation graphique de la fonction de densité de loi normale centrée réduite et sur la transcription du problème d'un point de vue géométrique.

Séance n° 4

n°	Thème principal de l'épisode	Commentaires
1	« Mathoral » exercice sur la loi uniforme	Détails dans ce qui suit
2	« Questions Rapides » sur la loi normale centrée réduite (questions de cours, propriétés géométriques de la densité, T-CalcSimple)	Détails dans ce qui suit
3	Travail en autonomie sur l'exercice 78 du manuel puis correction au tableau	Détails dans ce qui suit
4	Cours : « IV. Lois Normales » : introduction avec géogébra, définition et premières propriétés. Premier exemple de T-CalcSimple et T-Det(a) pour la loi normale centrée réduite (utilisation d'une calculatrice)	Voir polycopié de cours en annexe B.1
5	Cours : « IV.2. Déterminer l'espérance ou l'écart-type d'une loi normale » : exemple avec T-Det(σ) et T-Det(μ)	Détails dans ce qui suit
6	Travail en autonomie sur les exercices 89 et 95 du manuel	Détails dans ce qui suit

FIGURE 5.28 – Synopsis de la séance n°4 – *Lois continues*

Détails à propos de l'épisode n° 1 :

La séance commence par un « Mathoral », une élève présente au tableau sa correction d'un exercice qu'elle a préparé avant la séance. Il s'agit d'un extrait de sujet de baccalauréat sur la loi uniforme, issu du manuel *Indice*, dont l'énoncé est présenté dans la figure 5.29.

La première question relève du type de tâches **T-CalcEsp** suivi du type de tâches **T-Interp** (interpréter le résultat dans le cas d'une espérance) avec les variations suivantes :

- étant donné que c'est la première question
- étant donné que la densité n'a pas été donnée dans l'énoncé
- à partir d'un énoncé en langage naturel (ex : calculer le prix moyen)
- étant donné qu'il s'agit, d'après l'énoncé, d'une loi uniforme

La seconde question relève du type de tâches **T-CalcSimple** avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné qu'il s'agit, d'après l'énoncé, d'une loi uniforme

Nous ne présentons pas l'analyse des trois questions suivantes (2.a ; 2.b et 2.c) car celles-ci relèvent des types de tâches **T-DonLoi** et **T-CalcSimple** mais dans le

Sujet C

Capacités mises en œuvre

- Calculer avec une loi uniforme
- Calculer avec une loi binomiale

Les deux rives d'un estuaire sont reliées par des bateaux qui quittent la rive nord exactement toutes les 10 minutes.

M. Dulac séjourne sur la rive nord et traverse l'estuaire une fois par jour pour se rendre dans la partie sud. Son arrivage au point d'embarquement sur la rive nord se fait au hasard.

1. Le temps, en minutes, séparant l'arrivée de M. Dulac à l'embarcadère du prochain départ du bateau, définit une variable aléatoire T qui suit une loi uniforme.

- a. Quel est le temps d'attente moyen de M. Dulac à l'embarcadère ?
- b. Montrer que la probabilité qu'un jour donné M. Dulac attende plus de 7 minutes à l'embarcadère est $p = 0,3$.

2. M. Dulac séjourne 10 jours sur la rive nord. Le nombre de jours où son attente, pour prendre le bateau, est supérieure à 7 minutes définit une variable aléatoire X .

On suppose que l'arrivée de M. Dulac à l'embarcadère se fait de façon indépendante d'un jour à l'autre.

- a. Quelle est la loi suivie par X ? Déterminer $E(X)$.
- b. Calculer, à 10^{-3} près, la probabilité que M. Dulac n'attende jamais plus de 7 minutes à l'embarcadère.
- c. Calculer $P(X \leq 5)$ à 10^{-3} près.

Sujet C

1. T suit la loi uniforme sur $[0; 10]$.
 - a. Le temps d'attente moyen de Monsieur Dulac est de 5 minutes.
 - b. $P(T > 7) = 1 - P(T \leq 7) = 1 - \frac{7}{10} = 0,3$.
2. X suit la loi binomiale de paramètres $n = 10$ et $p = 0,3$.
 - a. $P(X = 0) = 0,7^{10} \approx 0,028$.
 - b. $P(X \leq 5) \approx 0,953$.

FIGURE 5.29 – Séance n°4 – Extraits du manuel *Indice*
Énoncé du sujet C (à g.) – Corrigé issu du manuel du professeur (à d.)

cas d'une variable aléatoire discrète (loi binomiale à établir question 2.a).

L'élève a écrit un corrigé qu'elle choisi de projeter au tableau (corrigé à retrouver dans la figure 5.30), elle lit les questions une à une puis lit la correction affichée au tableau. Elle ne commente pas ce qui est écrit.

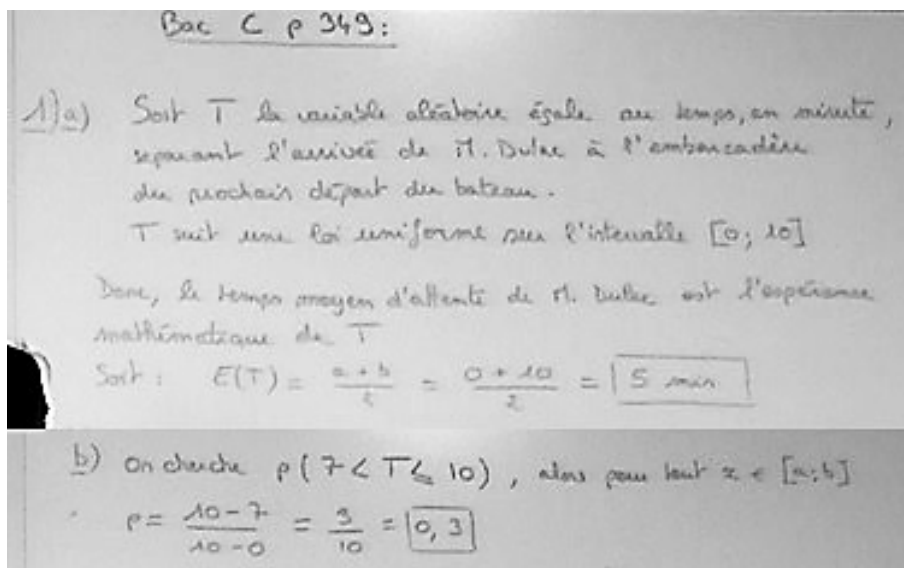


FIGURE 5.30 – Séance n°4 – Correction de l'élève projetée au tableau

Conclusion :

Les types de tâches **T-CalcEsp** suivi de **T-Interp** ainsi que **T-CalcSimple** ont déjà été proposés et traités par les élèves dans le cas d'une loi uniforme. L'adaptation *l'énoncé est en langage naturel*, à la charge de l'élève, ne semble pas avoir posé de difficulté ici.

On retrouve dans cet exercice des questions sur la loi uniforme entremêlées avec des questions sur la loi binomiale.

Détails à propos de l'épisode n° 2 :

La séance se poursuit par un moment de « Questions Rapides ». Ces questions sont posées par l'enseignante sur la loi normale centrée réduite (certaines questions sont écrites au tableau).

- Première question : « Soit X une variable aléatoire qui suit la loi normale centrée réduite, j'appelle f la fonction de densité de cette loi de probabilité. Donnez-moi l'expression algébrique de $f(x)$ et son intervalle de définition ». Il s'agit du type de tâches **T-DonF** dans le cas d'une loi centrée réduite.

L'enseignante passe dans les rangs en précisant aux élèves qu'ils doivent savoir cette formule par coeur. Mais ils ne semblent pas connaître l'expression de f , l'enseignante leur demande alors de commencer par dessiner l'allure de la courbe de la fonction dans un repère. Un élève (E) passe au tableau et trace correctement l'allure de la courbe en cloche de la loi normale. L'enseignante (A) s'adresse ensuite à la classe :

A : « Quelle est la propriété géométrique fondamentale de cette courbe ? Pour quelle valeur de x est-elle définie ? »

E : « Pour tout x de \mathbb{R} »

A : « Oui, on n'oublie jamais de préciser les ensembles de définition et de dérivabilité. Est-ce que tu te rappelles d'une valeur approchée de l'ordonnée de $x = 0$? »

E : « 0,33 ? »

A : « Non je crois pas, c'est plutôt 0,4. »

L'élève complète le graphe de la fonction en inscrivant la valeur 0,4.

A : « Une dernière question, quelle est la propriété principale de cette courbe ? »

E : « Symétrie par rapport à l'axe des ordonnées. »

- Seconde question : « Quelle est la probabilité que X soit positif ou nul ? Que X soit strictement positif ? » à laquelle un élève répond simplement : « $1/2$ et $1/2$ ».

L'enseignante complète cette réponse par des éléments de technologie : « *Par symétrie et parce que la probabilité que X soit égal à 0 c'est 0 car on est sur les lois continues.* »

- Troisième question : « *X suit toujours une loi normale centrée réduite, calculer $P(-0,3154 \leq X \leq 1,5779)$. Je veux un dessin et une valeur approchée à la calculatrice, au millième près* ». Cette question relève du type de tâche, déjà rencontré, **T-CalcSimple**.

L'enseignante passe dans les rangs et envoie une élève (que l'on appellera Romane (R)) au tableau à qui elle demande d'illustrer sur le graphe de f la probabilité. L'élève pointe le sommet de la courbe et dit : « *Ici ça sera 0,4* » elle reprend ici ce qui vient d'être rappelé lors de la première question, à savoir l'ordonnée en 0. L'enseignante (A) ne se satisfait pas de cette réponse et reprend en s'adressant cette fois-ci à la classe :

A : « *Quel est le lien, entre la fonction de densité dont la courbe est dessinée au tableau et le calcul d'une probabilité ? Que faut-il faire sur le dessin ? Géométriquement comment est représentée cette probabilité ?* »

Un autre élève (E2) répond :

E2 : « *J'ai dessiné les bornes.* »

A : « *Sur quel axe ?* »

E2 : « *Des abscisses.* »

A : « *Il faut donc Romane que tu places tes deux bornes qui encadrent X .* »

L'élève (R) au tableau s'exécute.

A : « *Et qu'est-ce qui est égal à la probabilité demandée ?* »

L'élève (R) hachure une zone.

A : « *Qu'est-ce que tu as fais ? Tu m'expliques ?* »

R : « *L'aire sous la courbe.* »

A : « *Explique-moi comme si je ne voyais pas ton dessin.* »

R : « *C'est égal à l'aire sous la courbe et au-dessus de l'axe des abscisses, comprise entre l'axe des abscisses et la courbe.* »

A : « *Il manque la borne de gauche et celle de droite, entre les droites d'équations ...* »

R : « *$x = -0,3154$ et $1,5779$* »

- Quatrième question : calculer $P(X \leq 1,28)$. L'enseignante demande à nouveau aux élèves de faire le dessin avant de faire le calcul. Il s'agit ici aussi du type de tâches déjà rencontré **T-CalcSimple** étant donné qu'il s'agit d'une loi normale centrée réduite. Il n'y a pas d'adaptations ici, il s'agit d'une question simple isolée. Un nouvel élève est au tableau, il écrit ce qui suit puis complète le graphe du tableau :

$$\begin{aligned} P(X \leq 1,28) &= P(X \geq 0) + P(0 \leq X \leq 1,28) \\ &= 0,5 + 0,400 \\ &= 0,900 \end{aligned}$$

L'enseignante s'étonne de $P(X \geq 0)$: « Ça fait bien 0,5 mais graphiquement c'est pas ça qu'on va faire, finis ton dessin. On va découper le domaine sous la courbe en deux aires », l'élève hachure le domaine correspondant à $P(X \leq 1,28)$ et l'enseignante lui dit : « L'aire que tu as hachurée en deuxième n'est pas égale, enfin elle est égale mais ce n'est pas la probabilité que X soit positif ou nul, c'est la probabilité que X soit négatif ou nul plutôt non ? », l'élève corrige.

- Cinquième question : calculer $P(X \leq -1,415)$. Un élève est envoyé au tableau et écrit :

$$\begin{aligned} P(X \leq -1,415) &= 0,5 - P(-1,415 \leq X \leq 0) \\ &\simeq 0,079 \end{aligned}$$

Il explique : « Comme on sait que la partie inférieure à 0 c'est 0,5, faut qu'on élimine la partie de -1,415 à 0. Du coup on fait 0,5 moins la probabilité de $-1,415 \leq X \leq 0$, finalement on trouve 0,079 », l'enseignante lui demande de compléter par un dessin puis elle ajoute : « Quand j'ai un intervalle avec une seule borne on ne peut pas utiliser la calculatrice directement ».

- Sixième question : « On cherche un réel a tel que la probabilité que $X \leq a$ soit égale à 0,654. Je veux l'arrondi au millième près de a et un dessin ». Il s'agit ici du type de tâches déjà rencontré dans un exemple du cours lors de la séance n°3, **T-Det(a)** : déterminer le réel a tel que $P(X < a) =$ un nombre donné. Une fonctionnalité de la calculatrice permet d'obtenir directement le résultat. Un élève passe au tableau et écrit : $a = 0,396$. L'enseignante complète quelques éléments de technologie : « C'est l'aire du domaine qui est donnée, on trouve 0,396 avec le troisième choix du menu de la calculatrice associé à la loi normale »
- Septième question : « On cherche un réel b tel que la probabilité que $X \geq b$ soit égale à 0,573, un dessin avant tout ». Il s'agit du même type de tâches qu'à la question précédente mais avec l'adaptation *déterminer le réel b tel que $P(X > b) =$ un nombre donné*, à la charge des élèves et qu'ils rencontrent pour la première fois.

Un élève passe au tableau, il semble bloqué. L'enseignante lui demande donc de commencer par le dessin et l'aide : « *Alors qu'est-ce qu'il faut faire pour avoir b ? Il faut se ramener à la recherche d'une probabilité* ». L'élève finit par retourner s'asseoir et l'enseignante (A) donne à la classe des aides procédurales : « *Il faut se ramener à l'événement contraire puisqu'on a la probabilité de l'évènement X plus grand que b , on peut connaître la probabilité de l'évènement X plus petit que b* ». Un autre élève (M) est envoyé au tableau.

A : « *C'est toi qui as la parole pour nous expliquer comment utiliser l'évènement contraire.* »

M : « *On fait 1 moins ...* »

A : « *Ecris en même temps que tu nous le dis* »

M : « *C'est 1 moins $P(X \leq b) = 0,573$, on passe le 1 de l'autre côté. Et du coup après on fait *FracNormale** » il écrit en même temps au tableau ce qui suit :

$$\begin{aligned}1 - P(X < b) &= 0,573 \\P(X < b) &= 0,427 \\b &= -0,184\end{aligned}$$

Conclusion :

Les élèves doivent connaître la densité d'une loi normale centrée réduite et les propriétés de cette dernière, le type de tâches **T-DonF** est à traiter d'abord avec un dessin (demande de l'enseignante).

Il s'agit de la première rencontre avec **T-CalcSimple** dans le cas d'une loi centrée réduite. L'enseignante attend des éléments de justifications pour les calculs (éléments de technologie) : *par symétrie, c'est une loi continue donc l'intégrale sur le domaine vaut 1*. **T-CalcSimple** est d'abord proposé avec la variation : calculer une probabilité du type $P(a \leq X \leq b)$ ce qui est immédiat à la calculatrice. Puis avec la variation : $P(X \leq b)$ et enfin avec la variation $P(X \leq -c)$ où c est positif, ces formes doivent être transformées pour être calculées. C'est exactement le même enchaînement que dans l'exemple du cours de la séance n°3, certaines formes de probabilités se calculent immédiatement à la calculatrice, d'autres doivent être transformées.

Concernant le type de tâches **T-Det(a)** : déterminer le réel a tel que $P(X < a) =$ un nombre, une fonctionnalité de la calculatrice permet d'obtenir directement le résultat et les élèves ont déjà rencontré ce type de tâche, pas de difficulté. En revanche, pour déterminer le réel b tel que $P(X > b) =$ un nombre donné, il faut transformer l'expression et pas juste rentrer un nombre dans la calculatrice, c'est plus difficile pour les élèves.

Du point de vue des changements de registres, les élèves semblent aisément hachurer les aires correspondant aux probabilités lorsque cela leur est demandé. Ce n'est plus le cas lorsqu'ils doivent écrire les probabilités de domaines hachurés. Le passage *probabilités* vers *aires* semble ne pas poser de difficulté, mais le passage *aires* vers *probabilités* semble plus difficile pour les élèves.

De façon générale, l'enseignante demande de compléter tous les calculs par un

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

dessin et indique que tous les exercices qu'ils vont rencontrer sur ce thème seront autour de ces types de tâches là (**T-CalcSimple** et **T-Det(a)**).

Ce moment d'échanges a également permis à l'enseignante de faire revoir aux élèves des éléments fondamentaux du cours sur la loi normale centrée réduite (symétrie notamment), vus à la séance précédente.

Détails à propos de l'épisode n°3 :

Les élèves travaillent en autonomie sur l'exercice n°78 du manuel dont l'énoncé se trouve dans la figure 5.31.

78 Lors de l'étude d'une population constituée par des lecteurs d'une revue, on construit un indice de satisfaction S qui suit la loi normale centrée réduite (plus la valeur de l'indice est élevée, plus le lecteur est satisfait).

Déterminer les indices de satisfaction suivants :

- a. S_1 tel que 5 % des lecteurs aient un indice supérieur à S_1 ;
- b. S_2 tel que 95 % des lecteurs aient un indice inférieur à S_2 ;
- c. S_3 tel que 5 % des lecteurs aient un indice inférieur à S_3 ;
- d. S_4 tel que 95 % des lecteurs aient un indice supérieur à S_4 ;
- e. S_5 tel que 95 % des lecteurs aient un indice compris entre $-S_5$ et S_5 ;
- f. S_6 tel que 5 % des lecteurs aient un indice en dehors de l'intervalle $[-S_6 ; S_6]$.

FIGURE 5.31 – Séance n°4 – Énoncé exercice 78

Les quatre premières questions (a. b. c. et d.) relèvent du type de tâches déjà rencontré, **T-Det(a)**, avec les variations suivantes :

- à partir d'un énoncé en langage naturel (ex : déterminer la température t telle que 10% des journées aient une température supérieure à t)
- étant donné qu'il s'agit, d'après l'énoncé, d'une loi normale centrée réduite
- déterminer le réel b tel que $P(X > b) =$ un nombre donné

L'enseignante explique : « Quand on a fait le (a) on fait le (b) sans calculatrice. Le (e) et le (f) on les fait sans calculatrice », une élève est envoyée au tableau et écrit :

$$\begin{aligned} \text{a) } P(X \geq S_1) &= 0,05 \\ &\Leftrightarrow 1 - P(X \leq S_1) = 0,05 \\ S_1 &\simeq 1,645 \\ \text{b) } P(X \leq S_2) &= 0,95 \\ S_2 &\simeq 1,645 \end{aligned}$$

L'élève ne dit rien à l'oral mais l'adaptation, l'énoncé étant en langage naturel, ne semble pas lui avoir posé de difficulté. L'enseignante (A) demande au reste de la classe :

A : « Quel est le lien entre S_1 et S_2 ? »

E1 : « Ils sont au même endroit. »

A : « Ils sont donc égaux. »

E2 : « Est-ce qu'on peut justifier avec un dessin ? »

A : « Un dessin au bac c'est un peu léger, un dessin plus les égalités... »

L'enseignante trace au tableau le graphe de la loi normale centrée réduite et y représente S_1 et S_2 (voir figure ci-dessous), elle s'adresse à la classe : « Toutes ces réponses on peut les écrire en termes de probabilité. Pour justifier proprement il faut détailler votre raisonnement, la calculatrice n'a que deux fonctions donc il faut détailler proprement sur votre cahier comment on se ramène à une de ces deux fonctions. Pour ça on peut utiliser un argument de géométrie parce qu'une propriété fondamentale de la loi normale centrée réduite c'est sa symétrie par rapport à l'axe des ordonnées. »

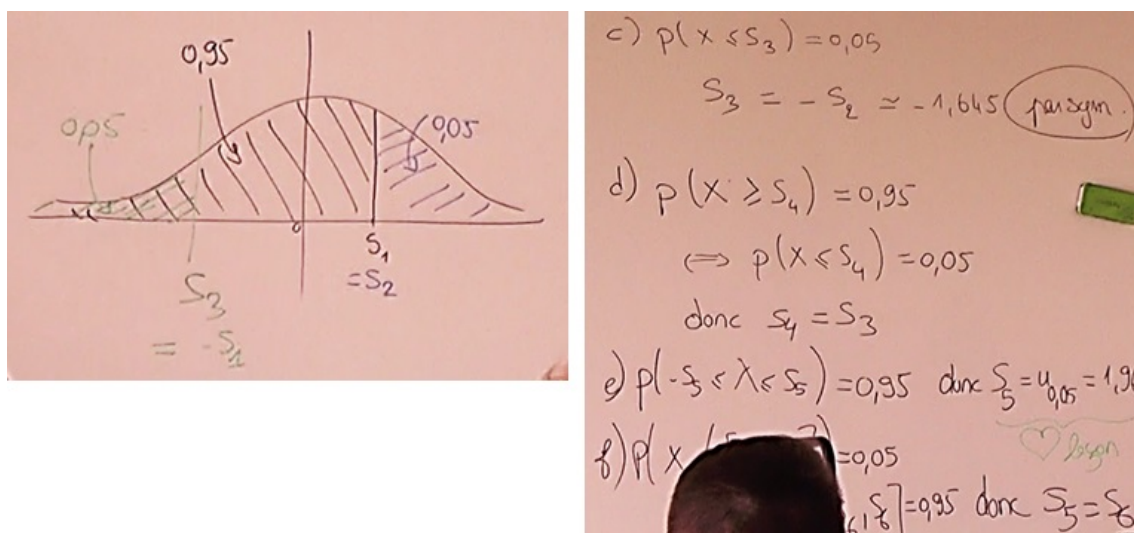


FIGURE 5.32 – Séance n°4 – Correction de l'exercice 78 écrite au tableau par l'enseignante

L'enseignante corrige au tableau questions (c), (d) et (e), elle indique : « Pour (c), par symétrie on va être en valeur absolue sur la même valeur que S_2 et S_1 mais avec un signe négatif. Sur une copie vous pouvez le justifier en disant « par symétrie » et avec un dessin qui illustre ou vous pouvez, si vous préférez, l'écrire en termes de probabilités en passant par l'événement contraire et en utilisant que l'aire sur $[0; +\infty[$ c'est 0,5. La question (d) c'est $1 - P(X > S_4)$, ça correspond à l'événement contraire de (c), il y a le même lien entre (c) et (d) qu'entre (a) et (b) »

La question (e) relève du type de tâche, lui aussi déjà rencontré, **T-Det**(u_α) : déterminer u_α pour α un nombre donné, à partir d'un énoncé en langage naturel. L'enseignante explique : « Ici il faut reconnaître une des deux valeurs que je vous ai dit d'apprendre par coeur dans la leçon, parce qu'on cherche ici deux bornes, on

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

cherche un intervalle centré en 0 tel que la probabilité ... On reconnaît ici le $u_{0,05}$ et on a appris dans la leçon que c'est 1,96. La dernière question (f) c'est l'évènement contraire. »

Conclusion :

Face au type de tâches **T-Det(a)** : déterminer le réel b tel que $P(X > b) =$ un nombre donné, l'enseignante explique que deux justifications sont possibles, soit un dessin et la mention *par symétrie*, soit une rédaction plus théorique c'est-à-dire expliciter le découpage de l'expression faisant intervenir l'évènement contraire. Cet exercice a permis de mettre en application la méthode vue lors de la séance n°3 pour traiter **T-Det(u_α)**, même si ici $\alpha = 0,05$ et qu'il s'agit en fait d'une valeur à connaître par coeur par les élèves et non à calculer.

Détails à propos de l'épisode n°5 :

La séance se poursuit par un moment de cours durant lequel l'enseignante introduit les lois normales (voir polycopié de cours en annexe B.1). Page 5 du polycopié se trouve un exemple dont l'énoncé est redonné dans la figure 5.33.

n°1 : Soit X une variable aléatoire qui suit une loi normale $N(3; \sigma^2)$ où σ est un réel strictement positif à déterminer sachant que $P(X < 2) = 0,4$.

n°2 : Soit X une variable aléatoire qui suit la loi normale $N(\mu; 10^2)$ où μ est un réel à déterminer sachant que $P(X < 30) = 0,7$.

Méthode : Ces problèmes peuvent être résolus grâce à la calculatrice, en se ramenant au cas d'une variable aléatoire qui suit la loi normale centrée réduite $N(0; 1)$.

FIGURE 5.33 – Séance n°4 – IV.2. Déterminer l'espérance ou l'écart-type d'une loi normale – Exemple

La première question relève du type de tâches **T-Det(σ)** : déterminer l'écart-type d'une variable aléatoire X de loi normale, avec les variations suivantes :

- sans qu'une variable aléatoire Z centrée réduite associée à X soit définie/donnée comme $\frac{X - \mu_X}{\sigma_X}$
- à partir d'une probabilité du type $P(X < b)$ donnée

La correction est détaillée dans le polycopié de cours, l'enseignante explique pourquoi et comment se ramener à un problème du type **T-Det(a)** : « Je m'intéresse à $X < 2$ ici car c'est l'indice que j'ai, je veux que la probabilité de cet évènement soit plus petite que 0,4. Et je vais me ramener à la variable aléatoire centrée réduite. Ces 2 événements sont équivalents donc ils ont la même probabilité. J'ai renommé $\frac{X-3}{\sigma_X}$ par Z . » Après avoir commenté la méthode de résolution, l'enseignante invite les élèves à faire eux-mêmes les calculs à l'aide de leur calculatrice.

Un élève demande des précisions à propos de la dernière étape du raisonnement :

$$\frac{-1}{\sigma} \simeq -0,253 \text{ donc } \sigma \simeq 3,95$$

Il s'agit ici d'une difficulté dans la résolution d'équations algébriques, l'enseignante explique : « *C'est comme résoudre une équation mais ici avec une valeur approchée.* »

La seconde question relève du type de tâches **T-Det**(μ) : déterminer la moyenne d'une variable aléatoire X suivant une loi normale, avec les variations suivantes :

- sans qu'une VA Z centrée réduite associée à X soit définie/donnée comme $\frac{X-\mu_X}{\sigma_X}$
- à partir d'une probabilité du type $P(X < b)$ donnée

Un élève lit la méthode de résolution de la seconde question. L'enseignante ajoute simplement : « *Avec ces lois d'espérance et d'écart-type quelconques vous allez pouvoir étudier tout un tas de situations qui représentent des choses très courantes.* »

Conclusion :

Il s'agit d'une première rencontre, pour les élèves, avec **T-Det**(σ) et **T-Det**(μ) sous forme d'un exemple de cours très détaillé. L'enseignante prend en charge la variation « *sans qu'une nouvelle variable aléatoire Z (de loi normale centrée réduite) soit définie comme $\frac{X-\mu_X}{\sigma_X}$* », en effet c'est la première rencontre pour les élèves avec ce type de tâches et *a fortiori* avec cette variation.

Détails à propos de l'épisode n°6 :

Travail en autonomie sur les exercices 89 et 95 du manuel dont les énoncés sont à retrouver dans la figure 5.34.

89 Un lot de kiwis a été calibré. On admet que la variable aléatoire qui à chaque kiwi, pris au hasard, associe sa masse en grammes, est une variable aléatoire X dont la loi de probabilité est la loi normale $\mathcal{N}(90; 6,25)$.
On accepte les fruits dont la masse est comprise entre 85 g et 95 g.
Parmi 10 000 fruits pris au hasard, combien en moyenne seront acceptés ?

95 Les résultats seront arrondis au millième le plus proche.
Dans un pays, la taille des femmes peut être modélisée par une variable aléatoire X suivant la loi normale de moyenne 165 cm et d'écart-type 6 cm alors que celle des hommes peut être modélisée par une variable aléatoire Y suivant la loi normale de moyenne 174 cm et d'écart-type 8 cm.
De plus, on sait que les femmes représentent 51,3 % de la population de ce pays.
On considère une personne prise au hasard dans cette population. On note A l'événement « la taille de cette personne dépasse 1,70 m » et F l'événement « cette personne est une femme ».

1. Calculer les probabilités $P_F(A)$ et $P_{\bar{F}}(A)$.
2. Utiliser un arbre pondéré pour déterminer la valeur de $P(A)$.
3. Quelle est la probabilité pour qu'une personne mesurant plus de 1,70 m soit une femme ?

FIGURE 5.34 – Séance n°4 – Énoncés des exercices 89 et 95

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

L'exercice 89 propose le type de tâches **T-CalcSimple** : calculer la probabilité d'un évènement simple du type $\{ a < X < b \}$ pour une loi normale, avec les variations :

- étant donné que la variable aléatoire est identifiée dans l'énoncé
- à partir d'un énoncé en langage naturel
- il s'agit d'une loi normale $\mathcal{N}(\mu; \sigma^2)$ de paramètres μ et σ donnés

La première question de l'exercice 95 relève du type de tâches **T-CalcSimple** : calculer la probabilité d'un évènement simple du type $\{X > b\}$ pour une loi normale, avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- c'est la première question de l'énoncé
- **étant donné que la probabilité recherchée est présentée comme une probabilité conditionnelle** (il s'agit d'une nouveauté par rapport au chapitre 4)

Cette première question demande donc un travail d'interprétation initial afin notamment d'identifier que l'évènement $\{A \text{ sachant } F\}$ est en réalité l'évènement $\{X > 170\}$ où X suit une loi normale.

Les deux questions suivantes correspondent aux types de tâches (et variations principales) rencontrés lors du chapitre sur les probabilités conditionnelles :

- **T-CalcSimple** : calculer la probabilité d'un évènement simple (étant donné que les évènements ont déjà été décrits et nommés dans le texte et qu'il est conseillé de représenter la situation par un arbre ; les données sont en pourcentages)
- **T-CalcProbaCond** : calculer une probabilité conditionnelle (à partir d'un énoncé en langage naturel, étant donné que la situation a déjà été représentée par un arbre mais sans que l'évènement recherché n'apparaisse sur l'arbre)

A propos de la notation $\mathcal{N}(90; 6, 25)$ de l'exercice 89, l'enseignante explique : « Attention à 6,25. Demandez-vous ce que ça représente pour cette loi normale. Quel est l'écart-type de cette loi normale ? C'est pas 6,25 on est bien d'accord, c'est une difficulté qui ne sert à rien. Il faut y penser. » L'enseignante s'exprime de nouveau à ce sujet : « Il faut une trace écrite sur votre cahier, si vous avez trouvé que l'écart-type c'était bien 2,5 il faut écrire quelque part racine de 6,25 ».

L'enseignante passe dans les rangs et regarde ce que font les élèves, ces exercices seront corrigés à la séance suivante mais voici quelques remarques de l'enseignante que nous avons relevées lorsque celle-ci se déplace dans la classe pour aider les élèves :

- En passant devant un élève (exercice 89) : « *Non là tu as calculé une probabilité, on te demande une proportion* ».
- A propos de l'exercice 95 : « *Essaye de traduire en lien avec les évènements A et F donnés de l'énoncé. Est-ce que l'on peut exprimer ça avec une probabilité liée à X ou à Y?* »
- A une autre élève (exercice 95) : « *X ne concerne que les femmes. Ici ça va être P(A sachant F). C'est seulement une traduction de l'énoncé en termes de variable aléatoire.* »
- A une autre élève encore (exercice 95) : « *La probabilité de A sachant F c'est la probabilité que X soit compris entre quoi et quoi? Je traduis ...* ».
- L'enseignante s'adresse aux élèves d'un autre îlot qui semblent avoir les mêmes difficultés : « *Ecrivez la probabilité que X soit plus grand que 170, cette probabilité exprimez-la avec les événements A et F.* »

De ces extraits nous pouvons dire que l'exercice 95 semble particulièrement difficile pour les élèves, notamment reconnaître l'évènement à la première question. Ils ont des difficultés à passer des évènements A et F à un évènement du type $\{X > t\}$. A la fin de la séance et pour aider les élèves à avancer ces exercices chez eux, l'enseignante écrit au tableau :

$$P_F(A) = P(X \geq 170)$$

Conclusion :

C'est la première rencontre (en exercice) pour les élèves avec **T-CalcSimple** dans le cas d'une loi normale. Complexité supplémentaire vis-à-vis de l'exemple du cours, l'énoncé (de l'exercice et de la question) est en langage naturel et l'évènement n'est pas nommé ni identifié, il y a donc un travail de modélisation à faire (voir détail dans la partie 5.2.3). L'enseignante prend ici en charge la reconnaissance et l'organisation de la première étape qui n'est pas indiquée dans l'énoncé : identifier la moyenne et l'écart-type de la loi normale proposée dans l'exercice.

Le second exercice mélange probabilité conditionnelle, arbre de probabilités et loi normale. La première question fait apparaître une nouvelle variation à la charge des élèves : *étant donné que la probabilité recherchée est présentée comme une probabilité conditionnelle*, cette question demande donc un travail d'interprétation initial afin notamment d'identifier les évènements en jeu.

Séance n°5

n°	Thème principal de l'épisode	Commentaires
1	Exercice 94 : T-CalcSimple calculer une probabilité du type $P(X < a)$ pour une loi normale de paramètres donnés	Détails dans ce qui suit
2	« Questions Rapides » sur la loi normale T-Det (u_α)	Détails dans ce qui suit
3	Fin du travail sur l'exercice 89 (entamé séance 4)	Détails dans ce qui suit
4	Travail sur les exercices 95, 100, 109 et 134 avec corrigés de l'enseignante : T-CalcSimple , T-DetInterv , T-Det ($\sigma + \mu$) et T-CalcProbaCond pour la loi normale ; T-Det ([a ; b]), T-CalcSimple et T-CalcEsp pour une densité f quelconque	Détails dans ce qui suit
5	Cours : « IV.3. Intervalles à 1, 2 ou 3 sigmas » comparaison avec la loi normale centrée réduite ($u_{0,05}$ et $u_{0,01}$) Théorème de Moivre-Laplace illustré sur géogebra	Voir polycopié de cours en annexe B.1

FIGURE 5.35 – Synopsis de la séance n°5 – *Lois continues*

Détails à propos de l'épisode n°1 :

La séance n°5 commence par l'exercice corrigé du manuel n°94 (voir énoncé figure 5.36).

94 La variable aléatoire X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$ avec $\mu = 2 \times 10^{100}$ et $\sigma = 10^{100}$.
Déterminer, à 0,01 près, la valeur de $P(X < 1,5 \times 10^{100})$.

94 $\mu = 2 \times 10^{100}$ et $\sigma = 10^{100}$.

Le calcul direct n'est pas possible à la calculatrice car les nombres sont trop grands.

$$\begin{aligned} P(X < 1,5 \times 10^{100}) &= \\ &P\left(\frac{X - \mu}{\sigma} < \frac{1,5 \times 10^{100} - \mu}{\sigma}\right) \\ &= P\left(T < \frac{(1,5 - 2) \times 10^{100}}{10^{100}}\right) \end{aligned}$$

où T suit la loi normale centrée réduite.

On en déduit :

$$\begin{aligned} P(X < 1,5 \times 10^{100}) &= P(T < -0,5) \\ &= 0,5 - P(-0,5 \leq T \leq 0) \\ &\approx 0,31. \end{aligned}$$

FIGURE 5.36 – Séance n°5 – Énoncé et corrigé de l'exercice 94

Un élève (C) fait la remarque suivante : « *Les nombres étaient trop grands* ». L'enseignante (A) lui répond : « *Et donc ? Ça plantait ! Qu'est-ce qu'on faisait néanmoins pour trouver la solution ?* »

A : « On avait une variable aléatoire X qui suit une loi normale d'écart-type et de moyenne donnés. Qu'est-ce qu'on faisait dans le cours ? »

C : « On ... sur $[0, 1]$ du coup on faisait plus petit que $P(X \dots 1,5)$ »

A : « Alors attention ce n'est plus X on change de variable aléatoire, tu dois l'appeler autrement, souvent on l'appelle Z »

L'enseignante s'adresse alors à toute la classe :

A : « On fabrique à partir de X la variable aléatoire centrée réduite qui lui est associée, donc si on se rappelle du cours qu'est-ce qu'on va faire subir à X pour obtenir la variable aléatoire dont parle Carl ? Qu'est-ce que l'on fait à une variable aléatoire X qui suit une loi normale, peu importe laquelle, pour passer à une autre variable aléatoire Z qui va être centrée réduite ? »

E2 : « On lui soustrait μ et on divise par σ »

A : « Soustraire μ c'est pour centrer sur l'axe des ordonnées, c'est pour que la moyenne soit à 0 et diviser par σ pour réduire pour que l'écart-type soit égal à 1 »

Conclusion :

Cet exercice permet à l'enseignante de proposer aux élèves le type de tâches **T- CalcSimple** calculer une probabilité du type $P(X < a)$, pour une loi normale de paramètres donnés, où a est tellement grand que la calculatrice, bien que ce soit une forme prise en charge par l'une de ses fonctionnalités, ne peut faire le calcul. Il faut donc centrer et réduire l'expression malgré tout.

Détails à propos de l'épisode n°2 :

La séance se poursuit par un moment de « Questions rapides », voici la question posée par l'enseignante : « On considère X une variable aléatoire qui suit la loi normale centrée réduite $\mathcal{N}(0, 1)$. En utilisant les notations de la leçon déterminer un réel positif $u_{0,2}$ tel que $P(-u_{0,2} \leq X \leq u_{0,2})$ soit égale à 0,8. Vous devez illustrer cette égalité par un dessin et déterminer une valeur approchée du réel $u_{0,2}$ au millième près. »

Il s'agit du type de tâche, déjà rencontré, **T-Det**(u_α) : déterminer u_α pour α un nombre donné. L'enseignante s'adresse à la classe : « L'aire, c'est-à-dire la probabilité, est donnée et on cherche les bornes. C'est forcément le 3e choix de la calculatrice, celui qui vous donne une borne quand on a l'aire. Il faut transformer l'expression en $P(X \leq \dots)$. »

Un élève demande à l'enseignante : « On cherche quoi en fait ? » ce à quoi l'enseignante répond : « On cherche la valeur de $u_{0,2}$, on cherche la borne. C'est une question de cours en principe. » Une élève est envoyée au tableau par l'enseignante

pour faire le dessin, voir figure 5.37.

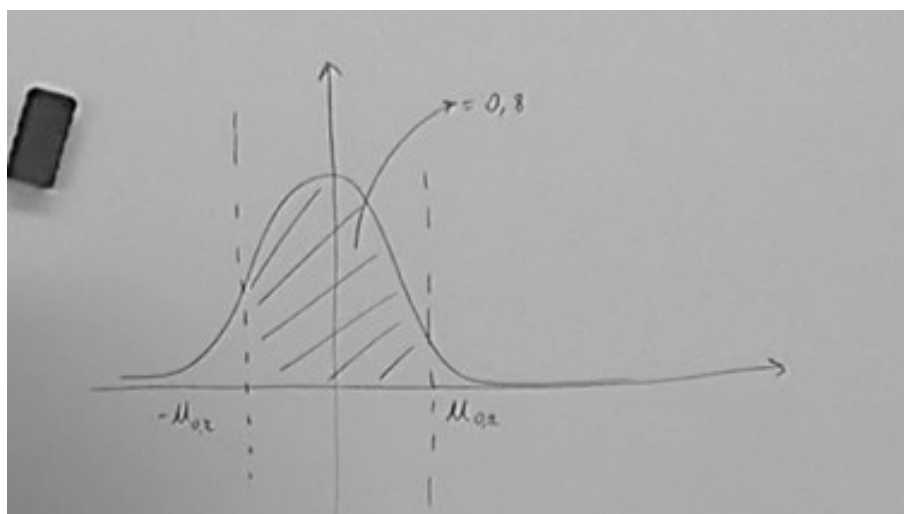


FIGURE 5.37 – Séance n°5 – « Questions rapides » – Correction d’une élève

L’enseignante s’adresse à l’élève qui est au tableau et lui demande d’indiquer que l’aire hachurée est égale à 0,8. Pour répondre à la question posée initialement un élève propose de faire *FracNormale(0,8)*, or cette fonctionnalité ne renvoie la borne a que pour des probabilités de la forme $P(X < a) = \text{nombre donné}$.

L’enseignante détaille la technique pour répondre ici : « *FracNormale de 0,8 ça vous donnerait une borne telle que toute cette aire là (elle hachure en rouge l’aire de moins l’infini à $u_{0,2}$) ça vous donne juste une borne supérieure; donc il faut connaître effectivement l’aire en rouge et si l’aire entre les 2 bornes on veut qu’elle soit égale à 0,8 en utilisant des propriétés de symétries on doit pouvoir trouver l’aire en rouge. Et là on pourra utiliser la calculatrice avec *FracNormale* », elle écrit au tableau :*

$$\text{FracNormale}(\quad ; 0, 1)$$

L’enseignante circule dans la classe et un nouvel élève (E) est envoyé au tableau pour finir de corriger cette question, il écrit :

$$\Phi(u_{0,2}) - 1 = 0,8$$

L’enseignante lui demande d’expliquer cette notation qui certes apparaît dans le manuel dans la partie de cours mais n’a pas été vue en classe avec l’enseignante. L’élève s’embrouille et n’arrive pas à répondre.

A : « *Qu’est-ce que c’est que cette notation ?* »

E : « *Perso, j’avais regardé des cours ... et j’avais vu des formules...* »

A : « *Oui mais ça représente quoi? Cette notation elle est dans votre manuel mais elle est pas expressément au programme donc je ne l’ai pas introduite dans la leçon, mais elle est très fréquente et vous allez la retrouver dans des exercices que j’ai choisis.* »

E : « Ça (il montre l'expression $P(-u_{0,2} \leq X \leq u_{0,2})$ inscrite au tableau) on sait que c'est égal à ça (il montre $\Phi(u_{0,2}) - 1$ qu'il vient d'écrire). »

A : « Mais pourquoi ? En fait tu ne comprends pas ce que tu as fait. »

L'élève rajoute un 2 devant $\Phi(u_{0,2})$ dans l'expression inscrite au tableau et explique : « Ensuite on fait passer ... en fait on veut juste le calcul de ... je ne sais plus son nom ... divisé par 2 est égal à 0,9 », il souligne $\Phi(u_{0,2})$ puis écrit à la suite :

$$\Phi(u_{0,2}) = \frac{1,8}{2} = 0,9$$

L'enseignante lui demande de compléter *FracNormale*(; 0, 1). L'élève complète par 0,9 et dit : « Du coup ce que l'on doit mettre ici c'est 0,9. Et après ça nous donne 1,282 ».

A : « Et donc $\Phi(u_{0,2})$ tu sais pas nous dire ce que c'est sur le dessin ? »

E : « Je sais juste qu'ils l'utilisent et voilà. »

A : « $\Phi(u_{0,2})$ est une notation qui veut dire, c'est la probabilité que $X \leq u_{0,2}$ »

Elle écrit au tableau : $\Phi(u_{0,2}) = P(X \leq u_{0,2})$ et explique que c'est la probabilité hachurée en rouge (i.e. l'aire sous la courbe de moins l'infini à $u_{0,2}$). Elle conclut : « Je peux alors utiliser la calculatrice avec le choix numéro 3 : *FracNormale*. Donc il faut réussir à passer de ça (elle montre $P(-u_{0,2} \leq X \leq u_{0,2})$) à ça (elle montre $P(X \leq u_{0,2})$) »

Un élève demande : « Mais 0,9 ça correspond à quelle aire ? », ce à quoi l'enseignante répond : « Justement à l'aire rouge, c'est ce que j'ai écrit là et c'est ce qu'il a écrit là » l'élève qui est toujours au tableau ajoute : « Vu qu'elle est centrée $u_{0,2}$ et $-u_{0,2}$ il y a autant de chaque côté. »

L'enseignante hachure sur le dessin l'aire qui correspond à la probabilité d'être inférieure ou égale à 0 et dit aux élèves : « Normalement vous savez combien ça fait » un élève répond 0,5. L'enseignante poursuit : « Et l'aire en rouge elle est égale à 0,5 plus la probabilité d'être entre 0 et $u_{0,2}$ que je cherche. Et donc par symétrie c'est 0,8 sur 2 donc 0,4. En additionnant ces 2 choses-là on retrouve le 0,9. On peut aussi le faire avec un dessin. » Et elle complète le dessin à retrouver figure 5.38.

En écrivant la dernière ligne : $P(X \leq u_{0,2}) = 0,4 + 0,5 = 0,9$, elle dit : « La calculatrice ne peut me donner que les bornes correspondant à ce genre de probabilités. On fait ça parce que c'est associé à notre outil calculatrice et c'est une fonction importante en probabilités. » L'enseignante termine cet exemple en faisant référence à une méthode alternative qui fait intervenir une formule du manuel : $\Phi(u_\alpha) = P(X \leq u_{0,2}) = 1 - \frac{\alpha}{2}$

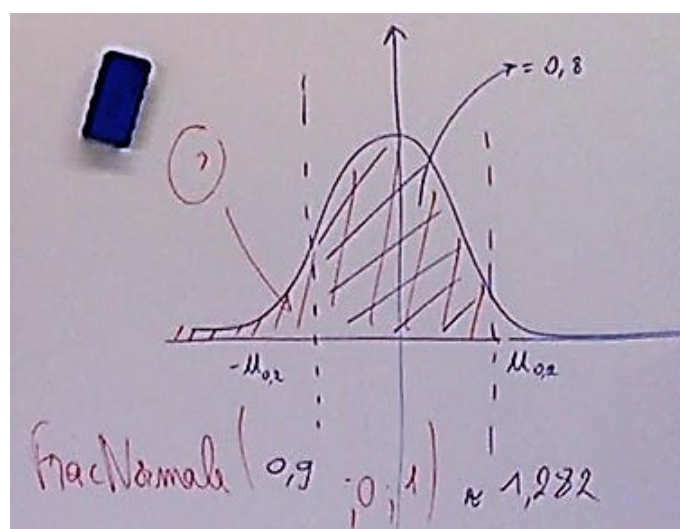


FIGURE 5.38 – Séance n°5 – Correction de l'exercice "Questions Rapides"

Conclusion :

Nouvelle rencontre avec **T-Det**(u_a), même configuration que dans l'exemple du cours et ici l'enseignante demande directement de traduire avec une « aire ». Ce type de tâches semble poser problème car aucune fonctionnalité de la calculatrice ne permet de répondre directement, il faut d'abord transformer l'expression.

L'enseignante prend ici en charge la reconnaissance de la tâche (l'aire est donnée, c'est-à-dire la probabilité, et on cherche les bornes), la reconnaissance de la méthode (c'est le troisième choix de la calculatrice, il faut transformer l'expression en $P(X \leq \dots)$) ainsi que le traitement de la tâche. L'enseignante explique non seulement comment transformer l'expression initiale mais aussi pourquoi, pour pouvoir utiliser la fonction de la calculatrice.

Cet exercice permet l'introduction involontaire (par un élève) de la notation $\Phi(u_a)$, qui sera plutôt utilisée à l'Université et qu'un élève essaye de manipuler (sans succès).

Détails à propos de l'épisode n°3 :

La séance se poursuit par l'exercice n°89 entamé lors de la séance précédente (voir figure 5.34).

L'enseignante ne revient pas sur le calcul de $P(85 < X < 95)$ mais elle souhaite détailler la justification du calcul final pour obtenir le nombre moyen de fruits acceptés parmi les 10000 fruits. Ci-dessous nous avons retranscrit l'échange entre l'enseignante et 5 élèves (noté R, E2, E3 et E4).

A : « Donc pour $P(85 < X < 95)$ vous aviez trouvé environ 0,9545. On n'avait pas corrigé cet exercice donc je voudrais revenir sur comment on finit de corriger la question : parmi 10000 fruits pris au hasard combien en moyenne seront

acceptés. Romain (R) je te laisse la parole pour finir de répondre à la question.
»

R : « *Fallait faire fois 10000.* »

A : « *Pourquoi ?* »

R : « *Parce qu'il y a 10000 fruits.* »

A : « *Ah oui seulement ?* »

R : « *Non parmi les 10000 pris au hasard, c'est pour la probabilité justement, on a fait fois 10000 pour savoir pour 10000 fruits qui seraient pris au hasard.* »

A : « *Tu m'avais parlé d'autre chose ?* »

R : « *J'allais faire la loi binomiale mais je pensais que ça n'allait pas marcher.* »

A : « *Pourquoi ça n'aurait pas marché ? Pourquoi il s'agit d'une loi binomiale ici ?* »

R : « *C'est la répétition de paramètre p d'expériences de Bernoulli totalement identiques et indépendantes pour chaque fruit.* »

A : « *Comment on justifie que c'est du hasard ? Il y a combien de répétitions ?* »

E2 : « *C'est du hasard.* »

A : « *Oui c'est comme une prise d'échantillons.* »

E3 : « *C'est indépendant.* »

A : « *Qu'est-ce qui est indépendant ?* »

E3 : « *On a pris 10000 fruits au hasard donc c'est-à-dire qu'on a pris un fruit puis un autre puis un autre indépendamment des uns des autres* »

A : « *Est-ce vraiment indépendant ? Est-ce que tu peux prendre deux fois le même fruit ?* »

E3 : « *Non.* »

A : « *Donc normalement quand on fait ça on n'aboutit pas à une loi binomiale, on a une loi binomiale quand on fait des tirages avec remise, c'est-à-dire qu'on peut prendre deux fois le même fruit. Pourquoi est-ce que dans cette situation d'échantillonnage, pourquoi on modélise quand même ça par une loi binomiale ?* »

E3 : « *Peut être parce qu'on peut avoir pris deux fruits qui ont la même masse ? C'est pas comme s'ils avaient tous une masse différente on peut tomber sur deux fruits qui ont la même masse.* »

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

E4 : « Sinon on peut introduire une nouvelle variable aléatoire qui va compter le nombre de fruits qui sont corrects du coup c'est une variable qui serait discrète et du coup ça rentre dans le cadre d'une loi binomiale.»

E3 : « Je voulais dire que c'est négligeable.»

A : « Effectivement c'est ça qui compte. Que l'on fasse ou non des remises, ce qu'on appelle des remises, quand le nombre total de fruits est suffisamment important on considère que ça ne va pas changer grand-chose, que la différence va être négligeable. En général il faut avoir 10 % je crois, il faut que la taille de l'échantillon représente moins de 10 % de la taille de la population totale pour qu'on puisse considérer que c'est un tirage avec remise même s'il est sans remise. Faut que vous ayez conscience de ça. Dans ce cas on peut considérer que l'on a 10000 fruits pris au hasard donc 10000 répétitions indépendante si l'on suppose que le nombre de fruits est suffisamment important au total ; donc soit Y la variable aléatoire égale au nombre de fruits qui sont acceptés alors Y suit la loi binomiale de paramètres, redites-moi ? »

R : « 10000 »

A : « Que vaut p , la probabilité de succès et le succès c'est que le fruit soit accepté ? »

E2 : « 0,9545 »

L'enseignante dicte finalement aux élèves la justification qu'elle attendait :

Y suit la loi binomiale de paramètres (10000; 0,9545) car il s'agit de 10000 répétitions identiques et indépendantes de la même épreuve de Bernoulli de paramètre 0,9545. Remarque : on considère que le nombre de fruits est suffisamment grand pour supposer que les répétitions sont indépendantes.

Elle termine en expliquant : « Et donc ici quand on vous demande le nombre moyen de fruits qu'est-ce qu'on vous demandait de calculer ? », un élève répond : « L'espérance » et l'enseignante conclut : « Qui est donnée par n fois p ! Voilà l'explication de pourquoi il fallait multiplier le résultat par 10000 » elle écrit ainsi au tableau :

$$Y \text{ suit } \mathcal{B}(10000; 0,9545) \text{ donc } E(Y) = n \times p$$

Conclusion :

Cet exercice, déjà présenté séance n°4 épisode n°6, combine à la fois **T-CalcSimple** pour une loi normale et **T-CalcEsp** pour une loi binomiale. Il s'agit de commencer par calculer la probabilité de succès au moyen d'une loi normale puis de conclure par des arguments relevant du chapitre sur les variables aléatoires discrètes et de calculer l'espérance de la loi binomiale adéquate.

Le type de tâches **T-CalcSimple** pour une loi normale, proposé ici ne semble pas poser de difficulté aux étudiants mais le traitement de la tâche est pris en charge par l'enseignante car elle demande des arguments et justifications plus détaillés que ce que propose l'élève.

Détails à propos de l'épisode n°4 :

La séance se poursuit par un travail en autonomie sur les exercices du manuel suivants : 95, 100, 109 et 134 dont les énoncés sont figure 5.39. L'enseignante a préparé un document « Parcours Lois Normales » regroupant les corrections de ces exercices et qui est distribué à chaque élève (voir polycopié en annexe B.2).

95 Les résultats seront arrondis au millième le plus proche. Dans un pays, la taille des femmes peut être modélisée par une variable aléatoire X suivant la loi normale de moyenne 165 cm et d'écart-type 6 cm alors que celle des hommes peut être modélisée par une variable aléatoire Y suivant la loi normale de moyenne 174 cm et d'écart-type 8 cm.

De plus, on sait que les femmes représentent 51,3 % de la population de ce pays.

On considère une personne prise au hasard dans cette population. On note A l'événement « la taille de cette personne dépasse 1,70 m » et F l'événement « cette personne est une femme ».

1. Calculer les probabilités $P_F(A)$ et $P_{\bar{F}}(A)$.
2. Utiliser un arbre pondéré pour déterminer la valeur de $P(A)$.
3. Quelle est la probabilité pour qu'une personne mesurant plus de 1,70 m soit une femme ?

109 Dans une usine de conditionnement, une machine remplit à la chaîne des bouteilles de jus de fruit

On note Z la variable aléatoire qui, à une bouteille prélevée au hasard dans la production associe la quantité versée par la machine, en centilitres.

On suppose que Z suit la loi normale de moyenne 70 et d'écart-type 1.

Les résultats seront arrondis au millième le plus proche.

1. Quelle est la probabilité pour que la bouteille prélevée contienne entre 68 et 72 centilitres de produit ?
2. Déterminer le réel h tel que :
$$P(70 - h \leq Z \leq 70 + h) = 0,99.$$
3. Calculer $P_{Z \geq 68}(Z \leq 72)$.

134 Densité de probabilité

Soit a un réel strictement positif et f la fonction définie sur $[-a; a]$ par $f(x) = e^{-|x|}$.

1. Déterminer a pour que f soit une densité de probabilité.
2. Soit X la variable aléatoire continue dont la fonction de densité de probabilité est la fonction f déterminée à la question 1.
 - a. Calculer $P(X > -0,5)$.
 - b. Soit g la fonction définie par $g(x) = xf(x)$.
Montrer que pour tout réel x appartenant à l'intervalle $[-a; a]$, on a $g(-x) = -g(x)$.
 - c. En déduire par des considérations géométriques que l'espérance mathématique de X est nulle.

PISTE : question 1. $\int_{-a}^a f(x) dx = \int_{-a}^0 f(x) dx + \int_0^a f(x) dx.$

100 La variable aléatoire X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$. Déterminer les valeurs arrondies à l'entier le plus proche de μ et σ sachant que :

$$P(X \geq 1100) = 0,9332 \quad \text{et} \quad P(X \leq 1600) = 0,8413.$$

FIGURE 5.39 – Séance n°5 – Énoncés des exercices 95, 100, 109 et 134

L'exercice 95 a déjà été proposé lors de la séance n°4 et analysé dans l'épisode n°6 de cette séance.

Les types de tâches (et leurs variations principales) présents dans l'exercice n°109 dans lequel il s'agit d'une loi normale de paramètres donnés, sont les suivants :

- La première question relève du type de tâche, déjà rencontré et traité par les élèves, **T-CalcSimple** : calculer la probabilité d'un événement simple du type $\{a < X < b\}$ pour une loi normale, avec les variations suivantes :
 - étant donné que la variable aléatoire est identifiée dans l'énoncé
 - à partir d'un énoncé en langage naturel
- La seconde question relève du type de tâche, jamais rencontré pour le moment, **T-DetInterv** : déterminer un intervalle I centré en l'espérance de X tel que $P(X \in I) =$ un nombre donné.

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

- La troisième question relève du type de tâche, jamais rencontré pour une loi normale, **T-CalcProbaCond** : calculer une probabilité conditionnelle du type $P_{X < a}(X > b)$.

L'exercice n°100 relève du type de tâche **T-Det($\sigma + \mu$)** : déterminer la moyenne et l'écart-type d'une variable aléatoire X suivant une loi normale. Ce type de tâches a déjà été rencontré en partie dans un exemple du cours lors de la séance n°3, il y avait d'abord un exemple avec **T-Det(σ)** puis un second avec **T-Det(μ)**, mais déterminer simultanément μ et σ est nouveau pour les élèves.

Nous avons relevé dans l'exercice n°134 les types de tâches et variations suivants :

- Question 1 : **T-Det([a;b])** : déterminer les bornes a et b de l'ensemble de définition d'une fonction f donnée afin que celle-ci représente une densité de probabilité. Type de tâches jamais rencontré pour une densité f quelconque.
- Question 2.a : **T-CalcSimple** : calculer la probabilité d'un événement simple du type $\{X > b\}$, étant donné que la variable aléatoire est identifiée dans l'énoncé et que sa densité a été donnée dans l'énoncé ou établie dans une question précédente.
- Question 2.b relève d'un type de tâches relatif au chapitre d'étude de fonctions, montrer qu'une certaine fonction g est impaire.
- Question 2.c : **T-CalcEsp** : calculer l'espérance mathématique d'une variable aléatoire continue dont la fonction de densité f est connue, à partir de considérations géométriques. Type de tâches déjà rencontré mais pas avec cette variation.

Conclusion :

Première rencontre avec les types de tâches **T-DetInterv** et avec **T-Det($\sigma + \mu$)** dans le cas d'une loi normale, avec **T-Det([a;b])** pour une densité f quelconque et avec **T-CalcEsp** à partir de considérations géométriques (densité quelconque).

Le document « Parcours Lois Normales » est, selon le discours de l'enseignante, à la fois un recueil d'exercices jugés importants, que les élèves doivent travailler chez eux s'ils n'ont pas eu le temps de les terminer en classe, et à la fois le polycopié des corrigés.

Séance n°6

n°	Thème principal de l'épisode	Commentaires
1	« Mathoral » extrait de sujet de baccalauréat n°D du manuel sur les lois normales T-CalcSimple et T-Det ($\mu + \sigma$)	Détails dans ce qui suit
2	« Questions Rapides » pour introduire la loi exponentielle : 1) Soit $A \in [0; +\infty[$, calculer en fonction de A : $\int_0^A 0,04e^{-0,04x} dx$ 2) Déterminer $\lim_{A \rightarrow +\infty} \int_0^A 0,04e^{-0,04x} dx$	X
3	Cours : « V. Loi exponentielle » : introduction avec Géogebra, lecture du polycopié, premières définitions, propriétés et exemples détaillés : T-CalcSimple et T-CalcEsp	Voir polycopié de cours en annexe B.1
4	Exercice 67 sur la loi exponentielle T-CalcSimple , T-CalcProbaCond et T-CalcEsp + T-Interp	Détails dans ce qui suit

FIGURE 5.40 – Synopsis de la séance n°6 – *Lois continues*

Détails à propos de l'épisode n°1 :

Une élève présente au tableau sa correction d'un exercice du manuel dans le cadre d'un « Mathoral ». L'énoncé de cet exercice se trouve dans la figure 5.41.

La première question correspond au type de tâches déjà rencontré dans le cadre de la loi normale, avec les variations suivantes :

- étant donné que la variable aléatoire est identifiée dans l'énoncé
- à partir d'un énoncé en langage naturel
- étant donné qu'il s'agit d'une loi normale de paramètres **non** donnés
- **calculer** $P(X < b)$ **étant donné qu'on connaît** $P(\mu < X < b)$

Les trois questions suivantes correspondent aux ingrédients de la technique pour résoudre **T-Det**($\mu + \sigma$). Pour rappel **T-Det**($\mu + \sigma$) est le type de tâches suivant : déterminer la moyenne et l'écart-type d'une variable aléatoire X suivant une loi normale lorsque deux probabilités du type $P(X < b)$ et $P(X > a)$ sont connues. C'est un type de tâches qui a déjà été traité par les élèves dans l'exercice n°100 lors de la séance n°5.

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

Sujet D

Capacité mise en œuvre

→ Déterminer les paramètres d'une loi normale

Dans une population, le taux de cholestérol en grammes par litre peut être assimilé à une variable aléatoire X suivant une loi normale $\mathcal{N}(\mu, \sigma^2)$. On note Φ la fonction définie pour tout réel t par $\Phi(t) = P(T < t)$, où T suit la loi normale centrée réduite.

1. On sait que, dans la population étudiée, la probabilité pour que le taux de cholestérol soit inférieur à 1,95 grammes par litre est 0,58 et la probabilité pour que le taux de cholestérol soit compris entre 1,95 et 2,10 grammes par litre est 0,38.

a. Calculer $P(X < 2,10)$.

b. Montrer que $\Phi\left(\frac{1,95 - \mu}{\sigma}\right) = 0,58$ et $\Phi\left(\frac{2,10 - \mu}{\sigma}\right) = 0,96$.

c. En déduire que μ et σ sont solutions du système :

$\begin{cases} 1,95 - \mu = a\sigma \\ 2,10 - \mu = b\sigma \end{cases}$ où a et b sont deux réels dont on donnera une valeur approchée à 10^{-4} près.

d. En déduire les valeurs approchées de μ et σ arrondies au centième le plus proche.

2. On prend désormais $\mu = 1,9$ et $\sigma = 0,1$. En admettant que les personnes dont le taux de cholestérol est supérieur à 2,15 grammes par litre doivent subir un traitement, quel est le marché potentiel pour un médicament permettant de réduire ce taux si cette population est composée de 100 000 individus ?

Sujet D

1. a. $P(X < 2,10) = P(X < 1,95) + P(1,95 < X < 2,10) = 0,96$.

b. $P(X < 1,95) = 0,58 \Leftrightarrow P\left(\frac{X - \mu}{\sigma} < \frac{1,95 - \mu}{\sigma}\right) = 0,58$
 $\Leftrightarrow \Phi\left(\frac{1,95 - \mu}{\sigma}\right) = 0,58$

car $\frac{X - \mu}{\sigma}$ suit la loi normale centrée réduite.

De même $P(X < 2,10) = 0,96 \Leftrightarrow P\left(\frac{X - \mu}{\sigma} < \frac{2,10 - \mu}{\sigma}\right) = 0,96$
 $\Leftrightarrow \Phi\left(\frac{2,10 - \mu}{\sigma}\right) = 0,96$.

c. On déduit de la question précédente $\frac{1,95 - \mu}{\sigma} = a$ et $\frac{2,10 - \mu}{\sigma} = b$ avec $a \approx 0,2019$ et $b \approx 1,7507$ d'où le système :

$$\begin{cases} 1,95 - \mu = a\sigma \\ 2,10 - \mu = b\sigma \end{cases}$$

d. $\mu \approx 1,93$ et $\sigma \approx 0,10$.

2. $P(X > 2,15) \approx 0,00621$ d'où un marché potentiel de 621 personnes.

FIGURE 5.41 – Séance n°6 – Extraits du manuel Indice

Énoncé du sujet D (à g.) – Corrigé issu du manuel du professeur (à d.)

L'élève commence par faire le dessin à retrouver figure 5.42 et explique comment elle a procédé pour répondre à la première question : « Dans l'énoncé on nous dit que la probabilité que le taux de cholestérol soit inférieur à 1,95 est de 0,58 et la probabilité que le taux de cholestérol soit compris entre 1,95 et 2,10 est 0,38. Du coup la probabilité que le taux de cholestérol soit inférieur à 2,10 c'est les 2 probabilités additionnées comme on le voit sur le dessin, du coup ça donne 0,96. »

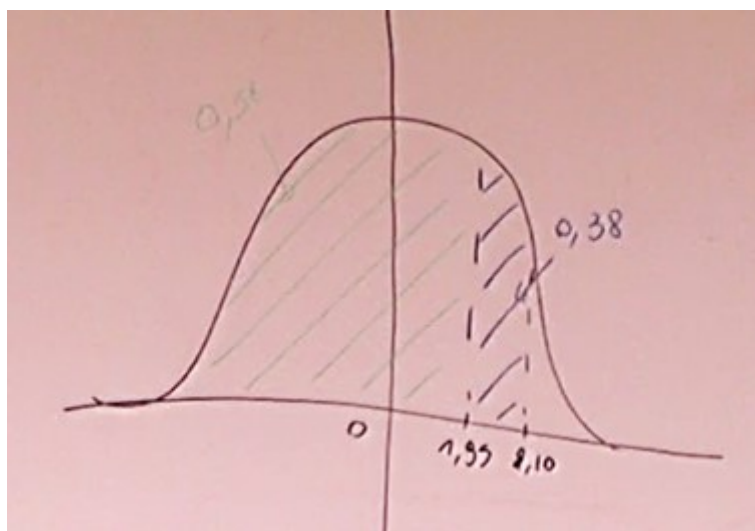


FIGURE 5.42 – Séance n°6 – Correction du sujet D

A propos de la deuxième question, elle explique : « On sait que $P(X < 1,95)$ c'est égal à 0,58 et que $P(X < 2,10)$ c'est 0,96. On sait que la variable aléatoire X suit la loi normale d'espérance μ et d'écart-type σ et que par définition $\frac{X-\mu}{\sigma}$ c'est la variable centrée réduite associée à X . Sachant que Φ représente la fonction $P(T < t)$, on peut dire que $\Phi\left(\frac{1,95-\mu}{\sigma}\right) = 0,58$ et pareil pour l'autre » et elle écrit :

$$\Phi\left(\frac{1,95-\mu}{\sigma}\right) = P(T < t)$$

L'enseignante la corrige : « Là c'est pas petit t », l'élève efface et reprend :

$$\Phi\left(\frac{1,95-\mu}{\sigma}\right) = P(T < X)$$

L'enseignante (A) : « C'est pas clair pour toi... », l'élève efface et reprend à nouveau :

$$\Phi\left(\frac{1,95-\mu}{\sigma}\right) = P(T < 1,95)$$

L'enseignante : « Alors, par définition... », l'élève (E) efface et corrige une nouvelle fois :

$$\Phi\left(\frac{1,95-\mu}{\sigma}\right) = P\left(T < \frac{X-\mu}{\sigma}\right)$$

A : « Regarde la définition de Φ dans l'énoncé »

E : « Du coup c'était 1,95 ? »

A : « Oui »

Elle écrit alors : $\Phi\left(\frac{1,95-\mu}{\sigma}\right) = P(T < 1,95)$ un autre élève l'aide et elle finit par corriger :

$$\Phi\left(\frac{1,95-\mu}{\sigma}\right) = P\left(T < \frac{1,95-\mu}{\sigma}\right) = 0,58$$

A : « Pourquoi est-ce que c'est 0,58 ? »

E : « Parce que c'est dit dans la question précédente »

A : « Non ce qui est dit c'est que c'est la probabilité de $X < 1,95$. Ça serait équivalent à quoi d'autre ? Est ce que tu peux l'écrire ? », l'élève écrit :

$$X < 1,95$$

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

A : « Il faut que tu fasses le lien entre X et t . Cet événement à quoi c'est équivalent ? »

E : « A la variable $\frac{X-\mu}{\sigma}$ » et elle écrit :

$$X < 1,95 \Leftrightarrow \frac{X - \mu}{\sigma} < 1,95$$

A : « Tu as soustrait μ au membre de gauche de l'inégalité et pas au membre de droite, et tu as divisé par σ à gauche mais pas à droite. Est-ce que tu as le droit de diviser par σ ? »

E : « Oui il est plus grand que 0 »

A : « Strictement »

L'élève corrige ce qu'elle a écrit :

$$X < 1,95 \Leftrightarrow \frac{X - \mu}{\sigma} < \frac{1,95 - \mu}{\sigma}$$

A : « Et donc ça c'est équivalent à quoi ? Tu n'as plus qu'à remplacer $\frac{X-\mu}{\sigma}$ par son autre nom, c'est-à-dire T », l'élève écrit.

A : « Je veux bien qu'au début de cette ligne tu écrives "car", c'est des questions qui ne sont pas faciles à rédiger pour vous. On a donc des événements qui sont équivalents donc les probabilités sont égales »

Dans la figure 5.43 se trouve ce que l'élève a écrit au tableau.

$$\Phi\left(\frac{1,95-\mu}{\sigma}\right) = P\left(T < \frac{1,95-\mu}{\sigma}\right) = P(X < 1,95) = 0,958$$
$$\text{car } X < 1,95 \Leftrightarrow \frac{X-\mu}{\sigma} < \frac{1,95-\mu}{\sigma} \Leftrightarrow T < \frac{1,95-\mu}{\sigma}$$

FIGURE 5.43 – Séance n°6 – Correction du sujet D

L'enseignante fait quelques corrections sur ce qui vient d'être écrit, en parlant du passage avec (*) elle explique : « C'est ça qu'il faut montrer. Là vous avez une équivalence parce qu'on a soustrait les deux membres d'une inégalité par un même nombre et divisé par un nombre strictement positif donc on ne change pas le sens. »

L'élève lit la troisième question (1.c) puis explique : « Pour trouver à peu près la valeur de $\frac{1,95-\mu}{\sigma}$ j'ai calculé avec la calculatrice et le choix 3 : FracNormale de

0,58 avec le μ égal à 0 et σ égal à 1. J'ai fait pareil pour la deuxième avec 2,10 », l'enseignante lui demande d'écrire au tableau et lui dicte ce qu'elle doit écrire :

$$\frac{1,95 - \mu}{\sigma} \simeq 0,2019$$
$$\frac{2,1 - \mu}{\sigma} \simeq 1,7507$$

L'enseignante demande aux élèves de bien préciser sur leur copies qu'ils ont utilisé la calculatrice et avec quelle fonction. L'élève poursuit ses explications : « Donc je suis repartie de ces deux équations là pour faire le système, ce qui me donne $a \simeq 0,2019$ et $b \simeq 1,7507$. Question suivante il fallait en déduire les valeurs approchées de μ et σ au centième, du coup j'ai résolu le système en soustrayant les deux lignes, ça me donne $\mu \simeq 1,9304$ et $\sigma \simeq 0,0968$ », elle écrit ces deux résultats numériques au tableau.

La cinquième question combine à la fois **T-CalcSimple** pour une loi normale et **T-CalcEsp** pour une loi binomiale (idem ex89 séance n°5), où **T-CalcSimple** est proposé avec les variations suivantes :

- étant donné que la variable aléatoire est identifiée dans l'énoncé
- à partir d'un énoncé en langage naturel
- étant donné qu'il s'agit d'une loi normale de paramètres μ et σ donnés

Pour cette question il s'agit de commencer par calculer la probabilité de succès au moyen d'une loi normale puis de conclure par des arguments relevant du chapitre sur les variables aléatoires discrètes et de calculer l'espérance de la loi binomiale adéquate.

L'élève explique comment elle a calculé $P(X \geq 2,15)$ puis représente cette probabilité sur le dessin. L'enseignante lui demande de détailler le calcul et l'élève écrit au tableau :

$$\begin{aligned} P(X \geq 2,15) &= 0,00621 \\ &= 1 - P(X \leq 2,15) \\ &= 1 - (0,5 + P(1,9 \leq X \leq 2,15)) \end{aligned}$$

L'enseignante indique que l'expression pouvait se décomposer d'une autre manière : « vous pouviez le faire plus rapidement, c'est 0,5 moins $P(1,9 \leq X \leq 2,15)$ »

L'élève poursuit : « On nous demandait pour une population de 100000 habitants du coup ça ramenait à une loi binomiale. Donc j'ai pris une variable aléatoire Y qui suit la loi binomiale de paramètres $n=100.000$ et $p=0,0061$ et j'ai calculé l'espérance », l'élève a refait comme dans l'exercice 89 et elle écrit au tableau :

$$E(X) = n \times p = 100000 \times 0,00621 = 621$$

L'élève conclut et interprète le résultat final trouvé : « Du coup l'espérance elle donnait 621, donc sur une population de 100000 individus, on peut espérer que le

marché potentiel soit d'environ 0,621 % de la population.»

L'enseignante demande à l'élève de détailler au tableau la résolution du système de la question précédente (1.d). Pendant ce temps un autre élève (E2) demande à l'enseignante : « *Est-ce ça change quelque chose d'écrire supérieur ou égal à 2,15 ou supérieur strict ?* »

L'enseignante l'envoie au tableau écrire les deux probabilités :

$$P(X \geq 2,15) \text{ et } P(X > 2,15)$$

L'élève (E2) repose sa question oralement à toute la classe. Un élève de la classe lui dit que c'est la même chose, E2 demande alors à l'enseignante laquelle il faut écrire sur sa copie, l'enseignante lui répond :

A : « *C'est la même chose mais ça dépend de X. C'est quoi X? Si c'est une variable aléatoire dans quel cas ces deux probabilités sont-elles égales? Ou dans quel cadre ?* »

E3 : « *Si c'est une variable aléatoire discrète* »

A : « *Non tu n'as pas de bol c'est si elle est continue justement !* »

E3 : « *Mais discret c'est pas continue ?* »

A : « *Non, justement ! C'est ... est-ce que je peux dire que c'est le contraire ? Enfin dans ta vie de lycéen tu as vu 2 types de variables aléatoires, les variables aléatoires discrètes et les continues. Donc il y a les variables aléatoires discrètes qui prennent des valeurs séparées et dans ce cas c'est a priori différent d'inclure ou non une borne. Par contre si ta variable aléatoire est continue ça va être parfaitement égal, parce qu'on est sur du calcul d'aire. Je te laisse revoir le premier paragraphe de la leçon.»*

Conclusion :

T-CalcSimple est un type de tâches qui a déjà été rencontré de nombreuses fois pour la loi normale, le fait que les paramètres ne soient pas donnés ici provoque une adaptation à la charge des élèves. Ils ne peuvent pas utiliser la calculatrice comme ils en ont l'habitude pour ce type de tâches mais doivent utiliser les valeurs de probabilités données dans l'énoncé et la relation de Chasles.

La première étape de **T-Det**($\mu + \sigma$) (centrer réduire et transformer l'expression) pose des difficultés aux élèves qui semblent liées à l'utilisation de la notation $\Phi(t)$. Nous relevons des confusions chez certains élèves entre variables aléatoires discrètes et continues et ce que cela implique sur les événements $X < a$ ou $X \leq a$.

A nouveau les élèves rencontrent **T-CalcSimple** (loi normale) combiné avec **T-CalcEsp** (loi binomiale), ça ne semble pas leur poser de problème.

L'enseignante demande aux élèves de préciser sur leur copies s'ils ont utilisé ou non la calculatrice et avec quelle fonction.

Détails à propos de l'épisode n° 4 :

La séance se termine par un travail en autonomie sur l'exercice 67 du manuel, dont l'énoncé se trouve dans la figure 5.44.

67 Un fabricant de jeux électroniques fabrique des consoles de jeux. On suppose que la variable aléatoire T qui représente la durée de vie en jours d'un composant d'une console de jeux suit la loi exponentielle de paramètre $\lambda = \frac{1}{700}$.

1. Quelle est la durée de vie moyenne d'un composant ?
2. Calculer la probabilité que le composant n'ait pas de défaillance durant les 120 premiers jours.
3. Calculer la probabilité que le composant soit encore en fonctionnement au bout de 2 ans.
4. Calculer la probabilité que le composant fonctionne encore au bout de 5 ans, sachant qu'il fonctionne encore au bout de 2 ans.
5. Au bout de combien de temps aura-t-on 10 % des composants en panne ?
6. En fait deux composants électroniques A et B de ce type sont montés sur la console.

On note T_A et T_B les variables aléatoires représentant les durées de vie de ces composants.

On suppose que les événements « $T_A > 300$ » et « $T_B > 300$ » sont indépendants.

Calculer la probabilité que la console fonctionne après 300 jours, les composants A et B étant montés en série.

AIDE : Lorsque les composants électroniques sont montés en série, il faut que les deux composants fonctionnent pour que la console fonctionne.

FIGURE 5.44 – Séance n°6 – Énoncé de l'exercice 67

La première question correspond au type de tâches **T-CalcEsp** + **T-Interp** : calculer l'espérance mathématique d'une variable aléatoire continue qui suit une loi exponentielle de paramètre λ donné, à partir d'un énoncé en langage naturel. Un élève est envoyé au tableau pour corriger cette question, il écrit :

$$\lambda = \frac{1}{700}$$
$$E(T) = \frac{1}{\lambda} = \frac{1}{\frac{1}{700}} = 700$$

La durée moyenne d'un composant est de 700 jours

La seconde question correspond au type de tâches **T-CalcSimple** d'une loi exponentielle de paramètre λ donné, à partir d'un énoncé en langage naturel. L'enseignante explique à une élève : « *S'il n'a pas de défaillance c'est qu'il continue à fonctionner* », cette élève semble avoir des difficultés à interpréter la seconde question et à identifier l'évènement recherché. Un élève passe au tableau écrire sa réponse à cette question :

$$2) P(T \geq 120) = e^{-\lambda \times 120} \simeq 0,842$$

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

L'enseignante lui demande : « *N'écris pas lambda ici, ou alors si tu écris avec lambda, rajoute une ligne avec la valeur de lambda* » l'élève complète :

$$= e^{-\lambda \times 120} = e^{-\frac{1}{700} \times 120} \simeq 0,842$$

Certains élèves s'interrogent sur le modèle à choisir du nombre de jours dans une année. L'enseignante tranche et explique que sur trois années on peut considérer 365 jours, mais que sur cinq années il faut considérer au moins une année bissextile.

L'enseignante circule dans la classe et un élève lui pose une question (inaudible) à laquelle elle répond : « *Si on cherchait plus petit que 120 oui, mais là on cherche plus grand que 120 ... qui n'a pas de défaillance, donc il n'est pas cassé, donc sa durée de vie elle est plus grande que 120 jours. Qu'est-ce qui se passe de 0 à 120 jours ?* » (suite inaudible du fond de la salle où se trouve la caméra). Il s'agit à nouveau de problèmes de reconnaissance de l'évènement dont on cherche à calculer la probabilité.

La troisième question est très similaire à la question précédente, il faut là encore identifier correctement, à partir de la variable aléatoire T de loi exponentielle, l'évènement dont on cherche la probabilité. Un autre élève est envoyé au tableau pour corriger cette question et il écrit :

$$3) P(T \geq 730) = e^{-\lambda \times 730} = e^{-\frac{1}{700} \times 730} \simeq 0,352$$

L'enseignante ajoute : « *On trouve environ 0,352 pour la probabilité que la durée de vie soit supérieure, où supérieure ou égale. Même si l'énoncé n'est pas clair on s'en fiche on est sur une loi continue* » elle fait allusion ici à une difficulté que rencontrent les élèves : distinguer ou non les évènements $\{X < a\}$ et $\{X \leq a\}$.

La quatrième question correspond au type de tâches **T-CalcProbaCond** : calculer une probabilité conditionnelle avec les variations suivantes :

- calculer : $P_{X>a}(X > b)$, étant donné qu'il s'agit, d'après l'énoncé, d'une loi exponentielle et en utilisant la propriété du vieillissement
- à partir d'un énoncé en langage naturel

L'élève qui est toujours au tableau, écrit ce qui suit :

$$4) P(T \geq 1826) = e^{-\lambda \times 1826} = e^{-\frac{1}{700} \times 1826}$$

L'enseignante l'interrompt : « *Je t'arrête Malo (M) parce que c'est 1826 que je voulais que tu expliques, et après je pense que tu n'as pas bien traduit la question mais on va refaire ça tous ensemble. Alors pourquoi 1826 ?* »

M : « *Parce qu'il y a cinq années et dans les cinq ans il y a obligatoirement une année bissextile* » pour justifier son calcul il écrit alors

$$4 \times 365 + 366$$

Sur demande de l'enseignante un autre élève (E) relit cette question.

A : « Avec une question libellée comme ça vous tombez tous dans le panneau j'y crois pas ! Relis en insistant sur LE mot »

L'élève (E) relit en insistant sur « sachant que », et explique :

E : « Du coup moi j'ai fait comme si c'était $P(A)$ fois $P(B)$ sur $P(A)$ »

A : « Alors c'est pas exactement ça la formule, il s'agit de quel type de probabilité ? »

E : « Conditionnelle »

A : « Revois ta formule alors ! »

E : « $P_A(B)$ c'est $P(A)$ fois $P(B)$ sur $P(A)$ non ? »

A : « C'est $P(A \text{ et } B)$ »

E : « Oui mais ça revient au même »

A : « Oui, quand les évènements sont ind... »

E : « Indépendants, mais ici c'est bon ils sont indépendants »

A : « Qu'est-ce que tu en sais ? Qu'est-ce qui te dit qu'ils sont indépendants ? »

E : « S'ils ne le disent pas c'est que c'est vrai »

L'enseignante met un terme à cet aparté. L'élève au tableau écrit ce qui suit :

$$4) P_{(T \geq 739)}(T \geq 1826) = \frac{P(730 \leq T \leq 1826)}{P(T \geq 730)} \simeq \frac{\quad}{0,352}$$

L'élève n'arrive pas à exprimer l'intersection et l'enseignante tente de l'aider, l'élève remplace $P(730 \leq T \leq 1826)$ par :

$$P(730 \leq T \leq 1826)$$

L'enseignante lui dit : « Ça ne correspond toujours pas à l'intersection des deux évènements. On le fait en français ? Ou on peut faire un dessin. Le plus facile c'est peut être en français, si T est plus grand que 730 et qu'il est aussi plus grand que 1826 comment se situe-t-il ? » L'élève est bloqué et l'enseignante le renvoie à sa place, un nouvel élève vient au tableau et complète :

$$\frac{P(T > 1826)}{P(T \geq 730)} \simeq \frac{0,926}{0,352} \simeq 0,209$$

Un élève semble ne pas avoir compris l'étape de l'intersection et l'enseignante lui ré-explique : « Pour l'intersection de ces deux événements en fait tu en as un inclus dans l'autre. »

L'enseignante termine en expliquant qu'il vaut mieux retarder de passer en valeur approchée le plus tard possible et garder l'expression avec des exponentielles jusqu'à l'avant-dernière étape (ce qui n'a pas été fait ici). Les élèves continuent à travailler en autonomie sur les dernières questions, c'est la fin de la séance et l'enseignante écrit au tableau :

$$5) P(T \geq t) = 0,90$$

Conclusion :

Première rencontre avec **T-CalcEsp** pour la loi exponentielle, d'abord dans un exemple du cours puis dans un exercice. La démonstration de $E(X) = 1/\lambda$ est une démonstration exigible qui fait partie des attentes institutionnelles de classe de terminale scientifique, mais dans les exercices la formule est utilisée directement.

Pour calculer une probabilité de la forme $P(a \leq X \leq b)$ dans le cas d'une loi exponentielle, les élèves ont vu en cours la propriété suivante qui fera office de technique de résolution pour le type de tâches **T-CalcSimple** :

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

Ici se pose le problème de la reconnaissance de l'évènement dont on cherche à calculer la probabilité et plus particulièrement la confusion entre probabilité conditionnelle et intersection d'évènements. Pour aider les élèves, l'enseignante traduit en français et insiste sur des mots bien particuliers pour activer des réflexes chez les élèves : « *sachant que* », reformulation de l'intersection, etc.

Nous avons relevé des confusions sur la définition d'une probabilité conditionnelle, sur le lien entre $P(A \cap B)$ et $P(A)P(B)$ lorsqu'il y a indépendance des évènements. L'enseignante souhaite que les élèves ne se servent de la calculatrice que pour obtenir le résultat final, or les élèves utilisent la calculatrice dès qu'ils le peuvent et sans réduire les fractions au préalable (avec des exponentielles par exemple).

Synthèse

Dans cette partie nous présentons une synthèse des résultats issus des analyses précédentes pour le thème des lois continues en classe de terminale scientifique au lycée.

Types de tâches et variations observées

Au cours de nos observations nous avons relevé, au sein des exemples et exercices traités en classe, les types de tâches et variations principales suivants :

- **T-Veriff** : vérifier qu'une fonction donnée est bien une densité de probabilités
- **T-DonF** : définir/donner la fonction de densité d'une variable aléatoire continue

- **T-CalcSimple** : calculer la probabilité d'un événement simple du type $\{X < a\}$, $\{X > b\}$ ou $\{a < X < b\}$
- **T-CalcProbaCond** : calculer une probabilité conditionnelle du type $P_{X < a}(X > b)$
- **T-CalcEsp** : calculer l'espérance mathématique d'une variable aléatoire continue
- **T-Det(a)** : déterminer le réel a tel que $P(X < a) = \text{nombre donné}$
- **T-Det(u_α)** : déterminer u_α pour α un nombre donné
- **T-DetInterv** : déterminer un intervalle I centré en l'espérance de X tel que $P(X \in I) = \text{nombre donné}$
- **T-Det(σ)** : déterminer l'écart-type d'une variable aléatoire X de loi normale
- **T-Det($\sigma + \mu$)** : déterminer la moyenne et l'écart-type d'une variable aléatoire X suivant une loi normale
- **T-Hach** : hachurer une certaine probabilité à partir de la représentation graphique de la fonction de densité
- **T-Interp** : interpréter le résultat

En comparant avec la grande variété de praxéologies présentées dans le chapitre 4 page 49, nous n'avons pas relevé ici les types de tâches suivants :

- **T-DonLoi** : définir/donner la loi d'une variable aléatoire continue
- **T-ExprEx** : exprimer $P(X > a)$ en fonction de λ le paramètre d'une loi exponentielle
- **T-Det(λ)** : déterminer le paramètre λ d'une loi exponentielle
- **T-CalcMoy** : calculer la moyenne de $P(X < a)$ et $P(Y < b)$ lorsque X suit une loi uniforme donnée et Y suit une loi normale donnée

En revanche nous avons relevé des nouveaux types de tâches et de nouvelles variations que nous n'avions pas obtenus à la suite des analyses menées dans le chapitre 4 page 49. Il s'agit des types de tâches (en caractères gras) et des variations (pour T-DonF et T-CalcProbaCond) suivants recensés comme "nouveau" par la suite :

- **T-DonF** étant donné qu'il s'agit d'une variable aléatoire qui suit une loi normale (ce qui s'apparente finalement à ce que l'on peut appeler une question de cours)
- **T-CalcProbaCond** dans le cas d'une loi uniforme et sans que la densité de la variable aléatoire ne soit donnée
- **T-CalcInter** : calculer une probabilité d'intersection de deux événements (cas d'une loi uniforme)
- **T-Det([a;b])** : déterminer les bornes a et b de l'ensemble de définition d'une fonction afin que celle-ci représente une densité de probabilités (cas d'une loi uniforme et d'une fonction f quelconque)

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

- **T-Det(μ)** : déterminer l'espérance d'une variable aléatoire X de loi normale (à partir d'une probabilité du type $P(a < X < b)$ donnée)

Ces résultats sont à retrouver dans le tableau 5.45 qui suit.

	T-Veriff	T-DonF	T-DonLoi	T-CalcSimple	T-Hach	T-ExprEx	T-CalcProbaCond	T-CalculInter	T-CalcEsp	T-Det([a,b])	T-Det(λ)	T-Det(a)	T-Det(u_{α})	T-DetInterv	T-Det(σ)	T-Det(μ)	T-Det($\sigma + \mu$)	T-Interp	T-CalcMoy
A partir d'un énoncé en langage naturel				X			-		X			X	X		-				-
Etant donné que c'est la première question de l'exercice				X					X										
Etant donné que la VA est identifiée dans l'énoncé et que sa densité est connue (énoncé ou établie à la question précédente)	X			X					X	X									
Etant donné que sa densité N'a PAS été donnée dans l'énoncé				X			X	X	X										
A partir de la représentation graphique de la fonction de densité	X			X	X				X										
Etant donné qu'il s'agit d'une loi uniforme		X		X			X	X	X	X									X
Etant donné qu'il s'agit d'une loi exponentielle		-		X		-	X		X				-						-
Etant donné qu'il s'agit d'une loi exponentielle de paramètre λ inconnu				-		-			-										
Etant donné qu'il s'agit d'une loi normale centrée réduite		X		X			-					X	X						-
Etant donné qu'il s'agit d'une loi normale $N(\mu; \sigma^2)$ de paramètres donnés				X			X					X		X					X
Etant donné qu'il s'agit d'une loi normale de paramètres μ et σ NON donnés				X											X	X	X		
Avec Z la VA centrée réduite associée à X de loi normale $N(\mu; \sigma^2)$				-											-				
Sans qu'une VA Z centrée réduite associée à X de loi normale $N(\mu; \sigma^2)$ ne soit définie/donnée															X	X			
Calculer $P(X > a)$ ou $P(a < X < b)$ étant donné qu'on connaît $P(X < a)$ et/ou $P(X > b)$				X															
Calculer $P(X > a)$ étant donné qu'on connaît $P(X < a - \mu)$ où X suit $N(\mu; \sigma^2)$				X															
Il s'agit d'un intervalle à 1, 2 ou 3 sigmas				-															
En utilisant un tableur				-								-	-						
En utilisant la propriété de vieillissement d'une loi exponentielle							-												
A partir de probabilités du type $P(a < X < b)$ ou $P(a < X)$ données ou déjà calculées									-	X	-				X	X	X	X	
A partir d'une espérance donnée ou déjà calculée				-		-													X

FIGURE 5.45 – "x" : variations de type de tâches relevées. "x" : nouveautés. "-" : variations absentes

Techniques instrumentées

La variation "*en utilisant un tableur*" relevée dans l'analyse du manuel Indice dans le chapitre 4 n'a pas été proposée aux élèves durant cette séquence. Les élèves n'ont pas utilisé de tableur durant les séances observées.

La place de la calculatrice est ici centrale. L'enseignante va même jusqu'à donner et écrire du savoir procédural (comment on s'en sert, quelles sont les fonctionnalités, etc.). Lors des premières rencontres avec cette technique instrumentée l'enseignante prend en charge la reconnaissance de la méthode qui parfois se résume au choix de la fonction de la calculatrice. Concernant la justification, l'enseignante attend des élèves qu'ils précisent sur leur copie s'ils ont utilisé ou non la calculatrice et avec quelle fonction.

La calculatrice sert aux calculs de probabilités, aux calculs d'espérances et de façon générale aux calculs d'intégrales. Elle est indispensable au début de la séquence car les élèves ne savent pas déterminer de primitives (uniquement pour les séances 1 et 2) et même lorsqu'ils pourraient s'en passer (séance n°6 loi exponentielle : les primitives ont été vues) la calculatrice est tout de même utilisée. Les élèves utilisent la calculatrice dès qu'ils le peuvent et sans réduire les fractions au préalable (avec des exponentielles par exemple).

Certaines fonctionnalités de la calculatrice permettent d'obtenir directement le résultat numérique lors du calcul de probabilités. En revanche, certaines formes de probabilités doivent être transformées pour être calculées au moyen de cet instrument (par exemple **T-CalcSimple**, **T-Det(u_α)** ou **T-Det(a)**). D'après ce que nous avons observé, ces transformations devront être justifiées par un dessin (représentation graphique de la densité) ou par un calcul (par exemple la relation de Chasles). Transformer l'expression, et ne pas juste rentrer un nombre dans la calculatrice, est fréquemment l'étape qui pose des difficultés aux élèves. Cette étape du traitement de la tâche est généralement prise en charge par l'enseignante qui explique non seulement comment transformer l'expression initiale, mais aussi pourquoi (pour pouvoir utiliser la fonction de la calculatrice).

L'enseignante accompagne toujours la technique instrumentée (calculatrice) d'explications géométriques, à partir du graphe de la densité. Elle insiste fortement sur l'utilisation de la calculatrice car les élèves devront être à l'aise avec cette technique, notamment dans le cas de la loi normale.

Nous retenons l'intervention de l'enseignante quant à l'utilisation de la table de calcul de la loi normale. En effet, à l'Université la technique pour calculer la probabilité d'un évènement du type $P(X < b)$ se fait par l'intermédiaire d'une table de calculs et plus d'une calculatrice comme dans le secondaire. Ainsi pour un même type de tâches il y a des techniques instrumentées différentes (nous y reviendrons dans la partie 5.3 page 194). C'est une forme d'accompagnement de la transition.

Changements de registres

L'enseignante prend également en charge les changements de registres. En effet, elle a pour habitude de relier le calcul de probabilités au calcul d'intégrales. Au-delà des variations, qu'elle prend en charge ou non, elle rajoute dans son discours systématiquement des éléments concernant la représentation graphique de la fonction et/ou le calcul d'intégrale qui en découle (le positionnement des bornes sur l'axe des abscisses, à quoi correspond l'aire sous la courbe etc.).

Au travers des exercices de ce chapitre, l'enseignante fait travailler les élèves sur le lien entre calculs d'aires, représentations graphiques et probabilités. Ce lien est présent tout le long du chapitre, notamment dans le discours de l'enseignante. L'enseignante fait notamment travailler les élèves sur le passage *intégrales vers probabilités* (et réciproquement), cette activité de traitement est par la suite à la charge des élèves (reconnaissance des bornes, hachurer sur le graphique la probabilité recherchée, etc.).

Les élèves semblent identifier facilement les aires correspondantes aux probabilités recherchées. Mais il semblerait que le passage *aires vers probabilités* soit plus difficile pour les élèves (associer une probabilité à une aire sous la courbe).

L'enseignante insiste sur la représentation graphique et la première étape du traitement, à savoir, reconnaître les bornes et hachurer sur le graphique la probabilité recherchée (**T-Calcsimple** avec densité quelconque) et cette étape du traitement semble devenir automatique pour les élèves. C'est encore plus le cas avec des exercices portant sur la loi normale et sur la transcription du problème d'un point de vue géométrique, les élèves doivent connaître la densité d'une loi normale et les propriétés de cette dernière, certains types de tâches sont à traiter d'abord avec un dessin.

De façon générale, l'enseignante demande de compléter tous les calculs par un dessin. La justification peut se faire soit avec un dessin et une phrase d'explication, soit avec une rédaction plus mathématique (expliciter le découpage de l'expression faisant intervenir l'évènement contraire par exemple). L'enseignante insiste fortement sur la résolution géométrique dans le cadre d'exercices relevant de la loi uniforme mais elle attend également des élèves une justification et un calcul détaillé. De façon générale elle attend des justifications pour les calculs de probabilités (éléments de technologie), on peut citer : "*par symétrie, c'est une loi continue donc l'intégrale sur le domaine vaut 1*".

Éléments saillants pour cet enseignement

Tout comme pour les probabilités conditionnelles, l'enseignante prend généralement en charge les variations lorsqu'elles sont rencontrées pour la première fois par les élèves. Du moins elle prend a minima en charge l'activité de reconnaissance de la tâche et de la méthode (commencer par un dessin, transformer

l'expression, utiliser la calculatrice).

Nous avons relevé plusieurs exercices qui mêlent questions sur la loi uniforme et questions sur la loi binomiale ou qui combinent à la fois **T-CalcSimple** pour une loi normale et **T-CalcEsp** pour une loi binomiale avec une forte modélisation à la charge des élèves (par exemple l'exercice 89 voir figure 5.34 page 164). Certains exercices mélangent probabilité conditionnelle, arbre de probabilités et loi normale.

Des analyses précédentes nous relevons certaines difficultés chez les élèves :

- dans une situation où l'énoncé est en langage naturel, difficultés à identifier l'évènement dont on cherche la probabilité (difficultés à traduire l'énoncé en langage naturel).
- difficultés quant à la reconnaissance de l'évènement dont on cherche à calculer la probabilité et plus particulièrement la confusion entre probabilités conditionnelles et intersections d'évènements.
- confusions sur la définition d'une probabilité conditionnelle, sur le lien entre $P(A \cap B)$ et $P(A)P(B)$ lorsqu'il y a indépendance des évènements.
- confusions concernant la distinction entre variables aléatoires discrètes et continues et ce que cela implique sur les évènements $X < a$ ou $X \leq a$.

5.2.3 Le cas particulier de la modélisation

Nous considérons ici l'activité de modélisation mathématique en général pour le thème des probabilités (ce peut être l'activité attendue d'après l'énoncé d'un exercice ou l'activité réelle des élèves). C'est ce que nous appelons la « **modélisation mathématique en probabilités** » ou plus brièvement « modélisation probabiliste ».

Nous en proposons une définition dans ce qui suit en nous appuyant sur les résultats de la partie 1.2 page 17 et notamment sur la description du processus de modélisation en sous-compétences proposée par Blomhøj et Jensen (2003) dont voici les principales étapes :

- (1) formulation de la tâche
- (2) production d'une représentation mathématique (modèle)
- (3) traduction des objets et des relations en mathématiques
- (4) utilisation de méthodes mathématiques pour obtenir un résultat mathématique
- (5) interprétation des résultats
- (6) validation du modèle

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

La modélisation probabiliste, comme nous la définissons dans ce travail, commence généralement par une situation aléatoire décrite en langage naturel. Il faut ensuite identifier les événements en jeu, les nommer et déterminer leurs probabilités. Selon la situation, l'utilisation d'un arbre ou d'un tableau de probabilités peut être pertinente ou non. Ainsi à l'étape (3) de la description de Blomhøj et Jensen (2003) nous ajoutons : *traduction des objets et des relations en mathématiques incluant éventuellement la représentation par un arbre ou un tableau de probabilités.*

La modélisation probabiliste mélange activité de reconnaissance (reconnaître la tâche à accomplir et la relier à un certain type de tâches pour laquelle une technique est connue), changements de registres (passer du langage naturel au formalisme des probabilités ou inversement) et enchevêtrements de techniques.

Le modèle est généralement déjà donné dans les exemples et exercices proposés en classe de terminale scientifique pour les thèmes des probabilités conditionnelles ou des lois continues. Ainsi les étapes (2) et (6) du processus de modélisation de Blomhøj et Jensen (2003) n'apparaissent pas dans l'activité de modélisation à ce niveau.

Nous revenons dans cette partie sur les interventions de l'enseignante du point de vue de la modélisation et particulièrement des praxéologies de modélisation, mais aussi sur l'activité des élèves dans des situations comme décrites ci-dessus.

En effet, la modélisation de situations de la vie réelle est une étape indispensable dans le fonctionnement de la connaissance scientifique. De plus, la probabilité est un domaine pour lequel des modèles mathématiques simples peuvent permettre de modéliser de manière puissante des situations parfois complexes.

Nous nous intéresserons dans la suite (partie 5.3 page 194) à la place qu'occupe l'activité de modélisation dans l'enseignement des probabilités en première année de biologie à l'Université.

L'enseignante du lycée ne parle pas explicitement de (processus de) modélisation à la classe (explication du choix du modèle, description des étapes comme évoqué ci-dessus, etc).

L'enseignante accompagne l'activité de modélisation probabiliste, face notamment à des exercices sur les probabilités conditionnelles comme l'exemple n°2 du photocopié de cours dont l'énoncé se trouve ci-dessous. Cet exercice a déjà été présenté plus haut (séance n°2 page 110).

Lors d'une épidémie chez des bovins, on s'est aperçu que si la maladie est diagnostiquée suffisamment tôt chez un animal, on peut le guérir, sinon la maladie est mortelle. Un test est mis au point et essayé sur un échantillon d'animaux dont 2 % sont porteur de la maladie. On obtient les résultats suivants :

- si un animal est porteur de la maladie, le test est positif dans 85 % des cas,
- si un animal est sain, le test est négatif dans 95 % des cas.

On choisit de prendre ces fréquences observées comme probabilités pour toute la population et d'utiliser le test pour un dépistage préventif de la maladie. Soient respectivement les événements M et T « être porteur de la maladie » et « avoir un test positif. »

1. Un animal est choisi au hasard, quelle est la probabilité que son test soit positif ?
2. Si le test du bovin est positif, quelle est la probabilité qu'il soit malade ?

Dans cet exemple on retrouve les différentes étapes de la modélisation probabiliste : formuler la tâche (identifier la probabilité recherchée), traduire l'énoncé pour associer aux événements identifiés leurs probabilités, mais aussi utiliser des méthodes mathématiques pour résoudre la tâche.

L'enseignante recommande aux élèves de commencer par modéliser la situation à l'aide d'un arbre, même lorsque ce n'est pas demandé dans l'exercice. C'est une aide procédurale que l'enseignante donne aux élèves régulièrement pour des exercices sur les probabilités conditionnelles qui s'y prêtent. Elle explicite comment construire l'arbre et comment le compléter. Enfin la dernière étape de la modélisation probabiliste pour des exercices de ce type est l'interprétation des résultats.

Dans le cas des lois continues, les exercices qui demandent un travail de modélisation probabiliste sont de la forme de ceux présentés dans la figure 5.46 (exercices issus de la séance n°4).

Dans ces exercices la loi est toujours donnée et la modélisation probabiliste revient généralement à formuler la tâche à partir de l'énoncé de la question (identifier ce qui est recherché), utiliser des méthodes mathématiques pour obtenir un résultat puis interpréter le résultat numérique.

Plusieurs échanges entre les élèves et l'enseignante nous permettent de dire que l'interprétation probabiliste des énoncés en langage naturel pose des difficultés aux élèves. En particulier lorsqu'il s'agit d'identifier les événements en jeu, certains élèves ont des difficultés à distinguer une probabilité conditionnelle et une probabilité d'intersection.

5.2. Enseignement des probabilités au secondaire : analyse du savoir enseigné

Sujet C

Capacités mises en œuvre

→ Calculer avec une loi uniforme

→ Calculer avec une loi binomiale

Les deux rives d'un estuaire sont reliées par des bateaux qui quittent la rive nord exactement toutes les 10 minutes.

M. Dulac séjourne sur la rive nord et traverse l'estuaire une fois par jour pour se rendre dans la partie sud. Son arrivé au point d'embarquement sur la rive nord se fait au hasard.

1. Le temps, en minutes, séparant l'arrivée de M. Dulac à l'embarcadère du prochain départ du bateau, définit une variable aléatoire T qui suit une loi uniforme.

a. Quel est le temps d'attente moyen de M. Dulac à l'embarcadère ?

b. Montrer que la probabilité qu'un jour donné M. Dulac attende plus de 7 minutes à l'embarcadère est $p = 0,3$.

2. M. Dulac séjourne 10 jours sur la rive nord.

Le nombre de jours où son attente, pour prendre le bateau, est supérieure à 7 minutes définit une variable aléatoire X .

On suppose que l'arrivée de M. Dulac à l'embarcadère se fait de façon indépendante d'un jour à l'autre.

a. Quelle est la loi suivie par X ? Déterminer $E(X)$.

b. Calculer, à 10^{-3} près, la probabilité que M. Dulac n'attende jamais plus de 7 minutes à l'embarcadère.

c. Calculer $P(X \leq 5)$ à 10^{-3} près.

Sujet C

1. T suit la loi uniforme sur $[0 ; 10]$.

a. Le temps d'attente moyen de Monsieur Dulac est de 5 minutes.

b. $P(T > 7) = 1 - P(T \leq 7) = 1 - \frac{7}{10} = 0,3$.

2. a. X suit la loi binomiale de paramètres $n = 10$ et $p = 0,3$.

b. $P(X = 0) = 0,7^{10} \approx 0,028$.

c. $P(X \leq 5) \approx 0,953$.

89 Un lot de kiwis a été calibré. On admet que la variable aléatoire qui à chaque kiwi, pris au hasard, associe sa masse en grammes, est une variable aléatoire X dont la loi de probabilité est la loi normale $\mathcal{N}(90 ; 6,25)$.

On accepte les fruits dont la masse est comprise entre 85 g et 95 g.

Parmi 10 000 fruits pris au hasard, combien en moyenne seront acceptés ?

FIGURE 5.46 – Extraits du manuel *Indice*
Énoncé du sujet C et son corrigé – Énoncé de l'exercice 89

Pour résumer

La modélisation d'un énoncé (partiellement) en langage naturel pose des difficultés aux élèves de terminale scientifique, notamment dans la reconnaissance des événements en jeu.

Nous soulignons le rôle médiateur de l'arbre dans l'activité de modélisation pour le champ des probabilités conditionnelles.

La reconnaissance de la tâche, c'est-à-dire lier une tâche proposée à un type de tâches dont on connaît une technique, par les étudiants est complexe et pas toujours réussie. En général les étapes de l'activité de modélisation sont prises en charge par l'enseignante du lycée.

5.2.4 Éléments caractéristiques indépendants des contenus

Certains éléments dans les pratiques de l'enseignante et dans les déroulements nous semblent caractéristiques de l'enseignement secondaire indépendamment du contenu de probabilités, c'est l'objet de cette courte partie.

Concernant l'enseignante, nous notons que nombreuses sont ses interventions et remarques en lien avec l'examen de fin d'année (baccalauréat) et ses attendus. Par exemple, elle dit : "vous avez des formules à connaître pour faire un exercice

de bac sur les probabilités" et oriente le travail vers des extraits de sujets de baccalauréat. Cette pratique semble propre au secondaire puisqu'à l'Université les enseignants ont généralement la main sur les sujets d'examens (ils les connaissent et même parfois les conçoivent). C'est donc assez différent du baccalauréat qui est un examen national dont les enseignants du secondaire ne connaissent pas au préalable le sujet.

En ce qui concerne les déroulements, les trois périodes « Mathoral », « Questions rapides », moments d'exercices alternés avec du cours, nous semblent caractéristiques de l'enseignement secondaire.

De la même manière, le photocopié de cours à trous est une modalité de déroulement propre au lycée qui a aussi un impact sur l'activité des élèves au secondaire. La pratique de photocopiés à trous observée dans cette classe de terminale n'a pas été observée à l'Université (comme nous le verrons dans la partie 5.3.3 page 252). Dans certains pays il y a des expérimentations (voir la recherche en Allemagne de Feudel et Panse (2021)) menées dans le supérieur, avec des photocopiés à trous, qui semblent porter leurs fruits pour aider les étudiants dans la prise de notes mais ce n'est pas une pratique que nous avons déjà observée en France.

5.3 Enseignement des probabilités en première année d'Université : le cas des étudiants de biologie

Nous avons eu l'opportunité d'observer les séances d'un module de probabilités à l'Université en première année de filière biologie. Le module de probabilités (28h au total) qui est proposé à ces étudiants se compose de sept séances de cours magistraux (CM) alternées avec sept séances de travaux dirigés (TD).

- Un CM dure deux heures et se déroule dans un amphithéâtre de 300 étudiants. L'enseignant projette des diapositives et les commente oralement à l'aide d'un micro, il y a assez peu d'interactions entre l'enseignant et les étudiants durant ces séances de cours. Il n'y a pas de photocopié de cours.
- Un TD dure également deux heures et les étudiants sont répartis en groupe de 30 à 40 individus. En fonction des groupes de TD les étudiants n'ont pas forcément le même enseignant en CM qu'en TD (il y a seulement deux ou trois groupes qui ont le même enseignant de CM et de TD). Le fonctionnement général des TD est le même pour tous les groupes car le déroulement s'appuie sur une feuille commune d'exercices. Certains enseignants corrigent tout au tableau, d'autres laissent plutôt les étudiants travailler en autonomie et s'assurent que les exercices sont compris par tous en circulant dans les rangs.

Le plan du cours de ce module est à retrouver en annexe C. Le programme de cours est assez ambitieux et dans les faits les deux dernières parties (partie 6 et partie 7) sont à peine abordées en CM et pas du tout en TD. Chaque cours magistral (CM) correspond à peu près à une ou deux grandes parties du plan. Dans le cadre de ce travail, nous avons assisté au CM n°2 et au TD n°2 concernant les probabilités conditionnelles et au CM n°4 et au TD n°4 concernant les lois continues.

Nous n'avons pas pu filmer ces séances mais à partir des notes prises au cours des séances nous avons produit des synopsis. Nous présentons dans ce qui suit les synopsis des séances, des analyses illustrées par des extraits de transcripts, puis une synthèse.

De la même façon que dans la partie précédente, les résultats du chapitre 4 page 49 guident notre analyse des séances à l'Université. Nous analysons les données recueillies selon la méthodologie présentée page 105 et nous complétons ces analyses en nous appuyant sur les résultats de la partie précédente, pour l'enseignement en classe de terminale.

5.3.1 Analyse du savoir enseigné pour le thème des probabilités conditionnelles

Dans cette partie nous présentons l'analyse du savoir enseigné pour le thème des probabilités conditionnelles en première année de biologie à l'Université.

Le synopsis du cours magistral auquel nous avons assisté se trouve figure 5.47. L'ensemble des diapositives de ce cours est à retrouver en annexe D.1.

Nous avons assisté à une séance de TD dont le synopsis se trouve figure 5.48. Cette séance a eu lieu trois jours après le cours magistral, l'enseignant qui est en charge de ce groupe n'est pas le même que celui qui a fait le cours magistral. Le groupe observé durant la séance de TD est constitué de 26 étudiants. L'enseignant procède de la façon suivante, il laisse les étudiants chercher en autonomie, circule dans les rangs et intervient auprès des étudiants qui le sollicitent. Il écrit au fur et à mesure une correction détaillée des exercices au tableau. Le TD sur les probabilités conditionnelles associé à ce cours correspond à la feuille d'exercices suivante (figure 5.49).

Nous présentons dans ce qui suit l'analyse des épisodes du CM et ceux du TD qui correspondent aux lignes grisées dans les synopsis (figure 5.47 et figure 5.48). En regard des résultats de la partie précédente sur l'enseignement des probabilités dans le secondaire, nous proposons, dans une dernière sous-partie, une synthèse en termes de comparaison et de transition pour le thème des probabilités conditionnelles à l'Université.

n°	Thème principal de l'épisode	Commentaires
1	2.1 Probabilités conditionnelles : définition, notation (diapo 1)	Détails dans ce qui suit
2	Exemples, modélisation et T-CalcProbaCond (diapos 2 et 3)	Détails dans ce qui suit
3	Formule des probabilités totales : définitions d'une partition de l'univers, proposition et démonstration. Cas $n=2$ (diapos 4, 5, 6)	X
4	Exemples et première rencontre avec T-CalcSimple . Utilisation de la formule des probabilités totales (diapo 7)	Détails dans ce qui suit
5	Formule de Bayes (diapo 8)	X
6	Exemples d'utilisation de la formule de Bayes (diapos 9 et 10)	Détails dans ce qui suit
7	2.2. Indépendance de deux événements (diapo 11)	X
8	Exemples d'étude d'indépendance de deux événements T-EtudInd (diapos 12, 13 et 14)	Détails dans ce qui suit
9	Indépendance et complémentaire (diapo 15)	X
10	Évènements mutuellement indépendants et exemples (diapos 16 et 17)	Détails dans ce qui suit

FIGURE 5.47 – Synopsis de la séance de cours magistral – *Probabilités conditionnelles*

n°	Thème principal de l'épisode	Commentaires
11	Exercice 1 : T-EtudInd pour deux configurations plutôt similaires d'une même situation (tirages parmi 12 boules puis parmi 13 boules). Ressemblances avec les exemples diapo 14 et 15 du cours.	Détails dans ce qui suit
12	Exercice 3 : T-CalcInter , T-CalcSimple (union de deux événements) et T-CalcProbaCond . Exercice proposé dans le test 2018 distribué à des lycéens et à des étudiants de biologie.	Détails dans ce qui suit
13	Exercice 5 : T-Assoc et T-CalcProbaCond	Détails dans ce qui suit
14	Exercice 7 : T-CalcProbaCond . Ressemblances avec l'exemple diapo 10 du cours.	Détails dans ce qui suit

FIGURE 5.48 – Synopsis de la séance de travaux dirigés – *Probabilités conditionnelles*

Feuille d'exercices 2

Exercice 1 Une urne contient 12 boules numérotées de 1 à 12. On en tire une au hasard, et on considère les événements : A = « tirage d'un nombre pair », B = « tirage d'un multiple de 3 » :

- (a) Les événements A et B sont-ils indépendants ?
- (b) Reprendre la question avec une urne contenant 13 boules.

Exercice 2. Une urne contient des jetons, dont chacun est rouge ou noir, et marqué A ou B. Lors du tirage d'un jeton, la probabilité d'en tirer un rouge est $\frac{3}{5}$, et d'en tirer un qui est rouge et A est p .

- (a) Que vaut la probabilité d'en tirer un qui est noir et marqué B ?
- (b) Pour quelles valeurs de p les événements «noir» et «B» sont-ils indépendants ?

Exercice 3 Xavier et Yvette ont deux durées de vie indépendantes, telles que $P(\text{Xavier vit encore 9 ans}) = \frac{4}{5}$, $P(\text{Yvette vit encore 9 ans}) = \frac{3}{5}$. Calculer les probabilités que :

- (a) Xavier et Yvette vivent encore 9 ans
- (b) L'un des 2 au moins vit encore 9 ans
- (b) Xavier seulement vit encore 9 ans
- (b) Xavier vit encore 9 ans sachant que l'un des 2 au moins vivra encore 9 ans.

Exercice 4. Une urne contient 3 boules blanches et 7 boules noires. Quand une boule est tirée, on la remet dans l'urne, avec 5 boules de la même couleur. On effectue ainsi deux tirages au hasard. Quelle est la probabilité que la première boule tirée soit noire sachant que la seconde est blanche ?

Exercice 5 Le gérant d'un magasin d'informatique a reçu un lot de boîtes de CD-ROM. 5% des boîtes sont abîmées. Le gérant estime que 60% des boîtes abîmées contiennent au moins un CD-ROM défectueux et que 98% des boîtes non abîmées ne contiennent aucun CD-ROM défectueux. Un client achète une boîte du lot. On désigne par A l'évènement : «la boîte est abîmée» et par D l'évènement «la boîte achetée contient au moins un disque défectueux».

- (a) Calculer les probabilités $P(A)$, $P(A^c)$, $P(D|A)$, $P(D|A^c)$, $P(D^c|A)$ et $P(D^c|A^c)$.
- (b) Le client constate qu'un des CD-ROM acheté est défectueux. Quelle est la probabilité pour qu'il ait acheté une boîte abîmée ?

Exercice 6. Un livre a une probabilité de $\frac{1}{3}$ de se trouver dans une commode comportant 5 tiroirs, et des chances égales de se trouver dans chacun des tiroirs.

- (a) On ouvre les 4 premiers tiroirs, sans le trouver ; quelle est la probabilité de le trouver dans le dernier tiroir ?
- (b) On ouvre les 2 premiers tiroirs, sans le trouver ; quelle est la probabilité de le trouver dans le dernier tiroir ? dans l'un des 3 derniers tiroirs ?

Indication : On notera T_i l'évènement «Le livre est dans le tiroir i », ceci pour i allant de 1 à 5, et T_0 «Le livre n'est pas dans la commode». Noter que ces six évènements forment un partition de Ω .

Exercice 7. Le quart d'une population est vacciné contre le choléra. Au cours d'une épidémie, on constate qu'il y a parmi les malades un vacciné pour 4 non-vaccinés. En plus, parmi les vaccinés il y a une personne sur 12 qui tombe malade. Quelle est la probabilité qu'un non-vacciné tombe malade ?

Exercice 8. Dans une population on trouve une proportion de 1 pour 10000 individus qui portent un certain virus. Il y a un test pour la présence de ce virus. Ce test n'est pas parfait : si un individu porte le virus, alors le test le détecte avec une probabilité de 0,99. Si un individu ne porte pas le virus, le test donne un résultat positif (erroné) avec probabilité de 0,001.

On tire un individu au hasard de la population, on lui fait passer le test, et le résultat est positif ; quelle est la probabilité qu'il porte vraiment le virus ?

FIGURE 5.49 – Feuille d'exercices – TD sur les probabilités conditionnelles

Détails des épisodes

Détails à propos de l'épisode n° 1 :

A propos du premier exemple en haut de la première diapositive (figure 5.50), l'enseignant fait le commentaire suivant : « On se demande : comment une information partielle sur le résultat de notre expérience modifie les probabilités des

2.1. PROBABILITÉ CONDITIONNELLE

Exemple : Quelle est la probabilité d'avoir un diabète ? Que devient cette probabilité si vous buvez 12 cannettes de coca par jour ?

Définition

Étant donnés deux événements A et B , avec $P(B) > 0$, on définit la probabilité de A sachant B par :

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

Remarque : Si B est fixé l'application $A \mapsto P(A|B)$ est une probabilité. En partic. $P(\Omega|B) = 1$.

FIGURE 5.50 – Diapositive n°1 – Cours sur les probabilités conditionnelles

autres éléments ? C'est un exemple relatif aux probabilités conditionnelles, comment une information supplémentaire modifie la probabilité qu'un évènement se réalise. »

Concernant la nouvelle notation de la probabilité conditionnelle qui est la suivante :

$$P(A|B)$$

L'enseignant précise : « Vous avez vu dans les notations la semaine dernière que les barres obliques signifient autre chose, il ne faudra pas confondre », il écrit ensuite au tableau :

$$P_B(A)$$

et indique : « Certains d'entre vous vont utiliser cette notation là, mais je vous déconseille de l'utiliser. Elle a des avantages, mais plutôt pour des étudiants en maths. Pour vous, je pense qu'il est préférable de garder $P(A|B)$ qui vous rappelle qu'il y a une division dans la définition d'une probabilité conditionnelle et qui permet d'avoir des moyens mnémotechniques de se souvenir des formules qu'on va voir aujourd'hui. »

Détails à propos de l'épisode n°2 :

Dans les deux diapositives suivantes (voir figure 5.51) sont présentés les premiers exemples de probabilités conditionnelles. Si ces exemples proposent tous les deux le type de tâches **T-CalcProbaCond** déjà vu au lycée, c'est surtout pour l'aspect modélisation qu'ils sont ici détaillés par l'enseignant.

A propos de l'exemple de la diapositive 2, l'enseignant commence par décrire Ω l'univers de l'expérience : « C'est l'ensemble des numéros qui apparaissent sur les faces du dé, un ensemble à six éléments. Vous avez une phase "modélisation" de l'expérience avec Ω qui donne toutes les issues possibles et P votre application qui donne les probabilités de tous les évènements. Cette application on la connaît rarement et

5.3. Enseignement des probabilités en première année d'Université : le cas des étudiants de biologie

2.1. PROBABILITÉ CONDITIONNELLE EXEMPLE

On lance un dé. A =« Obtenir 6 », B =« Obtenir un chiffre pair ».

$$P(B) = \frac{1}{2}, \quad P(A \cap B) = P(A) = \frac{1}{6}, \quad P(A|B) = \frac{1}{3}.$$

Que vaut $P(B|A)$?

2.1. PROBABILITÉ CONDITIONNELLE EXEMPLE

Un lac contient 90 brochets, 9 truites, et 1 esturgeon. On pêche un poisson au hasard (l'équiprobabilité est supposée). Soient les événements

T =« Attraper une truite »,
 A =« Ne pas attraper un brochet ».

Alors

$$P(A) = \frac{10}{100}$$
$$P(A \cap T) = P(T) = \frac{9}{100}$$
$$P(T|A) = \frac{9}{10}.$$

$$P(A|T) = 1 \text{ car } T \subset A.$$

Rq.1 : On a considéré que $\Omega = \{B, T, E\}$, B =« On pêche un brochet » etc., avec $P(B) = 90\%$, $P(T) = 9\%$, et $P(E) = 1\%$.

Rq.2 : Le nombre exact de poisson est inutile, on aurait fait pareil Sous l'hypothèse « Un lac contient 90% de brochets, 9% de truites, et 1% d'esturgeons ».

FIGURE 5.51 – Diapositives n°2 et n°3 – Cours sur les probabilités conditionnelles

on veut connaître des choses dessus. Ici on sait qu'aucune face n'est privilégiée, sinon on nous aurait dit que le dé était pipé ou truqué. Donc P est la probabilité uniforme, ce qui veut dire que la probabilité d'obtenir 1 c'est la même que celle d'obtenir 2 etc. Comme il y a six issues et que la probabilité d'obtenir une des six issues c'est 1, la probabilité d'obtenir un nombre c'est $1/6$.» Il écrit au tableau :

Modélisation par :

$$\Omega = \{1, \dots, 6\}$$

P est la probabilité uniforme

Il poursuit : « *Tous les exercices commencent de la même manière, soit on vous donne dans l'énoncé Ω et P , soit vous essayez de les déterminer, on essaye de décrire au maximum. Dans ce modèle, l'évènement A est décrit par un ensemble à un élément, 6. L'évènement B est décrit par une partie de Ω qui contient seulement les nombres pairs : 2, 4 et 6. » Il écrit à la suite :*

$$A = \{6\} \text{ et } B = \{2, 4, 6\}$$

L'enseignant poursuit en expliquant comment retrouver $P(A|B)$: « *Et après vous faites les calculs, c'est un exemple très simple. Vous calculez l'intersection, les évènements à la fois dans A et B il n'y a que le 6, donc l'intersection de A et B c'est A . Et ensuite on applique notre formule, $P(A \cap B)$ divisé par $P(B)$, ça fait $1/3$. On a quelque chose de raisonnable malgré une définition abstraite. Sur un exemple facile, si on sait qu'on a un nombre pair on a le choix entre 2, 4 et 6, la probabilité d'avoir 6 ça doit donc être $1/3$. Si on n'avait pas trouvé $1/3$ par le calcul ça veut dire qu'on aurait dû changer la définition.»*

Il termine en demandant aux étudiants : « *Que vaut $P(B \text{ sachant } A)$? Ça vaut 1 ! Si on sait que le résultat c'est 6 alors c'est sûr qu'il est pair. Maintenant vous devez aussi être capable de l'écrire, c'est-à-dire le calculer avec la formule. Mais on constate que la définition est cohérente avec l'intuition. »*

Pour le second exemple, l'enseignant commence de la même façon, par traduire l'énoncé qui est en langage naturel : « *Ce qu'il faut comprendre dans cette phrase c'est que ce lac ne contient pas d'autres types de poissons, c'est pas très clair dans le langage usuel, mais pour nous ça voudra dire "ne contient que ces trois types de poissons dans ces proportions là". On aimerait connaître la probabilité d'attraper une truite quand on n'a pas attrapé un brochet. Les calculs sont des calculs que vous avez déjà faits au lycée, ce qui change c'est qu'on va vous demander d'abord de modéliser par un ensemble Ω et une probabilité P , avant de faire les calculs.* »

A propos des deux remarques de la fin de la diapositive, il explique : « *Ici on a deux manières de choisir un modèle, ce n'est pas unique. Il y a un modèle proposé dans la remarque 1 et un autre remarque 2. Remarque 1, on considère que le résultat de l'expérience c'est l'espèce de poisson, on a donc trois résultats possibles. L'énoncé nous donne les valeurs des probabilités des événements élémentaires, $P(T)$, $P(B)$ et $P(E)$, il faut encore décrire l'évènement A . Est-ce que quelqu'un peut me dire qui est A ?* »

Un étudiant répond : « *Ne pas attraper un brochet* »

L'enseignant : « *Oui, mais dans mon modèle, j'ai un ensemble avec B , T et E . Qui sera A ?* »

Un étudiant répond : « *T et E* » et l'enseignant écrit au tableau :

$$A = \{T, E\}$$

L'enseignant termine : « *Ensuite on finit les calculs comme vous avez l'habitude de faire. Si l'on regarde le modèle de la remarque 2, dans ce modèle Ω est l'ensemble de tous les poissons du lac, alors même si on ne connaît pas le nombre exact de poissons dans le lac on connaît les pourcentages, 90% sont des brochets, etc. Maintenant B , T et E se sont des ensembles, B c'est l'ensemble des poissons qui sont des brochets* » et il écrit :

Ω = l'ensemble de tous les poissons du lac

$B \subset \Omega$ = l'ensemble des brochets

$T \subset \Omega$

$E \subset \Omega$

L'enseignant : « *Maintenant qui est P ? Vous prenez un poisson dans un lac, vous n'avez aucune raison de privilégier un poisson plutôt qu'un autre, donc on va prendre la probabilité uniforme* » et il écrit :

P est la probabilité uniforme

L'énoncé donne $|B| = \frac{90}{100} |\Omega|$ donc $P(B) = \frac{90}{100} = 0,9$

L'enseignant conclut : « *Ce n'est pas du tout le même modèle qu'au-dessus, au-dessus Ω contient seulement trois éléments, ici il y en a vraiment beaucoup, d'après l'énoncé il y en a 100. Dans le premier modèle le résultat de l'expérience c'est*

une espèce, dans le second modèle c'est un poisson bien déterminé avec un génotype et tout ce qu'il faut. Une fois que le modèle est posé on peut faire les calculs, je vous laisse faire pour les esturgeons.»

L'enseignant ne détaille ici que la modélisation de la situation, le choix du modèle et les conséquences que cela a. On peut imaginer qu'il ne détaille pas le calcul de la probabilité conditionnelle car il considère que cela est suffisamment familier pour les étudiants qui sortent du lycée.

Détails à propos de l'épisode n°4 :

2.1. PROBABILITÉ CONDITIONNELLE

EXEMPLE

Un lac contient des gardons et des truites. Parmi les gardons, il y en a 30% en âge de se reproduire, et parmi les truites il y en a 60% en âge de se reproduire. On sait qu'il y a 80% de gardons dans le lac. On pêche un poisson (avec équiprobabilité). Quelle est la probabilité d'attraper un poisson en âge de se reproduire ?

Ω est l'ensemble des poissons du lac. Ils ont chacun la même probabilité d'être pêché. On pose

R = « Pêcher un poisson en âge de se reproduire »
G = « Pêcher un gardon »
T = « Pêcher une truite ».

Alors

$$P(R) = P(R|G)P(G) + P(R|T)P(T) = \frac{30}{100} \frac{80}{100} + \frac{60}{100} \frac{20}{100} = \frac{36}{100}$$

FIGURE 5.52 – Diapositive n°7 – Cours sur les probabilités conditionnelles

Un nouvel exemple est présenté diapositive n°7 (voir figure 5.52). L'enseignant insiste à nouveau sur l'interprétation de l'énoncé et le modèle qui en découle : « *La première phrase signifiera toujours pour nous que le lac ne contient QUE des gardons et des truites. On prend un poisson au hasard et on se demande quelle est la probabilité d'avoir pêché un poisson capable de se reproduire.* » Cet exemple propose aux étudiants, pour la première fois dans ce cours, le type de tâches **T-CalcSimple** rencontré au lycée de nombreuses fois. Les variations présentes ici sont les suivantes :

- à partir d'un énoncé en langage naturel
- sans que les événements aient été explicitement nommés dans le texte
- étant donné que c'est la première ou la seule question de l'exercice
- les données sont en pourcentages

Etant donné qu'il s'agit d'un exemple de cours, l'enseignant prend en charge l'organisation et le traitement de la tâche. Concernant la modélisation de la situation, il explique : « *On commence par expliciter le modèle que l'on va utiliser.*

L'énoncé nous dit qu'on attrape un poisson, donc le résultat de l'expérience c'est "un poisson bien particulier" et l'ensemble des résultats possibles c'est l'ensemble de tous les poissons. Vous pouvez vous intéresser seulement à l'espèce, et à sa capacité à se reproduire ou non, mais on va faire ça dans un second temps. La première chose c'est qu'est-ce que vous avez dans la main et la réponse c'est un poisson bien particulier. Il peut y avoir plein de caractéristiques, la longueur de ses nageoires, s'il a des parasites, ce que vous voulez, donc il est préférable de commencer par avoir le résultat le plus précis possible. Donc Ω ça va être l'ensemble des poissons et on nous précise dans l'énoncé qu'il s'agit d'équiprobabilité donc vous prenez la probabilité uniforme : tous les poissons peuvent être pêchés avec la même probabilité. »

L'enseignant prend du temps pour bien expliquer comment modéliser la situation. Il poursuit : « *Donc on a R , je vous rappelle qu'un événement c'est une partie de Ω , et c'est la partie qui est constituée des poissons en âge de se reproduire. Pareil pour G et T , ce sont des parties de Ω » et il écrit au tableau :*

$$\begin{aligned}R &\subset \Omega = \text{l'ensemble des poissons capables de se reproduire} \\G &\subset \Omega \\T &\subset \Omega\end{aligned}$$

L'enseignant continue : « *L'énoncé vous dit la chose suivante : la probabilité de pêcher un gardon, comme vous pêchez un poisson de façon uniforme, c'est le nombre de gardons divisé par le nombre de poissons. L'énoncé vous dit que c'est 0,8. L'énoncé nous dit aussi que la probabilité de pêcher une truite c'est le nombre de truites sur le nombre de poissons, et c'est 20 % » et il écrit :*

$$\begin{aligned}P(G) &= \frac{|G|}{|\Omega|} = \frac{80}{100} \\P(T) &= \frac{20}{100}\end{aligned}$$

L'enseignant passe des données en pourcentages, aux données en fréquences relatives, il parle de probabilité uniforme sans rappeler le lien avec l'équiprobabilité. Il poursuit l'interprétation de l'énoncé : « *Ensuite on nous dit dans la deuxième phrase "PARMI les gardons IL Y EN A 30 % en âge de se reproduire", donc si vous regardez le nombre de poissons qui sont en âge de se reproduire ET qui sont des gardons vous allez obtenir 30 % de la population de gardons, 30 % du nombre de gardons» et il écrit :*

$$|R \cap G| = \frac{30}{100}|G|$$

« *Ça signifie que la probabilité de R sachant G ... l'énoncé vous donne la proportion "PARMI les gardons", des poissons en âge de se reproduire » et il écrit :*

$$P(R|G) = \frac{P(R \cap G)}{P(G)} = \frac{\frac{|R \cap G|}{|\Omega|}}{\frac{|G|}{|\Omega|}} = \frac{30}{100}$$

D'après les résultats de la partie précédente (partie 5.2.1) ce n'est pas de cette façon que les étudiants ont appris à passer du registre de la langue naturelle

au registre symbolique probabiliste. On peut imaginer qu'ils auraient traduit directement la phrase "parmi les gardons, 30% sont en âge de se reproduire" comme étant : $P_G(R) = 0,3$

L'enseignant ne détaille pas la suite des calculs, et notamment celui de $P(R)$ par la formule des probabilités totales, il laisse cette application numérique aux étudiants.

Détails à propos de l'épisode n°6 :

2.1. PROBABILITÉ CONDITIONNELLE

EXEMPLE

Deux machines M_1 et M_2 produisent respectivement 100 et 200 pièces. M_1 produit 5% de pièces défectueuses et M_2 en produit 6%.
Quelle est la probabilité qu'une pièce défectueuse soit produite par M_1 .

Soit D l'événement « la pièce est défectueuse ». Que dit l'énoncé ?

$$P(M_1) = \frac{1}{3}, \quad P(M_2) = \frac{2}{3}, \quad P(D|M_1) = \frac{5}{100}, \quad P(D|M_2) = \frac{6}{100}$$

Alors

$$P(M_1|D) = \frac{P(D|M_1)P(M_1)}{P(D)} = \frac{P(D|M_1)P(M_1)}{P(D|M_1)P(M_1)+P(D|M_2)P(M_2)} = \frac{5}{17} \simeq 29\%.$$

FIGURE 5.53 – Diapositive n°9 – Cours sur les probabilités conditionnelles

Un exemple d'utilisation de la formule de Bayes est présenté aux étudiants diapositive n°9 (voir figure 5.53). L'enseignant commence par modéliser la situation : « Voici un problème avec des machines qui construisent des pièces défectueuses ou non, on tire une pièce défectueuse et on veut savoir de quelle machine elle provient. Je vais à nouveau rappeler qui sont Ω et P dans ce contexte. Donc quelle expérience on fait ici ? On prend une pièce au hasard dans une grande boîte de 300 pièces, donc le résultat de notre expérience c'est une pièce bien identifiée parmi ces pièces-là. Dans Ω , c'est-à-dire dans la boîte, il y a 300 pièces » et il écrit :

$$\Omega = \text{l'ensemble de toutes les pièces produites}$$
$$|\Omega| = 300$$

Il poursuit l'explication du modèle choisi pour représenter la situation : « Maintenant vous allez différencier certaines parties de Ω . Donc on a M_1 qui est à la fois le nom de la machine mais on va aussi donner le même nom à l'ensemble des pièces produites par la machine, et on vous dit qu'il y a 100 éléments dans M_1 . Je vous laisse faire pour M_2 . On a aussi un autre ensemble, l'ensemble D des pièces défectueuses » il écrit à la suite :

$M_1 \subset \Omega$ = l'ensemble des pièces produites par la première machine

$|M_1| = 100$

$M_2 \subset \Omega$ = l'ensemble des pièces produites par la deuxième machine

$|M_2| = 200$

$D \subset \Omega$ = l'ensemble des pièces défectueuses

A nouveau la modélisation de la situation fait apparaître la probabilité P uniforme : « *Donc ça c'est la description de vos pièces. Ensuite vous prenez une pièce au hasard, comme ici il n'y a pas d'indication sur la manière dont vous choisissez votre pièce, on considère qu'elle est choisie au hasard, totalement aléatoire, donc c'est la probabilité uniforme. Cela veut dire que je regarde le nombre d'éléments qu'il y a dans M_1 et je divise par le nombre total d'éléments. C'est la loi de probabilités la plus simple.* » L'enseignant passe ensuite très rapidement sur les quatre probabilités qui résultent de l'interprétation de l'énoncé : $P(M_1)$, $P(M_2)$, $P(D|M_1)$ et $P(D|M_2)$.

L'enseignant explique ensuite l'intérêt de s'intéresser ici à $P(M_1|D)$: « *Si je vous donne une pièce au hasard, vous ne pouvez pas savoir si elle vient de M_1 ou M_2 car elles sont toutes dans le même grand bac, par contre vous pouvez tester facilement si elle est défectueuse ou non. Du coup, si elle est défectueuse, vous vous demandez quelle est la probabilité qu'elle vienne de la machine M_1 ? C'est ce qu'on peut facilement tester ici.* »

Il passe très rapidement sur le détail des calculs pour aboutir à $P(M_1|D)$ et indique aux étudiants : « *Je vous laisse faire les calculs, étant donné que vous n'aurez pas le droit à la calculatrice au contrôle continu et à l'examen, entraînez-vous à obtenir ce résultat. Ça c'est très classique, la seule différence avec ce que vous faisiez pour le bac c'est qu'ici on veut que vous écriviez ces phrases là* » il fait référence ici aux phrases qui modélisent la situation : c'est-à-dire décrire l'univers et la probabilité P .

L'exemple de la diapositive n°10 (voir figure 5.54) est très similaire à l'activité d'approfondissement du manuel Indice qui présente justement à cette occasion la formule de Bayes. A nouveau, l'enseignant justifie l'intérêt de ce genre de question, pourquoi cette probabilité en particulier est intéressante à connaître dans ce type de situation et répond au problème posé : « *C'est un exemple classique que vous avez sans doute déjà rencontré, un représentant de commerce qui vous dit que son test est extraordinaire parce que si vous prenez 100 personnes atteintes par la maladie le test sera positif 99 fois et si vous prenez 1000 personnes qui ne sont pas atteintes par la maladie il ne sera positif que cinq fois. Ce test est très efficace du coup je le vends cher et il faut me l'acheter. C'est un exercice très classique, il faut savoir ce que vous voulez acheter, est-ce que vous voulez un test qui fonctionne sur les personnes déjà malades, que vous savez déjà malades? Mais si elle est déjà malade, ça ne sert à rien de lui faire passer le test. Donc l'information "le test est positif sachant que la personne est déjà malade" n'est d'aucune utilité. Ce qui nous intéresse c'est, si je fais*

2.1. PROBABILITÉ CONDITIONNELLE

EXEMPLE

Une maladie touche une personne sur 10 000. Un test de dépistage détecte 99% des personnes infectées avec seulement 0,5% de faux positifs. Quelle est la probabilité qu'une personne dont le test est positif soit effectivement infectée ?

Soient les événements M être infecté par la maladie et T avoir un test positif. Ces deux événements sont deux parties de la population totale Ω . On sait

$$P(M) = \frac{1}{10\,000}, \quad P(T|M) = \frac{99}{100}, \quad P(T|M^c) = \frac{5}{1000}.$$

Alors

$$P(M|T) = \frac{P(T|M)P(M)}{P(T)} = \frac{P(T|M)P(M)}{P(T|M)P(M)+P(T|M^c)P(M^c)} = \frac{\frac{99}{100} \times 10^{-4}}{\frac{99}{100} \times 10^{-4} + \frac{5}{1000} \times (1-10^{-4})} \simeq 1,94\%.$$

FIGURE 5.54 – Diapositive n°10 – Cours sur les probabilités conditionnelles

le test à une personne et que je ne sais pas si elle est malade ou non, le test est positif, est-ce que la personne est réellement malade ou pas ? La probabilité que la personne soit malade sachant que le test est positif.»

Il décrit ensuite la modélisation probabiliste de la situation :

Ω : la population étudiée

$M \subset \Omega$: personnes porteuses de la maladie

$T \subset \Omega$: personnes qui réagissent + au test

L'enseignant, à nouveau, décrit la modélisation et l'interprétation de la question mais il laisse les calculs à la charge des étudiants : « *On aimerait que M et T ce soient les mêmes gens en fait ! Les données de l'énoncé sont écrites sur la diapo. Nous on cherche la probabilité d'être malade quand le test est positif. Donc on sort la formule de Bayes et vous remplacez par ce que ça vaut.* »

Il conclut en interprétant le résultat final : « *On trouve moins de 2% ! Donc malgré ce qu'on nous disait, ce test quand il est positif on a moins de 2% de chance d'être vraiment malade, donc ce test ne vaut rien finalement. En fait c'est le 1 sur 10000 qui fait que ça ne marche pas bien. Si c'était une maladie rare qui touche une personne sur deux ça serait très différent. Ici très peu de gens sont malades, c'est ça qui fausse le test »*

Détails à propos de l'épisode n°8 :

L'enseignant présente diapositive n°12 (voir figure 5.55) un second exemple de **T-EtudInd** : étudier l'indépendance de deux événements, avec les variations suivantes :

2.2. ÉVÉNEMENTS INDÉPENDANTS

EXEMPLE

On jette deux fois un dé. $\Omega = \{1, \dots, 6\}^2$. Soient les événements

$A = \text{« La somme des résultats vaut 4 »} = \{(1, 3), (2, 2), (3, 1)\}$

$B = \text{« Le premier lancé donne 1 »} = \{(1, 1), \dots, (1, 6)\}$

A et B sont-ils indépendants ?

$$P(A)P(B) = \frac{3}{36} \frac{6}{36} = \frac{1}{72}.$$

$$P(A \cap B) = P((1, 3)) = \frac{1}{36}.$$

FIGURE 5.55 – Diapositive n°12 – Cours sur les probabilités conditionnelles

- à partir d'un énoncé en langage naturel
- étant donné que les événements ont déjà été décrits et nommés dans le texte

L'enseignant décrit l'expérience de la diapositive n°12, un premier travail d'interprétation de l'énoncé et de modélisation de la situation est nécessaire, il faut décrire les événements en jeu selon le modèle choisi, déterminer la probabilité de ces événements, puis enfin utiliser le critère (ou la définition) d'indépendance de deux événements.

Il explique : « *Il y a six éléments dans B . Est-ce que les événements sont indépendants ? La probabilité ici c'est la probabilité uniforme même si elle n'est pas explicitée, il n'y a pas de résultats favorisés par rapport à un autre. A contient trois éléments parmi les 36, donc sa probabilité c'est 3 sur 36 et B c'est 6. Vous faites les calculs. Quels sont les éléments de A qui commencent par 1 ? Il n'y en a qu'un ! Et la probabilité d'obtenir $(1, 3)$ c'est 1 sur 36 donc on compare les deux et on voit que les événements ne sont pas indépendants. Par contre ils sont positivement corrélés, la réalisation de l'un des deux va augmenter la probabilité que l'autre se réalise. C'est un exemple avec des dés c'est facile.* » A nouveau, de nombreuses choses importantes sont dites oralement et tout n'est pas écrit au tableau, bien qu'il s'agisse de nouveautés (dans le cadre de ce module) pour les étudiants.

A propos de deux exemples similaires (diapos n°13 et n°14, voir annexe D.1) qui présentent **T-EtudInd** dans une configuration légèrement différente (fratrie de 2 et fratrie de 3) l'enseignant indique : « *Les événements A et B sont décrits par les mêmes phrases et sont ici indépendants alors qu'ils ne l'étaient pas dans l'exemple précédent, parce que l'on n'est pas tout à fait dans la même situation. C'est pour ce genre de petites choses qu'il est important de décrire avant un modèle, et de savoir ce sur quoi on travaille.* »

Détails à propos de l'épisode n° 10 :

<p style="text-align: center; color: #008000;">2.2. EVÉNEMENTS INDÉPENDANTS</p> <p style="text-align: center; color: #4169E1;">INDÉPENDANCE DE n ÉVÉNEMENTS</p> <div style="background-color: #E0FFE0; padding: 5px; border: 1px solid #808080; margin-top: 10px;"> <p>Définition</p> <p>Soit $(A_n)_{n \geq 1}$ une suite d'événements, ils sont indépendants si pour tout ensemble fini d'indices (i_1, \dots, i_k) distincts, on a</p> $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k}).$ <p>Attention, cette propriété est plus forte que l'indépendance deux à deux qui demande simplement $A_i \perp\!\!\!\perp A_j$ pour tous $i \neq j$.</p> </div>	<p style="text-align: center; color: #008000;">2.2. EVÉNEMENTS INDÉPENDANTS</p> <p style="text-align: center; color: #4169E1;">EXEMPLE</p> <p>On lance un dé deux fois de suite.</p> <p>A = « Première fois pair » B = « Deuxième fois pair » C = « La somme est paire »</p> <p>Alors $A \perp\!\!\!\perp B$, $A \perp\!\!\!\perp C$ et $B \perp\!\!\!\perp C$ mais les événements ne sont pas indépendants car</p> $P(A \cap B \cap C) = P(A \cap B) = \frac{1}{4}$ <p>mais</p> $P(A)P(B)P(C) = \frac{1}{2} \frac{1}{2} \frac{1}{2} = \frac{1}{8}.$
---	--

FIGURE 5.56 – Diapos n°16 et n°17 – Cours sur les probabilités conditionnelles

Le cours se termine avec une dernière définition présentée diapositive n°16 (voir figure 5.56). Il s'agit de l'indépendance de n événements, c'est nouveau pour les étudiants au vu de ce qu'ils ont fait au lycée. L'enseignant insiste sur la remarque (être 2 à 2 indépendants c'est moins fort que mutuellement indépendants) et il détaille le cas $n = 3$:

Dans Ω, \mathbb{P} , on considère trois événements : A_1, A_2 et A_3 . Ces trois événements sont indépendants si :

- $P(A_1 \cap A_2) = P(A_1)P(A_2)$
- $P(A_1 \cap A_3) = P(A_1)P(A_3)$
- $P(A_2 \cap A_3) = P(A_2)P(A_3)$
- $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$

Diapositive n°17 (voir figure 5.56) se trouve un exemple qui illustre ce nouveau concept. Le cours touche à sa fin et l'enseignant demande aux étudiants de décrire eux-mêmes le modèle " Ω et \mathbb{P} la probabilité uniforme", d'écrire la partie de Ω qui traduit l'évènement A (idem pour B et C), de calculer la probabilité de A, B et C , puis la probabilité de A inter B, A inter C et B inter C . Il termine : « *Ils sont indépendants deux à deux je vous laisse le vérifier. Ils vérifient les trois premières conditions, par contre, si A et B sont réalisés ça va très fortement influencer la réalisation de C , en fait l'intersection de A, B et C c'est seulement l'intersection de A et B . C'est l'exemple typique d'évènements qui ne sont pas mutuellement indépendants même s'ils sont deux à deux indépendants.* »

Exercice 1 Une urne contient 12 boules numérotées de 1 à 12. On en tire une au hasard, et on considère les événements : A = « tirage d'un nombre pair », B = « tirage d'un multiple de 3 » :

- (a) Les événements A et B sont-ils indépendants ?
- (b) Reprendre la question avec une urne contenant 13 boules.

FIGURE 5.57 – TD sur les probabilités conditionnelles – Énoncé de l'exercice 1

Rappel : Si A et B sont deux événements avec $P(B) > 0$ alors la probabilité de A sachant B est définie par :

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

Les événements A et B sont indépendants lorsque

$$P(A|B) = P(A)$$

$P(A \cap B) = P(A)P(B)$ Critère pour savoir si deux événements sont indépendants

On considère deux événements :

A = le tirage d'un nombre pair

B = le tirage est un multiple de 3

(a) le cas d'une urne contenant 12 boules :

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

On a $P(A) = \frac{6}{12} = \frac{1}{2}$ $P(B) = \frac{4}{12} = \frac{1}{3}$ et $P(A \cap B) = \frac{2}{12} = \frac{1}{6}$

Ainsi $P(A)P(B) = \frac{1}{6} = P(A \cap B)$ donc les événements A et B sont indépendants

(b) l'urne contient 13 boules :

$P(A) = \frac{6}{13}$, $P(B) = \frac{4}{13}$ et $P(A \cap B) = \frac{2}{13}$

Cette fois $P(A)P(B) \neq P(A \cap B)$ donc les événements ne sont pas indépendants

FIGURE 5.58 – TD sur les probabilités conditionnelles – Exercice 1 – Correction de l'enseignant

Détails à propos de l'épisode n° 11 :

La séance de TD associée à ce cours commence par l'exercice 1 (voir figure 5.57) pour lequel la correction écrite par l'enseignant au tableau est à retrouver figure 5.58. L'enseignant de ce TD n'est pas le même que celui qui a dispensé le cours magistral.

La première question de l'exercice 1 relève du type de tâches **T-EtudInd** : étudier l'indépendance de deux événements avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les événements ont déjà été décrits et nommés dans le texte

– les données sont en fréquences naturelles

Étant donné que les données sont en fréquences naturelles et que **T-EtudInd** est la première et unique question de l'énoncé, de nombreux ingrédients de techniques interviennent dans la réalisation de la tâche (d'après le chapitre 4) :

- Tout d'abord identifier les événements en jeu et déterminer leur probabilité selon les données de l'énoncé : **T-Assoc**, mais dans le cas où les données de l'énoncé sont en fréquences naturelles, un ingrédient de la technique est le type de tâches **T-CalcSimple-FN**
- **T-CalcSimple-FN** : calculer la probabilité d'un événement, à partir d'un énoncé en langage naturel (étant donné que les données sont en fréquences naturelles). Pour cela il faut identifier qu'il s'agit d'une situation d'équiprobabilité et calculer la probabilité des événements comme le quotient du nombre de cas favorables sur le nombre total de cas.
- Enfin, dernière étape, comparer le produit des probabilités des événements avec la probabilité de l'intersection des événements.

L'enseignant demande aux étudiants de reprendre la définition d'indépendance de leur cours puis explique la méthode à suivre pour résoudre **T-EtudInd**. L'enseignant prend ici en charge la sous-activité d'organisation de la tâche : « Ici, vous avez deux événements A et B , facilement vous déterminez la probabilité que A se réalise, la probabilité que B se réalise et la probabilité que A et B se réalisent. Vous regardez ensuite si dans le cas de la question (a) la probabilité c'est le produit des probabilités $P(A)$ fois $P(B)$, si c'est le cas ils sont indépendants, si ce n'est pas le cas ils ne sont pas indépendants. Le fait que A se réalise n'a pas d'influence sur le fait que B se réalise, ça se traduit par le fait que la probabilité conditionnelle de A sachant B c'est la même que la probabilité de A . »

Un étudiant a du mal à calculer $P(A \cap B)$ (**T-CalcSimple-FN**), l'enseignant lui demande : « Comment vous calculez la probabilité d'avoir à la fois un nombre pair et un multiple de 3 ? Ça se calcule très facilement »

Après avoir rappelé la définition d'une probabilité conditionnelle l'enseignant fait quelques rappels sur la notion d'indépendance (définition + critère) et notamment sur son interprétation : « Intuitivement, le fait que deux événements soient indépendants c'est le fait que l'un des deux événements se produise ou pas n'a pas d'impact sur la probabilité que l'autre se réalise, ça veut dire quoi ? C'est une notion qui apparaît dans le cours et qui est cruciale pour ce qui concerne les probabilités conditionnelles : les événements A et B sont indépendants lorsque la probabilité que A se réalise sachant que B se réalise c'est la probabilité de A »

L'enseignant prend en charge la reconnaissance de la méthode, l'organisation et le traitement pour **T-EtudInd**. En effet, il commence par décrire les événements : « On tire une boule au hasard, on suppose qu'on a autant de chance de tirer une boule plutôt qu'une autre car les boules sont identiques, on considère deux événements... » et il écrit :

On considère deux évènements :

A = le tirage d'un nombre pair

B = le tirage est un multiple de 3

(a) le cas d'une urne contenant 12 boules :

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

Il entoure ensuite en rouge les nombres de la liste qui sont des multiples de 3 et en bleu les multiples de 2, il prend en charge le traitement de **T-Calcsimple-FN** car il explique : « *On a six nombres pairs et 4 multiples de 3, alors comme il n'y a pas plus de chance de tirer une boule plutôt qu'une autre, quelle est la probabilité d'obtenir une boule dont le numéro est pair ? Bah c'est 6 chances sur 12. Et un multiple de 3 ? 4 chances sur 12. Ceux qui sont à la fois des multiples de 2 et des multiples de 3, cette probabilité là c'est 2 sur 12* » puis il écrit ce qui suit. Il ne parle pas ici d'univers Ω ni de probabilité uniforme P .

$$\text{On a } P(A) = \frac{6}{12} = \frac{1}{2} \quad P(B) = \frac{4}{12} = \frac{1}{3} \quad \text{et} \quad P(A \cap B) = \frac{2}{12} = \frac{1}{6}$$

Ainsi $P(A)P(B) = \frac{1}{6} = P(A \cap B)$ donc les évènements A et B sont indépendants

L'enseignant corrige ensuite la seconde question qui relève du même type de tâches mais avec une configuration légèrement différente (treize boules au lieu de douze), il faut donc reprendre la modélisation. Il prend à nouveau en charge les sous-activités d'organisation et de traitement, il conclut par : « *Dans cette histoire-là, même si on pourrait avoir le réflexe de se dire que ça ne change pas, et bien en fait si car vous changez complètement la configuration numérique. Cet exercice illustre la notion d'évènements indépendants.* »

Détails à propos de l'épisode n° 12 :

Exercice 3 Arthur et Brigitte ont deux durées de vie indépendantes, telles que :

$$P(\text{Arthur vit encore 9 ans}) = \frac{4}{5}, \quad P(\text{Brigitte vit encore 9 ans}) = \frac{3}{5}.$$

Calculer les probabilités que :

- (a) Arthur et Brigitte vivent encore 9 ans
- (b) L'un des 2 au moins vive encore 9 ans
- (c) Arthur seulement vive encore 9 ans
- (d) Arthur vive encore 9 ans sachant que l'un des 2 au moins vivra encore 9 ans.

FIGURE 5.59 – TD sur les probabilités conditionnelles – Énoncé de l'exercice 3

Suite du TD avec l'exercice 3 (voir figure 5.59) pour lequel la correction écrite par l'enseignant au tableau est à retrouver figure 5.60.

Les deux variations suivantes sont communes aux quatre types de tâches proposés dans l'exercice :

5.3. Enseignement des probabilités en première année d'Université : le cas des étudiants de biologie

On considère deux évènements :

A = Xavier vit encore 9 ans

B = Yvette vit encore 9 ans

On sait que $P(A) = \frac{4}{5}$ et $P(B) = \frac{3}{5}$ et on suppose que A et B sont indépendants.

(a) On veut déterminer $P(A \cap B)$, comme les évènements sont indépendants on a :

$$P(A \cap B) = P(A)P(B) = \frac{12}{25}$$

(b) On a $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$= \frac{4}{5} + \frac{3}{5} - \frac{12}{25} = \frac{23}{25}$$

(c) $P(A \cap \bar{B}) = P(A) \times P(\bar{B})$ puisque les évènements sont indépendants :

$$P(A \cap \bar{B}) = P(A) \times P(\bar{B}) = \frac{4}{5} \times \frac{2}{5} = \frac{8}{25}$$

(d) On cherche à déterminer $P(A|A \cup B)$, par définition on a :

$$P(A|A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)} = \frac{P(A)}{P(A \cup B)} = \frac{20}{23}$$

Car on a $A \subset A \cup B$ et donc $A \cap (A \cup B) = A$

FIGURE 5.60 – TD sur les probabilités conditionnelles – Exercice 3 – Correction de l'enseignant

- à partir d'un énoncé en langage naturel
- sans que les évènements aient été explicitement nommés dans le texte

La première et la troisième questions relèvent du type de tâches **T-CalcInter** : calculer la probabilité d'une intersection d'évènements étant donné que les évènements sont indépendants. La seconde question relève du type de tâche, calculer la probabilité d'une union d'évènements non-disjoints à partir d'un énoncé en langage naturel. Ce type de tâche, que l'on nommera **T-CalcUnion**, est une nouveauté car il n'apparaît pas dans les résultats du chapitre 4 (on avait relevé dans les analyses de manuel T-CalcSimple mais pour une union d'évènements disjoints). La quatrième question relève du type de tâches **T-CalcProbaCond** : calculer une probabilité conditionnelle.

Un étudiant demande : « Est-ce que ça veut dire que les deux évènements sont indépendants ? » l'enseignant répond : « Oui, le fait que Xavier vive est totalement indépendant du fait que Yvette vive ou pas ». L'enseignant prend en charge la reconnaissance de la méthode car il donne l'aide constructive suivante aux étudiants : « La caractérisation de l'indépendance c'est dans les deux sens, c'est un "si et seulement si", donc si vous savez que les évènements sont indépendants, c'est le cas ici, vous pouvez faire les calculs avec cette hypothèse. Et vice versa, si vous pouvez calculer la probabilité, vous pouvez vérifier s'ils sont indépendants ou non, c'est ce qu'on vient de faire dans l'exercice 1. »

L'enseignant prend en charge la première étape du travail de modélisation, identifier et nommer les événements en jeu car ils ne sont pas identifiés clairement dans l'énoncé. Il corrige au fur et à mesure au tableau.

Première question, il faut interpréter l'énoncé pour utiliser l'indépendance des deux événements puis interpréter la première question pour identifier l'évènement dont on cherche la probabilité. Pour **T-CalcInter** à la première question, l'enseignant prend en charge toutes ces étapes. Il explique : « *La première question c'est la probabilité que A ET B se réalisent, les deux vivent 9 ans. Alors comment on détermine cette probabilité? Ici on a une information cruciale, c'est qu'on a supposé les événements A et B indépendants. Comme ils sont indépendants, on a $P(A \cap B) = P(A)P(B)$.* »

Pour la seconde question, **T-CalcUnion**, il faut interpréter "l'un des deux" comme l'union de deux événements, l'enseignant prend en charge à nouveau l'interprétation de l'énoncé et explique : « *Quand on dit "ou", en général, en math, si on ne précise pas, ce n'est pas un "ou exclusif", c'est "au moins", donc les deux peuvent vivre. L'un des deux vit encore 9 ans c'est la probabilité que A ou B se réalise, l'un n'excluant pas l'autre, les deux peuvent se réaliser aussi. C'est ce qu'on appelle A union B.* »

L'enseignant demande aux étudiants comment calculer la probabilité de A ou B en termes de probabilités de A et B, il laisse ici une partie du traitement à la charge des étudiants. Un étudiant répond : « *On ajoute, on met $P(A) + P(B) - P(A \cap B)$* », l'enseignant complète le tableau et termine le calcul. La formule de Poincaré, assez peu utilisée au lycée (elle est demandée par l'enseignante dans les questions rapides mais pas utilisée en exercices) a été revue dans le chapitre 1 de ce module.

Pour répondre à la troisième question, **T-CalcInter**, il faut interpréter "seulement" comme la réalisation de A et \bar{B} , c'est-à-dire l'intersection de ces deux événements. L'enseignant (F) laisse la reconnaissance du type de tâches aux étudiants et leur demande d'interpréter l'énoncé afin d'identifier l'évènement recherché : « *Quelle est la probabilité que vous cherchez à déterminer ?* », un échange avec trois étudiants (E1, E2 et E3) s'ensuit :

E1 : « *A union \bar{B}* »

F : « *Est-ce que vous êtes sûr que c'est cet événement-là ?* »

E1 : « *Heu non... intersection ?* »

F : « *Ce que l'on veut déterminer c'est $A \cap \bar{B}$, c'est-à-dire la probabilité que A se réalise : Xavier vit encore 9 ans, mais que B ne se réalise pas : Yvette ne vive pas. Comment est-ce que vous calculez cette probabilité là ? Vous avez un moyen de la calculer puisque les événements sont indépendants* »

E2 : « *$P(A)$ fois $P(\bar{B})$* »

F : « *Pourquoi ?* »

E3 : « *Parce que les événements sont indépendants* »

F : « *On sait que A et B sont indépendants, donc les événements A et \bar{B} sont aussi indépendants : que Yvette vive encore 9 ans n'a pas d'influence sur le fait que Xavier vive encore 9 ans et ça implique que si elle meurt ça n'a pas d'influence non plus* »

L'enseignant, en évoquant l'indépendance des deux événements ci-dessus, donne une aide procédurale aux étudiants quant au traitement de la tâche (calculer la probabilité d'une intersection lorsque les deux événements sont indépendants).

L'enseignant procède de la même façon pour la dernière question, **T- CalcProbaCond**. Il laisse aux étudiants la reconnaissance du type de tâches et leur demande de traduire l'énoncé en langage naturel : « *Comment ça se traduit en termes de probabilités d'événements? Un point crucial dans ce type d'énoncé c'est de réussir à traduire en termes des événements considérés et de formules, les événements qu'on cherche à déterminer.* » Un étudiant répond : « *$P(A \text{ sachant } A \text{ union } B)$* » l'enseignant prend en charge le traitement de la suite de cette question.

L'enseignant conclut à propos de cet exercice : « *Vous avez ici une illustration de calculs de probabilités lorsque vous avez comme information supplémentaire que certains événements sont indépendants.* »

Détails à propos de l'épisode n°13 :

Exercice 5 Le gérant d'un magasin d'informatique a reçu un lot de boîtes de CD-ROM. 5% des boîtes sont abîmées. Le gérant estime que 60% des boîtes abîmées contiennent au moins un CD-ROM défectueux et que 98% des boîtes non abîmées ne contiennent aucun CD-ROM défectueux. Un client achète une boîte du lot. On désigne par A l'événement : «la boîte est abîmée» et par D l'événement «la boîte achetée contient au moins un disque défectueux».

- Calculer les probabilités $P(A)$, $P(\bar{A})$, $P(D|A)$, $P(D|\bar{A})$, $P(\bar{D}|A)$ et $P(\bar{D}|\bar{A})$. On commencera par caractériser celles qui sont explicitement dans l'énoncé.
- Le client constate qu'un des CD-ROM acheté est défectueux. Quelle est la probabilité pour qu'il ait acheté une boîte abîmée?

FIGURE 5.61 – TD sur les probabilités conditionnelles – Énoncé de l'exercice 5

Le TD se poursuit avec l'exercice 5 (voir figure 5.61) dont la correction écrite par l'enseignant au tableau est à retrouver figure 5.62. L'enseignant laisse les étudiants chercher l'exercice et passe dans les rangs s'arrêtant ponctuellement ou quand un étudiant a une question.

On considère deux évènements :

A = la boîte est abimée

B = la boîte contient un CD défectueux

On dispose des informations suivantes :

$$P(A) = \frac{5}{100} = \frac{1}{20} ; P(\overline{D}|\overline{A}) = \frac{98}{100} ; P(D|A) = \frac{60}{100} = \frac{3}{5}$$

(a) $P(\overline{A}) = 1 - P(A) = \frac{95}{100}$

$$P(D|\overline{A}) = 1 - P(\overline{D}|\overline{A}) = \frac{2}{100} \quad \text{et} \quad P(\overline{D}|A) = 1 - P(D|A) = \frac{40}{100}$$

(b) On cherche à déterminer $P(A|D) = \frac{P(A \cap D)}{P(D)}$ et $P(D|A) = \frac{P(A \cap D)}{P(A)}$ donc :

$$P(A|D) = \frac{P(D|A)P(A)}{P(D)} \quad \text{Formule de Bayes}$$

$$P(D) = P(D|A)P(A) + P(D|\overline{A})P(\overline{A})$$

Finalement $P(A|D) = \frac{3}{22}$

FIGURE 5.62 – TD sur les probabilités conditionnelles – Exercice 5 – Correction de l’enseignant

A la suite d’une question d’un étudiant il écrit au tableau :

$$A^c = \overline{A}$$

Il s’agit ici d’un changement de notation qui semble poser problème aux étudiants qui sortent du lycée (d’après des entretiens menés avec des étudiants de biologie voir partie 6.3 page 283).

La première question relève du type de tâches **T-Assoc** : associer les valeurs de l’énoncé à des probabilités d’évènements et **T-CalcBar** : calculer, à partir de la probabilité d’un évènement, la valeur de la probabilité de son évènement contraire. Les variations présentes sont les suivantes :

- à partir d’un énoncé en langage naturel
- étant donné que c’est la première question de l’exercice
- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- les données sont en pourcentages

L’enseignant prend en charge la reconnaissance de la tâche, l’organisation et le traitement : « On nous donne un certain nombre d’informations et il faut traduire ces informations en termes des évènements A et B. On nous dit 5% des boîtes sont abimées, donc la probabilité d’avoir une boîte abimée est de 5/100. On nous dit ensuite que 60% des boîtes abimées contiennent un CD défectueux, ça se traduit comment? C’est la probabilité que le CD soit défectueux sachant que la boîte est abimée, donc c’est D sachant A. Et 98% des boîtes non abimées ne contiennent aucun

CD défectueux, ça veut dire que la probabilité de ne pas avoir un CD défectueux sachant que la boîte n'est pas abîmée est de 98%. Donc D barre sachant A barre. »

A propos de la première question il indique : « *La probabilité de D sachant \bar{A} , ça fait appel aux propriétés de base des probabilités conditionnelles, ça résume la manière dont on peut calculer des probabilités conditionnelles. La vous regardez l'évènement D sachant que \bar{A} est réalisé, et vous connaissez la probabilité que \bar{D} sachant \bar{A} , ça c'est un truc qu'il ne faut pas oublier, c'est 1 moins ... Une probabilité conditionnelle c'est une probabilité, donc si vous fixez \bar{A} c'est les mêmes règles opératoires que les probabilités. Donc la probabilité d'un évènement c'est 1 moins la probabilité de la négation de cet évènement.* » et il corrige cette question au tableau.

La seconde question relève du type de tâches **T-CalcProbaCond**, l'enseignant laisse la reconnaissance de la tâche aux étudiants : « *Quel est l'évènement dont on cherche à connaître la probabilité ici ? Il faut toujours que vous ayez à l'esprit, quelles sont les probabilités dont vous disposez et quelle est la probabilité que vous cherchez. On cherche quelle probabilité ? Là il faut passer du français aux maths »*

Une étudiante répond : « *$P(A \text{ sachant } \bar{D})$ »*

L'enseignant complète et prend en charge l'organisation et le traitement de la tâche : « *On dispose d'un certain nombre de probabilités, alors comment on fait le calcul ? Le point crucial, vous avez une formule très utile ... Si on se souvient plus on reprend la définition de $P(D|A)$ » il écrit :*

On cherche à déterminer $P(A|D) = \frac{P(A \cap D)}{P(D)}$ et $P(D|A) = \frac{P(A \cap D)}{P(A)}$ donc :

$$P(A|D) = \frac{P(D|A)P(A)}{P(D)} \text{ Formule de Bayes}$$

L'enseignant explique pourquoi la formule de Bayes est importante : « *Cette relation est très utile parce que si vous ne connaissez pas la probabilité que l'évènement A se réalise sachant que D se produit mais que vous connaissez la probabilité que D se produit sachant que A se réalise, ça vous permet quand même de calculer cette probabilité là. Ici dans cette formule, on connaît $P(D|A)$ et $P(A)$. Finalement la probabilité que je cherche c'est $P(D)$. La probabilité totale qu'un évènement se produise indépendamment de tout autre, elle peut s'écrire aussi en termes de probabilités conditionnelles. Comment on calcule $P(D)$? Dans votre cours vous avez le résultat suivant » Il écrit la formule des probabilités totales et termine l'exercice.*

L'enseignant conclut quant à cet exercice : « *Cet exercice est important car la première question c'est la manipulation des définitions de base et la seconde est caractéristique du calcul de probabilités conditionnelles avec les deux formules que l'on utilise sans cesse. »*

Détails à propos de l'épisode n° 14 :

Exercice 7. Le quart d'une population est vacciné contre le choléra. Au cours d'une épidémie, on constate qu'il y a parmi les malades un vacciné pour 4 non-vaccinés. En plus, parmi les vaccinés il y a une personne sur 12 qui tombe malade. Quelle est la probabilité qu'un non-vacciné tombe malade ?

FIGURE 5.63 – TD sur les probabilités conditionnelles – Énoncé de l'exercice 7

On considère deux évènements :

M = le patient est malade

V = le patient est vacciné

On dispose des informations suivantes :

$$P(V) = \frac{1}{4} \quad P(M|V) = \frac{1}{12} \quad P(V|M) = \frac{1}{5}$$

On cherche à déterminer $P(M|\bar{V}) = \frac{P(\bar{V}|M)P(M)}{P(V)}$

On connaît $P(V)$ et $P(\bar{V}|M) = 1 - P(V|M) = \frac{4}{5}$ Or :

$$P(M) = P(M|V)P(V) + P(M|\bar{V})P(\bar{V})$$

En remplaçant les valeurs numériques et en posant $p = P(M|\bar{V})$ on obtient :

$$p = \frac{\frac{4}{5}(\frac{1}{12} \times \frac{1}{4} + p\frac{3}{4})}{\frac{3}{4}}$$

Cela donne $p = \frac{1}{45} + \frac{4}{5}p$ donc $p = \frac{1}{9}$

FIGURE 5.64 – TD sur les probabilités conditionnelles – Exercice 7 – Correction de l'enseignant

L'exercice n°7 (voir énoncé figure 5.63 et la correction écrite par l'enseignant au tableau figure 5.64) relève du type de tâches **T-CalcProbaCond** avec une forte modélisation laissée à la charge de l'élève (identifier les évènements, les nommer, déterminer leurs probabilités, interpréter la question et formuler la tâche).

L'enseignant commence par expliquer : « C'est un exercice typique car c'est l'utilisation du calcul de probabilités pour avoir des statistiques en épistémologie. On cherche à savoir, si vous êtes vaccinés, quelles sont les chances pour que vous tombiez malade face à cette épidémie de choléra. En fait ça peut être utile parce que si vous avez des données statistiques, notamment dans des modèles bio-médico, dans des enquêtes épidémiologiques, savoir calculer certaines probabilités d'une certaine manière plutôt qu'une autre ça peut être intéressant parce que le coût d'implémentation de votre étude n'est pas forcément le même. »

L'enseignant commence la correction et demande aux étudiants quels sont les évènements que l'on peut considérer d'après l'énoncé. Un étudiant répond : « Pro-

babilité d'être vacciné c'est $\frac{1}{4}$ » L'enseignant attendait ici une réponse de modélisation car les événements ne sont ni nommés ni décrits explicitement, il lui répond : « *Quels sont les événements qu'il faut considérer ? Avant de mettre les nombres il faut les événements. Typiquement on a quoi dans cet énoncé ? On a ceux qui sont vaccinés et ceux qui sont malades, donc on a deux événements* » Il prend en charge l'identification des événements, de les nommer (organisation), il écrit au tableau :

On considère deux événements

M = le patient est malade

V = le patient est vacciné

On dispose des informations suivantes : $P(V) = \frac{1}{4}$

Un autre étudiant répond : « *La probabilité d'être malade sachant qu'on est vacciné c'est $1/12$* » et l'enseignant complète : « *Et sachant qu'on est malade la probabilité d'être vacciné c'est $1/5$* » et il écrit :

$$P(M|V) = \frac{1}{12} \quad \text{et} \quad P(V|M) = \frac{1}{5}$$

L'enseignant donne directement la valeur mais il y a une adaptation ici car $P(V|M)$ n'est pas évidente à donner d'après l'énoncé (contrairement à $P(M|V)$). Il reste à interpréter la question et à identifier la probabilité recherchée, un étudiant propose : « *Qu'on soit vacciné sachant qu'on est malade* » puis se reprend : « *Qu'on soit non vacciné sachant qu'on est malade* », une étudiante le corrige : « *Non c'est l'inverse !* » et l'enseignant ajoute : « *C'est le problème, il faut réussir à traduire le français dans les maths ce qui n'est pas toujours évident.* »

Une étudiante propose : « $P(M|\bar{V})$ »

L'enseignant conclut : « *Quelqu'un qui n'est pas vacciné quelles sont ses chances d'être malade ? Comme ça vous pourrez comparer à la probabilité d'être malade sachant qu'on est vacciné, et savoir quel est l'efficacité du vaccin dans la population. Typiquement vous avez un certain nombre de données épidémiologiques, vous savez parmi les gens vaccinés quelle est la proportion de ceux qui sont malades et vous voulez connaître la probabilité d'être malade quand on n'est pas vacciné, de façon à comparer l'efficacité du vaccin par rapport au taux de prévalence dans la population* » et écrit au tableau :

On cherche à déterminer $P(M|\bar{V})$

$$\text{On a } P(M|\bar{V}) = \frac{P(\bar{V}|M)P(M)}{P(\bar{V})}$$

On connaît $P(\bar{V}|M)$ et $P(V)$

L'enseignant entoure dans la formule les probabilités qui sont connues d'après l'énoncé et celles qui ne le sont pas, il écrit les valeurs des probabilités qui sont connues.

« Ici on cherche $P(M)$, comment on fait? Il faut aussi se souvenir qu'on peut utiliser la formule des probabilités totales. Est-ce que vous pouvez m'exprimer la probabilité de M en fonction de l'évènement V ? » et il écrit :

$$\text{Or } P(M) = P(M|V)P(V) + P(M|\bar{V})P(\bar{V})$$

L'enseignant entoure d'une certaine couleur les probabilités, dans l'expression ci-dessus, qui sont connues d'après l'énoncé et termine la résolution de l'exercice en posant $p = P(M|\bar{V})$ la probabilité recherchée.

Synthèse

Nous présentons dans cette partie la synthèse des résultats et des analyses détaillées ci-dessus, dans le cas des probabilités conditionnelles à l'Université. Il s'agit d'une synthèse en termes de comparaison et de transition car les analyses sont faites en regard de ce que nous avons observé au lycée. La comparaison est donc intégrée dans ce que nous décrivons dans la suite de cette partie.

La définition d'une probabilité conditionnelle est la même que celle vue au lycée mais l'enseignant présente également cette notion par le prisme de la nouvelle définition d'une probabilité (application $A \mapsto P(A|B)$) qui a été vue dans le premier chapitre. On peut imaginer que c'est assez abstrait pour les étudiants.

Notations :

Une nouvelle notation apparaît au tout début du cours. Il s'agit de la notation de la probabilité conditionnelle notée $P(A|B)$ au lieu de $P_B(A)$. Il est intéressant de remarquer que l'enseignant motive ce changement du fait qu'il s'agisse d'étudiants de biologie, l'enseignant s'adapte au public et argumente en faveur de cette notation car elle représente, selon lui, un moyen mnémotechnique de se souvenir de la définition d'une probabilité conditionnelle.

Une autre notation présente sur les diapositives du cours et nouvelle pour les étudiants qui sortent du lycée, est A^c . La notation A^c n'est pas celle des programmes de lycée. L'enseignant, dans son discours, dit parfois " A barre " pour " A^c ". Lors de la séance de TD un étudiant demande ce que signifie la notation A^c présente dans l'énoncé d'un exercice.

Nouveautés :

La plupart des formules et définitions de ce cours ont déjà été vues au lycée (comme par exemple la définition d'une probabilité conditionnelle et la formule des probabilités totales) et l'enseignant le souligne à plusieurs reprises. Il fait également référence à la formule des probabilités en cascade, nous relevons ici cette remarque car cette formule n'est pas au programme de lycée, contrairement à

celle des probabilités totales. Les étudiants ne l'ont sans doute jamais rencontrée.

Les nouveautés en termes de contenus et de types de tâches, pour ce cours, par rapport à ce que nous avons observé en terminale scientifique sont les suivantes :

- les notations $P(A|B)$ et A^c
- la notion de partition de l'univers
- la formule des probabilités totales dans le cas $n > 2$
- la formule de Bayes
- **T-CalcUnion** : calculer la probabilité d'une union d'évènements disjoints
- positivement corrélé et négativement corrélé
- indépendance de n évènements

Concernant la formule des probabilités totales dans le cas de deux évènements, qui est principalement celle que les étudiants ont utilisée au lycée, l'enseignant indique que ce sera surtout celle qu'ils utiliseront cette année.

L'enseignant considère que les élèves ont déjà vu la formule de Bayes au lycée, mais celle-ci n'est pas au programme de lycée bien qu'elle puisse apparaître parfois dans des manuels ou au gré des calculs (activité d'approfondissement manuel Indice).

La définition d'indépendance ($P(A)P(B) = P(A \cap B)$) et le critère d'indépendance ($P(A|B) = P(A)$) présentés dans le cours ont déjà été vus au lycée, la nouveauté ici ce sont les notions de positivement corrélé ou négativement corrélé. La définition donnée en cours n'est pas la même que celle rappelée par l'enseignant de TD (intervertion critère et définition), ce qui peut créer une confusion chez les étudiants.

Exemples et modélisation :

Dès le début du cours l'enseignant motive l'intérêt des probabilités conditionnelles par des exemples issus de situations de la vie réelle (relation entre diabète et consommation de soda, population de poissons dans un lac etc).

Dans les exemples et exercices, l'enseignant accorde beaucoup d'importance à l'étape de modélisation de l'expérience aléatoire. Il explique comment il souhaite que les étudiants procèdent dans une situation similaire. Parfois les situations présentées auraient pu être modélisées à l'aide d'un arbre de probabilités, comme au lycée, mais ce n'est pas le cas ici à l'Université. Ces situations sont présentées avec deux nouveaux concepts (non utilisés dans le secondaire) : Ω l'univers et P la mesure de probabilité. C'est tout à fait nouveau pour les étudiants, un évènement est désormais présenté comme un sous-ensemble de l'ensemble Ω qui doit être décrit préalablement au début de chaque exercice.

L'enseignant insiste sur l'importance du choix du modèle, il le justifie en expliquant que premièrement c'est pour ne pas faire des calculs dans le vide, deuxièmement pour savoir à quoi se réfèrent les calculs d'un point de vue mathématique et enfin pour être capable d'analyser *a posteriori* si cela correspond à l'expérience de biologie que l'on souhaite modéliser.

Types de tâches, variations et techniques :

Les exercices du TD demandent tous un certain travail de modélisation, identifier les événements ou au moins associer aux événements décrits dans l'énoncé leurs probabilités, interpréter la question posée pour identifier la probabilité recherchée, interpréter le résultat final quand cela est demandé. De façon générale, l'enseignant de TD prend en charge la première étape du travail de modélisation, identifier et nommer les événements en jeu lorsqu'ils ne sont pas identifiés clairement dans l'énoncé. Il prend en charge les sous-activités de traitement et d'organisation quasiment à chaque fois.

Voici dans le détail les types de tâches, variations et techniques rencontrés dans ce cours. Nous comparons systématiquement si les tâches proposées aux étudiants sont des nouveautés ou si les techniques sont différentes de celles vues au lycée :

- **T-CalcProbaCond** : *calculer une probabilité conditionnelle.*
 - Il s'agit d'un type de tâches rencontré dans plusieurs exemples du cours, généralement le calcul n'est pas détaillé car seulement la modélisation de l'énoncé est présentée par l'enseignant. Par exemple diapositive n°9, l'enseignant passe du temps sur la modélisation de la situation (création du modèle) mais ne détaille pas comment déterminer les probabilités conditionnelles qui découlent de l'interprétation de l'énoncé. Il motive, dans le contexte de la situation de l'exemple, l'intérêt de s'intéresser à cette probabilité en particulier, car elle représente réellement quelque chose qu'il est important de savoir. Il détaille la méthode à suivre : traduire les données de l'énoncé et utiliser la formule de Bayes (nouveauté par rapport au lycée).
 - Ce type de tâches est à nouveau rencontré dans les exercices du TD à partir d'un énoncé en langage naturel, sans que les événements ne soient explicitement nommés dans le texte (ex3 et ex7). Dans un autre exercice il faut à la fois utiliser la formule de Bayes puis celle des probabilités totales pour calculer le dénominateur (ex5) ou pour développer le numérateur (ex7) (nouveauté par rapport au lycée).
- **T-CalcSimple** : *calculer une probabilité simple.*

Il s'agit d'un type de tâches rencontré dans un exemple du cours. La formule des probabilités totales n'est pas détaillée. Les variations *l'énoncé est en langage naturel* et *les événements ne sont pas donnés dans l'énoncé* sont prises en charge par l'enseignant qui insiste surtout sur l'interprétation de l'énoncé mais ne détaille pas les calculs.
- **T-EtudInd** : *étudier l'indépendance de deux événements.*

- Il s'agit d'un type de tâches rencontré plusieurs fois dans des exemples du cours. Concernant la technique, il est attendu de faire un calcul pour montrer une indépendance d'évènements. L'enseignant ne veut pas de « *discours métaphysique* » même s'il incite les étudiants à avoir toujours l'intuition de l'indépendance ou non des événements lorsqu'ils rencontrent une situation probabiliste. Il détaille la technique de cette façon : « *Il y a trois nombres à calculer et il faut juste vérifier si l'un est le produit des deux autres.* »
- Ce type de tâches est également rencontré en TD (ex1). L'enseignant prend en charge les sous-activités d'organisation et de traitement de la tâche : calculer la probabilité de A et celle de B, calculer la probabilité que A et B se réalisent, vérifier si c'est le produit ou non. L'enseignant prend en charge l'identification des événements, puis, étant donné que les données sont en fréquences naturelles, une étape supplémentaire de la modélisation qui est de calculer la probabilité d'apparition des événements identifiés : **T-CalcSimpleFN** (nouveau par rapport au lycée). Cet enchevêtrement de techniques peut être une source de difficultés pour les étudiants.
- **T-CalcInter** : *calculer la probabilité d'une intersection d'évènements.*
Il s'agit d'un type de tâches rencontré en TD (ex3), avec les variations suivantes : étant donné que les événements sont indépendants, à partir d'un énoncé en langage naturel, sans que les événements ne soient explicitement nommés dans le texte. Ce type de tâches a déjà été vu au lycée.
- **T-CalcUnion** : *calculer la probabilité d'une union d'évènements non disjoints.*
Il s'agit d'un nouveau type de tâches rencontré en TD (ex3), avec les variations suivantes : à partir d'un énoncé en langage naturel (nouveau), sans que les événements ne soient explicitement nommés dans le texte. Il faut utiliser la formule de Poincaré qui a été rencontrée en cours au lycée mais qui n'a pas été mise en fonctionnement directement dans des exercices (d'après nos observations).
- **T-Assoc + T-CalcBar** : *associer les valeurs de l'énoncé à des probabilités d'évènements et calculer; à partir de la probabilité d'un événement, la probabilité de son événement contraire.* Il s'agit d'un type de tâches rencontré en TD dans l'exercice 5, l'enseignant prend en charge l'interprétation de l'énoncé, et traduit chacune des informations disponibles en termes de probabilités.

Hormis **T-CalcUnion** l'ensemble de ces types de tâches a déjà été rencontrés au lycée, d'après l'analyse des manuels (chapitre 4 page 49) et les observations menées en terminale scientifique (partie 5.2).

5.3.2 Analyse du savoir enseigné pour le thème des lois continues

Dans cette partie nous présentons l'analyse du savoir enseigné pour le thème des lois continues en première année de biologie à l'Université.

Le synopsis du cours magistral auquel nous avons assisté se trouve figure 5.65. L'ensemble des diapositives de ce cours est à retrouver en annexe E.1. Nous avons également assisté à une séance de TD dont le synopsis se trouve figure 5.66. Cette séance a eu lieu trois semaines après le cours magistral, l'enseignant qui est en charge de ce groupe a également fait le CM. Le groupe observé durant la séance de TD est constitué d'une quinzaine d'étudiants. L'enseignant procède de la façon suivante, il laisse les étudiants chercher en autonomie, circule dans les rangs et intervient auprès des étudiants qui le sollicitent. Il écrit au fur et à mesure une correction détaillée des exercices au tableau. Le TD associé à ce cours correspond à la feuille d'exercices suivante (figure 5.67).

n°	Thème principal de l'épisode	Commentaires
1	4.1 Densité : première définition, propriétés de la fonction de densité (diapo 1)	Détails dans ce qui suit
2	Premiers exemples de densité (graphiquement) (diapos 2 et 3)	Détails dans ce qui suit
3	Exemples de densités et premiers calculs de probabilités T-CalcSimple et T-VerifF (diapos 4 et 5)	Détails dans ce qui suit
4	Mode et médiane (diapo 6 et 7)	X
5	Espérance et variance (diapos 8, 9 et 10)	Détails dans ce qui suit
6	Loi uniforme (diapos 11 et 12)	Détails dans ce qui suit
7	Loi exponentielle (diapos 13 et 14)	Détails dans ce qui suit
8	Loi normale (diapos 15, 16, 17 et 18)	Détails dans ce qui suit
9	Fonction de répartition (diapos 19 et 20)	X

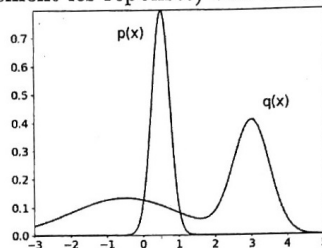
FIGURE 5.65 – Synopsis de la séance de cours magistral – *Lois continues*

n°	Thème principal de l'épisode	Commentaires
10	Exercice 4 : T-EtudInd pour deux configurations plutôt similaires d'une même situation (tirages parmi 12 boules puis parmi 13 boules). Ressemblance avec les exemples diapo 14 et 15 du cours.	Détails dans ce qui suit
11	Exercice 3 : T-CalcInter , T-CalcSimple (union de deux événements) et T-CalcProbaCond . Exercice proposé dans le test 2018 distribué à des lycéens et à des étudiants de biologie.	Détails dans ce qui suit
12	Exercice 2 : T-Assoc et T-CalcProbaCond	Détails dans ce qui suit

FIGURE 5.66 – Synopsis de la séance de travaux dirigés – *Lois continues*

Feuille d'exercices 4

Exercice 1 (On justifiera brièvement les réponses) Soit les deux densités suivantes.



1. Les deux ensembles de données suivants : $A = \{0, 0.3, 0.7, 1\}$ et $B = \{-1, 2, 2.5, 3\}$ représentent chacun des v.a. tirées selon p ou q , mais on ne sait plus lequel a été tiré selon quelle loi. Doit-on attribuer (selon toute vraisemblance) A à p et B à q ou l'inverse ?
2. L'une des deux lois est gaussienne. Laquelle ?
3. Laquelle a la plus grande variance ? Laquelle a la plus grande moyenne (espérance) ?

Exercice 2 Pour cet exercice, on utilisera la table au dos. Soit $X \sim \mathcal{N}(0, 1)$, $Y \sim \mathcal{N}(5, 9)$.

- (a) Trouver $P(-1 \leq X \leq 1)$, $P(-2 \leq X \leq 2)$, $P(0 \leq X \leq 1)$.
- (b) Trouver $P(2 \leq Y \leq 8)$, $P(Y \geq 2)$, $P(Y \leq 8)$.

Exercice 3 La durée de vie en années d'une machine à laver suit une distribution $\mathcal{E}(\frac{1}{10})$.

- (a) Quelle est son espérance de vie ?
- (b) Quelle est la probabilité qu'elle tienne au moins une année ?
- (c) Quelle est la probabilité qu'elle tienne au moins 10 ans ?

Exercice 4 Soit X une variable aléatoire de densité $f(x) = cx^2 1_{0 \leq x \leq 1} + \frac{1}{2} 1_{2 \leq x \leq 3}$.

- (a) Que vaut c ?
- (b) Calculer $P(\frac{1}{4} \leq X \leq \frac{3}{4})$?
- (c) Calculer l'espérance et la variance de X .

Exercice 5. Soit $a < 0 < b$, et X une variable aléatoire de loi $\mathcal{U}([a, b])$, quelle est la loi du signe de X ? Soit $t > 0$, quelle est la loi de $1_{X \leq t}$? *Indication.* Considérer les cas $t < a$, $a \leq t \leq b$, et $t > b$.

Exercice 6.

- (a) On tire 1000 nombres aléatoires qui sont distribués selon une loi exponentielle $\mathcal{E}(2)$. Combien des nombres tirés, à peu près, sont plus petits que 1 ?
- (b) On tire 1000 nombres aléatoires qui sont distribués selon une loi normale $\mathcal{N}(2, 9)$. Combien des nombres tirés, à peu près, sont entre 2 et 3 ? (Il faudra utiliser la table des valeurs de Φ .)

Exercice 7. Une pièce de monnaie porte sur une face l'inscription 17 et sur l'autre face 20. On modélise l'expérience par une variable aléatoire X de loi $P(X = 17) = 0,5$ et $P(X = 20) = 0,5$.

- (a) Résoudre le système $a \times 0 + b = 17$, $a \times 1 + b = 20$. (Élémentaire)
- (b) Quelle est la loi de la v.a. Y telle que $aY + b = X$?
- (c) En déduire l'espérance et la variance de X .

Exercice 8. À un péage autoroutier, n voitures empruntent au hasard et indépendamment les unes des autres, l'un des trois passages mis à leur disposition. Soient X , Y et Z les v.a. dénombrant les voitures ayant franchi ces passages.

- (a) Déterminer la loi de X , celle de $X + Y$.
 (b) Calculer les variances de X , Y et de $X + Y$. En déduire la covariance de X et Y .
 Trouvez-vous l'énoncé réaliste ?

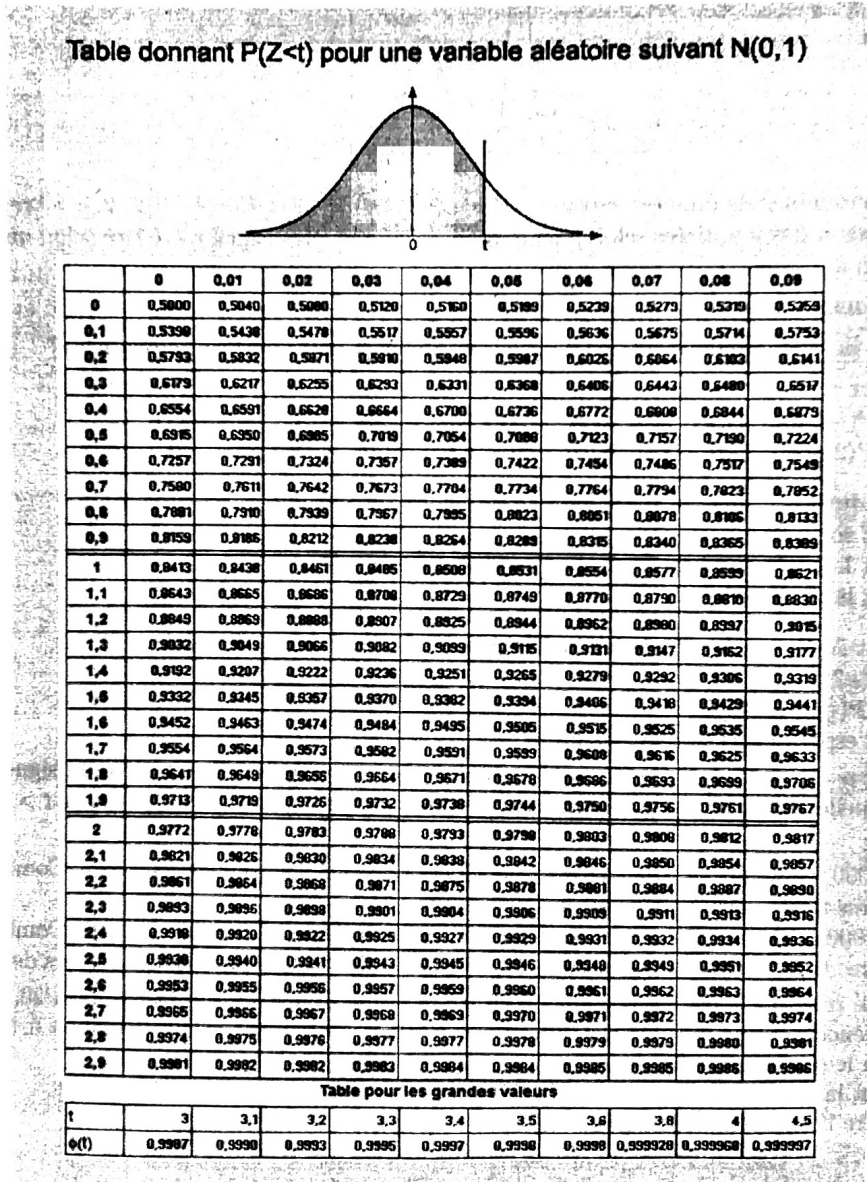


FIGURE 5.67 – Feuille d'exercices – TD sur les lois continues

Nous présentons dans ce qui suit l'analyse des épisodes du CM et ceux du TD qui correspondent aux lignes grisées dans les synopsis. En regard de la partie précédente sur l'enseignement dans le secondaire, nous proposons, dans une dernière sous-partie, une synthèse en termes de comparaison et de transition pour le thème des lois continues à l'Université.

Détails des épisodes

Détails à propos de l'épisode n° 1 :

4.1. DENSITÉ

Définition
Une variable aléatoire (v.a.) $X : \Omega \rightarrow \mathbb{R}$ admet une densité s'il existe une fonction f telle que pour tout intervalle $[a, b]$:

$$P(X \in [a, b]) = \int_a^b f(x)dx.$$

On dit que X est une v.a. continue.

Remarques. Nécessairement,

$$f \geq 0$$
$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

Noter que sous ces deux conditions, on a bien que l'application $A \mapsto \int_A f(x)dx$ est une mesure de probabilité. On dit qu'une fonction qui satisfait ces deux conditions est une densité de probabilité.

FIGURE 5.68 – Diapositive n°1 – Cours sur les lois continues

Après avoir présenté la première définition de ce cours l'enseignant explique aux étudiants les liens entre probabilité, densité et intégrale : « *La probabilité que votre variable aléatoire appartienne à un intervalle $[a; b]$ va être donnée par l'intégrale de cette fonction. Vous pouvez penser f comme étant une densité de masse sur une droite, il y a des parties qui pèsent plus lourd et il y a des parties qui pèsent moins lourd. Et la probabilité que X appartienne à un segment $[a; b]$, donc à une partie de cette droite, c'est en fait le poids de cette partie. Vous avez une densité, le poids d'une partie est donné par une intégrale.* »

A propos des deux caractéristiques d'une densité de probabilités rappelées sur la diapositive (voir figure 5.68) (il est important de relever que, contrairement au lycée, on ne s'intéresse plus à savoir si la fonction f est continue ou définie) l'enseignant ajoute : « *X c'est toujours un nombre, un nombre aléatoire, quel que soit l'élément ω de Ω , $X(\omega)$ c'est un nombre. Donc la probabilité que ce nombre soit plus grand que $-\infty$ et plus petit que $+\infty$, c'est 1, toujours.* »

L'enseignant donne des indications sur comment traiter des exercices qui demanderont de manipuler $\int_{-\infty}^{+\infty} f(x)dx = 1$: « *En général quand vous allez étendre cette chose par linéarité, par la relation de Chasles, vous pouvez définir la probabilité d'appartenir à une réunion avec deux intégrales par linéarité. Vous voyez, ça vous redonne la même chose quand les intervalles se chevauchent.* »

Détails à propos de l'épisode n°2 :

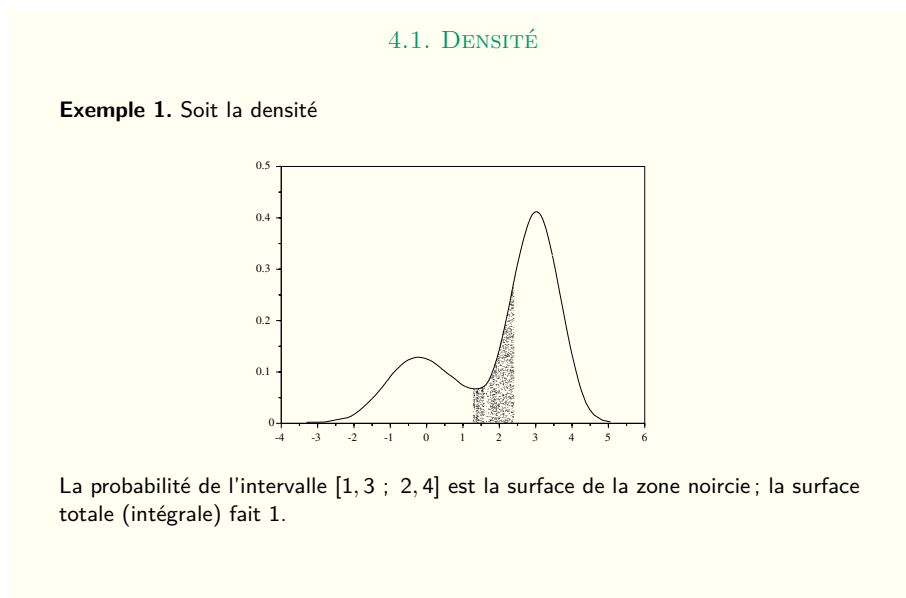


FIGURE 5.69 – Diapositive n°2 – Cours sur les lois continues

Le cours se poursuit avec des exemples de fonctions de densité.

A propos de l'exemple 1 (diapositive n°2 voir figure 5.69) l'enseignant explique comment calculer l'aire hachurée sous la courbe : « Pour calculer la probabilité que votre variable aléatoire soit contenue entre 1, 3 et 2, 4 vous n'avez qu'à mesurer l'aire qui est grisée. Pour calculer cette aire, vous avez plusieurs outils à votre disposition. D'abord vous savez calculer des aires de rectangles, et quand la fonction est courbée comme ça, et que vous n'avez pas de formule pour cette fonction f vous faites appel à votre logiciel de calcul préféré. Il va faire le calcul numérique avec des petits rectangles et il va calculer ce truc là. Dans les exercices que l'on va vous donner vous pouvez calculer l'intégrale en calculant une primitive. »

L'enseignant présente un second exemple (diapositive n°3 voir figure 5.70) qu'il décrit aux étudiants : « On a des exemples qui sont issus des statistiques. Ici vous pouvez voir deux variables aléatoires, une qui vous donne l'âge d'un marié le jour de son mariage et une autre qui donne l'âge de la mariée le jour du mariage. Ce sont de vraies données, récupérées auprès des bureaux de statistiques de l'INSEE. D'après les densités de ces deux courbes, la plupart des gens semblent se marier autour de 20 ans, on voit plein de choses que l'on ne verrait pas si je vous avais donné la liste des nombres dans un tableau Excel à 150000 lignes avec 150000 nombres. Il faut arriver à représenter ces données soit graphiquement comme ça, soit, et c'est ce qu'on va voir après, avec des nombres caractéristiques : l'espérance et la variance. »

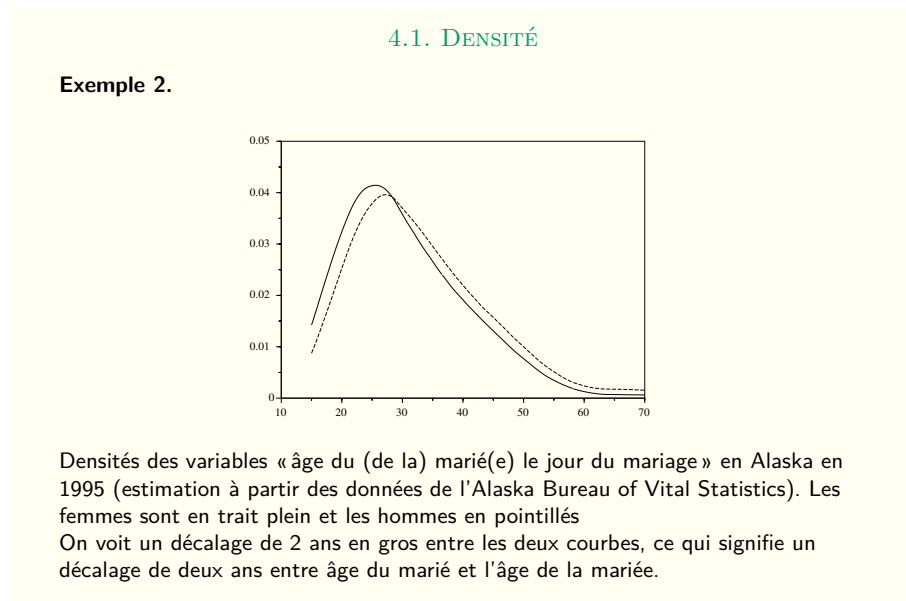


FIGURE 5.70 – Diapositive n°3 – Cours sur les lois continues

Détails à propos de l'épisode n°3

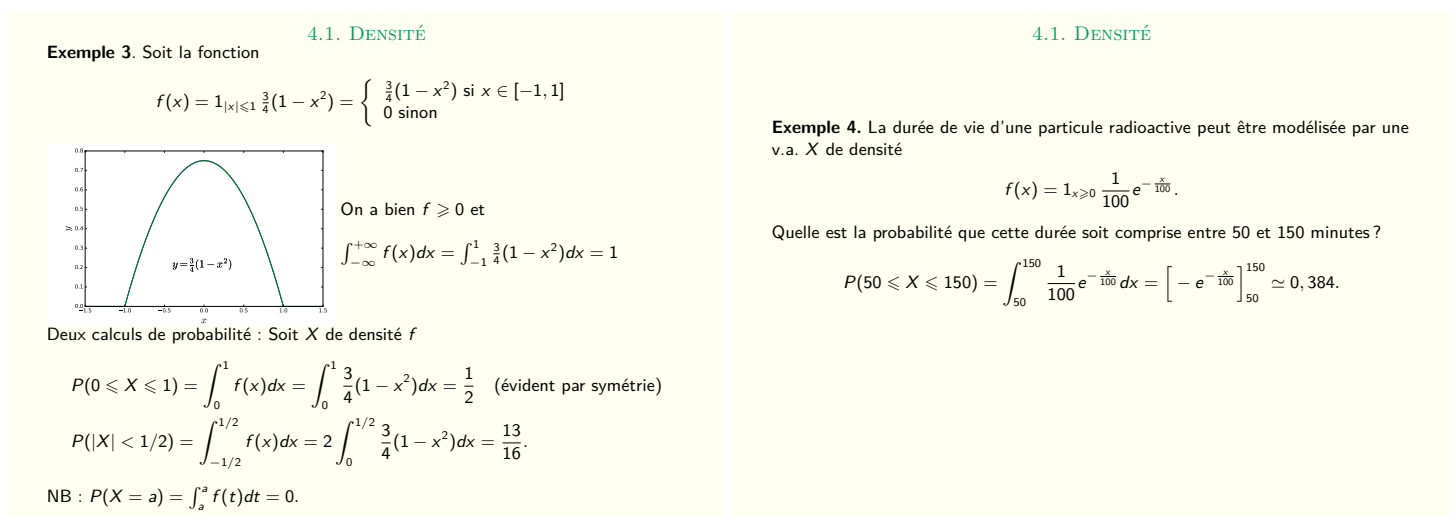


FIGURE 5.71 – Diapositives n°4 et n°5 – Cours sur les lois continues

Deux nouveaux exemples sont présentés aux étudiants diapositives 4 et 5 (voir figure 5.71), la fonction de densité f est explicitée. Il s'agit de la première rencontre (dans le cadre de ce module) avec les types de tâches **T-CalcSimple** et **T-Veriff** (vérifier qu'une fonction donnée est bien une densité de probabilité), les variations présentes ici sont les suivantes :

- étant donné que la variable aléatoire est identifiée dans l'énoncé et que sa densité a été donnée dans l'énoncé
- étant donné que l'énoncé est en langage naturel (ex4)

- à partir de la représentation graphique de la fonction de densité (ex3)
- calculer $P(|X| < a)$ (ex3)

Étant donné qu'il s'agit d'un cours magistral ces variations sont toutes prises en charge par l'enseignant. L'enseignant commence par définir une des nouveautés de ce cours, la fonction indicatrice : « Vous voyez apparaître cette indicatrice, il va falloir apprendre à travailler avec, une indicatrice vaut 1 quand la condition est vérifiée et vaut 0 quand la condition n'est pas vérifiée. Donc quand x est compris entre -1 et 1 vous remplacez ça (il parle de l'indicatrice) par 1. Quand x est ailleurs vous remplacez ça par 0. » Il écrit au tableau l'expression de f et ajoute : : « Vous voyez à gauche de l'accolade il y a deux conditions. Soit l'indicatrice vaut 1 soit elle vaut 0. Soit on est entre -1 et 1 soit on y est pas, quand on est entre -1 et 1 il faut utiliser la première formule pour trouver la valeur de la fonction, sinon il faut utiliser la seconde formule, et la deuxième formule elle est facile car c'est 0. »

L'enseignant représente au tableau le graphe de la fonction f en insistant sur chacun des morceaux définis par la fonction indicatrice. Il a conscience des difficultés que peut poser l'utilisation d'une telle fonction, de la difficulté à utiliser la définition de l'indicatrice afin d'obtenir une expression de la fonction f : « Tous les ans ça pose problème, la différence entre une fonction et une formule. Pour définir une fonction, je peux utiliser une formule ou une autre, je peux aussi définir des fonctions sans même avoir de formule pour les définir. Ici ceux qui sont à l'aise vont utiliser la notation avec l'indicatrice, ceux qui ne sont pas à l'aise, je vous encourage à tout de suite traduire les choses avec l'accolade, si x est là la fonction vaut ça, sinon elle vaut ça. »

Il termine en expliquant le lien entre fonction définie par une indicatrice et calculs d'intégrales : « Quand on a à calculer l'intégrale d'une fonction on aime bien calculer les intégrales avec des formules à l'intérieur. Vous allez utiliser des primitives, donc vous aimez bien avoir des formules. Et donc on fera attention, l'intégrale va d'abord devoir être découpée sur chaque intervalle où on va pouvoir utiliser une formule. »

Il détaille ensuite oralement les quelques éléments de réponse à **T-VerifF** inscrits sur la diapositive, à ce sujet un étudiant (E) interrompt l'enseignant (G) et demande :

E : *Pourquoi on est entre -1 et 1 et pas -2 par exemple ?* »

G : « *Ici la formule, si x est compris entre -1 et 1 il faut appliquer la formule du haut, si x vaut -2 on applique la deuxième formule et ça vaut 0, $f(-2)$ vaut 0.* »

E : « *Mais qui a déterminé l'intervalle $[-1; 1]$?* »

G : « *C'est moi ! J'ai dit f c'est cette chose-là. L'indicatrice vaut 1 quand la condition est vérifiée, elle vaut 0 sinon. L'intégrale qui est là, elle paraît grosse car elle va de moins l'infini à plus l'infini, c'est énorme. Sauf que regardez, l'aire sous*

la courbe (il montre en dehors de $[-1; 1]$) est très facile à calculer car c'est 0, la fonction est nulle. Tout ce qui reste à calculer c'est l'aire qui est là, entre -1 et 1, vous pouvez remplacer f par $\frac{3}{4}(1 - x^2)$. Et après c'est un calcul que vous savez faire depuis presque un an. Intégrer un polynôme de degré 2 avec une primitive vous devez savoir le faire. »

Montrer qu'une fonction est bien une densité est un type de tâches que l'on retrouvera en TD, l'enseignant à ce propos insiste sur la technique de résolution : « Si à un moment, je vous demande de vérifier que telle fonction est une densité de probabilité, vous devez vérifier ces deux conditions-là. C'est toujours positif, quel que soit la formule utilisée, et l'intégrale, l'aire sous la courbe vaut 1. »

Un étudiant demande à l'enseignant : « Je ne comprends pas comment on trouve que l'intégrale vaut 1, comment on fait le calcul avec la formule ? » l'enseignant refait le calcul au tableau en expliquant qu'il découpe l'intégrale en trois parties grâce à la relation de Chasles :

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{-1} f(x)dx + \int_{-1}^1 f(x)dx + \int_1^{+\infty} f(x)dx = 0 + \int_{-1}^1 \frac{3}{4}(1 - x^2)dx + 0$$

Un autre étudiant demande : « Mais j'ai pas compris, l'indicateur sert à savoir quelle formule il faut utiliser ? » l'enseignant ré-explique le fonctionnement de la fonction indicatrice.

L'enseignant poursuit de la même manière l'explication du calcul de probabilités, il donne encore beaucoup d'éléments oralement, il ne refait pas les calculs : « Comment on calcule la probabilité que X soit compris entre 0 et 1, on fait l'intégrale de la densité entre 0 et 1. On regarde X qui est entre 0 et 1 et on se dit, quelle formule je vais utiliser pour calculer $f(x)$? Je remplace ensuite $f(x)$ par la formule, et ensuite l'intégrale d'un polynôme de degré 2 vous savez faire ça. »

A propos de l'exemple 4 et de la fonction f , l'enseignant explique : « il va falloir d'abord déterminer la formule que vous devez utiliser. C'est une fonction définie par deux formules et assez souvent, il y a une des deux formules qui est 0. Ici, quand x est négatif la fonction est toujours nulle » il écrit au tableau ce qui suit et trace le graphe de f :

$$f(x) = \begin{cases} \frac{e^{-x/100}}{100} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

C'est la première rencontre pour les étudiants dans ce cours avec la loi exponentielle vue au lycée, l'enseignant explique : « La loi exponentielle mesure des durées de vie, quelle est la durée de vie de votre tortue de compagnie ? Est-ce qu'elle va vivre entre 50 et 150 ans ? Vous calculez l'intégrale entre 50 et 150. Vous l'avez déjà fait des centaines de fois, une primitive de $e^{\lambda x}$ c'est $\frac{1}{\lambda}e^{\lambda x}$, vous mettez tout ça vous calculez avec une calculatrice, vous obtenez une probabilité qui est à peine plus grande que 1/3. » Ici l'enseignant parle de calculatrice pour obtenir une

approximation du résultat final, mais les étudiants avaient pour habitude au lycée d'utiliser la calculatrice dès la première étape.

Un étudiant demande à l'enseignant : « *On peut avoir des densités qui ne sont pas continues ?* » ce à quoi il répond : « *Oui bien sûr. Les densités continues il y en a que dans les exercices de math, il n'y en a jamais dans la vraie vie, à part la Gaussienne, c'est le point faible du programme de lycée, vous imposez des densités continues.* »

Détails à propos de l'épisode n°5

4.3. ESPÉRANCE. VARIANCE

Définition

Soit X une v.a. de densité f , on définit, espérance, variance et écart-type par

- ① $E[X] = \int_{-\infty}^{+\infty} xf(x)dx$
- ② $Var(X) = E[(X - m)^2] = \int_{-\infty}^{+\infty} (x - m)^2 f(x)dx, \quad m = E[X].$
- ③ $\sigma(X) = \sqrt{Var(X)}.$

RAPPEL : Pour les v.a. discrètes, on avait la formule analogue

$$E[X] = \sum_i x_i P(X = x_i).$$

Comme pour les v.a. discrètes on a

$$E[h(X)] = \int h(x)f(x)dx$$

$$E[aX + b] = aE[X] + b, \quad a, b \in \mathbb{R}$$

$$Var(X) = E[X^2] - E[X]^2$$

$$Var(X + a) = Var(X)$$

$$Var(aX) = a^2 Var(X)$$

FIGURE 5.72 – Diapositive n°8 – Cours sur les lois continues

La diapositive n°8 (voir figure 5.72) présente l'espérance et la variance. L'enseignant indique que ce sont des indicateurs qui n'ont pas beaucoup de significations directes et que pour comprendre ce qu'ils signifient il faut utiliser des théorèmes et pas juste décrire les objets. Tout est nouveau sur cette diapositive sauf la définition de l'espérance. L'enseignant explique l'analogie avec les formules vues pour les variables aléatoires discrètes, il rappelle au tableau les formules vues dans le chapitre précédent (VA discrètes) et commente les similitudes et les différences. Il rajoute la formule suivante :

$$E(X + Y) = E(X) + E(Y)$$

A partir de l'exemple précédent (diapositive n°3 5.69) l'enseignant illustre, diapositive n°9 (voir figure 5.73), les nouvelles définitions qui viennent d'être énoncées : « *Le mode, la médiane et l'espérance mesurent, en quelque sorte, les positions de votre variable X de 3 manières différentes. Ces trois nombres coïncident très rarement.* »

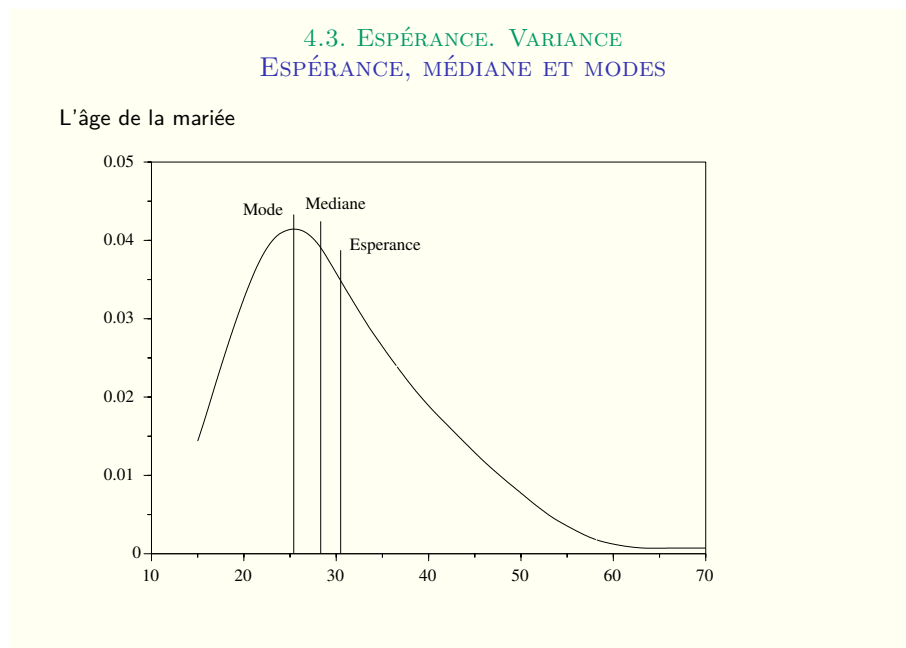


FIGURE 5.73 – Diapositive n°9 – Cours sur les lois continues

Afin de faire le lien avec des notions de statistiques qui seront vues les années suivantes, il explique, toujours oralement, en quoi des valeurs extrêmes dans les données peuvent modifier ces trois nombres caractéristiques : « *S'il y a beaucoup de mariages à 70 ans l'âge moyen de mariage va augmenter, mais ça va sûrement pas changer votre mode. Le mode est insensible aux phénomènes extrêmes, mais l'espérance, elle est très sensible. Donc c'est délicat, il arrive qu'on fasse des erreurs de mesure et qu'on introduise des valeurs extrêmes qui n'existent pas. Du coup ça va fausser l'espérance, mais le mode lui n'est pas du tout influencé et continue à avoir une vraie signification.* »

L'exemple de la diapositive n°10 (voir figure 5.74) va permettre de présenter des premiers calculs d'espérance et de variance aux étudiants : **T-CalcEsp** et **T-CalcVar** (nouveau) dans le cas d'une fonction de densité définie par morceaux.

L'enseignant commence par écrire l'expression de f car il y a une erreur dans l'énoncé :

$$f(x) = \mathbb{1}_{0 \leq x \leq 20} (20 - x) \times \frac{1}{200}$$

L'enseignant explique comment vérifier que la fonction est bien une densité (**T-Veriff**) et donne les détails oralement. La technique attendue ici est de reconnaître une forme géométrique caractéristique (triangle) et de calculer son aire à l'aide d'une formule élémentaire : « *Votre courbe vous dessine un triangle. Il n'y a pas besoin de calculer des intégrales pour vérifier que l'intégrale de moins l'infini à plus l'infini de $f(x)dx$ vaut 1. Malgré tout, si vous ne reconnaissez pas l'aire d'un triangle, utilisez des outils vus au lycée.* » Étant donné qu'il s'agit d'un cours magistral les sous-activités d'organisation et de traitement sont prises en charge

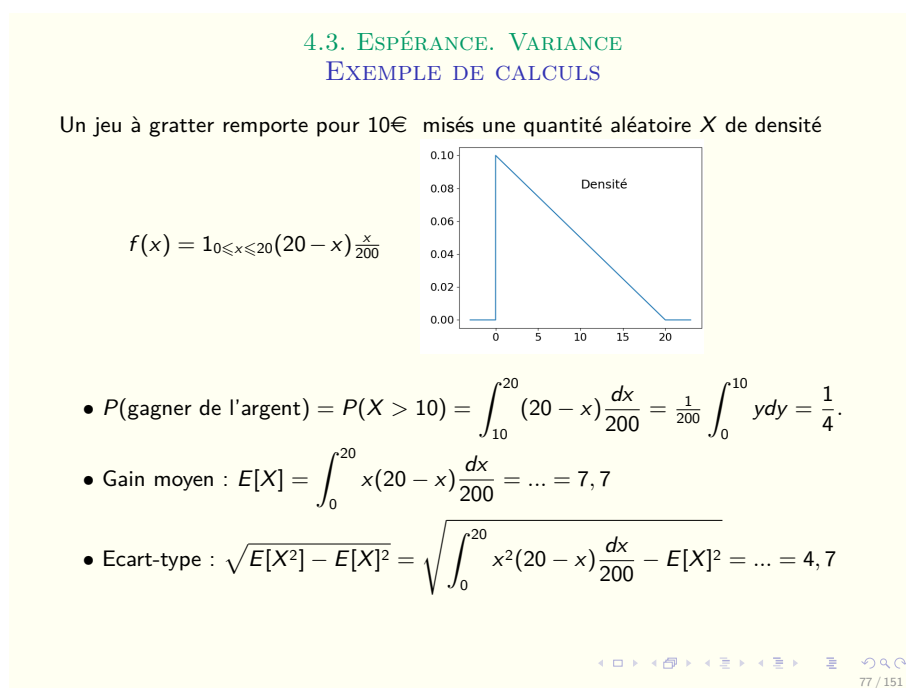


FIGURE 5.74 – Diapositive n°10 – Cours sur les lois continues

par l'enseignant.

La même technique est attendue pour la question suivante, **T-CalcSimple**. L'enseignant explique le passage de l'évènement $\{\text{gagner de l'argent}\}$ à l'évènement $P(X > 10)$: « *La probabilité de gagner de l'argent, c'est la probabilité de gagner au moins vos 10 euros que vous avez misés au départ. Et dans ce cas là c'est encore un triangle mais on peut aussi faire le calcul avec un changement de variables, on pose $y = 20 - x$, vous faites votre changement de variable et vous calculez une primitive* » L'enseignant propose une seconde technique pour calculer cette intégrale, un changement de variables, c'est une technique vue dans le module d'analyse du premier semestre.

L'enseignant détaille ensuite la méthode que les étudiants devront suivre pour calculer une espérance **T-CalcEsp** : « *Pour la moyenne, je vous conseille de toujours réécrire la formule que vous allez utiliser avant de faire des calculs, par quoi je remplace $f(x)$? Ça va dépendre de l'endroit où vit x . Ici vous découpez en trois intervalles et après vous remplacez. Puis vous n'avez plus qu'un polynôme et vous utilisez vos techniques habituelles* » il écrit au tableau le détail du calcul :

$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^0 x f(x) dx + \int_0^{20} x f(x) dx + \int_{20}^{+\infty} x f(x) dx \\ &= 0 + \int_0^{20} x \frac{20 - x}{200} dx + 0 = \dots = 7,7 \end{aligned}$$

Il termine avec le calcul de l'écart-type **T-CalcVar**, détaille la méthode et les

étapes successives : « Pour l'écart-type, vous allez d'abord calculer l'espérance de X^2 . Faites-le par étapes, on ne calcule pas l'écart-type directement. On découpe, prenez le temps de l'écrire, après je remplace f par ce que ça vaut, vous obtenez un polynôme, vous mettez tous les monômes, vous développez, vous distribuez votre intégrale, puis primitives, vous simplifiez vos fractions, et une fois que vous avez vos fractions simplifiées vous rentrez ça dans vos calculatrices » et il écrit au tableau :

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_{-\infty}^0 x^2 f(x) dx + \int_0^{20} x^2 f(x) dx + \int_{20}^{+\infty} x^2 f(x) dx$$

$$= 0 + \int_0^{20} x^2 \frac{20-x}{200} dx + 0 = \dots$$

$$Var(X) = E(X^2) - E(X)^2 \quad \text{et} \quad \sigma_X = \sqrt{Var(X)}$$

Étant donné qu'il s'agit d'un cours magistral les sous-activités d'organisation et de traitement sont prises en charge par l'enseignant.

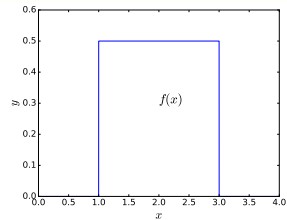
Détails à propos de l'épisode n°6

4.4. LOI CONTINUES CLASSIQUES
LOI UNIFORME

La loi uniforme sur l'intervalle $[a, b]$, $\mathcal{U}([a, b])$ est la loi de densité

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

Exemple $a = 1, b = 3$:



Infinité de modes. Médiane à $(b-a)/2$.

78 / 151

4.4. LOI CONTINUES CLASSIQUES
LOI UNIFORME

$$E[X] = \int x f(x) dx = \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} \left[\frac{1}{2} x^2 \right]_a^b = \frac{1}{2} \frac{b^2 - a^2}{b-a} = \frac{a+b}{2}$$

$$E[X^2] = \frac{1}{b-a} \int_a^b x^2 dx = \frac{1}{b-a} \left[\frac{1}{3} x^3 \right]_a^b = \frac{1}{3} \frac{b^3 - a^3}{b-a} = \frac{1}{3} (b^2 + ab + a^2)$$

$$Var(X) = E[X^2] - E[X]^2 = \dots = \frac{1}{12} (b-a)^2$$

$$\sigma(X) \simeq 0,3 (b-a).$$

79 / 151

FIGURE 5.75 – Diapositives n°11 et n°12 – Cours sur les lois continues

Le cours se poursuit, diapositives 11 et 12 (voir figure 5.75), par la présentation de la loi uniforme déjà vue au lycée, l'enseignant explique ce que modélise cette loi et ce qu'elle représente : « La loi uniforme, vous prenez un nombre au hasard entre a et b , cette expérience-là est modélisée par la variable dont la densité est représentée ici en bleu. C'est l'indicatrice de l'intervalle divisée par la longueur de l'intervalle. Vous pouvez calculer l'aire sous la courbe, vous avez affaire à un rectangle et on montre facilement que c'est bien une densité. »

Il passe ensuite rapidement sur les éléments caractéristiques de la loi uniforme (modes et médiane) puis détaille oralement la méthode à suivre pour le calcul de l'espérance **T-CalcEsp** et de l'écart-type **T-CalcVar** : « Pour l'espérance vous reprenez les formules, vous calculez l'intégrale de moins l'infini à plus l'infini de $xf(x)dx$. Comme votre fonction vaut 0 sauf entre a et b, pas besoin de choses compliquées ici pour calculer son intégrale, c'est un trapèze. Calculer l'aire d'un trapèze, ce sont des petits calculs que vous êtes censés maîtriser. Comme tout à l'heure, ce qui nous intéresse c'est l'écart-type, mais on le calcule par étapes. Ce qui est important ici, c'est que l'écart-type est proportionnel à la largeur de votre intervalle. Je vous laisse finir les calculs tout seul avant le prochain TD. »

Détails à propos de l'épisode n°7

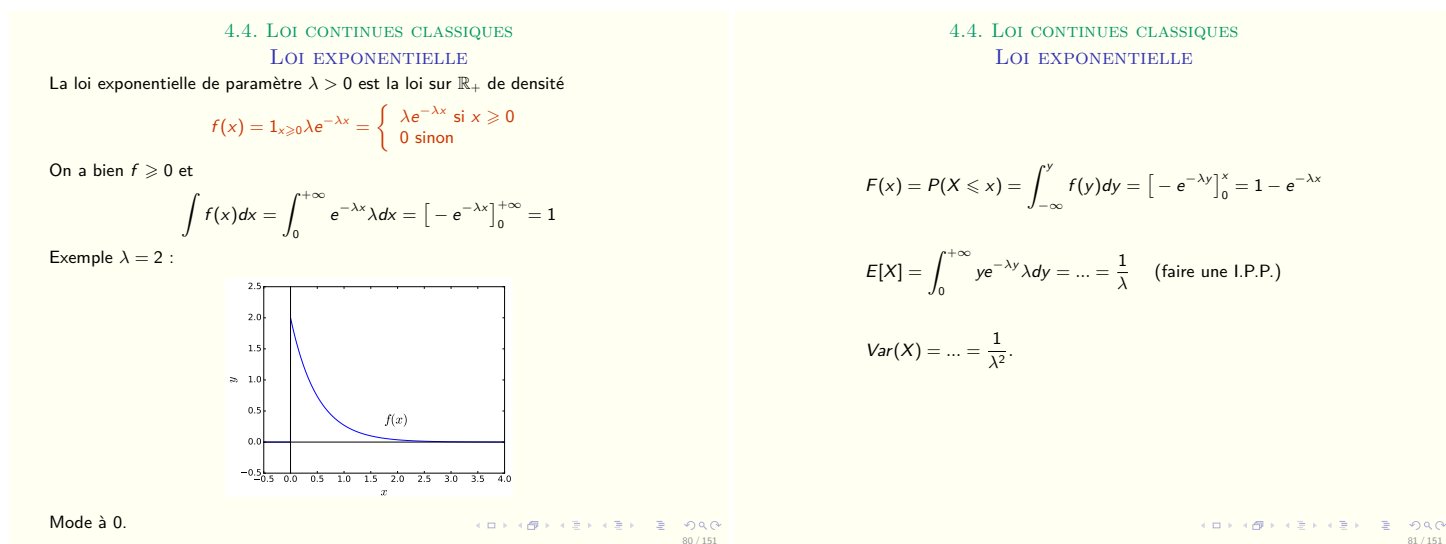


FIGURE 5.76 – Diapositives n°13 et n°14 – Cours sur les lois continues

Une autre loi, elle aussi vue au lycée, est présentée dans ce cours diapositives 13 et 14 (voir figure 5.76). Il s'agit de la loi exponentielle. L'enseignant explique que cette loi modélise des phénomènes de durée de vie, sans usure et à quoi correspond son espérance : « Une loi que vous n'avez peut être pas vue au lycée, c'est la loi exponentielle. En général, les lois exponentielles mesurent des durées de vie, de particules ou d'autres choses. Le λ qui est là c'est en fait l'inverse de l'espérance de votre variable aléatoire, donc ça sera l'âge moyen à la mort, la durée de vie moyenne. Après vous avez une décroissance qui est là pour vous dire que vous regardez un phénomène de durée de vie. Ce phénomène est sans usure, donc ce n'est pas comme j'ai dit tout à l'heure la durée de vie d'une tortue, c'est un phénomène sans mémoire, comme par exemple pour un atome de carbone 14. »

L'enseignant passe ensuite un peu de temps sur le type de tâches **T-VerifF** (montrer que f est bien une densité) présenté sur la diapositive, il détaille la dernière étape qui est une limite, et fait remarquer aux étudiants qu'ils doivent

connaître les limites remarquables de la fonction exponentielle.

L'enseignant poursuit : « *Ce qui est important avec la loi exponentielle c'est d'être capable de faire une intégration par parties. Vous avez vu ça au semestre précédent* », il parle ici du calcul de l'espérance d'une loi exponentielle qui fait intervenir une intégration par parties. C'est une grande différence avec le lycée où l'espérance d'une loi exponentielle n'était pas calculée mais juste attendue en tant que formule : $E(X) = \frac{1}{\lambda}$. Il détaille au tableau la méthode pour le calcul de l'espérance (**T-CalcEsp**) : « *Vous devez d'abord écrire la formule du cours. Ici, vous devez remarquer que l'intégrale de moins l'infini à 0 ça fait 0 parce que la loi est à densité et la densité s'annule quand x est négatif. Vous nommez votre variable d'intégration, je vous conseille de garder x , même si j'ai pris y sur la diapo du cours. Et après, j'ai calculé cette chose-là : quand vous avez un monôme multiplié par une exponentielle, la seule manière raisonnable de calculer une primitive, c'est de faire des intégrations par parties* » et il écrit :

$$\begin{aligned} E[X] &= \int_{-\infty}^{+\infty} x f(x) dx = \int_0^{+\infty} x \lambda e^{-\lambda x} dx \\ &= [-x e^{-\lambda x}]_0^{+\infty} - \int_0^{+\infty} -e^{-\lambda x} dx = 0 + \int_0^{+\infty} e^{-\lambda x} dx \\ &= \left[-\frac{1}{\lambda} e^{-\lambda x}\right]_0^{+\infty} = \frac{1}{\lambda} \end{aligned}$$

Détails à propos de l'épisode n°8

Le cours continue avec la loi normale centrée réduite, l'enseignant commence par décrire ce que modélise une telle loi : « *C'est une loi que vous avez tous vue, la formule qui donne la densité il faut la connaître par cœur. C'est le genre de lois avec lesquelles il est assez difficile de faire des calculs, mais ce sont des lois qui ont de très bonnes propriétés mathématiques et qui apparaissent un peu partout dès qu'on regarde des phénomènes que l'on moyennise. Quand on en regarde dix mille expériences et qu'on fait la moyenne des dix mille résultats on voit toujours des gaussiennes. Écrivez ça pour ceux qui n'ont jamais vu la fonction de répartition d'une gaussienne* » puis il explicite au tableau la fonction Φ :

$$P(X \leq t) = \Phi(t) = \int_{-\infty}^t \dots$$

Diapositive n°16 (voir figure 5.77), l'enseignant présente rapidement les lois normales, il insiste sur le lien entre loi normale et loi normale centrée réduite et explique le lien entre μ et σ et l'allure de la courbe : « *Une loi normale qui n'est pas centrée ni réduite, a une certaine moyenne, une certaine espérance μ , une certaine variance σ^2 et une certaine densité qui est celle écrite ici. Le μ c'est l'abscisse du haut de la cloche et le σ c'est quelque chose qui vous donne une information sur la largeur de la cloche. La chose la plus importante c'est la dernière ligne, si Y suit une loi normale, alors le X qui est défini par cette formule-là c'est une loi normale centrée*

4.4. LOI CONTINUES CLASSIQUES

LOI NORMALE

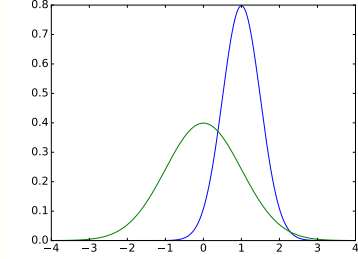
Si $X \sim \mathcal{N}(0, 1)$ alors la v.a. $Y = \sigma X + \mu$ satisfait

$$E[Y] = \mu, \quad \text{Var}(Y) = \sigma^2.$$

On vérifie, en calculant $P(a \leq Y \leq b) = P\left(\frac{a-\mu}{\sigma} \leq X \leq \frac{b-\mu}{\sigma}\right)$, que la densité de Y vaut

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

Cette loi est notée $\mathcal{N}(\mu, \sigma^2)$. Densités de $\mathcal{N}(0, 1)$ et $\mathcal{N}(1, \frac{1}{4})$



Réciproquement, si $Y \sim \mathcal{N}(\mu, \sigma^2)$, alors $\frac{Y-\mu}{\sigma} \sim \mathcal{N}(0, 1)$.

FIGURE 5.77 – Diapositive n°16 – Cours sur les lois continues

réduite. »

Il poursuit en introduisant les tables de lois normales qui sont tout à fait nouvelles pour les étudiants qui arrivent du lycée : « Est-ce que parmi vous il y en a qui n'ont jamais vu de tables de la loi normale ? Vous en avez une sur votre feuille de TD4. Cette table vous permet de calculer la probabilité que X soit plus petit qu'un certain nombre. Pour ceux qui ne l'ont jamais utilisée, on apprendra à l'utiliser et on apprendra à calculer la probabilité que Y soit plus petit qu'un certain nombre en utilisant cette table-là. »

Le cours se poursuit avec les intervalles à 1, 2 et 3 sigmas, déjà vus au lycée, l'enseignant explique l'interprétation des trois expressions écrites sur la diapositive n°17 (voir figure 5.78) : « Y va s'éloigner de μ de moins de σ avec une probabilité à peu près égale à $2/3$, la probabilité que Y soit proche de μ à σ près, c'est environ $2/3$. La probabilité que cet écart soit plus petit que 2σ devient à peu près égale à $0,95$, on prend donc un intervalle deux fois plus grand. Vous avez appris que pour obtenir exactement $0,95$ c'est avec $1,96$ mais si on met un 2 c'est plus facile de s'en souvenir et la probabilité est meilleure. Enfin, la probabilité d'être dans un intervalle trois fois plus grand, cette fois-ci avec 3σ à droite et 3σ à gauche, c'est quasiment 1 , c'est donc presque sûr, vous avez 99% de chance que votre Y soit dans cet intervalle. »

L'exemple présenté diapositive n°18 (voir figure 5.79) relève du type de tâches **T-Calcsimple** pour une loi normale de paramètres μ et σ donnés. L'enseignant indique qu'il s'agit d'un exemple qu'ils retrouveront dans les exercices de TD,

4.4. LOI CONTINUES CLASSIQUES

LOI NORMALE : ORDRE DE GRANDEUR

Si $X \sim \mathcal{N}(0, 1)$, $Y \sim \mathcal{N}(\mu, \sigma^2)$, alors

$$P(|Y - \mu| \leq \sigma) = P(|X| \leq 1) \simeq 0,68$$
$$P(|Y - \mu| \leq 2\sigma) = P(|X| \leq 2) \simeq 0,954$$
$$P(|Y - \mu| \leq 3\sigma) = P(|X| \leq 3) \simeq 0,997.$$

La probabilité que la variable sorte de l'intervalle $\mu \pm 3\sigma$ est donc extrêmement faible.
La variable tombe en gros deux fois sur trois dans l'intervalle $[\mu - \sigma, \mu + \sigma]$.




FIGURE 5.78 – Diapositive n°17 – Cours sur les lois continues

4.4. LOI CONTINUES CLASSIQUES

LOI NORMALE

Soit $Y \sim \mathcal{N}(175, 36)$, que vaut $P(Y > 185)$?

Posons $X = \frac{Y - \mu}{\sigma} \sim \mathcal{N}(0, 1)$, alors

$$P(Y > 185) = P\left(\frac{Y - \mu}{\sigma} > \frac{185 - 175}{6}\right) = P\left(X > \frac{5}{3}\right) \simeq 4,75\%.$$

Le dernière égalité étant obtenue en utilisant un ordinateur (ou une table pour les nostalgiques).




FIGURE 5.79 – Diapositive n°18 – Cours sur les lois continues

il précise oralement : « Ici l'espérance c'est 175 et la variance c'est 36, que l'on choisi presque toujours comme un carré histoire de simplifier les calculs, et on vous demande de calculer une probabilité. »

Étant donné qu'il s'agit d'un cours magistral les sous-activités d'organisation

et de traitement sont prises en charge par l'enseignant qui explique brièvement la méthode à suivre : « On commence par centrer et réduire la variable aléatoire, donc $Y > 185$ c'est la même chose que $\frac{Y-\mu}{\sigma} > \frac{185-\mu}{\sigma}$. Vous remplacez et vous appelez X la variable $\frac{Y-\mu}{\sigma}$. Donc si Y suit cette loi normale alors X suit une loi normale centrée réduite et vous utilisez votre table pour calculer cette probabilité, ou votre calculatrice. Vous allez refaire plein de fois des exemples de ce genre, c'est l'utilisation la plus basique que l'on fait de la loi normale. »

Détails à propos de l'épisode n°9

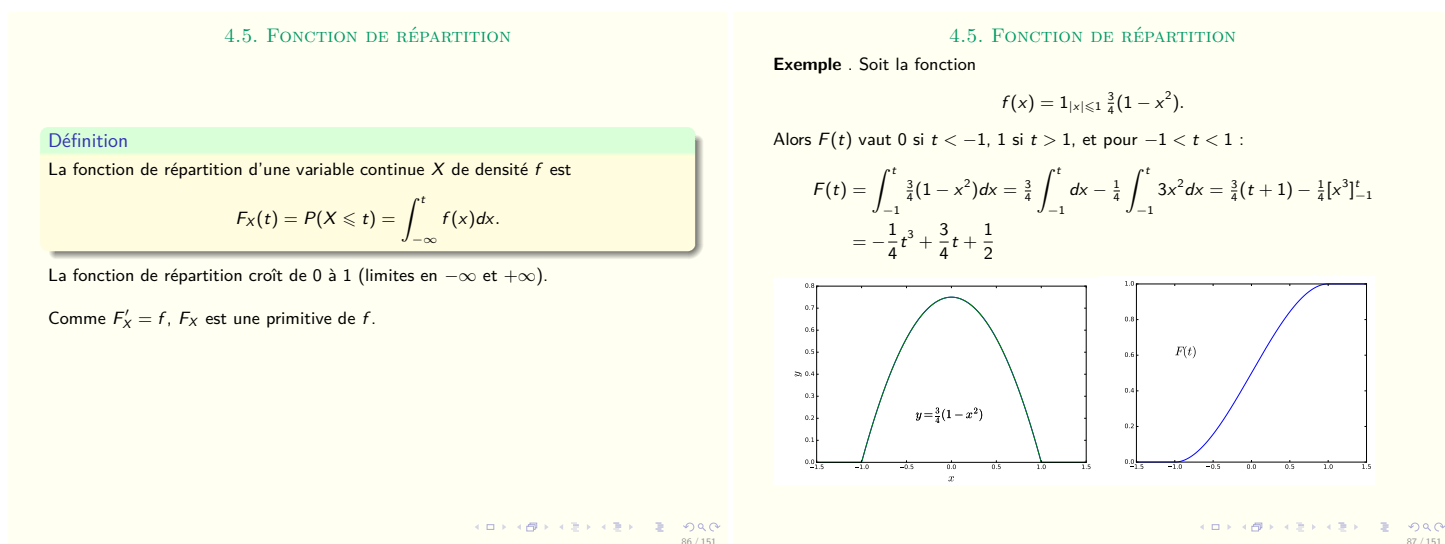


FIGURE 5.80 – Diapositives n°19 et n°20 – Cours sur les lois continues

Le cours se termine avec la présentation de la fonction de répartition F_X diapositives 19 et 20 (voir figure 5.80) : « La fonction de répartition permet de faire certains calculs sans avoir à calculer d'intégrale. Ici on ne veut pas savoir ce que vaut Ω , ce sont les issues d'une certaine expérience. f est une mesure sur les issues d'une certaine expérience, ça prend donc des valeurs qui sont des nombres. La fonction de répartition c'est la probabilité que ce nombre mesuré soit plus petit que t , c'est une fonction de t . En général f est une fonction de x , et votre fonction de répartition, il est préférable de l'écrire en fonction d'une autre variable. En particulier, il est important d'encadrer $F'_X = f$ et de voir que l'on a changé le nom des variables entre F et f , donc il ne faudra pas se tromper, c'est une relation de fonctions. » L'enseignant reprend un exemple déjà traité (exemple 3 diapositive n°4 figure 5.71 page 227) pour expliquer comment déterminer la fonction de répartition et il écrit :

$$F(t) = \int_{-\infty}^t f(x)dx = \begin{cases} \text{si } t \leq -1, F(t) = 0 \\ \text{si } t \geq 1, F(t) = 1 \\ \text{si } t \in]-1; 1[, F(t) = \int_{-1}^t f(x)dx \end{cases}$$

Détails à propos de l'épisode n° 10

Ce TD sur les lois continues est donné par le même enseignant que celui qui a fait le cours magistral. Il débute par l'exercice 4 dont l'énoncé se trouve figure 5.81). La correction faite par l'enseignant au tableau est à retrouver figure 5.82.

- Exercice 4** Soit X une variable aléatoire de densité $f(x) = cx^2 \mathbb{1}_{0 \leq x \leq 1} + \left(\frac{1}{2}\mathbb{1}_{2 \leq x \leq 3}\right)$
- Que vaut c ?
 - Calculer $P\left(\frac{1}{4} \leq X \leq \frac{3}{4}\right)$?
 - Calculer l'espérance et la variance de X .

FIGURE 5.81 – TD sur les lois continues – Énoncé de l'exercice 4

X de densité $f : \mathbb{R} \rightarrow [0; +\infty[$
 $x \mapsto cx^2 \mathbb{1}_{[0,1]}(x)$

(a) On cherche les valeurs de c telles que :

- f soit positive
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

La première condition est vérifiée lorsque $c \geq 0$ et la seconde donne :

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x)dx &= \int_{-\infty}^0 f(x)dx + \int_0^1 f(x)dx + \int_1^{+\infty} f(x)dx \\ &= 0 + \int_0^1 cx^2 dx + 0 = c \int_0^1 x^2 dx = \frac{c}{3} \end{aligned}$$

qui doit être égal à 1. Donc une seule valeur de c convient, $c = 3$

(b) Calculons $P\left(\frac{1}{4} \leq X \leq \frac{3}{4}\right) = P(X \in \left[\frac{1}{4}; \frac{3}{4}\right])$

$$= \int_{\frac{1}{4}}^{\frac{3}{4}} f(x)dx = \int_{\frac{1}{4}}^{\frac{3}{4}} 3x^2 dx = \left[x^3\right]_{\frac{1}{4}}^{\frac{3}{4}} = \frac{3^3}{4^3} - \frac{1}{4^3} = \frac{26}{32} \times 2 = \frac{13}{32}$$

(c) $E[X] = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^0 xf(x)dx + \int_0^1 xf(x)dx + \int_1^{+\infty} xf(x)dx$

$$= 0 + \int_0^1 x3x^2 dx + 0 = \int_0^1 3x^3 dx = \left[3\frac{x^4}{4}\right]_0^1 = \frac{3}{4}$$

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f(x)dx$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{(48 - 45)}{80} = \frac{3}{80}$$

FIGURE 5.82 – TD sur les lois continues – Correction de l'exercice 4 par l'enseignant

Cet exercice présente un type de tâches tout nouveau pour les étudiants qui sortent du lycée. Il s'agit de déterminer les valeurs possibles d'une constante c qui

apparaît dans l'expression de f , une densité de probabilité. On nommera ce type de tâches **T-Det(cf)**.

On retrouve dans cet exercice les types de tâches suivants déjà rencontrés lors du cours magistral sous forme d'exemples : **T-CalcSimple**, **T-CalcEsp** et **T-CalcVar** pour une densité f quelconque définie par une indicatrice.

L'enseignant donne des indications (orales et écrites au tableau) : « *Si vous ne savez pas comment commencer, essayez de tracer le graphe de la fonction f . c est un paramètre que l'on vous demande de déterminer dans la première question* » et il écrit l'expression de la densité au tableau.

L'indicatrice est la première difficulté de cet exercice. L'enseignant passe dans les rangs et aide les étudiants en les incitant à commencer par le graphe de f : « *Vous devez savoir ce que signifie ce 1. S'il y avait pas le 1 est-ce que vous sauriez me dessiner le graphe de cette fonction ? Dessinez sans le 1 et ensuite vous regardez dans votre cours.* »

Rapidement l'enseignant inscrit des aides procédurales au tableau, il fait des rappels de cours et trace l'allure de la fonction f . Les étudiants ne semblent pas à l'aise avec cette question, et l'enseignant finit par prendre en charge la reconnaissance de la méthode pour cette première question : « *Est-ce que quelqu'un peut me dire ce que vaut c ? Il faut que vous vous souveniez des deux propriétés qu'ont les densités, c'est une fonction positive, ici c'est le cas, et son intégrale sur toute la droite doit être égale à 1. Donc vous m'écrivez ce que ça veut dire* »

L'enseignant prend en charge l'organisation, la reconnaissance et le traitement de cette tâche mais laisse les étudiants faire le calcul de l'intégrale : « *Si c est positif la parabole est orientée vers le haut, je vous laisse regarder la seconde condition, il va y avoir du c qui traîne, vous me calculez cette intégrale.* »

Il circule dans les rangs mais les étudiants ont des difficultés avec la fonction f qui est définie par morceaux au moyen d'une indicatrice. L'enseignant donne des aides aux étudiants : « *Regardez le graphe plutôt que la formule. L'intégrale entre moins l'infini et 0 elle est facile à calculer, celle entre 1 et plus l'infini elle est facile à calculer et il vous en reste une troisième. Vous coupez en trois bouts, donc ça fait 0 si x est négatif, cx^2 si $0 < x < 1$ et 0 si $x > 1$. Vous remplacez, sur chacun des segments, f est définie par une seule formule, on remplace donc f par sa formule dans chacun des cas* » il poursuit la correction au tableau :

$$\begin{aligned}\int_{-\infty}^{+\infty} f(x)dx &= \int_{-\infty}^0 f(x)dx + \int_0^1 f(x)dx + \int_1^{+\infty} f(x)dx \\ &= 0 + \int_0^1 cx^2dx + 0 = c \int_0^1 x^2dx = c/3\end{aligned}$$

Il passe dans les rangs pour accompagner les étudiants dans la fin du calcul : « *Vous savez faire ça, vous avez déjà calculé plein de fois cette intégrale* ». Un étudiant

propose pourtant comme résultat $c/\sqrt{6}$ et un autre 6. L'enseignant redonne la formule générale pour la primitive de x^n mais il ne donne pas le détail du calcul pour l'intégrale de 0 à 1 de $x^2 dx$. Il explique : « Vous avez obtenu $x^3/3$ évalué en 1 ça vaut un tiers et en 0 ça vaut 0. Puis ce qu'on obtient ça doit être égal à 1 » et il termine de corriger cette première question.

Un étudiant demande des précisions sur le calcul de $\int_0^1 cx^2 dx$: « On ne met pas c au carré ? On n'intègre pas les constantes ? » Outre l'activité de reconnaissance de la tâche, l'activité de traitement pose elle aussi des difficultés aux étudiants.

La seconde question qui relève du type de tâches **T-CalcSimple**, est traitée de la même façon. L'enseignant prend en charge la reconnaissance de la méthode et l'organisation. En effet, il écrit au tableau $P(1/4 \leq X \leq 3/4) = P(X \in [1/4; 3/4])$ puis la traduction en termes d'intégrales. Il laisse quelques instants aux étudiants pour finir les calculs (traitement), mais les étudiants semblent en difficulté et l'enseignant donne des aides procédurales : « Comment on calcule cette probabilité à partir de la densité de X ? Utiliser vos formules, commencer par me dire "pour X une variable aléatoire de densité f ce nombre-là se calcule de cette manière". Regardez quelle formule on utilise, est-ce qu'on en utilise 1, 2 ou 3 ? Est-ce qu'on doit redécouper ? Remplacez f par les formules qui conviennent et faites les calculs. »

L'enseignant circule dans la classe puis prend en charge la sous-activité de traitement en corrigeant cette question au tableau. Il détaille le choix de la formule à utiliser pour f suivant les bornes de l'intégrale et passe un peu de temps sur la simplification des fractions car à l'examen il attend les résultats sous forme de fractions irréductibles (il le dit oralement) et la calculatrice ne sera pas autorisée.

Pour la troisième question, qui relève de **T-CalcEsp** et **T-CalcVar**, l'enseignant rappelle la définition de l'espérance mathématique au tableau, puis laisse les étudiants faire les calculs (organisation + traitement). Il prend en charge la sous-activité d'organisation car il détaille, à nouveau, la méthode qu'il souhaite que les étudiants appliquent : « Vous commencez par écrire la définition de l'espérance, qui se trouve dans votre cours. Ensuite vous venez couper ça en trois bouts et remplacer f par ce qu'il vaut sur chacun des trois bouts et vous faites le calcul. Normalement si vous avez compris ensuite vous enchaînez avec le calcul de l'espérance de X^2 . Je veux voir le découpage, vous devez faire la même chose qu'au-dessus, découper en trois bouts. »

Un étudiant essaye de découper en faisant intervenir $1/4$ et $3/4$, l'enseignant ré-explique le découpage en s'appuyant sur le graphe de f . Il termine la correction au tableau et pour la fin du calcul il ajoute : « Vous n'avez plus qu'à calculer la primitive d'un monôme, vous savez le faire ».

Il poursuit avec le calcul de la variance, qui, rappelons le, est une nouveauté par rapport aux contenus du lycée, dans le cas d'une variable aléatoire continue.

Il prend en charge les sous-activités de reconnaissance et d'organisation en rappelant au tableau la formule du cours. C'est encore le même procédé qu'il explique aux étudiants : « *Maintenant pour l'espérance de X^2 c'est exactement de la même façon, je vous rappelle la formule du cours, le carré ne porte pas sur f attention* » et il prend finalement en charge la sous-activité de traitement en corrigeant rapidement ce dernier calcul.

L'enseignant demande aux étudiants de refaire cet exercice mais avec une autre fonction f . En effet, il y a deux versions différentes de la feuille de TD4 donc deux fonctions f différentes selon les feuilles. La nouvelle fonction f est la suivante :

$$X \text{ de densité } f : \mathbb{R} \rightarrow [0; +\infty[\\ x \mapsto cx^2 \mathbb{1}_{[0,1]}(x) + \frac{1}{2} \mathbb{1}_{[2,3]}(x)$$

Un étudiant remplace directement c par 3 qui est la valeur obtenue dans la version précédente de l'exercice, au lieu de chercher la nouvelle valeur de c . Les étudiants semblent encore en difficultés car il s'agit désormais de la somme de deux fonctions continues par morceaux. L'enseignant passe dans les rangs pour accompagner chaque étudiant dans les calculs. Il passe du temps sur la simplification des fractions et écrit le détail de cette simplification au tableau.

Détails à propos de l'épisode n° 11

Le TD se poursuit par l'exercice 3 dont l'énoncé se trouve figure 5.83). La correction faite par l'enseignant au tableau est à retrouver figure 5.84.

Exercice 3 La durée de vie en années d'une machine à laver suit une distribution $\mathcal{E}(\frac{1}{10})$.

- Quelle est son espérance de vie?
- Quelle est la probabilité qu'elle tienne au moins une année?
- Quelle est la probabilité qu'elle tienne au moins 10 ans?

FIGURE 5.83 – TD sur les lois continues – Énoncé de l'exercice 3

L'exercice 3 propose les types de tâches suivants : **T-CalcEsp** et **T-CalcSimple** dans le cas d'une loi exponentielle.

Avant de laisser les étudiants en autonomie sur l'exercice 3 l'enseignant s'adresse à eux : « *Vous avez trois questions dans cet exercice, vous avez une question : "Quelle est son espérance" à laquelle vous pouvez répondre sans faire de calcul. La loi exponentielle est une loi qu'on a regardée en cours, il y a un certain nombre de résultats que vous pouvez utiliser et en particulier l'espérance. Mais si vous ne connaissez pas la formule vous faites le calcul. Mais d'abord vous me trouvez dans votre cours la densité d'une loi exponentielle.* »

Pour la première question il attend juste la réponse, qu'elle soit donnée ou calculée, il choisit ici de refaire le calcul (et l'intégration par parties) avec les

La durée de vie de la machine est notée X , elle suit une loi exponentielle. $E(\frac{1}{10})$ signifie que X est une VA à densité et que sa densité est ...

$$f(x) = \frac{1}{10} e^{-\frac{x}{10}} \mathbb{1}_{x \geq 0}$$

Calculons $E[X]$

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx = \int_0^{+\infty} x f(x) dx = \int_0^{+\infty} \frac{x}{10} e^{-\frac{x}{10}}$$

On fait une intégration par parties avec :

$$u = \frac{x}{10} \quad u' = \frac{1}{10} \quad v' = e^{-\frac{x}{10}} \quad \text{et} \quad v = 10e^{-\frac{x}{10}}$$

$$\begin{aligned} E[X] &= \left[\frac{x}{10} (-10e^{-\frac{x}{10}}) \right]_0^{+\infty} - \int_0^{+\infty} \frac{1}{10} - 10e^{-\frac{x}{10}} dx \\ &= [-xe^{-\frac{x}{10}}]_0^{+\infty} + \int_0^{+\infty} e^{-\frac{x}{10}} dx \\ &= 0 + \int_0^{+\infty} e^{-\frac{x}{10}} dx \\ &= [-10e^{-\frac{x}{10}}]_0^{+\infty} = 10 \end{aligned}$$

(b) Calculons $P(X \geq 1) =$

$$= \int_1^{+\infty} f(x) dx = \int_1^{+\infty} \frac{1}{10} e^{-\frac{x}{10}} dx = [-e^{-\frac{x}{10}}]_1^{+\infty} = e^{-1/10}$$

(c) Calculons $P(X \geq 10) =$

$$= \int_{10}^{+\infty} f(x) dx = \int_{10}^{+\infty} \frac{1}{10} e^{-\frac{x}{10}} dx = [-e^{-\frac{x}{10}}]_{10}^{+\infty} = e^{-1}$$

FIGURE 5.84 – TD sur les lois continues – Correction de l'exercice 3 par l'enseignant

étudiants.

L'enseignant écrit la densité de la fonction exponentielle avec une indicatrice et ajoute : « C'est cette fonction-là quand x est positif. C'est une durée de vie à partir d'un moment fixé, donc ça ne peut pas être un nombre négatif. C'est pour ça que l'on utilise l'indicatrice, vous devez connaître cette formule. On vous demande ensuite un certain nombre de calculs. On commence par la définition, je vous laisse remplacer et faire les calculs. » Il écrit au tableau la définition de $E[X]$ et le début du calcul de l'espérance, il prend ainsi en charge la reconnaissance de la méthode, l'organisation et ne laisse que l'activité de traitement à la charge des élèves.

Il donne des aides procédurales aux étudiants quant à ce calcul (faire une intégration par parties) : « C'est cette quantité que l'on veut calculer, c'est un nombre. Une exponentielle contre une puissance de x c'est toujours avec une intégration par parties. Qui est u , u' , qui est v et v' ? C'est un calcul qu'on a fait en cours mais je

veux que vous le refassiez. On veut se débarrasser du x . »

L'enseignant détaille étape par étape l'intégration par parties puis passe du temps sur la détermination des limites lorsque x tend vers plus l'infini. Il conclut ce calcul en donnant une interprétation de la valeur trouvée : « *La loi exponentielle de paramètre λ ça modélise quasiment toujours des durées de vie, alors pas de machines à laver parce que ça modélise des durées de vie de phénomènes sans usure, et une machine à laver ça s'use malheureusement. Le paramètre c'est l'inverse de la durée de vie moyenne.* »

Pour les deux questions suivantes il prend en charge la reconnaissance de l'évènement et à travers elle l'interprétation de la question en langage naturel. Il précise aux étudiants : « *On cherche maintenant la probabilité que X soit plus grand que 1, c'est-à-dire la durée de vie soit d'au moins un an. La probabilité que la machine tienne au moins une année. Donc vous rappelez la formule du cours, vous remplacez puis vous calculez.* »

Il conclut sur cet exercice en donnant l'interprétation des valeurs trouvées : « *C'est le même calcul, ça fait 0,368. Donc un peu moins d'une chance sur deux que votre machine ait une durée de vie supérieure à dix ans. Donc en moyenne on a une durée de vie de dix ans, mais malgré tout si vous prenez plein de machines et que vous les faites fonctionner y en a qu'un tiers qui vont survivre plus de dix ans, les autres vont survivre moins de dix ans.* »

Détails à propos de l'épisode n° 12

Le TD se termine avec l'exercice 2 (voir figure 5.85) qui est traité très rapidement car c'est la fin de la séance de TD. La correction faite par l'enseignant au tableau est à retrouver figure 5.86.

Exercice 2 Pour cet exercice, on utilisera la table au dos. Soit $X \sim \mathcal{N}(0, 1)$, $Y \sim \mathcal{N}(5, 9)$.
(a) Trouver $P(-1 \leq X \leq 1)$, $P(-2 \leq X \leq 2)$, $P(0 \leq X \leq 1)$.
(b) Trouver $P(2 \leq Y \leq 8)$, $P(Y \geq 2)$, $P(Y \leq 8)$.

FIGURE 5.85 – TD sur les lois continues – Énoncé de l'exercice 2

Cet exercice propose **T-CalcSimple** pour une loi normale. Les variations présentes ici sont les suivantes :

- étant donné que la variable aléatoire est identifiée dans l'énoncé
- étant donné qu'il s'agit d'une loi normale (centrée réduite ou non)
- étant donné qu'il s'agit d'une loi normale de paramètres μ et σ donnés
- calculer $P(X > a)$ ou $P(a < X < b)$ étant donné qu'on connaît $P(X < a)$ et/ou $P(X > b)$

Soit $X \sim \mathcal{N}(0, 1)$ on a $P(X \in [a, b]) = ?$: on ne peut pas calculer ça seul

$P(X \leq a, bc) = ?$: a, b vous donne la ligne dans la table et $0,0c$ vous donne la colonne.

$P(X \leq 1, 41) \simeq 0, 9207$

$P(X \leq ?) \simeq 0, 975$ alors $? = 1, 96$

(a) $P(0 \leq X \leq 1) = P(X \leq 1) - P(X \leq 0)$

Dans l'exercice :

$$\begin{aligned}P(|X| \leq 1) &= P(X \leq 1) - P(X \leq -1) \\ &= P(X \leq 1) - P(X \geq 1) \\ &= P(X \leq 1) - (1 - P(X \leq 1))\end{aligned}$$

FIGURE 5.86 – TD sur les lois continues – Correction de l'exercice 2 par l'enseignant

L'enseignant commence par demander aux étudiants si certains n'ont jamais utilisé de tables de lois normales. Ils sont nombreux à répondre positivement et il prend alors du temps pour expliquer aux étudiants comment utiliser une table de loi normale centrée réduite : « *Ce qui est donné ici c'est la probabilité que notre X soit plus petit qu'un nombre qui s'écrit a, bc . Vous lisez le a, b en ordonnée, sur la colonne, ça vous donne donc la ligne et le $0,0c$ vous donne la colonne. Si vous voulez la probabilité que X soit plus petit que $1, 41$, vous regardez la ligne $1, 4$ et la colonne $0, 01$ et vous obtenez $0, 9207$. »*

Il poursuit avec un petit exemple indépendant de l'exercice n°2, il s'agit de **T-Det(a)** : déterminer le réel a tel que $P(X < a) = \text{nombre donné}$. Il demande aux étudiants : « *Quelle est le nombre tel que la probabilité d'être inférieur ou égal à ce nombre, soit égale à $0, 975$? Vous cherchez $0, 975$ dans la table et puis vous regardez quelle ligne et quelle colonne vous donnent $0,975$. Ici c'est $1, 9$ et $0, 06$. La table s'utilise dans les deux sens mais on a mis que les nombres positifs car on peut tout faire avec. »*

Concernant le premier calcul de l'exercice n°2 (**T-CalcSimple** pour une loi normale centrée réduite), l'enseignant prend entièrement en charge les sous-activités associées à la tâche (reconnaissance, organisation et traitement). Il ne détaille pas le découpage de l'expression (relation de Chasles) mais explique rapidement comment trouver les valeurs dans la table des deux probabilités suivantes : $P(X \leq 1)$ et $P(X \leq 0)$.

A propos du calcul suivant, il explique : « *Là on a un problème car -1 n'est pas dans le tableau, car on a que les nombres positifs. Donc on va regarder les propriétés de symétrie de la gaussienne. On va dire que cette aire-là c'est la même que cette aire-*

là, on lit dans le tableau, et ici je dis que c'est la même chose que cette probabilité-là par symétrie. Maintenant j'ai bien un nombre positif mais l'égalité n'est pas dans le bon sens, j'utilise l'identité des événements complémentaires. Ici on a des lois continues donc mettre inférieur strict ou inférieur ou égal ça ne change rien. Maintenant on a deux nombres qui sont dans le tableau» et il écrit :

$$\begin{aligned}P(|X| \leq 1) &= P(X \leq 1) - P(X \leq -1) \\ &= P(X \leq 1) - P(X \geq 1) \\ &= P(X \leq 1) - (1 - P(X \leq 1))\end{aligned}$$

L'enseignant prend ici en charge à nouveau les sous-activités relatives à la tâche, il détaille le découpage de la probabilité. A aucun moment les étudiants n'ont la possibilité de se demander comment utiliser la table dans ce cas et pourquoi il est nécessaire de transformer l'expression.

Synthèse

Nous présentons dans cette partie la synthèse des résultats et des analyses détaillés ci-dessus, dans le cas des lois continues à l'Université. Il s'agit d'une synthèse en termes de comparaison et de transition car les analyses sont faites en regard de ce que nous avons vu au lycée (partie 5.2). La comparaison est donc intégrée dans ce que nous décrivons dans la suite de cette partie.

Nouveautés :

Les trois lois principales de ce cours (uniforme, exponentielle et normale) ont déjà été vues en classe de terminale au lycée. Les nouveautés de ce cours, en termes de contenus et de types de tâches, par rapport au cours observé en terminale, sont les suivantes :

- la fonction de densité peut être discontinue et définie avec une indicatrice
- T-Det(cf) : déterminer les valeurs possibles d'une constante c qui apparaît dans l'expression de f , fonction de densité définie par morceaux.
- le mode (maximum local de la densité) et la médiane (plus petite valeur vérifiant $P(X \leq m) = \frac{1}{2}$)
- les propriétés de l'espérance (transfert, linéarité)
- la variance et ses propriétés
- T-CalcVar : calculer la variance ou l'écart-type d'une variable aléatoire continue
- T-CalcEsp : calculer explicitement l'espérance mathématique d'une variable aléatoire de loi exponentielle (intégration par parties)

- la table de la loi normale centrée réduite
- les intervalles caractéristiques pour la loi normale centrée réduite (valeurs différentes de celles vues au lycée)
- la fonction de répartition F_X dans un cas général (Φ pour la loi normale a été vue dans des exercices au lycée)

Les fonctions de densité discontinues sont une grande nouveauté pour les étudiants qui sortent du lycée puisque dans le secondaire seules les fonctions de densité continues sont au programme. Ces fonctions sont généralement définies au moyen d'une fonction indicatrice dans le cours de l'Université. L'enseignant va passer du temps à expliquer ce que représente la fonction indicatrice et comment l'utiliser pour calculer une probabilité ou une espérance : « *Vous allez utiliser des primitives, donc l'intégrale va d'abord devoir être découpée sur chaque intervalle où on va pouvoir utiliser une formule.* »

L'utilisation des tables de lois normales pour des calculs relevant des types de tâches **T-CalcSimple** ou **T-Det(a)** est caractéristique de cet enseignement à l'Université. C'est tout à fait nouveau pour les étudiants qui utilisaient auparavant la calculatrice pour ce genre de tâche, mais l'enseignant ne semble pas totalement en avoir conscience, il dit lors du TD : « *Est-ce que quelqu'un n'a jamais utilisé de tables de lois normales ?* »

L'enseignant motive l'utilisation de ce nouvel instrument pour le cas des lois normales : « *On a un problème avec ces lois-là c'est qu'on ne peut pas calculer avec des formules, donc il faut trouver des valeurs approchées* » et étant donné que la calculatrice n'est pas autorisée lors de l'examen, il n'y a plus que la table de loi normale centrée réduite qui permet de faire les calculs.

Déroulement :

Nous présentons dans ce qui suit les nouveautés en termes de contrat didactique, ou de responsabilités vis-à-vis du savoir en jeu pour le thème des lois continues.

Durant le cours magistral, l'enseignant ne refait pas systématiquement les calculs d'intégrales. Généralement les détails ne sont pas écrits sur les diapositives et les étudiants doivent faire eux-mêmes ces calculs. Il semblerait que ces calculs ne sont pas détaillés en CM car ils relèvent de compétences acquises au lycée (par exemple : techniques géométriques de calculs d'intégrales) ou au semestre précédent dans le module d'analyse (par exemple : techniques d'intégration par parties et de changement de variables, notion de fonction indicatrice). L'enseignant souhaite que les étudiants s'exercent par eux-mêmes : « *Normalement vous avez revu ça au semestre précédent, nous on part du principe que vous savez faire ces calculs-là.* »

En revanche, l'enseignant prend un peu de temps sur la simplification des fractions car à l'examen il attend les résultats sous forme de fractions irréductibles (il

le dit oralement) et la calculatrice ne sera pas autorisée donc les étudiants doivent savoir le faire « à la main ».

Statistiques :

Il s'agit ici de remarques propres à un contenu spécifique, les statistiques. Nous nous plaçons, dans ce paragraphe, dans une perspective de ce qui se passe à l'Université sans comparaison avec ce qui se fait au lycée. En effet, les statistiques sont enseignées en terminale scientifique dans un chapitre à part entière, nous n'avons pas assisté aux séances dédiées ni analysé les chapitres de manuels associés.

Afin de faire le lien avec des notions de statistiques qui seront vues les années suivantes, l'enseignant présente le mode et la médiane par leur intérêt dans l'analyse qualitative de grands jeux de données : « *Ces variables aléatoires, dès que vous allez faire des statistiques l'année prochaine, elles sont là pour représenter des ensembles de données gigantesques. Donc il faut avoir quelques nombres qui caractérisent les comportements de notre variable aléatoire.* »

Durant le CM, l'enseignant présente certains liens entre données statistiques et variables aléatoires continues. Par exemple, lorsqu'il explique, d'un point de vue qualitatif, en quoi des valeurs extrêmes dans les données sont susceptibles de modifier les nombres caractéristiques mode, médiane et moyenne.

Modélisation :

Durant le cours magistral, l'enseignant fait de nombreuses références à des situations de modélisation, s'appuyant sur l'analyse de données par exemple. Il présente aux étudiants le lien entre statistiques, données réelles et variables aléatoires continues, mais ces thèmes ne seront pas travaillés en TD ni évalués lors de l'examen final. Ce sont des exercices plus mathématiques et pas directement liés à un phénomène que l'on modélise (selon les termes de l'enseignant), sur lesquels les étudiants vont s'exercer en TD.

A propos de la loi exponentielle, l'enseignant explique durant le cours que cette loi modélise des phénomènes de durée de vie, sans usure. Il propose une interprétation assez générale de l'espérance d'une telle loi : l'espérance sera l'âge moyen à la mort, la durée de vie moyenne.

A propos des lois normales, l'enseignant explique que ces lois apparaissent dès que l'on regarde des phénomènes que l'on moyennise. L'interprétation de la variance et de l'écart-type pour ces lois est importante et significative selon l'enseignant.

Types de tâches, variations et techniques :

De façon générale, les techniques attendues et présentées par l'enseignant pour **T-CalcSimple**, **T-CalcEsp** et **T-CalcVar** sont très similaires. Les ingrédients de la technique qui représentent ici les actions successives pour effectuer la tâche sont les suivants :

- premièrement, réécrire la définition de l'espérance ou de la probabilité recherchée,
- deuxièmement, découper l'intervalle selon la fonction de densité f (si elle est définie par morceaux ou non),
- enfin, calculer chacune des intégrales (logiciel de calcul, techniques géométriques, primitives, changements de variables ou intégration par parties)

Cette technique n'est pas celle que nous avons observée dans le secondaire, d'une part car les fonctions de densité ne sont jamais continues par morceaux, d'autre part car les changements de variables et intégrations par parties ne sont pas au programme de la classe de terminale scientifique. Enfin, pour la loi exponentielle qui pourrait se traiter au lycée par des techniques de calculs d'intégrales, c'est la technique instrumentée (calculatrice) qui est quasiment toujours privilégiée.

L'enseignant attend un graphe de la fonction de densité au début de chaque exercice où une fonction f quelconque est définie (ex4 du TD par exemple) même si cela n'est pas demandé.

Voici dans le détail les types de tâches, variations et techniques rencontrés dans ce cours. Nous comparons systématiquement si les tâches proposées aux étudiants sont des nouveautés ou si les techniques sont différentes de celles vues au lycée :

- **T-Veriff** : *vérifier qu'une fonction donnée est bien une densité de probabilité.*
 - La technique présentée par l'enseignant, pour ce type de tâche, est la même que celle vue au lycée : vérifier les deux conditions du cours qui sont : être positif sur l'ensemble de définition (quel que soit la formule utilisée), et justifier que l'aire sous la courbe vaut 1.
 - Vérifier la seconde condition se fait au moyen de techniques d'intégration que nous n'avons pas observées au lycée. En revanche, selon l'expression de la fonction de densité, la technique attendue peut être de reconnaître une forme géométrique caractéristique (triangle, rectangle, trapèze) et de calculer son aire à l'aide d'une formule de géométrie élémentaire. Nous avons observé la mise en oeuvre de cette technique géométrique au lycée.
 - Ce type de tâches est traité dans le cas d'une loi exponentielle (nouveau par rapport au lycée). L'enseignant attend des étudiants qu'ils soient capables d'établir les limites remarquables de la loi exponentielle, nécessaires au calcul de l'intégrale sur le domaine de définition.
- **T-CalcSimple** : *calculer la probabilité d'un évènement simple.*

- Dans le cas d'une loi quelconque la technique présentée par l'enseignant est la suivante :
 - * commencer par interpréter la question en langage naturel lorsque cela est nécessaire
 - * traduire la probabilité recherchée en termes d'intégrales
 - * choisir la formule adéquate à utiliser selon les bornes de l'intégrale (par exemple fonction définie par morceaux)
 - * découper l'intégrale en fonction de l'expression de f et des bornes
 - * calculer la (ou les) intégrale(s)

Finalement, hormis les deux premières étapes, le reste relève de techniques d'analyse.

- Selon l'expression de la fonction de densité, les techniques de calculs d'intégrales proposées par l'enseignant sont diverses. La technique attendue peut être géométrique, tout comme pour **T-Veriff**. L'enseignant propose également la technique de changement de variables qui a été vue dans le module d'analyse du premier semestre mais qui n'est pas élémentaire pour les étudiants (nouveau par rapport au lycée).
 - Lors du TD (ex3 - loi exponentielle voir figure 5.83 page 242) l'enseignant prend en charge la reconnaissance de l'évènement et l'interprétation de la question en langage naturel, il précise aux étudiants que chercher la probabilité que X soit plus grand que 1, revient à s'intéresser à ce que la durée de vie soit d'au moins 1 an. Il attend ensuite un calcul d'intégrales et non pas l'application d'une formule comme cela était fait au lycée.
 - Une autre nouveauté par rapport au lycée concerne les lois normales, il s'agit de l'utilisation de la table de loi normale centrée réduite. Pour le calcul $P(0 \leq X \leq 1)$ l'enseignant ne détaille pas le découpage de l'expression mais explique comment trouver les valeurs des deux probabilités : $P(X \leq 1)$ et $P(X \leq 0)$ dans la table.
Face à une probabilité de la forme $P(X \leq -1)$ l'enseignant explique aux étudiants comment transformer l'expression (utiliser les propriétés de symétrie de la gaussienne et de l'évènement contraire) afin de pouvoir utiliser la table de loi normale et illustre ces propriétés sur un dessin.
- **T-CalcEsp** : *calculer l'espérance d'une variable aléatoire continue.*
 - La technique présentée par l'enseignant est très similaire à celle présentée pour **T-CalcSimple** :
 - * écrire la formule du cours à utiliser
 - * remplacer $f(x)$ par la bonne formule (fonction définie par une indicatrice)
 - * découper l'intervalle si besoin
 - * calculer chacune des intégrales (souvent il s'agit de polynômes ou d'exponentielles)
 - Concernant la loi uniforme, la technique attendue par l'enseignant est de reconnaître une forme géométrique caractéristique (trapèze) et de calculer son aire à l'aide d'une formule de géométrie élémentaire (nouveau par rapport au lycée où il s'agit juste de l'application de la formule $(a+b)/2$).

- Concernant la loi exponentielle, l'enseignant attend un calcul explicite au moyen d'une intégration par parties : « *Ce qui est important avec la loi exponentielle c'est d'être capable de faire une intégration par parties. Vous avez vu ça au semestre précédent.* » C'est une grande différence avec le lycée où l'espérance d'une loi exponentielle n'était pas calculée mais juste attendue en tant que formule : $1/\lambda$. Dans le TD l'enseignant traite dans le détail ce type de tâche, il commence par réécrire la définition de $E[X]$, puis prend en charge la reconnaissance de la méthode, l'organisation et laisse l'activité de traitement (intégration par parties) à la charge des étudiants, l'enseignant conclut le calcul en donnant une interprétation de la valeur trouvée.
- **T-CalcVar** : *calculer la variance ou l'écart-type d'une variable aléatoire continue.*
Il s'agit d'un nouveau type de tâches pour les étudiants. Les ingrédients de la technique présentée par l'enseignant sont les suivants : calculer l'espérance de X et celle de X^2 , pour cela, et en fonction de comment est définie la fonction de densité, découpage, identification de la formule à utiliser selon l'intervalle, remplacer f par la bonne formule, calculer la (ou les) intégrale(s) puis simplifier les fractions.
Ce type de tâches est présenté pour la première fois durant le CM pour une loi quelconque, puis rapidement pour la loi uniforme (calcul non détaillé). T-CalcVar sera revu lors du TD dans l'exercice 4.
- **T-Det(a)** : *déterminer le réel a tel que $P(X < a) = \text{nombre donné (loi normale)}$.*
Ce type de tâches est présenté rapidement lors d'un exercice du TD. Les étudiants ont déjà rencontré ce type de tâches au lycée mais désormais il sera à traiter au moyen d'une table de loi normale et non plus avec une calculatrice : « *La table s'utilise dans les deux sens mais on a mis que les nombre positifs car on peut tout faire avec.* »
- **T-Det(cf)** : *déterminer les valeurs possibles d'une constante c qui apparaît dans l'expression de f , fonction de densité définie par morceaux.*
 - Il s'agit d'un nouveau type de tâches pour les étudiants qui sortent du lycée. L'enseignant demande de commencer par le graphe, il prend en charge l'organisation, la reconnaissance et le traitement de cette tâche. La sous-activité de traitement est en partie laissée à la charge des étudiants (calcul de l'intégrale). Les ingrédients de la technique pour une telle tâche sont les suivants : utiliser les deux conditions pour être une densité, puis faire les calculs associés.
 - D'après nos observations, l'enseignant considère que cette tâche ne pose pas de problème particulier aux étudiants car ce sont des calculs déjà faits en terminale. Pourtant les étudiants ont des difficultés avec la reconnaissance de la tâche (car nouveauté) mais aussi avec l'organisation (fonction définie par morceaux par une indicatrice) et le traitement de la tâche. De façon générale les calculs d'intégrales posent problème.

5.3.3 Éléments caractéristiques en termes de déroulement

Nous présentons dans ce qui suit des éléments caractéristiques indépendants des contenus enseignés pour ce cours de probabilités, en termes de contrat didactique ou de déroulement. Il s'agit de nouveautés vis-à-vis de la classe de terminale scientifique, ou alors d'éléments notables que nous avons relevés.

Premier point, désormais et contrairement à ce que nous avons pu observer dans le manuel (chapitre 4) et dans la classe de terminale scientifique, le cours est d'un seul bloc et sous forme de diapositives projetées au tableau (un cours magistral pour tout un thème suivi d'une séance de TD pour ce thème). Il n'y a plus cette alternance au sein d'une même séance, entre moments de cours et exercices en autonomie, il s'agit sans doute d'une nouveauté pour au moins certains étudiants.

De nombreuses informations sont données aux étudiants oralement mais non écrites au tableau.

- Pour le thème des probabilités conditionnelles on peut citer par exemple les informations suivantes données par l'enseignant : *une situation peut être modélisée de plusieurs façons, il est préférable de choisir un critère le plus général possible (pour Ω), l'équiprobabilité signifie la distribution uniforme des événements élémentaires.*
- Dans le cas des lois continues nous avons relevé par exemple des liens entre probabilité, densité et intégrale, les propriétés de la fonction de densité, les découpages à effectuer sur une intégrale.

L'enseignant détaille peu durant le CM les calculs de probabilités des exemples qui sont affichés sur les diapositives. Suite à nos observations on peut émettre plusieurs hypothèses quant à cela, premièrement l'enseignant considère que ce sont des calculs qui ont déjà été faits de nombreuses fois dans le secondaire (notamment pour les probabilités conditionnelles). Deuxièmement, l'enseignant attend des étudiants qu'ils fassent eux-mêmes ces calculs, notamment effectuer les calculs numériques à la main pour s'entraîner à réduire des fractions sans l'aide de la calculatrice. Troisièmement, l'enseignant choisit de laisser à la charge des étudiants ce qu'ils ont vu au premier semestre (techniques de calcul d'intégrales au semestre 1) et dans des cours antérieurs (calculs de probabilités dans le secondaire).

- ↪ Par exemple, pour le thème des probabilités conditionnelles, lors de la première rencontre avec T-EtudInd, l'enseignant ne détaille pas le calcul des probabilités des événements. En revanche, il décrit avec précision la modélisation de la situation, et notamment le choix du modèle et interprète les résultats.
- ↪ Pour le cas des lois continues, l'enseignant laisse de côté en CM les calculs d'intégrales. Or, comme nous l'avons vu dans le chapitre 4 mais aussi suite

aux observations de classe en terminale scientifique (partie 5.2.2), les techniques d'intégration ont été très peu vues et mises en fonctionnement par les élèves dans le secondaire pour les contenus de lois à densité. Il s'agit bien souvent de techniques, soit instrumentée, soit faisant appel à d'autres types de raisonnements (géométrique) que les étudiants ont pratiqués au lycée.

L'enseignant émet les recommandations suivantes quant aux attendus dans ce cours, pour le thème des probabilités conditionnelles : « *Pour ce chapitre, on suppose que les calculs et appliquer les formules vous savez le faire, vous avez vu ça au lycée. Ce qu'on vous demande en plus que l'année dernière, c'est la terminale "++", c'est de dire qui est Ω , P , A et B autrement que par une phrase, en me décrivant des ensembles.* »

Dans le chapitre 6 qui suit, nous allons nous intéresser plus en détails aux élèves et aux étudiants. Au travers de ce que nous venons de faire parties 5.2 et 5.3 nous allons essayer d'avoir des éléments de compréhension de l'impact de ces caractéristiques de l'enseignement sur les pratiques mathématiques des élèves et des étudiants.

Chapitre 6

Probabilités à la transition secondaire-supérieur : analyses relatives aux élèves et étudiants

Après avoir observé et analysé des enseignements de probabilités au lycée (partie 5.2) et à l'Université en première année de biologie (partie 5.3) ; et dans l'objectif de construire un MPR étendu adapté aux attentes et aux besoins des étudiants biologistes débutants à l'Université (MPR-ET) ; nous allons dans cette partie nous intéresser plus particulièrement aux élèves et aux étudiants et à leurs apprentissages pour les deux thèmes de probabilités qui nous intéressent.

Dans ce chapitre 6, nous présentons tout d'abord un test que nous avons proposé à des élèves de lycée et à des étudiants de première année de biologie. Nous exposons, dans une seconde partie, les résultats d'un questionnaire en ligne que nous avons conçu à destination d'étudiants biologistes débutants à l'Université. Dans une troisième partie nous décrivons et analysons des entretiens menés auprès de quatre étudiants de biologie. Enfin, dans une quatrième et dernière partie, nous présentons le sujet d'examen de probabilités proposé à ces étudiants et l'analyse de certaines copies d'examen. Ce travail va nous permettre d'apporter des éléments de réponse à la question de recherche n°3 : *Quelles sont les spécificités des étudiants non-spécialistes dans l'apprentissage des probabilités ? (le cas d'étudiants de première année de biologie)*

6.1 Test sur les probabilités conditionnelles

Nous avons élaboré un test sur le thème des probabilités conditionnelles et nous l'avons soumis à des élèves de lycée et à des étudiants de première année de biologie en décembre 2018 (voir tableau chronologique de recueil de données 3.1 page 47). Les élèves de lycée sont ceux de la classe de terminale scientifique dans laquelle nous avons fait nos observations et les étudiants en première année de biologie sont ceux de l'Université où se déroulent nos observations.

L'énoncé de ce test se trouve figure 6.1. Nous avons proposé ce test à tous les

Test

Exercice 1 :

Soient X et Y deux individus dont les durées de vie sont indépendantes et sont telles que $P(X \text{ vive encore 9 ans}) = \frac{2}{5}$ et $P(Y \text{ vive encore 9 ans}) = \frac{3}{5}$. Calculer les probabilités que :

1. X et Y vivent encore 9 ans ;
2. l'un des 2 au moins vive encore 9 ans ;
3. X seulement vive encore 9 ans ;
4. X vive encore 9 ans sachant que l'un des 2 au moins vivra encore 9 ans.

Exercice 2 :

Vous êtes directrice ou directeur du cabinet du ministre de la santé. Une maladie est présente dans la population, dans la proportion d'une personne malade sur 10000. Le responsable d'un grand laboratoire pharmaceutique vient vous vanter son nouveau test de dépistage : si une personne est malade, le test est positif à 99% ; si une personne n'est pas malade, le test est négatif à 98%. Autorisez-vous la commercialisation de ce test ?

Exercice 3 :

Le service marketing d'un magasin de téléphonie a procédé à une étude du comportement de sa clientèle. Il a ainsi observé que celle-ci est composée de 42% de femmes. 35% des femmes qui entrent dans le magasin y effectuent un achat alors que cette proportion est de 55% pour les hommes. Une personne entre dans le magasin. On note les événements :

- F : "la personne est une femme".
- R : "la personne repart sans rien acheter".

Dans tout l'exercice, donner des valeurs approchées des résultats au millième.

1. Construire un arbre pondéré illustrant la situation.
2. Calculer la probabilité que la personne qui est entrée dans le magasin soit une femme et qu'elle reparte sans rien acheter.
3. Montrer que $P(R) = 0,534$.
4. Les événements F et R sont-ils indépendants ?

FIGURE 6.1 – Énoncé du test

élèves (29) de la classe de terminale, à la fin de la cinquième et dernière séance de cours sur les probabilités conditionnelles. Les élèves ont eu une demi-heure pour faire le test.

Le test a également été proposé à des étudiants de première année de biologie à la fin du module d'analyse (MATH1) du premier semestre. Le test a donc été réalisé par ces étudiants avant qu'ils n'assistent au cours de probabilités du second semestre (MATH2) dont nous avons observé et analysé certaines des séances (partie 5.3 page 194). Vingt-cinq étudiants volontaires ont participé pour une durée similaire. Nous attirons l'attention du lecteur sur ce dernier point pour souligner qu'ici se trouve une des limites de cette étude. En effet, ces 25 étudiants volontaires ne sont pas forcément représentatifs des étudiants de biologie en général.

Dans ce qui suit nous commençons par présenter nos choix pour l'élaboration

de ce test puis dans un second temps l'analyse *a priori* des trois exercices de ce test en regard des analyses et résultats présentés pour l'enseignement des probabilités au lycée (chapitre 4 et chapitre 5, partie 5.2). Dans une troisième partie nous présentons les grilles d'analyse des productions d'élèves et d'étudiants. Enfin, dans une quatrième et dernière partie, nous analysons les copies recueillies au moyen de ces grilles puis nous présentons les conclusions issues de ces analyses.

6.1.1 Présentation du test

Le test que nous proposons se trouve figure 6.1, il consiste en trois exercices sur le thème des probabilités conditionnelles.

Nous avons choisi l'exercice 1 car il est issu de la feuille de TD du module de probabilités de l'Université (l'exercice 3 est proposé dans le TD2 voir figure 5.59 page 210) mais réalisable avec des techniques vues au lycée. Les étudiants qui répondent à ce test n'ont, *a priori*, jamais rencontré cet exercice car le test leur a été proposé avant le début du module de probabilités.

L'exercice 2 est lui aussi issu d'une feuille de TD (d'une année antérieure) de ce même module de probabilités. Nous avons choisi cet exercice car il nécessite un important travail de modélisation probabiliste et parce qu'il est lié à un contexte de biologie.

Enfin, le troisième exercice est issu d'un manuel de terminale scientifique (Barbazo et al., 2016). Nous l'avons choisi car, tout en étant classique, il propose un travail de modélisation probabiliste à la charge de l'élève.

6.1.2 Analyse a priori

Nous présentons dans cette partie une analyse *a priori* de ce test en nous référant aux types de tâches, variations et techniques, observés ou attendus dans le secondaire d'après les chapitres 4 et 5 (partie 5.2). Nous complétons cette analyse *a priori* en présentant ces exercices du point de vue de l'élève, c'est-à-dire, en évoquant la difficulté de chaque type de tâches (ou technique) selon le contexte précis de l'enseignement reçu. S'agit-il de connaissances mobilisables, disponibles ou en cours d'acquisition ? Quelles sont les adaptations à la charge de l'élève ? S'agit-il d'applications immédiates ?

Ce procédé a des limites car les étudiants de l'Université à qui nous avons proposé le test proviennent pour la plupart de terminale scientifique mais n'ont pas assisté, *a priori* en terminale, aux séances de cours que nous analysons dans la partie 5.2.

Exercice 1

Comme nous l'avons dit ci-dessus, l'exercice 1 est très similaire à l'exercice 3 proposé dans le TD2 (voir partie 5.3.1), les valeurs numériques et la description des événements sont les seuls changements. Cependant les étudiants qui vont faire ce test n'ont, pour rappel, jamais rencontré cet exercice 3 car ils ont fait le test avant de commencer le module de probabilités du second semestre.

Dans cet exercice il s'agit d'effectuer une modélisation probabiliste à partir d'un énoncé partiellement en langage naturel et de calculer une probabilité (conditionnelle, d'une intersection ou d'une union de deux événements).

La première question propose le type de tâches **T-CalcInter**, calculer la probabilité d'une intersection, avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les événements sont indépendants
- sans que les événements aient été explicitement nommés dans le texte
- étant donné que c'est la première de l'exercice
- les données sont en fréquences relatives

Les ingrédients d'une technique, qui représentent ici les actions successives pour effectuer la tâche, sont : identifier les événements (par exemple on note A l'événement : " X vive encore 9 ans", et B l'événement : " Y vive encore 9 ans"), interpréter l'énoncé de l'exercice pour en déduire que les événements A et B sont indépendants, interpréter le "et" comme l'intersection des événements A et B , en utilisant l'indépendance des deux événements, calculer $P(A \cap B)$ comme le produit de $P(A)$ et de $P(B)$, ce qui donne :

$$P(X \text{ et } Y \text{ vivent encore 9 ans}) = P(A \cap B) = P(A) \times P(B) = \frac{2}{5} \times \frac{3}{5} = \frac{6}{25}$$

D'après les observations de la partie 5.2, identifier les événements A et B est une connaissance disponible pour les élèves du lycée qui ont suivi un enseignement sur les probabilités conditionnelles. Interpréter le "et" comme l'intersection des événements A et B est une connaissance en cours d'acquisition déjà mise en fonctionnement directement dans les exercices. Il s'agit ici d'une application immédiate qui ne demande pas d'adaptation.

La seconde question propose le type de tâches **T-CalcUnion**, calculer la probabilité d'une union d'événements non disjoints, avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- sans que les événements aient été explicitement nommés dans le texte
- les données sont en fréquences relatives

Les ingrédients d'une technique pour effectuer la tâche sont : interpréter "l'un des deux" comme l'union des événements A et B , utiliser la formule de Poincaré

pour calculer $P(A \cup B)$, à partir de la question précédente qui donne $P(A \cap B)$ déduire le résultat numérique :

$$\begin{aligned} P(\text{l'un des deux au moins vive encore 9 ans}) &= P(A \cup B) \\ &= P(A) + P(B) - P(A \cap B) \\ &= \frac{2}{5} + \frac{3}{5} - \frac{6}{25} = \frac{19}{25} \end{aligned}$$

D'après nos analyses de la partie 5.2 interpréter l'énoncé de la question comme l'union des événements A et B n'a jamais été rencontré directement dans des exercices. Utiliser la formule de Poincaré est une connaissance en cours d'acquisition qui a été rencontrée en cours mais qui n'a pas été mise en fonctionnement directement dans des exercices. Cette tâche demande un peu d'adaptation aux élèves.

La troisième question propose le type de tâches **T-CalcInter**, calculer la probabilité d'une intersection d'évènements, avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- sans que les évènements aient été explicitement nommés dans le texte
- les données sont en fréquences relatives

Les ingrédients d'une technique sont : interpréter le mot "*seulement*" comme la réalisation des événements " X vive encore 9 ans" et " Y ne vive pas 9 ans", identifier qu'il s'agit de l'intersection des événements A et \bar{B} , utiliser l'indépendance de A et B pour déduire l'indépendance de A et \bar{B} , par indépendance des événements, calculer l'intersection $P(A \cap \bar{B})$ comme le produit de $P(A)$ et de $P(\bar{B})$. Ce qui donne :

$$\begin{aligned} P(X \text{ seulement vive encore 9 ans}) &= P(X \text{ vive encore 9 ans et } Y \text{ ne vive pas 9 ans}) \\ &= P(A \cap \bar{B}) = P(A) \times P(\bar{B}) \\ &= P(A) \times (1 - P(B)) \\ &= \frac{2}{5} \times \frac{2}{5} = \frac{4}{25} \end{aligned}$$

Interpréter l'énoncé comme la réalisation des événements A et \bar{B} , c'est-à-dire l'intersection de ces évènements, est une connaissance en cours d'acquisition déjà mise en fonctionnement directement dans des exercices. Si deux événements A et B de Ω sont indépendants, alors A et \bar{B} sont indépendants, est également une connaissance en cours d'acquisition déjà mise en fonctionnement directement dans des exercices. Enfin, calculer la probabilité d'un évènement connaissant la probabilité de son évènement contraire (**T-CalcBar**) est une connaissance acquise déjà mise en fonctionnement directement dans des exercices.

La quatrième question propose le type de tâches **T-CalcProbaCond**, calculer une probabilité conditionnelle, avec les variations suivantes :

- à partir d'un énoncé en langage naturel

- sans que les évènements aient été explicitement nommés dans le texte
- les données sont en fréquences relatives

Les ingrédients d'une technique sont les suivants : interpréter l'énoncé pour identifier l'événement "*l'un des deux au moins vivra encore 9 ans*" comme étant $A \cup B$, interpréter "*sachant que*" comme la réalisation de l'événement A sachant que l'événement $A \cup B$ est réalisé, c'est-à-dire la probabilité conditionnelle de A sachant $A \cup B$, utiliser la définition d'une probabilité conditionnelle, calculer le numérateur **T-CalcInter** et remplacer le dénominateur par la valeur obtenue question (2). Ce qui donne :

$$\begin{aligned} P(X \text{ vive encore 9 ans sachant que l'un des deux au moins vivra encore 9 ans}) &= P_{A \cup B}(A) \\ &= \frac{P(A \cap (A \cup B))}{P(A \cup B)} \\ &= \frac{P(A)}{P(A \cup B)} \\ &= \frac{2}{5} \times \frac{25}{19} = \frac{10}{19} \end{aligned}$$

Interpréter "*sachant que*" comme la réalisation de l'événement A sachant que l'événement $A \cup B$ est réalisé et utiliser la définition de la probabilité conditionnelle sont des connaissances en cours d'acquisition déjà mises en fonctionnement directement dans des exercices. En revanche, le fait que l'évènement qui conditionne soit une union de deux évènements n'est pas quelque chose qui a été rencontré dans la partie 5.2.

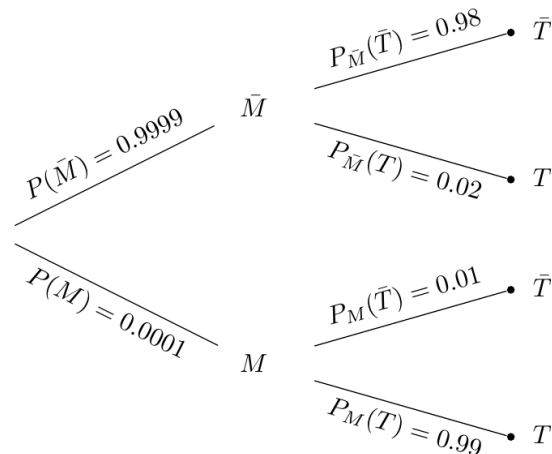
Exercice 2

L'exercice 2 demande un travail de modélisation préalable à la résolution de la tâche demandée. Les principaux types de tâches que nous avons identifiés pour cet exercice sont, premièrement, d'effectuer une modélisation probabiliste à partir d'un énoncé en langage naturel (**T-Modé**), deuxièmement, de calculer une probabilité conditionnelle (**T-CalcProbaCond**) et enfin, d'interpréter le résultat numérique afin de répondre à la question en langage naturel (**T-Interp** sans que les évènements ne soient nommés ou décrits).

Les ingrédients d'une technique pour **T-Modé** sont les suivants :

- identifier les évènements probabilistes : on peut noter M l'événement : "la personne est malade", et T l'événement : "le test est positif"
- associer leurs probabilités à chacun des évènements (T-Assoc) :
 $P(M) = \frac{1}{10000}$, $P_M(T) = 0,99$ et $P_{\bar{M}}(\bar{T}) = 0,98$
- identifier les évènements contraires et calculer leurs probabilités (T-CalcBar) :
 $P(\bar{M}) = 0,9999$, $P_M(\bar{T}) = 0,01$ et $P_{\bar{M}}(T) = 0,02$

Tous ces ingrédients entrent dans le cadre d'une activité de modélisation probabiliste telle que définie dans la partie 5.2.3. La technique de représentation de la situation par un arbre de probabilités est une technique attendue dans le secondaire mais qui ne fait plus partie de la praxéologie à l'Université (partie 5.3). Voici une représentation par un arbre que l'on pourrait attendre ici :



On peut associer au deuxième type de tâches **T-CalcProbaCond** (*calculer la probabilité d'être malade en sachant que le test est positif*), deux sous-types de tâches qui sont les suivants : interpréter la question en termes de probabilité et calculer la probabilité conditionnelle correspondante. Chacun de ces deux types de tâches sont réalisés par des techniques que l'on peut décrire par un ensemble d'ingrédients : utiliser la définition d'une probabilité conditionnelle, détailler le dénominateur selon la formule des probabilités totales, à partir de l'arbre déterminer les valeurs des probabilités d'intersections présentes dans la formule.

$$\begin{aligned}
 P_T(M) &= \frac{P(M \cap T)}{P(T)} \\
 &= \frac{P(M \cap T)}{P(M \cap T) + P(\bar{M} \cap T)} \quad (\text{d'après la formule des probabilités totales}) \\
 &= \frac{0.000099}{0.000099 + 0.019998} \\
 &\simeq 0.0049
 \end{aligned}$$

Enfin, il faut interpréter le résultat obtenu pour répondre au problème posé, une formulation possible est par exemple : *il n'y a que 0,49% de chances qu'une personne positive au test soit effectivement malade, le test n'est donc pas efficace.*

La tâche est ici complexe. En effet, dans cet exercice, la reconnaissance des sous-types de tâches (ou ingrédients d'une technique) et l'organisation de leurs raisonnements (succession des sous-types de tâches puis traitement) sont de la responsabilité de l'étudiant. La complexité de la tâche a un impact sur


l'organisation praxéologique et en particulier sur la complexité de la technique (et des ingrédients de cette technique) à mettre en oeuvre.

Voici quelques éléments technologiques qui justifient le choix de calculer $P_T(M)$. On cherche ici à savoir si le test est efficace du point de vue d'un soignant. Un patient vient faire un test pour savoir s'il est malade ou non. Si son test est positif (respectivement négatif), il est important de savoir s'il est vraiment malade (respectivement pas malade), c'est-à-dire si le test est fiable. Le test est fiable si la probabilité, sachant que le test est positif, qu'une personne soit effectivement malade, est proche de 1. Il s'agit donc de calculer $P_T(M)$, pour cela on peut utiliser la formule de Bayes. Ces éléments technologiques vont au-delà des mathématiques, le fait de calculer $P_T(M)$ est entièrement à la charge des étudiants et repose sur des éléments technologiques liés à un contexte socio-médical.

Le contexte de l'exercice, qui est l'étude de l'efficacité des tests sur des populations malades et non malades, est plutôt un contexte familier aux étudiants d'après l'analyse du manuel de lycée réalisée chapitre 4. Ici ils doivent interpréter la question pour identifier l'évènement dont ils doivent calculer la probabilité puis interpréter le résultat obtenu pour répondre à la question initiale. Dans un exercice de terminale, il y aurait une première question "calculer la probabilité d'être malade sachant que le test est positif" et une deuxième question "interpréter le résultat", comme dans l'exercice suivant (figure 6.2) issu du manuel indice.

80 PROBLÈME DE SYNTHÈSE

On étudie une maladie dans la population d'un pays. Pour dépister cette maladie chez une personne, 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 $45 \text{ ng} \cdot \text{mL}^{-1}$.



Une personne étant choisie au hasard dans la population, on appelle les événements :

- M « le patient est atteint par la maladie étudiée » ;
- D « 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. Déterminer la probabilité qu'un patient ait un dépistage positif.
2. Calculer $P_D(M)$. 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 ?

FIGURE 6.2 – Exemple d'exercice issu du manuel Indice

D'après nos analyses en classe de terminale (partie 5.2), les techniques à utiliser ici par les élèves sont toutes basées sur des connaissances acquises ou en cours d'acquisition. En revanche, reconnaître cette structure praxéologique complexe et s'organiser en conséquence est entièrement à la charge de l'élève car il n'y a pas de sous-question intermédiaire associée à chacun des sous-types de tâches décrits ci-dessus. La très forte activité de modélisation laissée à l'élève n'est pas quelque chose d'habituel d'après l'analyse de manuel (chapitre 4) et de séances que nous avons pu réaliser .

Exercice 3

L'exercice 3 est plus classique pour les élèves issus de terminale scientifique. En effet il provient d'un manuel de terminale scientifique (Barbazo et al., 2016). Les techniques à utiliser par les élèves sont toutes basées sur des connaissances acquises ou en cours d'acquisition.

La première question relève du type de tâches **T-ModéArb**, modéliser une situation probabiliste décrite en langage naturel par un arbre de probabilités, avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que c'est la première question de l'exercice
- les données sont en pourcentages

Les ingrédients d'une technique sont les suivants :

- associer les probabilités de l'énoncé à chacun des événements **T-Assoc** : $P(F) = 0,42$, $P_F(\bar{R}) = 0,35$ et $P_{\bar{F}}(\bar{R}) = 0,55$
- identifier les événements contraires et calculer leurs probabilités **T-CalcBar** : $P(\bar{F}) = 0,58$, $P_F(R) = 0,65$ et $P_{\bar{F}}(R) = 0,45$
- tracer un arbre de probabilités : d'un point placé arbitrairement faire partir deux branches, placer les événements F et \bar{F} au bout de ces deux premières branches, faire partir deux branches de chacun des noeuds F et \bar{F} , au bout de ces branches placer les événements R et \bar{R}
- compléter l'arbre à l'aide des trois règles de construction de l'arbre : inscrire sur la branche allant du premier noeud au sommet F (respectivement \bar{F}) $P(F)$ (respectivement $P(\bar{F})$), inscrire sur la branche allant de F à R (respectivement \bar{R}) $P_F(R)$ (respectivement $P_F(\bar{R})$).

La seconde question question relève du type de tâches **T-CalcInter**, calculer la probabilité d'une intersection d'évènements, avec les variations suivantes :

- à partir d'un énoncé en langage naturel
- étant donné que les évènements ont déjà été décrits et nommés dans le texte

- étant donné que la situation a déjà été représentée par un arbre d'après une précédente question

Les ingrédients d'une technique sont les suivants : interpréter l'énoncé pour identifier que l'événement "*la personne est une femme et elle part sans rien acheter*" est $F \cap R$ et utiliser l'arbre de probabilités avec la règle du chemin. Ce qui donne : $P(F \cap R) = P(F) \times P_F(R) = 0.273$.

La troisième question question relève du type de tâches **T-CalcSimple**, calculer la probabilité d'un évènement simple, avec les variations suivantes :

- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que la situation a déjà été représentée par un arbre d'après une précédente question

Les ingrédients d'une technique sont les suivants : utiliser l'arbre et la formule des probabilités totales. Ce qui donne :

$$P(R) = P(F \cap R) + P(\overline{F}) \times P_{\overline{F}}(R) = 0.273 + 0.261 = 0.534$$

La quatrième question question relève du type de tâches **T-EtudInd** : étudier l'indépendance de deux évènements, avec les variations suivantes :

- étant donné que les évènements ont déjà été décrits et nommés dans le texte
- étant donné que la situation a déjà été représentée par un arbre
- étant donné que les probabilités ont déjà toutes été calculées

Les ingrédients d'une technique sont les suivants : calculer le produit $P(F) \times P(R) = 0,224$ et comparer avec la probabilité de l'intersection $P(F \cap R)$ calculée à la question 1. Ce qui donne : *Les deux résultats sont différents donc les évènements F et R ne sont pas indépendants.*

D'après nos observations en classe de terminale scientifique (partie 5.2), construire un arbre pondéré à partir d'un énoncé en langage naturel, interpréter le "*et*" comme l'intersection de deux évènements, calculer la probabilité d'une intersection de deux évènements en utilisant un arbre, calculer la probabilité d'un évènement en utilisant la formule des probabilités totales et déterminer si deux évènements sont indépendants ou non sont des connaissances en cours d'acquisition déjà mises en fonctionnement directement dans les exercices et exemples.

6.1.3 Méthodologie de l'analyse des productions d'élèves et d'étudiants

A partir de l'analyse *a priori* du test nous avons conçu des grilles d'étude des productions d'élèves et d'étudiants. Ces grilles, à retrouver en annexe F et dont un extrait est proposé figure 6.3, ont été adaptées durant toute l'analyse des copies, notamment pour référencer les erreurs faites par les élèves et étudiants.

Au moyen de ces grilles, nous avons effectué une analyse quantitative de l'ensemble des productions des élèves et étudiants, ce qui nous a permis de mettre en évidence des premiers résultats concernant à la fois les difficultés pour ce thème mais aussi la modélisation mathématique en probabilité à la transition secondaire-supérieur. Nous présentons dans la partie suivante l'analyse des copies d'élèves et d'étudiants.

Exercice 1
la situation est représentée par un arbre
Question1 : T-CalcInter
la réponse est correcte justification par l'indépendance des événements compréhension : la probabilité demandée est une intersection d'évènements erreur : l'énoncé n'est pas interprété comme une intersection d'évènements notations inhabituelles
Question2 : T-CalcUnion
la réponse est correcte la réponse est incomplète compréhension : la probabilité demandée est celle d'une union de 2 événements autre technique : somme des probas de toutes les issues possibles technique : utilisation correcte de la formule de Poincaré e1 : erreur sur la formule de Poincaré e2 : l'énoncé n'est pas interprété comme une union d'évènements e3 : addition de $P(A)$ et $P(B)$ e4 : erreur de calcul

FIGURE 6.3 – Extrait des grilles d'analyse du test

6.1.4 Analyse a posteriori

Dans cette sous-partie nous exposons l'analyse des productions recueillies au moyen des grilles présentées dans la sous-partie précédente. Il s'agit d'une analyse quantitative illustrée par des extraits de copies d'élèves ou d'étudiants. Ces grilles d'analyse sont à retrouver en annexe F.

Le test a été traité par 29 élèves de terminale scientifique et par 25 étudiants en première année de biologie à l'Université.

Exercice 1

L'exercice 1 a été abordé par tous les élèves et étudiants. Ils sont quatre à faire un arbre de probabilités pour cet exercice, qui ne s'y prête pas, et deux étudiants proposent une représentation avec un diagramme de Venn. Voici un extrait de

copie dans la figure 6.4.

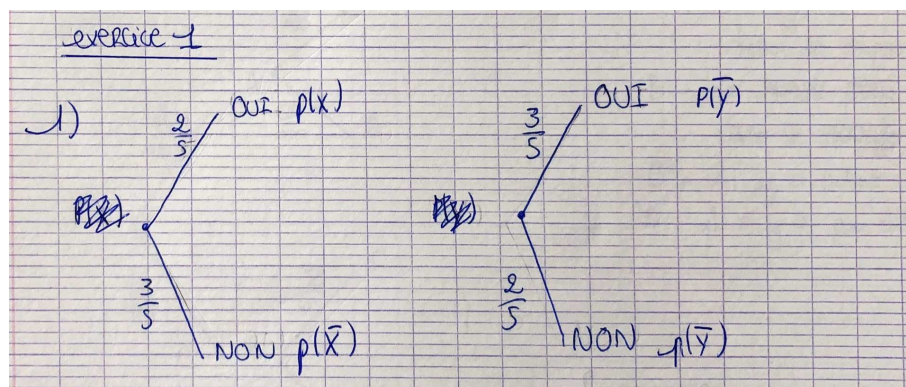


FIGURE 6.4 – Représentation de la situation probabiliste de l'exercice 1 par deux arbres – copie d'élève

Question n°1 : T-CalcInter	élèves	étudiants	TOTAL
la réponse est correcte	29	22	51
identification de l'intersection	25	16	41
justification par l'indépendance	20	1	21

FIGURE 6.5 – Extrait n°1 de la grille d'analyse de l'exercice 1

Tous les élèves de terminale et 22 étudiants en biologie ont donné la bonne valeur pour la première question (voir grille d'analyse figure 6.5), à savoir, le produit des deux valeurs de l'énoncé qui est $\frac{6}{25}$. Il n'est pas clair que tous aient bien identifié qu'il s'agit d'une intersection d'évènements car parmi eux, 25 élèves de terminale et 16 étudiants écrivent distinctement qu'ils cherchent la probabilité d'une intersection. Ils ne sont que 20 élèves et 1 étudiant à justifier le produit des probabilités par l'indépendance des évènements.

Nous ne relevons plus que 26 bonnes réponses à la seconde question (21 élèves et 5 étudiants) (voir grille d'analyse figure 6.6).

Ils sont 34 (24 élèves et 10 étudiants) à avoir identifié que l'évènement recherché question n°2 est une union de deux évènements. Tous les élèves de terminale utilisent la technique de la formule de Poincaré pour calculer $P(A \cup B)$. En revanche à l'Université certains proposent le découpage suivant (voir extrait de copie dans la figure 6.7) :

$$P(A \cup B) = P(A \cap B) + P(A \cap \bar{B}) + P(\bar{A} \cap B)$$

6.1. Test sur les probabilités conditionnelles

Question n°2 : T-CalculUnion	élèves	étudiants	TOTAL
la réponse est correcte	21	5	26
identification d'une union	24	10	34
technique : formule de Poincaré	24	3	27
autre technique	0	2	2

FIGURE 6.6 – Extrait n°2 de la grille d'analyse de l'exercice 1

Exercice 1:

1) $p(X \text{ et } Y \text{ vivant}) = \frac{2}{5} \times \frac{3}{5} = \frac{6}{25}$

2) $p = p(X_v \cap Y_v) + p(X_v \cap Y_{\bar{v}}) + p(X_{\bar{v}} \cap Y_v)$

$$= \frac{6}{25} + \frac{2}{5} \times \frac{2}{5} + \frac{3}{5} \times \frac{3}{5} = \frac{6}{25} + \frac{4}{25} + \frac{9}{25}$$

$$= \frac{19}{25}$$

FIGURE 6.7 – Réponse à la question n°2 par une technique autre que la formule de Poincaré – copie d'étudiant

Les principales erreurs que nous avons relevées pour cette seconde question sont les suivantes :

- l'énoncé n'est pas interprété comme une union
- les deux nombres de l'énoncé sont simplement additionnés ($P(A) + P(B)$)
- erreur dans la formule de Poincaré.

Question n°3 : T-CalculInter	élèves	étudiants	TOTAL
la réponse est correcte	9	5	14
identification que le probabilité recherchée fait intervenir l'évènement B	13	8	21
identification d'une intersection	6	6	12
identification de la bonne intersection	5	5	10
justification par l'indépendance	4	1	5
erreur : répondre $P(A)$	14	13	27

FIGURE 6.8 – Extrait n°3 de la grille d'analyse de l'exercice 1

Ils ne sont que 14 (9 élèves et 5 étudiants) à avoir répondu correctement à la troisième question : $P(A \cap \bar{B}) = \frac{4}{25}$ (voir grille d'analyse figure 6.8). Parmi eux 10 (5 élèves et 5 étudiants) ont clairement explicité qu'il s'agit d'une intersection d'évènements et ils ne sont plus que 5 (4 élèves et 1 étudiants) à avoir justifié le produit des probabilités par l'indépendance des évènements.

Parmi ceux qui n'ont pas abouti (38), ils sont 7 (4 élèves et 3 étudiants) à avoir identifié que la probabilité recherchée fait intervenir l'évènement B et seulement 2 à avoir identifié une intersection d'évènements. L'erreur la plus courante est que l'énoncé n'a pas été identifié comme une intersection d'évènements, en effet, ils sont 27 (14 élèves et 13 étudiants) à avoir répondu $P(A)$ (valeur donnée de l'énoncé).

Question n°4 : T-CalcProbaCond	élèves	étudiants	TOTAL
la réponse est correcte	0	0	0
identification d'une probabilité conditionnelle	19	13	32
identification que la condition est $A \cup B$	12	6	18
identification de la bonne intersection	5	5	10
erreur : évènements mal identifiés	7	6	13

FIGURE 6.9 – Extrait n°4 de la grille d'analyse de l'exercice 1

Aucun élève ou étudiant n'a répondu correctement à la dernière question (réponses incomplètes ou fausses) (voir grille d'analyse figure 6.9). Ils sont 32 (19 élèves et 13 étudiants) à avoir identifié qu'il s'agit bien d'une probabilité conditionnelle (voir extrait de copie ci-dessous figure 6.10), et parmi eux 18 (12 élèves et 6 étudiants) ont identifié que l'évènement qui conditionne est $A \cup B$.

4. $P(X \text{ arrive encore 9 ans}) = \frac{P(X \text{ arrive encore 9 ans}) \times P(X \cap Y)}{P(A \text{ des 2 arriva encore 9 ans})}$

$P\left(\frac{A}{B}\right) = \frac{P(A) + P(A \cap B)}{P(B)}$

FIGURE 6.10 – Réponse à la question n°4 de l'exercice 1 – copie d'étudiant

Nous avons relevé les erreurs suivantes pour cette quatrième question :

- les évènements sont mal identifiés (7 élèves et 6 étudiants)

- erreurs dans la formule qui définit une probabilité conditionnelle (5 élèves et 5 étudiants) (voir extrait de copie ci-dessus)
- l'énoncé n'est pas interprété comme une probabilité conditionnelle (2 élèves et 1 étudiants)
- l'utilisation à tort de l'indépendance de A et $A \cup B$ pour le calcul du numérateur : $P(A \cap (A \cup B))$ (7 élèves)

Exercice 2

L'exercice 2 a été abordé par 49 élèves et étudiants ayant répondu au test (soit 91%) (voir grille d'analyse figure 6.11).

Les élèves de terminale et les étudiants en biologie ont rencontré des difficultés et ont dans l'ensemble proposé des solutions fausses (une seule réponse correcte donnée par une étudiante qui sort de deux années de PACES). D'après l'analyse *a priori* ci-dessus nous nous attendions à de tels résultats au vu de la difficulté que représente l'étape de modélisation ici. Ce qui nous intéresse plus particulièrement c'est de savoir à quelles étapes les répondants ont eu des difficultés, quelles étapes de la résolution prennent-ils en charge et de quelle façon.

La diversité des réponses est grande, la probabilité la plus souvent calculée par les élèves de terminale est $P(M \cap T)$ (5 réponses); il n'y a pas de réponse dominante parmi les étudiants en biologie.

Exercice 2	élèves	étudiants	TOTAL
l'exercice est traité	27	22	49
la réponse est correcte	0	1	1
identification des événements en jeu	18	14	32
T-Assoc et T-CalcBar correctement accomplis	13	13	26
représentation de la situation par un arbre	13	13	26

FIGURE 6.11 – Extrait n°1 de la grille d'analyse de l'exercice 2

Même si les élèves et étudiants n'ont pas répondu correctement, beaucoup d'entre eux ont correctement identifié les événements en jeu (32 d'entre eux, soit environ 65%). Ils sont un peu moins nombreux à représenter la situation par un arbre de probabilités (26 d'entre eux, soit environ 53%) comme on peut le voir dans l'extrait de copie suivant 6.12.

Ce qui semble avoir posé des difficultés aux élèves et étudiants c'est une fois la situation représentée par un arbre, quoi calculer pour répondre à la question

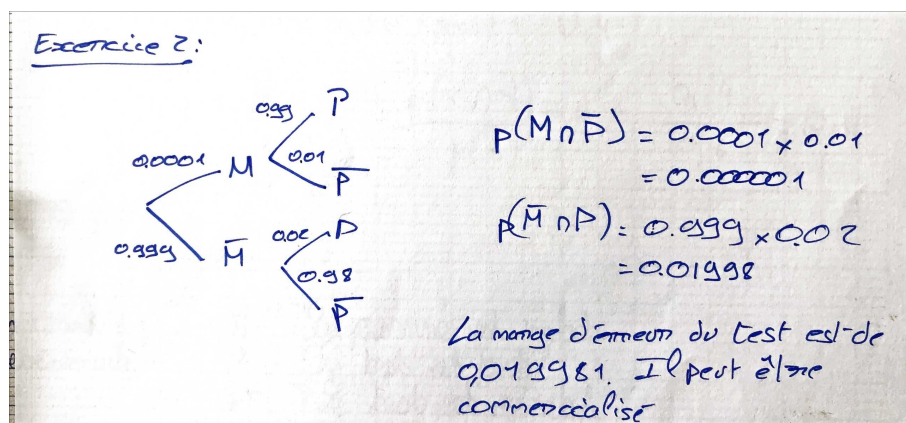


FIGURE 6.12 – Représentation de la situation probabiliste de l'exercice 2 par un arbre – copie d'étudiant

posée, voici un exemple de copie d'élève dans la figure 6.13.

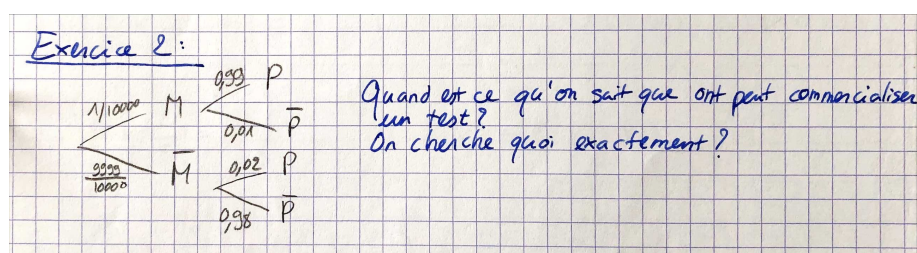


FIGURE 6.13 – Réponse d'un étudiant à l'exercice 2 après avoir représenté la situation par un arbre

En revanche, tous les élèves et étudiants qui ont choisi de représenter la situation par un arbre ont correctement associé les valeurs numériques de l'énoncé aux événements correspondants (T-Assoc) et ont correctement calculé les probabilités des événements complémentaires (T-CalcBar).

Etant donné que nous considérons que modéliser correctement la situation signifie identifier les événements en jeu et leur associer leurs probabilités, dans cet exercice, seule la moitié des étudiants ayant répondu ont fait une modélisation correcte de la situation.

Concernant l'interprétation de la question et la réponse donnée par les étudiants et élèves, 41 d'entre eux (soit environ 84%) ont formulé une réponse en langage naturel. Parmi eux, 19 se sont appuyés sur leurs calculs de probabilités antérieurs pour répondre. En revanche, 8 élèves et étudiants ont répondu à la question en langage naturel en se basant uniquement sur leur représentation de la situation par l'arbre de probabilités (pas de calcul).

Nous notons que 16 d'entre eux (environ 33%), 10 élèves de terminale et 6 étudiants en biologie, ont répondu à la question sans avoir préalablement effectué

6.1. Test sur les probabilités conditionnelles

Exercice 2	élèves	étudiants	TOTAL
réponse appuyée par un calcul	11	8	19
réponse appuyée uniquement par un arbre	4	4	8
réponse appuyée ni par un calcul ni par un arbre	10	6	16

FIGURE 6.14 – Extrait n°2 de la grille d'analyse de l'exercice 3

ni calcul de probabilité ni modélisation probabiliste (comme une représentation par un arbre de probabilités) (voir grille figure 6.14). Ces élèves et étudiants n'ont pas été en mesure d'identifier dans la tâche qui leur était prescrite les différents types de tâches à réaliser. Voici deux exemples d'une telle réponse figure 6.15.

"Non, car la marge d'erreur est énorme pour une population de 10000 personnes. Sur 10000, il peut y avoir 200 personnes qui sont déclarées comme malades alors qu'elles ne le sont pas du tout. Ceci est lié aux 98%. Si la maladie concernait 1 personne sur 100, cela aurait été plus intéressant."

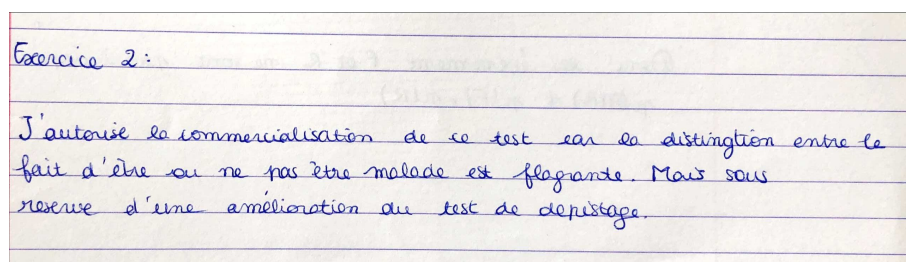


FIGURE 6.15 – Réponses à l'exercice 2 sans calcul – copies d'élèves

En ce qui concerne l'activité de modélisation probabiliste, nous relevons que les élèves et étudiants utilisent ici des arbres de probabilités alors qu'il n'est indiqué nulle part dans l'énoncé qu'un arbre pourrait être utilisé pour répondre. Il s'agit d'une observation importante qui est liée à une problématique de transition secondaire-supérieur. En effet, depuis le lycée, les élèves ont pris l'habitude d'utiliser ce type de représentations (voir partie 5.2). La construction et l'utilisation d'un arbre de probabilités sont des compétences largement développées dans les manuels et dans le programme officiel de mathématiques de la classe de terminale (partie 4). Pour rappel, le programme officiel de mathématiques de terminale indique que : "un arbre de probabilités correctement construit est une preuve", ce qui n'est plus concevable à l'Université (voir partie 5.3).

Exercice 3

L'exercice 3 a été abordé par 26 élèves et par tous les étudiants (soit 94,4% des répondants au test).

Question n°1 : T-ModéArb	élèves	étudiants	TOTAL
l'arbre est complet et correct	18	24	42
erreur : T-Assoc	6	0	6
erreur : T-CalcBar	1	0	1

FIGURE 6.16 – Extrait n°1 de la grille d'analyse de l'exercice 3

Tous les élèves (26) et étudiants (24) à avoir traité cet exercice ont représenté la situation par un arbre de probabilités à la première question (voir grille d'analyse figure 6.16) comme il était demandé. Mais seulement 18 des 26 élèves de terminale ont produit un arbre complet et correct. Les erreurs relevées sont des difficultés à interpréter correctement les données numériques de l'énoncé : confusion entre R et \bar{R} ou des erreurs dans le calcul de la probabilité du complémentaire.

Question n°2 : T-CalcInter	élèves	étudiants	TOTAL
la réponse est correcte	17	20	37
identification de la bonne intersection	19	20	39
justification par la formule $P(F \cap R) = P(F) \times P_F(R)$	20	3	23
erreur : identification de l'évènement	1	4	5

FIGURE 6.17 – Extrait n°2 de la grille d'analyse de l'exercice 3

Nous relevons 37 bonnes réponses à la seconde question (17 élèves et 20 étudiants) (voir grille d'analyse figure 6.17). Tous ont bien écrit qu'il s'agit de la probabilité de $F \cap R$. Parmi eux, 15 élèves mais seulement 3 étudiants justifient le résultat en écrivant :

$$P(F \cap R) = P(F) \times P_F(R)$$

Les erreurs que nous avons relevées dans les autres copies sont les suivantes : utilisation à tort de l'indépendance de F et R , , évènements mal identifiés dès le départ (question n°1), l'énoncé n'est pas interprété comme une intersection (1 élève et 4 étudiants) ou confusion avec $P_F(R)$.

Question n°3 : T-CalcSimple	élèves	étudiants	TOTAL
la réponse est correcte	14	20	34
la réponse est incomplète	3	0	3
justification par la formule des probabilités totales	17	16	33
erreur : formule des probabilités totales	1	1	2
erreur : utilisation à tort de l'indépendance de R et \bar{F}	0	2	2
erreur à la Q1 ou à la Q2	4	0	4

FIGURE 6.18 – Extrait n°3 de la grille d'analyse de l'exercice 3

Nous relevons 34 bonnes réponses à la troisième question (14 élèves et 20 étudiants), et 3 réponses incomplètes (voir grille d'analyse figure 6.18).

Tous les élèves de terminale et 16 étudiants justifient correctement en utilisant la formule des probabilités totales, soit ils écrivent "d'après la formule des probabilités totales", soit ils écrivent $P(R) = P(F \cap R) + P(\bar{F} \cap R)$.

Quelques répondants se trompent dans l'utilisation de la formule des probabilités totales ou écrivent $P(\bar{F} \cap R) = P(\bar{F}) \times P(R)$ pensant à une indépendance.

Question n°4 : T-EtudInd	élèves	étudiants	TOTAL
la question est traitée	15	20	35
la réponse est correcte	12	11	23
la réponse est correctement justifiée	12	1	13

FIGURE 6.19 – Extrait n°4 de la grille d'analyse de l'exercice 3

Ils ne sont plus que 35 à avoir répondu à la dernière question (15 élèves et 20 étudiants), contre 51 (26 élèves et 25 étudiants) au début de l'exercice (voir grille d'analyse figure 6.19). Parmi eux, 23 réponses sont correctes (12 élèves et 11 étudiants) et seules 13 (12 élèves et 1 étudiant) sont correctement justifiées, c'est-à-dire par le calcul adéquat : $P(F) \times P(R) = P(F \cap R)$.

Dans la figure suivante 6.20 on retrouve un extrait d'une telle copie.

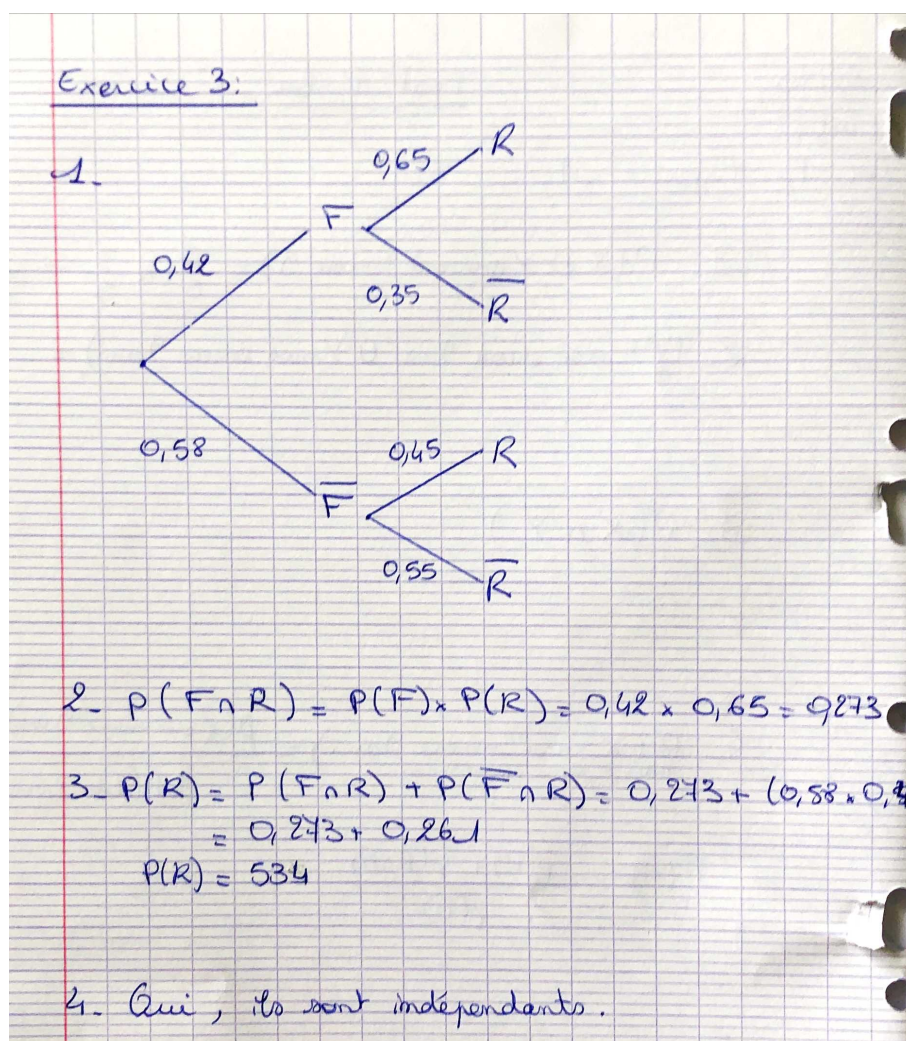


FIGURE 6.20 – Copie d'étudiant avec à la question n°2 une mauvaise justification du calcul et aucune justification à la question n°4

Synthèse

Ce test nous a permis de recueillir des éléments sur ce que les élèves de terminale savent faire et quelles techniques ils sont capables de mettre en oeuvre suite à l'enseignement qu'ils ont reçu en probabilités conditionnelles. Par exemple, ils sont très à l'aise avec la représentation par un arbre mais l'identification, par l'interprétation d'un énoncé en langage naturel, des événements en jeu peut être un obstacle à une représentation juste de la situation.

Les élèves et étudiants sont très à l'aise sur des tâches rencontrées et traitées au lycée (exercice 3). De façon générale, ils savent identifier des intersections ou des probabilités conditionnelles lorsque l'énoncé est en langage naturel (reconnaissance de "et" ou "sachant que" dans l'exercice 1 ou à la deuxième question de l'exercice 3).

Les élèves et étudiants rencontrent des difficultés à mobiliser les connaissances qui viennent d'être vues pour des tâches qu'ils n'ont jamais rencontrées (exercice 1 : T-CalcProbaCond avec l'évènement qui conditionne qui est une union). Il s'agit d'un type de tâches dont les ingrédients d'une technique sont pourtant des types de tâches qu'ils ont déjà rencontrées dans des cas un peu différents.

Comme nous l'avons rappelé, les étudiants de biologie dont nous venons d'analyser les productions n'ont pas encore suivi le module de probabilités, ces premiers résultats nous ont conforté dans l'idée de recueillir des éléments concernant les pratiques de ces étudiants mais cette fois-ci durant leur enseignement de probabilités à l'Université. Pour ce faire nous avons réalisé un questionnaire à destination de ces étudiants puis des entretiens. Dans les deux parties qui suivent nous présentons le travail d'enquête (questionnaire et entretiens) que nous avons mené auprès de ces étudiants, au moment où ils suivent le module de probabilités.

6.2 Questionnaire en ligne à destination des étudiants de biologie

Nous avons élaboré un questionnaire en ligne à destination des étudiants inscrits en première année de biologie dans l'Université où se déroule nos observations. Ce questionnaire a été proposé au second semestre, en mars 2019 (voir tableau chronologique de recueil de données 3.1 page 47), alors que les étudiants sont en plein dans le module de probabilités, appelé "MATH2". Nous rappelons que ces étudiants ont suivi un module d'analyse au premier semestre, appelé "MATH1".

6.2.1 Présentation du questionnaire

Dans cette sous-partie, nous présentons le questionnaire élaboré à destination des étudiants de biologie, que nous avons intitulé : "Méthodologie en mathématiques chez les étudiants de L1 Biologie".

Nous avons construit ce questionnaire avec trois objectifs. Premièrement, obtenir quelques éléments sur les méthodes de travail en mathématiques des étudiants de biologie, deuxièmement, faire un état des lieux sur le ressenti et le rapport de ces étudiants aux probabilités, enfin, sélectionner des étudiants pour des entretiens approfondis. Il s'agit d'un questionnaire de seulement quelques questions et nous avons conscience des limites de celui-ci, notamment si l'on souhaite creuser l'aspect des méthodes de travail personnel chez les étudiants débutants à l'Université. C'est pourquoi nous proposons dans la partie 6.3, des entretiens afin de compléter les réponses obtenues à ce questionnaire.

Afin de motiver le plus d'étudiants possible à répondre à ce questionnaire nous avons souhaité le rendre anonyme et nous l'avons mentionné dans le texte introductif. Seule la dernière question permet de récolter les adresses email des

participants qui le souhaitent, en vue d'un entretien ultérieur.

Nous avons élaboré ce questionnaire en ligne au moyen de Google Forms. Le lien de ce questionnaire a été diffusé par email à l'ensemble des étudiants de la promotion de première année de biologie (environ 600 étudiants) au moyen d'une liste de diffusion.

Ce questionnaire est composé des sept questions suivantes :

1. *Que faisiez-vous en 2017-2018 ?*

Il s'agit d'une question fermée à choix unique qui nous permettra de choisir parmi les répondants, volontaires pour un entretien, un étudiant provenant de terminale scientifique.

2. *Quelle note avez-vous eue en mathématiques au baccalauréat ?*

Il s'agit d'une question ouverte à réponse courte qui permet de situer le niveau en mathématiques, d'après les attentes de l'institution, des étudiants à leur sortie du lycée.

3. *A l'issue du semestre précédent, quelle note avez-vous obtenue à l'examen de Mathématiques 1 ?*

Il s'agit d'une question ouverte à réponse courte qui permet d'identifier les étudiants qui ont rencontré d'ores-et-déjà des difficultés dans le premier module de mathématiques (MATH1) de la première année de licence.

4. *Combien d'heures avez-vous passées à réviser l'examen de Mathématiques 1 ? (préparation de fiches, entraînement sur des annales, relecture du cours, refaire les exercices de TD, etc...)*

Il s'agit d'une question fermée à choix unique qui nous permet d'avoir une idée du temps que consacrent les étudiants de biologie aux mathématiques, et dans notre cas à la préparation de l'examen du module d'analyse (MATH1). En effet, les étudiants non-spécialistes accordent généralement peu de temps au travail personnel en mathématiques.

5. *Concernant les probabilités enseignées en Mathématiques 2, est-ce que vous trouvez ça plutôt facile, difficile ?*

Il s'agit d'une question fermée à choix unique permettant d'avoir une première idée sur le rapport de ces étudiants aux probabilités.

6. *Seriez-vous d'accord de me rencontrer afin de poursuivre par un entretien plus approfondi concernant vos méthodes de travail en mathématiques ?*

Il s'agit d'une question fermée à choix unique qui nous permettra de connaître les étudiants volontaires pour un entretien ultérieur.

7. *Si vous avez répondu "oui" à la question 6 merci de me laisser une adresse mail sur laquelle je vous contacterai personnellement.*

Il s'agit d'une question ouverte à réponse courte dans laquelle les étudiants peuvent inscrire leur adresse email.

Six semaines après la première diffusion aux étudiants du lien vers ce questionnaire nous avons obtenu 134 réponses. Parmi ces réponses, 28 étudiants (soit 20,9% d'entre eux) donnent leur accord pour poursuivre par un entretien plus approfondi. Nous présentons dans ce qui suit l'analyse des réponses aux cinq premières questions des 134 répondants.

6.2.2 Analyse des réponses au questionnaire

Nous présentons dans ce qui suit l'analyse des réponses des étudiants de biologie à chacune des cinq premières questions présentées ci-dessus.

Q1 : « Que faisiez-vous en 2017-2018 ? »

La première question est une question fermée à choix unique mais avec un item "autre" pour lequel la réponse est libre. Les choix proposés sont les suivants : terminale S, terminale ES ou L, terminale technologique, terminale professionnelle, L1 biologie.

Nous avons obtenu les réponses suivantes :

- 99 (73,9%) étaient en terminale scientifique (notée TaleS).
- 12 (9,4%) étaient en PACES : première année commune aux études de santé.
- 8 (6%) étaient déjà en première année de biologie à l'Université (notée L1 bio).
- 3 (2,2%) étaient en terminale ES ou terminale L (notée TaleES/TaleL).
- 3 (2,2%) étaient en terminale technologique (notée Tale techno).
- 1 (0,7%) était en terminale professionnelle (notée Tale pro).
- 1 (0,7%) était en école d'infirmière.
- 1 (0,7%) était en prépa concours orthophoniste.
- 1 (0,7%) était en prépa concours infirmière.
- 1 (0,7%) était en première année d'informatique à l'Université (notée L1 info).
- 1 (0,7%) était en classe préparatoire hypokhâgne.
- 1 (0,7%) était en classe préparatoire MPSI.
- 1 (0,7%) était en remise à niveau scientifique RNS à l'Université (notée L1 RNS).
- 1 (0,7%) était technicienne de laboratoire après avoir obtenu un BTS (notée TL).

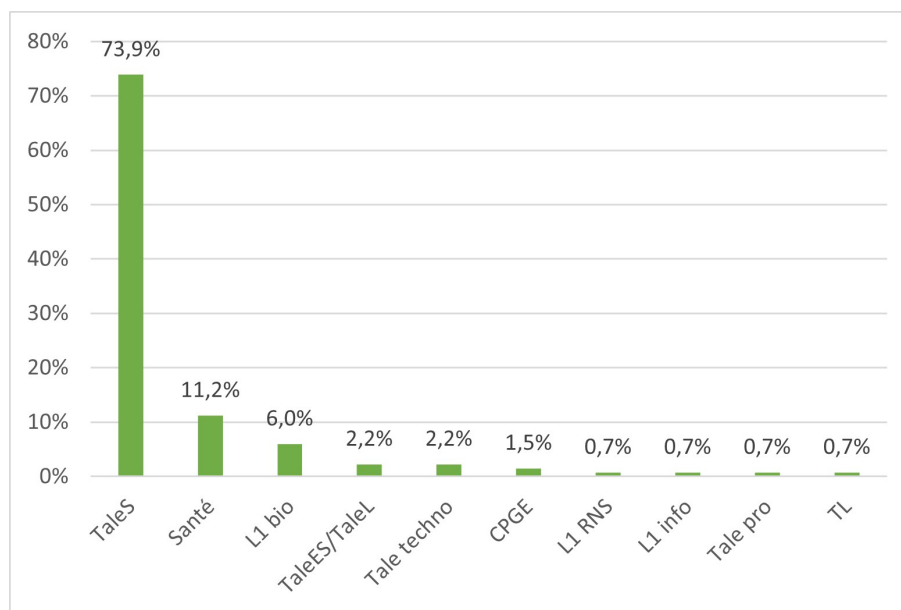


FIGURE 6.21 – Provenance des étudiants de L1 biologie

Nous regroupons les 15 individus provenant de PACES, d'école d'infirmière, de prépa concours orthophoniste et de prépa concours infirmière sous le nom "Santé". De même nous regroupons les 2 individus issus de classes préparatoires sous le nom "CPGE". Les résultats sont résumés dans la figure suivante 6.21.

Etant donné que nous nous intéressons à la transition entre lycée et Université, et plus particulièrement aux élèves qui sortent d'une terminale scientifique, dans la suite nous distinguons deux catégories, les réponses des 99 étudiants provenant de terminale scientifique et les réponses de l'ensemble de la population.

Q2 : « Quelle note avez-vous eue en mathématiques au baccalauréat ? »

Il s'agit d'une question ouverte à réponse courte. Nous avons obtenu les réponses suivantes :

classes de note	[0; 4[[4; 8[[8; 12[[12; 16[[16; 20]	total
effectif population entière	1	20	40	55	18	134
pourcentage population entière	0,8 %	14,9 %	29,9 %	41,0 %	13,4 %	100 %
effectif terminale S	1	18	29	38	13	99
pourcentage terminale S	1,0 %	18,2 %	29,3 %	38,4 %	13,1 %	100 %

La moyenne au baccalauréat des 134 répondants est de 11,72, celle des étudiants sortant de terminale scientifique est de 11,38. Concernant la note en ma-

6.2. Questionnaire en ligne à destination des étudiants de biologie

thématiques que les répondants ont obtenue au baccalauréat, 15,7 % d'entre eux ont eu moins de 8/20, 29,9 % d'entre eux ont eu entre 8 et 12 sur 20, 41 % d'entre eux ont eu entre 12 et 16 sur 20 et 13,4 % d'entre eux ont eu plus de 16/20. Si on ne regarde que les étudiants qui sortent de terminale scientifique les proportions sont très similaires. Ces résultats sont résumés dans la figure suivante 6.22.

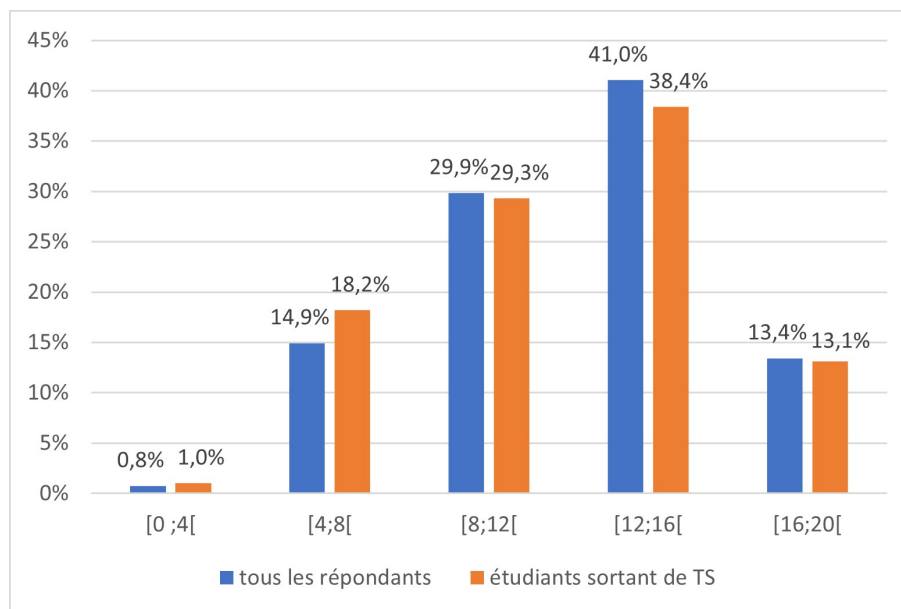


FIGURE 6.22 – Répartition des notes en mathématiques au baccalauréat

Q3 : « A l'issue du semestre précédent, quelle note avez-vous obtenue à l'examen de Mathématiques 1 ? »

Il s'agit d'une question ouverte à réponse courte. Nous avons obtenu les réponses suivantes :

classes de note	[0 ; 4[[4 ; 8[[8 ; 12[[12 ; 16[[16 ; 20]	total
effectif population entière	13	23	40	35	23	134
pourcentage population entière	9,7 %	17,2 %	29,9 %	26,1 %	17,2 %	100 %
effectif terminale S	4	15	32	31	17	99
pourcentage terminale S	4,0 %	15,2 %	32,3 %	31,3 %	14,2 %	100 %

La moyenne des notes à l'examen de MATH1 de la population entière est de 10,53 tandis qu'elle est de 11,28 chez les étudiants provenant de terminale scientifique. 26,9 % des répondants déclarent avoir eu moins de 8/20 et ils sont 43,3 %

à avoir eu plus de 12/20. Les notes des étudiants sortant de terminale scientifique sont légèrement meilleures, 19,2 % d'entre eux ont eu strictement moins de 8/20 et 48,5 % d'entre eux ont eu plus de 12/20. Ces résultats sont résumés dans la figure suivante 6.23.

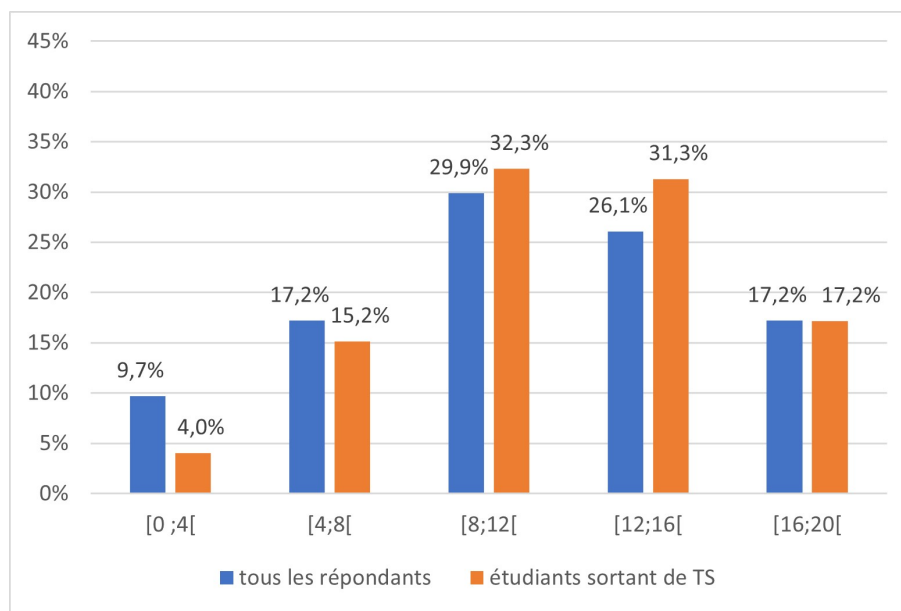


FIGURE 6.23 – Répartition des notes à l'examen de MATH1 (analyse)

Nous avons ensuite regardé pour chaque étudiant, la différence entre la note obtenue en mathématiques au baccalauréat et celle obtenue à l'examen du module d'analyse MATH1. Cette comparaison a bien sûr des limites, les attendus (rédaction, justification, contenus) sont différents, les correcteurs sont différents (enseignants de lycée vs enseignants d'Université) et les contenus sont plus divers au baccalauréat que pour l'examen de MATH1 (analyse fonctionnelle, intégration, équations différentielles). Voici les résultats que nous avons obtenus en comparant ces deux notes pour chaque étudiant :

- Parmi les 134 répondants autant ont progressé que baissé (50,7 % d'entre eux ont eu une meilleure note au partiel de MATH1 qu'au baccalauréat de maths, 49,3 % d'entre eux ont eu une meilleure note au bac).
- Les 99 étudiants sortant de terminale scientifique ont plutôt progressé (59,6 % d'entre eux ont eu une meilleure note au partiel de MATH1 qu'au baccalauréat de maths, 40,4 % d'entre eux ont eu une meilleure note au bac).

Q4 : « Combien d'heures avez-vous passées à réviser l'examen de MATH1 ? »

Nous avons précisé à la suite de la question : (préparation de fiches, entraînement sur des annales, relecture du cours, refaire les exercices de TD, etc...). Il s'agit

6.2. Questionnaire en ligne à destination des étudiants de biologie

d'une question fermée à réponse unique parmi les items proposés : moins d'1h, entre 1h et 3h, entre 3h et 6h, plus de 6h. Nous avons obtenu les réponses suivantes :

	moins d'1h	1h à 3h	3h à 6h	plus de 6h	total
effectif population entière	10	43	56	25	134
pourcentage population entière	7,5 %	32,1 %	41,8 %	18,7 %	100 %
effectif terminale S	7	34	40	18	99
pourcentage terminale S	7,1 %	34,3 %	40,4 %	18,2 %	100 %

7,5 % des répondants déclarent avoir travaillé moins d'une heure pour préparer le partiel de MATH1, 32,1 % des répondants déclarent avoir travaillé entre 1h et 3h, 41,8 % des répondants déclarent avoir révisé entre 3h et 6h, enfin 18,7 % des répondants déclarent avoir travaillé plus de 6h. Les proportions sont très similaires lorsque l'on s'intéresse uniquement aux répondants qui sortent de terminale scientifique. Ces résultats sont résumés dans la figure suivante 6.24.

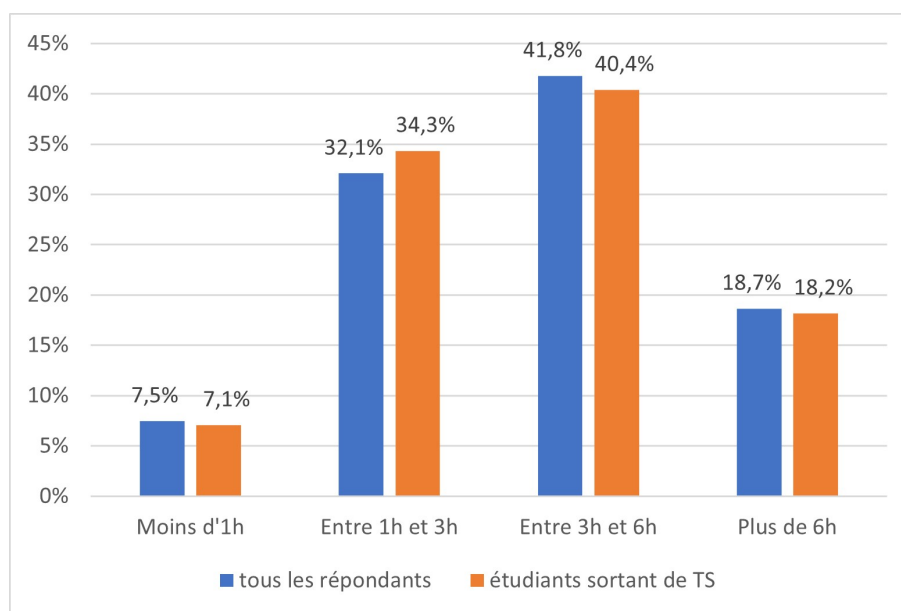


FIGURE 6.24 – Temps consacré à préparer et réviser l'examen de MATH1 (analyse)

Q5 : « Concernant les probabilités enseignées en Mathématiques 2, est-ce que vous trouvez ça plutôt facile, difficile ? »

Il s'agit d'une question fermée à réponse unique avec une échelle représentant la difficulté ressentie (voir figure suivante).

FIGURE 6.25 – Enoncé de la question n°5

Nous avons obtenu les réponses suivantes :

	1	2	3	4	5
difficulté	<i>très facile</i>	<i>facile</i>	<i>normal</i>	<i>difficile</i>	<i>très difficile</i>
effectif population entière	5	16	51	52	10
pourcentage population entière	3,7 %	11,9 %	38,1 %	38,8 %	7,5 %
effectif terminale S	4	13	37	38	2
pourcentage terminale S	4,3 %	13,8 %	39,4 %	40,4 %	2,1 %

A nouveau, les résultats des étudiants sortant de terminale scientifique sont très proches des réponses de la population entière. Seulement 15,6 % des étudiants considèrent le module de probabilités comme très facile ou facile, ils sont 46,3% à trouver les probabilités enseignées dans ce cours difficiles ou très difficiles.

Ces résultats sont résumés dans la figure suivante 6.26.

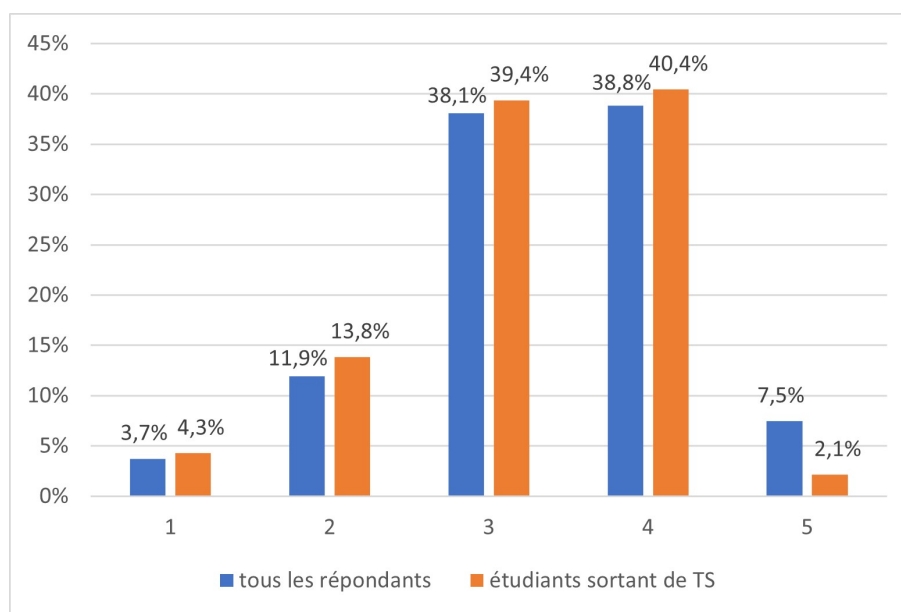


FIGURE 6.26 – Ressenti quant à la difficulté du module MATH2 (probabilités) : très facile (1) ... très difficile (5)

Synthèse

Ce questionnaire nous a permis de faire un premier état des lieux du ressenti des étudiants de première année de biologie qui sont en plein dans le module de probabilités du second semestre. Il est intéressant de relever que les étudiants déclarent trouver assez difficile les contenus du module de probabilités.

Ce questionnaire nous a également servi à contacter plusieurs étudiants et étudiantes pour des entretiens plus approfondis que nous allons détailler dans la partie suivante.

6.3 Entretiens avec des étudiants de première année de biologie

L'objectif de ces entretiens est d'approfondir certains points évoqués dans le questionnaire, notamment le travail personnel consacré aux mathématiques et plus particulièrement aux probabilités; ou encore le ressenti et les difficultés rencontrées pour les contenus de probabilités enseignés dans ce module. En effet, des auteurs ont mis en évidence que les étudiants non-spécialistes accordent peu de temps au travail personnel en mathématiques.

Parmi les 28 étudiants prêts à nous accorder un entretien suite au questionnaire, nous avons sélectionné des profils d'étudiants sortant de terminale scientifique dont l'écart de notes entre le baccalauréat et l'examen de MATH1 est négatif,

ce qui ne faisait plus que neuf étudiants. En effet, notre problématique nous amène à nous intéresser aux étudiants qui sont sortis du lycée avec un niveau correct en mathématiques et qui rencontrent des difficultés lors de leur arrivée à l'Université. Nous avons contacté par email ces étudiants pour leur proposer un entretien et finalement nous avons rencontré deux étudiants (Adam et Ben) et deux étudiantes (Cloé et Maud) en avril et mai 2019 (voir tableau chronologique de recueil de données 3.1 page 47).

6.3.1 Présentation de la grille d'entretien

La grille d'entretien que nous avons élaborée se trouve en annexe G. Nous avons conçu cette grille pour orienter l'échange avec les étudiants sélectionnés, mais au sein d'une même partie les questions n'ont pas forcément été abordées dans l'ordre présenté dans la grille.

Dans la grille nous notons "CC" pour "contrôle continu", il s'agit à l'Université d'une modalité de contrôle de connaissances. Par exemple pour le module de probabilités (MATH2) il y a un contrôle continu (CC) au milieu du module et un contrôle terminal (examen) à la fin du module. Voici une présentation des différentes parties de cette grille d'entretien.

- ▷ La première partie de l'entretien va permettre de mieux connaître le profil de l'étudiant.
- ▷ Dans une seconde partie nous cherchons à savoir quelles ressources l'étudiant utilise pour les contenus de probabilités à l'Université, comment et à quelles occasions.
- ▷ Dans une troisième partie nous voulons aborder ces mêmes questions mais pour les mathématiques en général.
- ▷ La quatrième partie permet de développer la question n°5 du questionnaire en ligne à propos du ressenti de l'étudiant et des difficultés rencontrées pour les contenus de probabilités.
- ▷ Enfin, dans une cinquième et dernière partie, nous avons souhaité revenir avec l'étudiant sur deux exercices vus en TD (l'exercice 5 du TD2 pour le thème des probabilités conditionnelles et l'exercice 4 du TD4 sur les lois continues, voir énoncés figure 6.27).

Exercice 5 Le gérant d'un magasin d'informatique a reçu un lot de boîtes de CD-ROM. 5% des boîtes sont abîmées. Le gérant estime que 60% des boîtes abîmées contiennent au moins un CD-ROM défectueux et que 98% des boîtes non abîmées ne contiennent aucun CD-ROM défectueux. Un client achète une boîte du lot. On désigne par A l'évènement : «la boîte est abîmée» et par D l'évènement «la boîte achetée contient au moins un disque défectueux».

- (a) Calculer les probabilités $P(A)$, $P(\bar{A})$, $P(D|A)$, $P(D|\bar{A})$, $P(\bar{D}|A)$ et $P(\bar{D}|\bar{A})$. On commencera par caractériser celles qui sont explicitement dans l'énoncé.
- (b) Le client constate qu'un des CD-ROM acheté est défectueux. Quelle est la probabilité pour qu'il ait acheté une boîte abîmée?

Exercice 4 Soit X une variable aléatoire de densité $f(x) = cx^2 1_{0 \leq x \leq 1} + \left(\frac{1}{2}\right) 1_{2 \leq x \leq 3}$.

- (a) Que vaut c ?
- (b) Calculer $P(\frac{1}{4} \leq X \leq \frac{3}{4})$?
- (c) Calculer l'espérance et la variance de X .

FIGURE 6.27 – Exercice 5 du TD2 (probabilités conditionnelles)
– Exercice 4 du TD4 (lois continues)

Nous avons également envoyé aux étudiants par email, en amont de l'entretien, deux exercices à faire en avance pour en discuter ensemble lors de l'entretien. Il s'agit d'un exercice pour le thème des probabilités conditionnelles "exercice 1" et d'un second sur les lois continues "exercice 2". Les énoncés de ces deux exercices sont à retrouver figure 6.28.

Exercice 1

Un grossiste achète des boîtes de thé vert chez deux fournisseurs. Il achète 80% de ses boîtes chez le fournisseur A et 20% chez le fournisseur B. 10% des boîtes provenant du fournisseur A présentent des traces de pesticides et 20% de celles provenant du fournisseur B présentent aussi des traces de pesticides.

On prélève au hasard une boîte du stock du grossiste et on considère les évènements suivants :

- événement A : « la boîte provient du fournisseur A » ;
- événement B : « la boîte provient du fournisseur B » ;
- événement S : « la boîte présente des traces de pesticides ».

1. Quelle est la probabilité de l'évènement $B \cap S$?
2. Justifier que la probabilité que la boîte prélevée ne présente aucune trace de pesticides est égale à 0,88.
3. On constate que la boîte prélevée présente des traces de pesticides. Quelle est la probabilité que cette boîte provienne du fournisseur B ?

Exercice 2

La production quotidienne X d'un produit en tonnes est une variable aléatoire continue qui prend ses valeurs dans l'intervalle $[0; 10]$ avec la densité de probabilité f définie par :

$$f(x) = 0,006(10x - x^2).$$

1. Vérifier que f est bien une densité de probabilité sur $[0; 10]$.
2. Calculer la probabilité des évènements suivants :
A : « $X \leq 7$ » ; B : « la production quotidienne dépasse 6 tonnes ».
3. Calculer l'espérance mathématique $E(X)$.

FIGURE 6.28 – Exercices envoyés en avance – Exercice 1 (probabilités conditionnelles) – Exercice 2 (lois continues)

Pour le thème des probabilités conditionnelles, nous avons choisi l'exercice 5 du TD2 (figure 6.27) car la question n°2 qui relève du type de tâches **T-CalcProbaCond**, fait intervenir la formule de Bayes non vue dans le secondaire. Nous avons choisi l'exercice 1 (figure 6.28) qui est issu du baccalauréat Asie 2013 car il propose également **T-CalcProbaCond** (question n°3) mais dans cet exercice pas besoin de la formule de Bayes car les étapes successives font l'objet des questions n°1 et n°2. Nous détaillons ci-dessous les types de tâches proposés dans l'exercice 1 (figure 6.28).

- La question n°1 de l'exercice 1 relève du type de tâches **T-CalcInter** avec les variations suivantes :
 - étant donné que les évènements ont déjà été décrits et nommés dans le texte
 - les données sont en pourcentages
- La seconde question de l'exercice 1 relève du type de tâches **T-CalcSimple** avec les variations suivantes :
 - à partir d'un énoncé en langage naturel
 - étant donné que les évènements ont déjà été décrits et nommés dans le texte
 - les données sont en pourcentages
- La troisième question de l'exercice 1 relève du type de tâches **T-CalcProbaCond** avec les variations suivantes :
 - à partir d'un énoncé en langage naturel
 - étant donné que les évènements ont déjà été décrits et nommés dans le texte
 - étant donné que les probabilités qui apparaissent dans la formule ont déjà été calculées
 - les données sont en pourcentages

En ce qui concerne le thème des lois continues, nous avons choisi l'exercice 4 du TD4 (figure 6.27) parce qu'il présente une fonction de densité continue par morceaux, définie par une indicatrice, qui sont deux grandes nouveautés du cours. De plus la première question relève de **T-Det(cf)** un nouveau type de tâches. Nous avons sélectionné l'exercice 2 (figure 6.28) qui est issu du manuel Indice car il propose des types de tâches similaires mais pour une fonction de densité plus classique, nous détaillons ci-dessous les types de tâches proposés dans l'exercice 2 (figure 6.28).

- La question n°1 de l'exercice 2 relève du type de tâches **T-Veriff** étant donné que la variable aléatoire est identifiée dans l'énoncé et que sa densité a été donnée.

- La seconde question de l'exercice 2 relève du type de tâches **T-CalcSimple** avec les variations suivantes :
 - étant donné que la variable aléatoire est identifiée dans l'énoncé et que sa densité a été donnée
 - étant donné que l'énoncé est en langage naturel
- La troisième question de l'exercice 2 relève du type de tâches **T-CalcEsp** étant donné que la variable aléatoire est définie et que sa densité a été donnée.

Concernant les deux exercices envoyés en avance par email (figure 6.28, nous avons demandé à l'étudiant de chercher à les résoudre si possible et s'il en a le temps. Il faut qu'il puisse y réfléchir sans forcément les résoudre entièrement et puisse expliquer ce qui le bloque ou comment il procède. Nous prévoyons de reprendre durant l'entretien la solution (partielle ou complète) proposée.

Comme mentionné précédemment, nous avons finalement rencontré quatre étudiantes et étudiants de première année de biologie : Cloé, Maud, Adam et Ben pour un entretien selon la grille présentée ci-dessus. Les entretiens ont eu lieu après le TD4 sur les lois continues (et donc après le TD sur les probabilités conditionnelles). Les transcriptions de ces entretiens sont à retrouver en annexe I (I.1, I.2, I.3 et I.4). Dans ce qui suit nous présentons l'analyse de ces entretiens et des réponses obtenues.

6.3.2 Analyse des entretiens

Pour commencer, nous présentons dans le tableau suivant (voir figure 6.29), selon les premières questions de la grille d'entretien et les question n°2 et n°3 du questionnaire en ligne, les profils des quatre étudiants interrogés.

	Adam	Cloé	Maud	Ben
Note de maths au baccalauréat	15	11	18	15
Mention au baccalauréat	AB	AB	TB	B
Note au CC de MATH1	15,5	14,5	20	17,5
Note à l'examen de MATH1	11	10	12,5	10
Note au CC de MATH2	2,5	11	15,5	14,5
Note à l'examen de MATH2	7	6,5	13	3,5

FIGURE 6.29 – Profils des étudiants interrogés

Pour rappel, les entretiens ont eu lieu après le TD sur les probabilités conditionnelles et après celui sur les lois continues. Le contrôle continu (CC) de MATH2

a déjà eu lieu mais les notes à l'examen ne nous ont été communiquées par les étudiants qu'en toute fin d'année. Nous pouvons constater dès à présent que pour trois de ces quatre étudiants le module de MATH2 (probabilités) a été largement moins bien réussi, que le module de MATH1 (analyse) du premier semestre.

En ce qui concerne le temps passé à réviser et travailler le module de MATH1 (Q4 du questionnaire) et le module de MATH2 (Q1 et Q2 de la grille d'entretien), les réponses que nous avons obtenues sont à retrouver dans la figure 6.30.

	Adam	Cloé	Maud	Ben
Nombre d'heures à réviser l'examen de MATH1	1h à 3h	3h à 6h	1h à 3h	1h à 3h
Nombre d'heures de travail pour MATH2 par semaine	0	moins d'1h	0	1/4h
Nombre d'heures à réviser le CC de MATH 2	4h	3h à 4h	10h	3h

FIGURE 6.30 – Temps consacré aux modules de mathématiques en L1

Ces quatre étudiants déclarent travailler davantage pour réviser et préparer un examen (ou contrôle) pour le module de probabilités MATH2 que pour celui de MATH1 (analyse). Ce qui confirme que ce module est difficile pour les étudiants, sans doute davantage que le module d'analyse du premier semestre.

En ce qui concerne les ressources utilisées pour réviser et préparer le contrôle continu de probabilités (Q3 de la grille d'entretien), les réponses que nous avons obtenues sont résumées dans le tableau suivant 6.31.

	Adam	Cloé	Maud	Ben
Cours	x	x		x
Refaire les TD		x	x	x
Relire les TD	x	x	x	x
Fiches de révisions ou assimilé	x		x	
Internet		x	x	
Feuille d'entraînement	x	x	x	x
Ressources de terminale		x		

FIGURE 6.31 – Ressources utilisées par les étudiants

Les quatre étudiants déclarent avoir préparé le contrôle continu à partir de la feuille d'entraînement (voir annexe H) qui a été mise à disposition sur Moodle. Tous sauf Adam déclarent refaire les exercices de TD, Adam se contente de relire ses notes prises en TD. Concernant le cours, hormis Maud, les étudiants interrogés déclarent relire leurs notes de cours ou les utiliser pour se rappeler des formules. Adam déclare faire des fiches de révisions tandis que Maud écrit les formules du cours sur des feuilles qu'elle affiche au mur. Cloé et Maud déclarent utiliser internet, soit pour trouver des exercices sur lesquels s'entraîner, soit pour trouver des éléments de cours ou des explications complémentaires. Enfin, Cloé déclare utiliser son cours de terminale pour s'entraîner sur des exercices et exemples similaires à ceux rencontrés en TD.

Concernant le travail durant les séances de TD (Q4 de la grille d'entretien), précisons que les réponses des étudiants dépendent aussi de l'enseignant de TD (les étudiants n'ont pas tous le même enseignant de TD) et de l'organisation que cet enseignant a choisi pour ces séances. Adam déclare que son activité en TD consiste seulement à prendre la correction en note car aucune autonomie n'est laissée aux étudiants. En revanche, Cloé, Maud et Bastien expliquent chercher les exercices du TD durant la séance et compléter leurs notes selon les remarques orales de l'enseignant et la correction inscrite au tableau.

Comme nous l'avons vu dans les réponses à la question n°3 les étudiants se servent tous des notes de TD pour préparer l'examen (ou le contrôle continu), nous leur avons donc demandé s'ils reprenaient tous les exercices de ces feuilles de TD et comment ils retravaillaient leurs notes (Q5 de la grille d'entretien). Adam nous a confirmé qu'il ne faisait que lire ses notes et que tous les exercices étaient corrigés lors des séances de TD. Cloé et Maud déclarent refaire tous les exercices des TD puis vérifier ou compléter avec leurs notes. Maud explique que tous les exercices ne sont pas faits en séances de TD et qu'elle complète avec des notes d'autres étudiants qui sont dans des groupes différents. Ben quant à lui refait certains exercices mais pas à l'écrit : « *J'ai l'exercice devant moi, je fais la réponse dans ma tête puis je compare avec ce que j'avais écrit en TD* ».

A propos des contenus de probabilités du lycée (Q7 de la grille d'entretien) Adam et Cloé déclarent qu'ils étaient plus à l'aise avec les probabilités discrètes de seconde ou de première qu'avec "*les intégrales de terminale*". Maud n'aimait pas les probabilités qu'elle trouvait trop théoriques et Ben déclare qu'il travaillait plus les probabilités au lycée que les autres contenus de mathématiques car il trouvait cela moins intuitif.

Les étudiants interrogés sont unanimes pour dire que les contenus du module de MATH2 sont en continuité avec le lycée (Q8 de la grille d'entretien). Cloé précise que c'est "*plus poussé*" quant à Ben il déclare : "*On avait déjà vu tout ça au lycée mais ici c'est avec d'autres méthodes, d'autres manières*".

A la question n°5 du questionnaire en ligne concernant la difficulté (générale)

des contenus enseignés dans le module de MATH2, Adam a répondu (5) très difficile, Cloé (4) difficile, Maud (3) normal, et Ben (1) très facile.

La question 9 et la suite de l'entretien (échanges à propos des exercices de TD et des exercices n°1 et n°2 envoyés en avance 6.28) nous permettent de présenter dans ce qui suit une synthèse de ce qui ressort de ces échanges concernant les difficultés qu'éprouvent les étudiants en probabilités à l'Université. Les résultats présentés ci-dessous ne sont pas nécessairement représentatifs de l'ensemble des étudiants de biologie car nous n'avons interrogé que quatre étudiants. Lorsque nous écrivons "*les étudiants*" c'est qu'au-moins trois d'entre eux ont évoqué cette difficulté lors de l'entretien.

- Les étudiants avaient l'habitude de travailler avec des arbres de probabilités au lycée, ils continuent à en faire au brouillon et déclarent éprouver des difficultés à s'en passer dans le cadre d'exercices s'y prêtant bien (le TD2 par exemple, voir 5.49 page 197). Tous sauf Ben représentent la situation de l'exercice n°1 par un arbre de probabilités.
- Les étudiants sont perturbés par certains termes de vocabulaire et certaines nouvelles notations du cours. Par exemple, concernant la notation de la probabilité conditionnelle, ils étaient habitués à écrire $P_B(A)$ et déclarent avoir des difficultés à utiliser $P(A|B)$. Adam confond $P(A|B)$ avec $P(A \setminus B)$. Seule Maud utilise la nouvelle notation $P(A|B)$ dans l'exercice n°1. Autre exemple, pour l'événement contraire, les étudiants avaient l'habitude d'utiliser la notation \bar{A} et déclarent unanimement que la notation A^c leur pose des difficultés.
- Le CM pose des difficultés aux étudiants, il serait selon eux très dense, trop rapide et Cloé regrette que ce soit sous forme de diapositives. Les étudiants éprouvent des difficultés à suivre, Maud explique que c'est : "trop rapide, on a juste le temps de copier". Adam regrette que les exemples présents sur les diapositives ne soient pas toujours détaillés. Les étudiants reprochent au CM d'être trop théorique et en décalage avec ce qu'on attend d'eux dans les exercices proposés en TD.
- De façon générale ils déclarent éprouver des difficultés à identifier la loi qu'ils rencontrent ou à reconnaître s'ils sont dans le cas discret ou continu. Adam et Cloé déclarent qu'ils ne savent pas toujours, face à une question, ce qu'on attend d'eux pour répondre. La multiplicité des lois à connaître est quelque chose qu'Adam et Ben trouvent particulièrement difficile.
- Cloé et Maud déclarent qu'utiliser la fonction indicatrice leur pose des difficultés. La première question de l'exercice 4 du TD4 pose de grosses difficultés aux étudiants : identifier ce qu'il faut faire, comment il faut le faire ou encore utiliser correctement l'indicatrice.
- Adam et Cloé déclarent spontanément ne pas savoir se servir d'une table de loi normale à la suite du TD4.

Synthèse

De ces entretiens nous retenons qu'effectivement ce module de probabilités pose des difficultés aux étudiants qui, au lycée, n'avaient pas de difficultés particulières avec ces contenus. Les élèves sont perturbés par les changements, qu'il s'agisse des attendus, des notations ou des modalités de fonctionnement.

Bien que les types de tâches soient très similaires à ceux rencontrés au lycée, les étudiants ont des difficultés à utiliser des techniques différentes de celles apprises dans le secondaire, par exemple utiliser une table pour la loi normale ou se passer des arbres pour les probabilités conditionnelles. Il s'agit d'un point important concernant la transition secondaire-supérieur dans le cas qui nous intéresse des probabilités.

Afin de compléter ces résultats et analyses nous avons souhaité avoir accès aux productions de ces étudiants lors de l'examen terminal de ce module de probabilités. Nous présentons l'analyse des copies d'examens dans la partie suivante.

6.4 Examen du module de probabilités en première année de biologie

L'examen du module de probabilités (MATH2) a eu lieu en fin d'année (mai 2019), plusieurs semaines après les entretiens détaillés précédemment. Nous avons pu récupérer le sujet d'examen (à retrouver figure 6.32), photographier les copies de Adam et de Maud et nous entretenir avec ces derniers à propos de leur copie respective et quant à leur ressenti.

Nous commençons par exposer dans quelles circonstances nous avons pu avoir accès aux copies d'examens de ces étudiants. Dans une deuxième sous-partie nous présentons l'analyse *a priori* des exercices relevant des thèmes qui nous intéressent : probabilités conditionnelles et lois continues, ces exercices se trouvent figure 6.32.

Nous ne détaillons donc pas l'analyse *a priori* des exercices 3 et 5 du sujet qui relèvent du thème des variables aléatoires discrètes, thème pour lequel nous n'avons pas observé de séance au lycée, ni à l'Université. Enfin nous présentons, au regard de l'analyse *a priori*, une analyse des copies d'examen de Adam et de Maud.

MAT 2

CONTRÔLE TERMINAL – DURÉE 2H

L'utilisation de calculatrice, de téléphone ainsi que de tout document est interdite. Les réponses doivent être rédigées. Les résultats doivent être donnés sous forme de fractions simplifiées

Question de cours [2pts]

Soient A, B et C trois événements. À quelles conditions ces trois événements sont-ils indépendants ?

2 points pour toute la réponse ou 0 : pas de demi mesure sur une question de cours

Pour que ces événements soit indépendants il faut que :

- $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$, $\mathbf{P}(A \cap C) = \mathbf{P}(A)\mathbf{P}(C)$, $\mathbf{P}(C \cap B) = \mathbf{P}(C)\mathbf{P}(B)$
- et $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)$

Exercice 1 [4 pts]

Soient c un nombre et X une variable aléatoire de densité $f(x) = c\mathbf{1}_{-1 \leq x < 0} + x\mathbf{1}_{0 \leq x \leq 1}$

- (a) Déterminez la valeur de c . **1 pt**
- (b) Donnez la valeur médiane de X . **1 pt**
- (c) Calculez $\mathbf{E}[X]$. **1 pt**
- (d) Calculez $\text{Var}(X)$. **1 pt**

- (a) La fonction f étant une densité de probabilité, $1 = \int_{-\infty}^{+\infty} f(x)dx = \int_{-1}^0 cdx + \int_0^1 xdx = c + \frac{1}{2}$, c'est-à-dire $c = \frac{1}{2}$.
- (b) Puisque $\mathbf{P}(X < 0) = \mathbf{P}(X > 0) = \frac{1}{2}$, la valeur médiane de X est 0.
- (c) Calculons $\mathbf{E}[X] = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-1}^0 \frac{x}{2}dx + \int_0^1 x^2dx = -\frac{1}{4} + \frac{1}{3} = \frac{1}{12}$.
- (d) Calculons $\mathbf{E}[X^2] = \int_{-\infty}^{+\infty} x^2f(x)dx = \int_{-1}^0 \frac{x^2}{2}dx + \int_0^1 x^3dx = \frac{1}{6} + \frac{1}{4} = \frac{5}{12}$. En utilisant la formule de Huygens $\text{Var}(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{5}{12} - \frac{1}{12^2} = \frac{59}{144}$.

Exercice 2 [3pts]

Soit X une v.a. d'espérance 3 et de variance 1 et Z une v.a. d'espérance 2 et de variance 4.

- (a) Quel est l'écart-type de $Y = 9X + 1$? **1 pt**
 - (b) La variance de $X + Z$ vaut 7. Que vaut le covariance de X et Z ? **1 pt**
 - (c) Que vaut la corrélation de X et Z ? **1 pt**
- (a) D'après les résultats du cours $\sigma_Y = 9\sigma_X = 9$
 - (b) Nous savons que $\text{Var}(X + Z) = \text{Var}(X) + 2\text{Cov}(X, Z) + \text{Var}(Z)$. Dans notre cas : $7 = 1 + 2\text{Cov}(X, Z) + 4$, c'est-à-dire $\text{Cov}(X, Z) = 1$.
 - (c) $\text{Corr}(X, Z) = \frac{\text{Cov}(X, Z)}{\sigma_X \sigma_Z} = \frac{1}{\sqrt{1} \sqrt{4}} = \frac{1}{2}$

6.4. Examen du module de probabilités en première année de biologie

Exercice 4 [5pts]

Au RU, deux types de légumes sont proposés : haricots verts ou épinards. Les six cents clients quotidiens du RU choisissent indépendamment les uns des autres les haricots avec probabilité 0.6 et les épinards avec probabilité 0.4.

On note S le nombre d'étudiants choisissant les épinards.

- (a) Quelle est la loi de S ? **1pt**
 - (b) Quel est l'écart-type de S ? **1pt**
 - (c) En utilisant le TLC, déterminez N tel que $\mathbf{P}(S > N)$ soit environ égal à $1/20$ (une table de la loi normale se trouve au dos). **1pt pour trouver la va normale + 1pt pour els calculs**
 - (d) Pour éviter le gaspillage, on décide de prévoir assez de parts d'épinards pour qu'il y ait moins d'une chance sur vingt que les étudiants choisissant des épinards n'en aient pas. Combien de plats d'épinards doit-on prévoir? **1pt**
- (a) Le choix d'un étudiant est modélisé par une expérience de Bernoulli dont le succès a une probabilité 0.4. La v.a. S compte le nombre de succès lors de la répétitions de 600 expériences indépendantes : $S \sim \mathcal{B}(600, 0.4)$
 - (b) Nous savons que $Var(S) = 600(0.4)(1 - 0.4) = 6 \cdot 4 \cdot 6 = 12^2$, ainsi $\sigma_S = \sqrt{Var(S)} = 12$.
 - (c) Nous cherchons N tel que $\mathbf{P}(S < N) = 0.95$, c'est-à-dire $\mathbf{P}\left(\frac{S - \mathbf{E}[S]}{\sigma_S} < \frac{N - 240}{12}\right) = 0.95$. d'après le TCL, la v.a. $\frac{S - \mathbf{E}[S]}{\sigma_S}$ suit à peu près une loi normale centrée réduite. La table donne $\frac{N - 240}{12} = 1,65$ et $N = 240 + 12 \cdot 1,65 = 240 + 16,5 + 3,3 = 259,8$ (**260 est aussi considéré comme juste**)
 - (d) D'après la question précédente nous devons prévoir 260 plats d'épinards.

$$\Phi(t) = P(X \leq t) \text{ pour } X \sim \mathcal{N}(0, 1)$$

t	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
1	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,8665	0,86864	0,87076	0,87286	0,87493	0,87698	0,879	0,881	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,9032	0,9049	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,9222	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,9452	0,9463	0,94738	0,94845	0,9485	0,95003	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,9608	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,9732	0,97381	0,97441	0,975	0,97558	0,97615	0,9767
2	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,9803	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,983	0,98341	0,98382	0,98422	0,98461	0,985	0,98537	0,98574
2,2	0,9861	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,9884	0,9887	0,98899
2,3	0,98928	0,98956	0,98983	0,9901	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,9918	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,9943	0,99446	0,99461	0,99477	0,99492	0,99506	0,9952
2,6	0,99534	0,99547	0,9956	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,9972	0,99728	0,99736
2,8	0,99744	0,99752	0,9976	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861

FIGURE 6.32 – Enoncé du sujet d'examen et corrigé de l'enseignant

6.4.1 Méthodologie du recueil et de l'analyse des copies

Une fois les examens terminés et les notes communiquées aux étudiants, a lieu en toute fin d'année la consultation des copies d'examen. Nous avons pu assister en juin 2019 (voir tableau chronologique de recueil de données 3.1 page 47), avec Adam et Maud. Il s'agit d'un moment qui a lieu à l'Université, où les étudiants peuvent, s'ils le souhaitent, venir consulter leurs copies d'examen qui sont notées et corrigées. Lors de cette consultation nous avons pu prendre en photo les copies de Maud et Adam et recueillir leurs réactions face à celles-ci. Durant ces échanges nous avons pris des notes de leurs commentaires.

Nous exposons dans ce qui suit l'analyse *a priori* faite en regard des observations menées à l'Université (partie 5.3 page 194). Systématiquement nous comparons le sujet d'examen avec les types de tâches et variations qui ont été rencontrés lors des séances de CM et de TD auxquelles nous avons assisté.

Enfin, nous présentons l'analyse des copies de Adam et de Maud, appuyée sur l'analyse *a priori* du sujet et complétée par les commentaires recueillis lors de la consultation des copies. L'analyse que nous proposons de leurs copies a des limites car entre en jeu également la façon dont les étudiants ont travaillé pour l'examen, les séances auxquelles ils ont assisté et comment les exercices ont été traités durant ces séances.

6.4.2 Analyse a priori du sujet d'examen

Nous présentons dans ce qui suit l'analyse *a priori* de la « Question de cours » et des exercices 1, 2 et 4 du sujet d'examen terminal de première année de biologie à l'Université en regard des résultats issus des observations menées en classe, présentés partie 5.3 page 194. Cette analyse a des limites car nous n'avons observé que deux séances de TD pour un seul groupe, il peut y avoir des enseignants de TD qui ont procédé différemment ou des exercices relevant des deux thèmes qui nous concernent qui ont été traités lors de séances ultérieures.

Concernant la question de cours, l'indépendance de trois événements a été évoquée en CM en tant qu'exemple (diapositive n°17 figure 207 page 207) à la suite de la diapositive n°16 qui propose une définition générale dans le cas de n événements. La réponse demandée n'apparaît donc pas explicitement sur le photocopié de cours présent sur Moodle. Or vu le nombre d'absents en CM on peut imaginer que cette question, bien que simplement « de cours », n'a pas été réussie par beaucoup d'étudiants.

L'exercice 1 ressemble énormément à l'exercice 4 du TD4. Il propose les types de tâches et variations déjà rencontrés suivants : **T-Det(cf)**, **T-CalcEsp** et **T-CalcVar**, étant donné que la fonction de densité est donnée dans l'énoncé, et qu'il s'agit d'une fonction continue par morceaux définie par une indicatrice. Ces types de tâches ont déjà été proposés en TD et traités par les étudiants, la fonction indicatrice est une nouveauté vis-à-vis du lycée.

- La première question qui relève du type de tâches **T-Det(cf)** a certainement posé des difficultés étant donné qu'il s'agit d'un nouveau type de tâches vis-à-vis du lycée et que les sous-activités ont toutes été prises en charge par l'enseignant en séance de TD. Il faut d'abord que les étudiants reconnaissent la méthode : pour déterminer la constante c ils doivent utiliser les propriétés de f (son intégrale vaut 1 sur \mathbb{R} et la fonction est positive). Ils doivent ensuite calculer l'intégrale de f sur \mathbb{R} , mais pour cela il faut réduire l'intervalle d'intégration en utilisant correctement l'indicatrice, ce qui ne sera sans doute

pas réussi par tous étant donné qu'ils éprouvent des difficultés à manipuler l'indicatrice (voir résultats des entretiens). La reconnaissance de la méthode, et donc des ingrédients d'une technique cités ci-dessus, est à l'initiative de l'étudiant.

- On peut imaginer que les étudiants qui n'ont pas abouti à la première question se sont arrêtés là pour cet exercice pensant ne pas être en mesure de pouvoir faire les questions suivantes sans la valeur de c .
- La seconde question relève du type de tâches, *a priori* jamais rencontré **T-Det(médiane)** : déterminer la valeur de la médiane d'une variable aléatoire continue. Les étudiants doivent ici transférer la définition de la médiane vue en CM (diapositive n°7 voir figure 6.33) à un exercice particulier où la fonction de densité est définie par une indicatrice. Cette question aura sans doute posé des difficultés aux étudiants d'autant qu'elle n'a jamais été traitée en séance de TD.

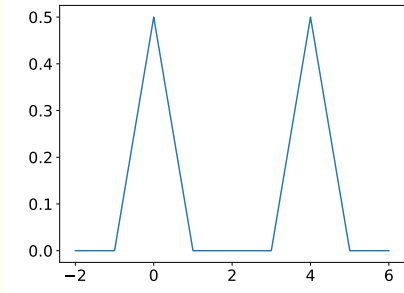
4.2. MÉDIANE
MÉDIANE

Définition

La médiane est une valeur m telle que

$$P(X \leq m) = \frac{1}{2}.$$

La v.a. X a donc une chance sur deux d'être inférieure à sa médiane. La médiane est généralement unique mais on peut fabriquer des exemples où ce n'est pas le cas :



74 / 151

FIGURE 6.33 – Diapositive n°7 – CM lois continues

- Les deux dernières questions relèvent de types de tâches déjà rencontrés et traités en TD : **T-CalcEsp** et **T-CalcVar**. Il faut utiliser les définitions d'espérance et de variance puis effectuer les calculs d'intégrales sur les bons intervalles selon l'expression de f . L'utilisation de l'indicatrice pour la réduction du domaine d'intégration posera peut-être des difficultés à ceux qui ne sont pas à l'aise avec cette fonction (voir résultats des entretiens). D'autant plus qu'en TD les sous-activités de reconnaissance, d'organisation et de traitement ont été prises en charge par l'enseignant.

L'exercice 2 est essentiellement un exercice de cours (propriétés de l'écart-type, lien entre variance et covariance, corrélation) qui porte sur le chapitre 5 : *Indépendance et loi des grands nombres*. Ce chapitre ne fait pas partie de notre objet d'étude mais la première question de l'exercice 2, qui relève de **T-CalcVar**, est une application des formules vues dans le chapitre 3 (variables aléatoires discrètes) et revues dans le chapitre 4 (variables aléatoires continues) (voir diapositive n°8 du CM 6.34 page 296). Cette première question n'a, *a priori*, pas été traitée en TD. Les propriétés à utiliser pour le calcul de l'écart-type n'ont été présentées en cours que pour la variance : $Var(X + a) = Var(X)$ et $Var(aX) = a^2Var(X)$. On peut donc imaginer que cette première question n'a pas forcément été bien réussie.

4.3. ESPÉRANCE. VARIANCE

Définition

Soit X une v.a. de densité f , on définit, espérance, variance et écart-type par

- ① $E[X] = \int_{-\infty}^{+\infty} xf(x)dx$
- ② $Var(X) = E[(X - m)^2] = \int_{-\infty}^{+\infty} (x - m)^2 f(x)dx, \quad m = E[X].$
- ③ $\sigma(X) = \sqrt{Var(X)}.$

RAPPEL : Pour les v.a. discrètes, on avait la formule analogue

$$E[X] = \sum_i x_i P(X = x_i).$$

Comme pour les v.a. discrètes on a

$$E[h(X)] = \int h(x)f(x)dx$$

$$E[aX + b] = aE[X] + b, \quad a, b \in \mathbb{R}$$

$$Var(X) = E[X^2] - E[X]^2$$

$$Var(X + a) = Var(X)$$

$$Var(aX) = a^2 Var(X)$$

75 / 151

FIGURE 6.34 – Diapositive n°8 – CM lois continues

Nous ne détaillons pas l'analyse *a priori* des exercices 3 et 5 qui relèvent du thème des variables aléatoires discrètes.

L'exercice 4 mélange à la fois le chapitre 3 sur les variables aléatoires discrètes, le chapitre 5 sur les théorèmes limites et le chapitre 4 sur les lois continues. Etant donné que nous n'avons pas assisté aux séances de TD associées aux chapitres 3 et 5 nous ne sommes pas en mesure de dire, pour chacune des questions de cet exercice, si les types de tâches ont déjà été rencontrés ou non. Les deux premières questions sont en lien avec les variables aléatoires discrètes, en revanche la troisième question propose **T-Det(a)** déterminer le réel a tel que $P(X < a) =$ nombre donné, avec les variations suivantes : X est une variable aléatoire de loi binomiale, il est indiqué d'utiliser le théorème centrale limite. Nous nous intéressons particulièrement à cette question car elle fait intervenir des ingrédients d'une technique, et donc des types de tâches, relatifs au chapitre 4 sur les lois continues.

- La première question correspond au type de tâches relatif au chapitre 3 sur les variables aléatoires discrètes : **T-DonLoi** pour une loi binomiale, à partir d'un énoncé en langage naturel. Cette question demande un certain travail préliminaire de modélisation et les étudiants sont habitués à modéliser une loi binomiale (c'est quelque chose que l'on retrouve dans les manuels de lycée, notamment dans le chapitre de lois continues).
- La deuxième question qui relève du type de tâches relatif au chapitre 3 sur les variables aléatoires discrètes, **T-DonVar**, n'a pas dû poser de difficulté aux étudiants qui ont su reconnaître la loi binomiale à la question précédente.
- La troisième question en revanche demande d'utiliser le TCL qui a été présenté en cours (chapitre 5) dans une version très générale, afin de déterminer le réel a tel que $P(X < a) = \text{nombre donné (T-Det(a))}$ où X suit une loi binomiale. Dans le cours, à la suite de ce théorème, les étudiants ont vu une proposition qui permet d'approximer la loi binomiale, ils doivent se servir ici de cette proposition afin de transformer X en une variable Y de loi normale $N(240, 144)$.

Les étudiants doivent ensuite transformer l'écriture, c'est-à-dire centrer et réduire la variable. Il s'agit d'un ingrédient de technique, et donc un type de tâche, qu'ils ont rencontré au lycée, qu'ils ont revu en CM mais qui n'a pas été retravaillé en TD (d'après la séance associée au chapitre 4 que nous avons observée). Enfin les étudiants doivent transformer l'expression afin de pouvoir utiliser la table de loi normale centrée réduite et trouver une valeur approchée de $\frac{N-240}{12}$. Il s'agit d'un ingrédient de technique, et donc d'un type de tâche, qui a été vu durant la séance de TD4 mais dont l'organisation et le traitement ont été pris en charge par l'enseignant.

Dans cette troisième question, la reconnaissance de la méthode et des ingrédients d'une technique est à la charge de l'étudiant. Les nouvelles variables aléatoires nécessitent sans doute d'être renommées par les étudiants afin qu'ils s'y retrouvent, par exemple : « Y la VA qui suit $N(240, 144)$ » puis « X la VA qui suit $N(0, 1)$ ».

6.4.3 Analyse de copies d'étudiants

Nous présentons dans ce qui suit l'analyse de la copie de Maud et de celle de Adam en regard des résultats des analyses menées à l'Université (partie 5.3) et de l'analyse *a priori* ci-dessus.

Analyse de la copie de Maud

La copie de Maud se trouve en annexe J.1.

La réponse de Maud à la question de cours est fautive (elle oublie une des conditions). Maud nous a expliqué le jour de la consultation des copies qu'elle n'avait pas révisé le cours pour l'examen. La notation qu'elle utilise pour

écrire l'intersection de deux événements A et B est inhabituelle, il s'agit de : "A ; B".

Maud a plutôt bien réussi l'exercice 1, elle a obtenu la note de 2,5/4. Il s'agit de types de tâches (hormis pour la médiane) déjà rencontrés en TD, il n'y pas d'adaptation à sa charge vis-à-vis de l'exercice similaire traité en TD (ex4 TD4).

- Sa réponse à la première question est correcte. Pour résoudre **T-Det(cf)** elle utilise la propriété de la densité qui est que son intégrale sur le domaine est 1, elle rappelle cette propriété au début de sa réponse. Elle transforme correctement l'intégrale selon l'expression de f (indicatrice).
- Maud n'a pas répondu à la deuxième question (**T-Det(médiane)**), elle nous a dit le jour de la consultation des copies qu'elle ne connaissait pas la définition de la médiane.
- Maud sait calculer l'espérance et la variance, les calculs sont corrects, il n'y pas d'adaptation à sa charge vis-à-vis de ce qu'elle a rencontré en TD. Elle suit la même méthode que celle présentée par l'enseignant en TD, pour le calcul de l'espérance par exemple : réécrire la définition de l'espérance, découper l'intervalle selon la fonction de densité f (lorsqu'elle est définie par morceaux), calculer chacune des intégrales à l'aide de primitives.

Maud n'a pas du tout réussi l'exercice 2. Ce sont pourtant des applications directes des formules du cours mais comme elle nous l'a dit lors de la consultation des copies, elle n'avait pas revu le cours. De plus, elle nous a confirmé que la première question ne correspond pas un type de tâches déjà rencontré ou traité en TD (aucun exercice fait là-dessus). A la première question il semblerait qu'elle ne sache pas que $\sigma(X+a) = \sigma(X)$ (pour X une VA et a un réel).

Maud a très bien réussi l'exercice 4 (5/5). Elle nous a dit lors de la consultation des copies qu'il s'agissait de questions qu'elle avait déjà faites au lycée. Ce que confirme l'enseignant qui nous a expliqué avoir choisi cet exercice d'après un sujet de baccalauréat. Voici quelques éléments remarquables issus de la copie de Maud :

- Concernant les deux premières questions (variables aléatoires discrètes), elle reconnaît la loi binomiale et le justifie correctement, puis elle calcule l'écart-type de la variable sans problème.
- Pour la question avec le TCL, Maud approxime la variable aléatoire S par une loi normale qu'elle nomme Y. Elle calcule ensuite l'espérance et la variance de cette nouvelle variable puis remplace S par Y dans l'expression de la probabilité.
A propos du dernier changement de variables (centrer et réduire la variable Y), Maud justifie sur sa copie : « *on centre et on réduit* ». Elle nomme cette nouvelle variable Z. Elle transforme l'expression, finit le calcul en se servant de la table de la loi normale et obtient le résultat correct pour N.
- Maud interprète correctement le résultat à la quatrième question.

Analyse de la copie de Adam

La copie de Adam se trouve en annexe J.2.

Pour la même raison que Maud, la réponse de Adam à la question de cours est fautive (oubli d'une des conditions). Adam nous a expliqué le jour de la consultation des copies ne pas avoir révisé le cours pour l'examen. Adam écrit dans sa copie d'examen l'intersection de deux événements A et B de la façon suivante : $A|B$. S'agit-il d'une confusion entre conditionnement et intersection ? Nous n'avons pas pu lui poser la question.

L'exercice 1 n'a pas été réussi par Adam (0/4). Voici quelques éléments remarquables issus de la copie de Adam, concernant cet exercice :

- Pour la première question qui relève de **T-Det(cf)**, Adam prend en charge la reconnaissance de la méthode et l'organisation correctement. Il utilise la même technique que celle observée en TD pour répondre (utilisation de la propriété de la densité qui est que son intégrale vaut 1) et semble à l'aise avec l'indicatrice et le découpage de l'intervalle d'intégration. C'est à la toute fin, au moment de résoudre l'équation algébrique $1 = c + \frac{1}{2}$ qu'il fait une erreur de calcul.
- A la seconde question, Adam donne la bonne valeur de la médiane mais ne justifie pas comment il l'a obtenue.
- Pour la troisième question, qui relève de **T-CalcEsp**, Adam ne prend en charge correctement ni la reconnaissance de la méthode, ni l'organisation, ni le traitement. Il semblerait que Adam calcule en fait l'intégrale de f sur son domaine de définition avec la valeur de c obtenue à la première question. Etant donné que la valeur qu'il a obtenue pour c est fautive, il ne retombe pas sur 1 (valeur de l'intégrale d'une densité sur son domaine de définition) mais sur 2.
- Pour répondre à la quatrième question, qui relève de **T-CalcVar**, Adam commence par réécrire la définition de la variance (comme préconisé en TD), il prend donc en charge la sous-activité de reconnaissance de la méthode correctement. Cependant il inverse $E(X)^2$ et $E(X^2)$ dans la formule. Mais c'est surtout dans le traitement de la tâche, pour le calcul de $E(X^2)$, que Adam se trompe. En effet, il écrit :

$$E(X^2) = 1,5 \int_1^0 1 + \int_0^1 X^2$$

au lieu de :

$$E(X^2) = 1,5 \int_{-1}^0 x^2 + \int_0^1 x^3$$

C'est en fait la formule de transfert $E[h(X)] = \int h(x)f(x)dx$ qu'il ne connaît pas et/ou ne sait pas appliquer ici.

A la première question de l'exercice 2 Adam donne la bonne valeur de l'écart-type mais sans donner aucune explication ni justification.

L'exercice 4 est partiellement réussi par Adam (2/5) :

- Il a réussi les deux premières questions qui relèvent du thème des variables aléatoires discrètes, il reconnaît la loi binomiale et le justifie correctement. Pour la question suivante il rappelle la formule littérale de l'écart-type puis effectue ensuite le calcul correctement.
- La troisième question, relève de **T-Det(a)** dans le cas d'une loi binomiale qu'il faut approximer au moyen du TCL. Nous rappelons que nous n'avons pas observé la séance de TD associée au chapitre 5 sur les lois limites, nous ne pouvons donc pas dire si ce type de tâches a déjà été rencontré ou traité, et de quelle façon. Ici Adam ne parle ni de TCL ni d'approximation de loi binomiale mais il utilise la bonne formule de changement de variables $\frac{X-np}{\sqrt{np(1-p)}}$ pour passer d'une variable aléatoire X de loi binomiale à une variable aléatoire (qu'il n'a pas définie) Z de loi normale centrée réduite. Il prend donc en charge correctement une partie de la reconnaissance et de l'organisation mais il s'arrête là dans sa réponse. A-t-il fait le changement similaire avec N dans la formule de la probabilité afin d'obtenir à droite de l'inégalité $N - 240/12$? La phrase suivante « *entre 1 et 1,28* » nous permet de dire qu'il ne sait pas se servir d'une table de loi normale, il a donc des difficultés dans le traitement de cette tâche.

Synthèse

Maud connaît les techniques à utiliser pour les différentes tâches qui lui sont demandées et elle justifie correctement ce qu'elle fait. Adam a des difficultés dans l'organisation et le traitement des tâches qu'il rencontre, par exemple pour **T-CalcEsp** et **T-CalcVar** à l'exercice 1, mais aussi pour **T-Det(a)** dans l'exercice 4.

Finalement, pour Maud et Adam ce sont les questions d'applications directes du cours qui n'ont pas été réussies (question de cours et exercice n°2). Dans l'exercice 2 il y a tout un travail des techniques qui est dévolu aux étudiants car ces notions n'ont, *a priori*, pas été retravaillées en TD sur des exercices. Au lycée, cela n'arriverait pas. Il y a d'une part une question de contrat didactique : ces deux étudiants n'ont pas compris qu'il fallait travailler le cours pour préparer l'examen. Et d'autre part, étant donné que ces notions n'ont pas été *a priori* manipulées en TD, les étudiants sont surpris de retrouver ce genre de tâches proposées à l'examen.

Chapitre 7

Conclusion et perspectives

Ce dernier chapitre présente une synthèse de notre travail et quelques perspectives que nous envisageons dans la continuité de cette thèse.

Nous exposons dans une première partie des éléments de réponse aux questions initialement posées (partie 2.5 page 42). Il ne s'agit pas ici de reprendre dans le détail les résultats de nos analyses qui sont à retrouver tout au long des trois derniers chapitres mais bien de présenter les résultats principaux qui émergent de notre étude.

Dans une seconde partie nous proposons de synthétiser les résultats de notre recherche sur la transition secondaire-supérieur en probabilités sous la forme d'un Modèle Praxéologique de Référence Étendu par la théorie de l'activité et adapté pour la Transition secondaire-supérieur, noté MPR-ET (voir cadre théorique partie 2.4 page 40). Nous présenterons dans une dernière partie de perspectives comment nous envisageons d'utiliser ce MPR-ET comme outil pour accompagner la transition. Avant cela nous présentons, dans une troisième partie, les limites dues à nos choix méthodologiques ou à l'échelle de notre étude que nous pouvons exprimer sur ce travail de recherche.

Enfin, dans une quatrième et dernière partie, nous nous appuyons sur ce MPR-ET pour proposer des pistes à cette recherche, il s'agit à la fois de travaux complémentaires déjà en cours et à la fois de futures possibilités d'action ou de recherche.

7.1 Réponses aux questions initiales

QR.1 : Quel est le savoir à enseigner en probabilités dans le secondaire en classe de terminale scientifique (selon une perspective épistémologique et institutionnelle) ?

L'analyse menée dans le chapitre 4 met en avant différents aspects qui permettent de répondre à cette première question selon une perspective épistémologique et historique. Parmi ces éléments de réponses, il y a la nécessité que les différentes approches des probabilités (classique, fréquentiste et bayésienne)

soient enseignées conjointement et articulées entre elles. Un enseignement qui combinerait ces approches permettrait aux élèves de donner du sens à l'usage des probabilités dans le monde réel mais aussi de réduire les difficultés qu'ils rencontrent dans ces apprentissages.

Le programme officiel de mathématiques pour la classe de terminale scientifique, en vigueur jusqu'en 2020, suit en partie ces recommandations. En effet, il présente l'enseignement des probabilités à travers l'approche classique et l'approche fréquentiste. Un exemple concret est le théorème de Moivre-Laplace, qui permet d'introduire la loi normale et fait le lien entre fréquences et probabilités.

Par ailleurs, l'analyse de ces programmes, pour le thème des probabilités, nous a permis de mettre en évidence les capacités attendues par l'institution. Parmi elles on trouve les capacités suivantes : savoir construire et utiliser un arbre de probabilités (pour le thème des probabilités conditionnelles) ou encore, savoir utiliser la calculatrice pour effectuer des calculs de probabilités (pour le thème des lois continues). Davantage de détails se trouvent partie 4.2 page 56. Les textes des programmes font également référence à la modélisation, à des situations concrètes ou à des problèmes issus de d'autres disciplines.

L'analyse de manuel et de sujets de baccalauréat que nous avons menée ensuite permet de compléter les éléments de réponses à cette première question de recherche QR.1 sur le savoir à enseigner en probabilités pour la classe de terminale scientifique, selon une perspective institutionnelle.

Nous avons produit dans le chapitre 4 deux représentations synthétiques du savoir à enseigner selon les manuels et les sujets de baccalauréat (voir figure 4.7 page 77 et figure 4.14 page 97). Nous invitons ainsi le lecteur à s'y référer pour entrer dans les détails car ici nous synthétisons brièvement les principaux aspects de cette étude.

Les exercices proposés pour le thème des probabilités conditionnelles sont très souvent issus de situations concrètes et réelles. En revanche, pour le thème des lois continues les exercices sont généralement dans un contexte théorique et les situations concrètes plutôt rares. Pour les lois continues, on retrouve des exercices issus des autres disciplines, comme le souhaite le programme officiel de mathématiques de la classe de terminale scientifique. Finalement, ces deux chapitres que nous avons choisis initialement pour des raisons de programme (uniques thèmes communs aux deux curricula), nous ont permis d'étudier la transition suivant des aspects qui s'avèrent différents selon l'un ou l'autre de ces chapitres, qu'il s'agisse de l'activité de modélisation ou de l'utilisation de techniques instrumentées.

Les analyses présentées dans le chapitre 4 page 49 ont également permis de mettre en évidence que de multiples types de tâches et variations font référence à une activité de modélisation probabiliste. Selon la situation, l'utilisation d'un

arbre ou d'un tableau de probabilités peut être pertinente ou non. Dans le cas où un arbre (ou un tableau) pourrait être une aide ou une étape - ingrédient de la technique - il n'est pas toujours suggéré par l'énoncé d'en faire un. Ainsi les changements de registres liés à une activité de modélisation du type passage du registre de la langue naturelle au registre de représentation *arbre* ou *tableau* ne sont pas toujours explicités et sont alors à la charge de l'élève.

En ce qui concerne les lois continues, les registres de représentations de type graphiques de fonctions sont présents dans seulement 10% des exercices, on pourrait imaginer qu'il y ait davantage ce genre de représentations au vu des recommandations du programme officiel.

Les techniques proposées dans les corrigés et savoir-faire pour les lois continues, reposent le plus souvent sur l'utilisation directe de formules (loi exponentielle) ou sur l'utilisation de la calculatrice (calculs avec la loi normale, déterminer des intervalles, etc.), même si des étapes préalables sont souvent nécessaires et donc des adaptations sont à la charge de l'élève (reconnaissance des modalités à mettre en oeuvre, utilisation de résultats issus de questions précédentes, transformation de l'expression).

Les éléments de réponses à cette première question (QR.1) permettent d'anticiper des résultats relatifs à l'activité des élèves, et aux problématiques de transition. C'est dans ce sens que nous abordons notre seconde question de recherche (QR.2).

QR.2 : Quelles sont les caractéristiques de la transition secondaire-supérieur en probabilités pour les étudiants non-spécialistes ? (le cas d'étudiants de première année de biologie)

C'est en appui sur les réponses à la première question de recherche (QR.1) que nous avons mené notre étude en vue de caractériser la transition secondaire-supérieur en probabilités, en termes de praxéologies enseignées, d'activité des étudiants et de pratiques des enseignants.

Dans cet objectif nous avons mené des analyses des données recueillies en classe de terminale scientifique et en première année de biologie, pour les deux thèmes de probabilités qui nous intéressent. Nous présentons ici les principaux éléments de réponses à cette question de recherche. Nous avons produit dans le chapitre 5 deux représentations synthétiques du savoir enseigné en classe de terminale scientifique (voir figure 5.16 page 129 et figure 5.45 page 187) et nous invitons le lecteur à s'y référer s'il souhaite davantage de détails.

Certains éléments dans les pratiques enseignantes, dans les déroulements ou en termes de contrat didactique, nous semblent caractéristiques de la transition secondaire-supérieur en probabilités pour les étudiants non-spécialistes. Dans le secondaire de nombreuses interventions, remarques et aides de l'enseignante sont en lien avec l'examen de fin d'année (baccalauréat) et ses attendus.

Le baccalauréat est un examen national dont les enseignants du secondaire ne connaissent pas au préalable le sujet. Cette pratique est propre au secondaire car à l'Université les enseignants connaissent et conçoivent l'examen.

En ce qui concerne les déroulements, il n'y a plus cette alternance, au sein d'une même séance, entre moment de cours et exercices en autonomie (en groupe ou individuel), il s'agit sans doute d'une nouveauté pour au moins certains étudiants. A l'Université désormais le cours est d'un seul bloc et sous forme de diapositives projetées au tableau (un cours magistral pour tout un thème suivi d'une séance de TD pour ce thème).

De la même manière, le polycopié de cours à trous est une modalité de déroulement propre au lycée (d'après nos observations) qui a aussi un impact sur l'activité des élèves au secondaire. Dans certains pays il y a des expérimentations (voir la recherche en Allemagne de Feudel et Panse (2021)) menées dans le supérieur, s'appuyant sur des polycopiés à trous, qui semblent fructueuses pour aider les étudiants dans la prise de notes mais ce n'est pas une pratique que nous avons déjà observée en France à l'Université.

Un autre élément qui nous semble caractéristique de cette transition est que, durant les cours magistraux à l'Université, les calculs de probabilités des exemples sont peu détaillés. En effet, l'enseignant choisit de laisser à la charge des étudiants les tâches qu'ils ont déjà rencontrées, soit au premier semestre soit dans le secondaire. Or, d'après nos observations en classe de terminale scientifique les techniques d'intégration ont par exemple été très peu mises en fonctionnement par les élèves pour les contenus de probabilités.

Dans ce paragraphe nous proposons de répondre à la question de recherche QR.2 avec des éléments plus spécifiques aux probabilités conditionnelles.

La disparition de la représentation par les arbres de probabilités, dans l'enseignement universitaire, pour ces contenus, est un premier point saillant de cette transition. En effet, dans le secondaire l'arbre a une place primordiale en tant qu'aide procédurale dans l'activité de traitement, en tant qu'intermédiaire dans l'activité de modélisation mais aussi comme élément de la technique lors du calcul d'une probabilité.

Les changements de notations pour des objets ou concepts déjà existants dans le secondaire sont une caractéristique de cette transition. Un exemple criant est la nouvelle notation de la probabilité conditionnelle $P(A|B)$, justifiée par l'enseignant qui explique qu'elle est plus facilement mémorisable pour un public de non-spécialistes.

Les nouveautés en termes de contenus et de types de tâches, pour ce cours, par rapport à ce que nous avons observé en terminale scientifique sont assez peu nombreuses. On peut tout de même citer la nouvelle technique qui utilise la formule de Bayes et pour laquelle l'enseignant considère qu'elle a été vue dans le secondaire, alors que celle-ci n'est pas au programme de lycée bien qu'elle puisse apparaître parfois dans des manuels ou au gré des exercices.

Nous exposons, dans ce troisième paragraphe, des éléments de réponses à la question de recherche QR.2 en lien cette fois-ci avec les contenus de lois

continues. Nous ne rentrons pas dans le détail des types de tâches, variations et techniques rencontrés dans ce cours mais nous présentons ce qui nous semble caractéristique de la transition secondaire-supérieur. Nous laissons le lecteur se référer au chapitre 5 s'il souhaite davantage de précisions.

Les fonctions de densité continues par morceaux sont une grande nouveauté pour les étudiants qui sortent du lycée puisque dans le secondaire seules les fonctions de densité continues sont au programme. Ces fonctions sont généralement définies au moyen d'une fonction indicatrice.

En termes de nouveautés de types de tâches il y a principalement T-Det(cf) : déterminer les valeurs possibles d'une constante c qui apparaît dans l'expression de f et T-CalcVar : calculer la variance ou l'écart-type d'une variable aléatoire continue.

La disparition de la technique utilisant la calculatrice (lois normales, lois exponentielles, autres calculs d'intégrales, etc) au profit de tables de lois normales ou de techniques d'intégration est caractéristique de cet enseignement à l'Université. En effet, dans le secondaire il s'agit bien souvent d'utilisation directe de formules (loi exponentielle), de techniques instrumentées (loi normale) ou bien de techniques géométriques (loi uniforme).

De façon générale, les techniques attendues et présentées ne sont plus celles que nous avons observées dans le secondaire, d'une part car les fonctions de densité sont désormais continues par morceaux et d'autre part car les techniques s'appuient sur des calculs et donc des changements de variables et des intégrations par parties qui ne sont pas au programme de la classe de terminale scientifique.

Enfin, voici des éléments qui nous semblent caractéristiques de la transition secondaire-supérieur du point de vue de la modélisation.

Dans le secondaire l'activité de modélisation, pour le thème des probabilités conditionnelles, se fait très souvent au moyen d'un arbre de probabilités. Cette représentation ne faisant plus partie du savoir enseigné à l'Université, la modélisation désormais ne s'appuie plus sur les arbres de probabilités.

À l'Université l'enseignant que nous avons observé accorde beaucoup d'importance à l'étape de modélisation. Les situations probabilistes sont à présent modélisées par deux nouveaux concepts : Ω l'univers et P la mesure de probabilité. Aussi, l'ensemble Ω doit être décrit préalablement au début de chaque exercice selon les recommandations de l'enseignant qui insiste sur l'importance du choix du modèle. L'enseignant fait, durant le cours, de nombreuses références à des situations de modélisation s'appuyant sur l'analyse de données afin notamment de préparer les étudiants aux cours de statistiques qui leur seront dispensés dans les années suivantes.

Cependant, au sein des exercices et des séances d'exercices, nous n'avons que très peu retrouvé ces notions de statistiques et ces tâches de modélisation à la charge des étudiants.

Après avoir caractérisé la transition secondaire-supérieur en probabilités en termes de praxéologies enseignées, d'activité des étudiants et de pratiques des enseignants nous nous sommes questionné sur les spécificités des étudiants non-

spécialistes débutants à l'Université dans l'apprentissage des probabilités, c'est l'objet de notre troisième question de recherche (QR.3).

QR.3 : Quelles sont les spécificités des étudiants non-spécialistes dans l'apprentissage des probabilités? (le cas d'étudiants de première année de biologie)

Les éléments de réponses à cette question (QR.3) s'appuient sur les résultats des analyses qui sont détaillées dans le chapitre 6, nous ne présentons ici que les résultats principaux qui ressortent de ces analyses.

Le test que nous avons proposé à des lycéens et à des étudiants de première année de biologie à l'Université nous a permis de recueillir des éléments à propos des sous-activités (reconnaissance, traitement, organisation) qu'ils sont capables de prendre en charge et des techniques qu'ils savent mettre en oeuvre suite à l'enseignement de probabilités qu'ils ont reçu dans le secondaire. De façon générale, ils utilisent avec aisance l'arbre de probabilités. Les étudiants savent identifier des intersections ou des probabilités conditionnelles lorsque l'énoncé est en langage naturel (reconnaissance de "et" ou "sachant que" dans les exercices du test). Face à un exercice plus exigeant en termes de modélisation, les étudiants rencontrent des difficultés dans l'identification des événements en jeu et de la probabilité recherchée. Ils rencontrent également des difficultés à mobiliser les connaissances qui viennent d'être vues pour des tâches qu'ils n'ont encore jamais rencontrées.

Le questionnaire soumis à des étudiants de première année de biologie nous a permis de faire un premier état des lieux de leur ressenti quant aux contenus et au déroulement du module de probabilités du second semestre. Nous relevons que ces étudiants déclarent trouver difficiles les probabilités enseignées dans ce module.

Les entretiens menés avec certains d'entre eux et les copies d'examens que nous avons pu recueillir et analyser nous permettent de préciser les éléments de réponses à la question de recherche QR.3 sur les spécificités des étudiants non-spécialistes pour le thème des probabilités.

Un premier élément saillant qui est ressorti très rapidement de ces entretiens est que les étudiants sont perturbés par certains termes de vocabulaire et certaines nouvelles notations du cours, notamment $P(A|B)$.

Concernant les déroulements, les séances de cours magistraux posent des difficultés aux étudiants. Ils seraient selon eux très denses, trop théoriques et trop rapides. Le fait qu'il s'agisse de diapositives projetées au tableau est une nouveauté pour eux. Les étudiants reprochent au cours magistral d'être en décalage avec ce que l'on attend d'eux dans les exercices proposés en séances d'exercices (TD). De façon générale ils sont perturbés par les changements, qu'il s'agisse des attendus, des notations ou des déroulements.

Pour le thème des probabilités conditionnelles, les étudiants déclarent avoir

l'habitude de travailler avec des arbres de probabilités et éprouver des difficultés à s'en passer dans le cadre d'exercices qui s'y prêtent.

Dans le cas des lois continues, ils déclarent éprouver des difficultés à identifier ce qu'il faut faire, à reconnaître la loi, ou encore à utiliser correctement la fonction indicatrice. Certains d'entre eux ne savent pas se servir de la table de loi normale à la suite du TD qui présente cette nouvelle technique.

Bien que les types de tâches proposés durant les TD ou à l'examen soient très similaires à ceux rencontrés au lycée, les étudiants ont des difficultés à utiliser des techniques différentes de celles apprises dans le secondaire, par exemple utiliser une table pour la loi normale ou se passer des arbres pour les probabilités conditionnelles. Il s'agit d'un point important concernant la transition secondaire-supérieur dans le cas qui nous intéresse des probabilités.

Au-delà des réponses synthétisées ci-dessus aux questions posées initialement, cette étude nous a conduit à nous demander quels outils théoriques peuvent nous permettre de mieux comprendre la transition? Aussi, ce travail nous amène à faire des propositions théoriques et méthodologiques particulières et que l'on peut appliquer au cas des probabilités pour les non-spécialistes. Il ne s'agit pas uniquement d'un outil pour représenter la transition telle qu'elle est, mais d'un outil qui permet de faire figurer des propositions en vue de construire un enseignement innovant qui réduirait les difficultés des étudiants à la transition.

7.2 Apport théorique pour l'étude de la transition et proposition pour le cas des probabilités pour les non-spécialistes

L'objectif de cette partie est de proposer un Modèle Praxéologique de Référence Étendu par la théorie de l'activité et adapté pour la Transition secondaire-supérieur, noté MPR-ET, dans le cas ici des probabilités pour les non-spécialistes. Cet outil que nous avons défini partie 2.4 (page 40) permet de mettre en évidence les évolutions entre le lycée et l'Université tout en faisant des propositions en vue de palier aux difficultés des étudiants rencontrées lors de la transition.

Nous ne visons pas l'exhaustivité des praxéologies à enseigner ni des adaptations car nous nous appuyons sur les éléments dont nous disposons suite aux observations et par l'analyse des données que nous avons recueillies auprès des institutions et des étudiants.

Dans un premier temps nous exposons la méthodologie de conception de notre MPR-ET puis dans une seconde sous-partie nous le présentons et l'analysons.

7.2.1 Méthodologie de construction du MPR-ET

Le MPR-ET que nous construisons va permettre dans un premier temps de synthétiser les réponses aux questions suivantes :

- Quelles organisations mathématiques présentes au lycée sont reprises et adaptées à l'Université ?
- Quelles techniques, nouvelles ou non, vise-t-on pour les étudiants ?
- Quelles sont les continuités ou ruptures entre institutions en terme de types de tâches et d'adaptations à la charge des étudiants ?
- Comment décrire les attentes institutionnelles envers les étudiants à la fin du cours de probabilités de première année de biologie ?

A partir des analyses menées et des résultats présentés tout le long de cette étude nous dressons notre MPR-ET sous la forme d'un schéma selon la méthodologie suivante.

Nous représentons les organisations mathématiques pour les deux thèmes de probabilités qui nous intéressent, à savoir, les probabilités conditionnelles et les lois continues. Il s'agit d'organisations mathématiques qui sont soit enseignées à l'Université, soit que nous recommandons d'enseigner d'après les analyses menées, soit qui étaient présentes dans le savoir enseigné du lycée et qui ne sont plus à l'Université.

De ces organisations mathématiques nous ne faisons apparaître dans le MPR-ET que le bloc *pratico-technique* (Chevallard, 2003) c'est-à-dire les types de tâches T (dans des encadrés de forme rectangulaire à trait plein) et les techniques τ (dans des encadrés de forme rectangulaire en pointillés).

Pour les probabilités conditionnelles nous présentons les types de tâches dans un seul et unique encadré du même nom mais en ce qui concerne les lois continues, nous proposons 4 encadrés qui représentent les contenus du chapitre : *loi uniforme, loi exponentielle, lois quelconques* et *lois normales*.

Sous ces encadrés décrivant la *praxis*, nous inscrivons, dans un encadré à coins arrondis, les éléments modifiant l'activité de l'étudiant ou provoquant des adaptations à la charge des étudiants. Il s'agit soit de variations de types de tâches, soit de notations nouvelles, soit d'éléments concernant le contexte de la situation ou encore le format des données.

Voici comment en détails nous allons compléter ces différents encadrés et ainsi construire le schéma du MPR-ET :

- en **vert**, les types de tâches identiques au lycée et à l'Université

7.2. Apport théorique pour l'étude de la transition et proposition pour le cas des probabilités pour les non-spécialistes

- en **bleu**, les types de tâches qui présentent des variations nouvelles vis-à-vis du lycée mais sans que ces variations n'impliquent de techniques jamais vues dans le secondaire. Dans ce cas-là, la nouvelle variation apparaît en bleu.
- en **rouge**, les types de tâches accomplis par des techniques différentes de celles vues au lycée ou bien les types de tâches dont une nouvelle variation implique la mise en oeuvre de nouvelles techniques. Dans ce cas-là, la nouvelle technique (ou la variation) apparaît en rouge.
- en **noir**, les types de tâches jamais rencontrés au lycée. Dans ce cas-là, la nouvelle technique apparaît en noir.
- en **orange**, les techniques qui existaient au lycée mais qui ne font plus partie du savoir enseigné à l'Université.
- en **violet**, les types de tâches, techniques et éléments provoquant des adaptations que l'on retrouve peu (voire pas du tout), à la charge des étudiants et dont il serait important qu'ils aient une place plus importante dans l'enseignement à l'Université.

Le MPR-ET ainsi construit est une contribution particulière de ce travail de recherche. Il s'agit d'un outil théorique permettant de mieux comprendre la transition, ici pour le cas des probabilités pour les non-spécialistes, et de faire des propositions en vue d'améliorer celle-ci.

7.2.2 MPR-ET pour les probabilités conditionnelles et les lois continues, le cas des étudiants de biologie

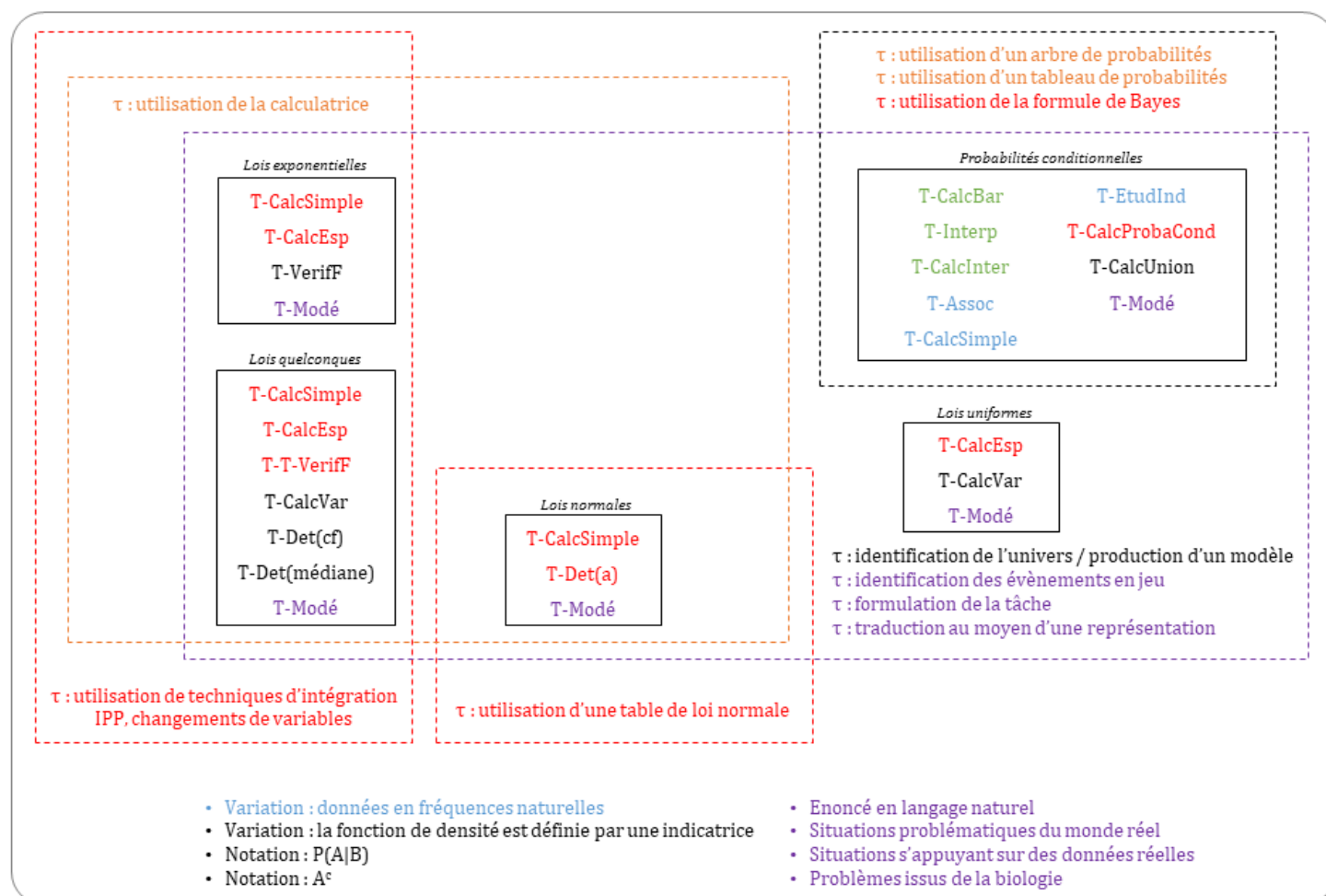


FIGURE 7.1 – MPR-ET pour les probabilités conditionnelles et lois continues

Le MPR-ET présenté ci-dessus figure 7.1, met en lumière de nombreux résultats que nous exposons et détaillons dans la suite.

La majorité des types de tâches relatifs au thème des probabilités conditionnelles, proposés et enseignés à l'Université auprès des étudiants de biologie ne sont pas nouveaux : T-CalculBar, T-Interp, T-CalculInter, T-Assoc, T-CalculSimple, T-EtudInd et T-CalculProbaCond. Ces derniers s'accomplissent au moyen de techniques que les étudiants ont déjà rencontrées et pratiquées dans le secondaire hormis pour T-CalculProbaCond où une nouvelle technique est désormais proposée par l'institution, il s'agit d'utiliser la formule de Bayes.

Pour les types de tâches précédemment mentionnés, les techniques du lycée s'appuyant sur l'utilisation d'un arbre ou d'un tableau de probabilités ne font plus

partie désormais du savoir enseigné ni des attendus en termes de procédures de résolution. Les étudiants sont pourtant très à l'aise avec la représentation de situations probabilistes par un arbre (voir résultats de l'analyse du test partie 6.1.4). Les étudiants avec lesquels nous nous sommes entretenus expliquent rencontrer des difficultés à se passer de cette représentation (voir partie 6.3.1). En effet, comme nous l'avons observé et décrit dans le chapitre 5, l'arbre de probabilités intervient dans l'enseignement secondaire comme aide procédurale dans l'activité de traitement mais également comme élément de technique et de technologie lors du calcul de probabilités.

Comme nous l'avons déjà mentionné partie 1.4, des auteurs comme Martignon et Wassner (2002), expliquent que l'utilisation de représentations permet une meilleure compréhension des problèmes par les étudiants et, concernant les arbres de probabilités en particulier, Diaz et De la Fuente (2007) ajoute qu'un intérêt de leur utilisation est qu'ils permettent de faire apparaître une grande quantité d'informations grâce à leur lisibilité. Ainsi nous recommandons l'utilisation de ces représentations dans l'enseignement des probabilités conditionnelles. Il faut cependant rester vigilant à ce que l'utilisation de celles-ci ne fasse pas perdre de sens aux concepts sous-jacents (formule des probabilités totales ou définition d'une probabilité conditionnelle par exemple).

La variation de types de tâches *les données sont en fréquences naturelles* est très présente dans les exercices proposés à l'Université. Nous l'avons relevé pour les types de tâches suivants : T-Assoc, T-CalcSimple, T-EtudInd, T-CalcProbaCond. Il s'agit d'une variation de types de tâches que nous n'avons pas rencontrée au cours des observations menées en classe de terminale scientifique (voir chapitre 5). Nous avons en revanche identifié dans le manuel de terminale (voir chapitre 4) le type de tâche T-CalcSimpleFN : *calculer la probabilité d'un évènement simple, à partir d'un énoncé en langage naturel, étant donné que les données sont en fréquences naturelles*. A l'Université, T-CalcSimpleFN apparaît en tant qu'ingrédient de technique permettant d'accomplir les types de tâches qui présentent la variation *les données sont en fréquences naturelles*.

Comme nous l'avons précédemment constaté dans le chapitre 5 (partie 5.3), de nouvelles notations sont apparues à l'Université pour des objets mathématiques déjà rencontrés au lycée. Il s'agit de la notation $P(A|B)$ qui remplace $P_B(A)$ et de A^c au lieu de \bar{A} . Les étudiants que nous avons interrogés (voir chapitre 6) nous ont indiqué que ces notations étaient perturbantes et cause de confusions. Ainsi nous rejoignons Totohasina (1992) (voir partie 1.4) qui explique que la notation $P(B|A)$ pourrait laisser penser - à tort - qu'il s'agit d'une fonction P appliquée à deux variables B et A , et recommandons d'utiliser la notation indicielle $P_A(B)$.

Le MPR-ET met en évidence, pour le thème des lois continues, que la majorité des types de tâches proposés à l'Université dans ce cours ne sont pas nouveaux. En revanche, ces types de tâches s'accomplissent désormais au moyen de techniques très peu mises en oeuvre dans le secondaire.

En effet, comme nous l'avons vu dans le chapitre 4, les techniques associées aux types de tâches T-CalcSimple, T-VerifF, T-Det(a) et T-CalcEsp, proposées au lycée reposent le plus souvent sur l'utilisation de la calculatrice lorsqu'il s'agit de lois normales ou de lois quelconques dès lors que la fonction de densité est connue. Il en va de même concernant la loi uniforme et la loi exponentielle pour lesquelles, en classe de terminale, les techniques reposent essentiellement sur l'utilisation directe de formules.

A l'Université ces types de tâches sont désormais accomplis suivant des techniques géométriques (quand la fonction de densité le permet) ou des techniques d'intégration, il peut s'agir de reconnaissance de primitives, de changement de variables ou d'intégrations par parties (voir chapitre 5). T-CalcEsp dans le cas d'une loi uniforme est traité avec une technique géométrique dans le secondaire (aire d'un trapèze) alors qu'à l'Université deux techniques sont proposées, une technique d'intégration classique (primitive) et la technique géométrique.

Ainsi la calculatrice, si présente dans le secondaire, n'a plus sa place dans le supérieur et la technique instrumentée associée ne fait plus partie du savoir enseigné. La table de loi normale apparaît alors comme un outil indispensable pour les types de tâches T-CalcSimple et T-Det(a). Il s'agit de la seule technique proposée aux étudiants dans ce cours pour les exercices faisant intervenir la loi normale. Cette remarque nous questionne vis-à-vis de la transition secondaire-supérieur car ces tables de lois normales ne font plus partie des programmes de lycée depuis de nombreuses années. On peut d'ailleurs lire à ce titre dans le document Ressources (2012) d'accompagnement des programmes de la classe de terminale : « *avec les calculatrices, ces tables ont aujourd'hui perdu leur intérêt* »

A l'Université où se déroulent nos observations, les étudiants de biologie ne font pas de programmation en parallèle du module de probabilités et n'ont pas de travaux pratiques associés à cet enseignement de probabilités. Ces étudiants n'ont donc pas accès aux logiciels spécifiques aux mathématiques qui auraient pu palier à l'absence de la calculatrice. On pourrait imaginer que les étudiants aient accès dans leur formation à des outils numériques de calculs remplaçant ainsi la calculatrice et permettant tout de même de pratiquer les calculs de probabilités en séances de travaux dirigés.

Plusieurs nouveaux types de tâches pour les contenus de lois à densité sont proposés à l'Université, des analyses menées au chapitre 5 (voir partie 5.3.2) nous relevons : T-CalcVar, T-Det(cf) et T-Det(médiane) dans le cas de lois quelconques, T-VerifF pour la loi exponentielle et T-CalcVar qui est rapidement évoqué pour la loi uniforme en CM.

T-CalcVar, T-Det(cf) et T-VerifF nécessitent de mettre en oeuvre des techniques d'intégration, nouvelles donc d'après ce que nous avons présenté plus haut. Il s'agit ici de tâches complexes pour lesquelles plusieurs sous-activités s'imbriquent,

il y a la reconnaissance de la tâche, l'organisation (fonction définie par morceaux par une indicatrice) et le traitement (techniques d'intégrations : intégrations par parties, changement de variables, etc.).

Nous l'avons évoqué rapidement mais une grande nouveauté de l'Université pour les contenus de lois à densité sont les fonctions indicatrices qui définissent les fonctions de densité. Face à des tâches relevant du type de tâche T-Det(cf) avec une fonction indicatrice, les étudiants que nous avons pu interroger (chapitre 6) expriment des difficultés à identifier ce qu'il faut faire (c'est-à-dire la reconnaissance de la tâche) puis à manipuler et utiliser correctement la fonction indicatrice. La complexité de lier une tâche à un type de tâche est une caractéristique de la transition secondaire-supérieur.

En appui sur les analyses menées tout au long de ce travail de thèse, nous recommandons d'accorder une place plus importante à la modélisation, notamment dans le champ des lois à densité où la modélisation reste anecdotique dans les séances que nous avons pu observer. En effet, comme nous l'avons vu dans le chapitre 1, des chercheurs, comme par exemple Chiel et al. (2010), ont établi que la modélisation pourrait permettre de remédier aux difficultés que de nombreux étudiants non-spécialistes rencontrent en mathématiques. Il pourrait ainsi être intéressant de repenser les mathématiques enseignées à ces étudiants en tant qu'outils de modélisation pour l'étude de problèmes issus de la biologie et d'imaginer un enseignement commun mathématiques/biologie.

De plus, d'après ce que nous avons présenté partie 4.1.3 il nous semble important que les différentes approches des probabilités (classique, fréquentiste, bayésienne) soient enseignées et articulées entre elles et l'approche par la modélisation est une façon de le faire comme l'explique Prodromou (2014). Cela aiderait à une meilleure compréhension des phénomènes aléatoires et permettrait aux étudiants de donner du sens à l'usage des probabilités dans le champ de la biologie.

Nous recommandons donc d'inclure dans cet enseignement universitaire des types de tâches (notés T-Modé) visant une activité de modélisation probabiliste complète, dans le sens que nous avons défini partie 5.2.3. Pour rappel voici les étapes de la modélisation probabiliste telle que nous l'avons définie dans le chapitre 5 :

- (1) formulation de la tâche
- (2) production d'une représentation mathématique (modèle)
- (3) traduction des objets et des relations en mathématiques incluant éventuellement la représentation par un arbre ou un tableau de probabilités
- (4) utilisation de méthodes mathématiques pour obtenir un résultat mathématique
- (5) interprétation des résultats
- (6) validation du modèle

Il serait donc intéressant de proposer par exemple des exercices en langage

naturel présentant une situation problématique issue de la biologie et de laisser à la charge des étudiants chacune des étapes de la modélisation (formulation de la tâche, production du modèle, identification des événements en jeu, etc). On pourrait imaginer des situations problèmes pour lesquelles les étudiants participeraient à la construction de modèles mathématiques. Aussi, selon nous, l'utilisation de représentations, comme les arbres dans le cas des probabilités conditionnelles, doit être envisagée comme intermédiaire dans l'activité de modélisation (à partir d'un énoncé en langage naturel, l'arbre permet de traduire les données avant de commencer à répondre aux questions).

Il pourrait également s'agir de proposer des situations issues de la biologie s'appuyant sur des données réelles et ainsi de présenter des exercices dans lesquels peuvent se combiner une perspective centrée sur les données et une perspective de modélisation. Ainsi nous rejoignons la recommandation de Bakker et al. (2018) qui est de ne pas enseigner les probabilités sans données.

7.3 Limites de cette étude

Dans cette partie, nous exposons certaines des limites dues à nos choix méthodologiques ou à l'échelle de notre étude que nous pouvons exprimer sur ce travail de recherche.

Nous avons réalisé l'analyse d'un seul manuel du secondaire. Comme nous l'avons mentionné dans le chapitre 4, c'est dans l'objectif d'obtenir un ensemble de praxéologies représentatives de ce qui est attendu à la fin du lycée que nous avons souhaité analyser ce manuel et, plutôt que d'étudier plusieurs manuels de terminale scientifique, nous avons préféré compléter cette analyse par celle de sujets de baccalauréat.

Nous n'avons observé qu'une seule enseignante de lycée expérimentée et experte sur le plan didactique et qu'un seul enseignement à l'Université, il pourrait être envisagé d'observer d'autres enseignants en particulier dans différents lycées et dans d'autres Universités en première année de biologie.

En complément des analyses de séances en classe nous aurions pu compléter ce travail en réalisant des entretiens détaillés avec l'enseignant de probabilités à l'Université afin de donner à voir plus expressément les pratiques enseignantes et les attentes pour ce cours envers les étudiants. Notre recherche présentée ici tient compte des déroulements. Ce n'est pas une recherche initialement axée sur les pratiques des enseignants bien qu'il s'agisse de futures pistes de travail (voir partie 7.4 suivante).

Concernant les limites d'échelle du test proposé et analysé dans le chapitre 6, partie 6.1 page 255, les résultats présentés à la suite de l'analyse des productions d'élèves et d'étudiants ne sont assurément pas représentatifs de l'ensemble des élèves de terminale scientifique ni des étudiants de biologie étant donné que nous

n'avons pu proposer ce test qu'à une cinquantaine d'individus. Il pourrait être envisagé d'étendre la proposition de ce test à davantage d'étudiants de biologie débutants à l'Université. Concernant les choix méthodologiques, le protocole d'analyse *a priori* que nous avons suivi (ie en regard des analyses de séances en classe, partie 5.2) a des limites car les étudiants de l'Université à qui nous avons proposé le test proviennent pour la plupart de terminale scientifique mais n'ont pas assisté, *a priori*, en terminale aux mêmes séances de cours que nous présentons dans le chapitre 5.

Une autre limite résultant de l'échelle de notre étude concerne les analyses des entretiens menés avec des étudiants de biologie. En effet, ces résultats ne sont pas nécessairement caractéristiques des étudiants de biologie car nous n'avons interrogé que quatre étudiants.

Enfin, nous souhaitons rappeler que tous les étudiants de biologie qui ont participé à cette étude, qu'il s'agisse de ceux ayant répondu au test, au questionnaire, ou de ceux avec lesquels nous avons mené des entretiens, ont tous été volontaires et ne sont donc pas obligatoirement représentatifs de l'ensemble des étudiants de biologie, il s'agit d'un biais de cette étude.

Nous présentons dans une dernière partie des possibilités d'actions et autres perspectives qui pourraient permettre de lever en partie les limites précédemment évoquées.

7.4 Perspectives

Ce MPR-ET a permis de synthétiser l'ensemble des résultats de ce travail de recherche et donne également des pistes pour envisager des perspectives à cette étude. Certains travaux complémentaires sont déjà en cours, nous les présentons dans une première sous-partie.

7.4.1 Travaux complémentaires en cours

Nous souhaitons poursuivre ce travail en approfondissant la compréhension des besoins mathématiques des biologistes. Pour cela nous avons d'ores-et-déjà réalisé un entretien avec un chercheur en biostatistiques qui enseigne les statistiques appliquées aux étudiants de biologie en fin de licence. Nous sommes en train de terminer d'analyser cet entretien que nous souhaitons compléter avec d'autres entretiens auprès d'enseignants-chercheurs en biologie. Ce travail pourra conduire à compléter et affiner le MPR-ET, en particulier les éléments qui ne sont pas présents à l'Université et qui pourraient utilement y être introduits (en violet dans le schéma).

Un autre projet dans la continuité de ce travail de thèse est de concevoir, en appui sur le MPR-ET, des exercices en ligne qui peuvent être utiles pour ces

étudiants de biologie. Ce travail a déjà été entamé, nous avons en effet conçu et implémenté quelques exercices sur le logiciel BRAISE (<http://braise.univ-rennes1.fr/braise.cgi>) qui est une Base RAISONnée d'EXercices de mathématiques conçue par une équipe d'enseignants-chercheurs de l'Université Rennes 1. Ces exercices sont accompagnés de modules d'aides à la résolution (indications, éléments de cours, méthodes et techniques) et de modules pour conclure (éléments de solutions, idées à retenir). Nous avons fait le choix de proposer des exercices qui présentent une forte modélisation laissée à la charge de l'étudiant et sont destinés aux étudiants de première année de biologie. Un exemple d'un tel exercice se trouve figure 7.2. Les exercices sur le thème des probabilités conditionnelles contiennent volontairement des indications qui s'appuient sur des arbres de probabilités afin d'aider les étudiants dans leurs activités de modélisation.

The screenshot displays the BRAISE software interface. At the top, it reads 'Braise Base raisonnée d'exercices de mathématiques Probabilités discrètes' and 'UNIVERSITÉ DE RENNES 1'. The main content area is titled 'Exercice numéro 4.8' and contains the following text under the heading 'Énoncé':

Une étude menée sur des souris possédant la même anomalie génétique a démontré que chez 30 % d'entre elles cette anomalie provoquait des lésions irréversibles ; et parmi ces 30 %, 5 % des souris mourraient prématurément des conséquences de cette anomalie. Quelle est la probabilité que l'anomalie génétique provoquant des lésions irréversibles mène à la mort prématurée de la souris ? Quelle est la probabilité que l'anomalie génétique n'entraîne pas la mort prématurée de la souris ? Que pensez-vous de la dangerosité de cette anomalie génétique ?

Below the text, there are two sections of buttons:

- Aides à la résolution**: Three buttons labeled 'Éléments de cours', 'Méthodes et techniques', and 'Indications'.
- Pour conclure**: Two buttons labeled 'Éléments de solution' and 'Idées à retenir'.

The left sidebar contains a list of navigation options: 'S'identifier', 'choix d'exercices par mots clés', 'choix d'exercices parmi ceux déjà consultés', 'la liste des exercices du chapitre', 'Retour aux résultats de votre recherche', and 'Caractéristiques de l'exercice'. At the bottom of the sidebar, there is a dropdown menu labeled 'Choisir un chapitre' with 'Probabilités discrètes' selected. The footer of the interface includes links for 'Mode d'emploi', 'Préférences', 'Accès prof.', 'Auteurs', and 'À propos'.

FIGURE 7.2 – Exemple d'un exercice et de son environnement BRAISE

Ces exercices ont été testés lors de plusieurs séances de travaux pratiques sur ordinateurs. Actuellement nous terminons d'analyser les données récoltées lors de ces séances, ces données sont de plusieurs sortes : un pré-test distribué au début du semestre, des grilles d'utilisation permettant d'avoir des éléments concernant le raisonnement suivi et les difficultés qu'ils ont rencontrées et enfin leurs réponses à un questionnaire en ligne ayant pour but d'obtenir un retour général de leur part sur BRAISE et leur usage de ce logiciel. Ces analyses vont nous permettre d'une part d'approfondir notre compréhension des difficultés des étudiants et des apports possibles de la modélisation ; d'autre part, d'affiner la conception de ces exercices en ligne.

7.4.2 Travaux envisagés pour la suite

Le travail mené durant cette thèse nous a permis de produire le MPR-ET précédemment décrit. Celui-ci va jouer un rôle puissant dans la suite de cette recherche car nous avons comme projet de concevoir un enseignement global et innovant qui se base sur ce MPR-ET et qui utiliserait les exercices conçus pour la plateforme BRAISE.

Dans ce sens, nous souhaitons approfondir le concept de MPR-ET, en considérant davantage le bloc *logos* et en portant une attention particulière aux éléments de technologies attendus et proposés par les institutions.

Nous envisageons également de poursuivre nos recherches autour des pratiques des enseignants du supérieur. En effet, les enseignants qui interviennent en mathématiques auprès des étudiants de biologie de première année ne sont pas forcément les mêmes en probabilités et en analyse, en CM et en TD. De plus, comme nous l'avons précédemment décrit (partie 1.1), les difficultés à la transition peuvent être dues aux nouvelles attentes liées aux pratiques expertes des mathématiciens. Il pourrait par exemple s'agir de filmer plusieurs enseignants ou enseignants-chercheurs, dans plusieurs universités.

Enfin, devant le changement majeur opéré récemment dans les programmes de lycée, il nous semble important de considérer l'impact de ces nouveaux programmes sur la transition. Cette perspective sera envisagée dans nos futures recherches. En effet, parmi les biologistes, certains auront suivi en terminale l'option "Spécialité maths", et d'autres l'option "Maths complémentaires", on peut donc supposer que la problématique de la transition sera très différente selon qu'ils aient suivi l'une ou l'autre de ces options.

D'autre part, soulignons que dans le programme de la nouvelle option "Maths complémentaires" il y a un intitulé "Inférence Bayésienne", aussi dans ce cours la formule de Bayes apparaît explicitement. A l'opposé, dans le programme de l'option "Spécialité maths", les lois continues ne sont plus travaillées. Il serait donc très intéressant d'analyser les conséquences de la réforme du lycée avec un focus sur le cas des probabilités.

Références

- Amrani, H., & Zaki, M. (2015). Student's conceptual difficulties with respect to the notion of random variable. *International Journal of Education, Learning and Development*, 3(9), 65–81.
- Artigue, M. (2004). *Le défi de la transition secondaire-supérieur : que peuvent nous apporter les recherches didactiques et les innovations développées dans ce domaine ?* (Communication présentée au premier congrès Canada-France de sciences mathématiques, Toulouse)
- Bakker, A., Hahn, C., Kazak, S., & Pratt, D. (2018). Research on probability and statistics education. In T. Dreyfus, M. Artigue, D. Potari, S. Prediger, K. Ruthven (Ed.), *Developing Research in Mathematics Education : twenty years of communication, cooperation and collaboration in Europe* (pp. 46–59). Abingdon, UK : Routledge.
- Barbazo, E., Barnet, C., Baheux, M., Castagnos, N., Cigana, M., Daniel, A., ... Racadot, C. (2016). *Mathématiques Barbazo Tle S spécifique*. Hachette Education.
- Batanero, C., Chernoff, E.-J., Engel, J., Lee, H., & Sanchez, E. (2016). *Research on teaching and learning probability*. Springer International Publishing.
- Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability. In Jones G.A. (Ed.), *Exploring probability in school. Mathematics Education Library* (Vol. 40). Boston : Springer.
- Batanero, C., Lopez-Martin, M., Arteaga, P., & Gea, M. (2018). Characterizing probability problems posed in university entrance tests in Andalusia. In C. Batanero and E. J. Chernoff (Ed.), *Teaching and learning stochastic* (pp. 103–123). Cham : Springer.
- Batanero, C., Tauber, L., & Sanchez, V. (2004). Students' reasoning about the normal distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 257–276). Dordrecht : Springer.
- Biza, I., Giraldo, V., Hochmuth, R., Khakbaz, A., & Rasmussen, C. (2016). *Research on teaching and learning mathematics at the tertiary level : State-of-the-art and looking ahead. ICME-13 topical surveys*. Switzerland : Springer International Publishing AG.
- Blomhøj, M., & Jensen, T. H. (2003). Developing mathematical modelling competence : conceptual clarification and educational planning. *Teaching Mathematics and its Applications*, 22(3), 123–139.
- 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 (A Joint ICMI/IASE Study)* (pp. 57–69). Springer, Dordrecht.
- Carranza, P., & Kuzniak, A. (2006). Dualité de la notion de probabilité et enseignement de la statistique au Lycée en France. In N. Bednarz & C. Mary (Eds.), *Actes du colloque EMF 2006*. Sherbrooke, Canada : Editions du CRP.
- Carrasco, E., & Elicer, R. (2017). Conditional probability as a decision-making tool : a didactic sequence. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the tenth Congress of the European society for Research in Mathematics Education (CERME10)*. Dublin, Ireland : DCU Institute of Education & ERME.
- Chaachoua, H. (2014). *Le rôle de l'analyse des manuels dans la théorie anthropologique du didactique*. Consulté sur <https://hal.archives-ouvertes.fr/hal-01519339>
- Chaachoua, H. (2018). T4TEL un cadre de référence didactique pour la conception des EIAH. In J. Pilet & C. Vendeira (Eds.), *Actes du séminaire de didactique des mathématiques de l'ARDM*. France : IREM de Paris – Université Paris Diderot.
- Chaachoua, H., Ferraton, G., & Desmoulin, C. (2017). Utilisation d'un modèle praxéologique de référence dans un EIAH. In G. Cirade et al. (Ed.), *Évolutions contemporaines du rapport aux mathématiques et aux autres savoirs à l'école et dans la société* (pp. 301–324). Consulté sur <https://citad4.sciencesconf.org>
- Chaput, B., Girard, J.-C., & Henry, M. (2011). Frequentist approach : Modeling and simulation in statistics and probability teaching. In C. Batanero, G. Burrill, and C. Reading (Ed.), *Teaching statistics in school mathematics - challenges for teaching and teacher education* (pp. 85–95). New York : Springer.
- Chevallard, Y. (1989). Le passage de l'arithmétique à l'algèbre dans l'enseignement des mathématiques au collège. *Petit x*, 19, 43–72.
- Chevallard, Y. (1998). Analyse des pratiques enseignantes et didactique des mathématiques : l'approche anthropologique. In R. Noirfalise (Ed.), *Actes de l'université d'Été de didactique de la rochelle* (pp. 91–118). France : IREM de Clermont-Ferrand.
- Chevallard, Y. (2003). Approche anthropologique du rapport au savoir et didactique des mathématiques. In S. Maury & M. Caillot (Eds.), *Rapport du savoir et didactiques* (pp. 81–104). Fabert.
- Chiel, H. J., McManus, J. M., & Shaw, K. M. (2010). From biology to mathematical models and back : Teaching modeling to biology students, and biology to math and engineering students. *CBE-Life Sciences Education*, 9(3), 248–265.
- Derouet, C. (2016). *La fonction de densité au carrefour entre probabilités et analyse en terminale S. Etude de la conception et de la mise en oeuvre de tâches d'introduction articulant lois à densité et calcul intégral* (Thèse). Université Paris Diderot (Paris 7) Sorbonne Paris Cité.
- Diaz, C., & De la Fuente, I. (2007). Assessing students' difficulties with conditional probability and bayesian reasoning. *International Electronic Journal of Mathematics Education*, 2(3), 128–148.
- Dorier, J.-L. (2000a). *On the Teaching of Linear Algebra* (Vol. 23). Kluwer Academic Publishers, Dordrecht.

- Dorier, J.-L. (2000b). Recherches en Histoire et en Didactique des Mathématiques sur l'Algèbre linéaire - Perspective théorique sur leurs interactions. *Les cahiers du laboratoire Leibniz*, 12.
- Doukhan, C., & Gueudet, G. (2019). Students' difficulties at the secondary-tertiary transition : the case of random variables. In U. Jankvist & M. van den Heuvel Panhuizen (Eds.), *Proceedings of the eleventh Conference for European Research in Mathematics Education* (pp. 2464–2471). Utrecht, Netherlands : Freudenthal Group.
- Duran, P., & Marshall, J. (2019). Mathematics for biological sciences undergraduates : a need assessment. *International Journal of Mathematical Education in Science and Technology*, 50(6), 807–824.
- Duval, R. (1993). Registres de représentation sémiotique et fonctionnement cognitif de la pensée. In R. Duval (Ed.), *Annales de didactique et de sciences cognitives* (Vol. 5, pp. 37–65). France : IREM de Strasbourg.
- Eichler, A. (2007). The impact of a typical classroom practice on students' statistical knowledge. In D. Pitta-Pantazi & G. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 722–731). Larnaca, Chypre : Department of Education - University of Cyprus.
- Eichler, A., & Vogel, M. (2014). Three approaches for modelling situations with randomness. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking : presenting plural perspectives* (pp. 75–99). Dordrecht, The Netherlands : Springer.
- Feudel, F., & Panse, A. (2021). Can Guided Notes Support Students' Note-taking in Mathematics Lectures? *International Journal of Research in Undergraduate Mathematics Education*.
- Girard, J.-C. (2003). Probabilités au Lycée. Difficultés et obstacles dans l'enseignement des probabilités. *Association des Professeurs de Mathématiques de l'Enseignement Public (APMEP)*, 143, 35–44.
- Girard, J.-C. (2005). Quelques hypothèses sur les difficultés rencontrées dans l'enseignement des probabilités. In M. Henry (Ed.), *Autour de la modélisation en probabilités* (pp. 189–200). Besançon : Presses universitaires de France-Comté.
- Gonzalez-Martin, A., Gueudet, G., Barquero, B., & Romo-Vázquez, A. (2021). Mathematics and other disciplines, and the role of modelling : advances and challenges. *Research and Development in University Mathematics Education*, 169–189.
- Gueudet, G. (2008). Investigating the secondary-tertiary transition. *Educational Studies in Mathematics*, 67(3), 237–254.
- Gueudet, G. (2017). University teachers' resources systems and documents. *International Journal of Research in Undergraduate Mathematics Education*, 3(1), 198–224.
- Gueudet, G., & Thomas, M. (2019). Secondary-tertiary transition in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 762–766). New York : Springer.
- Gueudet, G., & Vandebrouck, F. (2019). *Entrée dans l'enseignement supérieur :*

- éclairages en didactique des mathématiques* (Rapport de recherche). CNESCO (Conseil national d'évaluation du système scolaire).
- Henry, M. (2003). Des lois de probabilités continues en Terminale Scientifique, pourquoi et pour quoi faire? *Repères-IREM*, 51, 5–25.
- Henry, M. (2005a). Modélisation en probabilités conditionnelles. In M. Henry (Ed.), *Autour de la modélisation en probabilités* (pp. 173–185). Besançon : Presses universitaires de France-Comté.
- Henry, M. (2005b). Notion d'expérience aléatoire. Vocabulaire et modèle probabiliste. In M. Henry (Ed.), *Autour de la modélisation en probabilités* (pp. 161–171). Besançon : Presses universitaires de France-Comté.
- Henry, M. (2010). Evolution de l'enseignement secondaire français en statistique et probabilités. *Statistique et Enseignement*, 1(1), 35–45.
- Heublein, U. (2014). Student drop-out from german higher education institutions. *European Journal of Education*, 49(4), 497–513.
- Huerta, M. P. (2014). Researching conditional probability problem solving. In E. Chernoff & B. Sriraman (Eds.), *Probabilistic Thinking. Advances in Mathematics Education* (pp. 613–639). Netherlands : Springer.
- Larsen, M. S., Kornbeck, K. P., Kristensen, R. M., Larsen, M. R., & Sommersel, H. B. (2013). Dropout phenomena at Universities : what is dropout? why does dropout occur? what can be done by the universities to prevent or reduce it? *Danish Clearinghouse for Educational Research*.
- Leontiev, A. N. (1984). *Activité, conscience, personnalité*. Moscou : Editions du Progrès.
- Martignon, L., & Wassner, C. (2002). Teaching decision making and statistical thinking with natural frequencies. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town, South Africa : International Statistical Institute and International Association for Statistical Education.
- Ministère de l'Education Nationale de la Jeunesse et de la Vie Associative (MEN-JVA), & DGESCO. (2001). Statistiques et probabilités dans l'enseignement secondaire.
- Ministère de l'Education Nationale de la Jeunesse et de la Vie Associative (MEN-JVA), & DGESCO. (2012). Ressources pour la classe terminale générale et technologique. probabilités et statistique.
- Ministère de l'Education Nationale de la Jeunesse et de la Vie Associative (MEN-JVA), & DGESCO. (2013). Ressources pour le lycée général et technologique. Les compétences mathématiques au lycée.
- Ministère de l'Education Nationale (MEN). (2010). Programme de l'enseignement des mathématiques. Classe première de la série scientifique. *Le Bulletin Officiel*.
- Ministère de l'Education Nationale (MEN). (2011). Programme de l'enseignement des mathématiques. Classe terminale de la série scientifique. *Le Bulletin Officiel*.
- Nankep, M. J. N. (2018). *Modélisation stochastique de systèmes biologiques multi-échelles et inhomogènes en espace* (Thèse). Ecole Normale Supérieure de Rennes.

- Nechache, A. (2016). *La validation dans l'enseignement des probabilités au niveau secondaire* (Thèse). Université Diderot Paris 7.
- Parzysz, B. (2011). Quelques questions didactiques de la statistique et des probabilités. In A. Kuzniak & F. Pluvinage (Eds.), *Annales de didactique et de sciences cognitives* (Vol. 16, pp. 127–147). IREM de Strasbourg.
- Poncy, M., Meny, J., Mounier, F., Vieudrin, D., Vincerot, F., Bonnafet, J., & Russier, M. (2012). *Indice Mathématiques Spécialité Tle S*. Bordas.
- Prodromou, T. (2014). Developing a modelling approach to probability using computer-based simulations. *Probabilistic thinking : presenting multiple perspectives.*, 417–439.
- Robert, A. (1998). Outils d'analyse des contenus mathématiques à enseigner au lycée et à l'Université. *Recherche en didactique des mathématiques*, 18(2), 139–190.
- Robert, A., & Vandebrouck, F. (2014). Proximités en acte mises en jeu en classe par les enseignants du secondaire et ZPD des élèves : analyses de séances sur des tâches complexes. *Recherches en Didactique des Mathématiques*, 34(2-3), 239–285.
- 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.
- Sentis, P. (2005). La notion de hasard : ses différentes définitions et leurs utilisations. *Laval théologique et philosophique*, 61(3), 463–496.
- Soury-Lavergne, S., & Bessot, A. (2012). Modélisation de phénomènes variables à l'aide de la géométrie dynamique. In J. Dorier & S. Coutat (Eds.), *Enseignement des mathématiques et contrat social : enjeux et défis pour le 21e siècle - Actes du colloque EMF 2012* (pp. 742–753). Genève. Consulté sur <http://emf.unige.ch/actes-emf-2012/>
- Séré, A., & Bassy, A.-M. (2010). *Le manuel scolaire à l'heure du numérique, une "nouvelle donne" de la politique des ressources pour l'enseignement* (Rapport N° 2010-08). France : Ministère de l'enseignement supérieur et de la recherche et Ministère de l'éducation nationale. Consulté sur <http://eduscol.education.fr/numerique/dossier/telechargement/rapport-ig-manuels-scolaires-2010.pdf>
- Tall, D. (1991). *Advanced Mathematical Thinking*. Dordrecht : Kluwer Academic Publishers.
- Totohasina, A. (1992). *Méthode implicative en analyse de données et application à l'analyse de conceptions d'étudiants sur la notion de probabilité conditionnelle* (Thèse). Université Rennes 1.
- Vandebrouck, F. (Ed.). (2008). *La classe de mathématiques : activités des élèves et pratiques des enseignants*. Toulouse : Octarès Editions.
- Vandebrouck, F. (2018). Activity theory in French didactic research. In G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt, & B. Xu (Eds.), *Invited Lectures from the 13th International Congress on Mathematical Education* (pp. 679–698). Cham, Switzerland : Springer.
- Vuurman, O., & Nardi, E. (2018). Negotiating different disciplinary discourses :

- biology students' ritualized and exploratory participation in mathematical modeling activities. *Educational Studies in Mathematics*, 101(2).
- Vivier, L., & Durand-Guerrier, V. (2016). Densité de D, complétude de R et analyse réelle - Première approche. In E. Nardi, C. Winslow, & T. Hausberger (Eds.), *Proceedings of the first conference of International Network for Didactic Research in University Mathematics (INDRUM)* (pp. 143–152). Montpellier, France : University of Montpellier and INDRUM.
- Weiss, H. H. (2013). The sir model and the foundations of public health. *Materials Matemàtics*, 1–71.
- Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, 33, 171–202.

Titre : Modèles praxéologiques dans la transition secondaire-supérieur : le cas des probabilités en filière biologie

Mot clés : Activité, Didactique pour l'enseignement supérieur, Mathématiques pour les non-spécialistes, Modèle praxéologique, Modélisation, Probabilités

Résumé : Le premier objectif de cette thèse est d'analyser les attendus institutionnels et les déroulements de classe en probabilités pour la première année de biologie à l'Université et de les comparer à ceux du secondaire en classe de terminale scientifique. Nous mettons en lumière les spécificités des étudiants non-spécialistes dans l'apprentissage des probabilités afin d'atteindre le second objectif de cette thèse qui est de faire l'état des lieux des difficultés rencontrées par les étudiants de biologie, pour ces enseignements de probabilités, dans le but d'y remédier.

Nous nous référons principalement à la Théorie Anthropologique du Didactique, que nous articulons avec des éléments de la Théorie de l'Activité adaptés à la didactique des mathématiques. De cette façon nous construisons un outil théorique appelé Modèle Praxéologique de Référence

Étendu par la théorie de l'activité et adapté pour la Transition secondaire-supérieur, noté MPR-ET, dans le cas ici des probabilités pour les non-spécialistes.

Cet outil est adapté aux besoins des étudiants biologistes débutants à l'Université et permet à la fois d'explicitier les attentes institutionnelles envers les étudiants à l'Université en tenant compte de ce qui est fait au secondaire mais également de mieux appréhender la transition. En effet, ce MPR-ET fait figurer des propositions en vue de construire un enseignement innovant qui réduirait les difficultés des étudiants à la transition, notamment de repenser les mathématiques enseignées à ces étudiants en tant qu'outils de modélisation pour l'étude de problèmes issus de la biologie et d'imaginer un enseignement commun mathématiques/biologie.

Title: Praxeological models in the secondary-tertiary transition: the case of probability in biology

Keywords: Activity, Didactics for higher education, Mathematics for non-specialists, Praxeological model, Modeling, Probabilities

Abstract: The first objective of this thesis is to analyze the institutional expectations and the class routines in probability for the first year of biology at the University and to compare them to those of Grade 12 in the science section of the secondary school. We highlight the specificities of non-specialist students in the learning of probability in order to reach the second objective of this thesis which is to take stock of the difficulties encountered by biology students in the teaching of probability, in order to remedy them.

We refer mainly to the Anthropological Theory of Didactics, which we articulate with elements of the Activity Theory adapted to the didactics of mathematics. In this way we build a theoretical tool called Praxeological Reference Model Extended by Activ-

ity Theory and adapted for the secondary-tertiary Transition, noted MPR-ET, in the case of probability for non-specialists.

This tool is adapted to the needs of beginner biology students at the University and allows both to make explicit the institutional expectations towards students at the University by considering what is done in high school but also to better understand the transition. Indeed, this MPR-ET includes proposals to build an innovative teaching that would reduce the difficulties of students at the transition, notably to rethink the mathematics taught to these students as modeling tools for the study of problems arising from biology and to imagine a joint mathematics/biology teaching.