



HAL
open science

Les initiatives des élèves dans l'enseignement et l'apprentissage des séries statistiques doubles : exemple de la classe de terminale SE en Centrafrique

Fabrice Delon Rodoumdje

► To cite this version:

Fabrice Delon Rodoumdje. Les initiatives des élèves dans l'enseignement et l'apprentissage des séries statistiques doubles : exemple de la classe de terminale SE en Centrafrique. Education. Université de Bretagne occidentale - Brest, 2022. Français. NNT : 2022BRES0078 . tel-04086930

HAL Id: tel-04086930

<https://theses.hal.science/tel-04086930>

Submitted on 2 May 2023

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.

THESE DE DOCTORAT DE

L'UNIVERSITE
DE BRETAGNE OCCIDENTALE

ECOLE DOCTORALE N° 603
Education, Langages, Interaction, Cognition, Clinique
Spécialité : Didactique des mathématiques

Par

Fabrice Delon RODOUMDJE

Les initiatives des élèves dans l'enseignement et l'apprentissage des séries statistiques doubles

Exemple de la classe de Terminale SE en Centrafrique

Thèse présentée et soutenue à Rennes, le 1 décembre 2022

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

Rapporteurs avant soutenance :

Jean Claude RÉGNIER
Rahim KOUKI

Professeur émérite des Universités, Université Lumière – LYON 2
Professeur (HDR), Université de Tunis – El Manar, Tunis - TUNISIE

Composition du Jury :

Président : Fabrice VANDEBROUCK
Examineurs : Jean Claude RÉGNIER

Rahim KOUKI

Dir. de thèse : Ghislaine GUEUDET
Co-dir. de thèse : Caroline POISSARD

Professeur (HDR), Université de Paris Cité
Professeur émérite des Universités, Université Lumière – LYON 2
Professeur (HDR), Université de Tunis – El Manar, Tunis - TUNISIE
Professeur (HDR), Université de Bretagne Occidentale
Professeur (HDR), Université Paris-Saclay
Maître de conférences, Université Bretagne Occidentale

Table des matières

Dédicace.....	9
Remerciements.....	10
INTRODUCTION.....	11
1 CONTEXTE DE LA RECHERCHE.....	14
1.1 Contexte institutionnel : présentation du système éducatif Centrafricain.....	14
1.2 Contexte mathématique : ajustement affine.....	16
1.2.1 Ajustement par la méthode au jugé ou à la règle.....	17
1.2.2 Ajustement par la méthode de Mayer.....	17
1.2.3 Ajustement par la méthode des Moindres carrés.....	18
1.3 L'enseignement des prérequis pour la série statistique à deux variables en Centrafrique.....	21
2 REVUE DE TRAVAUX.....	23
2.1 Histoire de l'ajustement affine et de la méthode des moindres carrés.....	23
2.2 Travaux en didactique sur l'enseignement et l'apprentissage de la statistique.....	25
2.2.1 Formation à l'esprit statistique, statistique dans les curricula, synthèses de recherches en didactique de la statistique.....	26
2.2.2 Pratiques des élèves et difficultés.....	30
2.2.3 Enseignement de la statistique sous forme de projet, activité statistique « authentique », données réelles.....	31
2.2.4 Enseignement des séries statistiques doubles et de l'ajustement.....	34
2.3 Travaux sur la modélisation.....	36
2.4 Synthèse du chapitre 2.....	39
3 CADRES THÉORIQUES ET QUESTIONS DE RECHERCHE.....	41
3.1 Théorie des situations didactiques.....	41
3.1.1 Le contrat didactique.....	41
3.1.2 Situation a-didactique.....	42
3.2 Théorie Anthropologique du didactique.....	43
3.3 Les registres de représentations sémiotiques.....	44
3.4 Les initiatives des élèves.....	45
3.4.1 Les <i>niveaux de mise en fonctionnement des connaissances</i>	45
3.4.2 Résolution de problème dans l'enseignement secondaire.....	45
3.4.3 Notion de topos des élèves.....	46
3.4.4 Démarche d'investigation.....	46
3.4.5 <i>Proposition de définition des initiatives des élèves</i>	46
3.5 Questions de recherche.....	47
4 MÉTHODOLOGIE DE CONSTRUCTION ET D'ANALYSE DE DONNÉES.....	49
4.1 Méthodologie de la construction d'un MPR.....	49
4.2 Méthodologie de l'analyse du manuel de Centrafrique.....	49
4.3 Méthodologie de l'analyse d'un enseignement ordinaire: méthodologie spécifique à l'enseignement ordinaire en RCA.....	50
4.3.1 Observation de séances de classe.....	50
4.3.2 Questionnaire et entretiens avec des enseignants.....	51
4.4 Méthodologie de conception et de l'analyse pour l'expérimentation en classe.....	52
4.4.1 Méthodologie de conception et de l'analyse de la première expérimentation en classe « Prévission du ménage 1 ».....	53
4.4.2 Méthodologie de conception et de l'analyse de la deuxième expérimentation en classe.....	53
4.4.3 Méthodologie de conception et de l'analyse de la troisième expérimentation en classe.....	54

5	ANALYSE DE L'ENSEIGNEMENT HABITUEL DE L'AJUSTEMENT AFFINE.....	56
5.1	Construction du modèle praxéologique de référence.....	56
5.1.1	Premiers types de tâches.....	56
5.1.2	Repérage des types de tâches dans le manuel de BTS	59
5.1.3	Construction du modèle praxéologique de référence à partir de l'étude historique, de la revue des travaux, du programme Centrafricain et du manuel BTS.....	73
5.2	Analyse du manuel de Centrafrique en utilisant le MPR.....	83
5.2.1	Analyse de la partie cours.....	83
5.2.2	Analyse de la partie exercice.....	85
5.2.3	Analyse praxéologique d'un exercice résolu du manuel en utilisant le MPR.....	88
5.2.4	MPR complété suite à l'analyse du manuel de RCA.....	92
5.3	Analyse d'un enseignement « ordinaire » en RCA.....	94
5.3.1	Description et analyse de la première séance de l'enseignement ordinaire.....	94
5.3.2	Description et analyse de la deuxième séance de l'enseignement ordinaire.....	104
5.4	Analyse du questionnaire aux enseignants.....	113
5.4.1	Analyse a priori du questionnaire.....	113
5.4.2	Analyse des réponses aux questionnaires.....	118
5.5	Analyse des entretiens avec les enseignants.....	123
5.6	Synthèse du chapitre 5.....	127
6	MISE EN PLACE ET ANALYSE D'EXPERIMENTATIONS.....	130
6.1	Analyse de la première expérimentation en classe « Préviation du ménage 1 » (2019).....	130
6.1.1	Présentation de l'exercice.....	130
6.1.2	Analyse a priori de l'exercice.....	131
6.1.3	Analyse de productions des élèves.....	136
6.1.3.1	Analyse globale.....	137
6.1.3.2	Analyse détaillée.....	138
6.2	Analyse de la deuxième expérimentation en classe (2020).....	145
6.2.1	Analyse a priori des problèmes.....	146
6.2.2	Analyse des productions des élèves.....	153
6.2.2.1	Analyse globale.....	154
6.2.2.2	Analyse détaillée.....	157
6.3	Analyse de la troisième expérimentation en classe (2021).....	168
6.3.1	Analyse du pré-test dans la classe 1.....	168
6.3.2	Présentation de l'enseignement conçu et analyse de son implémentation.....	174
6.3.2.1	Présentation et analyse du premier épisode (séance 1).....	176
6.3.2.2	Présentation et analyse du deuxième épisode (séance 2).....	180
6.3.2.3	Présentation et analyse du troisième épisode (séance 3).....	182
6.3.3	Analyse du post-test dans la classe 1.....	185
6.3.3.1	Analyse globale des réponses des élèves de la classe 1.....	186
6.3.3.2	Analyse détaillée des réponses des élèves de la classe 1.....	188
6.3.4	Analyse du post-test dans la classe 2 (classe témoin).....	194
6.3.4.1	Analyse globale.....	194
6.3.4.2	Analyse détaillée.....	196
6.3.5	Conclusion à propos de la troisième expérimentation.....	199
6.4	Synthèse du chapitre 6.....	200
	CONCLUSION GÉNÉRALE ET PERSPECTIVES.....	203
	Synthèse d'éléments de réponse aux questions de recherche.....	203
	Perspectives à la suite de cette thèse.....	206
	Perspectives pour la recherche.....	206
	Perspectives pour la formation.....	207

Bibliographie.....	208
Annexes.....	214
A. Transcription d'un enseignement ordinaire en Centrafrique.....	215
Transcription de la première séance de cours.....	215
Transcription de la deuxième séance de cours.....	221
B Le questionnaire.....	231
Questionnaire Professeur P2.....	235
Questionnaire Professeur P4.....	239
Questionnaire Professeur P10 (Aristide).....	242
C Guide d'entretien.....	246
Transcription d'entretien avec Professeur P2.....	247
Transcription d'entretien avec Professeur P4.....	251
Transcription d'entretien avec P10 (Aristide).....	256
D Première expérimentation en classe.....	261
Énoncé de la première expérimentation en classe.....	261
Copie des élèves du groupe 5.....	262
Copie des élèves du groupe 6.....	267
Copie des élèves du groupe 7.....	269
Copie des élèves du groupe 10.....	271
Copie des élèves du groupe 12.....	273
E. Deuxième expérimentation en classe.....	275
Énoncé de la deuxième expérimentation en classe.....	275
Copie des élèves du groupe 1.....	277
Copie des élèves du groupe 4.....	280
Copie des élèves du groupe 7.....	286
F. Troisième expérimentation en classe.....	290
Copie des élèves du groupe 1 de la classe 1 : pré-test.....	290
Trame de cours implémenté dans la classe 1.....	292
Copie des élèves du groupe 4 de la classe 1 : post-test.....	293
Copie des élèves du groupe 1 de la classe 2 : post-test.....	296

Liste des tableaux

Tableau 1: Méthodologie générale - données recueillies.....	54
Tableau 2: Extrait du programme Centrafricain de Mathématiques (Programmes de Mathématiques des classes de terminale SM, E & SE, 2002, p.4).....	57
Tableau 3: LEs types de tâches inférés du programme ou rencontrés dans le manuel.....	70
Tableau 4: analyse praxéologique d'un exercice résolu.....	77
Tableau 5: Tableau récapitulatif des types de tâches des exercices du manuel CIAM.....	86
Tableau 6: Grille récapitulative de l'analyse praxéologique de l'exercice.....	90
Tableau 7: Synopsis de la première séance.....	95
Tableau 8: Synopsis de la deuxième séance.....	105
Tableau 9: Réponses des enseignants à la première partie du questionnaire.....	117
Tableau 10: Tableau statistique pour l'exercice « <i>prévision du ménage 1</i> ».....	132
Tableau 11: Sous- série 1.....	133
Tableau 12: Sous- série 2.....	133
Tableau 13: Récapitulatif de l'analyse globale.....	137
Tableau 14 : Tableau récapitulatif des réponses des élèves à l'exercice 1.....	153
Tableau 15: Réponses des 12 groupes à l'exercice 2.....	154
Tableau 16: Réponses des 12 groupes à l'exercice 3.....	155
Tableau 17 : Tableau récapitulatif des réponses des élèves de la classe 1 au pré-test.....	171
Tableau 18 : Tableau récapitulatif de la troisième expérimentation dans la classe 1.....	174
Tableau 19: Synopsis du premier épisode.....	177
Tableau 20: Synopsis du deuxième épisode.....	180
Tableau 21: Synopsis du troisième épisode.....	182
Tableau 22 : Tableau récapitulatif des réponses des élèves de la classe 1 au problème 1.....	185
Tableau 23: Réponses des 10 groupes à l'exercice 2.....	186
Tableau 24: Réponses des 10 groupes au problème 3.....	186
Tableau 25: Tableau récapitulatif des réponses des élèves de la classe 2 au problème 1.....	194

Liste des figures

Figure 1: Exemple illustrant le nuage de points et la droite d'équation $y=a+bx$	19
Figure 2: Illustration de différentes situations vis-à-vis de l'ajustement.....	61
Figure 3: Modèle praxéologique de référence : Série statistique double - Etude historique - Revue des travaux - Programme de 2002 et manuel BTS.....	74
Figure 4: Un exercice corrigé extrait du manuel de BTS.....	76
Figure 5: Types de tâches dans l'exercice résolu du manuel BTS et prises d'initiative possibles de l'élève.....	82
Figure 6: Introduction du registre « tableau linéaire » et du vocabulaire caractère, modalité, effectif, distribution.....	84
Figure 7: Extrait du manuel Mathématiques Terminale SE.....	89
Figure 8: Représentation des types de tâches présents dans l'exercice résolu du manuel Mathématiques Terminale SE, mise en évidence des prises d'initiative.....	92
Figure 9: Modèle praxéologique de référence complété.....	93
Figure 10: Une élève au tableau pour remplir le 'tableau statistique'.....	101
Figure 11: Droite de Mayer.....	110
Figure 12: L'enseignant écrit l'énoncé de l'exemple au tableau.....	112
Figure 13: Extrait de la réponse de P1.....	121
Figure 14: Extrait de la réponse de P4.....	121
Figure 15: Extrait de la réponse de P3.....	122
Figure 16: Énoncé de l'exercice « prévision du ménage 1 » soumis à des élèves de Terminale SE en 2019.....	131
Figure 17: Extrait de copie du groupe 6.....	140
Figure 18: Extrait de copie du groupe 11.....	140
Figure 19: Extrait de copie du groupe 5.....	141
Figure 20: Extrait de copie du groupe 7.....	142
Figure 21: Extrait de copie du groupe 10.....	142
Figure 22: Extrait de copie du groupe 13.....	143
Figure 23: Extrait de copie du groupe 2.....	144
Figure 24: Extrait de copie du groupe 8.....	145
Figure 25: Représentation graphique de l'exercice 3.....	151
Figure 26: Extrait de copie du groupe 5.....	158
Figure 27: Extrait de copie du groupe 5.....	160
Figure 28: Extrait de copie du groupe 3.....	161
Figure 29: Extrait de copie du groupe 6.....	162
Figure 30: Extrait de copie du groupe 1.....	163
Figure 31: Extrait de copie du groupe 6.....	165
Figure 32: Extrait de copie du groupe 1.....	166
Figure 33: Extrait de copie du groupe 10.....	173
Figure 34: Exercice introductif.....	178
Figure 35: Extrait du cahier d'un élève.....	184
Figure 36: Extrait de copie du groupe 1.....	188
Figure 37: Extrait de copie du groupe 4.....	189
Figure 38: Extrait de copie du groupe 10.....	190
Figure 39: Extrait de copie du groupe 5.....	191
Figure 40: Extrait de copie du groupe 1.....	192
Figure 41: Extrait de copie du groupe 8.....	193
Figure 42: Extrait de copie du groupe 10.....	194
Figure 43: Extrait de copie du groupe G1.....	196

Figure 44: Extrait de copie du groupe G1.....197
Figure 45: Extrait de copie du groupe G1.....197
Figure 46: Extrait de copie du groupe G5.....198
Figure 47: Extrait de copie du groupe G4.....199

DÉDICACE

A mon père Amos

A ma Mère Ruth

A ma femme Élisabeth

A ma fille adorée Préfina Bath-shéba Christine

A mes frères Gildas et Vianney

A mes sœurs Marlène, Flavie, Ornella et Ninette.

Je dédie ce travail

REMERCIEMENTS

En premier lieu, je tiens à adresser toute ma gratitude et à exprimer toute ma reconnaissance à l'égard de Ghislaine Gueudet et Caroline Poisard, qui ont accepté de diriger ce travail de recherche. Je les remercie pour leurs disponibilités, la pertinence de leurs remarques, leurs encouragements et conseils pendant des moments difficiles de la rédaction de cette thèse.

Mes remerciements s'adressent ensuite à Monsieur Jean-Claude Régnier et Monsieur Rahim Kouki pour avoir accepté d'être les rapporteurs de cette thèse, et pour l'intérêt qu'ils ont porté à mon travail.

Je voudrais remercier vivement :

- Monsieur Fabrice Vandebrouck président du jury et Madame Patricia Marzin-Janvier qui m'ont fait l'honneur de se rendre disponible et faire partie du jury de cette thèse ;
- Monsieur Jean-Marie Boilevin et Madame Lætitia Bueno-Ravel pour leurs conseils et encouragements ;
- A l'équipe du laboratoire CREAD, Service Recherches & Relations Internationales et l'INSPE de Rennes qui m'ont toujours bien accueilli durant mes séjours de recherche par leur hospitalité et leur sympathie à la réussite de ce travail ;
- L'Université de Bangui pour leur soutien et pour l'intérêt qu'ils ont porté à mon travail ;
- La direction du lycée d'application de l'école normale supérieure (LAENS) pour m'avoir accueilli au sein de l'établissement, ainsi qu'aux collègues qui ont bien voulu m'ouvrir les portes de leurs classes dans le cadre des différentes enregistrements et d'entretiens faites tout au long de cette recherche.

Enfin, je remercie tous les membres de ma famille, ainsi que mes amis pour leur soutien.

INTRODUCTION

Notre recherche en didactique des mathématiques se situe dans le contexte de l'enseignement secondaire en République Centrafricaine (RCA). La RCA est un pays qui fait face à de nombreuses difficultés politiques et économiques. L'une des conséquences de cette situation est que les classes dans l'enseignement secondaire sont particulièrement surchargées : il n'est pas rare d'avoir environ cent-cinquante élèves dans une classe de Terminale. Notons que la RCA est située au cœur de l'Afrique.

Étant nous-même formateur d'enseignants à l'école normale supérieure (ENS) de Bangui la capitale, nous avons constaté que la formation à l'ENS est beaucoup basée sur l'aspect mathématique et non professionnel. Cela est dû au fait que la plupart des enseignants qui interviennent à l'ENS sont des enseignants permanents à la faculté. Notre motivation est de contribuer à une amélioration de l'enseignement des mathématiques en RCA, en développant les compétences didactiques au sein de la formation à l'ENS.

L'idée de ce sujet sur les séries statistiques doubles et l'ajustement affine part d'un constat : pendant notre cours de probabilité/statistique en deuxième année de mathématiques informatiques (MI), physique chimie (PC) et sciences de la vie et de la terre (SVT) à la faculté des sciences à l'Université de Bangui, nous constatons que beaucoup d'étudiants ont tendance à oublier déjà certains concepts des séries statistiques doubles et l'ajustement affine appris en Terminale. Et à chaque fois, il faut réviser ces savoirs statistiques. Alors, ce constat amène à s'interroger sur l'enseignement/ apprentissage des séries statistiques doubles et l'ajustement affine au lycée, plus particulièrement en Terminale.

Pendant la formation des enseignants de mathématiques de collège et de lycée à l'ENS de Bangui, on aborde l'enseignement de la statistique pendant un semestre (sur les trois ans de formation). Pour certaines années, l'enseignement de la statistique était retiré du programme de la formation des futurs enseignants de mathématiques. Par conséquent, ces enseignants ne se sentent pas à l'aise avec l'enseignement de la statistique, par manque de confiance d'enseigner une leçon mal maîtrisée.

Nous faisons deux constats, en tant que formateur à l'ENS de Bangui.

- Certains enseignants de mathématiques en centrafricaine se sentent moins à l'aise avec le chapitre sur les séries statistiques doubles et l'ajustement affine, que dans les autres chapitres. Pour cela, ils reportent ce chapitre vers la fin du programme, dans le but de le traiter superficiellement ou pas du tout faute de temps. D'autres enseignants préfèrent sauter ce chapitre sous prétexte que le temps leur fait défaut et qu'ils n'ont pas le temps nécessaire de tout faire.
- Pour les élèves de terminale, ils sont beaucoup plus préoccupés par l'examen du baccalauréat. Par manque d'organisation à l'approche de l'examen, certains élèves en centrafricaine n'ont pas eu le temps nécessaire pour bien s'exercer sur les derniers chapitres étudiés en classe. Et ces élèves n'ont pas pu acquérir des compétences calculatoires

nécessaires pour appliquer les techniques statistiques. Ils ont encore moins pu développer des compétences dans l'interprétation statistique.

Aucune recherche en didactique des mathématiques n'a encore eu lieu à notre connaissance en RCA. Il s'agit donc aussi pour nous de contribuer à faire émerger une communauté de recherche en didactique dans ce pays, en mathématiques en ce qui nous concerne mais aussi dans d'autres disciplines. Nous rappelons que cette thèse s'inscrit dans le cadre d'un partenariat entre l'école normale supérieure (ENS) de Bangui et l'institut national supérieur du professorat et de l'éducation (INSPE) de Bretagne.

Nous souhaitons effectuer une recherche donnant un rôle central aux contenus mathématiques en jeu. Nous aurions pu choisir de nombreux contenus mathématiques, puisque le contexte de la RCA faisait de chaque questionnement possible une question nouvelle du point de vue de la recherche en didactique des mathématiques.

Cependant, il nous a semblé que le thème de la statistique était particulièrement intéressant, car important dans de multiples professions et pour la formation des citoyens de la RCA. Vue l'importance de l'enseignement de la statistique et la nécessité pour un citoyen de comprendre et d'interpréter les données qu'il croise dans divers domaines (Gattuso, 2011), il est donc important pour les futurs cadres de développer une pensée et un raisonnement statistique (Régnier, 2005).

Plus précisément, nous nous proposons d'étudier un volet resté encore peu étudié par la recherche en didactique, à savoir l'enseignement et l'apprentissage des séries statistiques doubles au niveau de la classe de Terminale scientifique. Ainsi, il nous semble intéressant de centrer notre étude sur l'enseignement et l'apprentissage des séries statistiques à deux variables et l'ajustement affine en classe de Terminale SE (D) en RCA.

Nous voudrions comprendre comment se déroule cet enseignement en RCA, en particulier concernant le sens donné à l'ajustement affine, et la possibilité pour les élèves de construire ce sens au-delà de l'application d'une technique mémorisée. Cette construction du sens est pour nous directement liée aux initiatives que les élèves peuvent prendre lors de l'enseignement. Nous nous intéressons à l'enseignement « ordinaire », et nous tentons de mettre en place un enseignement donnant un rôle important au sens de l'ajustement et aux initiatives des élèves.

Les questions initiales posées de notre étude sont : Quelles sont précisément les difficultés des élèves avec les séries statistiques doubles et l'ajustement affine, en RCA ? Est-ce que l'enseignement est responsable de ces difficultés, et comment pourrait-on les surmonter ?

Ces questions naïves vont guider les choix théoriques et méthodologiques que nous allons présenter dans la suite de notre étude.

Plan de la thèse

Nous organisons notre étude autour de 6 chapitres.

Dans le premier chapitre consacré à la présentation du contexte de la recherche, nous évoquons le contexte institutionnel centrafricain, le contexte mathématique et l'enseignement et l'apprentissage des séries statistiques doubles en RCA. Dans le contexte mathématique, il s'agit d'explicitier les

trois méthodes d'ajustement affine (l'ajustement à la règle ou au jugé, l'ajustement par la méthode de Mayer et l'ajustement par la méthode des moindres carrés) apprises en terminale en RCA.

Dans le chapitre deux, nous présentons une revue de la littérature de recherche en lien avec notre questionnement. Ces travaux concernent l'historique de l'ajustement affine et en particulier de la méthode des moindres carrés, des travaux de recherches en didactique sur l'enseignement et l'apprentissage de la statistique, et les travaux portant sur la modélisation mathématique et plus particulièrement la modélisation dans le domaine des probabilités et des statistiques.

Le troisième chapitre expose les cadres théoriques dans lesquels nous nous plaçons et nous formulons également la problématique de notre recherche dans ce chapitre. Concernant les cadres théoriques, nous évoquons la théorie des situations didactiques (TSD, Brousseau, 1989) en particulier les concepts de contrat didactique et la situation a-didactique, la théorie anthropologique du didactique (TAD, Chevallard, 1998) à travers le concept de la praxéologie et celui de modèle praxéologique de référence (MPR), et les registres de représentations sémiotiques (Duval, 1993). En fin de ce chapitre nous présentons nos questions de recherche, appuyées sur ces cadres théoriques.

Les principaux éléments méthodologiques pour le recueil et l'analyse des données retenus pour l'étude feront l'objet du quatrième chapitre.

Le cinquième chapitre est consacré dans une première partie à la construction du MPR pour les séries statistiques doubles et l'ajustement affine. Nous nous basons sur l'étude historique, la revue de travaux (chapitre 2), l'analyse du programme de Centrafrique et l'analyse d'un manuel de BTS de France pour concevoir le MPR. Dans une deuxième partie, nous allons nous servir du MPR précédemment construit pour analyser le manuel utilisé en Centrafrique. Dans une troisième partie, nous faisons l'analyse d'un enseignement ordinaire fait en Centrafrique. La quatrième et la cinquième partie sont consacrées respectivement à l'analyse d'un questionnaire et d'entretiens avec des enseignants de RCA.

Nous présentons dans le sixième chapitre, l'analyse des trois expérimentations mises en œuvre dans la classe de terminale SE en Centrafrique. Chaque expérimentation a été construite à partir des résultats de la précédente. La première partie est consacrée à l'analyse de la première expérimentation qui est centrée sur la production des élèves, dont l'objectif est de laisser le choix de la méthode d'ajustement affine à la charge des élèves dans un problème proposé à ces derniers. La deuxième partie concerne l'analyse de la deuxième expérimentation. La deuxième expérimentation consiste à observer où se situe la prise d'initiatives des élèves à travers trois problèmes qui leur sont proposés. Et la troisième partie du chapitre est consacrée à l'analyse de la troisième expérimentation. Dans la troisième expérimentation, nous avons conçu et implémenté un enseignement sur le chapitre de la série statistique double et l'ajustement affine dans une classe de terminale SE que nous nommons classe 1. Nous avons soumis les élèves de cette classe à un pré-test et un post-test. Et nous avons aussi proposé le même post-test dans une autre classe que nous nommons classe 2. Une synthèse du chapitre y est aussi réalisée.

La dernière partie présente la conclusion de cette étude. Elle fait le bilan des résultats obtenus pour répondre aux questions de recherche concernant les prises d'initiatives des élèves de Terminale SE

en Centrafrique. Nous proposons également des perspectives pour la formation des enseignants et pour de futures recherches qui découlent de cette étude.

1 CONTEXTE DE LA RECHERCHE

L'objectif de ce chapitre est de présenter le contexte de la recherche dans son ensemble. Une première partie est consacrée au contexte institutionnel, c'est-à-dire, à la présentation du système éducatif centrafricain. Cependant, l'objectif n'est pas de faire l'historique du système éducatif centrafricain mais de présenter les faits pour que le lecteur n'ayant pas ou très peu de connaissances sur le système éducatif Centrafricain puisse en comprendre le fonctionnement. Une deuxième partie est consacrée au contexte mathématique, nous présenterons la notion d'ajustement affine pour les séries statistiques doubles afin de pouvoir par la suite l'étudier. La troisième partie est consacrée à la présentation des notions qui constituent des prérequis pour l'ajustement affine, dans les programmes scolaires de la RCA (l'enseignement de l'ajustement lui-même fera l'objet d'une présentation détaillée dans le chapitre 5).

1.1 Contexte institutionnel : présentation du système éducatif Centrafricain

Le système éducatif centrafricain relève de quatre ministères qui sont respectivement, le Ministère de l'Enseignement Primaire et Secondaire (MEPS), le Ministère de l'Enseignement Technique et de l'Alphabétisation (META), le Ministère de l'Enseignement Supérieur (MES) et le Ministère de la Recherche Scientifique et de l'Innovation Technologique (MRSIT).

Le système éducatif centrafricain fonctionne sur la base d'une seule langue officielle d'enseignement : le français. Le Sango, langue nationale a été instituée seconde langue officielle par la loi N° 97.014 du 10 Décembre 1997, portant orientation de l'Education en République Centrafricaine. Cependant, elle n'est pas encore utilisée comme langue d'enseignement. Le système éducatif centrafricain comprend deux secteurs, public et privé. Les modalités de création et de fonctionnement des établissements d'enseignement et de formation sont déterminées au niveau central de l'État, par les ministères en charge de l'Enseignement Primaire, Secondaire, Technique et de l'Alphabétisation ainsi que celui de l'Enseignement Supérieur. C'est ainsi que l'État par voie législative ou réglementaire,

- définit le régime de l'enseignement, sur le plan d'organisation et de formation;
- arrête les programmes et les manuels scolaires;
- fixe les modalités de création, d'ouverture, de fonctionnement et de financement des établissements et institutions privées de formation;
- contrôle les établissements et institutions privées de formation;
- régit les systèmes et les modalités d'évaluation des élèves et des étudiants et organise les examens officiels nationaux et l'année académique sur toute l'étendue du territoire national.

En définitive, l'État apparaît comme la principale institution organisatrice du système éducatif centrafricain.

En 1982, une tentative de réforme a opposé les classiques et les modernes. Les premiers réclamaient la pérennisation du modèle colonial alors que les seconds optaient pour des innovations relatives au passage à l'échelle des expériences des écoles de promotion collective. Même si ce débat a

débouché sur une première réorganisation du système éducatif en quatre niveaux (préscolaire, fondamental 1 et 2, enseignement secondaire général, technique et professionnel, enseignement supérieur), il a fallu attendre les mois de Mai et Juin 1994, pour qu'à travers les États Généraux de l'Éducation et de la Formation, une réforme profonde du système soit amorcée, suivie de l'élaboration d'un Plan National du Développement de l'Éducation et de la Formation (PNDE). Ce processus de planification a été repris autour des objectifs de l'Éducation Pour Tous (EPT) en 2004. Mais le Plan National d'Action n'a pu être effectivement mis en œuvre en raison des contraintes budgétaires et financières. Il était donc impérieux d'établir un nouveau diagnostic en 2006 et de développer un véritable plan de stratégie sectorielle de l'éducation en vue de répondre aux nombreux défis du développement de la République Centrafricaine.

La structure globale du système éducatif centrafricain, du secteur public comme celui du privé, s'organise autour des ordres d'enseignement suivants: le Préscolaire, le Fondamental 1 (F1), le Fondamental 2 (F2), l'Enseignement Secondaire Général, Technique et Professionnel (SGTP), l'Enseignement Supérieur ainsi que l'Alphabétisation et l'Éducation Non Formelle. L'Enseignement Supérieur est dispensé dans l'unique Université de Bangui ainsi que les établissements (privés) qui ont signé des conventions avec elle. L'Enseignement Préscolaire est le premier niveau et dure en principe trois ans mais est généralement fait sur deux ans. L'Enseignement Fondamental 1, deuxième niveau quant à lui, dure six ans et est sanctionné par le Certificat d'Étude Fondamental 1 (CEF 1). Ce cycle est suivi de l'enseignement Fondamental 2 qui dure sept ans, repartit sur deux cycles. Le premier cycle qui dure quatre (4) ans est sanctionné par le Brevet des collèges puis le second cycle quant à lui, dure trois ans et sanctionné par le Baccalauréat. La formation Technique et Professionnelle se fait sur deux ou trois ans, après la fin du collège c'est-à-dire après l'obtention du brevet des collèges. Enfin, l'enseignement supérieur couvre une période de deux à huit ans.

L'accès à l'Enseignement Supérieur est libre pour les titulaires du baccalauréat et l'accès aux écoles de formation ainsi qu'à certains établissements privés est subordonné à la réussite à un concours. S'agissant de l'enseignement technique et de la formation professionnelle, elle est structurée autour de deux cycles. Le cycle court et l'Enseignement post-primaire durent deux ans pour les Brevet d'Étude Professionnel (BEP). Les élèves de niveau 5ème sont formés en trois ans et la formation est sanctionnée par un Certificat d'Aptitude Professionnel (CAP). Les travailleurs possédant le CAP préparent en deux ans le Brevet d'Étude Technique (BET). Le cycle long dure trois ans et est sanctionné par un Baccalauréat Technique ou par un Baccalauréat de Technicien en Économie Familiale (BTEF). Quant à l'enseignement supérieur et l'université, il est structuré autour de trois axes: le premier axe forme en deux ans pour l'obtention du Diplôme Universitaire de Technologie (DUT), ou du Brevet de Technicien Supérieur (BTS); le deuxième axe forme en trois ans les titulaires de la licence, et le troisième axe forme en cinq ans pour l'obtention du Master et en sept ans pour le Doctorat en Médecine.

La politique du système éducatif centrafricain en matière de formation et recrutement des enseignants du collège a permis d'augmenter le nombre des enseignants estimés à 974 en 2005 à 1957 en 2010 d'une part, et les enseignants du secondaire général et de l'enseignement technique et professionnel estimés à 461 en 2005 à 842 en 2010.

Concernant la formation initiale, les professeurs sont recrutés sur concours national au niveau du baccalauréat pour les professeurs du collège et niveau licence pour les professeurs du lycée. Le cursus de formation pour les professeurs du collège dure trois années et celui des professeurs du lycée est organisé sur deux années à l'École Normale Supérieure (ENS) de Bangui .

Concernant la formation des professeurs du lycée, la première année est consacrée à un renforcement disciplinaire par matières d'enseignements et à des contenus transversaux (des théories didactiques et des concepts-clés communs à plusieurs disciplines, pédagogie générale, technique d'expression, anglais, informatique, etc). Quant à la deuxième année, elle se déroule en alternance, cours de didactiques disciplinaires, pédagogie d'une part et stage dans les lycées d'autre part. En effet, le stage se déroule en deux phases : la première phase est consacrée au stage d'observation qui dure deux semaines, suivi de la deuxième phase qui est un stage pratique en pleine responsabilité. A la fin de la deuxième année chaque étudiant réalise un mémoire de fin de formation.

Dans les classes dont nous avons connaissance en Centrafrique, l'effectif des élèves par classe dans les lycées publics avoisine 100 et dans les lycées privés 40. Nous n'avons pas de chiffres officiels, mais selon les enseignants, c'est une situation fréquente. C'est l'une des conditions qui ne permet pas aux élèves d'être dans de bonnes conditions d'apprentissage, dans le public.

Comme nous l'avons dit ci-dessus en introduction, nous sommes nous-même formateur à l'ENS de Bangui. Dans ce cadre nous dispensons des cours de didactique des mathématiques et de probabilité /statistique.

C'est cette implication en tant que formateur qui a été notre point de départ pour la recherche qui est présentée ici.

1.2 Contexte mathématique : ajustement affine

Dans cette partie nous présentons le contenu mathématique « expert » qui est en jeu dans notre travail.

Il s'agit d'étudier certaines caractéristiques d'une population Ω , qui est un ensemble d'individus : $\Omega = \{\omega_1, \dots, \omega_n\}$. Ces caractéristiques peuvent être mesurées, ce qui conduit à des données numériques associées à ces individus. On a donc des fonctions de Ω dans \mathbb{R} que l'on va nommer des variables statistiques (quantitatives). Une série statistique double correspond à la donnée de deux variables : X définie par $X(\omega_i) = x_i$ et Y définie par $Y(\omega_i) = y_i$, les x_i et les y_i étant des nombres réels. Nous notons que ces variables sont bien des variables statistiques, et non des variables aléatoires : nous nous plaçons dans le contexte de l'étude statistique d'une population particulière et non d'un échantillon choisi aléatoirement.

Soit le plan \mathcal{S} muni d'un repère orthogonal $(0, \vec{i}, \vec{j})$. A chaque couple (x_i, y_i) , on peut associer le point $M_i(x_i, y_i)$.

On appelle *nuage de points* associé à la série statistique double (X, Y) l'ensemble des points M_i appartenant au plan \mathcal{S} de coordonnées (x_i, y_i) . Le point de coordonnées (\bar{x}, \bar{y}) est le *point moyen* et est noté G , où \bar{x} (respectivement \bar{y}) est la moyenne de la variable X (respectivement Y).

Nous allons dans ce qui suit considérer un nuage de points dont les abscisses sont classées en ordre croissant, et nous intéresser à la recherche d'une fonction affine qui donnerait Y en fonction de X.

Une série statistique double étant donnée, la première question qu'on se pose est, est-ce qu'il y a un lien entre les deux variables statistiques composant cette série? Et après si on a vu qu'il y avait un lien, on peut avoir envie d'utiliser l'une des variables dont les valeurs sont connues (par exemple le temps) pour dire quelque chose sur l'autre variable qu'on ne connaît pas. Pour cela, on se demande s'il est possible de trouver une fonction qui donnerait la variable Y en fonction de la variable X, voire la fonction la plus simple possible donc affine.

Lorsque les points du nuage, associés à cette série statistique double sont sensiblement alignés, on peut alors construire une droite passant au plus près possible de ces points. On dit alors que cette droite réalise un ajustement affine du nuage de points de la série statistique double. Le sens attribué à l'expression « au plus près possible » doit être précisé. On peut, dans un premier temps, conserver un sens intuitif, graphique. On s'intéressera en particulier à réduire au maximum les écarts verticaux entre les valeurs observées et les valeurs théoriques que l'on va trouver grâce à la droite.

On définit algébriquement l'écart comme étant la différence entre la valeur observée et la valeur ajustée, $e_1 = y_1 - (ax_1 + b)$, $e_2 = y_2 - (ax_2 + b)$, ..., $e_n = y_n - (ax_n + b)$. Cette différence peut être positive, négative ou nulle (si les valeurs théoriques et observées sont confondues). Pour résoudre le problème de signe et faire en sorte que la sommation de ces écarts ne se traduise pas par une compensation quand ils sont de signes contraires, on les élève au carré, ce qui les rend tous positifs.

Il s'agit donc de minimiser la somme des carrés des distances euclidiennes entre les points du nuage et les points de la droite d'ajustement ayant la même abscisse, c'est-à-dire, si la droite d'ajustement a pour équation $y = ax + b$, la quantité S :

$$S(a, b) = \sum_{i=1}^n [y_i - (ax_i + b)]^2$$

Cette quantité est nommée somme des résidus.

Ci-dessous nous présentons les trois méthodes d'ajustement qui jouent un rôle essentiel dans cette thèse.

1.2.1 Ajustement par la méthode au jugé ou à la règle

C'est une méthode qui consiste à tracer une droite passant, le plus près possible des points du nuage, mais sans donner de définition formelle de « plus près ». Ainsi, on peut choisir par exemple d'équilibrer le nombre de points situés au-dessus et au-dessous de la droite, ou de joindre deux points extrêmes, etc. On dit alors que l'on a déterminé la droite d'ajustement graphiquement à la règle. Cette méthode n'est pas mathématiquement optimale (au sens de l'erreur mentionnée ci-dessus), mais elle permet de mettre en avant le sens de l'ajustement, nous y reviendrons par la suite.

1.2.2 Ajustement par la méthode de Mayer

C'est une méthode qui consiste à déterminer la droite passant par deux points moyens de deux sous-ensembles constituant le nuage de points. Elle consiste à partager l'ensemble des points (x_i, y_i) en deux sous-ensembles M_1 et M_2 de même effectif (si possible, sinon un des ensembles aura un élément de plus que l'autre). Notons que si les points sont associés à une pondération, il faut de plus tenir compte de cette pondération pour la détermination de deux sous-ensembles et des points moyens.

Il s'agit ensuite de tracer une droite d'équation $y=ax+b$, qui passe par les points moyens G_1 et G_2 , appelée droite de Mayer. Cette droite (G_1G_2) de Mayer est censée ainsi ajuster au mieux le nuage de points. Autrement dit :

- On commence par « découper » la série statistique double en deux sous séries bien distinctes, c'est-à-dire que l'on découpe le nuage de points $M_i(x_i, y_i)$ en deux sous-ensembles distincts et de même effectif (ou presque : si le nombre de points est impair, on peut mettre le point surnuméraire dans n'importe lequel des deux sous-ensembles) ;
- On calcule les coordonnées des deux points moyens G_1 et G_2 associés à ces deux sous-ensembles, et on place ces deux points sur le graphique ;
- On trace la droite (G_1G_2) appelée droite de Mayer du nuage de points $M_i(x_i, y_i)$, qui doit passer par le point moyen G du nuage de points $M_i(x_i, y_i)$. C'est cette droite qui constitue avec cette méthode un ajustement affine la série double (x_i, y_i) .

Dans la pratique, la méthode de Mayer est simple et fournit une droite « qui approche le nuage de points » lorsque les points du nuage sont sensiblement alignés. Cependant, cette méthode n'est pas optimale (toujours au sens de l'erreur évoquée ci-dessus), contrairement à la méthode des moindres carrés que nous présentons ci-après.

1.2.3 Ajustement par la méthode des Moindres carrés

La méthode des moindres carrés consiste à déterminer la droite dite de régression de y en x qui rend minimale la somme des carrés des résidus $S(a, b)$. On appelle cette droite la droite de régression de y en x .

Nous allons ci-dessous démontrer l'existence de cette droite et dans le même temps établir son équation.

Dans la pratique, pour minimiser $S(a, b)$, cela revient à calculer les dérivées partielles de $S(a, b)$ par rapport à a et b , et de les égaliser à zéro.

Démonstration de la méthode des moindres carrés

La méthode des moindres carrés propose d'ajuster le nuage de points $M_i(x_i, y_i)$ par une droite d'équation $y=ax+b$, avec a et b qui rendent minimale la somme des carrés :

$$S(a, b) = \sum_{i=1}^n [y_i - (ax_i + b)]^2 \text{ (Figure 1).}$$

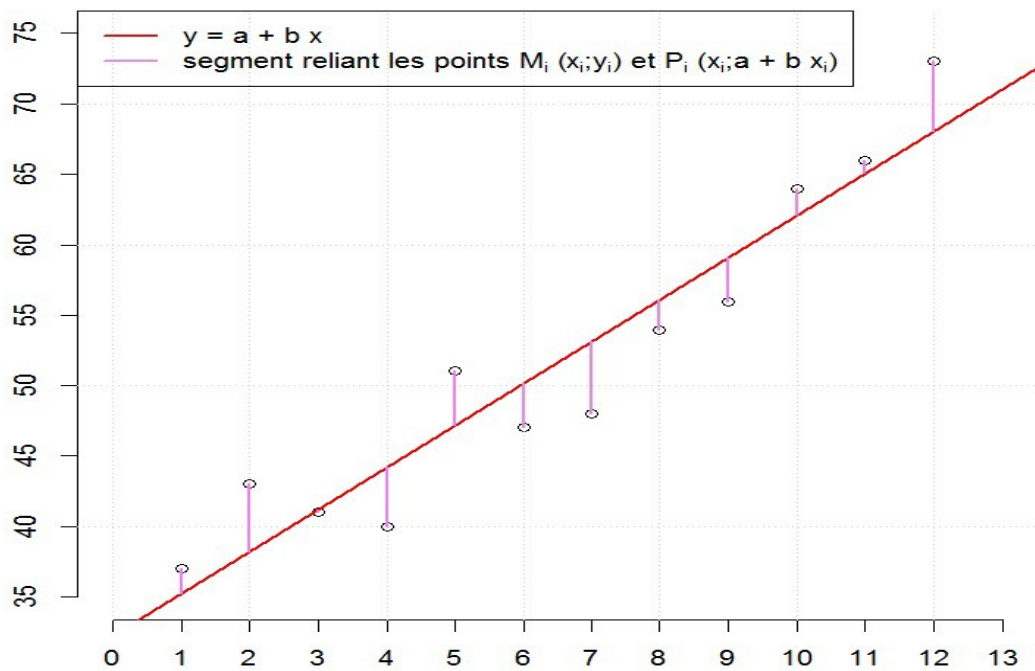


Figure 1: Exemple illustrant le nuage de points et la droite d'équation $y=a+bx$

On veut donc minimiser la quantité $S(a,b) = \sum_{i=1}^n [y_i - (ax_i + b)]^2$.

La somme $S(a,b)$ est une fonction de deux variables a et b qui va de \mathbb{R}^2 dans \mathbb{R} .

Pour trouver a et b , on regarde les deux dérivées partielles, c'est-à-dire, on cherche un point critique :

$$\frac{\partial S(a,b)}{\partial a} = 0$$

$$\frac{\partial S(a,b)}{\partial b} = 0$$

$$\frac{\partial S(a,b)}{\partial a} = -2 \sum_{i=1}^n x_i (y_i - ax_i - b) = 0 \quad (1)$$

$$\frac{\partial S(a,b)}{\partial b} = -2 \sum_{i=1}^n (y_i - ax_i - b) = 0 \quad (2)$$

On en déduit $\sum_{i=1}^n x_i y_i = a \sum_{i=1}^n x_i^2 + b n \bar{x}$ (3)

$$n \bar{y} = a n \bar{x} + n b \quad (4)$$

Divisons la relation (4) par n , on a : $\bar{y} = a \bar{x} + b$

$$\text{D'où } b = \bar{y} - a \bar{x} \quad (5)$$

Ce résultat indique que la droite passe par le point moyen G de coordonnées (\bar{x}, \bar{y}) .

Il donne également b en fonction de a .

En remplaçant b par son expression dans (3), pour trouver a :

$$\sum_{i=1}^n x_i y_i = a \sum_{i=1}^n x_i^2 + (\bar{y} - a \bar{x}) n \bar{x}$$

$$\sum_{i=1}^n x_i y_i = a \sum_{i=1}^n x_i^2 + n \bar{y} \bar{x} - a n \bar{x}^2$$

$$a \sum_{i=1}^n x_i^2 - a n \bar{x}^2 = \sum_{i=1}^n x_i y_i - n \bar{y} \bar{x}$$

$$a = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2}$$

$$a = \frac{\frac{1}{n} \sum x_i y_i - \bar{x} \bar{y}}{\frac{1}{n} \sum x_i^2 - \bar{x}^2} = \frac{\text{Cov}(X, Y)}{V(X)}$$

En définitive, les valeurs de a et de b qui minimisent $S(a, b)$ sont :

$$a_0 = \frac{\text{Cov}(X, Y)}{V(X)} \text{ et } b_0 = \bar{y} - a_0 \bar{x}$$

Remarque : Le point moyen G du nuage de la série appartient toujours à la droite de régression de y en x .

On appelle *coefficient de corrélation linéaire* entre X et Y le réel noté r égal à : $r = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$ où σ_{XY} est la covariance du couple (X, Y), σ_X est la variance de X et σ_Y est la variance de Y.

Propriété du coefficient de corrélation

Comme $S(a_0, b_0) = n V(Y)(1 - r^2)$, alors l'ajustement affine donne une bonne estimation lorsque r est proche de 1.

1.3 L'enseignement des prérequis pour la série statistique à deux variables en Centrafrique

Dans le programme centrafricain, la série statistique à une variable est introduite d'une manière progressive depuis le collège jusqu'à la classe de première. Dans le programme officiel de 2017, en classe de cinquième la statistique est déclinée en trois points : effectifs et fréquences, regroupement en classes et diagrammes. Voici quelques commentaires et instructions prévus dans le programme que l'on peut lire sur la statistique.

« On initiera les élèves à organiser les données statistiques sous forme de tableau en présentant les effectifs et les fréquences en pourcentage.

Amener les élèves à savoir lire, interpréter et construire un diagramme conforme à ce qui est dans le programme (diagramme en bâtons, à bandes, circulaire, semi-circulaire). On amènera les élèves à passer d'un type de représentation à un autre. » (programme officiel, 2017)

Pour la classe de quatrième, la statistique est répartie de la manière suivante : moyennes et pourcentage. On peut lire dans la partie commentaire et instruction sur la statistique :

« Les élèves sont confrontés à des situations familières où deux procédés de calcul différent de la moyenne sont mis en œuvre : somme des n données divisée par n , moyenne pondérée des valeurs par leurs effectifs.

Pour des petits effectifs, les élèves doivent savoir calculer une moyenne par la procédure de leur choix. Pour des grands effectifs, l'utilisation du tableur ou de la calculatrice est recommandée. » (programme officiel, 2017)

Et en classe de troisième, la statistique est répartie comme suit : médiane d'une série statistique, étendue d'une série statistique et quartiles d'une série statistique. Il est mentionné en commentaires et instructions :

« Dans le processus d'apprentissage de la statistique, il s'agit de savoir organiser, traiter et représenter des données fournies à l'état brut.

On entraînera les élèves à la démarche propre aux statistiques : organisation des données, avec la collecte ou la lecture d'informations et la présentation des données à l'aide d'un tableau des effectifs.

Traitement des données, déterminer la médiane, les quartiles, l'étendue et la moyenne d'une série statistique. Présentation des résultats, avec un diagramme circulaire, en bâtons ou à bandes. » (programme officiel, 2017)

Pour la classe de seconde en Centrafrique, le programme est réparti comme suit: le cas discret (effectifs cumulés, fréquences cumulées) ; cas de regroupement par classe de même amplitude ; différents modes de représentation d'une distribution statistique ; diagramme cumulatif ; caractéristiques de position (mode, moyenne, médiane) ; caractéristiques de dispersion (écart moyen, variance, écart type).

Le programme de la classe de première SE (Sciences Expérimentales) en Centrafrique porte sur la notion des séries statistiques groupées en classes et les caractéristiques des séries groupées (de position et de dispersion). Ce chapitre, limité à l'étude statistique d'un caractère qualitatif continu renferme des notions déjà abordées en classe de troisième et seconde.

Par ailleurs, en ce qui concerne les équations de droites qui constituent aussi un prérequis important pour l'ajustement affine, elles sont enseignées à la fin du collège (classe de troisième). Nous rappelons ici la définition.

Soit le repère (O, \vec{i}, \vec{j}) et une fonction affine $f(x) = ax + b$ définie sur \mathbb{R} .

On appelle la droite ID la représentation graphique de cette fonction f , de coefficient directeur a , passant le point b de coordonnées $(0, b)$ dans le repère (O, \vec{i}, \vec{j}) .

- a s'appelle le coefficient directeur de la droite ;
- b s'appelle l'ordonnée à l'origine de la droite ;
- La droite d'équation $y = ax + b$ est appelée équation réduite de la droite ID.

Les élèves apprennent à déterminer l'équation cartésienne d'une droite passant par deux points. Les méthodes qu'ils apprennent : avec le système 2×2 ou avec la pente de la droite.

2 REVUE DE TRAVAUX

Dans cette partie, nous présentons des travaux épistémologiques et didactiques concernant notre sujet d'étude. Ces travaux peuvent être directement liés à la statistique, et plus précisément à l'ajustement affine. Cependant, ils peuvent aussi concerner la modélisation mathématique. En effet, cet aspect de modélisation est important dans la problématique de l'ajustement, et nous nous intéressons particulièrement aux initiatives laissées aux élèves dans cet aspect de modélisation. Pour ce faire, il nous semble nécessaire de commencer d'abord par l'historique de l'ajustement affine et en particulier de la méthode des moindres carrés, ensuite les recherches didactiques en lien avec l'enseignement de la statistique, enfin les travaux portant sur la modélisation mathématique.

2.1 Histoire de l'ajustement affine et de la méthode des moindres carrés

La problématique de l'ajustement, et en particulier la méthode des moindres carrés, ont joué un rôle essentiel dans l'histoire de la statistique (Stigler 1990). Nous n'allons pas entrer ici dans le détail de cette histoire, mais nous allons présenter certains points de ce processus historique qui jouent un rôle important pour bien comprendre les concepts en jeu dans notre étude didactique.

Stigler (1990) souligne qu'au XVIII^{ème} siècle, le développement de la méthode de moindres carrés a été étroitement associée à trois problèmes scientifiques majeurs : (1) la détermination et la représentation mathématique des mouvements de la lune ; (2) compte-rendu d'une inégalité non périodique qui a été observé dans les mouvements des planètes Jupiter et Saturne ; et (3) la détermination de la forme et des mesures de la terre (géodésie). Il convient de noter plusieurs éléments à propos de ce contexte. Il est important du point de vue économique : en effet, la position des planètes est essentielle pour la navigation. Ainsi de nombreux scientifiques s'intéressent à ces questions. Les scientifiques se trouvent dans un contexte de physique. Ils réalisent des observations, prennent des données avec différents outils. Ils savent que ce processus physique de prise de données est susceptible d'engendrer des erreurs : manque de précision de l'observateur, manque de précision de l'appareil etc. Intuitivement, certains vont penser à faire des moyennes d'observations recueillies pour accroître leur fiabilité. Cependant, à cette époque les probabilités ne sont pas connues, et le fait de faire des moyennes pour améliorer les mesures prises n'est pas du tout fondé par une idée de type « loi des grands nombres ». L'ajustement recherché est purement physique, il n'y a pas d'idée de mesurer l'incertitude : on est loin de ce qu'on pourrait appeler une pensée statistique (Régnier 2005). Il s'agit seulement de réduire les erreurs de mesure et le lien entre ces deux perspectives, qui est important de nos jours et peut être exploité dans l'enseignement, (Munier, Chesnais & Molvinger 2018) n'est pas encore fait.

L'un des scientifiques ayant joué un rôle essentiel dans ce processus est Johann Tobias Mayer (1723-1762) qui s'intéressait aux mouvements de la lune. Son travail est important pour nous, car une des méthodes d'ajustement actuellement enseignée est nommée « méthode de Mayer ». Tobias Mayer cherche à déterminer les coordonnées d'un point situé à la surface de la Lune (le cratère Malinius). Avec un raisonnement que nous ne présentons pas ici (voir Stigler 1986 pour des détails) et en réalisant des observations étalées du 11 avril au 1748 au 4 mars 1749, il obtient 27 équations linéaires, avec lesquelles il veut déterminer seulement trois inconnues. Il décide alors de former

trois groupes de neuf équations chacune. Pour chaque groupe, il fait la moyenne des neuf équations. Il dit qu'en procédant ainsi il divise par neuf les erreurs de mesure (ce qui n'est pas juste, et il ne donne pas de justification mathématique associée). Il obtient ainsi trois équations, et il résout le système de trois équations pour obtenir les trois inconnues cherchées.

De nos jours, la méthode dite « méthode de Mayer » pour l'ajustement affine consiste à séparer en deux un nuage de points, et joindre les deux centres de gravité des sous-nuages. Bien que cette démarche soit très différente de ce que nous avons décrit ci-dessus, on retrouve deux aspects importants : il n'y a pas un objectif de minimiser une distance (objectif qui historiquement apparaît plus tard, voir ci-dessous) ; et il y a une idée de faire des « groupements ».

Stigler compare la démarche de Mayer et celle de Euler (1707-1783) qui s'intéresse lui au mouvement de Saturne et de Jupiter. Euler est lui aussi face à un trop grand nombre d'équations : 75 équations pour trouver six inconnues. Mais il ne fait pas de groupes d'équations. Il détermine de manière exacte les deux premières inconnues et reste bloqué pour les quatre autres. En effet, par un raisonnement mathématique il montre que regrouper des équations et en faire la moyenne ferait augmenter l'erreur. Il ne semble pas avoir comme Mayer une intuition liée à la prise de mesures.

Par la suite, des travaux concernant la mesure de la Terre vont rencontrer les mêmes problématiques. C'est le cas tout d'abord des travaux de Boscovich (1711-1787). Il travaille à partir de mesures : mesure de longueur d'arc à la surface de la terre, et d'angles correspondants. Il semble réfléchir d'emblée à des caractéristiques générales de sa méthode, en cherchant quelles corrections doivent être effectuées sur ses mesures de manière à se rapprocher des « vraies valeurs » en respectant un ensemble de conditions données par des relations physiques (ce travail est publié en 1770). Cette étude l'amène à étudier un ensemble (nuage) de points caractérisés par deux coordonnées, et à essayer d'approcher cet ensemble de points par une droite. L'une des conditions physiques considérées conduit à dire que la droite doit passer par le centre de gravité G du nuage. Boscovich calcule la somme des distances entre la droite et les points du nuage en faisant pivoter la droite autour du centre de gravité. Laplace (1749-1827), s'intéressant lui aussi à la mesure de la Terre, reprend (entre 1799 et 1805) le travail de Boscovich et le modifie, d'une part, en construisant un modèle analytique (alors que Boscovich en était resté à un raisonnement géométrique) et, d'autre part, en introduisant une pondération des mesures d'angles, en fonction des longueurs des arcs correspondants. De cette manière, il formalise et améliore la méthode de Boscovich.

Les différentes avancées présentées ci-dessus ont contribué à l'élaboration de la méthode des moindres carrés par Legendre (1752-1833). Legendre travaillait lui aussi sur des problèmes d'astronomie et de géodésie. En 1805, il publie un traité intitulé « Nouvelles méthodes pour la détermination des orbites des comètes », dans lequel figure la méthode des moindres carrés. Il présente son objectif : déduire le plus précisément possible de « vraies » valeurs à partir de mesures qui sont approximatives. Ce problème conduit à des équations du type $E = a + bx + cy + dz + \dots$; où les valeurs des coefficients a, b, c, d, \dots sont connues, x, y, z etc. sont les inconnues à déterminer. E représente l'erreur due aux mesures, et il y a un certain nombre des équations (et donc des erreurs $E', E'', E''' \dots$). Legendre propose de minimiser la somme des carrés des erreurs : $E^2 + E'^2 + E''^2 \dots$. Il montre alors que les calculs à effectuer pour cet objectif de minimiser la somme sont relativement élémentaires. Cette méthode est présentée de manière très générale, elle peut s'appliquer dans de nombreux contextes, et son succès est immédiat. D'ailleurs il s'ensuit une controverse entre Gauss

et Legendre sur la paternité de la méthode, qui atteste de l'importance de cette méthode. Gauss affirme qu'il utilisait cette méthode depuis 1795 (Samueli, 2010), sans l'avoir publiée. Sa publication majeure sur le sujet date de 1809. Elle concerne également le mouvement des corps célestes ; toutefois Gauss adopte d'emblée une perspective probabiliste. Pour une quantité X inconnue où il dispose de plusieurs mesures x_1, x_2, \dots, x_n les erreurs sont $e_1 = x_1 - X, e_2 = x_2 - X, \dots, e_n = x_n - X$. Il considère ces erreurs comme des réalisations de n variables aléatoires indépendantes de même loi continue. Gauss cherche la valeur de la quantité inconnue X qui rendra maximale la probabilité d'observation des mesures et des erreurs (ce qui dans le vocabulaire moderne serait nommé le principe du maximum de vraisemblance). En utilisant la moyenne arithmétique, il obtient ainsi la loi normale et le principe des moindres carrés. Donc Gauss avait publié en 1809 la loi normale en utilisant l'hypothèse que si une quantité a été obtenue par plusieurs observations, la moyenne arithmétique des valeurs observées sera la valeur la plus probable.

La problématique de l'ajustement se retrouve ensuite dans le contexte de l'économie, où Pareto publie en 1895 la loi des revenus permettant de représenter les distributions de salaires. Il s'agit de produire une courbe qui représente les logarithmes des revenus x en fonction des logarithmes du nombre $N(x)$ des revenus supérieurs à x . Selon Armate (2005), chez Pareto, l'ajustement est totalement empirique, il ne dérive d'aucun principe du type probabiliste. Pareto rejette l'usage de la méthode de moindres carrés pour ajuster une droite sur un nuage de points. Pareto (1909) explique son rejet de cette méthode en soulignant notamment le fait que, dans son cas, les écarts ne sont pas des erreurs de mesure. Nous voyons ainsi qu'au début du 20^{ème} siècle, certaines personnes remettaient en cause la méthode des moindres carrés telle que introduite par Legendre.

Du point de vue historique, la question du choix de la méthode d'ajustement la plus pertinente est importante. Nous regarderons dans les programmes et les manuels si cette question du choix est évoquée. (chapitre 5)

2.2 Travaux en didactique sur l'enseignement et l'apprentissage de la statistique

Dans cette partie nous nous intéressons à la recherche en didactique de la statistique, en nous centrant sur les aspects qui nous semblent les plus pertinents pour notre travail sur les séries statistiques doubles.

Rolland (2020) fait le point sur les publications en didactique de la statistique et l'enseignement de la statistique pendant dix ans, à travers des thèses de doctorat et articles. D'après ce dernier, il existe peu des travaux traitant de l'enseignement de la statistique en France durant la période 2009-2019. Cependant, au niveau international de nombreux travaux sont menés, avec des revues spécifiquement dédiées à l'enseignement et l'apprentissage de la statistique. Dans ce qui suit, nous avons sélectionné des travaux qui sont particulièrement en lien avec notre recherche pour diverses raisons :

(1) des travaux qui concernent la statistique en général, avec des perspectives assez variées entre ce qui touche aux curricula, ce qui touche à la formation à l'esprit statistique et des synthèses sur les recherches en didactique de la statistique.

(2) Pratiques et difficultés des élèves, nous permettent de cerner certaines difficultés des élèves et étudiants en statistiques.

(3) L'enseignement de la statistique sous forme de projet est lié à ce que nous faisons, parce que nous voulons que les élèves soient actifs. Et c'est en lien avec des idées de donner des responsabilités aux élèves.

(4) L'enseignement des séries statistiques doubles et de l'ajustement.

2.2.1 Formation à l'esprit statistique, statistique dans les curricula, synthèses de recherches en didactique de la statistique

Il s'agit ici de présenter des résultats issus des travaux en lien avec la formation à l'esprit statistique et les recherches en didactique de la statistique.

La formation à l'esprit statistique et la statistique dans les curricula

Dans la recherche française les travaux de Régnier (par exemples 2002, 2005, 2012) jouent un rôle essentiel en didactique de la statistique. Le point de départ de sa réflexion est le constat d'un double manque : manque de formation statistique en France, depuis le primaire à la formation professionnelle (et notamment celle de professeurs de mathématiques) ; et manque de recherches dans le champ de la didactique de la statistique.

Régnier (2002) explicite de la manière suivante les objectifs visés pour l'apprenant par une formation statistique :

- *construire un modèle requis par une approche statistique congruent au modèle dans lequel une problématique est posée.*
- *décrire, traiter, analyser des données de manière pertinente dans le cadre du modèle construit pour conduire une étude.*
- *tenir un raisonnement intégrant l'idée de risque dans l'énoncé des conclusions.*
- *interpréter les phénomènes étudiés sur la base de données statistiques recueillies sur des faits et leurs relations.*
- *communiquer des résultats des analyses de données en faisant une distinction entre le modèle utilisé et la réalité qu'il est supposé représenter, entre les traitements conduits au sein du modèle et les interprétations reformulées dans le contexte au sein duquel est posée la problématique.*
- *poursuivre de façon autonome et personnalisée un apprentissage en statistique afin d'enrichir les acquis personnels actuels. (Régnier 2002, p.6)*

Il s'agit selon lui de former l'apprenant au développement d'un *esprit statistique*. Selon l'auteur (Régnier 2005), l'esprit statistique se caractérise par un renoncement à la recherche de la vérité, qui est remplacée par des notions de vraisemblance et de plausibilité. Pour l'auteur, développer l'esprit statistique demande l'acquisition et la manipulation de trois modes de raisonnement : la déduction,

l'induction et l'éduction. En ce qui concerne ce dernier concept, Régnier se référant à Lalande (1926) définit comme une opération de pensée par laquelle « une cause efficiente, agissant sur une matière, y fait apparaître une forme déterminée ». Ces trois modes sont nécessaires à la conduite d'un raisonnement statistique, et à la production d'une interprétation statistique. Régnier souligne également l'importance en statistique des « pratiques sociales professionnelles de référence », et donc du travail des statisticiens professionnels. Il montre la complexité de la transposition didactique de la statistique, et met en évidence l'importance de la modélisation dans la pratique de la statistique. Dans la modélisation statistique d'une situation donnée il importe d'identifier ce que représentent les trois concepts premiers de la statistique : la population, l'individu, et la (ou les) variable(s) statistique(s). Dans une approche inspirée par la théorie des situations didactiques, il s'agit de concevoir et de mettre en œuvre avec des élèves ou étudiants des situations problèmes de statistique permettant ce travail de modélisation, allant de l'établissement d'un modèle à l'interprétation statistique.

Régnier (2012) étudie une question épistémologique fondamentale présentée comme suit : « dans quelle mesure l'enseignement et l'apprentissage de la statistique peuvent-ils être objets de recherches scientifiques dont les apports pourraient bénéficier aux pratiques ordinaires en salle de classe ? » L'auteur s'intéresse en particulier aux erreurs des élèves, difficultés, obstacles, comme objets d'étude de la didactique de la statistique dans l'apprentissage de la statistique. Il montre que les obstacles et les difficultés dans l'apprentissage de la statistique sont des objets utiles pour la pratique enseignante. Selon l'auteur, les erreurs produites par l'apprenant constituent des traces précieuses observables des problèmes auxquels il est confronté. Dans notre propre démarche, les erreurs des élèves observées dès le début de la recherche ont joué un rôle essentiel pour l'ensemble de notre étude, depuis l'élaboration de notre questionnement de recherche jusqu'à la conception d'un enseignement expérimental.

Gattuso (2011) montre l'importance de l'enseignement de la statistique de base dans un collège d'enseignement général et professionnel (cegep) québécois (17-18 ans). Selon l'auteur, l'enseignement de la statistique favorise le développement de la pensée statistique et la construction de concepts statistiques, en ce sens, qu'il ne faut pas rester sur l'aspect calculatoire mais développer la compréhension conceptuelle. Henry (2010) souligne que l'apprentissage des outils statistiques permet aux futurs citoyens de faire la lecture de tableaux, de graphiques, d'analyses économiques, etc. Gattuso (2011) souligne que l'aspect calculatoire des concepts statistiques n'assure pas la compréhension de ces concepts statistiques. Les enseignants doivent aider les élèves à développer une compréhension conceptuelle. Selon Gattuso (2011), il est important d'organiser les données recueillies sous forme de tableau afin de traiter des informations.

Burrill et Biehler (2013) font des recommandations pour enseigner les idées statistiques et font aussi quelques suggestions pour soutenir l'apprentissage des élèves. Ainsi, les auteurs caractérisent les neuf concepts suivants comme idées fondamentales en statistique dans le développement d'un curriculum statistique cohérent: les données, la variabilité, les distributions, les représentations, les associations et la modélisation des liens entre deux variables, les modèles probabilistes pour les processus de génération de données, l'échantillonnage et l'inférence. Concernant les données, selon les auteurs un curriculum soigneusement conçu tient à ce que les élèves participent à la construction des données. Ils soulignent que les représentations graphiques sont importantes pour la compréhension et l'analyse des données. Ils expliquent l'utilisation de la modélisation

mathématique comme un pont entre la statistique et la mathématique. D'après eux, le contenu des curricula devrait aider les élèves à se servir des données collectées en comparant des modèles mathématiques exprimés par des fonctions et à distinguer le graphe d'une fonction d'un modèle statistique. En plus, le curriculum peut aider les élèves à remettre en cause les mauvaises conceptions comme la confusion entre la corrélation et la causalité.

Burrill et Biehler (2013) recommandent la prise en compte des points suivants pour enseigner la statistique :

- Activer explicitement et s'appuyer sur les connaissances antérieures.
- Utiliser les outils technologiques avec les séquences d'activités pour développer des concepts importants.
- Faire participer les élèves au travail collaboratif et sur des projets.
- Fournir une rétroaction, dans le but d'aider les élèves à anticiper les obstacles les plus courants qu'ils doivent éviter.

Batanero et al (2013) présentent la situation de la statistique dans le programme de l'enseignement du secondaire sur le plan international. Ces auteurs notent, en France, la volonté d'initier les élèves à la statistique inférentielle dans les filières scientifiques. Les enseignants français sont peu outillés pour faire des traitements autres que formels des données proposées en situations problèmes. Et cela n'aide pas à développer l'esprit critique des élèves face aux données. Pour la situation en Belgique, le programme est centré sur le traitement de données, il ne se préoccupe pas de l'origine ou de la méthode de recueil des données. La situation en Espagne se réfère à la classification des sept idées fondamentales en statistique (Burrill et Biehler, 2013). Au Mexique, il en est de même, sauf en ce qui concerne l'association et corrélation entre variables. Batanero et al (2013) font les mêmes constats que certains auteurs en ce qui concerne l'enseignement de la statistique. D'une part, ils revendiquent la nécessité d'enseigner la statistique dans l'enseignement des mathématiques. D'autre part, ils soulèvent la question de la formation initiale et continue des professeurs de mathématiques en statistique. Souvent, ces professeurs sont peu préparés aux spécificités de la statistique.

Chevallard et Wozniak (2003), dans une étude de la transposition didactique de la statistique, montrent les liens qui existent entre la statistique et les objets du monde naturel et social. En effet, ils ont considéré d'abord la statistique comme un continent qui interagit avec les autres disciplines. Cette interaction est illustrée, d'une part, par des exemples sur la modélisation des données statistiques et, d'autre part, par des exemples sur la posture d'émancipation citoyenne. Ensuite ils mettent en évidence la notion de variabilité, en montrant d'abord l'importance de la notion de variabilité sur le calcul de la mesure et l'incertitude en sciences physiques (à travers un exemple en classe de troisième), et sur la modélisation d'une variable aléatoire continue dans les sciences expérimentales.

Synthèses de recherches en didactique de la statistique

Hahn (2015) dans son article intitulé : « la recherche internationale en éducation statistique : état des lieux et questions vives » présente une synthèse des travaux de recherche menés en didactique de la statistique. En effet, il s'agit de présenter d'abord un historique du champ, puis les principales questions étudiées par les travaux actuels. Cet article consiste en une présentation des origines du champ de la recherche, une présentation des principaux auteurs de ce champ et enfin une réflexion sur le lien entre statistique et probabilité. Pour les origines du champ de la recherche, ce champ s'est construit sur le groupe international de réflexion sur l'apprentissage des probabilités et de la statistique.

Dans les années, 1980 les enseignants majoritairement américains issus de trois disciplines : la statistique, les mathématiques et la psychologie, créent un groupe international de réflexion sur les probabilités et la statistique, qui devient plus tard un groupe de recherche.

A partir de 2000, le groupe publie sous le nom Statistique Education Research Newsletter (SERN) et devient SERJ vers 2002. SERJ est édité par International Association for Statistics Education (IASE), IASE est un groupe qui s'adresse aux personnes qui s'intéressent à l'enseignement de la statistique. Et à partir de 2003, le groupe de recherche sur l'enseignement de la statistique créa un groupe de travail dans le Congress of European Research in Mathematics Education (CERME).

Selon Hahn (2015) le terrain d'étude considéré par les chercheurs didacticiens de la statistique a évolué, passant de l'enseignement supérieur à l'enseignement secondaire et primaire. Cela est dû au profil des chercheurs qui sont pour la plupart des enseignants du supérieur et des didacticiens qui s'intéressent à la formation des maîtres.

Concernant les thèmes étudiés : l'auteur (Hahn, 2015) parle d'abord des conceptions fausses des élèves et de la remédiation sur les concepts statistiques de base. Ainsi la compréhension de la notion de moyenne et de variation ont été abordés mais également celle des histogrammes et leur effet sur l'accès à l'inférence. L'auteur souligne l'influence des conceptions préalables ainsi que la persistance de certaines conceptions fausses des élèves. A l'exemple de la conception transitive du coefficient de Pearson, c'est-à-dire si X est corrélé linéairement à Y et Y corrélé linéairement à Z , alors X est corrélé linéairement à Z (Castro Sotos et al, 2009).

Hahn (2015) note également des questions vives pour la recherche en didactique de la statistique. Nous retenons deux de ces questions, particulièrement liées à notre recherche. La première question parle de la relation entre probabilité et statistique. Dans les années 1990, les réformateurs de l'enseignement pensaient que les difficultés des élèves provenaient des enseignements trop formels des probabilités. D'où l'idée de se limiter à l'enseignement des notions de probabilité qui sont indispensables. Concernant la deuxième question qui parle du rapport à la réalité, l'auteur parle de l'importance de la pratique dans l'enseignement de la statistique.

Parzys (2011) présente un travail sur les problèmes didactiques qui se posent dans l'enseignement de la statistique et des probabilités. Il constate qu'en statistique comme en probabilité les représentations sémiotiques (Duval 1996) peuvent servir comme outils de résolution de problèmes : divers graphiques en statistique et l'arbre en probabilité. Selon l'auteur, en statistique descriptive, l'élaboration des graphiques se fait en respectant deux objectifs : la fidélité et la clarté. La fidélité à

l'ensemble des données, c'est-à-dire le rapport entre les données et leur représentation. Et la clarté permet de rendre compte, de présenter les résultats et d'en tirer des conclusions.

2.2.2 Pratiques des élèves et difficultés

Dans cette section nous allons nous intéresser aux travaux de recherche en didactique de la statistique qui concernent les pratiques des élèves et les difficultés qu'ils rencontrent au cours de leur apprentissage.

Batanero et al (1994) s'intéressent aux erreurs et aux difficultés des élèves sur la compréhension des concepts élémentaires en statistique. Il ressort de leurs travaux que certaines difficultés des élèves concernent la lecture des données statistiques, et l'interprétation des graphiques. Ils font également des erreurs de compréhension conceptuelle dans l'application de la notion de la moyenne dans les exercices.

Gattuso et Mary (1997), mènent une étude sur la compréhension du concept de moyenne et les stratégies utilisées, auprès de 241 élèves et étudiants au Québec, où il y a 40 élèves de l'enseignement secondaire qui n'ont pas encore vu la leçon sur la moyenne, 41 élèves de l'enseignement secondaire qui ont déjà vu la leçon sur la moyenne, 74 élèves du collège inscrits à un cours de méthodes quantitatives et 86 futurs maîtres en mathématiques en début de formation à l'université. Ainsi les auteurs constatent que, le calcul de moyenne de données groupées semble s'améliorer avec le niveau de scolarité des élèves et certains élèves ont des difficultés avec la moyenne pondérée. La performance des élèves varie d'un niveau à l'autre et diffère selon les questions. Pour les stratégies, les explications fournies par les élèves et étudiants ne présentent pas toujours la procédure utilisée pour trouver la réponse mais plutôt une justification.

Mary et Theis (2003) mènent une étude dans le cas d'une expérimentation dans une classe de l'enseignement primaire. Les auteurs s'intéressent plus précisément à la compréhension des statistiques d'élèves en difficultés graves d'apprentissages (c'est-à-dire : des enfants dyslexiques, dysorthographiques et dysphasie). L'expérimentation a été réalisée auprès d'une vingtaine d'élèves de deux classes de l'enseignement primaire au (Québec). La plupart ont 13 ans et sont en dernière année de l'enseignement primaire, dans une classe multi-niveaux. Une grande partie d'entre eux présente des troubles spécifiques d'apprentissage. L'expérimentation est réalisée à partir d'entrevues collectives. Les entrevues sont réalisées en quatre séances d'environ une heure chacune autour d'un thème. Chaque groupe est pris en charge par un expérimentateur, chargé de poser des questions qui suscitent l'interaction des membres du groupe. Les élèves ont élaboré un questionnaire portant sur des éléments qui les concernent directement : leur âge, leur sexe, la couleur de leurs cheveux, s'ils portaient des lunettes, le temps pris pour se rendre à l'école, etc. Et vingt élèves ont répondu au questionnaire. Ensuite les élèves analysent les données recueillies, et réfléchissent en même temps à leur prise de données. Les résultats ont révélé que certains élèves ont fait recours à certaines connaissances comme la notion de moyenne et de pourcentage, pour reformuler leurs questions. Ils ont tendance à se prononcer sur une relation hypothétique de cause à effet pour répondre aux questions. Selon les auteurs, l'expérimentation par questionnaire a permis d'engager les élèves dans une discussion entre eux et de faire émerger différentes stratégies utiles au raisonnement statistique.

Poursuivant l'étude de ce même thème, Mary et Theis (2007) décrivent des stratégies d'élèves à risque (les élèves à risque présentent des retards d'apprentissage ou des difficultés pouvant mener à l'échec) de niveau fin primaire dans des situations de problèmes statistiques. L'expérimentation a eu lieu dans deux classes de l'enseignement primaire, d'une école à Montréal accueillant exclusivement des élèves présentant un retard sur le plan des apprentissages. Ces élèves avaient reçu un enseignement sur la notion de moyenne dans la classe 1 (12 élèves) et sur les diagrammes à lignes brisées dans la classe 2 (8 élèves). Ils étaient répartis en sous-groupes de quatre ou cinq élèves. Et chaque sous-groupe est animé par une enseignante ou un des chercheurs. Un problème comportant deux questions est donné aux 20 élèves à risques âgés pour la plupart de 13 ans. Les données ont été récoltées par l'enseignante une semaine avant, auprès des élèves. Ils devaient répondre à un certain nombre de questions portant sur des caractéristiques personnelles et certaines habitudes de vie. Les variables du questionnaire ont été choisies par les chercheurs. Les deux questions retenues pour cette étude ont été formulées par les élèves. Il ressort de cette étude que la plupart des élèves qui ont appris en classe le concept de moyenne n'ont pas eu de difficulté à entrer dans la résolution de manière assurée. Nous retenons que, pour répondre aux questions issues de leurs propres interrogations, les élèves ont proposé des stratégies pertinentes sur la possibilité de comparer des populations à effectifs différents et sur les moyens d'y parvenir. L'approche utilisée par un certain groupes d'élèves pour répondre aux tâches demandées, s'appuie sur un raisonnement qui a du sens, même s'ils ne parviennent pas aux résultats. Nous notons que les tâches proposées aux élèves proviennent des données réelles et des questions provenant des élèves.

Concernant les connaissances des futurs enseignants en statistique, Eichlet et Zapata-Cardona (2016) se réfèrent à Shaughnessy (2007) qui suggère que l'accent devrait être mis sur les connaissances liées au contexte pour la pratique quotidienne des enseignants plutôt que sur la connaissance du contenu, qui a pour longtemps été le centre de la recherche sur les connaissances des enseignants. En ce qui concerne la connaissance statistique des enseignants en poste, Jacobbe et Horton (2010) ont étudié la compréhension des affichages de données sur trois professeurs de mathématiques. Ils ont constaté que les professeurs étaient compétents sur des questions liées à l'affichage de données, mais qu'ils ne réussissaient pas à répondre à des questions liées à la compréhension graphique. En outre Jacobbe (2012) a étudié la compréhension de trois enseignants du primaire en poste en ce qui concerne les concepts de moyenne et la médiane. Les résultats ont révélé que ces enseignants n'avaient pas une bonne connaissance conceptuelle de ces concepts statistiques de base.

Casey (2010) a étudié les connaissances nécessaires pour l'enseignement de la statistique. Il souligne que les enseignants ont besoin d'une base de connaissance substantielle pour enseigner le concept de coefficient de corrélation (calcul, interprétation, sensibilité, estimation et terminologie).

Les résultats des travaux de Castro-Sotos et al (2009) sur l'étude de la corrélation, ont révélé que les étudiants avaient des idées fausses sur la corrélation, ils sont incertains lorsqu'ils tirent des conclusions de deux corrélations entre trois variables (transition de la corrélation). En effet, les étudiants croient que la corrélation entre X et Z sera positive étant donné que les corrélations entre X et Y, et Y et Z sont positives.

2.2.3 *Enseignement de la statistique sous forme de projet, activité statistique « authentique », données réelles*

Dans cette section, nous nous intéressons aux travaux portant sur des enseignements de la statistique orientés vers des projets. Ces enseignements sont des enseignements expérimentaux qui ont été conçus par les chercheurs eux-mêmes. Batanero et al. (2013) considèrent que le meilleur moyen d'aider l'élève à développer son sens statistique c'est d'enseigner la statistique sous forme du travail en projet. Ces références sont très pertinentes pour nous, car elles sont en lien avec l'idée de donner des responsabilités aux élèves. A propos de l'enseignement orienté vers des projets, Markulin et al. (2021) dans leur article intitulé : *Project-based learning in statistics : a critical analysis* font un bilan de travaux en lien avec l'enseignement de la statistique sous forme de projet ou de résolution de problèmes. En effet, ils se réfèrent aux travaux de Carver et al. (2016), qui présentent le rapport GAISE (Guidelines for Assessment and Instruction in Statistics Education) dont ce dernier formule des recommandations en six points sur l'enseignement de la statistique. Concernant l'enseignement de la pensée statistique, le rapport propose d'enseigner la statistique comme processus d'investigation de la résolution des problèmes et de prise de décision. L'un des objectifs fixés dans la recommandation est d'intégrer le processus d'enquête et de proposer aux étudiants de réaliser des projets qui impliquent la conception d'études, la collecte de données, l'analyse et l'interprétation des données.

Brousseau (2009) présente pendant la conférence prononcée au 41^e journées de la Société française de la Statistique, une réflexion sur la statistique dans les programmes de mathématiques, ainsi qu'une expérimentation qui s'est déroulée entre 1973 et 1974 à l'école Jules Michelet de Talence, dans le cadre du COREM (Centre d'Observation sur l'Enseignement des Mathématiques) dans des classes de CM2 (10 - 11 ans). L'ingénierie didactique qu'il a conçue s'est étendue dans chaque classe sur 32 séances, soit 24h d'activités, 12h en éveil et 12h en mathématiques. Il s'agit d'une situation fondamentale qui amène les élèves à découvrir la statistique et son mode de pensée, connue sous le nom de « bouteille de Brousseau ». Une bouteille opaque contient 5 billes de deux couleurs, par exemple rouges et blanches. Le bouchon de la bouteille est transparent, il laisse apparaître une seule boule à la fois. Le but est de déterminer la proportion de boules blanches et de boules rouges contenues dans la bouteille. Brousseau constate que les élèves font progressivement des estimations, en notant ce qu'ils voient en retournant la bouteille. Ici il s'agit d'une initiation aux statistiques inférentielles et aux probabilités, avec l'idée de fréquence, et la construction intuitive d'un principe du type « loi des grands nombres ». En fin d'expérimentation, les élèves savent qu'ils peuvent se prononcer sur la proportion de billes rouges, sans jamais ouvrir la bouteille mais en faisant un grand nombre d'observations. Nous retiendrons que les élèves de CM2 (10 – 11 ans) sont capables dans le contexte de cette ingénierie de construire des rudiments de méthodes statistiques et de calcul des probabilités, sans faire recours à la rhétorique du hasard, mais il est nécessaire d'institutionnaliser et de stabiliser ces connaissances.

A ce propos, Zembrera et al (2017) proposent une réflexion sur l'importance et la place accordée aux données et à leur recueil dans les curricula de statistique en France dans les filières de sciences humaines à l'université. Ces auteurs formalisent leur réflexion sur l'enseignement de la statistique de la manière suivante:

- Dans le supérieur, le contenu du programme de la statistique se compose de deux volets, la statistique descriptive et la statistique inductive.
- En situation de classe dans le supérieur, les enseignants se basent sur des données préalablement fournies. La nature, le choix, la validité, le mode de recueil, l'origine, la production et la qualité des données ne font pas l'objet d'un questionnement critique de la part des enseignants. Cela s'explique par la faible masse horaire consacrée aux cours de statistique, dans les maquettes pédagogiques des filières non spécialistes.
- Les manuels de statistique adressés aux filières non spécialistes dans le supérieur sont caractérisés par un contenu mathématique très réduit et une mise en avant de méthodes.
- Dans le secondaire, dans le système scolaire français, les enseignants ont beaucoup mis l'accent sur le traitement de données et non sur le recueil et la production des données.

En plus, l'inadéquation entre les contenus des cours de statistique et les véritables besoins des étudiants sont des causes de difficultés de compréhension, du manque d'intérêt et de motivation de ces étudiants.

Ces mêmes auteurs soulignent que la faible masse horaire attribuée à l'enseignement de la statistique pousse les intervenants à faire des choix de contenu, en privilégiant le traitement des données, au détriment des concepts liés aux données et du recueil des données. Dans ce choix on constate un glissement de l'enseignement de la statistique vers l'enseignement du traitement des données statistiques. En sciences humaines et sociales, les appellations quantitative et qualitative prêtent à confusion et amènent des conceptions erronées chez les étudiants. A titre d'exemple, les étudiants qui ont recueilli leurs données à travers un entretien, semblent restreindre le traitement statistique aux données qualitatives et ignorent qu'on peut tout à fait faire un traitement quantitatif de données d'entretien, par exemple à travers une analyse lexicale.

Visnovska et Cobb (2019) s'intéressent au point de vue d'un groupe d'enseignants de mathématiques du secondaire, lors d'une formation sur les activités statistiques en classe, aux États-Unis. Les auteurs partent du constat (fait dans d'autres recherches) que le contexte joue un rôle important dans le développement du raisonnement statistique par les élèves. Il est important que les élèves puissent aussi réfléchir à la collecte de données ou même la faire eux-mêmes. Or, l'examen des manuels scolaires montre que ce n'est pas pris en compte. Les statistiques sont introduites comme une série de procédures et de calculs. Parfois il y a des contextes concrets, mais seulement pour intéresser les élèves et les motiver plus, pas pour donner du sens au raisonnement statistique. Le but de l'étude est de chercher à aider les enseignants à approfondir leur compréhension des idées statistiques centrales ; à donner un sens aux interprétations et solutions statistiques individuelles des élèves et à adapter les séquences pédagogiques conçues aux exigences et aux contraintes de la situation locale, dans lequel ils ont travaillé. Au début de la formation, les enseignants sont centrés sur les méthodes et les calculs. Lors de la formation, ils travaillent en groupe pour concevoir des séances dans lesquelles les élèves sont plus actifs. A la fin de la formation, les auteurs notent que les points de vue des enseignants sur l'enseignement de la statistique ont évolué. Pour la plupart des enseignants à l'issue de la formation, pour qu'une activité pédagogique soit intéressante, les élèves doivent avoir une expérience personnelle préalable avec le sujet en main (par exemple les élèves doivent comparer les données sur le poids des pastèques de deux distributeurs de fruits différents

dans le but de conseiller une restauratrice sur l'entreprise auprès de laquelle elle devrait acheter des pastèques pour la fabrication du jus). Les enseignants se sont orientés vers des activités de la statistique qui s'approchent d'un véritable problème et laissent plus de responsabilités aux élèves.

Gattuso et Pannone (2000) mènent une recherche expérimentale d'enseignement de la statistique en Italie, où l'équipe de recherche a construit un cours de la statistique élémentaire en se basant sur l'utilisation de données réelles et officielles. Les élèves utilisent des calculatrices statistiques, ou un tableur Excel.

Ainsi le cours a été pensé dans le but de faire développer chez les élèves, les compétences pour lire et interpréter des données, de fournir les éléments du langage statistique. Il ressort de cet enseignement expérimental que les élèves partent d'un ensemble de données, ils apprennent à lire des tableaux statistiques, à faire des graphiques, ils peuvent voir des résultats, ils arrivent à extraire des informations et à donner un sens à ces données. Grâce à cela, ils peuvent voir en quoi la statistique aide à la prise de décision et à la prévision.

Koparan et Güven (2012) analysent un enseignement sous forme de projet, dans lequel les étudiants travaillent en groupes, ils sélectionnent leurs variables, ils recueillent et analysent leurs propres données. C'est une étude qui vise à évaluer l'attitude des étudiants envers les statistiques.

Jamie (2002) présente une revue des travaux sur la façon dont les méthodes de simulation par ordinateur (CSM) sont utilisées pendant l'apprentissage de la statistique, aux États-Unis. Selon l'auteur, l'intégration des ordinateurs dans les classes de statistiques post-secondaires a eu un impact majeur sur l'enseignement et l'apprentissage de la statistique au cours des dernières décennies. L'un des avantages de ces outils pendant l'apprentissage de la statistique, est qu'il permet aux étudiants de réaliser des tâches de calcul plus rapidement et plus efficacement, ce qui les libère pour se concentrer davantage sur les concepts statistiques. Jamie (2002) souligne que l'application d'apprentissage de la statistique à l'aide de CSM peut permettre aux étudiants de développer leur propre compréhension des concepts statistiques. Et c'est une occasion pour eux d'apprendre en construisant leurs propres idées et connaissances à partir des expériences de simulation par ordinateur, avec l'aide de l'enseignant.

En plus, l'auteur (Jamie) présente une revue des travaux liée à l'enseignement et l'apprentissage de la statistique à l'aide de CSM où il considère principalement les étudiants en post-secondaire. Il ressort de ces travaux que, de nombreux chercheurs dans le domaine de l'enseignement de la statistique recommandent l'utilisation des CSM pour enseigner des concepts abstraits. Pour montrer l'avantage de l'utilisation de la méthode par simulation, l'auteur s'est référé aux travaux de Kersten (1983) sur le théorème central limite, qui souligne que la méthode de simulation peut clarifier les concepts et les théorèmes de la statistique, et peut également permettre à l'étudiant non orienté vers les mathématiques en statistique élémentaire d'avoir des expériences inductives avec des concepts statistiques d'une manière très efficace en terme de temps.

Concernant la régression linéaire, Jamie (2002) fait référence à Franklin (1992) qui suggère d'utiliser la simulation pour illustrer certains aspects de base du modèle de la régression linéaire simple à l'aide du logiciel MINITAB pour simuler les erreurs aléatoires. Pour Jamie (2002) l'utilisation de CSM pour illustrer les concepts statistiques peut encourager les étudiants à identifier leurs idées fausses et à les modifier en fonction de leur propre construction de ces concepts.

2.2.4 Enseignement des séries statistiques doubles et de l'ajustement

Estepa et Sanchez-Cobo (2003) présentent une étude exploratoire de caractérisation des connaissances que les étudiants ont sur les notions de corrélation et de régression à partir de la résolution des problèmes. Les auteurs analysent les réponses de 193 étudiants universitaires sur deux problèmes à travers un questionnaire. Dans cette analyse, ils étudient les procédures des étudiants et discutent des réussites, des difficultés et des erreurs que les étudiants manifestent dans leur réponse aux questions posées. Il ressort de cette analyse que, sur le premier problème le choix d'utiliser la droite de régression de y sur x ou de x sur y est à la charge des étudiants et leur choix dépend de la procédure utilisée. Selon les auteurs, la principale difficulté rencontrée par les étudiants dans le premier problème est qu'ils ne tiennent pas compte du fait que la droite de régression passe par le centre de gravité (\bar{x}, \bar{y}) que nous appelons dans le contexte centrafricain le point moyen $G(\bar{x}, \bar{y})$. Ils ont tendance à confondre la variable explicative et la variable expliquée. Concernant le calcul du coefficient de corrélation, les étudiants connaissent l'algorithme approprié pour le calcul mais ils ne sont pas en mesure de faire correctement le calcul. Il ressort de cette analyse que, près de deux tiers des élèves calculent le coefficient de corrélation mais seulement la moitié arrive à une expression correcte de la droite de régression. Toutefois, ils ont le contrôle de l'ordre de grandeur pour le coefficient de corrélation.

Casey et Wasserman (2015) réalisent une étude dont le but est d'explorer les connaissances que les enseignants ont sur l'ajustement affine. Des entretiens basés sur les tâches ont été menés avec dix-neuf enseignants de mathématiques en formation initiale et continue au États-Unis. Les chercheurs ont demandé à ces enseignants de placer une droite d'ajustement « au jugé » sur un nuage de points, représentant des données contextualisées (par exemple le nombre moyen d'entrées annuelles au cinéma en fonction du prix moyen du ticket). Les enseignants devaient « penser à voix haute » en plaçant leur droite, ou expliquer leurs choix ensuite. Cinq cas avec des répartitions différentes des points sont présentés. Les enseignants doivent aussi trancher entre deux propositions différentes de droites d'ajustement, tracées par des élèves, en expliquant leur choix. Les chercheurs constatent que les enseignants avaient une capacité relativement forte à placer de manière cohérente une droite d'ajustement « au jugé ». En effet, leurs droites placées au jugé étaient assez proches de la droite de régression des moindres carrés sur la plupart des tâches. Cependant, les enseignants ont diverses conceptions de la droite d'ajustement : « elle doit représenter l'ensemble des données » ; « elle doit permettre de prédire ce qui peut se passer dans l'avenir » etc. Également ils utilisent des critères variés, selon la forme du nuage de points, par exemple : « la droite doit passer par le plus grand nombre possible de points » ou « s'il n'y a pas de forme identifiable du nuage, la droite doit être horizontale ». Chaque participant semblait toutefois avoir un critère dominant. De plus, ils constatent que les enseignants faisaient preuve d'un manque de sens statistique. En statistique, il est possible d'avoir une approche flexible, tenant compte du contexte. Certains enseignants adoptent une approche plus déterministe (par exemple, ils veulent que la droite d'ajustement passe par le point $(0,0)$ sans faire le lien avec le contexte). Casey et Wasserman (2015) mettent l'accent sur la nécessité de formation des enseignants à la statistique et à l'idée d'ajustement affine, afin que ceux-ci puissent l'enseigner à leur élèves. En particulier, ils considèrent que le lien entre ajustement affine et inférence doit être travaillé (dans quel cas une droite d'ajustement permet de faire des inférences sur la valeur d'une des variables).

Marasinghe et al. (2004) ont conçu et présenté des outils logiciels qui peuvent être utilisés pour enseigner la régression statistique à l'aide de graphiques dynamiques. Ces auteurs construisent des modules présentant des leçons utilisant leur outil logiciel. Il s'agit ici de régression statistique au sens large, c'est-à-dire avec une variable expliquée Y et un certain nombre N de variables explicatives X_1, X_2, \dots, X_N . Cependant les auteurs considèrent en particulier la régression linéaire simple et la méthode des moindres carrés. Le logiciel permet, pour un nuage de point donné, de placer une droite d'ajustement comme on le souhaite, et de faire afficher la somme des carrés des résidus. Il est aussi possible de faire afficher la droite des moindres carrés (pour laquelle on peut aussi bien sûr calculer la somme des carrés des résidus). On peut afficher simultanément plusieurs droites, et les résidus correspondants : avec un tel outil, les élèves peuvent accéder à des éléments numériques pour réfléchir à la meilleure droite d'ajustement.

Dans la section qui suit, nous allons nous intéresser aux travaux de recherche en didactique des mathématiques qui concernent la modélisation.

2.3 Travaux sur la modélisation

Les travaux de recherche en didactique de la statistique soulignent l'importance du lien avec le monde « réel », et donc une nécessité de modélisation. Mais cette modélisation elle-même est particulière, car elle est liée au mode de raisonnement statistique, qui n'est pas le même, par exemple, que le raisonnement géométrique. Ainsi, nous présentons des travaux de recherche en lien avec la modélisation mathématique, suivi des travaux en lien avec la modélisation dans le domaine des probabilités et des statistiques.

[A propos de la modélisation mathématique](#)

Dans cette partie nous nous intéressons aux travaux de recherche en lien avec la modélisation mathématique. Pour cela nous rappelons quelques définitions de la modélisation que l'on trouve dans les travaux de recherche en didactique des mathématiques.

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.

Coulange (1997) définit un modèle mathématique comme une interprétation mathématique liée à une situation réelle (qui peut éventuellement provenir d'autres disciplines de sciences expérimentales ou humaines) ou à sa description en langage naturel, relativement à des questions que l'on se pose sur cette situation.

Quant à Henry (2010), il définit un modèle comme une représentation simplifiée et idéaliste de la réalité.

De nombreux auteurs en didactique des mathématiques ont publié des travaux en lien avec la modélisation mathématique. Nous nous sommes référés en particulier à Kuzniak (2021). Dans son cours à la 21^{ème} École d'été de didactique des mathématiques, il propose un chemin possible pour faire la modélisation mathématique :

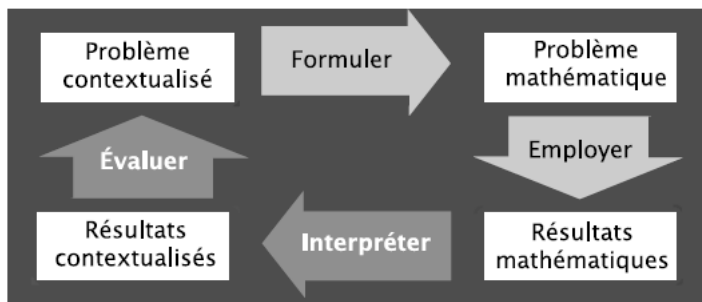
Problèmes venus du monde réel \Leftrightarrow travail sur les modèles \Leftrightarrow un traitement mathématique \Leftrightarrow application des résultats.

Il note historiquement un consensus dans les recherches menées à la fin des années 60, qui rend compte de l'application des mathématiques aux phénomènes naturels, de la manière suivante :

1. Description et formulation d'un modèle idéalisé en langage mathématique ;
2. Solution du problème mathématique et déduction de résultats susceptibles de vérification expérimentale ;
3. Comparaison des observations et de la théorie pour arriver à une évaluation de la validité du monde.

Ce modèle élémentaire est affiné par la suite, avec en particulier une prise en compte du caractère potentiellement cyclique de la modélisation. En effet, l'interprétation des résultats d'une première modélisation peut conduire à remettre en cause le modèle choisi, et donc à repartir sur un nouveau cycle. Deux exemples de tels cycles sont particulièrement connus :

- Le cycle simplifié de PISA (2012) décliné en quatre points entre le monde réel (ou problème contextualisé) et le monde mathématique :



- Le cycle de modélisation de Blum et LeiB (2005, reprise 2018) axé sur les compétences suivantes : comprendre la tâche, simplification/organiser, mathématiser, travail mathématique, interprétation, validation et présentation.

Ce cycle de modélisation est explicité en sept processus, orientés vers le cognitif comme suit :

1. Comprendre la tâche et construire la situation modèle ;
2. Simplifier et structurer la situation modèle et construire le modèle réel ;
3. Transformer le modèle réel en un modèle mathématique ;
4. Appliquer les techniques mathématiques pour produire des résultats mathématiques ;
5. Interpréter les résultats mathématiques ;
6. Valider ces résultats en relation avec la situation réelle ;
7. Présenter et expliquer le processus de solution ainsi que les résultats.

Les auteurs ont aussi rédigé un article (Blum & Ferri, 2009) dans lequel ils montrent des exemples de la façon dont les élèves et les enseignants gèrent les tâches de modélisation. Nous entendons par tâche de modélisation, une tâche avec une forte demande de modélisation.

Ils se réfèrent à des exemples pour définir la compétence de modélisation comme la capacité de construire des modèles en réalisant ces différentes étapes de manière appropriée ainsi que d'analyser ou comparer des modèles donnés. Selon Blum et Ferri (2009) les modèles mathématiques sont partout et souvent en lien avec des outils technologiques.

En ce qui concerne les pratiques des enseignants vis-à-vis de la modélisation en classe, Blum et Ferri (2009) soulignent que, pour un enseignement de qualité, il est crucial qu'un équilibre permanent entre le guidage et l'indépendance des élèves soient maintenus. Dans le contexte de la modélisation, les interventions du professeur qui donnent des indices aux élèves à un méta-niveau tels que : « Imaginez la situation ! Que visez-vous ? Que manque-t-il encore ? Ce résultat correspond-il à la situation réelle ?, etc. » sont les plus adéquats. Mais dans la réalité des classes les interventions de l'enseignant amènent souvent un guidage plus important

Pour effectuer une tâche de modélisation, Blum et Ferri (2009) proposent un plan en quatre étapes aux élèves :

1. Comprendre la tâche :

- Lisez le texte avec précision et imaginez la situation clairement ;
- Faites un croquis

2. Établissement du modèle :

- Recherchez les données dont vous avez besoin. Si nécessaire : faire des suppositions ;
- Rechercher des relations mathématiques

3. Utilisation mathématique :

- Utilisez les procédures appropriées ;
- Notez votre résultat mathématique

4. Expliquer le résultat :

- Relier le résultat à la tâche. Si nécessaire, revenez à 1 ;

Notez votre réponse finale. (Blum et Ferri 2009, p.12)

Piednoir (2004) propose la démarche de la modélisation de la manière suivante :

- prendre une situation dite réelle ou concrète ;
- en faire une description naïve qui permet de sélectionner les aspects les plus importants ;
- construire un objet mathématique censé rendre compte de la description précédente ;
- manipuler cet objet pour aboutir à des conclusions ;
- rapporter les conclusions à la situation réelle si possible.

Les étapes de la démarche de la modélisation citées ci-dessus cadrent bien avec une étude statistique auprès des élèves de terminale, dans le cas de notre étude où on se centre sur l'ajustement affine et l'estimation. Devant un problème de statistique avec des données réelles ou fictives, les élèves doivent construire un modèle pour faire une estimation, en se servant des concepts statistiques pour aboutir à une conclusion et confronter celle-ci à une situation réelle.

Il ressort de cette partie que construire un modèle de la réalité étudiée est une difficulté à différents niveaux scolaires, en particulier parce qu'il faut choisir les outils mathématiques appropriés et faire des aller-retours entre les mathématiques et la situation réelle.

A propos de la modélisation en statistique

Dans le cadre de notre étude, il est nécessaire de nous intéresser aussi aux travaux de recherche en didactique qui s'intéressent à la modélisation dans le cas particulier des probabilités et des statistiques. La modélisation en statistique a été largement abordée dans les travaux de J-C. Régnier comme nous l'avons mentionnés ci-dessus. Par ailleurs d'autres travaux abordent également ce point.

Dans sa thèse Doukhan (2021), étudie l'enseignement et l'apprentissage des probabilités à la transition entre le secondaire et le supérieur en France. Pour l'auteure, la modélisation doit faire partie de l'activité de l'élève et qu'il est important qu'une partie du processus de modélisation soit à la responsabilité de l'élève. Elle observe que l'activité de modélisation est plus présente au lycée qu'en première année d'université pour des étudiants inscrits dans des filières de biologie, en termes de nombre d'exercices dont l'énoncé est contextualisé. Elle note que, dans le secondaire l'activité de modélisation concernant la probabilités conditionnelles se fait souvent à travers un arbre de probabilités, mais cela se fait autrement à l'Université. A l'Université, il y a des exercices plus théoriques. En revanche lorsque les exercices sont contextualisés, l'activité de modélisation est plus à la charge des élèves : ils doivent identifier des événements, dans certains cas introduire une variable aléatoire alors que ces éléments souvent donnés d'emblée au lycée.

Sotura (2017) s'intéresse à l'enseignement des probabilités et de la statistique au lycée en France, en particulier aux nécessités de modélisation et le développement de la prise d'initiative. Ce texte relate un stage de l'IREM de Paris Diderot autour des enjeux de l'enseignement des probabilités et de la statistique au lycée en France. Lors du stage, les formateurs proposent des situations où les stagiaires vont prendre des initiatives de modélisation et concernant le mode de résolution. Pour Sotura (2017), c'est plus facile d'aider les élèves à appliquer une formule qu'à modéliser mais le fait d'apprendre aux élèves à exécuter des calculs sans qu'ils puissent en mesurer le sens et la portée réduit l'intérêt de leur activité. Pour elle, l'initiation des élèves à une démarche de modélisation peut se faire en leur proposant des situations problématisées. Elle souligne qu'il est nécessaire en dehors des évaluations de laisser aux élèves l'initiative de la modélisation du problème dans le cadre de séances en classe ou à la maison.

2.4 Synthèse du chapitre 2

Dans ce chapitre nous avons considéré différents travaux historiques et didactiques, qui éclairent notre étude. Nous notons ici brièvement les points qui nous semblent le plus importants.

- L'approche historique nous a montré que la problématique de l'ajustement est ancienne et importante. La méthode des moindres carrés a eu énormément de succès très rapidement dans différents domaines. Cependant, elle est aussi restée contestée, et l'histoire des mathématiques nous invite donc à nous interroger : est-ce que dans l'enseignement de l'ajustement, les élèves auront l'occasion de se questionner sur la meilleure méthode d'ajustement ? Est-ce que lorsque la méthode des moindres carrés est enseignée, une explication est fournie concernant les erreurs qui sont minimisées ? ,
- Les élèves rencontrent de nombreuses difficultés dans l'apprentissage de la statistique. Notamment, les différents types de représentations sont très importantes en statistiques et les élèves ont du mal à interpréter les représentations proposées, à passer d'une représentation à une autre etc. ;
- Les enseignements orientés vers l'investigation et basés sur des projets aident les élèves à construire le sens des concepts et des méthodes statistiques. Par exemple, lorsque les élèves ou étudiants réalisent leur propre recueil de données, ou au moins travaillent sur des données réelles. Ceci nous conforte dans notre intérêt pour l'étude des prises d'initiatives des élèves dans l'enseignement de la statistique.
- En lien avec le point précédent, se pose la question de la modélisation. Dans le contexte centrafricain, pour faire un ajustement affine, il faut réaliser une modélisation. En considérant les étapes du cycle de Blum, on peut penser que très probablement toutes ces étapes ne seront pas forcément à la charge des élèves au moins dans l'enseignement ordinaire, contrairement au cas de certains enseignements orientés vers l'investigation et analysés dans des travaux cités ci-dessus. Dans l'enseignement ordinaire en Centrafrique, la situation de départ est certainement déjà simplifiée. Est-ce que les élèves seront en charge de proposer le modèle mathématique ? De produire les résultats mathématiques ? De les interpréter, de les comparer à la situation réelle ? Est-ce que le choix du modèle sera interrogé ? L'un des intérêts du thème de l'ajustement affine est qu'on peut discuter de savoir si réaliser un ajustement est pertinent ou non, nous regarderons de près dans notre travail si cette étape est présente et comment les élèves peuvent la réaliser.
- Finalement, nous notons que de nombreux travaux montrent l'intérêt d'outils logiciels spécifiques qui peuvent soutenir la modélisation statistique. Malheureusement dans notre contexte cette piste ne peut pas être exploitée, l'environnement matériel ne permettant pas le recours à des logiciels.

3 CADRES THÉORIQUES ET QUESTIONS DE RECHERCHE

Dans ce chapitre nous présentons les cadres théoriques que nous utilisons et les questions de recherche de notre travail. Il s'agit des concepts et outils théoriques issus de différentes théories de didactique des mathématiques. Ces concepts et outils théoriques nous servent ensuite à élaborer nos questions de recherche, à guider nos choix méthodologiques et une grande partie de nos analyses repose sur ces outils et concepts.

Pour réaliser cette étude, nous prenons appui sur des concepts issus de la théorie des situations didactiques TSD de Brousseau (1998), de la théorie anthropologique du didactique (TAD) de Chevallard (1998) et sur les registres de représentations sémiotiques de Duval (1993).

Nous allons d'abord présenter les différents concepts issus de la théorie des situations didactiques, en lien avec le partage de responsabilité (contrat didactique) ou pour définir l'initiative des élèves (situation a-didactique). Nous mobilisons également certains concepts de la théorie anthropologique du didactique qui nous permettront d'analyser le contenu des manuels, des problèmes proposés aux élèves et le programme officiel de la RCA (l'approche praxéologique). L'analyse praxéologique nous sert aussi à la conception de notre modèle praxéologique de référence (MPR), qui nous permet de répondre à nos questions de recherche, soit pour définir l'initiative des élèves. Enfin, nous faisons recours aux registres de représentations sémiotiques de Duval (1993), car ils sont nécessaires pour pouvoir prendre en compte les représentations qui jouent un grand rôle dans l'enseignement et l'apprentissage de la statistique, comme nous l'avons vu dans notre revue de travaux au chapitre 2.

3.1 Théorie des situations didactiques

3.1.1 *Le contrat didactique*

Pour analyser la responsabilité laissée aux élèves, pendant l'apprentissage des séries statistiques doubles et l'ajustement affine, nous nous appuyons en particulier sur les concepts de contrat didactique au sens de Brousseau (2004). Le contrat didactique trouve son origine vers les années 80, lorsque Brousseau essaie d'expliquer l'échec électif des élèves en mathématiques. Brousseau (1998) définit la notion de contrat didactique comme :

l'ensemble des obligations réciproques et des sanctions que chaque partenaire de la situation didactique

- impose ou croit imposer, explicitement ou implicitement, aux autres

- et celles qu'on lui impose ou qu'il croit qu'on lui impose (Brousseau 1998, p.6).

Ainsi l'enseignant attend de la part de l'élève la mise en œuvre de la connaissance visée et de voir l'élève s'engager dans le problème qui lui est dévolu (Brousseau, 1983). Selon l'auteur, le contrat est l'ensemble des règles entre l'enseignant et les élèves ; l'établissement de ces règles peut être vue comme une négociation :

Cette négociation produit une sorte de jeu, dont les règles provisoirement stables, permettent aux protagonistes et notamment à l'élève de prendre des décisions dans une certaine sécurité, nécessaire pour lui assurer l'indépendance caractéristique de l'appropriation. Des relations de pouvoir de chacun des protagonistes sur l'autre doivent nécessairement s'établir et s'accepter à titre de rétroactions. Ce pouvoir doit être fondamentalement limité et doit évoluer pour permettre le fonctionnement du système enseigné (Brousseau, 1986, p.55).

En terme de partage des responsabilités relatives aux savoirs entre le professeur et les élèves, cela peut changer et évoluer au cours de la construction et de la validation de la connaissance :

Répartition des responsabilités entre professeurs et les élèves est donc à comprendre comme la répartition entre ce que l'enseignant assure quant aux décisions mathématiques et ce qu'il laisse à la charge de l'élève. Ainsi, étudier la répartition des responsabilités par rapport aux connaissances mathématiques entre le professeur et les élèves consiste à étudier la part du travail mathématique que l'enseignant laisse à la charge des élèves pour leur permettre de construire, en confrontation avec la situation, leurs connaissances mathématiques (Hersant, 2013, p.90).

Par ailleurs, le contrat didactique définit également le partage des responsabilités entre les élèves et l'enseignant, vis-à-vis du savoir en jeu.

Le contrat didactique est l'ensemble des règles, souvent implicites, qui fixent les responsabilités respectives des élèves et du professeur, engagés dans une situation d'enseignement-apprentissage des mathématiques. On ne retient parmi ces règles que celles qui sont spécifiques des contenus mathématiques en jeu. (Dorier et al, 2018, p.501).

Pour notre travail, le concept de contrat didactique correspond au partage de responsabilités entre l'enseignant et les élèves par rapport au savoir (contenu mathématique) (Hersant, 2014). Tout ce qui est à la charge de l'élève, par exemple, pour prouver ou valider une proposition, ou produire la solution d'un problème, ou encore la validation d'un raisonnement au sein d'un groupe d'élèves sont considérés comme la part de la responsabilité de l'élève relative au savoir mathématique.

3.1.2 Situation a-didactique

La notion d'initiatives des élèves est en lien avec la notion de "situation a-didactique" introduite par Brousseau. Une situation non didactique est une situation qui n'est pas organisée avec un objectif d'apprentissage. Une situation est dite a-didactique lorsqu'il y a un objectif d'apprentissage, mais que celui-ci disparaît derrière la nécessité de mobiliser le savoir, nécessité qui provient de la situation :

Le maître se refuse d'intervenir comme possesseur des connaissances qu'il veut voir apparaître. L'élève sait bien que le problème a été choisi pour lui faire acquérir une connaissance nouvelle mais il doit savoir aussi que cette connaissance est entièrement justifiée par la logique interne de la situation. (Brousseau, 2004, p.59)

Dans une situation a-didactique, c'est à l'élève de formuler des stratégies, de prendre des décisions. Pour cela, l'enseignant fait en sorte que les élèves acceptent leur part de responsabilité dans l'apprentissage. Dans notre étude nous interrogeons le potentiel a-didactique des situations rencontrées par les élèves dans l'étude de séries statistiques doubles. Par exemple, lorsque les élèves

réalisent eux-mêmes une enquête et en traitent les données, il s'agit d'une situation a-didactique, dans laquelle l'élève a de grandes responsabilités et prises d'initiatives possibles. Mais nous verrons que ceci n'est pas présent dans l'enseignement en Centrafrique.

3.2 Théorie Anthropologique du didactique

Nous utilisons par ailleurs le cadre de la théorie anthropologique du didactique (TAD) de Chevallard (1998), à travers l'approche praxéologique. Selon la TAD, toute activité humaine consiste à accomplir une tâche d'un certain type T , dont-il existe une manière d'accomplir T (technique τ), justifiée par une technologie θ qui permet d'expliquer, voire de la produire, et qui à son tour est justifiable par une théorie notée Θ .

La praxéologie, qui peut également être nommée organisation praxéologique ou organisation mathématique (OM) est définie par un quadruplet $[T/\tau/\theta/\Theta]$ avec la composante praxis $[T/\tau]$, qui signifie « pratique » (savoir-faire), décrit les techniques τ permettant d'accomplir certains types de tâches T . et la composante logos $[\theta/\Theta]$ qui signifie « raison » (savoir). La mise en œuvre de cette approche praxéologique pour l'analyse de savoirs s'organise autour des types de tâches T . Pour un type de tâches T donné, une praxéologie relative à T précise une manière d'accomplir, de réaliser les tâches appartenant au type T : à une telle manière de faire, τ , on donne ici le nom de technique (du grec tekhnê, savoir-faire) (Chevallard, 1998). On entend par technologie, et on note généralement θ , un discours rationnel (logos) sur la technique la tekhnê τ , discours ayant pour objet premier de justifier « rationnellement » la technique τ , en nous assurant qu'elle permet bien d'accomplir les tâches du type T , c'est-à-dire de réaliser ce qui est prétendu (Chevallard, 1998). Donc la technologie a pour rôle de justifier, d'expliquer la technique utilisée. Et toute technologie a besoin à son tour d'une justification, que l'on appelle la théorie associée à cette technologie technique Θ .

Nous utilisons la TAD pour analyser les praxéologies proposées aux élèves, dans des manuels scolaires ou dans des sujets de devoirs. Nous utiliserons par ailleurs le concept de *sous-type de tâches* noté T' défini par Chaachoua (2018) dans son cadre théorique T4TEL. Soit T un type de tâches, T' est à la fois un sous-ensemble de T et un type de tâches. Chaachoua (2018) utilise plusieurs types de tâches pour décrire la technique, qu'il appelle *ingrédients de la technique*.

Par exemple : « réaliser un ajustement affine avec la méthode de Mayer » est un sous-type de tâches, du type de tâches « réaliser un ajustement affine ». Les types de tâches suivants : Calculer les coordonnées des deux points moyens G_1 et G_2 associés aux deux sous-nuages ; Placer les points moyens G_1 et G_2 sur le graphique; Tracer la droite (G_1G_2) ; Déterminer une équation de la droite (G_1G_2) , deviennent des ingrédients de techniques pour ce type de tâches.

Dans ce travail, nous allons concevoir un modèle praxéologique de référence (MPR) pour l'enseignement des séries statistiques doubles et de l'ajustement affine. Le concept de MPR a été développé au sein de la TAD.

Une Organisation Mathématique 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. Notons que ce dernier, pour des raisons liées à sa problématique, peut bien entendu procéder à un autre découpage que celui de l'institution voire compléter ; il construit alors un modèle praxéologique de référence (MPR) regroupant les praxéologies à enseigner, enseignées mais également enseignables. (Chaachoua, Ferraton, & Desmoulins, 2017, p.302).

Dans le cadre de notre travail, notre institution est la classe de terminale SE en Centrafrique.

Comme le soulignent Chaachoua et al. (2017) ci-dessus, un MPR est construit par le chercheur en fonction de ses objectifs. Dans notre cas, nous allons utiliser l'étude historique, la revue des travaux, le programme officiel de mathématique de Centrafrique mais également un manuel de BTS de France (nous expliquons ce choix méthodologique au chapitre 4) pour déterminer les praxéologies à enseigner mais également les praxéologies enseignables. Ce MPR conçu pour notre analyse, nous servira à répondre à nos questions de recherche, nous le comparerons aux praxéologies effectivement enseignées en Centrafrique.

3.3 Les registres de représentations sémiotiques

Duval (1993) appelle registre de représentation tout système sémiotique permettant d'accomplir les trois activités cognitives fondamentales à la sémiosis : la formation d'une représentation identifiable, le traitement et la conversion. La formation d'une représentation identifiable, constitue une trace ou un assemblage de traces perceptibles qui soient identifiables comme une représentation de quelque chose dans un système déterminé. Le traitement, c'est une transformation interne à un même registre de représentation. Et convertir les représentations produites dans un système en représentation d'un autre système, de telle façon que ces dernières permettent d'explicitier d'autres significations relatives à ce qui est représenté ; c'est ce que Duval appelle la conversion.

En effet dans le domaine des séries statistiques à deux variables, les représentations jouent un rôle très important. La série est souvent donnée au départ sous forme d'un tableau. A ce niveau, elle n'est pas accessible pour faire des prévisions. Par conséquent, des conversions sont nécessaires pour pouvoir passer du registre tableau au registre graphique ou algébrique, en vue de faire des prévisions.

Block et al (2012) ont essayé de mettre en évidence l'importance des registres de représentation, lors d'une étude sur les problèmes numériques en fin de première année dans trois pays (France, Grèce, Mexique). A ce propos, ils ont étudié l'influence que joue le registre numérique des écritures chiffrées et le registre graphique de représentation des quantités discrètes utilisées dans l'énoncé d'une tâche sur les démarches de résolution, c'est-à-dire, ils essaient de faire un lien entre les registres de représentation en jeu dans l'énoncé et les démarches des élèves. Il ressort de cette étude que, si le type de tâches est déjà travaillé, ce qui importe, c'est la maîtrise de la conversion du registre graphique de représentation des quantités discrètes en registre numérique des écritures chiffrées. Quand le type de tâches n'est pas encore travaillé, ce qui importe, c'est de bien coordonner les deux registres, qui vont au delà de la maîtrise des conversions. Dans notre travail, nous tenterons de même de mettre en relation le ou les registres employés dans les énoncés et les stratégies des élèves. Par ailleurs, nous considérons qu'une part importante des prises d'initiatives

des élèves peut concerner la conversion de registres : est-elle à leur charge, ou indiquée par l'énoncé ? Est-ce que les élèves doivent manipuler plusieurs registres, faire preuve de flexibilité d'un registre à l'autre ?

3.4 Les initiatives des élèves

Dans cette partie, nous essayons de définir ce que nous appelons ici "*initiatives des élèves*". Pour ce faire nous considérons différents travaux, en lien avec ces initiatives : soit parce qu'ils donnent des outils pour évaluer le potentiel de prise d'initiative que recèle un énoncé (Robert, 1998), soit parce qu'ils concernent des modalités d'enseignement permettant des prises d'initiatives.

3.4.1 Les niveaux de mise en fonctionnement des connaissances

Dans le contexte du développement de recherches en didactique des mathématiques portant sur des savoirs « avancés » à enseigner au lycée et à l'université. Robert (1998) propose de nouveaux outils théoriques et méthodologiques pour analyser des notions mathématiques à enseigner au lycée et à l'Université. Dans cet article, l'auteure présente différentes dimensions pour l'analyse des savoirs à enseigner. Nous retenons particulièrement les niveaux de mise en fonctionnement des connaissances. L'article présente trois niveaux de mise en fonctionnement des connaissances des élèves. Ils s'agit d'abord du niveau technique. Ce niveau consiste à mettre en jeu de manières immédiates les théorèmes, des formules, des propriétés et des définitions. Il concerne le fonctionnement des outils. Ensuite le niveau des connaissances mobilisables. Il correspond à des mises en fonctionnement larges, impliquant les caractères outil et objet. Il met en jeu ce qui est explicite. Et enfin, le niveau des connaissances disponibles ; ce niveau consiste à appliquer des méthodes non prévues, c'est-à-dire chercher soi-même dans ses connaissances ce qui peut intervenir. Donc la logique de la réussite des élèves se situe au niveau technique, tandis que la logique d'apprentissage se place au niveau des connaissances mobilisables et disponibles qui nécessite une réorganisation du savoir. Il ressort de cet article que, les différents niveaux de mise en fonctionnement peuvent contribuer à définir les initiatives des élèves.

3.4.2 Résolution de problème dans l'enseignement secondaire

Quand l'élève arrive à comprendre les consignes et à s'engager vers une solution pour faire naître des connaissances, cela peut être qualifié d'une prise d'initiative de la part de l'élève. La résolution d'un problème peut se faire après avoir appris une nouvelle connaissance ou bien avant, ou pendant une séquence. Donc nous supposons que le processus de résolution de problèmes constitue un objet d'apprentissage en soi.

La résolution de problème peut être vue comme un moyen d'apprentissage de savoirs mathématiques. D'après Coppé et Houdement (2009), la résolution de problèmes est un moyen de contrôler les connaissances acquises des élèves. Elle permet la construction de nouveaux outils

mathématiques. Et elle occupe une place très importante dans les programmes français. Pour donner du sens aux savoirs, nous pouvons soumettre les élèves à la résolution d'un problème, et ils peuvent acquérir diverses compétences mathématiques par la résolution de problèmes.

Les initiatives des élèves dans la résolution des problèmes supposent une appropriation et la résolution du problème par l'élève, selon ses propres représentations. On retrouve ici une perspective proche de celle présente dans la théorie des situations didactiques avec la notion de situation a-didactique. Cependant tous les problèmes proposés aux élèves n'ont pas forcément les caractéristiques d'une situation a-didactique. En revanche, en utilisant la classification de Robert (1998), on peut dire qu'ils dépassent un simple niveau technique pour la mise en fonctionnement des connaissances.

3.4.3 Notion de topos des élèves

Nous pouvons utiliser aussi la TAD pour décrire les activités de l'élève dans la situation d'apprentissage dans une institution. Les élèves doivent accomplir des tâches relevant de certains types de tâches, associés à certaines praxéologies. Il y a une négociation de la répartition des responsabilités entre le professeur et les élèves (topogénese). Ainsi la notion de topos de l'élève, définie comme l'ensemble des gestes que l'élève aura à accomplir en autonomie didactique (Chevallard, 2001) peut aussi nous informer sur les initiatives des élèves.

3.4.4 Démarche d'investigation

La démarche d'investigation vise à développer l'autonomie des élèves, en construisant leurs apprentissages dans la résolution de problèmes. Pour résoudre un problème, ils confrontent leurs points de vue, ils partagent leurs idées, ils déploient selon des choix qui leur sont propres et en autonomie, des représentations qui favorisent la prise d'initiative. La résolution d'un exercice permet à l'élève de s'investir dans son apprentissage personnel. Et pour être acteur de son apprentissage, il apprend à trouver par lui-même ou en équipe les techniques, les stratégies, les méthodes, les modèles adéquats et les concepts utiles pour construire sa connaissance.

D'après Coppé (2012), la démarche d'investigation est un outil pédagogique qui permet de développer l'autonomie des élèves, laissant une grande autonomie aux élèves et la responsabilité de faire avancer le savoir dans le processus d'enseignement/ apprentissage.

Pour Monod, et al. (2010), le facteur qui influence beaucoup plus l'autonomie est la motivation des élèves. Et il y a aussi, les qualités propres des élèves, leur imagination, leur inventivité. L'organisation d'interactions entre élèves et groupes d'élèves est importante dans la mise en œuvre de démarches d'investigation dans la classe. Elle peut contribuer à développer l'aspect social de l'autonomie.

3.4.5 Proposition de définition des initiatives des élèves

Dans cette partie, nous précisons ce que nous appelons ici *initiatives des élèves*. Nous avons vu que le « topos » de l'élève désigne dans la TAD ce qui est à la charge de l'élève, sous sa responsabilité dans une praxéologie : le choix d'une technique parmi plusieurs, ou simplement l'application d'une

technique qui est indiquée ; la justification de cette technique etc. Cependant le topos, les responsabilités des élèves et la notion de prise d'initiatives ne coïncident pas. En effet, d'une part dans ce topos il peut y avoir des activités techniques qui ne requièrent pas d'initiatives de la part des élèves. Par exemple, la représentation graphique du nuage de points où l'élève n'a pas à faire le choix de l'échelle ne relève pas des prises d'initiatives mais de l'application d'une technique. Par ailleurs, il peut y avoir des prises d'initiatives des élèves qui ne relèvent pas des responsabilités qui leur sont attribuées en termes de praxéologies : par exemple proposer une comparaison entre un énoncé et une situation familière, qui n'était pas nécessairement attendu par l'institution. C'est pourquoi nous proposons une définition spécifique.

En lien avec la modélisation, nous considérerons qu'il y a des initiatives des élèves possibles à différentes étapes. Lorsqu'il faut choisir un modèle mathématique ; lorsqu'il faut prendre des informations dans l'énoncé pour leur appliquer le modèle mathématique choisi, parfois en transformant les données de l'énoncé ; lorsqu'il faut interpréter les résultats fournis par le modèle mathématique, voire changer ensuite de modèle, selon le cycle de modélisation de Blum et LeïB (2005, reprise 2018).

Du point de vue des registres, nous considérerons les changements de registre également comme des prises d'initiatives des élèves. Ces prises d'initiatives sont limitées si la nécessité de changement de registre est indiquée dans l'énoncé, ou habituelle pour un certain type d'exercices. Elle est plus importante si l'élève décide par lui-même de changer de registre.

Nous proposons de considérer les initiatives des élèves comme des processus issus des actions et décisions de l'élève, en rapport avec le savoir. La prise d'initiative des élèves consiste à mobiliser des connaissances, et à formuler des stratégies lors de la résolution des problèmes.

Nous distinguerons la prise d'initiatives potentielle que l'on peut observer par exemple en analysant le manuel. D'autre part, la prise d'initiatives effective qui relève de la responsabilité des élèves, pendant des activités.

Nous venons de présenter notre cadre théorique et de justifier sa pertinence au regard des phénomènes d'enseignement/apprentissage que nous voulons analyser. Ce faisant, nous présentons notre problématique et les questions de recherche en lien avec les différents éléments théoriques présentées ci-dessus.

3.5 Questions de recherche

Les questions initiales posées au départ de notre étude étaient : Quelles sont précisément les difficultés des élèves avec les séries statistiques doubles et l'ajustement affine, en Centrafrique ? Est-ce que l'enseignement est responsable de ces difficultés, et comment pourrait-on les surmonter ? Nous faisons l'hypothèse que l'enseignement existant en Centrafrique laisse peu de place à la construction du sens et aux initiatives des élèves, et qu'en leur laissant plus d'initiatives les élèves maîtriseraient mieux la statistique en arrivant dans l'enseignement supérieur.

Les cadres théoriques présentés ci-dessus nous amènent à choisir un regard particulier sur ces questions de départ, qui va nous mener à nos questions de recherche. Dans cette étude, nous nous intéressons aux initiatives des élèves dans les organisations praxéologiques pour le thème de la

statistique et plus précisément de l'ajustement affine pour des séries statistiques doubles. En particulier nous nous intéressons aussi aux initiatives laissées aux élèves dans le travail de modélisation qui est très souvent présent en séries statistiques doubles et l'ajustement affine. Nous considérons ce sujet sous deux angles : celui de l'enseignement tel qu'il est habituellement pratiqué et des initiatives qu'il peut induire ; et celui des initiatives effectives des élèves, dans le cadre de problèmes conçus pour leur laisser des responsabilités significatives.

C'est en nous appuyant sur les différents cadres théoriques précédemment présentés, que nous en arrivons aux questions suivantes :

- QR.1 : Quel est le savoir à enseigner au sujet des séries statistiques doubles et de l'ajustement affine en Centrafrique terminale SE.
- QR.2 : Quelles sont les prises d'initiatives possibles des élèves relativement à ce savoir ? Quel est le savoir effectivement enseigné à ce sujet dans l'enseignement ordinaire ?
- QR.3 : Quelles sont les prises d'initiatives effectives des élèves de terminale en Centrafrique, lorsqu'on leur propose un exercice sur le thème de l'ajustement affine qui permettrait une telle initiative ? A quelles conditions peut-on mettre en œuvre en Centrafrique un enseignement qui laisse plus d'initiative aux élèves que ce qui est pratiqué dans l'enseignement ordinaire ?

Dans le chapitre suivant, nous présentons la méthodologie que nous avons élaborée pour répondre à ces questions.

4 MÉTHODOLOGIE DE CONSTRUCTION ET D'ANALYSE DE DONNÉES

Dans ce chapitre, nous nous proposons de décrire les choix méthodologiques qui nous ont permis de recueillir et d'analyser des données en lien avec la notion des séries statistiques doubles et l'ajustement affine. Nous considérons successivement les différents aspects de notre étude à travers lesquels nous cherchons à répondre aux questions de la recherche.

4.1 Méthodologie de la construction d'un MPR

Pour construire notre Modèle praxéologique de référence (MPR), notre méthodologie est basée sur l'étude historique, la revue des travaux, l'analyse du programme scolaire officiel de Centrafrique de 2002, ainsi que l'analyse d'un manuel de BTS utilisé en France. En effet le programme de mathématiques de la classe de terminale SE de 2002 n'est pas détaillé, il ne nous permet pas d'aller plus loin dans nos analyses. Le choix du manuel mathématique comptabilité et gestion informatique de gestion (2000), utilisé en France en année de BTS est dû au fait que, le programme du manuel cadre bien avec le programme de la statistique descriptive enseigné en Centrafrique en classe de terminale SE. Nous considérons le contenu de ce manuel sous l'angle de la Théorie anthropologique du Didactique TAD (Chevallard, 1998), pour décrire en termes de types de tâches, des techniques nécessaires à la réalisation des types de tâches, les technologies et les théories. Ces praxéologies nous permettent de concevoir notre modèle praxéologique de référence, que nous allons ensuite employer dans l'analyse du manuel CIAM (Collection Inter Africaine de Mathématiques) utilisé en Centrafrique.

Pour le manuel de BTS de France, nous avons fait une analyse globale des contenus du chapitre 11, qui porte sur les séries statistiques à deux variables, ainsi que l'analyse des exercices en terme de praxéologies. Pour l'analyse praxéologique des exercices, nous cherchons à identifier dans chaque question le type de tâches en lien avec notre question de recherche, la technique associée à ce type de tâches et la technologie qui justifie la technique utilisée.

4.2 Méthodologie de l'analyse du manuel de Centrafrique

Nous avons retenu pour notre étude le manuel Mathématiques Terminale SE (CIAM) utilisé en Centrafrique. Ce choix est dû au fait que, le programme du manuel de mathématiques CIAM est le manuel officiel utilisé en Centrafrique.

Nous nous intéressons au chapitre sur les séries statistiques doubles. Nous avons fait l'usage du MPR construit précédemment pour l'analyse de la partie cours et de la partie exercices. Il s'agit pour nous de voir si les types de tâches et sous-types de tâches apparaissant dans le MPR sont présents dans le manuel, et si certains types de tâches du manuel ne seraient pas dans le MPR. Par la suite, nous faisons l'analyse d'un exercice résolu en terme de praxéologie (Chevallard, 1998), et nous revenons sur le MPR. Nous précisons là où se situent les prises d'initiatives laissées à l'élève dans les types de tâches et sous-types de tâches lors d'un exercice.

4.3 Méthodologie de l'analyse d'un enseignement ordinaire: méthodologie spécifique à l'enseignement ordinaire en RCA

4.3.1 Observation de séances de classe

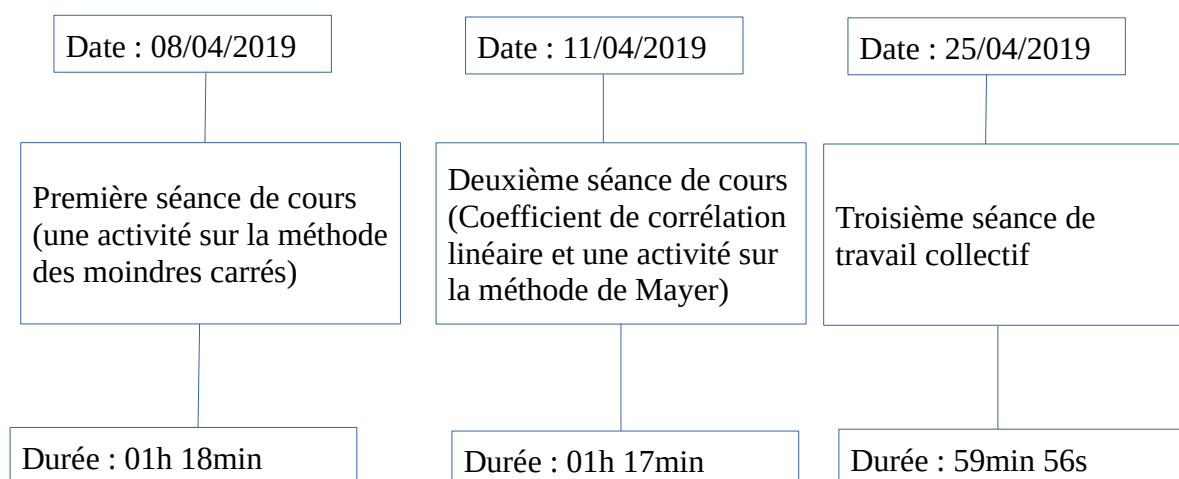
Pour cette partie relative à l'enseignement « ordinaire », nous avons tout d'abord voulu observer les pratiques d'un enseignant. Ces observations ont eu lieu en 2018-2019. L'enseignant observé est un enseignant volontaire, il détient un CAPES en Mathématiques. Il a une expérience de plus de quinze ans d'enseignement, il assure les cours de mathématiques en terminale depuis plus de dix ans. C'est un enseignant que nous jugeons expérimenté, dans la suite de la thèse nous le nommons Aristide (pseudonyme). L'effectif de la classe est de 146 élèves.

Les séances d'enseignement observées, dédiées aux séries doubles sont d'un volume horaire total de trois heures dix-huit minutes, dont deux séances de cours et une séance de travail collectif en petits groupes d'élèves. Dans notre analyse, nous ne nous focalisons que sur les deux premières séances. En effet, la troisième séance concerne un problème que nous avons nous-même proposé, elle ne relève donc pas de l'enseignement « ordinaire ».

Aristide organise les trois séances de la manière suivante :

- La première séance de cours est faite sous forme d'activité. L'enseignant propose un problème sur la méthode des moindres carrés, problème représente une situation de la vie réelle. Les concepts comme nuage de points, point moyen, série double sont définis pendant l'activité. Durant la séance, par moment l'enseignant envoie un élève au tableau pour réaliser une tâche ou il pose oralement des questions aux élèves ;
- La deuxième séance de cours est répartie sur deux moments. Un premier moment où l'enseignant introduit la notion du coefficient de corrélation linéaire, cette dernière est associée à des connaissances déjà vues pendant la première séance et le deuxième moment est consacré à un problème portant sur la méthode de Mayer et dans un contexte non réel ;
- La troisième séance est consacrée à un travail collectif en petits groupes sur un exercice, qui a duré environ une heure.

Nous résumons ci-dessous ces séances observées en classe de terminale SE avec Aristide.



Le recueil des données est basé sur l'enregistrement vidéo des trois séances, que nous avons recueillies en classe de terminale SE en Centrafrique, au lycée d'application de l'école normale supérieure (LAENS) au cours de l'année scolaire 2018-2019. Nous avons transcrit les vidéos des deux premières séances, et réalisé un synopsis de ces séances. Les transcripts et les synopsis figurent dans les **annexes A** (à retrouver en annexe A.1 et A.2) de la thèse. Nous avons de plus choisi des épisodes significatifs dont nous faisons l'analyse détaillée.

4.3.2 Questionnaire et entretiens avec des enseignants

Le questionnaire est un outil de recueil des données produit par et pour le chercheur, qui repose sur du déclaratif. Il convient de garder à l'esprit que le questionnaire n'a pas pour but de questionner la pratique enseignante, mais plutôt d'appréhender les points de vue des enseignants, au moyen des questionnaires déclaratifs.

Nous avons conçu un questionnaire pour des enseignants de mathématiques en Centrafrique. Le questionnaire est constitué de douze questions ouvertes et fermées (Annexe B). Le questionnaire a été soumis début mai 2020. En tout, nous avons distribué seize questionnaires et dix enseignants nous ont restitué leur questionnaire. Nous avons laissé aux enseignants environ deux semaines pour répondre aux questions. Dans le cas de l'enseignant Aristide, il nous a rendu le questionnaire un peu plus tard. Trois de ces professeurs nous ont accordé un entretien plus tard après avoir répondu au questionnaire, dans le but d'avoir des informations complémentaires sur leurs réponses et sur certaines raisons de leurs choix d'enseignement.

Présentation du questionnaire

Tout d'abord, les questions du questionnaire concernent les caractéristiques professionnelles de ces professeurs. Ensuite, elles sont centrées sur l'avis personnel des professeurs sur un problème proposé aux élèves, ce qui nous permettra de mieux comprendre les décisions des professeurs. Et enfin, nous interrogeons les professeurs sur l'enseignement de la statistique à deux variables.

Le questionnaire est constitué de douze questions réparties en trois thèmes. Le premier thème est constitué de 7 questions. 6 sont des questions fermées, certaines sont à choix multiple, où on demande aux enseignants l'année d'expérience en enseignement, le niveau enseigné actuellement ou s'ils utilisent le manuel CIAM (nous rappelons que c'est le manuel mathématiques terminale SE utilisé en Centrafrique) pour préparer leurs cours. La cinquième question est ouverte, et concerne d'autres ressources utilisées par les professeurs. Ces questions sont relatives aux caractéristiques professionnelles des enseignants, comme l'ancienneté dans le métier d'enseignant ou le niveau de la classe enseigné.

Le deuxième thème du questionnaire est constitué de 3 questions, centrées sur la pratique de l'enseignant dans les statistiques à deux variables, avec l'emploi de l'exemple d'un exercice. Dans cette partie du questionnaire, les personnes enquêtées (enseignants) doivent donner leurs points de vue sur l'exercice et sur la production d'élèves. Pour cela, on leur demande s'ils peuvent proposer

un tel type de problème à leurs élèves. Il s'agit d'une question pour savoir le point de vue de l'enseignant sur l'exercice proposé aux élèves. Ensuite, nous avons une question où on demande aux enseignants d'analyser la réussite et les erreurs d'élèves. Cette question nous permet d'avoir une idée sur la manière dont les enseignants évaluent leur élèves. Et la dernière question renvoie à l'avis personnel de l'enseignant dans sa pratique vis à vis des élèves.

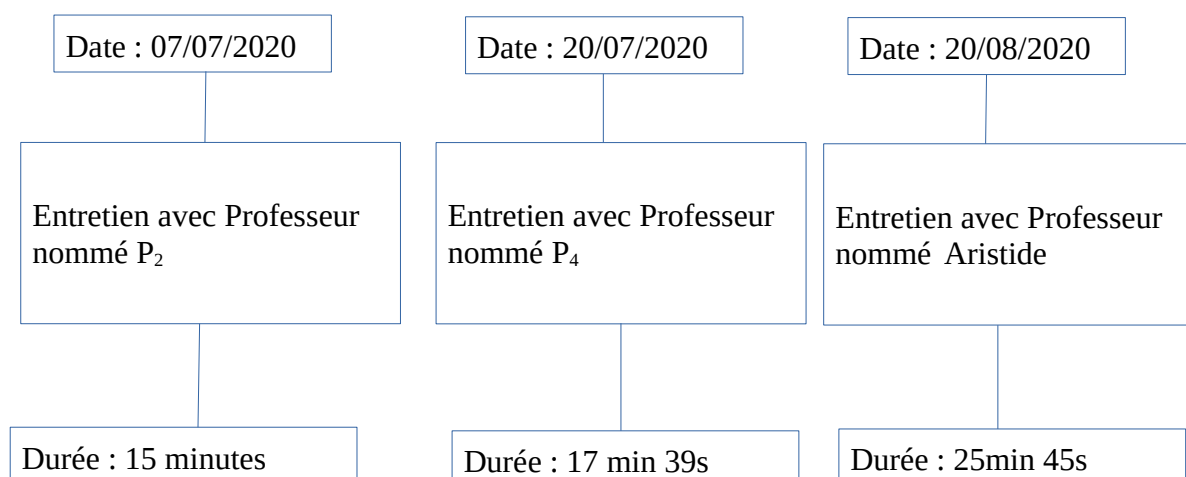
Le troisième et le dernier thème est composé de 2 questions. La première question est à propos de ce qui est important pour l'élève en terme d'objectif à atteindre concernant la série statistique à deux variables. C'est une question à choix multiple où on demande aux enseignants de choisir une réponse parmi plusieurs. Il s'agit de savoir l'objectif fixé par l'enseignant pendant son cours de la série statistique à deux variables. Quant à la dernière question, elle concerne les difficultés éprouvées par les élèves lors de l'apprentissage de la série statistique à deux variables.

Bien que nous n'ayons pas pour objectif de questionner ou d'analyser la pratique de l'enseignant, les réponses à ces questions nous semblent importantes et nous aident à mieux comprendre les points de vue des enseignants sur l'enseignement des séries statistiques doubles et de l'ajustement, et sur les prises d'initiatives d'élèves.

En ce qui concerne l'entretien, il nous semble adéquat d'avoir un complément d'informations afin d'explicitier les réponses des professeurs issues du questionnaire. Nous avons établi une liste de questions permettant d'orienter l'entretien, répartie en trois parties. La première partie fait un retour sur le questionnaire rempli par l'enseignant, la seconde est relative à la pratique d'enseignement de la statistique à deux variables, et la troisième partie concerne les initiatives des élèves en statistique à deux variables.

Le choix de ces trois professeurs est basé sur leur ancienneté (17 à 31 ans d'enseignement). Nous avons réalisé les entretiens de manière individuelle. Nous avons interviewé trois enseignants, nommés P₂, P₄, au mois de Juillet et Aristide au mois d'Août.

Nous résumons les détails des entretiens avec ces enseignants dans le tableau ci-dessus.



4.4 Méthodologie de conception et de l'analyse pour l'expérimentation en classe

Nous avons mené successivement trois expérimentations, au cours des années 2019 ; 2020 et 2021. Chaque expérimentation a été construite à partir des résultats de la précédente.

4.4.1 Méthodologie de conception et de l'analyse de la première expérimentation en classe « Préviation du ménage 1 »

Suite aux analyses de manuels, nous avons conçu en 2019 un devoir centré sur la notion d'estimation. Ce devoir comportait un unique exercice, dont le but était l'estimation des dépenses d'un ménage en 2019. L'énoncé est présenté ci-dessous, nous en ferons l'analyse a priori en partie 6.1.2

Exercice

Le tableau suivant indique en euros les montants des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense qu'en 2019, il dépensera moins de 10000 euros. Qu'en pensez-vous de cette prévision ? vous expliquerez en détail la méthode que vous avez choisie. Puis justifier pourquoi un ajustement affine est correct (ou approprié).

L'activité que nous avons élaborée a été mise en œuvre le 25 avril 2019 dans la classe de terminale SE, au lycée d'application de l'école normale supérieure de Bangui (LAENS). L'effectif des élèves était de 146, mais seulement 84 élèves ont accepté de prendre part à l'activité. Les élèves interagissent en groupes pour résoudre un exercice en temps limité de 1 heure. Nous avons 16 groupes. Les élèves sont répartis en nombres de 5 ou 6 par groupes. Nous avons ainsi récupéré les productions écrites des élèves. Nous avons analysé ces productions d'élèves en utilisant des critères qui seront présentés dans le chapitre 6.

4.4.2 Méthodologie de conception et de l'analyse de la deuxième expérimentation en classe

Après avoir analysé les productions des élèves de la première expérimentation de 2019, nous avons conçu un second devoir qui s'appuyait sur les résultats de nos analyses. Ce devoir comportait trois exercices, le texte du devoir se trouve en (**annexe E.1**)

En 2020, ce nouveau devoir a été proposé à des élèves du lycée d'Application de l'École Normale Supérieure de Bangui (LAENS), le lundi 31 Août 2020 à 7h30min (durée : 2h). L'effectif total des élèves est de 76, mais 66 élèves sont présents à ce devoir. Ils sont répartis en groupe de cinq pour certains et six pour d'autres. Les élèves interagissent en groupes pour résoudre ensemble les tâches, et se mettre d'accord sur leurs réponses communes. Dans certains groupes, les élèves discutent sur ce qu'ils vont mettre sur la feuille qui sera rendue avant que le secrétaire du groupe n'écrive sur la feuille. Et dans d'autres groupes s'ils ne s'entendent pas bien sur une question, alors chaque élève du groupe écrit quelque chose et ils mettent en commun leur résultat, avant d'avancer. Nous avons en tout 12 groupes.

La surveillance a été assurée par nous-même (le chercheur) et le professeur. Remarquons que nous avons autorisé aux élèves qu'ils pouvaient commencer par n'importe quel exercice qui leur semble facile. Pour analyser ces productions d'élèves nous avons élaboré une liste de critères que nous présentons dans le chapitre 6.

4.4.3 Méthodologie de conception et de l'analyse de la troisième expérimentation en classe

La troisième expérimentation menée en 2021 a eu lieu dans deux établissements (un établissement public et un établissement privé) en Centrafrique. Nous nommons la classe de l'établissement public classe 1 et celle de l'établissement privée classe 2. L'effectif de la classe 1 est de 69 élèves et celui de la classe 2 est de 10 élèves.

Nous avons conçu et implémenté l'enseignement sur la notion des séries statistiques à deux variables dans la classe 1. Pour des raisons d'organisation, avant de lancer notre expérimentation, nous avons établi une trame (**annexe F.2**) pour avoir la possibilité de contrôler le temps didactique et de nous adapter en fonction des réactions des élèves. La classe 2 a suivi un cours « ordinaire ».

Dans la classe 1, nous avons d'abord soumis les élèves à un pré-test avant de commencer le cours, et à la fin de notre expérimentation nous avons de nouveau soumis les élèves à un post-test. Par la suite, nous avons passé le même post-test dans la classe 2.

Nous présentons la conception de ce cours expérimental dans le chapitre 6. Nous avons filmé nos propres séances, et nous avons réalisé un synopsis sur la base de ces films.

Dans le cadre de notre travail, nous faisons le choix d'analyser des extraits des trois épisodes dans différentes séances de cours implémenté dans la classe 1. Ce choix est motivé par le fait que ces épisodes différent de ce que nous avons observé pendant l'enseignement ordinaire au cours de l'année scolaire 2018-2019.

Le post-test proposé aux élèves des deux classes était le même que nous avons soumis aux élèves pendant la deuxième expérimentation en 2020. Dans la classe 1 au lycée d'application de l'école normale supérieure (LAENS), le post-test a eu lieu, le 29 avril 2021 et a duré 2 heures. Nous avons réussi aussi à passer ce post-test le 14 août 2021, dans la classe 2 (Lycée Saint Philippe) car nous connaissons le collègue qui enseigne dans cet établissement et qui était volontaire pour le faire passer auprès de ses élèves de la classe de terminale SE.

Date	Données recueillies
Avril 2019	Observations de séances en terminale SE
Avril 2019	Test soumis aux élèves de terminale SE
Mai 2020	Questionnaire soumis aux enseignants de terminale SE
Juillet - Août 2020	Entretien avec trois enseignants de terminale SE
Août 2020	Test soumis aux élèves de terminale SE
Avril 2021	Pré-test soumis aux élèves de terminale SE (lycée public)
Avril 2021	Implémentations de séances en terminale SE
Avril et Août 2021	Post-test soumis aux élèves de terminale SE dans deux établissements (public et privé)

Tableau 1: Méthodologie générale - données recueillies

5 ANALYSE DE L'ENSEIGNEMENT HABITUEL DE L'AJUSTEMENT AFFINE

Dans le chapitre quatre, nous avons présenté les éléments méthodologiques de la thèse. Dans ce chapitre cinq, la question fondamentale de notre travail est celle d'étudier le savoir à enseigner et le savoir effectivement enseigné au sujet des séries statistiques doubles et de l'ajustement affine en Terminale SE en Centrafrique. Nous nous intéresserons aussi aux prises d'initiatives possibles des élèves définies précédemment.

Les questions de recherche que nous étudions ici est :

QR.1 : Quel est le savoir à enseigner au sujet des séries statistiques doubles et de l'ajustement affine en Centrafrique (terminale SE).

Q.R2 : Quelles sont les prises d'initiatives possibles des élèves relativement à ce savoir ? Quel est le savoir effectivement enseigné à ce sujet dans l'enseignement ordinaire ?

Pour apporter des éléments de réponse à QR1 et QR2, nous allons d'abord construire un MPR à l'aide de nos analyses épistémologiques, du programme de Centrafrique et d'un manuel français. Nous comparons aussi ce MPR aux résultats issus de l'analyse du manuel utilisé en Centrafrique. Ensuite nous utilisons notre MPR pour analyser un enseignement fait en Centrafrique. Et enfin nous analysons le point de vue des enseignants. Dans le même temps, ces analyses doivent nous permettre de répondre à notre question de recherche sur la prise d'initiative des élèves.

5.1 Construction du modèle praxéologique de référence

Dans le but de concevoir un MPR, nous nous basons sur les résultats du chapitre 2 : l'étude historique et la revue de travaux. Nous complétons ces résultats par l'analyse du programme de Centrafrique ; ceci nous amène à inférer un certain nombre de types de tâches. Ensuite, nous présentons l'analyse du manuel de BTS où nous repérons d'autres types de tâches. Enfin, nous construisons notre MPR issu du programme de Centrafrique et du manuel de BTS.

5.1.1 *Premiers types de tâches*

Notre étude historique nous a permis d'observer l'importance du choix d'une méthode pour l'ajustement affine ainsi que la nécessité de justifier la pertinence de l'emploi d'une telle méthode. La revue de travaux nous a aussi montré l'intérêt d'un ajustement à la règle pour construire le sens de l'ajustement affine, et l'intérêt de comparer plusieurs droites d'ajustement. De plus, les travaux consacrés à des enseignements orientés vers l'investigation ont souligné l'intérêt que les élèves collectent eux-mêmes, les données ou qu'ils travaillent sur des données réelles, qui doivent être « nettoyées » avant d'être traitées.

Nous complétons ces constats par des éléments issus de programme de mathématiques de la classe de terminale SE de 2002 en Centrafrique. Il s'agissait en effet du programme en vigueur lors de notre recueil de données. Ce programme est présenté sous forme d'un tableau en plusieurs points : ensemble de nombres, arithmétique, calcul littéral, étude de fonctions numériques, suites numériques, intégration, organisation des données, angles, trigonométrie, calculs vectoriels dans le plan, géométrie analytique plane, applications du plan, coniques et la géométrie de l'espace. La partie consacrée à **l'organisation des données** est déclinée en deux sous parties : séries doubles et probabilité (analyse combinatoire, probabilités et variables aléatoires). La partie qui nous intéresse est celle consacrée aux séries statistiques doubles, elle concerne uniquement la Terminale SE et est très succincte. Les contenus sont regroupés en trois points de la manière suivante :

- Nuage de points, point moyen
- Ajustement linéaire (Mayer, moindres carrés)
- Droites de régression, covariance et coefficient de corrélation.

ORGANISATION DES DONNÉES	Terminales SM & E	Terminale SE
		<ul style="list-style-type: none"> • Séries doubles : <ul style="list-style-type: none"> – Nuage de points, point moyen. – Ajustement linéaire (Mayer, moindres carrés). – Droites de régression, covariance et coefficient de corrélation.
	<ul style="list-style-type: none"> • Analyse combinatoire <ul style="list-style-type: none"> – Consolidation des acquis de 1^{er}. – Formule du binôme et triangle de Pascal. • Probabilités <ul style="list-style-type: none"> – Vocabulaire – Propriétés – Équiprobabilité – Indépendance 	<ul style="list-style-type: none"> • Analyse combinatoire <ul style="list-style-type: none"> – Consolidation des acquis de 1^{er}. – Formule du binôme et triangle de Pascal. • Probabilités <ul style="list-style-type: none"> – Vocabulaire – Propriétés – Équiprobabilité – Probabilités conditionnelle,

	<ul style="list-style-type: none"> – Épreuve de Bernoulli • Variable aléatoire – Loi de probabilité – Fonction de répartition – Espérance mathématique, variance et écart type. – Loi binomiale. 	<ul style="list-style-type: none"> événements indépendants, probabilités totales • Variable aléatoire – Loi de probabilité – Fonction de répartition – Espérance mathématique, variance et écart type. – Épreuve de Bernoulli – Loi binomiale.
--	--	---

Tableau 2: Extrait du programme Centrafricain de Mathématiques (Programmes de Mathématiques des classes de terminale SM, E & SE, 2002, p.4)

En utilisant notre étude historique, notre revue de travaux et ce programme, nous retenons dans un premier temps dix types de tâches généraux qui nous permettent de construire un modèle praxéologique de référence. Voici ces dix premiers types de tâches concernant l'enseignement de l'ajustement affine :

- T1 : Faire la représentation graphique de la série double et de son point moyen ;
- T2 : Étudier si un ajustement affine est pertinent ;
- T3 : Choisir une méthode d'ajustement affine ;
- T4.1 : Réaliser un ajustement affine avec la méthode « à la règle » ;
- T4.2 : Réaliser un ajustement affine avec la méthode de Mayer ;
- T4.3 : Réaliser un ajustement affine avec la méthode des moindres carrés ;
- T5 : Donner une estimation ;
- T6 : Interpréter l'estimation ;
- T7 : Recueillir des données ;
- T8 : Nettoyer des données réelles.

Pour compléter ces premiers types de tâches, nous avons sélectionné un manuel français de BTS que nous allons analyser. En effet, traditionnellement les programmes de Centrafrique s'inspirent des programmes français avec un léger décalage temporel. En 2018, les séries statistiques doubles n'étaient pas enseignées en France en Terminale ; mais on trouvait dans le programme de BTS de Comptabilité et Gestion Informatique l'ensemble des contenus figurant dans le programme de Centrafrique. C'est donc ce manuel de BTS en France que nous allons utiliser pour construire notre MPR. Ceci nous permettra par la suite d'utiliser ce MPR pour analyser le manuel de Terminale SE

de Centrafrique. Nous notons que ceci induit un certain biais : en effet, en classe de BTS il y a peu de démonstrations mathématiques. Nous notons dans notre analyse du manuel de BTS les justifications qui sont présentes ou non ; il est a priori possible de trouver plus de démonstrations mathématiques dans le manuel de Centrafrique.

5.1.2 **Repérage des types de tâches dans le manuel de BTS**

Présentation et structuration du manuel de mathématiques BTS

Le manuel 'mathématiques Comptabilité et Gestion Informatique de gestion 2000' comporte 16 chapitres, dont huit d'analyse, trois de statistiques, trois de probabilité et deux d'algèbre. Notre travail porte sur le chapitre onze, intitulé séries statistiques à deux variables.

Le cours sur les séries statistiques à deux variables va de la page 240 à la page 248, et les exercices vont de la page 249 à la page 261.

Nous présentons dans ce qui suit, la structure détaillée du chapitre intitulée : *Séries statistiques à deux variables*. Le chapitre comprend quatre rubriques : 1.Cours, 2.Travaux pratiques, 3.Exercices résolus, 4.Exercices et problèmes.

Nous recourons, lorsque cela nous semble utile, à la notion de *registres de représentation sémiotique* (Duval, 1993) et à la notion de *praxéologie* (Chevallard, 1998).

Rubrique 1 : cours

La rubrique cours est répartie de la manière suivante :

- Série statistique double ;
- Nuage de points - Point moyen du nuage ;
- Ajustement affine ;
- Coefficient de corrélation linéaire ;
- Les lissages .

La notion de lissage est développée dans ce manuel. Cependant, elle n'est pas enseignée dans les classes de terminales SE en Centrafrique. Pour cela que nous ne faisons pas l'analyse de cette notion.

En ouverture du chapitre les auteurs annoncent les objectifs suivants :

- Consolidation et approfondissement des acquis des années antérieures ;
- Calcul d'un coefficient de corrélation ;
- Estimation à l'aide de modèles mathématiques.

Séries statistiques doubles

Registre de la langue naturelle

Pour définir la série statistique double, les auteurs utilisent le registre de la langue naturelle :

« On appelle série statistique double (X, Y) l'ensemble des couples (x_i, y_i) , chaque couple étant associé à un individu de la population. »

Dans cette définition, les termes mathématiques sont des prérequis pour les élèves (ensemble, couple) et sont utilisés, dans un registre dit naturel. La définition est illustrée par un exemple dans le contexte économique.

Registre des tableaux

Concernant la représentation des données, la série statistique double est représentée généralement par le registre des tableaux. Deux exemples sont donnés dans un contexte économique pour illustrer la représentation des données.

Exemple 1 (p.241)

Le tableau ci-dessous représente pour six voitures le prix de vente ainsi que la cote de revente au bout d'un an (les prix sont donnés en euro).

Prix du véhicule neuf : x_i	7680	8835	9435	11240	15900	16350
Cote de revente au bout d'un an : y_i	6240	7080	7080	8130	12180	12435

Exemple 2 (p.241)

La société CBT vient de commercialiser un nouvel article A. Le tableau suivant représente la série chronologique : x_i est le rang du trimestre et y_i le nombre, en centaines, d'articles A vendus.

Rang du trimestre : x_i	1	2	3	4	5	6
Nombre d'articles : y_i	9,8	12,7	18	20,7	28,2	32,5

Ces tableaux à deux lignes permettent d'étudier simultanément deux caractères d'une même population.

Nuage de points – Point moyen du nuage

La définition du nuage de points est mise en évidence dans un encadré et illustrée par deux exemples dans le contexte économique. Cette notion est directement liée à un type de tâches qui est : « Représenter le nuage de points associé à la série statistique ». Que nous appellerons par la suite **T_Rep_Nuage**. Ce type de tâches amène donc, ici à faire la conversion du registre de représentation du tableau des valeurs des variables quantitatives au registre de représentation graphique avec des points dans un repère judicieusement choisi.

Nous avons ensuite, un discours technologique sur le choix des unités du repère :

« On choisit les unités sur les axes de façon à ce que le dessin ait de bonnes proportions. Il convient parfois de choisir seulement une fenêtre ne contenant pas l'origine du repère. » (p.241)

Le point moyen du nuage est défini de la même manière que précédemment, c'est-à-dire mis en évidence dans un encadré et illustré par un exemple.

« On appelle point moyen d'un nuage le point G de coordonnées (\bar{x}, \bar{y}) , où \bar{x} (respectivement \bar{y}) est la moyenne de la variable X (respectivement Y) ».

L'exemple d'illustration est donné préalablement sous forme de tableau à double entrée, dans le contexte économique. Pour cette partie, le type de tâche est: « calculer les coordonnées du point moyen G » que nous nommons aussi par la suite **T_Cal_CoordG**. La technique relative à ce type de tâche, consiste à calculer les moyennes \bar{x} et \bar{y} représentant respectivement les séries simples X et Y . Ces moyennes \bar{x} et \bar{y} constituent les coordonnées du point moyen G . Et la technologie qui justifie cette technique est la définition de la moyenne.

Une fois que les coordonnées du point moyen G sont calculées, on peut le placer sur le graphique. Ce placement du point moyen G sur le graphique que nous nommons : **T_Pla_MoyG**, est illustré à l'aide d'un exemple. Et il n'y a pas de discours technologique sur le fait que le point moyen doit se trouver en gros au milieu du nuage. Ces trois premiers types de tâches constituent le type de tâches général que nous avons nommé T1 faire la représentation graphique de la série double et de son point moyen. Par ailleurs, pour passer de **T_Cal_CoordG** à **T_Pla_MoyG** on utilise le registre numérique et le registre graphique.

Ajustement affine

Cette partie commence par une illustration de trois différentes représentations graphiques.

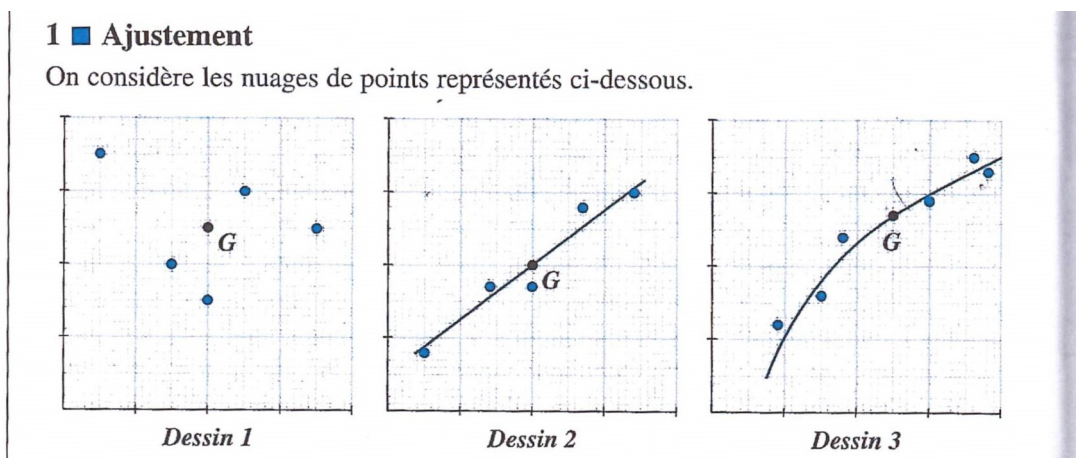


Figure 2: Illustration de différentes situations vis-à-vis de l'ajustement

Dans ce registre graphique, le type de tâches qui est : « Étudier si un ajustement affine est pertinent », n'est pas explicitement cité. Nous avons nommé ce type de tâches T2. Par contre, on représente les nuages de points, puis on cherche l'existence d'un lien éventuel entre les deux variables (la notion de corrélation sera abordée après la notion des ajustements). Graphiquement cela revient à chercher s'il existe une droite qui passe au plus près des points du nuage. On dit que l'on effectue un ajustement.

On remarque que, sur le premier graphique (dessin 1), le nuage de points est dispersé de façon quelconque. Par contre sur le deuxième graphique (dessin 2), le nuage de points présente une forme rectiligne et la courbe d'ajustement est une droite. Pour le troisième graphique (dessin 3), le nuage de point peut être approché par une courbe parabolique. Nous avons ensuite la définition de la droite d'ajustement, suivie d'une remarque.

« Une droite d'ajustement est une droite qui passe au plus près des points du nuage.

Il est généralement admis que la droite d'ajustement affine passe par le point moyen G du nuage »

Le fait de dire que la droite passe « au plus près » ou « approche au mieux » des points du nuage, demeure intuitif et n'est pas justifié mathématiquement. On parle de 'la droite d'ajustement affine' comme si il en avait une seule.

En remarque les auteurs mentionnent que le point moyen G appartient à la droite d'ajustement affine. C'est une forme de discours technologique mais qui n'est pas fondée sur une preuve mathématique. Les auteurs invoquent un consensus. Comme nous l'avons dit ci-dessus, dans le contexte du BTS il y a peu de démonstrations mathématiques : il s'agit de donner aux étudiants des techniques utiles pour leur future activité professionnelle.

Ajustement à la règle (au jugé)

Nous avons ici une praxéologie introduite, dont le type de tâches évoqué est : « Réaliser un ajustement affine avec la méthode à la règle ». Nous nommons ce type de tâches : **T4.1**. Nous signalons que ce type de tâches n'est pas demandé dans la rubrique des exercices dans le manuel du BTS de France, elle est seulement évoquée dans le cours. La technique proposée dans le manuel est de tracer une telle droite à la règle, en équilibrant les points situés au-dessus et les points situés en dessous de la droite. Il est également possible de tracer à l'œil à l'aide d'une règle une telle droite. Et les éléments technologiques sont dans ce cas de placer la règle graduée le plus « près » possible des points du nuage, et de tracer une droite. Cependant le sens du terme « près » n'est pas explicité ici, les élèves ne sont pas invités à réfléchir à différentes manières de réaliser cette proximité, à la distance qui est sous-entendue etc. A ce stade du cours, on ne sait pas non plus à quoi peut servir un ajustement affine. Pourquoi veut-on faire passer une droite « près » du nuage n'est pas expliqué. On place une droite ; l'équation de cette droite est de la forme $y=ax+b$. Pour déterminer cette équation, il suffit de connaître les coordonnées de deux points de la droite, on peut le faire par lecture graphique. Nous nommons le types de tâches qui permet de déterminer l'équation cartésienne : **T_Dét_EquatC**

Ajustement par la droite de Mayer

Le type de tâches **T4.2** fait référence au contenu sur l'ajustement par la droite de Mayer.

Le principe de la méthode est :

«on partage le nuage en deux sous-nuages puis on détermine les points moyens G_1 et G_2 des deux sous-nuages. »

Suivi de la définition de la droite de Mayer :

« La droite de Mayer est la droite (G_1G_2) ; elle constitue une droite d'ajustement dans le cas d'un nuage présentant une forme longiligne. »

La technique consiste d'abord à répartir la série en deux sous-séries et à calculer les coordonnées des points moyens $G_1(\bar{x}_1, \bar{x}_2)$ et $G_2(\bar{x}_1, \bar{x}_2)$, nous nommons le type de tâches **T_Cal_CoordG1G2**. Ensuite il s'agit de déterminer l'équation de la droite de Mayer $y=ax+b$, c'est ce que nous nommons le type de tâches **T_Dét_EquatG1G2**. L'élève aura recours au registre de représentation algébrique pour déterminer l'équation de la droite et au registre de représentation graphique pour représenter la droite (G_1G_2) dans le repère. Les éléments technologiques qui justifient la technique concernent le calcul de l'équation (moyenne arithmétique, le vecteur directeur d'une droite, résolution d'une équation de premier degré). En revanche il n'y a aucune justification du fait que la droite est bien une droite d'ajustement affine, au sens qu'elle passe « près » des points du nuage. Ainsi la praxéologie est incomplète.

Le principe de la méthode est mis en application par un exemple dans le contexte économique.

Ajustement par la méthode des moindres carrés : droites de régression

Cette rubrique fait référence au type de tâches **T4.3** : Réaliser un ajustement affine avec la méthode des moindres carrés. Pour réaliser **T4.3**, on se sert de **T_Dét_EquatMCyx**: Déterminer une équation de la droite par la méthode des moindres comme technique, et **T_Cal_Moyxy** : Calculer la moyenne \bar{x} de X et la moyenne \bar{y} de Y , **T_Cal_Cov** : Calculer la covariance et **T_Tra_DroiteD** : Tracer la droite D de régression de y en x , comme des ingrédients de technique.

Cette partie commence par une représentation graphique d'une droite D non parallèle aux axes d'équation $y=a \cdot x+b$ dans un repère. Où les P_i représentent les points de la droite D de même abscisse x_i que les points M_i . Donc $M_iP_i=|y_i-(ax_i+b)|$

Nous avons ensuite la définition de la droite de régression de y en x , et aussi celle de la droite de régression de x en y :

« La droite de régression de y en x est la droite telle que la somme $\sum_{i=1}^{i=n} M_iP_i^2$ soit minimale ».

Le fait de dire que la somme des carrés des écarts entre les points M_i et la droite soit minimale, correspond au contenu mathématiques que nous avons présenté dans la partie 1.2.1.3 du chapitre 1. Il n'y a pas de démonstration mathématique de ce résultat.

La droite de régression de x en y est définie de manière similaire. Il n'y a pas d'explication dans le livre concernant le fait qu'on pourra parfois faire une régression de y en x , et parfois de x en y .

Par ailleurs, nous avons aussi la définition de la covariance qui joue un rôle important dans la technique pour réaliser **T4.3** :

« on appelle covariance de la série statistique double le réel :

$$\text{cov}(x, y) = \sigma_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{\sum_{i=1}^n x_i y_i}{n} - \bar{x} \bar{y} . \gg$$

Et l'expression de la formule de covariance est donnée de deux manières différentes : nous avons la forme factorisée et la forme développée.

Propriété : Équation des droites de régression

Dans cette partie, nous avons les résultats qui permettent de déterminer l'équation de la droite de régression de y en x que nous avons nommé T_Dét_EquatMCyx :

- les droites de régression passent par le point moyen $G(\bar{x}, \bar{y})$ du nuage ;
- la droite de régression D de y en x a pour équation : $y = ax + b$ avec :

$$a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\sigma_{xy}}{V(x)} \text{ et } b = \bar{y} - a\bar{x} .$$

Et la démonstration de cette propriété n'est pas donnée.

Deux exemples sont donnés sur la détermination d'équation des droites de régression, et les exemples sont dans le contexte économique. Nous verrons les praxéologies utilisées dans les deux exemples avec les données de l'exemple 2, où le type de tâches est T4.3 : Réaliser un ajustement affine avec la méthode des moindres carrés.

La droite de régression D de y en x a pour équation : $y = ax + b$. Cette partie est illustrée directement par les données de l'exemple 2. La technique consiste à déterminer le coefficient

directeur a par la relation $a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\text{cov}(x, y)}{V(x)}$ où $\text{cov}(x, y) = \frac{\sum_{i=1}^n x_i y_i}{n} - \bar{x} \bar{y} .$, $V(X) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$ et

$b = \bar{y} - a\bar{x}$. qui est un ingrédient de technique de T4.3.

La technique pour la détermination du coefficient directeur a , est de procéder directement à l'application des formules, en utilisant les données de l'exercice.

$$\bar{x} = \frac{1}{6} \sum_{i=1}^6 x_i = \frac{1}{6} \cdot (1+2+3+4+5+6) = \frac{21}{6} = 3,5$$

$$\bar{y} = \frac{1}{6} \sum_{i=1}^6 y_i = \frac{1}{6} (9,8+12,7+18+20,7+28,2+32,5) = \frac{121,1}{6} = 20,1$$

$$\sigma_x^2 = V(X) = \frac{1}{6} \sum_{i=1}^6 x_i^2 - \bar{x}^2 = \frac{1}{6} (1^2+2^2+3^2+4^2+5^2+6^2) - (3,5)^2 = \frac{91}{6} - 12,25 = 2,92$$

puis, on en déduit $a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\text{cov}(x, y)}{V(x)} = \frac{13,02}{2,92} = 4,46 .$

Et b est déterminé par la relation : $b = \bar{y} - a\bar{x} = 20,1 - 4,46 \times 3,5 = 4,48$

D'où la droite d'équation : $y = 4,46x + 4,48$

Le deuxième exemple est de trouver la droite de régression D' de x en y et qui a pour équation : $x = a'y + b'$. Pour cela, nous procédons de la même manière que dans l'exemple précédent,

$$a' = \frac{\sigma_{xy}}{\sigma_y^2} = \frac{\text{cov}(x, y)}{V(y)} \text{ et } b' = \bar{x} - a' \bar{y}$$

La technique est de calculer d'abord la variance de y :

$$\sigma_y^2 = \frac{1}{6} \sum_{i=1}^6 y_i^2 - \bar{y}^2 = \frac{1}{6} ((9,8)^2 + (12,7)^2 + (18)^2 + (20,7)^2 + (28,2)^2 + (32,5)^2) - (20,1)^2 = 72,87,$$

$$a' = \frac{\sigma_{xy}}{\sigma_y^2} = \frac{\text{cov}(x, y)}{V(y)} = \frac{13,02}{72,87} = 0,22 \text{ et } b' = \bar{x} - a' \bar{y} = 3,5 - 0,22 \times 20,1 = -0,93$$

D'où, l'équation de la droite : $x = 0,22y - 0,93$

Il n'y a pas d'explication sur le fait de choisir une régression de y en x ou de x en y , comme nous l'avons dit ci-dessus. C'est en fait le contexte des exercices qui permet de choisir : si on veut déterminer une valeur de y , on fait une régression de y en x . Et inversement si on veut déterminer une valeur de x .

Coefficient de corrélation linéaire

Le paragraphe commence par la définition du coefficient de corrélation linéaire, suivi d'un exemple dans le contexte économique. Dans cet exemple, on étudie aussi le signe du coefficient de corrélation linéaire.

Propriété :

r , $\text{cov}(x, y)$, a et a' ont le même signe.
 $-1 \leq r \leq +1$

$$aa' = r^2$$

L'interprétation graphique de cette propriété est la suivante :

1^{er} cas : $|r| = 1$, les droites D et D' ont même coefficient directeur, on dit aussi que la corrélation est parfaite.

2^{èm} cas : $|r|$ est proche de 1, on dit que la corrélation est forte, cela veut dire que, le nuage de points est aplati autour de la droite de régression.

Cette notion permet aussi de justifier la pertinence de l'ajustement : elle fournit un élément de discours technologique, pour le type de tâches T2.

Après cette analyse du cours, nous allons poursuivre par l'analyse des exercices, afin de compléter notre liste de praxéologies d'une part ; et d'autre part, de chercher ce qui est confié à la responsabilité de l'élève dans ces exercices.

Notre objectif dans cette partie consacrée aux exercices, est de faire une analyse en termes de praxéologies : types de tâches, techniques, technologies et éventuellement théorie, pour chercher ce qui est confié à la responsabilité de l'élève.

Rubrique 2 : Travaux pratiques

La rubrique travaux pratiques comporte trois exercices.

- Le premier exercice est intitulé : Un exemple d'ajustement par la droite de Mayer.

Dans cet exercice, les types de tâches qui sont demandés sont :

- **T_Rep_Nuage** : représenter le nuage de points associé à la série statistique double ;
- **T_Cal_CoordG** : calculer les coordonnées du point moyen G du nuage ;
- **T_Pla_MoyG** : Placer le point moyen G sur le graphique ;
- **T_Cal_CoordG1G2** : Déterminer les coordonnées de G_1 et G_2 ;
- **T_Plat_MoyG1G2** : Placer G_1 et G_2 sur le graphique ;
- **T_Dét_EquatG1G2** : Déterminer une équation de la droite par la méthode de Mayer ;
- **T_Vér_PointG** : Vérifier que G appartient à la droite (G_1G_2) ;
- **T5** : donner une estimation.

Il y a dans cet exercice une certaine responsabilité à la charge de l'élève. En effet il a, à sa charge le choix de l'échelle pour représenter le nuage de points. Et il a la responsabilité de trouver une estimation en l'an 1996. Pour ce faire, il doit d'abord déterminer une équation de la droite (G_1G_2) qui permettra de trouver l'estimation demandée.

- Le deuxième exercice est intitulé : Les droites de régression. Il s'agit ici, de faire l'ajustement, en utilisant la méthode des moindres carrés. Pour cela, les types de tâches qui sont demandés sont de :

- **T_Rep_Nuage** : Représenter le nuage de points associé à la série statistique double ;
- **T_Pla_MoyG** : Placer le point moyen G du nuage ;
- **T_Cal_Coef** : Calculer le coefficient de corrélation entre x et y ;
- **T2** : Étudier si un ajustement affine est pertinent ;
- **T_Dét_EquatMCyx** : Déterminer une équation de la droite par la méthode des moindres carrés ;
- **T5** : Donner une estimation.

Ce qui est confié à l'initiative de l'élève est : le choix de l'échelle pour représenter le nuage de points, de donner une estimation, en utilisant la méthode des moindres carrés pour la détermination de l'équation de la droite.

- Le troisième exercice est intitulé : Ajustement avec changement de variable. Les types de tâches qui sont demandés dans cet exercice sont :

- **T_Rep_Nuage** : Représenter le nuage de points ;
- **T_Dét_EquatMCyx** : Déterminer une équation de droite de régression, obtenue par la méthode des moindres carrés ;
- **T6** : Donner une interprétation de l'estimation ;
- **T_Cal_Coef** : Calculer le coefficient de corrélation;
- **T2** : Étudier si un ajustement affine est pertinent;

La technique qui permet de déterminer l'équation de la droite est la méthode des moindres carrés.

Initiative de l'élève : Dans un premier temps, il a la responsabilité du choix de l'échelle et d'interpréter la tendance décrite par la droite D, obtenue par la méthode des moindres carrés. Dans un second temps, la responsabilité confiée à l'élève consiste à faire un changement de variable sur l'axe des abscisses. Et d'interpréter la nouvelle tendance, obtenue par la méthode des moindres carrés.

Rubrique 3 : Exercices résolus

Dans cette rubrique, il y a trois exercices :

Exercice 1 : une droite d'ajustement

On retrouve certains types de tâches qui sont les mêmes que dans les exercices précédents (travaux pratiques), mais les données numériques diffèrent, tels que :

- **T_Rep_Nuage** : représenter le nuage de points ;
- **T_Cal_CoordG** : calculer les coordonnées du point moyen G ;
- **T_Pla_MoyG** : Placer le point G sur le graphique ;
- **T_Cal_Coef** : calculer le coefficient de corrélation ;
- **T2** : Étudier si un ajustement affine est pertinent;
- **T_Dét_EquatMCyx** : déterminer une équation de la droite de régression par la méthode des moindres carrés ;
- **T_Tra_DroiteRégyx** : droite de régression de y en x;
- **T5** : déduire une estimation.

L'analyse praxéologique approfondie de cet exercice sera faite dans la partie 5.1.3.

Exercice 2 : un ajustement après lissages

Dans cette partie, on voit une variété de registres de représentation. Et les méthodes en jeu sont, l'ajustement à la règle (au jugé), la méthode de Mayer et la méthode de lissage qui n'est pas enseignée dans les classes de terminales SE en Centrafrique.

Exercice 3 : une série pondérée

Pour cet exercice, trois tâches sont demandées qui relèvent des types suivants :

- **T_Cal_CoordG** : déterminer les coordonnées du point moyen G ;
- **T_Cal_Cov** : calculer la covariance ;
- **T_Dét_EquatMCyx** : déterminer une équation de la droite de régression par la méthode des moindres carrés.

Concernant le type de tâches : déterminer les coordonnées du point moyen G du nuage formé par les points $M_i(x_i, y_i)$. La technique relative pour ce type de tâches est de calculer \bar{x} et \bar{y} à l'aide des formules suivantes :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n n_i x_i \text{ et } \bar{y} = \frac{1}{n} \sum_{i=1}^n n_i y_i \text{ où } n \text{ est l'effectif total.}$$

La technologie qui justifie cette technique est la définition de la moyenne pondérée.

Pour le type de tâches : calculer la covariance de la série double (X, Y). La technique consiste à utiliser la formule : $cov(x, y) = \frac{1}{n} \sum_{i=1}^n n_i x_i y_i - \bar{x} \bar{y}$

La technologie qui justifie cette technique est la moyenne pondérée et la définition de la covariance.

Concernant le type de tâches : déterminer une équation de la droite de régression de y en x, par la méthode des moindres carrés. La technique consiste à trouver la variance de x, puis d'utiliser la relation : $y - \bar{y} = a(x - \bar{x})$ (1), avec $a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\sigma_{xy}}{V(x)}$ pour trouver l'équation de la droite.

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n n_i x_i^2 - \bar{x}^2, \quad a = \frac{\sigma_{xy}}{\sigma_x^2}, \text{ alors remplaçant } \bar{x}, \bar{y} \text{ et } a \text{ par leurs valeurs dans (1), on aura}$$

l'équation de la droite.

La technologie relative à la technique est la définition de la moyenne pondérée, la variance, la covariance et la résolution d'une équation à une inconnue.

Rubrique 4 : Exercices et problèmes

Cette rubrique est divisée en trois parties (A, B et C) et comprend vingt exercices en tout.

Nous analysons ici dix-huit des exercices proposés.

Partie A : Pour apprendre le cours (exercices avec solutions)

La partie A est répartie en séries d'exercices :

Première série d'exercices : la droite de Mayer

Cette partie comprend trois exercices (1, 2 et 3). Les tâches demandées sont liées. Pour ces types de tâches, on demande aux élèves de :

- **T_Rep_Nuage** : représenter le nuage de points ;
- **T_Cal_CoordG, T_Cal_CoordG1G2** : déterminer les coordonnées des points moyens G, G₁, G₂ ;

- **T6** : Interpréter le résultat.

Différents registres de représentation sont utilisés dans ces exercices, tels que : le registre de représentation graphique, le registre algébrique et le registre des tableaux.

Initiative de l'élève dans ces exercices : il a la responsabilité d'interpréter les résultats. Et c'est dans les interprétations qu'on voit les initiatives des élèves.

Deuxième série d'exercices : les droites de régression

Cette série contient deux exercices (4 et 5), dont les questions sont liées. Où on demande de :

- **T_Rep_Nuage** : représenter le nuage de points ;
- **T_Dét_EquatMCyx** : déterminer une équation de la droite par la méthode des moindres carrés ;
- **T5** : faire une estimation.

Pour réaliser les types tâches qui sont demandés, les élèves doivent manipuler différents registres de représentations.

Initiative de l'élève dans ces exercices : l'élève a la responsabilité de chercher une estimation.

Troisième série d'exercices : avec changement de variable

Cette partie contient un seul exercice (exercice 6). Cet exercice relève des types de tâches :

- **T_Rep_Nuage** : représenter le nuage de points ;
- **T_Cal_Coef** : calculer le coefficient de corrélation ;
- **T4.3** : Réaliser un ajustement affine avec la méthode des moindres carrés ;
- **T5** : déterminer une estimation.

Initiative de l'élève dans cet exercice : l'élève a la responsabilité du choix de l'échelle de représentation graphique du nuage de points, et la recherche de l'estimation. Et la prise d'initiatives de l'élève est moyenne dans certains types de tâches comme calculer le coefficient de corrélation.

Quatrième série d'exercices : un lissage

Nous ne traitons pas dans cette thèse les exercices de lissage.

Partie B : Solutions des exercices de la partie A

Cette partie est consacrée au corrigé des sept exercices de la partie A.

Partie C : Pour approfondir le cours

La partie C est organisée de la même manière que la partie A. Par ailleurs, le nombre d'exercices est plus dense que dans la partie A.

Première série d'exercices : la droite de Mayer

Il y a trois exercices (8, 9 et 10) sur la droite de Mayer. Les deux premières questions sont presque similaires dans les trois exercices. Pour ces deux questions, les types de tâches qui sont demandés à l'élève sont de :

- **T_Rep_Nuage** : représenter le nuage de points dans un repère ;
- **T_Cal_CoordG, T_Cal_CoordG1G2** : déterminer les coordonnées des points moyens G, G_1, G_2 ;
- **T_Pla_MoyG** : Placer le point G sur le graphique ;
- **T_Pla_MoyG1G2** : Placer les points G_1, G_2 sur le graphique ;
- **T_Dét_EquatG1G2**: déterminer une équation de la droite par la méthode de Mayer.
- **T5** : donner une estimation

Les différentes questions concernent l'estimation (exercices 9 et 10) ou l'interprétation (exercice 8).

Concernant l'exercice 8 il y a aussi le type de tâches :

- **T_Uti_GrapAlg** : Utiliser une résolution graphique puis algébrique, à partir de la droite (G_1G_2) pour faire une estimation ;

Initiatives des élèves dans ces exercices : l'élève a la responsabilité de donner une estimation.

Deuxième série d'exercices : droites de régression

Il y a aussi trois exercices (11, 12 et 13) sur la méthode des moindres carrés. Les types de tâches correspondants sont :

- **T_Dét_EquatMCyx** : déterminer une équation de la droite par la méthode des moindres carrés, puis de
- **T5** : faire l'estimation (exercices 12 et 13) ou ;
- **T6** : interpréter l'estimation (exercice 11).

Troisième série d'exercices : avec changement de variable

Dans cette série, il y a cinq exercices (14, 15, 16, 17 et 18) sur le changement de variable.

Pour les exercices 14, 15, 17 et 18 les données ne conviennent pas, pour faire l'ajustement affine. Pour cela, l'utilisation de la fonction logarithme (\ln) permet de faire un changement de variable pour modifier les données. Et ensuite de déterminer une équation de la droite par la méthode des moindres carrés.

Pour l'exercice 16, le changement de variable s'effectue par le biais de la fonction inverse (on pose $f = \frac{1}{x}$).

Quatrième série d'exercices : Cette série concernant les lissages, nous ne la prenons pas en compte dans le cadre de cette thèse.

Nous récapitulons ci-dessous l'ensemble des types des tâches inférés du programme ou rencontrés dans le manuel.

N°	Types de tâches
01	T_Rep_Nuage : Représenter le nuage de points d'une série donnée, dans un repère ;
02	T_Cal_CoordG : Calculer les coordonnées du point moyen G;
03	T_Pla_MoyG : Placer le point moyen G dans le nuage de points;
04	T_Cal_Coef : Calculer le coefficient de corrélation ;
05	T_Cal_CoordG1G2 : Calculer les coordonnées des deux points moyens G_1 et G_2 associés aux deux sous-nuages ;
06	T_Uti_GrapAlg : Utiliser une résolution graphique puis algébrique, à partir de la droite (G_1G_2) pour faire une estimation ;
07	T_Pla_MoyG1G2 : Placer G_1 et G_2 sur le graphique ;
08	T_Vér_PointG : Vérifier que G appartient à la droite (G_1G_2) ;
09	T_Cal_Var : Calculer la variance des séries statistiques associées aux variables X et Y ;
10	T_Dét_EquatC : Déterminer l'équation cartésienne ;
11	T_Dét_EquatMCyx : Déterminer une équation de la droite de régression D de y en x , obtenue par la méthode des moindres carrés ;
12	T_Dét_EquatMCxy : Déterminer une équation de la droite de régression D' de x en y , obtenue par la méthode des moindres carrés ;
13	T_Cal_Cov : Calculer la covariance de la série statistique double (X, Y) ;
14	T_Tra_DroiteRégxy : Tracer la droite D de régression de y en x ;
15	T_Ecr_Modx : Écrire l'ensemble des modalités du caractère X ;
16	T_Ecr_Mody : Écrire l'ensemble des modalités du caractère Y ;
17	T_Etab_Tabxy : Établir le tableau linéaire de la série double de caractère (x, y) ;
18	T_Cal_Moyxy : Calculer la moyenne \bar{x} de X et la moyenne \bar{y} de Y ;
19	T_Mont_Liaxy : Montrer l'existence d'une liaison entre X et Y ;

20	T1 : Faire la représentation graphique de la série double et de son point moyen ;
21	T2 : Étudier si un ajustement affine est pertinent ;
22	T3 : Choisir une méthode d'ajustement affine ;
23	T4.1 : Réaliser un ajustement affine avec la méthode « à la règle » ;
24	T4.2 : Réaliser un ajustement affine avec la méthode de Mayer ;
25	T4.3 : Réaliser un ajustement affine avec la méthode des moindres carrés ;
26	T5 : Donner une estimation ;
27	T6 : Interpréter l'estimation.

Tableau 3: *LEs types de tâches inférés du programme ou rencontrés dans le manuel*

Pour construire notre modèle praxéologique de référence, nous examinerons les relations entre les types de tâches cités ci-dessus. Le programme donne une certaine organisation du savoir, il y a plusieurs thèmes avec une succession.

Concernant les types de tâches généraux nous détaillons les types de tâches qui deviennent des techniques, des ingrédients de la technique nécessaires à la réalisation des types de tâches généraux.

Nous présentons si possible les technologies, les ingrédients de technologie, tout en gardant à l'esprit que, certains types de tâches n'ont ni technologie, ni ingrédient de technologie.

T1 : *faire la représentation graphique de la série double et de son point moyen.*

En appui sur les programmes et les manuels, nous avons dans la technique associée les types de tâches:

- **T_Rep_Nuage**: Représenter le nuage de points d'une série donnée, dans un repère ;
- **T_Pla_MoyG** : Placer le point moyen G dans le nuage de points. Ingrédient de technique :
T_Cal_CoordG : Calculer les coordonnées du point moyen G ;

Pour accomplir T1, la technique consiste à réaliser d'abord **T_Rep_Nuage**, ensuite pour la mise en œuvre de la technique **T_Pla_MoyG** on se sert de **T_Cal_CoordG** comme ingrédient de technique relatif à la réalisation de **T_Pla_MoyG**.

T2 : *Étudier si un ajustement affine est pertinent.*

Pour ce type de tâche, il y a deux techniques. Il y a la technique par raisonnement graphique pour voir la forme du nuage, justifiée par une technologie. Cette technologie consiste à dire que lorsque

le nuage de points présente une forme allongée, alors on peut proposer un ajustement affine. En revanche si les points sont dispersés on ne le peut pas. Ici T1 devient un ingrédient de technique.

Et la deuxième technique ou il faut calculer le coefficient de corrélation linéaire. Cette deuxième technique est justifiée par une technologie qui sont la définition et la propriété du coefficient de corrélation linéaire. **T_Cal_Coef** : Calculer le coefficient de corrélation entre x et y .

Puis une fois qu'on a dit que c'était justifié il y a le type de tâche suivant :

T3 : *Choisir une méthode d'ajustement affine.*

Dans le manuel de BTS de France, ce type de tâches n'est pas demandé. Souvent on précise la méthode à utiliser.

T4.1 : *Réaliser un ajustement affine avec la méthode « à la règle ».*

La technique pour réaliser le type de tâches T4.1, est explicitée ci-dessus au paragraphe sur « l'ajustement affine à la règle ».

T4.2 : *Réaliser un ajustement affine avec la méthode de Mayer.*

La technique associée est justement la méthode de Mayer, et tout ce qui est lié à G_1 et G_2 dans la liste citée ci-dessus, c'est à la fois des types de tâches et des ingrédients de cette technique. Pour réaliser ce type de tâche, les types de tâches suivants :

- **T_Cal_CoordG1G2** : Calculer les coordonnées des deux points moyens G_1 et G_2 associés aux deux sous-nuages ;
- **T_Pla_MoyG1G2** : Placer les points moyens G_1 et G_2 sur le graphique;
- **T_Tra_DroiteG1G2** : Tracer la droite (G_1G_2) ;
- **T_Dét_EquatG1G2** : Déterminer une équation de la droite (G_1G_2) .

Remarque : dans certains exercices on trouve aussi le type de tâches

- **T_Vér_PointG** : Vérifier que G est un point de cette droite.

Ces types de tâches deviennent des ingrédients de techniques pour ce type de tâches T4.2. Les technologies qui légitiment les types de tâches et techniques mises en œuvre sont le principe de la méthode de Mayer et la définition du point moyen.

T4.3 : *Réaliser un ajustement affine avec la méthode des moindres carrés .*

La technique qui permet de calculer l'équation de la droite, est le type de tâches **T_Dét_EquatMCyx** : Déterminer une équation de la droite de régression D de y en x (ou **T_Dét_EquatMCxy** : Déterminer une équation de la droite de régression D' de x en y) , obtenue par la méthode des moindres carrés ; avec les ingrédients de techniques :

- **T_Cal_Moyxy** : Calculer la moyenne \bar{x} de X et la moyenne \bar{y} de Y ;
- **T_Cal_Var** : Calculer la variance des séries statistiques associées aux variables X et Y ;
- **T_Cal_Cov** : Calculer la covariance de la série statistique double (X, Y) ;

- **T_Tra_DroiteRégxy** : Tracer la droite D de régression de y en x .

Les technologies qui permettent de produire l'équation de la droite de régression sont la définition de la covariance et la propriété des moindres carrés. Elles permettent de déterminer les coefficients a et b .

Concernant **T5** : *donner une estimation*, c'est un type de tâches général. Il y a la première technique associée, qui est le type de tâches T4.1. La deuxième technique associée est le type de tâches T4.2 avec les ingrédients de techniques : **T_Cal_CoordG1G2**, **T_Pla_MoyG1G2**, **T_Tra_DroiteG1G2**, **T_Dét_EquatG1G2**, **T_Uti_GrapAlg** et la troisième technique est le type de tâches T4.3 avec ses ingrédients de techniques : **T_Cal_Cov**, **T_Dét_EquatMCyx** (ou **T_Dét_EquatMCxy**). Il peut y avoir T2 comme ingrédient de technologie. Ici, T2 a un statut particulier, il joue le rôle de justification, c'est-à-dire le rôle de T2 c'est de justifier qu'on peut faire le T4.1 ou T4.2 ou le T4.3.

Pour **T6** : *interpréter l'estimation obtenue par le modèle*. Il n'y a pas de technique générale associée. En effet l'interprétation dépend du contexte. Parfois il s'agit simplement de traduire dans le registre de la langue naturelle le résultat numérique trouvé, par exemple « on peut estimer que leurs dépenses en 2019 seront environ de 10000 euros ». Mais parfois il est aussi attendu une plus grande prise de recul, sur la qualité de l'estimation, ou faisant appel à des connaissances culturelles.

5.1.3 Construction du modèle praxéologique de référence à partir de l'étude historique, de la revue des travaux, du programme Centrafricain et du manuel BTS

Dans ce qui suit, nous présentons dans une première partie notre modèle praxéologique de référence (MPR). Dans la seconde partie, nous faisons l'analyse praxéologique approfondie de l'exercice 1 du manuel de BTS (Exercice résolu, p.251, figure 4), en se référant à notre MPR, pour pouvoir analyser les responsabilités qui sont laissées aux élèves, dans le cas où cet exercice leur serait donné.

Pour ce qui est de la construction du MPR, nous sommes parti de la liste de types de tâches ci-dessus et des relations que nous avons présentées plus haut. Nous mettons en pointillé les flèches quand le type de tâches n'est pas obligatoire.

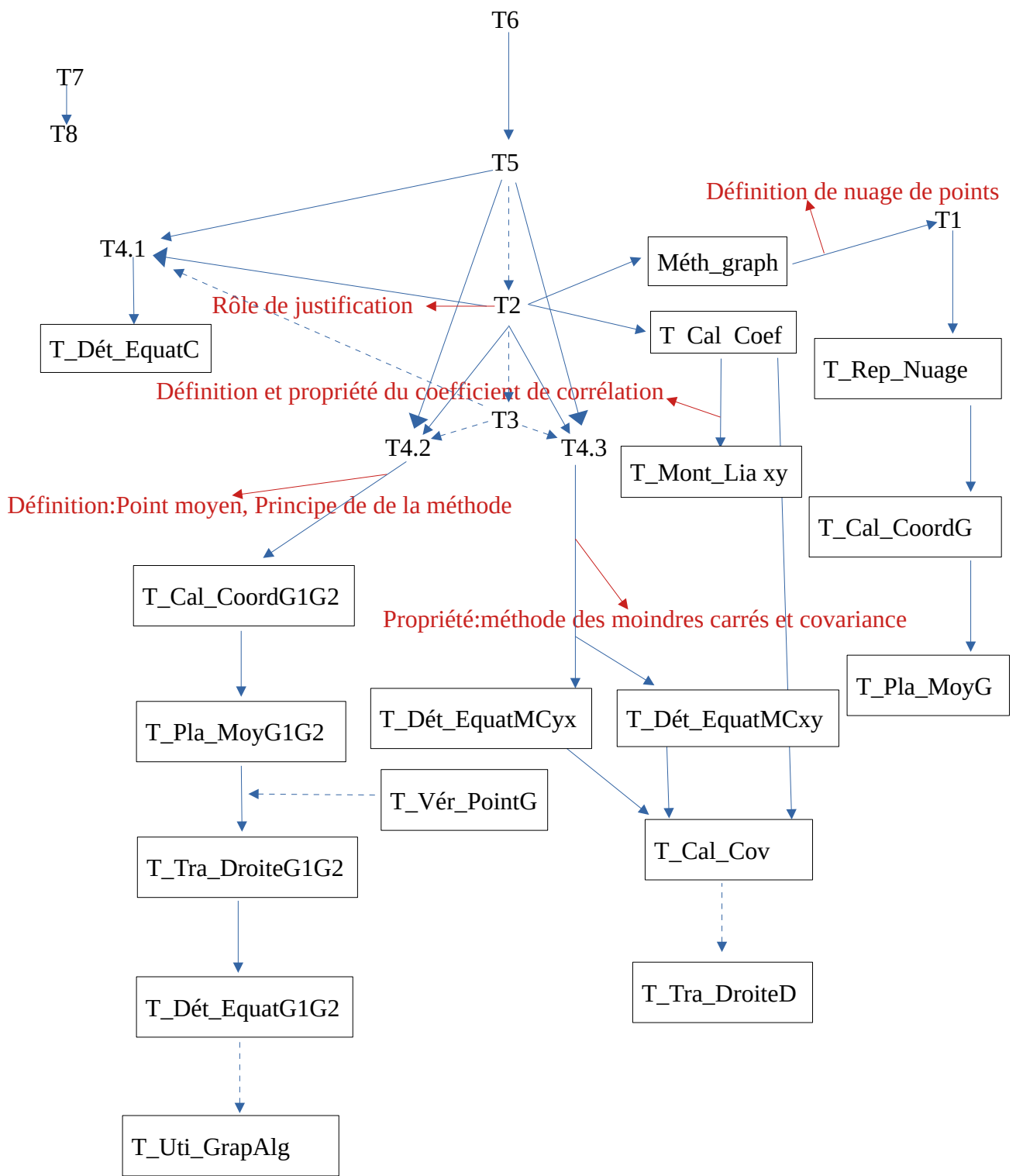


Figure 3: Modèle praxéologique de référence : Série statistique double - Etude historique - Revue des travaux - Programme de 2002 et manuel BTS

Notons que nous avons vu dans notre analyse épistémologique et dans notre revue de travaux qu'il aurait pu exister d'autres types de tâches, que nous n'avons pas trouvés dans le programme de Centrafrique ni dans le manuel de BTS. Nous pouvons ici citer T3 : choisir la méthode d'ajustement affine, T7 : recueillir des données et T8 : nettoyer des données réelles. Les deux derniers types de tâches (T7 et T8) peuvent-être réalisés indépendamment des huit premiers types de tâches généraux. Mais dans la plupart des cas, dans les énoncés des exercices des séries statistiques doubles dans le manuel de BTS, on propose toujours les données sous forme d'un tableau statistique à double entrée. Il n'y a pas de travaux pratiques où on proposerait aux étudiants de recueillir des données par exemple.

Analyse d'un exercice résolu du manuel BTS : MPR et initiatives des élèves

Nous proposons ici l'analyse d'un exercice corrigé qui se trouve à la page 251 du manuel BTS (figure 4). La notion abordée dans cet exercice est la notion d'ajustement. Pour cette analyse, nous utilisons les types de tâches répertoriés plus haut. Ceci nous permet de représenter cet exercice, et les lieux où l'initiative des élèves peut intervenir, en utilisant notre MPR.

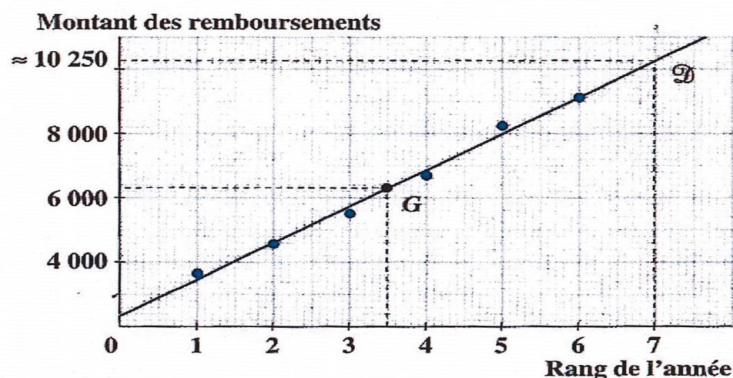
1 Une droite d'ajustement

Le tableau suivant indique en euros les montants des remboursements annuels y_i effectués de 1994 à 1999 par un ménage ayant contracté plusieurs emprunts.

Année	1994	1995	1996	1997	1998	1999
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3 568	4 561	5 502	6 693	8 241	9 120

- 1 Représenter le nuage de points associé à la série statistique double.
- 2 Calculer les coordonnées du point moyen G et le placer sur le dessin précédent.
- 3 Calculer le coefficient de corrélation.
Un ajustement affine est-il justifié ?
- 4 a) Déterminer une équation de la droite de régression \mathcal{D} de y en x , obtenue par la méthode des moindres carrés.
b) Tracer \mathcal{D} sur le dessin précédent.
c) En déduire une estimation des remboursements en l'an 2000.

■ 1



■ 2 G a pour coordonnées (\bar{x}, \bar{y}) donc $G(3,5 ; 6 295,83)$

■ 3 $r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$; $\sigma_{xy} = \frac{\sum_{i=1}^6 x_i y_i}{n} - \bar{x} \bar{y} = \frac{151 983}{6} - 3,5 \times 6 295,83 = 3 295,08$;

$\sigma_x = 1,708$ et $\sigma_y = 1 937 015$; donc $r \approx 0,996$.

r est proche de 1, on peut envisager un ajustement affine.

■ 4 a) \mathcal{D} a pour équation : $y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2} (x - \bar{x})$

On obtient : $y - 6 295,83 = \frac{3 295,08}{1,708^2} (x - 3,5) = 1 129,5(x - 3,5)$.

Une équation de \mathcal{D} est donc : $y = 1 129,5x + 2 342,6$.

c) L'an 2000 correspond à $x = 7$ d'où $y = 10 249$.

On peut estimer le montant des remboursements en l'an 2000 à 10 249 euros.

Figure 4: Un exercice corrigé extrait du manuel de BTS

Types de tâches	Techniques	Technologie
<ul style="list-style-type: none"> T_Rep_Nuage 	<p>Le rang de l'année x_i constitue les valeurs du premier caractère et le montant des remboursements y_i constitue les valeurs du second caractère, alors on place ces couples de points de coordonnées (x_i, y_i) sur le repère. Ces couples de points constituent le nuage de points.</p>	<p>Connaissances antérieures sur le repère, et le choix de l'échelle.</p>
<ul style="list-style-type: none"> T_Cal_CoordG; T_Pla_MoyG 	<p>- Pour calculer les coordonnées du point moyen G, on procède de la manière suivante : en calculant la moyenne de la série statistique x_i, qui représente le rang de l'année :</p> $\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^6 x_i = \frac{1+2+3+4+5+6}{6} = 3,5$ <p>et la moyenne de la série statistique y_i, représentant le montant des remboursements :</p> $\bar{y} = \frac{1}{6} \cdot \sum_{i=1}^6 y_i$ $\bar{y} = \frac{3568+4561+5502+6693+8241+9120}{6}$ $\bar{y} = 6295,83$ <p>\bar{x} et \bar{y} représentent les coordonnées de point moyen G.</p> <p>Pour placer le point G dans le repère, nous repérons \bar{x} sur l'axe des abscisses et \bar{y} sur l'axe des ordonnées. Puis nous plaçons le couple (\bar{x}, \bar{y}) sur le graphique qui correspond aux coordonnées de G</p>	<ul style="list-style-type: none"> Définition de la moyenne d'une série statistique ; Définition du point moyen.
<ul style="list-style-type: none"> T_Cal_Coef ; T2 	<p>Le coefficient de corrélation est égal à la covariance des variables x et y divisé par le produit de l'écart type de x et de y c'est-à-dire :</p> $r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$ <p>où la covariance est donnée par la relation suivante :</p> $\sigma_{xy} = \frac{1}{6} \cdot \sum_{i=1}^6 x_i \cdot y_i - \bar{x} \cdot \bar{y}$ $\sigma_{xy} = \frac{1 \times 3568 + 2 \times 4561 + 3 \times 5502 + 4 \times 6693 + 5 \times 8241 + 6 \times 9120}{6} - 3,5 \times 6295,83 = 3295,08$ <p>Et les écarts types de x et de y :</p> $\sigma_x = \sqrt{V(x)} = \sqrt{\frac{1}{6} \cdot \sum_{i=1}^6 x_i^2 - \bar{x}^2} = \sqrt{\frac{91}{6} - 12,25}$ $\sigma_x = \sqrt{2,91} = 1,708 ;$	<p>Définition de la : variance, écart type, covariance et coefficient de corrélation.</p>

	$\sigma_y = \sqrt{V(y)} = \sqrt{\frac{1}{6} \cdot \sum_{i=1}^6 y_i^2 - \bar{y}^2}$ $\sigma_y = \sqrt{\frac{83174400}{6} - 39637475,38} = 1937015$ <p>puis on déduit le coefficient de corrélation :</p> $r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = \frac{3295,08}{1,708 \times 1937015} = 0,996$	
<ul style="list-style-type: none"> • T_Dét_EquatM • C_{yx} ; • T_Tra_DroiteRé • gyx; • T5. 	<p>L'équation de la droite (D) est de la forme :</p> $y = ax + b = \frac{\sigma_{xy}}{\sigma_x^2} (x - \bar{x}) + \bar{y} \text{ où } a = \frac{\sigma_{xy}}{\sigma_x^2} \text{ et}$ $b = \bar{y} - a\bar{x}$ $y = \frac{3295,09}{(1,708)^2} (x - 3,5) + 6295,83$ <p>d'où l'équation de la droite (D) est</p> $y = 1129,5x + 2342,6 \text{ (1)}$ <p>- Pour représenter la droite (D), faisant le tableau de signe, avec deux points on peut tracer la droite (D) ;</p> <p>- D'après le tableau statistique, l'an 2000 correspond à $x=7$, puis remplaçant la valeur de x dans l'équation (1) :</p> $y = 1129,5 \times 7 + 2342,6 = 10249$	<ul style="list-style-type: none"> - Formules donnant l'équation des droites de régression ; - Définition de la : variance, covariance, la moyenne ; - Propriété géométrique de deux points de la droite ; - Résolution d'une équation de premier degré à une inconnue - l'ordonnée du point d'abscisse $x=7$ de la droite de régression nous donne une estimation des remboursements pour l'année future concernée.

Tableau 4: analyse praxéologique d'un exercice résolu

L'exercice comporte six questions avec différents types de tâches. Ici nous essayons d'identifier où peuvent se situer les initiatives de l'élève.

Dans la question une, le type de tâches est: représenter le nuage de points associé à la série statistique double. En se référant à notre MPR, c'est le type de tâches nommé **T_Rep_Nuage**. Dans cet exercice, la technique relative à ce type de tâches consiste à choisir d'abord l'échelle, ensuite, de placer les valeurs des deux caractères dans le repère. Pour ce faire, le rang de l'année x_i constitue les valeurs du premier caractère placées sur l'axe des abscisses et le montant des remboursements y_i constitue les valeurs du second caractère qui sont placées sur l'axe des ordonnées. Et à chaque couple (x_i, y_i) on associe un point. Et l'ensemble de ces points forme le nuage de points associé à la série statistique double. La technologie qui permet de justifier la technique est la connaissance de la notion de repère orthonormé.

Pour le type de tâches : calculer les coordonnées du point Mayer G. D'après notre MPR, c'est **T_Cal_CoordG**. La technique qui permet d'accomplir ce type de tâches, consiste à calculer d'abord la moyenne de la série statistique du premier caractère et ensuite celle de la deuxième série

statistique du second caractère, c'est-à-dire on fait le quotient de la somme des valeurs des x_i par 6, où 6 représente l'effectif. La technique mise en œuvre est le type de tâches **T_Cal_Moyxy** : Calculer la moyenne \bar{x} de X et la moyenne \bar{y} de Y .

$$\bar{x} = \frac{1}{n} \sum_{i=1}^6 x_i = \frac{1+2+3+4+5+6}{6} = 3,5$$

On procède de la même manière pour déterminer la moyenne de la deuxième série statistique,

$$\bar{y} = \frac{1}{6} \cdot \sum_{i=1}^6 y_i = \frac{3568+4561+5502+6693+8241+9120}{6} = 6295,83$$

Les deux moyennes \bar{x} et \bar{y} constituent les coordonnées du point moyen G. La technologie qui justifie cette technique est la définition de la moyenne arithmétique et du point moyen.

Pour le type de tâches : placer G sur le dessin. Il correspond au type de tâches nommé **T_Pla_MoyG**. La technique consiste à repérer la valeur de la moyenne $\bar{x}=3,5$ sur l'axe des abscisses et la moyenne $\bar{y}=6295,83$ sur l'axe des ordonnées. La projection de ces deux points dans le repère constitue le point moyen G. La technologie qui éclaire cette technique est la connaissance sur le repère.

Pour ces deux types de tâches, l'élève fait appel à ses connaissances antérieures pour calculer les coordonnées du point moyen G et le placer dans le repère. Par ailleurs ces trois types de tâches correspondent au type de tâches général nommé T1.

D'après notre MPR, le troisième type de tâches de cet exercice : calculer le coefficient de corrélation (**T_Cal_Coef**) est l'une des techniques pour réaliser T2 (Cf 5.1.2). La technique relative à **T_Cal_Coef** est de calculer le rapport de la covariance de x et y avec le produit des écarts types

$$\text{de } x \text{ et de } y : r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = \frac{\frac{1}{6} \sum_{i=1}^6 x_i \cdot y_i - \bar{x} \cdot \bar{y}}{\sqrt{\frac{1}{6} \cdot \sum_{i=1}^6 x_i^2 - \bar{x}^2} \cdot \sqrt{\frac{1}{6} \cdot \sum_{i=1}^6 y_i^2 - \bar{y}^2}}$$

Pour ce faire, calculons d'abord la covariance

$$\begin{aligned} \sigma_{xy} &= \frac{1}{6} \cdot \sum_{i=1}^6 x_i \cdot y_i - \bar{x} \cdot \bar{y} \\ &= \frac{1 \times 3568 + 2 \times 4561 + 3 \times 5502 + 4 \times 6693 + 5 \times 8241 + 6 \times 9120}{6} - 3,5 \times 6295,83 = 3295,08 \end{aligned}$$

Ensuite les écart-types :

$$\sigma_x = \sqrt{V(x)} = \sqrt{\frac{1}{6} \cdot \sum_{i=1}^6 x_i^2 - \bar{x}^2} = \sqrt{\frac{91}{6} - 12,25} = 1,708$$

$$\sigma_y = \sqrt{V(y)} = \sqrt{\frac{1}{6} \cdot \sum_{i=1}^6 y_i^2 - \bar{y}^2} = \sqrt{\frac{83174400}{6} - 39637475,38} = 1937015$$

Enfin, déduisons le coefficient de corrélation :

$$r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = \frac{3295,08}{1,708 \times 1937015} = 0,996$$

Et la technologie qui justifie cette technique est la définition de la : moyenne, variance, l'écart type, covariance.

Concernant l'application numérique, l'utilisation de la calculatrice est indispensable.

Pour ce type de tâches, il s'agit d'appliquer plusieurs concepts statistiques (la moyenne, la variance, l'écart type et la covariance) et d'avoir une bonne maîtrise de la manipulation des calculatrices scientifiques.

Pour le type de tâches : justifier si l'ajustement affine est approprié, il correspond à notre T2. La technique consiste à observer la valeur du coefficient de corrélation linéaire. Comme r est proche de 1, alors il existe une forte corrélation entre les deux variables. Donc la corrélation est jugée suffisante pour effectuer un ajustement affine. La technologie qui éclaire cette technique est la propriété du coefficient de corrélation et la notion du nuage de points.

Pour le type de tâche T2, dans la plus part des exercices où les données des exercices ne nécessitent pas un changement de variable, le contrat didactique prévoit qu'on répond toujours par oui.

La dernière question comporte trois types de tâches :

Le premier type de tâches est de déterminer une équation de la droite de régression (D) de y en x , par la méthode des moindres carrés. C'est le type de tâches **T_Dét_EquatMCyx** notre MPR,

Les techniques relatives à ce type de tâches sont :

- **T_Cal_Moyxy** : Calculer la moyenne \bar{x} de X et la moyenne \bar{y} de Y ;
- **T_Cal_Var** : Calculer la variance des séries statistiques associées aux variables X et Y ;
- **T_Cal_Cov** : Calculer la covariance de la série statistique double (X, Y) .

C'est-à-dire, d'utiliser l'équation des droites de régression de y en x :

$$y = ax + b \quad (1) \quad \text{où } a = \frac{\sigma_{xy}}{\sigma_x^2} \text{ et } b = \bar{y} - a\bar{x}, \text{ on remplace } a \text{ et } b \text{ dans (1) :}$$

$$y - \bar{y} = \frac{\sigma_{xy}}{\sigma_x^2} (x - \bar{x})$$

Les valeurs suivantes ont été déjà calculées dans les questions précédentes : $\sigma_{xy} = 3295,08$,

$$\sigma_x = 1,708, \quad \bar{x} = 3,5, \quad \bar{y} = 6295,83$$

$$\text{on a : } y = \frac{3295,08}{(1,708)^2} (x - 3,5) + 6295,83$$

Alors l'équation de la droite de régression (D) de y en x est : $y = 1129,5x + 2342,6$

La technologie qui justifie cette technique est la définition de la : moyenne, variance, et la covariance. L'utilisation de la calculatrice scientifique est important dans la réalisation de ce type de

tâches. Pour calculer par exemple le quotient de la covariance de x et y par la variance de x :

$$a = \frac{3295,09}{(1,708)^2} = 1129,5, \text{ les élèves ont besoin d'utiliser la calculatrice pour ce genre de calcul.}$$

Le deuxième type de tâches est de tracer la droite (D). C'est un ingrédient de technique (optionnel) de T4.3, nommé **T_Tra_DroiteRégxy**. Pour ce faire, la technique consiste à faire un tableau de valeur en donnant par exemple deux valeurs différentes à la variable x qui permettront de déterminer les valeurs de y dans l'équation de la droite (D). Puis nous traçons la droite (D) à la règle, passant par ces deux points. La technologie qui justifie cette technique est la représentation d'une droite avec deux points.

Pour le troisième type de tâches qui est de déduire une estimation des remboursements en l'an 2000. Ce type de tâches correspond à T5. Ici il y a deux techniques relatives à ce type de tâches, soit ils utilisent l'équation de la droite calculée en a), soit ils font une lecture graphique pour déterminer l'estimation. Cela nous permet d'avoir en l'an 2000, $x=7$ et $y=10249$

Donc le montant des remboursements est estimé à 10249 euros en l'an 2000.

La technologie qui justifie la première technique est la lecture des données dans un tableau statistique, pour avoir la valeur de $x=7$. Et la méthode de résolution des équations du premier degré à une inconnue.

Concernant les initiatives de l'élève dans cet exercice, elles se situent aux questions suivantes :

Question 1, nommé T_Rep_Nuage. Il a la responsabilité de choisir l'échelle.

Question 4 b) Pour la représentation graphique de la droite (D), l'élève a à sa charge de trouver deux points de la droite, et de tracer à la règle la droite (D) passant par ces deux points.

Question 4 c) Il a la responsabilité de déterminer la valeur de x et de la noter dans le tableau statistique.

Nous allons mettre graphiquement en évidence les endroits où peut s'exercer l'initiative des élèves dans cet exercice, en extrayant de notre MPR les types de tâches qui s'appliquent dans cet exercice, en tenant compte de l'ordre des questions posées, et en mettant en couleur bleue les types de tâches pour lesquels l'élève a l'occasion de prendre des initiatives.

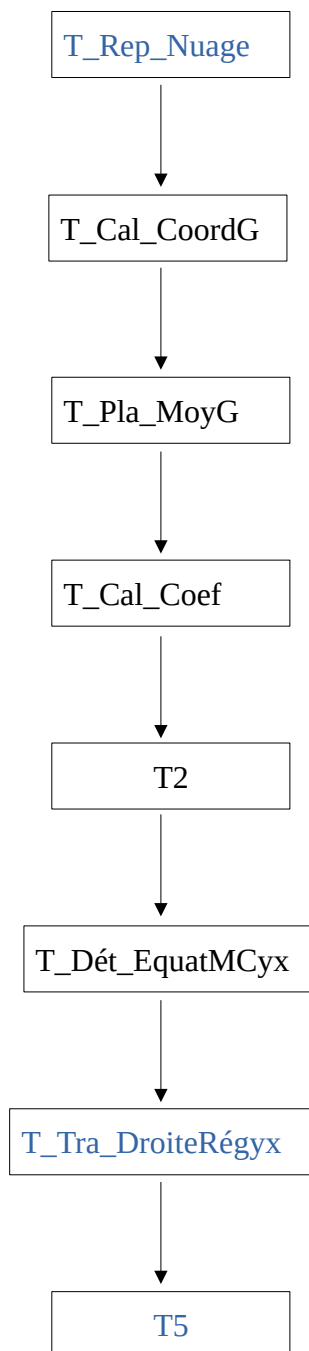


Figure 5: Types de tâches dans l'exercice résolu du manuel BTS et prises d'initiative possibles de l'élève

Cette représentation montre qu'il y a peu d'initiatives laissées dans l'exercice résolu du manuel de BTS. De plus, cette initiative concerne des savoirs précédents : choisir le bon repère, choisir deux points pour tracer une droite etc. Il est très peu lié à la démarche statistique elle-même.

5.2 Analyse du manuel de Centrafrique en utilisant le MPR

Dans cette section nous utilisons le MPR pour analyser le manuel de Centrafrique communément appelé CIAM. Nous allons notamment faire usage du MPR que nous avons précédemment construit dans la partie 5.1 pour nos analyses de la partie cours et de la partie exercices. Nous allons ensuite faire une analyse praxéologique d'un exercice résolu et revenir au MPR.

L'analyse du manuel de Centrafrique se fait uniquement sur le chapitre portant sur la statistique double. Ce chapitre est subdivisé en trois rubriques : cours, travaux pratiques, et exercices.

5.2.1 Analyse de la partie cours

La première rubrique consacrée au cours est répartie de la manière suivante :

- Séries statistiques doubles – Nuage de points ;
- Ajustement et corrélation linéaire.

Dans la première page de ce manuel, il est indiqué une introduction suivie du sommaire dans lequel, les auteurs décrivent l'organisation du contenu du chapitre.

Séries statistiques doubles – Nuage de points

Nous avons en introduction :

L'étude des séries statistiques à un seul caractère s'avère insuffisante lorsqu'il s'agit de certains phénomènes aléatoires dépendant d'au moins deux de caractères. Dans ces cas, on est contraint de considérer simultanément les caractères mis en exergue.

Toutefois, dans le cadre du programme de la terminale SE, nous allons cibler notre étude sur les séries statistiques à deux caractères quantitatifs appelés variables statistiques. Ces couples de variables statistiques sont représentés sous forme de tableaux statistiques à deux caractères ou séries statistiques doubles (Ouédraogo, p.270)

Cette introduction nous montre la limite de la série statistique à une variable et elle décrit la notion de séries statistiques doubles.

La notion de séries statistiques doubles dans le cas discret est abordée à l'aide d'un exemple sur les notes obtenues en mathématiques et en physique par 25 candidats à un concours (figure 6). L'exemple consiste à ordonner des données, dans le but d'établir d'abord un tableau statistique associé à la variable X qui correspond aux notes en mathématiques et un tableau statistique associé à la variable Y qui correspond aux notes en physique. Ensuite d'établir un tableau linéaire.

Exemple 1

On considère les notes obtenues en Mathématiques et en Physique par 25 candidats au concours d'admission en première année de BTS Électrotechnique Appliquée.

■ Tableaux linéaires

Le dépouillement et la mise en ordre des données nous fournissent une première représentation de la série statistique à deux caractères sous forme de **tableaux linéaires**.

TABLEAU I

Candidat	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Note de maths	6	9	13	9	10	11	9	11	12	14	16	11	14	7	9	12	14	16	11	12	12	10	8	9	12
Note de physique	11	13	12	13	9	10	13	10	9	7	8	13	9	4	13	11	14	8	13	16	11	9	17	13	11

2 candidats ont 10 en mathématiques et 9 en physique

Désignons par :

- \mathcal{P} la population étudiée c'est-à-dire l'ensemble des 25 candidats.
- X le caractère « note obtenue en Mathématiques »
- Y le caractère « note obtenue en Physique »
- M_X l'ensemble des modalités du caractère X
- M_Y l'ensemble des modalités du caractère Y

On a alors :

- $M_X = \{6 ; 7 ; 8 ; 9 ; 10 ; 11 ; 12 ; 13 ; 14 ; 16\}$; $M_Y = \{4 ; 7 ; 8 ; 9 ; 10 ; 11 ; 12 ; 13 ; 14 ; 16 ; 17\}$
- $X(B) = 9$; $X(F) = 11$; $Y(A) = 11$; $Y(C) = 12$
- l'effectif de la modalité 9 du caractère X est 5
- la distribution ou série statistique associée au caractère Y est définie par le tableau ci-dessous.

Candidat	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Note de physique	11	13	12	13	9	10	13	10	9	7	8	13	9	4	13	11	14	8	13	16	11	9	17	13	11

Figure 6: Introduction du registre « tableau linéaire » et du vocabulaire caractère, modalité, effectif, distribution.

Un deuxième exemple similaire est proposé dans le cas du regroupement en classes.

Ensuite nous avons la définition de la série statistique double qui commence par des explications sur les notations :

« P est une population d'effectif N, sur laquelle sont définis deux caractères X et Y,

$\{ x_1 ; x_2 ; \dots ; x_p \}$ l'ensemble des modalités du caractère X, noté M_X ,

$\{ y_1 ; y_2 ; \dots ; y_m \}$ l'ensemble des modalités du caractère Y, noté M_Y ,

φ l'application de P dans $R \times R$ définie par $\varphi(\alpha) = (X(\alpha) ; Y(\alpha))$,

x_i et y_j deux modalités respectives de X et Y.

On appelle effectif du couple $(x_i ; y_j)$ le nombre d'antécédents de ce couple par l'application φ . Cet effectif est noté n_{ij} .

On appelle série statistique double de caractère (X ; Y) l'ensemble des triplets $(x_i ; y_j ; n_{ij})$.

Le nuage de points est défini ainsi :

« Et on appelle nuage de points associé à la série statistique double de caractère $(X ; Y)$ la partie de F dont tous les points ont un couple de coordonnées $(x_i ; y_j)$ d'effectif non nul, où F est la représentation graphique de $M_X \times M_Y$ dans le plan muni d'un repère orthogonal ». (Haddad-ouédraogo, 1999)

Vu la structuration du chapitre, le premier point abordé ne constitue qu'une partie de T1, car la notion de point moyen est abordé dans la partie suivante.

En se servant de T2 par le biais de la méthode graphique, on aborde la deuxième partie du cours qui s'intéresse à la notion d'ajustement et du coefficient de corrélation linéaire.

Ajustement et corrélation linéaire

La deuxième partie aborde les notions d'ajustement et du coefficient de corrélation linéaire. Le manuel utilise deux terminologies de la notion d'ajustement tantôt : « ajustement affine » tantôt « ajustement linéaire ». Dans la suite, nous utiliserons la terminologie « ajustement affine ».

Elle débute par une introduction générale consacrée à la notion d'ajustement. Ce point est centré sur une présentation, l'objectif général de l'ajustement, l'ajustement graphique et le point moyen d'un nuage. En ce qui concerne la notion d'ajustement, il y a trois méthodes développées dans ce manuel :

La première méthode, nommée ajustement graphique, n'est rien d'autre que le type de tâches nommé **T4.1**, qui consiste à tracer une droite qui épouse le mieux la forme du nuage associée à la série statistique étudiée. Mais la qualité de l'estimation ou de l'écart valeur observée – valeur estimée. Concernant le point moyen, les auteurs ont donné la définition suivie d'un exemple.

Ensuite nous avons l'ajustement affine par la méthode de Mayer nommé **T4.2**, qui commence par le principe de la méthode qui consiste à partager l'ensemble des points (x_i, y_i) du tableau en deux sous ensembles G_1 et G_2 de même effectif. Nous nommons ce type de tâches « **T_PartagerG1G2** », s'ensuit une remarque puis un exemple.

Et enfin l'ajustement affine par la méthode des moindres carrés nommé **T4.3**. Dans cette partie, se trouve l'explication sur la notion de meilleur ajustement affine. Il s'agit d'explicitier que la droite d'ajustement d'un nuage qui réalise le meilleur ajustement est telle que la somme des carrés des écarts est minimale. Ensuite nous avons la définition de la droite de régression, de la covariance et la propriété de la méthode des moindres carrés.

La démonstration de l'équation de la droite des moindres carrés est donnée.

Un exemple suit cette propriété, il s'agirait de l'exemple dont on fera l'analyse à la partie 5.2.3.

La partie se termine par la notion de coefficient de corrélation linéaire et ses caractéristiques. Dans notre MPR, la notion de coefficient de corrélation se réfère au type de tâches **T_Cal_Coef**.

5.2.2 Analyse de la partie exercice

Travaux pratiques

Dans cette rubrique, nous avons deux travaux pratiques, le premier intitulé TP1 : exemple d'ajustement se ramenant à un ajustement affine et le deuxième intitulé TP2 : Ajustement affine et ajustement exponentiel.

Concernant le TP1, il y a plusieurs objets mathématiques utilisés dans cet exercice qui ne feront pas partie de notre étude, à l'exemple de l'ajustement du type exponentiel se ramenant à un ajustement affine. Par ailleurs, on retrouve les types de tâches tels que : T4.3, T_Cal_Coef et T5.

Le TP2 est subdivisé en trois parties, la première partie concerne la représentation et se réfère au type de tâches T1, s'ensuit la deuxième partie qui concerne l'ajustement affine. Il s'agit des types de tâches : T_Cal_Coef, T_Dét_EquatMCyx, T_Tra_DroiteRégyx et T5.

Pour réaliser les types de tâches T_Cal_Coef et T_Dét_EquatMCyx les auteurs ont utilisé un nouveau type de tâches qui ne figure pas dans notre MPR. Nous nommons ce type de tâches T_Etab_TabS : établir le tableau statistique ; nous faisons l'hypothèse que ce type de tâches est particulièrement présent en Centrafrique, et pas en France. Nous rediscutons ce point ci-dessous.

Quant à la troisième partie du TP2, elle porte sur l'ajustement exponentiel qui ne fait pas partie de notre objet d'étude.

Exercices : Entraînement et Problèmes

Dans cette partie nous mettons en avant, les exercices portant sur les notions : du nuage de points, le coefficient de corrélation linéaire, le point moyen, la méthode de Mayer, la méthode des moindres carrés, la représentation graphique et l'interprétation. Ce choix se justifie d'une part par la densité de ces notions dans les exercices et d'autre part ces notions mathématiques fait partie de notre objet d'étude. Le tableau 5 présente quinze exercices, qui abordent la notion du nuage de points, onze exercices la notion de coefficients de corrélation linéaire, trois exercices la notion de points moyen, trois exercices la méthodes de Mayer, dix exercices sur la méthodes des moindres de carrés, seize exercices sur la représentation graphique, quinze exercices sur l'estimation et deux exercices sur l'interprétation (c'est des exercices où on demande de commenter ou d'expliquer le résultat statistique trouvé dans un contexte bien défini).

Types de Tâches	Nombres d'exercices	Contexte des données
T_Rep_Nuage	15	<ul style="list-style-type: none"> • 8 exercices : contexte économique; • 4 exercices : contexte de la vie courante ; • 2 exercices : contexte scolaire ; • 1 exercice : contexte médical.
T_Cal_Coef	11	<ul style="list-style-type: none"> • 6 exercices : contexte économique; • 3 exercices : contexte de la vie courante ; • 1 exercice : contexte scolaire ; • 1 exercice : contexte médical.
T_Cal_CoordG	02	<ul style="list-style-type: none"> • 2 exercices : contexte économique

T_Pla_MoyG		
T_Cal_CoordG1G2 T_Pla_MoyG1G2	01	<ul style="list-style-type: none"> • 1 exercice : contexte scolaire
T4.2	03	<ul style="list-style-type: none"> • 2 exercices : contexte économique; • 1 exercice : contexte scolaire
T4.3	10	<ul style="list-style-type: none"> • 6 exercices : contexte économique; • 3 exercices: contexte de la vie courante ; • 1 exercice : contexte scolaire
T1	16	<ul style="list-style-type: none"> • 8 exercices : contexte économique; • 5 exercices: contexte de la vie courante ; • 2 exercices : contexte scolaire ; • 1 exercice : contexte médical.
T5	15	<ul style="list-style-type: none"> • 8 exercices : contexte économique; • 5 exercices: contexte de la vie courante ; • 1 exercices : contexte scolaire ; • 1 exercice : contexte médical.
T6	02	<ul style="list-style-type: none"> • 1 exercice : contexte économique; • 1 exercice : contexte médical.

Tableau 5: Tableau récapitulatif des types de tâches des exercices du manuel CIAM

Concernant les 15 exercices sur le nuage de points. Les tâches demandées à l'élève sont liées au type de tâches T_Rep_Nuage. Ce type de tâches laisse place à l'initiative de l'élève lorsque le choix de l'échelle est à la charge de l'élève.

Pour les onze exercices sur le coefficient de corrélation linéaire T_Cal_Coef ainsi que les trois exercices sur la notion de point moyen T_Cal_CoordG et T_Pla_MoyG. Les tâches que l'élève a à accomplir pour produire le résultat ne nécessitent pas beaucoup d'initiatives. En effet, pour produire le résultat, l'élève a besoin de connaître la formule du coefficient de corrélation linéaire et d'effectuer les calculs. Ces exercices peuvent poser des difficultés pour les étudiants qui n'ont pas de calculette.

Concernant les trois exercices portant sur l'ajustement affine par la méthode de Mayer T4.2. L'élève doit connaître le principe de la méthode pour pouvoir réaliser ce type de tâches. C'est le type de tâches travaillé en classe, il n'y a pas particulièrement d'initiative attendue de l'élève.

Pour la méthode des moindres carrés, nous avons dix exercices. Dans ce types de tâches T4.3, on demande à l'élève de faire l'ajustement de y en x ou de x en y . Cependant, il est toujours dit dans l'énoncé laquelle des deux il faut choisir. Pour la réalisation de ce type de tâches, l'élève doit mobiliser des techniques T_Dét_EquatMCyx (ou T_Dét_EquatMCxy) et des ingrédients de technique comme, T_Cal_Moyxy, T_Cal_Var, T_Cal_Cov et T_Tra_DroiteD. Ici non plus, il n'y a

pas d'initiative attendue (notamment, les élèves n'ont pas à justifier si il faut faire un ajustement de y en x ou de x en y).

Pour les quinze exercices liés à T5, où on demande à l'élève de faire une estimation de l'une des deux caractères étudiés. La prise initiative est importante, il doit tenir compte de l'hypothèse faite sur la variable potentiellement explicative pour estimer la variable à expliquer. Cela veut dire, que l'on a a priori accès à des valeurs exactes pour x à partir desquels, on cherche à estimer y (ou inversement)

Et enfin les exercices d'interprétations T6, sont des exercices où on demande à l'élève d'expliquer les résultats trouvés. La prise d'initiative de l'élève est importante, il a, à sa charge la responsabilité d'expliquer ou de justifier ses résultats mathématiques, en retournant dans le contexte concret de l'exercice.

5.2.3 Analyse praxéologique d'un exercice résolu du manuel en utilisant le MPR

Nous présentons un exemple d'un exercice résolu dans le manuel Mathématiques Terminale SE, à la page 283 (figure 7).

L'exercice porte sur la représentation graphique et l'ajustement. Nous nous sommes référé à cet exercice pour mettre en avant les diverses techniques d'ajustement utilisées.

■ Exemple

L'évolution de 1985 à 1989 du salaire horaire moyen d'un ouvrier, dans un pays en voie de développement donné, est consignée dans le tableau ci-dessous.

- 1) Représentons graphiquement les variations de ce salaire horaire moyen.
- 2) Ajustons à la règle le nuage de points obtenu par une droite dont on déterminera une équation cartésienne.
- 3) Ajustons le nuage de points par la méthode des moindres carrés.
- 4) En admettant que cette évolution se poursuive, donnons une estimation du salaire horaire moyen d'un tel ouvrier en l'an 2001.

Années	1985	1986	1987	1988	1989	1990	1991
Salaire horaire moyen en FCFA	1650	1760	1930	2020	2220	2450	2530

1) Représentation graphique des variations de ce salaire horaire moyen

On désigne les années par leurs rangs ; l'année 1985 aura le rang 1, l'année 1986 le rang 2, et ainsi de suite...

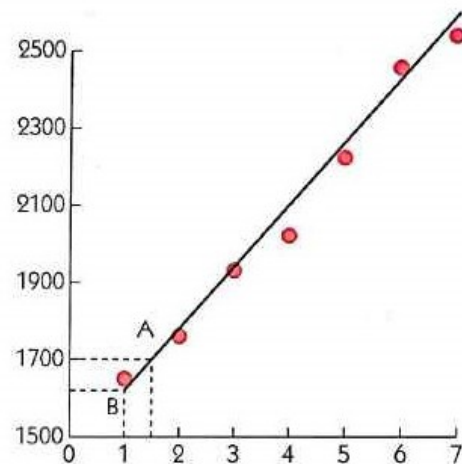
2) La droite d'ajustement passe par les points

$A(\frac{3}{2}; 170)$ et $B(1; 162)$.

On obtient une équation de la droite (AB) : $y = 16x + 146$.

3) Tableau des calculs

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$
1	1650	-3	-430	1290	9
2	1760	-2	-320	640	4
3	1930	-1	-150	150	1
4	2020	0	-60	0	0
5	2220	1	140	140	1
6	2450	2	370	740	4
7	2530	3	450	1350	9
28	14560	0	0	4310	28



On a : $\bar{x} = \frac{28}{7} = 4$; $\bar{y} = \frac{14560}{7} = 2080$;

d'où : $a = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{4310}{28} = 153,92$; $b = \bar{y} - a\bar{x} = 2080 - (4 \times 153,92) = 1464,32$.

La droite de régression a pour équation : $y = 153,92x + 1464,32$.

4) La droite des moindres carrés est la droite de régression de Y en fonction de X,

c'est aussi la droite d'estimation de Y à partir de X. L'an 2001 aura le rang 17, d'où l'estimation obtenue en remplaçant dans l'équation (1) x par 17. On obtient : $y = 4083,6 \approx 4084$.

Donc, 4 084 FCFA est une estimation du salaire horaire moyen de cet ouvrier en l'an 2001.

Figure 7: Extrait du manuel Mathématiques Terminale SE

La première question correspond au type de tâches T_Rep_Nuage. Ce type de tâches consiste à représenter graphiquement les variations du salaire (les données). La technique consiste à choisir l'échelle de la manière suivante : nous désignons les années par leurs rangs (les variables x_i), c'est-à-dire l'année 1985 correspond par exemple à 1cm, l'année 1986 correspond à 2cm et ainsi de suite... Et le salaire horaire moyen correspond aux variables y_i . Puis nous plaçons ces couples de

points dans le repère. La technologie qui justifie cette technique est la connaissance sur le repère orthonormé et le choix de l'échelle.

La deuxième question regroupe deux types de tâches. Le premier type de tâche est: ajuster à la règle le nuage de points obtenu par une droite que nous avons nommée : T4.1; et le deuxième type de tâches est : déterminer l'équation cartésienne nommé **T_Dét_EquatC**.

Le premier type de tâches est relatif à la résolution graphique qui fait référence au registre graphique, et le second type de tâches est relatif à la résolution algébrique qui fait aussi référence au registre algébrique. La technique de T4.1 est de tracer une droite à la règle qui passe au plus près du nuage de points. On ne sait pas ici comment les auteurs ont tracé. Et la technique en lien avec le deuxième type de tâches T_Dét_EquatC consiste à repérer deux points de la droite et utiliser leurs coordonnées pour trouver l'équation. Il y a une infinité de choix possibles pour cette droite, mais les auteurs du manuel ne le disent pas dans leur corrigé. La rédaction du corrigé n'explique pas qu'il y avait plusieurs choix, ni pourquoi les auteurs ont choisi de retenir les points notés A(3/2 ; 1700) et B(1 ; 1620). Nous notons d'ailleurs que le manuel parle par erreur de 170 et 162.

Les élèves ont a priori la responsabilité, dans cet exercice, de choisir la droite qui leur paraît approcher le mieux les points du nuage. Mais à la lecture du corrigé cette initiative possible n'est pas discutée, il n'y a aucun discours technologique dans le corrigé.

La troisième question appartient au type de tâches T4.3 : ajuster le nuage de point par la méthode des moindres carrés. La technique qui permet de déterminer l'équation de la droite qui ajuste mieux le nuage de points, comporte un ingrédient qui consiste à faire un tableau à double entrée. C'est une technique nommée T_Etab_TabS : établir le tableau statistique, relative à la réalisation de T4.3. Cette technique est couramment utilisée en Centrafrique pour réaliser T4.3. Nous notons que le tableau est nommé ici 'tableau des calculs' ceci explique bien son rôle. Ce tableau permet de déduire la valeur coefficient directeur

$$a = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{4310}{28} = 153,92 .$$

Ensuite de remplacer la valeur de a dans l'expression de $b = \bar{y} - a \cdot \bar{x} = 2080 - 4 \times 153,92 = 1464,32$.

Alors, nous avons la droite de régression qui ajuste le nuage de points et a pour équation : $y = 1532,92 \cdot x + 1464,32$. La technologie qui justifie cette technique est la définition de la moyenne, la variance, la covariance et la formule du coefficient directeur d'une droite.

La prise d'initiative de l'élève est réduite pour ce type de tâches, qui ne nécessite que l'application de la méthode des moindres carrés. Et dans les classes de terminale SE en Centrafrique, les calculatrices scientifiques avec les fonctions statistiques sont autorisées, mais peu nombreuses.

Quant à la quatrième question, le type de tâches est : T5 donner une estimation du salaire horaire moyen en 2001. La technique qui permet de réaliser ce type de tâches est de convertir l'année 2001 au rang 17. Puis en remplaçant $x=17$ dans l'équation de la droite de régression $y = 1532,92 \cdot x + 1464,32$, nous aurons : $y = 1532,92 \times 17 + 1464,32 = 4083,6 \simeq 4084$. L'estimation obtenue est 4084 franc CFA. La technologie relative à cette technique est la résolution algébrique de l'équation du premier degré. Ici l'élève applique la méthode apprise. Mais la conversion de l'année 2001 au rang $x=17$ nécessite d'une grande prise d'initiative de la part de l'élève.

Les initiatives possibles de l'élève dans cet exercice :

- La prise d'initiative de l'élève dans cette représentation graphique correspond à sa la responsabilité du choix de l'échelle.
- La prise d'initiative de l'élève se situe dans la conversion de l'année 2001 au rang $x=17$ pour trouver y .

Types de Tâches	Techniques	Technologies
T_Rep_Nuage	choisir l'échelle et représenter le graphique	Connaissances sur le choix de l'échelle et sur le repère orthonormé
T4.1 ; T_Dét_EquatC	Tracer une droite qui passe au plus près du nuage de points et l'on a le choix entre plusieurs droites; Si la droite d'ajustement tracée, passe par deux points $A(x_A, y_A)$ et $B(x_B, y_B)$ des nuages, alors l'équation est $y=ax+b$. Le coefficient directeur de la droite (AB) est $a = \frac{y_B - y_A}{x_B - x_A} = \frac{1620 - 1700}{1 - 3/2} = 160$ Nous déterminons son ordonnée à l'origine, en utilisant un des points de la droite. Choisissons le point B par exemple, on a $y_B = a \cdot x_B + b$ $1620 = 160 \times 1 + b$, on en déduit que $b = 1460$. D'où l'équation est : $y = 160 \cdot x + 1460$	Propriété géométrique : deux points de la droite ; Notion d'ajustement affine
T4.3	T_Etab_TabS : Faire un tableau de calculs statistiques. Puis nous déduisons le coefficient directeur a et b D'où l'équation de la droite de régression : $y = 1532,92 \cdot x + 1464,32$	Définition de la moyenne arithmétique, de la covariance et la formule de coefficient directeur.
T5	Faire la conversion de l'année 2001 en rang 17 (échelle équivaut à 17cm). Puis remplacer x par 17 dans l'équation $y = 1532,92 \cdot x + 1464,32$ L'estimation obtenue est $y = 1532,92 \times 17 + 1464,32 = 4083,6 \simeq 4084$ Soit 4084 fcfa de salaire en l'an 2001 pour cet ouvrier.	Résolution algébrique de l'équation du premier degré.

Tableau 6: Grille récapitulative de l'analyse praxéologique de l'exercice

Dans la représentation suivante, issue du MPR nous mettons en couleur bleue les types de tâches pour lesquels l'élève prend des initiatives.

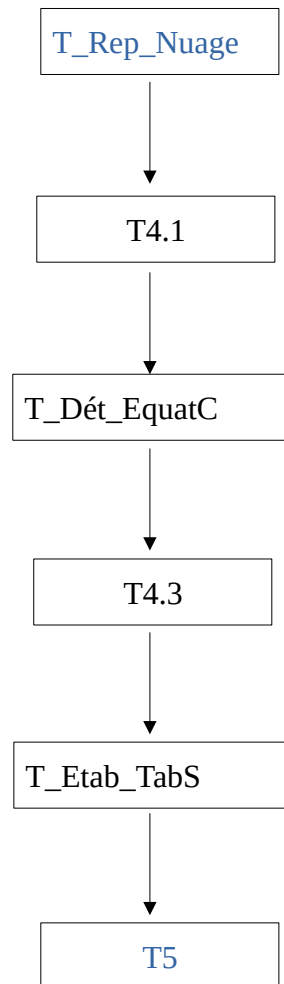


Figure 8: Représentation des types de tâches présents dans l'exercice résolu du manuel Mathématiques Terminale SE, mise en évidence des prises d'initiative

5.2.4 MPR complété suite à l'analyse du manuel de RCA

Dans ce qui suit, nous présentons le MPR complété suite à l'analyse du manuel Mathématique Terminale SE.

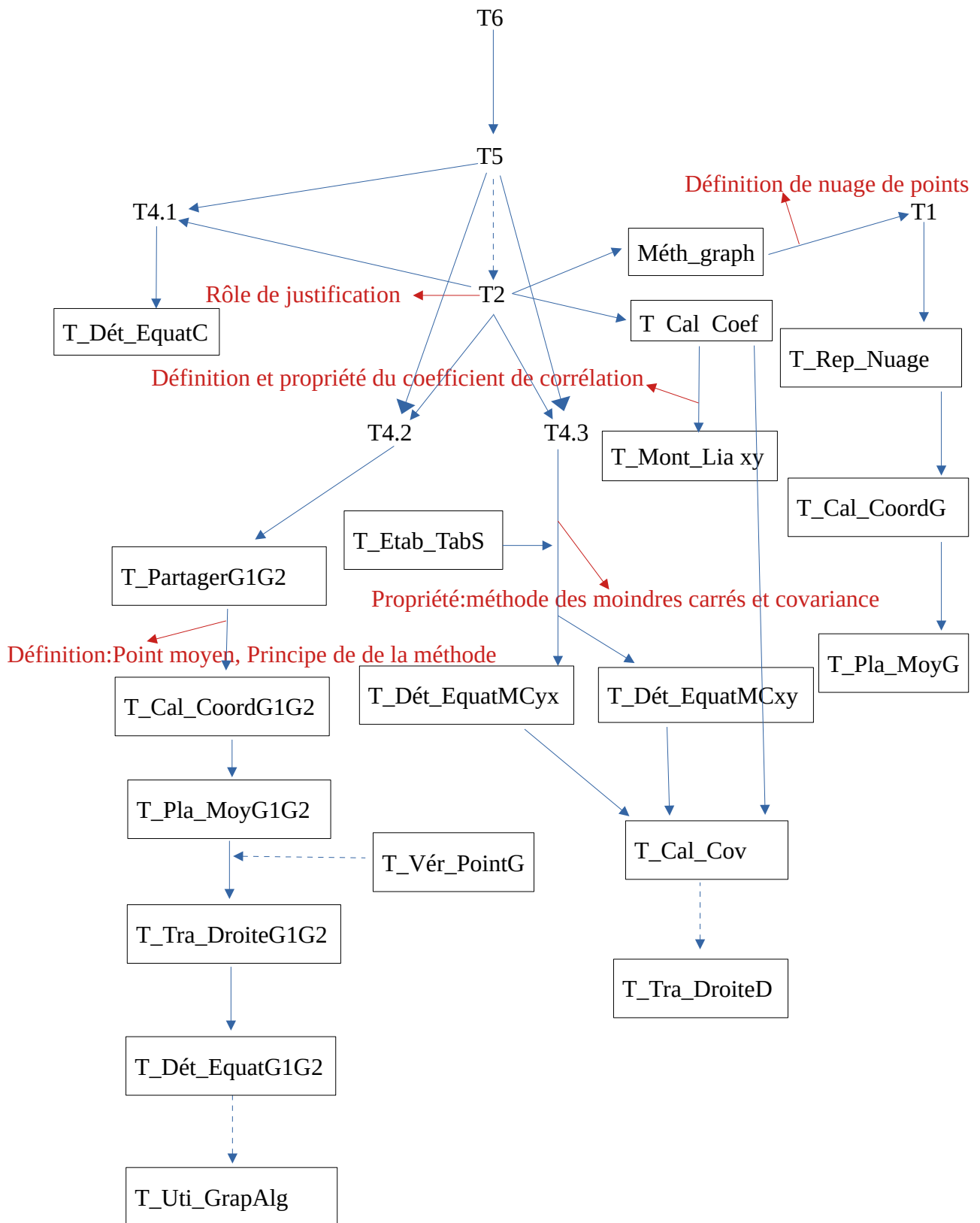


Figure 9: Modèle praxéologique de référence complété

Cette représentation montre l'ajout de trois nouveaux types de tâches : T_Etab_TabS, T_PartagerG1G2 et T_Dét_EquatC et la prise en compte du type de tâches T_Rea_AjReg. La réalisation du type de tâches T_Etab_TabS nécessite l'utilisation de calculatrice.

5.3 Analyse d'un enseignement « ordinaire » en RCA

Dans cette partie, nous souhaitons décrire et analyser un enseignement de la série statistique à deux variables, pendant un enseignement ordinaire dans une classe de terminale SE en Centrafrique.

Notre but est de compléter notre enquête sur l'enseignement « ordinaire » de l'ajustement linéaire pour des séries statistiques doubles, afin de déterminer où se situent les possibilités de prises d'initiatives des élèves, et si ces possibilités donnent lieu à des prises d'initiatives effectives. Nous avons donc observé un enseignant (que nous nommons Aristide), en 2018-2019.

Le recueil de données relatif à cette partie a été décrit dans la partie méthodologique de notre travail. Nous rappelons brièvement que nous avons observé trois séances en terminale SE en Centrafrique. Nous avons transcrit les vidéos des deux premières séances et réalisé un synopsis de ces séances.

Pour réaliser une analyse de l'activité du professeur et des élèves en classe, nous commençons d'abord par décrire ce qui se passe au cours de la séance d'enseignement. Par exemple :

- quand le professeur écrit au tableau et que les élèves prennent note sur leurs cahiers et/ou ils suivent l'explication ;
- lorsqu'un élève vient au tableau pour résoudre un problème ou pour répondre à une question ;
- lorsque le professeur et les élèves interagissent ensemble pour la réalisation d'une réponse à une question ou pour assurer l'avancée du savoir.

Nous analysons ensuite les organisations mathématiques (les types de tâches, les techniques, les technologies et les théories) proposées par le professeur, en se référant aussi au MPR.

Nous nous appuyons notamment sur les questions suivantes pour faire notre analyse :

- Pour chaque type de tâche, la technique introduite par l'enseignant est-elle accompagnée d'un discours technologique ?
- A quels moments la responsabilité des élèves est-elle importante ou réduite pendant cet enseignement de la statistique ?

Pour chacune des deux séances, nous présentons un synopsis sous forme de tableau. Ensuite nous faisons l'analyse d'extraits de la séance.

5.3.1 Description et analyse de la première séance de l'enseignement ordinaire

Nous donnons ci-dessous le synopsis de cette séance.

L'énoncé figure après le synopsis.

Minutage	Phases	Savoirs en jeu	Modalités de travail	Descriptions
00 :07	Présentation de l'énoncé de l'activité	Représentation graphique des séries doubles. Ajustement affine par la méthode des moindres carrés.	L'enseignant écrit l'intégralité de l'énoncé au tableau.	L'enseignant écrit au tableau.
<p><u>Commentaires</u> : Le professeur commence la leçon sur la série statistique double, par une «application». Il copie au tableau l'énoncé d'un exercice, qui se trouve dans le manuel Transmath ES, de France.</p>				
09 :50	Représentation du nuage des points	Nuage des points Repère, origine coordonnées, échelle	Travail collectif : L'enseignant donne des explications sur l'exercice. Les élèves désignés vont au tableau. Les élèves interrogés répondent aux questions. Échanges entre l'enseignant et élèves.	L'enseignant explique l'échelle du repère pour la représentation des nuages. L'enseignant commence à corriger l'exercice avec la participation des élèves. Les élèves doivent représenter le nuage des points sur leur cahier. L'enseignant affirme qu'à partir de la figure, le nuage de points permet d'effectuer un ajustement affine.
<p><u>Commentaires</u> : Le professeur corrige la première question, il représente le nuage des points dans un repère. Il représente sur le repère les coordonnées du premier point $M_1(36, 12)$, il laisse la latitude aux élèves de représenter dans leur cahier le reste des points qui forment le nuage des points. La question 2.a), on demande de constater que le nuage de points permet d'effectuer un ajustement affine. Le professeur affirme que le nuage des points permet d'effectuer un ajustement affine, sans finir de représenter les coordonnées des autres points. Or la disposition de ces points du nuage est importante pour savoir si un ajustement affine est approprié.</p>				
28 :36	Calcul des valeurs nécessaires à la détermination de l'équation de la droite de régression	La notion de la moyenne	Travail collectif : Deux élèves passent au tableau faire des calculs avec leur calculatrice. Les élèves interrogés répondent aux questions. Échanges entre l'enseignant et la	Le professeur fait un tableau de calculs statistique au tableau. Il demande aux élèves de compléter le tableau. C'est-à-dire qu'ils doivent recopier le tableau sur leurs cahiers et le compléter. Une élève est au tableau avec sa calculatrice pour compléter le tableau statistique. Une deuxième élève est au tableau pour compléter le reste du tableau statistique.


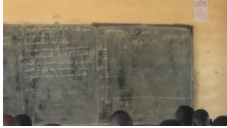
			deuxième élève.	L'enseignant complète le reste du tableau statistique.
<p><u>Commentaires</u> : La première élève qui passe au tableau, calcule la somme des x_i, les différentes valeurs des x_i^2 ainsi que leur somme.</p> <p>La deuxième élève calcule aussi les différentes valeurs des y_i^2 ainsi que leur somme. Mais elle a des difficultés à manipuler la calculatrice et à lire les chiffres qui s'affichent à l'écran de la calculatrice.</p>  <p>Les élèves dictent le reste des valeurs qu'ils ont déjà calculées dans leur cahiers. Et le professeur complète le reste du tableau.</p>				
52 :00	Détermination du point moyen.	La moyenne, le point moyen	Travail collectif : l'enseignant pose des questions, les élèves répondent.	Utilisation du tableau statistique pour calculer \bar{x} , \bar{y} , les coordonnées du point moyen G
<p><u>Commentaires</u> : Le professeur se sert du tableau statistique pour calculer : \bar{x}, \bar{y}, les coordonnées du point moyen G. L'objectif de ces calculs, c'est de les utiliser pour déterminer l'équation de la droite, mais le professeur ne l'a pas fait ; Il affirme que la droite d'ajustement passera par le point moyen G mais il ne donne pas de justification. Il est possible à ce moment qu'il lui manque un document avec les formules donnant l'équation de la droite des moindres carrés.</p> 				
53:26	Indication pour la troisième question.	Équation cartésienne	L'enseignant interroge un élève. Un élève pose la question à l'enseignant (57:17). L'enseignant répond à la question posée.	L'enseignant explique la troisième question de l'exercice. Il explicite comment faire pour trouver la valeur de y . Il donne des indications aux élèves. Comment ils doivent faire pour calculer la tension artérielle.
<p><u>Commentaires</u> : Les élèves doivent calculer la tension artérielle, en utilisant l'équation de la droite trouvée. Par conséquent, ils ne peuvent pas le faire, du fait qu'ils n'ont pas déterminé l'équation de la droite. L'enseignant donne des indications qui permettront de calculer cette tension artérielle, sans avoir déterminé l'équation de la droite.</p>				

Tableau 7: Synopsis de la première séance

L'enseignant introduit son cours par une activité. Notons que les élèves n'ont encore jamais rencontré de série statistique double, ni naturellement entendu parler d'ajustement ou de droite des moindres carrés. L'activité doit donc pouvoir servir à introduire les concepts et les méthodes ;

l'enseignant aura certainement un rôle important dans la résolution, on ne peut pas s'attendre à ce que les élèves « découvrent » la méthode des moindres carrés.

Ainsi nous commençons par une analyse a priori de cette activité, qui nous permettra de mettre en avant certains enjeux de l'enseignement de la série statistique à deux variables au cours d'un enseignement ordinaire en terminale SE.

Analyse a priori de l'activité « tension artérielle »

Nous présentons ci-dessous l'énoncé de l'activité choisie par le professeur pour construire son cours, puis nous en faisons une analyse a priori, du point de vue d'élèves qui connaîtraient déjà la méthode des moindres carrés.

Énoncé

Le tableau suivant donne la moyenne Y des maxima de la tension artérielle en fonction de l'âge X dans une population donnée.

Âge X	36	42	48	54	60	66	70
Tension Y	12	13	13,5	14,3	15,4	15,8	16

1) Représenter graphiquement le nuage de points $M(x, y)$. On prendra pour unités graphique 0,5cm pour 1 an en abscisse et 3cm pour l'unité de tension artérielle en ordonnée. L'origine sera le point de coordonnées (30 ;10).

2.a) Constater que le nuage de points permet d'effectuer un ajustement affine et donnez une équation de la droite des moindres carrés.

b) Représenter cette droite.

3) Une personne de 75 ans a une tension artérielle de 17,5. Quelle serait sa tension théorique en utilisant la droite des moindres carrés ? Comparez avec la tension réelle.

Analyse a priori

Cette activité contient des questions qui sont liées, aux diverses étapes où les productions demandées aux élèves sont un graphique à produire et à exploiter, des résultats numériques à calculer, l'équation de la droite à déterminer et à exploiter.

1) Cette activité commence par une tâche qui consiste à représenter graphiquement le nuage de points $M(x, y)$. Cette tâche relève du type de tâches « T_Rep_Nuage ».

Dans ce problème, il s'agit de réaliser une conversion entre le registre « tableau » de départ et le registre graphique. L'échelle est donnée, en abscisse 0,5cm correspond à un an et en ordonnée 3cm correspond à l'unité de tension. De plus, l'origine du repère est fixée au point de coordonnées (30 ; 10) . Pour représenter le nuage de points dans le repère, les élèves doivent mobiliser des connaissances antérieures : la notion d'échelle, de repère, d'origine et la représentation des points dans un repère. Il n'y a pas de prise d'initiative pour la construction du repère, mais on peut dire que ensuite, placer les points dans le repère demande une certaine forme d'initiative pour la conversion du registre tableau au registre graphique. Il peut y avoir des erreurs dans le placement des points si les élèves oublient que l'origine est (30;10).

La deuxième question comprend deux questions 2.a) et 2.b).

2.a) La question 2.a) relève du type de tâches T2 et le contrat prévoit de répondre par oui. La production demandée consiste d'abord à observer la forme du nuage de points. Comme le nuage de points présente une forme rectiligne, alors le nuage peut être approché par une droite. C'est cette observation de la forme du nuage de points qui est attendue comme justification de la pertinence d'un ajustement affine.

La détermination de l'équation de la droite des moindres carrés relève du type de tâches « T_Dét_EquatMCyx ». La technique consiste à utiliser les formules $a = \frac{cov(x, y)}{\sigma_x^2}$ et $b = \bar{y} - a\bar{x}$.

Pour calculer la valeur de a, deux variantes de la technique sont possibles.

La première variante consiste à rassembler les données dans un tableau de calculs statistique, qui permettrait de calculer la valeur de a . Cette technique est privilégiée dans les enseignements en Centrafrique, nous la présentons ci-dessous

									Totaux
x_i	36	42	48	54	60	66	70	376	376
y_i	12	13	13,5	14,3	15,4	15,8	16	100	100
x_i^2	1296	1764	2304	2916	3600	4356	4900	21136	21136
y_i^2	144	169	182,25	204,49	237,16	249,64	256	144254	144254
$x_i y_i$	432	546	648	772,2	924	1042,8	1120	5485	5485

$$\bar{x} = \frac{1}{n} \sum_1^7 x_i = \frac{1}{7} (376) = 53,71 \text{ et } \bar{y} = \frac{1}{n} \sum_1^7 y_i = \frac{1}{7} (100) = 14,28$$

$$cov(x, y) = \frac{1}{n} \sum_1^n x_i y_i - \bar{x} \bar{y} = 16,59 ; \sigma_x^2 = \frac{1}{n} \sum_1^7 x_i^2 - (\bar{x})^2 = 134,66$$

$$a = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{16,59}{134,66} = 0,12$$

Quant à la deuxième variante, elle consiste à calculer directement la valeur de a , en utilisant la

$$\text{formule : } a = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}}{\frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2} = 0,12$$

Ensuite de déduire la valeur de b , $b = \bar{y} - a\bar{x} = 14,28 - 0,12 \times 53,71 = 7,83$

D'où l'équation de la droite $y = 0,12x + 7,83$

La question 2.b) demande de représenter la droite d'équation $y = 0,12x + 7,83$. Ceci relève du type de tâches « T_Tra_DroiteD ». A priori, la question ne pose pas de problème puisque ces élèves sont supposés avoir une maîtrise de la technique de repérage des points dans un plan, où ils peuvent mobiliser des connaissances qui permettent de tracer la droite donnée par l'équation.

Pour cela, si nous attribuons par exemple deux valeurs supérieures à 30 à la variable x pour déterminer les valeurs de la variable y , ensuite nous plaçons les coordonnées (x, y) de ces deux points dans le repère d'origine $(30; 10)$. En joignant ces deux points par une droite, nous aurons notre droite d'équation $y = 0,12x + 7,83$. On peut noter ici qu'il s'agit uniquement de représenter une demi-droite, les valeurs de l'âge comme celles des tensions étant toujours positives.

Pour la troisième question, la tâche proposée consiste à déterminer la tension théorique sachant que l'âge de la personne est 75 ans (en utilisant la droite de moindres carrés), ça relève du type de tâches « T5 ». Puis de comparer avec la tension réelle qui est 17,5, ici cette tâche relève du type de tâches « T6 ».

Ici, 75 ans représente la valeur de la variable x qui permet de déterminer la valeur de la variable y correspondant à la tension théorique.

$y = 0,12x + 7,83$ l'équation de la droite de moindres carrés.

$x = 75$ on a : $y = 0,12 \times 75 + 7,83 = 16,83$. Alors la tension théorique obtenue par la droite des moindres carrés est $y = 16,83$

Donc la tension réelle est supérieure à la tension théorique ($17,5 > 16,83$).

Par ailleurs, nous pensons que les élèves peuvent proposer d'autres types de réponses qui ne mèneront pas nécessairement à la solution voulue pour cette tâche. Ces réponses peuvent être erronées soit à cause de l'équation de la droite erronée, ou soit à cause de la confusion faite sur la valeur de la variable x et y . Les initiatives des élèves sont réduites dans cette question, puisqu'on indique la méthode « en utilisant la droite des moindres carrés ». Il faut cependant qu'ils interprètent l'énoncé, pour identifier la valeur de la « tension réelle ». Ici l'enjeu de la question n'est pas très clair. Il n'est pas nécessaire de faire une estimation, puisqu'on connaît la tension réelle. Il s'agit donc de discuter sur la qualité de l'estimation : on obtient 16,83, au lieu de 17,5, est-ce que ça pose un problème ? Mais on ne dispose pas des moyens pour répondre à une telle question. Les

élèves pourraient tout de même dire que l'on conclut que la personne concernée ne semble pas souffrir d'hypertension ; mais il n'est pas clair qu'une telle réponse soit attendue par le professeur. Nous pouvons supposer que peu d'élèves iront dans ce sens, le contexte ne leur étant sans doute pas familier.

Analyse d'épisodes de la première séance

Nous allons regarder de plus près certains passages de cette séance, consacrée à l'activité « tension artérielle » décrite ci-dessus.

L'enseignant commence par écrire l'énoncé au tableau et les élèves prennent note sur leurs cahiers. Concernant le premier type de tâche « T_Rep_Nuage », l'enseignant donne des explications aux élèves sur l'origine du repère. L'enseignant demande aux élèves de représenter le nuage de points dans leurs cahiers.

Si on prend une unité comme étant le carreau, l'unité est égale à un carreau. Si on arrive à placer 36 quelque part ici, 36 à 70, ça va nous faire 34 et on voit que 34 carreaux, c'est trop grand, il faut le réduire. Nous disons qu'une unité va correspondre à un demi carreau.

Si une unité correspond à un demi carreau, alors un carreau va correspondre à deux unités. Un carreau sur l'axe des abscisses va correspondre à deux unités, parce qu'on prendra pour unité graphique 0,5cm pour un an ou pour une unité sur l'axe des abscisses. En ordonnée, on prendra 3cm puisque nous sommes entrain de convertir, les centimètres en carreau puisque, nous travaillons euh_ _ _ votre cahier... (Séance 1, 16:52).

Ensuite il interagit avec les élèves pour construire le repère et représenter le nuage de points. Ainsi, l'enseignant trace le repère au tableau, en expliquant comment est gradué le repère. Après avoir représenté le premier point du nuage, il laisse le temps aux élèves de placer les autres points du nuage sur leurs cahiers. Force est de constater, que les élèves n'ont pas pu finir la représentation du nuage sur leurs cahiers et l'enseignant passe au deuxième type de tâches. Cependant, la technique introduite consiste à convertir les unités en carreau et ensuite de les représenter sur le repère. Cette technique est accompagnée de discours technologique qui est le repérage dans un plan.

Le deuxième type de tâches T2 consiste à observer la forme du nuage. L'observation de la forme du nuage relève de la technologie, du sens de l'ajustement.

« En ce qui concerne la deuxième question, on demande en petit a, de constater que le nuage de points permet d'effectuer un ajustement affine. De constater à partir de la figure, mais la figure nous permet de constater effectivement qu'un ajustement affine est justifié » (Séance 1, 27:34).

Comme l'enseignant n'a pas fini la représentation du nuage de points, il explique oralement que la forme du nuage permet de confirmer qu'un ajustement affine est justifié. Mais il n'avait pas expliqué concrètement ce qu'est un ajustement affine.

Le troisième type de tâches T_Tra_DroiteD consiste à déterminer une équation de la droite des moindres carrés. Nous rappelons que la méthode des moindres carrés est un savoir nouveau pour ces élèves. L'enseignant leur explique qu'il faut faire un tableau statistique pour calculer la moyenne, la variance et la covariance, mais il n'a pas donné la définition de la covariance.

« vous complétez le tableau par les variables x_i^2 , y_i^2 , $x_i y_i$ et aussi les totaux. De calculer des moyennes \bar{x} , \bar{y} , de calculer la covariance de x , y , de calculer la variance de x » (Séance 1, 28:36).

Dans cette partie, l'enseignant s'appuie sur des savoirs anciens des élèves. Pour la statistique à une variable, les élèves ont vu les notions de moyenne et de variance.

L'enseignant désigne une première élève qui passe au tableau avec sa calculatrice en main. Elle commence par calculer la somme des x_i , des y_i , les différentes valeurs des x_i^2 et leur somme dans le tableau. Ensuite, pour y_i^2 l'enseignant fait passer un autre élève au tableau pour continuer.

de la tension arterielle en fonction de l'age x d'une population donnée.

Age x_i	36	42	48	54	60	66	70	376
Tension y_i	12	13	13,5	14,3	15,4	15		100
x_i^2	1296	1764	2304	2916	3600			
y_i^2								
$x_i y_i$								

Figure 10: Une élève au tableau pour remplir le 'tableau statistique'.

Comme on le voit à la figure 10, le « tableau de calculs statistique » ne comporte que les valeurs de x_i , de y_i , leurs carrés, leurs produits, et en fin de chaque ligne la somme des éléments de la ligne. Il n'y a pas dans le tableau les moyennes, les variances ni la covariance. Ce tableau est utile dans le contexte de la RCA parce que peu d'élèves ont des calculatrices, il faut s'assurer que les valeurs sont justes. Mais le remplissage du tableau est long, et se fait aux dépens de la compréhension du sens de la démarche statistique.

Pour remplir la ligne des produits $x_i y_i$ professeur interagit par oral avec les élèves sur les valeurs et il écrit leurs propositions dans le tableau statistique.

Pour calculer les moyennes, les élèves font les calculs et l'enseignant écrit la réponse proposée au tableau. Ici le professeur s'appuie sur une connaissance préalable des élèves, à propos de la notion de moyenne qu'ils ont vue dans l'étude des série statistiques à une variable.

« Et donc le point moyen G , c'est le point de coordonnées \bar{x} , \bar{y} , c'est-à-dire 53,71 et 14,28.

On aura également à placer dans le repère le point G. On verra également que cette droite de régression, la droite des moindres carrés. Cette droite (D) passe par le point G, c'est-à-dire que le point G appartient à la droite (D) » (Séance 1, 53:10).

Notons que la détermination du point moyen G n'est pas une question de l'activité. L'enseignant a choisi d'ajouter cette question, parce que cette activité devrait lui servir de support pour faire le cours sur la méthode des moindres carrés.

Par ailleurs, il demande aux élèves de placer le point moyen G sur le graphique précédent qui est dans leur cahier. Notons qu'en ce qui concerne la représentation graphique, l'enseignant demande aux élèves de le faire eux-même dans leurs cahiers. Durant cette activité, l'enseignant n'a pas fait la représentation graphique au tableau. Il ne pourra donc pas faire de commentaire sur la position du point G par rapport au nuage de point en s'appuyant sur un graphique.

Concernant le calcul des coordonnées du point moyen, il y a un discours technologique qui justifie la technique utilisée. Par contre, pour le types de tâches T_Dét_EquatMCyx l'enseignant n'a pas fait de discours technologiques qui permet de déterminer les coefficients a et b .

Pour la troisième question, dont le type de tâches se réfère à T5, l'enseignant explique aux élèves que la droite de régression c'est le lien qu'il y a entre y et x . Comme l'équation de la droite n'est pas déterminée, il ne sera pas possible d'estimer la valeur de y .

Nous notons que l'enseignant parle de la droite de régression sans la définir. Il n'a pas expliqué ce qui est une régression. Nous faisons l'hypothèse qu'il manque à l'enseignant un document où figurent les formules donnant l'équation de la droite des moindres carrés.

« Une personne de 75 ans a une tension artérielle 17,5, une personne de 75 ans a une tension artérielle réelle de 17,5, c'est-à-dire une personne de 75 ans doit avoir réellement une tension de 17,5. Mais théoriquement à partir de cette droite de régression, à partir de ce lien qu'il y a entre x et y . Quelle serait alors cette tension ? Est-ce que c'est la même chose ? Si je donne la valeur 75 à x , quelle serait alors la valeur de y ? Je calcule, est ce que ça va me donner effectivement 17,5? Quelle serait sa tension théorique, en utilisant la droite des moindres carrés ? » (Séance 1, 55:17).

C'est dans cette optique que l'enseignant donne des indications aux élèves pour qu'ils puissent finir l'activité eux mêmes à la maison.

En ce qui concerne l'analyse de la première séance, l'enseignant a une manière très particulière pour introduire des savoirs nouveaux. Il introduit des savoirs nouveaux à travers une activité au début de son cours. On peut penser que ceci vise à favoriser les initiatives des élèves, via une « activité d'introduction » qui va leur faire découvrir le savoir. Cependant les partages de responsabilités entre le professeur et les élèves fait que l'initiative de ceux-ci reste très limitée. Par ailleurs l'enseignant lui-même ne semble pas maîtriser suffisamment la méthode, ou disposer de documents lui donnant accès aux formules qu'il aurait dû introduire. Nous rappelons que nous nous basons sur le synopsis et le transcript pour faire notre analyse.

Au début de la séance l'enseignant écrit l'énoncé au tableau et les élèves prennent note, le niveau de responsabilité des élèves est très faible.

La première tâche consiste à faire la représentation des points du nuage dans un repère dont l'origine est le point de coordonnées (30, 10), elle relève du type de tâches T_Rep_Nuage. L'accomplissement de ce type de tâches repose sur des connaissances suivantes : repère cartésien du plan, coordonnées d'un point, lecture d'un tableau statistique. Il s'agit de savoirs déjà rencontrés antérieurement par les élèves. Dans ce type de tâches, la responsabilité des élèves peut être importante, s'ils ont à leur charge le choix de l'échelle de la représentation ou celui de l'origine. Ici elle est déjà réduite par l'énoncé de l'exercice, qui donne l'origine et l'échelle. Puis le professeur réduit encore, en donnant des conseils très concrets sur le nombre de carreaux du cahier qui correspond à la graduation.

D'où la responsabilité des élèves vis-à-vis du savoir est réduite. Pour établir le repère, l'enseignant représente le premier point du nuage $M_1(36,12)$, ensuite les élèves ont la responsabilité de représenter les autres points du nuage dans leurs cahiers.

La deuxième tâche consiste d'abord à observer la forme du nuage de points et en déduire qu'on peut faire un ajustement ceci relève du type de tâches T2. Pour observer la forme du nuage, les élèves ont besoin de la représentation graphique du nuage. Force est de constater que l'enseignant n'a pas fini la représentation du nuage de points au tableau. Il a représenté un seul point du nuage et le reste est confié à la charge des élèves. Qui, eux aussi n'ont pas pu finir la représentation avant que l'enseignant passe à un autre type de tâches. Donc ce point n'est pas totalement abordé, l'enseignant n'a pas donné le temps aux élèves de finir la représentation du nuage de points dans leur cahiers et de constater la forme de ce dernier, car ceci relève de la technologie qui permet de justifier que la forme du nuage permet d'effectuer un ajustement affine. Nous réitérons encore ici que l'enseignant n'a pas expliqué ce qu'est un ajustement affine aux élèves.

« Un ajustement affine est justifié, c'est un ajustement possible. Ajustement affine justifié, c'est-à-dire qu'on peut déterminer la droite des moindres carrés.

L'affirmation de l'ajustement affine permet de déterminer la droite de régression » (Séance 1, 26:48)

Il aurait dû leur expliquer que : lorsque les points du nuage sont sensiblement alignés, on appelle droite d'ajustement affine une droite qui passe au plus près des points du nuage. Et on dit que cette droite réalise un ajustement affine du nuage de points.

Pour le type de tâches : déterminer une équation de la droite par la méthode des moindres carrés, le savoir nouveau est important et son introduction n'a pas été faite lors de la séance. L'enseignant n'a pas donné les formules qui donnent l'équation de la droite des moindres carrés, il n'a pas non plus expliqué cette expression « moindres carrés » qu'il utilise cependant.

« La question b, c'est de représenter cette droite. Pour la représenter, il faut d'abord déterminer son équation. Pour déterminer l'équation de la droite de régression, vous complétez le tableau par les variables x_i^2 , y_i^2 , $x_i y_i$ et aussi les totaux. De calculer des moyennes \bar{x} , \bar{y} , de calculer la covariance de x , y , de calculer la variance de x . Complétez le tableau. Nous avons ici les totaux. Nous avons les x_i^2 , les y_i^2 , les $x_i y_i$. Complétez le tableau. Allez y ! »

Ici on voit que l'enseignant introduit seulement une partie de la technique, c'est-à-dire la manière d'accomplir ou de réaliser le type de tâches, il n'a aucun discours technologique qui accompagne

cette introduction de la technique. Par exemple le point moyen n'est pas défini, il n'a pas donné la formule pour le coefficient directeur de la droite de régression. Ensuite, l'enseignant passe au tableau et il interagit de nouveau avec les élèves pour remplir le reste du tableau. Après le calcul des moyennes, l'enseignant explique aux élèves qu'ils doivent placer le point moyen G sur le graphique précédent et ce point appartient à la droite de régression.

« P : Avec ce tableau, on aura à calculer \bar{x} . On sait que \bar{x} est égale à la somme des x_i sur n , n est égale à ? c'est-à-dire 376 divisé par 7 est égale à ? 53, ?

Élèves : 53,71

P : \bar{y} serait égale à ? 100 divisé par 7, ce qu'est égale à ?

Élèves : 14,28

P : Et donc le point moyen G, c'est le point de coordonnées \bar{x} , \bar{y} , c'est-à-dire 53,71 et 14,28. »
(Séance 1, 52:01)

Ici on voit la présence d'un discours technologique relatif à la technique qui permet de déterminer les moyennes, c'est-à-dire les coordonnées du point moyen G.

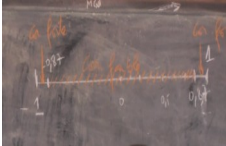
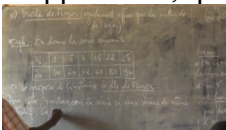
Pour le reste de l'activité l'enseignant explique aux élèves comment ils doivent faire l'estimation.

« Je calcule, est ce que ça va me donner effectivement 17,5 ? Quelle serait sa tension théorique, en utilisant la droite des moindres carrés ? Et après on vous demande de comparer avec la tension réelle. La tension réelle c'est celle qui est de 17,5. Ce qu'on va calculer et on va comparer avec la tension réelle. [On ne voit pas le professeur déterminer la droite d'équation des moindres carrés, qui devrait servir à déterminer la tension artérielle prévue pour un homme de 75 ans]. Voilà les indications par rapport à la suite de l'exercice à la question trois. » (Séance 1, 55 : 42)

Ce qui est censé être introduit dans cette partie, est de déterminer la droite de régression de y en x par la méthode des moindres carrés et de la représenter sur le graphique. Cette droite devrait ensuite être utilisée pour calculer la tension théorique d'un homme âgé de 75 ans. Nous rappelons que la notion de l'estimation comme la méthode des moindres carrés font partie des savoirs nouveaux. . L'enseignant n'a pas présenté la méthode des moindres carrés, et l'idée d'estimation est abordée très rapidement. Il finit la séance en expliquant aux élèves, comment exprimer par exemple y en fonction de x ou x en fonction de y quand on connaît l'équation de la droite. En plus, pendant la séance l'enseignant privilégie plus un travail sur la technique que le discours technologique relatif à la technique. De plus la question trois de l'activité n'est pas résolue.

5.3.2 Description et analyse de la deuxième séance de l'enseignement ordinaire

Pour cette deuxième séance de cours, les élèves ont déjà assisté à une première séance où Aristide a introduit la notion de la série double. La deuxième séance a duré une heure dix-sept minutes comme l'indique le synopsis suivant :

Minutage	Phases	Savoirs en jeu	Modalités de travail	Descriptions
00 :50	Définition de coefficient de corrélation linéaire .	Coefficient de corrélation linéaire.	L'enseignant écrit le cours au tableau et les élèves recopient dans leur cahier.	L'enseignant aborde : la notion du coefficient de corrélation linéaire et des propriétés.
<u>Commentaires</u> : Le professeur est entrain de recopier sa fiche au tableau.				
08 :54	Explication de l'enseignant à la question d'un élève.	coefficient de corrélation, corrélation forte et faible.	Un élève pose la question à l'enseignant. L'enseignant répond à la question de l'élève.	Un élève pose une question sur la remarque faite au tableau par l'enseignant sur la notion du coefficient de corrélation. <i>Si $r \geq 0,87$ ou $r \leq -0,87$, ce n'est pas le contraire, le contraire Monsieur ?</i> L'enseignant utilise le premier exercice d'application pour pour répondre à la question de l'élève.
<u>Commentaires</u> : Le professeur donne des explications sur les intervalles du coefficient de corrélation, en utilisant une droite graduée.				
				
24:01	Présentation de l'énoncé d'un exercice d'application sur la méthode de Mayer	Ajustement affine par la méthode de Mayer : droite de Mayer	Travail collectif : L'enseignant écrit l'exercice au tableau, les élèves recopient l'exercice dans leur cahier.	L'enseignant écrit l'exercice au tableau et il explique.
<u>Commentaires</u> : Le professeur recopie au tableau, l'énoncé de l'exercice qui est sur sa fiche. Et les élèves recopient dans leur cahier.				
26:32	Correction de l'exercice d'application sur la méthode de Mayer	Ajustement affine par la méthode de Mayer. La moyenne pondérée, point moyen.	Travail collectif : Les élèves interrogés répondent aux questions.	L'enseignant partage la série en deux sous séries égales, sous forme de tableau. Il calcule les moyennes et les points moyens de chaque sous séries. Le professeur corrige l'exercice avec la participation des élèves.
<u>Commentaires</u> : On voit ici que, le professeur introduit la méthode de Mayer à partir d'un exercice d'application, qu'il corrige lui même.				
				

34:27	Explication sur la résolution graphique de la droite (G_1G_2)	Une droite passant par deux points.	Travail collectif : Les élèves interrogés répondent aux questions.	L'enseignant donne des explications sur l'équation de la droite : <i>Dire que la droite (G_1G_2) passe par le point G_1, cela signifie que le point $G_1 \in (G_1G_2)$ et aussi le point $G_2 \in (G_1G_2)$.</i>
<u>Commentaires</u> : Le professeur explique oralement aux élèves qu'ils peuvent tracer une droite qui passe par les points G_1 et G_2 .				
43:16	Détermination de l'équation de la droite (G_1G_2)	Équation de la droite	Travail collectif : Les élèves interrogés répondent aux questions. L'enseignant écrit au tableau, avec la participation des élèves.	L'enseignant détermine les coordonnées de G_1 et de G_2 . Il déduit l'équation de la droite $(G_1G_2) : y = \frac{158}{49}x + \frac{1474}{147}$
<u>Commentaires</u> : Ici, le professeur met en jeu le savoir sur la détermination de l'équation de droite à partir deux points.				
51:35	Vérification de l'équation de la droite (G_1G_2)		Travail collectif : Les élèves interrogés répondent aux questions. L'enseignant écrit au tableau, avec la participation des élèves.	L'enseignant vérifie l'équation de la droite $(G_1G_2) : y = \frac{158}{49}x + \frac{1474}{147}$ par certaines valeurs des x_i et y_i et les coordonnées de G_1 et G_2
<u>Commentaires</u> : Le professeur commence par vérifier l'équation de la droite (G_1G_2) avec les coordonnées de certains points du nuage. Ces points ne vérifient pas forcément l'équation de la droite (G_1G_2) . Parce que la droite de cette équation peut ne pas passer par certains points du nuage ou peut passer au plus près de ces points. L'équation de la droite (G_1G_2) est déterminée à partir des points G_1 et G_2 , c'est pourquoi les coordonnées des points G_1 et G_2 vérifient l'équation.				
01:02:02	Détermination de l'équation de la droite (G_1G_2) , en utilisant les vecteurs	Vecteurs colinéaires, déterminant	Travail collectif : Les élèves interrogés répondent aux questions. L'enseignant écrit au tableau,	L'enseignant interroge un élève : <i>Les trois points sont alignés, ça signifie quoi ?</i> Puis il les interroge sur la notion des vecteurs colinéaires et le déterminant. Il utilise la méthode de déterminant pour trouver l'équation de la droite.

			avec la participation des élèves.	
<p><u>Commentaires</u> : Le professeur utilise une autre méthode pour déterminer l'équation de droite (G_1G_2).</p>				

Tableau 8: Synopsis de la deuxième séance

Ci-dessous nous analysons certains passages de cette séance.

Pour introduire la notion du coefficient de corrélation linéaire, l'enseignant demande aux élèves la définition de l'écart type, avant d'écrire au tableau la formule du coefficient de corrélation linéaire. Par la suite l'enseignant écrit au tableau les deux propriétés du coefficient de corrélation suivantes :

- le coefficient de corrélation est compris entre -1 et 1 ;
- Le carré du coefficient de corrélation linéaire est égal au produit des deux coefficients directeurs, $r^2 = a \times a'$

« Soit (D) , c'est la droite qui a pour équation $y = ax + b$, c'est la droite de régression de x en y . Et (D') qui a pour équation $x = a'y + b'$, la droite de régression de x en y . On a : $r^2 = a \times a'$ » (Séance 2, 04:50).

Notons que l'enseignant n'a pas donné les formules pour a et a' .

Ensuite l'enseignant fait des remarques sur la notion abordée mais il ne démontre pas que l'ajustement est de qualité si r^2 proche de 1. Pour répondre à la question posée par un élève sur la valeur du coefficient de corrélation, il trace une droite graduée pour expliquer aux élèves l'intervalle des valeurs du coefficient de corrélation.

L'enseignant a repris les données de l'activité de la première séance pour calculer le coefficient de corrélation. Pour l'exemple, l'enseignant interagit avec les élèves pour le faire. Il pose la question aux élèves, en demandant la valeur de la variance de X et de Y ainsi que leurs écart types. Ici il s'agit de connaissance anciennes que les élèves ont déjà de par leur travail dans les années antérieures sur la statistique à une variable. Ils ont aussi calculé la covariance de X et Y . Durant la première séance nous avons remarqué que l'enseignant n'a pas donné la formule de la covariance mais le travail qu'il avait demandé aux élèves de faire à la maison, consiste aussi à calculer la covariance.

L'enseignant calcule le coefficient de corrélation et il trouve 0,95, il demande aux élèves de l'interpréter. Un élève reprend en disant que « la corrélation est forte » et l'enseignant complète l'interprétation :

Dire que la corrélation linéaire entre x et y est forte, signifie qu'un ajustement affine est possible, un ajustement affine est justifié, c'est-à-dire on peut trouver une droite qui passera le plus près possible des points du nuage (Séance 2, 19:04).

L'enseignant explique ce qu'est un ajustement affine. Nous rappelons que la notion d'ajustement affine n'était pas expliquée pendant la première séance.

L'enseignant indique oralement aux élèves le rôle du coefficient de corrélation linéaire. La notion de coefficient de corrélation linéaire introduite par l'enseignant est accompagnée d'un discours technologique, mais l'expression « une droite qui passera le plus près possible des points du nuage » reste floue. Il est toujours possible, pour un nuage de points donné, de trouver une droite qui minimise la distance (choisie) de l'ensemble des points du nuage avec la droite. Ceci ne dépend pas de la forme du nuage, donc ce discours technologique repose sur une idée intuitive : « si le nuage n'a pas une forme allongée, la droite ne passe pas 'assez' près des points ».

Par la suite, l'enseignant aborde une nouvelle partie du cours,

« Nous allons donner un deuxième exemple. Pour cela nous allons continuer avec la deuxième méthode qui est la droite de Mayer, droite de Mayer. On parle également de l'ajustement affine par la méthode de Mayer » (Séance 2, 23:58).

Il introduit maintenant la méthode de Mayer à travers un exemple, dans un contexte non concret. Il écrit d'abord l'énoncé au tableau et les élèves notent l'énoncé dans leurs cahiers. Par ailleurs, il explique aux élèves comment découper une série, en deux sous-séries en fonction de l'effectif. Nous rappelons que la méthode de Mayer consiste à déterminer la droite passant par deux points moyens du nuage de points. Elle consiste à partager l'ensemble des points (x_i, y_i) en deux sous-ensembles G_1 et G_2 de même effectif (si possible, sinon un des ensembles aura un effectif de plus que l'autre). On peut déterminer l'équation de la droite de Mayer en résolvant le système suivant :

$$\begin{cases} \bar{y}_1 = a \bar{x}_1 + b \\ \bar{y}_2 = a \bar{x}_2 + b \end{cases}$$

Analyse a priori

Dans cette partie nous présentons l'analyse a priori de l'exemple utilisé par l'enseignant pour introduire la méthode de Mayer. Nous explicitons les différents types de tâches et les techniques attendues en se référant au MPR.

Énoncé

On donne la série suivante :

x_i	1	5	7	15	22	25
y_i	10	20	42	60	80	90

On se propose de construire la droite de Mayer. Cette type de tâches se réfère à T4.2.

Dans ce problème les types de tâches sont implicites, il s'agit de :

- Partager l'ensemble des points (x_i, y_i) du tableau en deux sous ensembles G_1 et G_2 de même effectif 3. Nous nommons ce type de tâches « **T_PartagerG1G2** ». C'est un ingrédient de technique de T3.
- calculer les coordonnées des points G_1 et G_2 . Cette tâche relève du type de tâches : T_Cal_CoordG1G2 ;
- Placer les deux points G_1 et G_2 sur le graphique. Cette tâche relève du type de tâches : T_Pla_MoyG1G2 ;
- tracer la droite (G_1G_2) passant par les points moyens G_1 et G_2 , c'est le type de tâches: T_Tra_DroiteG1G2;
- déterminer l'équation de la droite de Mayer (G_1G_2) , ça relève du type de tâches : T_Dét_EquatG1G2.

Pour G_1

x_i	1	5	7
y_i	10	20	42

Pour G_2

x_i	15	22	25
y_i	60	80	90

Le découpage de la série en deux sous ensembles de même effectif est évident. Pour ce type de tâche, si la consigne est comprise il n'y aura pas de difficulté à partager la série en deux sous séries.

Calculons les coordonnées de G_1

$$\bar{x}_1 = \frac{1}{3} \sum_{i=1}^3 x_i = \frac{13}{3} = 4,3 \text{ et } \bar{y}_1 = \frac{1}{3} \sum_{i=1}^3 y_i = \frac{72}{3} = 24 \text{ d'où } G_1(4,3; 24)$$

De la même manière on calcule les coordonnées de G_2

$$\bar{x}_2 = \frac{1}{3} \sum_{i=1}^3 x_i = \frac{62}{3} = 20,6 \text{ et } \bar{y}_2 = \frac{1}{3} \sum_{i=1}^3 y_i = \frac{203}{3} = 76,6 \text{ d'où } G_2(20,6; 76,6)$$

Le second type de tâches consiste à calculer les coordonnées des points moyens G_1 et G_2 . La notion de point moyen est abordé pendant la première séance, donc l'erreur possible des élèves peut être l'erreur de calcul.

En ce qui concerne les représentations graphiques (nuage de points, les points G_1 et G_2 , et la droite (G_1G_2)), il faut d'abord choisir l'échelle par exemple en abscisse 1 cm correspond à 2 et en ordonnée 1 cm correspond à 5. Le choix de l'échelle relève d'une prise d'initiative importante. Les élèves auront la responsabilité du choix de l'échelle en tenant compte de la plus petite et la plus grande valeur de chaque variable x et y pour graduer le repère. Pour les points G_1 et G_2 , les

élèves peuvent utiliser la technique de repérage de points dans le plan pour placer ces deux points et ensuite, ils joignent ces points par une droite.

Le troisième type de tâches consiste à déterminer l'équation de la droite de Mayer, en se servant des coordonnées de G_1 et G_2 .

Déterminons l'équation de la droite de Mayer,

$$\begin{cases} \bar{y}_1 = a \bar{x}_1 + b \\ \bar{y}_2 = a \bar{x}_2 + b \end{cases}$$

En remplaçant les coordonnées de G_1 et G_2 dans le système ci-dessus, on aura :

$$a = \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1} = \frac{76,6 - 24}{20,6 - 4,3} = 3,22 \text{ et } b = \bar{y}_2 - a \times \bar{x}_2 = 76,6 - 3,22 \times 20,6 = 10,26$$

Alors, l'équation de la droite de Mayer est : $y = 3,22x + 10,26$

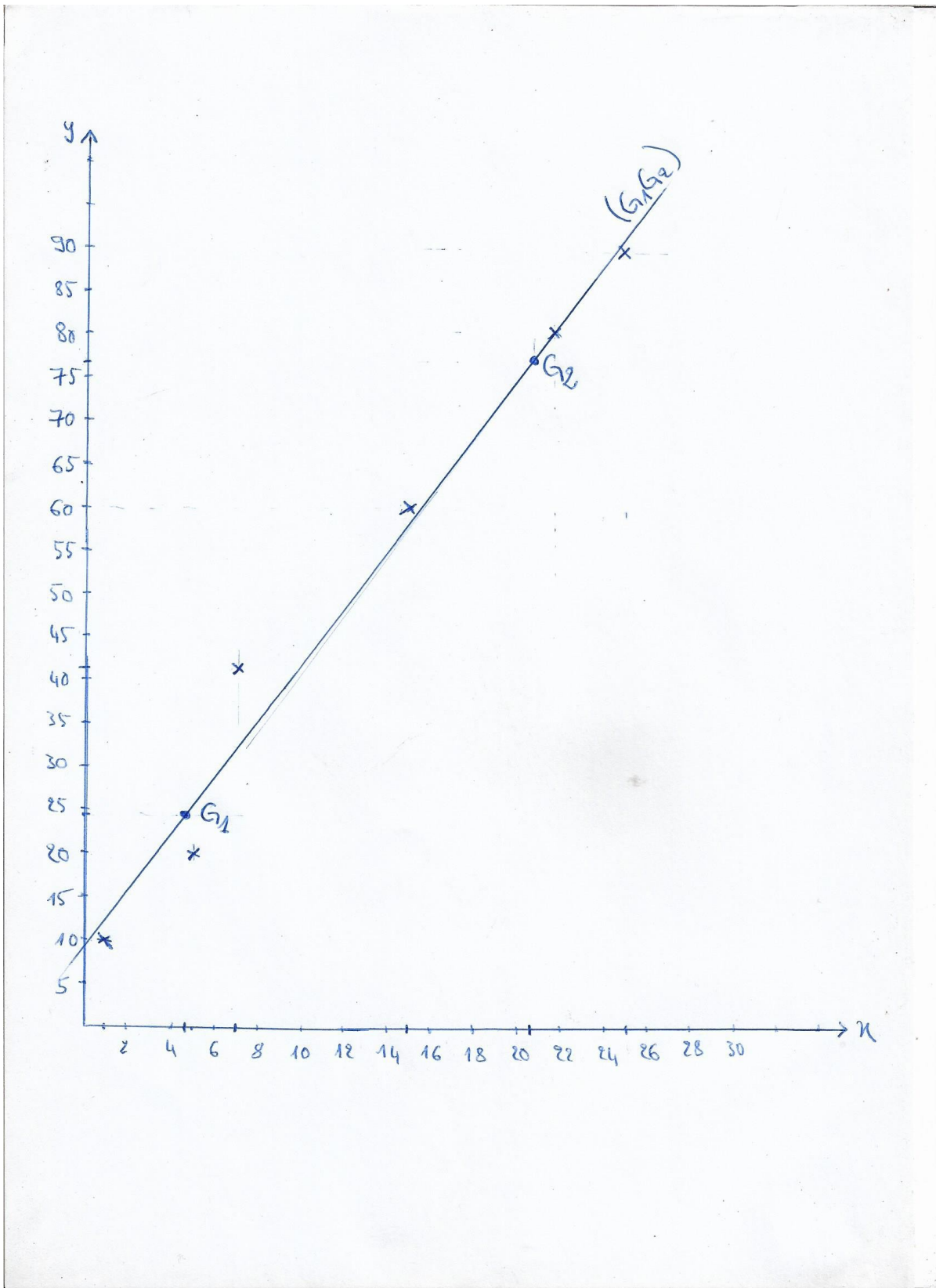


Figure 11: Droite de Mayer

Concernant la description et l'analyse de l'exemple tel qu'il s'est déroulé en classe, un élève désigné par l'enseignant propose de découper la série en deux sous séries et l'enseignant écrit sa

proposition au tableau. Ensuite, l'enseignant demande aux élèves de calculer respectivement les coordonnées de points moyens G_1 et G_2 . En interagissant avec les élèves, ils ont trouvé les coordonnées de ces points moyens. Concernant la représentation graphique, l'enseignant explique aux élèves qu'ils doivent d'abord placer les deux points moyens sur le repère et ensuite tracer une droite (G_1G_2) passant par ces deux points moyens :

« graphiquement, la droite de Mayer, la droite de Mayer c'est la droite (G_1G_2) qui passe par les points G_1 et G_2 » (Séance 2, 34:27), c'est un discours technologique justifiant la réalisation de T_Pla_MoyG1G2.

L'enseignant n'a pas fait la représentation graphique au tableau, il l'a laissée à la charge des élèves qui doivent le faire à la maison. Pour la détermination de l'équation de la droite, le professeur explique aux élèves la technique pour déterminer l'équation de la droite (G_1G_2) et ils interagissent ensemble pour trouver l'équation. Pour se rassurer de la réponse, l'enseignant propose aux élèves de faire la vérification de l'équation de la droite de Mayer trouvée. Pour cela, il demande aux élèves de remplacer dans l'équation, la valeur de x par x_1 pour vérifier qu'on trouve bien y_1 .

Durant cette deuxième séance, l'enseignant commence par rappeler ce qui a été fait pendant la première séance avant d'aborder le contenu de la deuxième séance. Il commence par définir la notion de coefficient de corrélation linéaire, suivi de propriétés et remarques sur le coefficient de corrélation linéaire. A ce moment, les élèves se contentent de prendre note et de poser des questions et l'enseignant leur répond. Par la suite l'enseignant s'appuie sur les données de l'activité de la première séance pour faire un exemple sur la notion de coefficient de corrélation linéaire. Pour cela, il interagit avec les élèves sur les calculs de la covariance, les variances et les écarts types, pour calculer le coefficient de corrélation linéaire. Alors, ce qui est à la responsabilité des élèves dans cet exemple, c'est le calcul des concepts statistiques cités précédemment. Concernant l'interprétation de la valeur du coefficient de corrélation linéaire, l'élève interrogé a des prises initiatives dans l'interprétation du résultat. La réponse proposée par l'élève est juste et l'enseignant a pris le soin d'explicitier sa réponse mais ce n'est pas développé.

En plus de cela, l'enseignant réitère une question qu'il a déjà demandé aux élèves de faire à la maison pendant la première séance. Il demande aux élèves de faire la représentation graphique de la droite de moindres carrés de l'activité précédente. Nous constatons que ce travail a été demandé à la fin de la première séance mais force est de constater que l'enseignant n'a pas pris le temps de vérifier les cahiers des élèves au début de la deuxième séance.

Pour aborder la technique de Mayer qui est un savoir nouveau, l'enseignant l'introduit sous forme d'un exemple.

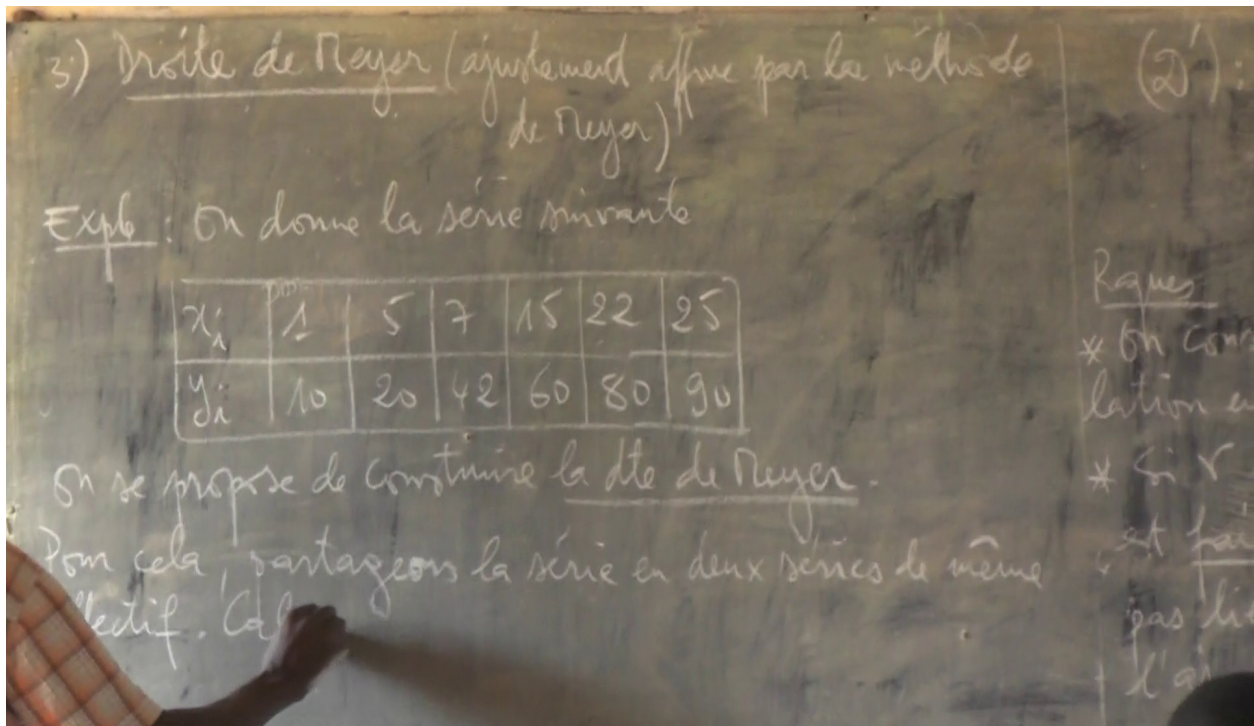


Figure 12: L'enseignant écrit l'énoncé de l'exemple au tableau

L'enseignant explique d'abord aux élèves la technique de Mayer, il commence par la technique qui consiste à découper la série en deux sous-séries sans donner des discours technologiques, c'est-à-dire il n'a pas expliqué pourquoi on a réparti la série en deux sous séries, ensuite il interagit avec les élèves et en posant des questions aux élèves sur chaque étape. Nous constatons que les élèves ont bien réparti la série en deux sous séries et l'enseignant leur demande de calculer les points moyens de ces deux sous séries. Ils ont réussi à réaliser la tâche demandée par l'enseignant, car la notion de point moyen a été déjà abordée pendant la première séance. Donc c'est des savoirs déjà rencontrés par les élèves.

Concernant la représentation graphique, l'enseignant responsabilise les élèves en demandant de le faire eux mêmes à la maison. Pour l'enseignant tout ce qui relève de la représentation graphique est à la charge des élèves. Il leur explique la procédure pour faire la représentation graphique de la droite de Mayer.

« graphiquement, la droite de Mayer, la droite de Mayer c'est la droite (G_1G_2) qui passe par les points G_1 et G_2 . Et il suffit tout simplement de les joindre. Raison pour laquelle, je dis graphiquement. Vous êtes assez grand, pour comprendre ce que je suis entrain de dire. Voilà ce qu'on voit graphiquement, on a construit la droite (G_1G_2) , mais quelle sera alors une équation de la droite (G_1G_2) ? Quelle est son équation ? Quelle sera alors équation ?

Déterminons une équation de la droite (G_1G_2) » (Séance 2, 34:27)

Pour la détermination d'une équation de droite (à partir des coordonnées de deux points de cette droite), l'enseignant donne des explications sur la procédure et ensemble avec les élèves, il détermine l'équation de la droite de Mayer et il leur montre la procédure pour vérifier si l'équation de droite trouvée est juste.

Concernant l'estimation, à la fin l'enseignant explique la procédure pour faire une estimation statistique, il précise aux élèves qu'il va aborder la notion d'estimation pendant les travaux dirigés.

Au regard de l'analyse, nous constatons que des savoirs nouveaux et importants comme : la définition de certains concepts (en particulier l'estimation), la procédure pour déterminer l'équation de la droite de Mayer, et l'introduction de ces savoirs est trop rapide, cela ne permet pas un suivi aisé par les élèves et ne donnant pas l'occasion aux élèves de bien suivre et de poser des questions.

5.4 Analyse du questionnaire aux enseignants

Dans cette section, nous nous proposons de poursuivre notre analyse à travers un questionnaire, articulant des questions centrées : sur l'enseignant, sur son point de vue sur ce qui est attendu des élèves et sur son enseignement de la série statistique à deux variables.

Pour ce faire, nous présentons l'analyse a priori du questionnaire s'ensuit l'analyse des réponses au questionnaire. Le texte complet du questionnaire est disponible en annexe (**annexe B**).

5.4.1 Analyse a priori du questionnaire

Le thème 1 intitulé « Questions centrées sur l'enseignant » comporte sept questions dont le but est de se renseigner sur l'enseignant.

Le thème 2 « Questions centrées sur la pratique de l'enseignant, exemple d'un exercice » a trois questions. Nous proposons un exercice à ces professeurs. Cet exercice, nommé « Préviation du ménage 1 », provient d'une expérimentation que nous avons menée en classe auparavant. Nous le rappelons ici, il est analysé en détail dans la partie 6.1 au chapitre suivant, nous renvoyons le lecteur à cette analyse.

Exercice « Préviation du ménage 1 »

Le tableau ci-dessous indique le montant des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3 568	4 561	5 502	6 693	8 241	9 120

Le ménage pense qu'en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette affirmation ?

A propos de cet problème, nous posons aux enseignants la question nommée question 8: « *Est-ce que vous proposeriez cet exercice à vos élèves ?* » nous attendons de la part des enseignants plusieurs types de réponses possibles.

- Ceux qui diront oui, nous pouvons proposer cet exercice à nos élèves.

Les élèves savent qu'il s'agit d'estimer la valeur dépensée en 2019 et de la comparer à 10000. Pour cela, ils peuvent justifier qu'on peut faire un ajustement affine et déterminer l'équation d'une droite d'ajustement, qui est de la forme $y=ax+b$, et après ils peuvent utiliser cette équation pour estimer la dépense en l'an 2019.

- Les enseignants qui diront oui, mais en le modifiant.

Ils peuvent notamment penser que l'exercice n'est pas assez guidé. Dans ce cas ils peuvent poser la question de la manière suivante :

1. Représenter les nuages de points de cette série double. Cette tâche se réfère au type de tâches T_Rep_Nuage ;
2. Déterminer l'équation de cette droite sous la forme $y=ax+b$ (en utilisant la méthode des moindres carrés, T4.2 ou la méthode de Mayer, T4.3). Cette tâche se réfère au type de tâches T_Dét_EquatMCyx ;
3. En utilisant l'équation de la droite trouvée à la question précédente, estimer la dépense du ménage en 2019. Cette tâche se réfère au type de tâches T5. Puis expliquez votre résultat. Cette tâche se réfère au type de tâches T6.

Pour ces enseignants, ils préfèrent expliciter ces sous-questions, pour guider le travail des élèves, cela permettrait aux élèves de s'engager dans la résolution de l'exercice.

- Certains enseignants diront non, ils ne peuvent pas proposer un tel exercice à leur élèves.

Pour eux la question posée est trop implicite, les élèves auront du mal à savoir de quoi il s'agit. Ils auront du mal à commencer l'exercice.

Nous proposons ensuite à ces enseignants deux productions d'élèves. La première production provient de la production d'un groupe d'élèves pendant la première expérimentation. La deuxième production est une production fictive, que nous avons conçue pour l'intégrer dans l'exercice nommé « Prévission du ménage 2 », utilisé lors de notre deuxième expérimentation.

Ci-dessous nous présentons les productions des élèves utilisées dans le questionnaire.

Pour chacune des productions d'élève :

9) Analyser la réussite et les erreurs éventuelles.

10) En cas d'erreurs identifiées, comment pourriez-vous aider ces élèves ?

$$y = ax + b \Leftrightarrow y = 10.000 x + b$$

D'après la formule.

$$a = \frac{\text{Cov}(x, y)}{V_x}$$

$$\Leftrightarrow a = \frac{3338,6}{2,91}$$

$$b = \bar{y} - a\bar{x}$$

$$b = 6280,83 - 1145,22 \times 3,5$$

$$\bar{x} = 1145,22$$

$$b = 2272,5$$

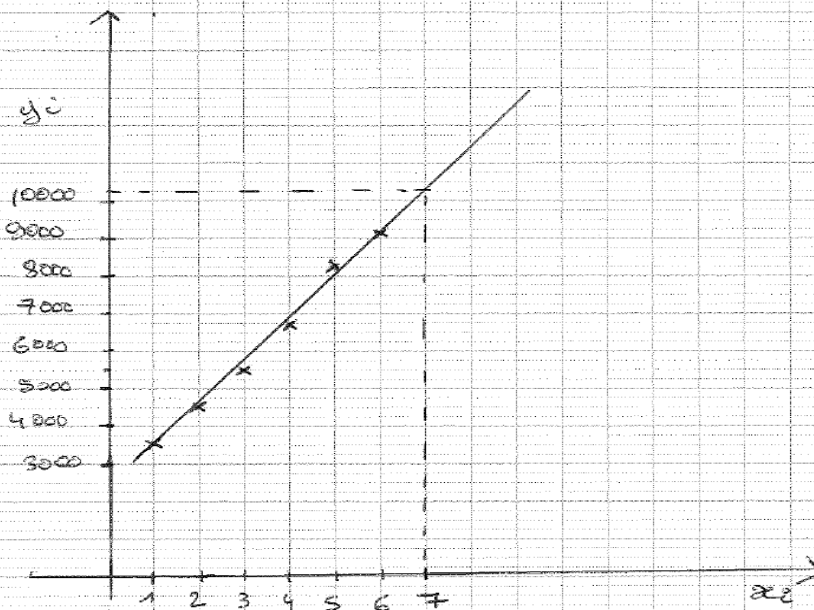
$$y = 1145,22x + 2272,5$$

$$\hat{y} = 1145,22 \times 10000 + 2272,5$$

$$\hat{y} = 11454472,5$$

le ménage pense que la prévision en 2019 est fautive car le ménage a dépensé plus

Élève A



La prévision du ménage est mauvaise. Nous avons tracé une droite d'ajustement et nous voyons que l'estimation des dépenses de 2019 dépasse 10000.

Élève B

Pour la question 9 « Analyser la réussite et les erreurs éventuelles », nous nous attendons aux réponses suivantes.

L'élève A a bien pensé à faire un ajustement affine, c'est une réussite. Il n'a pas justifié qu'on pouvait faire un ajustement affine dans cette situation, on peut le lui reprocher.

Il a utilisé la méthode des moindres carrés pour trouver l'équation de la droite $y = 1145,22x + 2272$. Cette équation lui servira à estimer la dépense du ménage en l'an 2019. Nous pouvons considérer la modélisation faite par l'élève A, comme une réussite. Il semble connaître la méthode des moindres carrés, en particulier comme trouver l'équation de la droite. Puis de déduire l'estimation en l'an 2019 (c'est-à-dire trouver le rang de l'année 2019, qui est $x=7$ et de le remplacer dans l'équation de la droite trouvée). Ensuite il interprète bien son résultat.

Concernant les erreurs faites par l'élève A, nous attendons que les enseignants formulent les observations suivantes. Son équation de la droite est erronée, il a fait une erreur de calcul lorsqu'il calculait le coefficient directeur a . Cette erreur de calcul joue aussi sur le résultat de l'équation de la droite. En plus de l'erreur de calcul, l'élève A, s'est trompé sur la valeur du rang de l'année 2019. En prenant 10000, comme la valeur du rang de l'année qui lui permettra de faire l'estimation.

La première erreur de l'élève A identifiée, est une erreur de calcul. Cette erreur se situe au niveau du calcul de la covariance. Pour l'aider, nous lui proposons de refaire le calcul de la covariance. Pour la deuxième erreur, il s'est trompé sur la valeur du rang de l'année 2019. On peut lui conseiller de réfléchir à l'ordre de grandeur du résultat, pour qu'il se rende compte qu'il a fait une erreur. Pour cela, il doit compléter le tableau jusqu'à l'année 2019. Et cela pourrait l'aider à identifier facilement le rang de l'année.

Pour l'élève B, certains enseignants vont trouver que sa production est correcte.

Certains enseignants diront qu'il a bien représenté les données de l'énoncé par un nuage de points. Et il a tracé une droite qui approche bien le nuage de points, donc il peut l'utiliser pour faire l'estimation. En plus il a bien vu que le rang de l'année est 7, et il a bien mis des pointillés pour indiquer où se trouve la valeur correspondant à 7. Il fait une estimation graphiquement, et ensuite il propose une interprétation en cohérence avec son estimation. Logiquement, ces enseignants ne proposeront pas d'aide à l'élève B.

Mais d'autres enseignants diront que la méthode n'est pas correcte, parce qu'elle n'est pas assez précise. Ces enseignants refuseront l'emploi d'une méthode graphique. Ils proposeront comme aide à l'élève un conseil du type « utiliser la méthode des moindres carrés ».

Le thème 3 « Questions centrées sur la pratique de l'enseignant pour la statistique à deux variables en général » du questionnaire est constitué des deux questions suivantes :

11) A propos du thème statistiques à deux variables, est-ce vous diriez que :

a) L'important est que l'élève apprenne des méthodes d'ajustement : moindre carrés, Méthode de Mayer

b) L'important est que l'élève sache interpréter les résultats de ses calculs

c) L'important est que l'élève parvienne à proposer un modèle adéquat pour faire une estimation

12) A propos des difficultés des élèves avec ce thème, est-ce que vous diriez que :

a) La principale difficulté est que les élèves ne connaissent pas bien les méthodes

b) La principale difficulté est que les élèves n'ont pas de calculatrice pour la plupart

c) Les élèves n'ont pas particulièrement de difficultés avec ce thème.

Concernant la question 11, nous attendons à avoir les différents types de réponses suivants :

- l'important est que l'élève apprenne des méthodes d'ajustement. Ces enseignants vont mettre en avant l'utilité de ces méthodes.
- Ceux qui diront que l'important n'est pas seulement d'apprendre des méthodes d'ajustement, ils peuvent dire que les méthodes d'ajustement ne sont que des éléments statistiques parmi d'autres qui servent à faire des prévisions ou à résoudre un problème statistique.
- Certains diront que l'important est que l'élève sache interpréter les résultats de ses calculs. Pour eux l'interprétation des résultats est une étape délicate nécessitant une bonne connaissance du phénomène étudié et des méthodes statistiques utilisées. Ils pensent que l'interprétation permet aux élèves de développer leur capacité et leur habilité à analyser les données.
- Les enseignants qui ne sont pas d'avis, pensent qu'il faut d'abord trouver les résultats avant de les interpréter. Pour ce faire, la première priorité est de trouver d'abord le résultat.
- Ceux qui pensent que ce qui est important est que l'élève parvienne à proposer un modèle adéquat pour faire l'estimation. Ces enseignants peuvent accepter différents modèles, laissant plus de place aux initiatives des élèves.

Pour la question 12, les enseignants qui pensent que :

- la principale difficulté est que les élèves ne connaissent pas bien les méthodes. Si l'élève n'arrive pas à mobiliser les concepts statistiques pendant l'application des méthodes, alors nous pourrions voir ça comme une difficulté.
- D'autres pensent que la connaissance des méthodes n'est pas la difficulté principale, du fait que ces méthodes sont apprises et utilisées en classe. Donc cela ne pourrait pas être la principale difficulté.
- D'autres également pensent que la principale difficulté est que les élèves n'ont pas de calculatrice pour la plupart. Pour faire des calculs élémentaires sur les données statistiques, il est nécessaire d'utiliser la calculatrice, donc sans calculatrice c'est difficile de faire un exercice de statistique.
- Enfin d'autres pensent au contraire que faire la statistique sans calculatrice n'est pas la principale difficulté, car sans la calculatrice on peut trouver avant tout un modèle adéquat pour faire la statistique.

- Il y a des enseignants qui pensent que, les élèves n'ont pas particulièrement de difficultés avec la statistique à deux variables. Ils vont probablement s'appuyer sur la connaissance des élèves sur les concepts statistiques et les méthodes d'ajustement, pour se justifier.

5.4.2 Analyse des réponses aux questionnaires

Les questionnaires ont été soumis début mai 2020. En tout, nous avons distribué dix-sept questionnaires et dix enseignants nous ont restitué leur questionnaire. Ces enseignants sont tous des hommes, il y a très peu de collègues femmes qui interviennent en mathématiques dans les lycées en Centrafrique. Nous avons laissé aux enseignants environ deux semaines pour répondre aux questions. Nous désignons les professeurs par P_1, P_2, \dots, P_{10} . Nous précisons que P_{10} est l'enseignant Aristide dont nous avons observé les cours.

Caractéristiques professionnelles des enseignants et ressources utilisées

	Nombre d'années d'expérience	A déjà enseigné en Terminale SE	Utilise le manuel CIAM	Autres ressources utilisées
P_1	10 ans	Oui	Oui	Bordas, Dimathème, Anabac, Prépacab, Déclic
P_2	17 ans	Oui	Oui	Déclic, Terracher
P_3	7 ans	Oui	Oui	Dimathème, Bordas, Triangle, Hachette Education, Web:Mathovore
P_4	31 ans	Oui	Oui	Dimathème, Déclic
P_5	5 ans	Oui	Oui	Uniquement CIAM
P_6	9 ans	Oui	Oui	Déclic, Hatier, Durandeu
P_7	27 ans	Oui	Non	Terracher, Déclic, Fractal
P_8	8 ans	Non	Oui	Uniquement CIAM
P_9	6 ans	Non	Oui	Dimathème, Hachette, Repère, Internet Triangle, Durandeu, Sesamath
P_{10} (Aristide)	25 ans	Oui	Oui	Déclic, Terracher, Fractal, Transmath, Annal Bac
		8/10 ont déjà enseigné en terminale SE	9/10 utilisent le manuel CIAM	6/10 utilisent Déclic, et 4/10 utilisent Dimathème

Tableau 9: Réponses des enseignants à la première partie du questionnaire

Les enseignants qui ont répondu au questionnaire ont des profils variés en termes d'expérience. Cependant ces enseignants ont tous exercé plus de quatre ans le métier de l'enseignement. Le plus jeune (P_5) en terme d'expérience à cinq ans d'ancienneté dans le métier et le plus ancien (P_4) a trente un ans d'ancienneté. Et la plupart de ces enseignants (huit) ont eu à enseigner déjà la classe de terminale SE.

En ce qui concerne les ressources utilisées, neuf enseignants sur dix se réfèrent au manuel CIAM pour préparer leur enseignement, dont deux utilisent CIAM comme l'unique manuel pour leur enseignement.

Les ressources sur lesquelles s'appuient ces enseignants en dehors du manuel CIAM sont des manuels français de différents éditeurs (éditions Didier, Bordas, Hachette Education, Hatier, Nathan, Magnard). Les manuels les plus utilisés par ces enseignants sont Déclic (Hachette Education) et Dimathème (Didier) ; ce sont des versions correspondant à d'anciens programmes français. En effet en France les séries statistiques doubles ne sont plus enseignées en Terminale depuis 2011.

Il nous semble indispensable de signaler un faible taux d'utilisation d'internet par ces enseignants pour préparer leur enseignement. Uniquement deux enseignants évoquent des ressources en ligne. Cela est dû à la non maîtrise des outils informatiques et aussi à la difficulté d'accès à l'internet.

Deux enseignants sur les dix n'ont pas encore enseigné une classe de terminale SE. Par contre les huit autres enseignants ont déjà enseigné la classe de terminale SE, et ont eu à enseigner la série statistique à deux variables en terminale SE. Seul un enseignant, P7, a déjà enseigné en terminale SE, mais n'utilise pas le manuel CIAM.

L'avis personnel des enseignants sur un exercice proposé aux élèves

Dans cette partie de notre analyse, nous allons donner à voir le point de vue des enseignants et leur appréciation, respectivement sur un exercice et sur la production de deux élèves (questions n°8, 9 et 10).

Nous analysons les réponses des huit enseignants qui sont déjà intervenus en terminale SE.

Un professeur (P₆) sur les huit a répondu qu'il proposerait l'exercice tel qu'il est, à ces élèves : « Cet exercice pousse l'élève à bien réfléchir, afin de choisir une méthode appropriée pour le résoudre ». Pour l'enseignant (P₆) c'est un exercice qui fait appel à la responsabilité de l'élève. C'est un exercice qui nécessite de la réflexion, l'élève doit développer des stratégies, il doit reformuler la situation afin de donner du sens à l'exercice. Donc l'élève a la responsabilité de structurer la situation qui lui est proposée. Cela nécessite une prise d'initiative de la part de l'élève, afin de choisir la méthode appropriée pour la résolution. Le point de vue de cet enseignant montre qu'il pense qu'il est important de faire avancer les connaissances des élèves sur le processus de la résolution des problèmes par le choix du modèle adéquat.

Par contre, sept enseignants (P₁ ; P₂ ; P₃ ; P₄ ; P₅ ; P₇ ; P₁₀) sur les huit ont répondu qu'ils modifieraient l'exercice avant de le proposer à leurs élèves. Nous donnons ci-dessous des exemples de leurs réponses.

« P₁ : Il est beaucoup préférable de reformuler le travail à faire, ce qui rendra plus clair l'exercice. Car si un élève n'arrive pas à comprendre le sens de la question comme l'indiqué ci-haut, alors il sera bloqué. Il est donc préférable d'éclater la question tout en orientant le travail. Par exemple :

1. Construire le nuage de points associé à cette série. Que constatez-vous ?
2. Déterminer et construire le point moyen G de ce nuage.
3. Déterminer une équation de la droite de régression de y en x par la méthode des moindres carrés.

4. En utilisant cette droite, estimer la dépense du ménage en 2019.
5. En déduire que le ménage ne pourrait pas dépenser moins de 10000 euros.

P₄: En statistique, les étapes importantes sont :

- 1-Dessiner le nuage des points,
- 2-Compléter le tableau ci-dessus par des informations données pendant le cours,
- 3-Déterminer l'équation de la droite de régression de y en x , puis la représenter graphiquement sur le nuage de points,
- 4-Conclure soit analytiquement, soit géométriquement.

P₁₀: 1- On pouvait supposer que les variables étaient en corrélation linéaire ;

- 2- On pouvait demander la représentation du nuage de points, cela permettrait une visibilité du lien entre les deux variables ;
- 3- On pouvait préciser la méthode d'ajustement linéaire (Mayer ou moindres carrés) ;
- 4- On pouvait demander une précision par rapport au nombre de décimales. ».

Selon les enseignants (P₁ ; P₂ ; P₄ ; P₇ ; P₁₀), pour que l'élève comprenne le sens de l'exercice, il est préférable de reformuler la question en plusieurs sous-questions bien articulées. Ces enseignants trouvent que l'exercice n'est pas assez guidé. Pour ces enseignants, en modifiant la question l'élève sera capable de résoudre lui-même l'exercice. Ces enseignants ont tous plus de dix ans d'ancienneté, si on analyse bien leur point de vue sur la question, ils préfèrent procéder, en utilisant les structurations des exercices rencontrées habituellement dans les manuels. Donc ils sont restés dans la même logique des structurations des manuels qu'ils utilisaient pendant leur enseignement. Ils ajoutent des questions intermédiaires pour guider l'élève, par crainte que celui-ci soit en difficulté. Avec ce choix ils réduisent les possibilités de prise d'initiative des élèves.

Quant aux enseignants P₃ et P₅, ils proposent une modification sur l'unité. De changer l'euro en franc CFA, la monnaie locale. C'est une modification de l'habillage et non sur le contenu mathématique de l'exercice. Ils mettent l'accent sur le contexte culturel qui peut donner du sens pour les élèves.

Question 9 et 10, à propos des copies d'élèves proposées.

Six enseignants (P₁, P₂, P₃, P₄, P₆, P₁₀) sur les huit qui ont déjà enseigné en terminale SE ont répondu à la neuvième et la dixième questions. Analysons leur point de vue.

Concernant la production de l'élève A, trois enseignants (P₁, P₃, P₆) sur les six pensent que l'élève A comprend l'exercice et propose un modèle adéquat. Donc l'élève A sait ce qu'on lui demande comme tâche à faire, mais il lui manque de la rigueur dans ses calculs. L'enseignant P₁ propose que l'élève utilise un tableau statistique pour déterminer les coefficients a et b de l'équation de la droite. Pour l'enseignant P₁₀, L'élève A a réussi à proposer un modèle adéquat, pour trouver l'équation de la droite de régression mais il a fait des erreurs de calculs. Pour les éventuelles erreurs de l'élève A, les enseignants (P₁, P₄, P₆) ont relevé des erreurs dues à un manque d'efficacité dans la pratique des calculs. Il y a aussi des erreurs des formules dues à la non maîtrise du cours.

Pour pallier ces erreurs, les enseignants (P₁, P₂, P₄) ont proposé l'utilisation d'un tableau statistique pour déterminer l'équation de la droite de régression. Pour eux, si l'élève A calcule les données dans un tableau statistique, il pourra éviter toutes ces erreurs. Nous rappelons que la technique

d'utilisation d'un tableau statistique que nous avons nommé T_Etab_TabS est la plus utilisée dans l'institution Terminale SE en Centrafrique.

Concernant l'estimation, deux enseignants (P_1 , P_6) ont souligné l'erreur commise par l'élève A sur la valeur du rang de l'année, permettant de faire l'estimation.

Quant à la production de l'élève B, cinq enseignants (P_1 , P_3 , P_4 , P_6 , P_{10}) ont analysé la production de l'élève B avec beaucoup de réserve. Pour eux, la méthode graphique manque de précision et elle est hasardeuse. Ces enseignants priorisent les deux méthodes les plus utilisées en classe T4.2 et T4.3. Donc pour eux, la méthode graphique n'est pas appropriée, car elle manque de précision analytique : le point moyen, l'équation de la droite. Pour l'enseignant P_3 l'élève B a bien représenté le nuage des points, la procédure utilisée ne donne pas une bonne précision. Pour lui, cet élève pouvait justifier cet ajustement affine en calculant la corrélation linéaire. L'enseignant P_{10} analyse la méthode graphique utilisée par l'élève B comme une réussite, mais pour lui c'est une méthode qui n'est pas précise, c'est une conjecture qui devra être démontrée par la suite.

Pour aider l'élève B, les enseignants (P_1 , P_2 , P_4) recommandent de déterminer d'abord l'équation de la droite de régression avant de faire l'estimation, c'est-à-dire utiliser les différentes étapes des méthodes enseignées en classe.

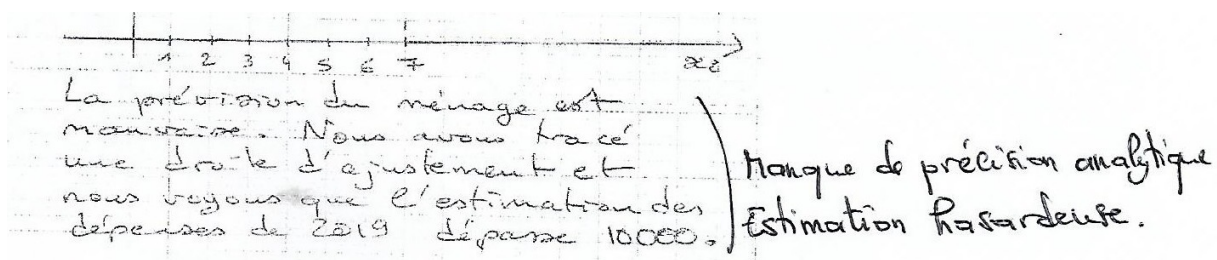


Figure 13: Extrait de la réponse de P_1

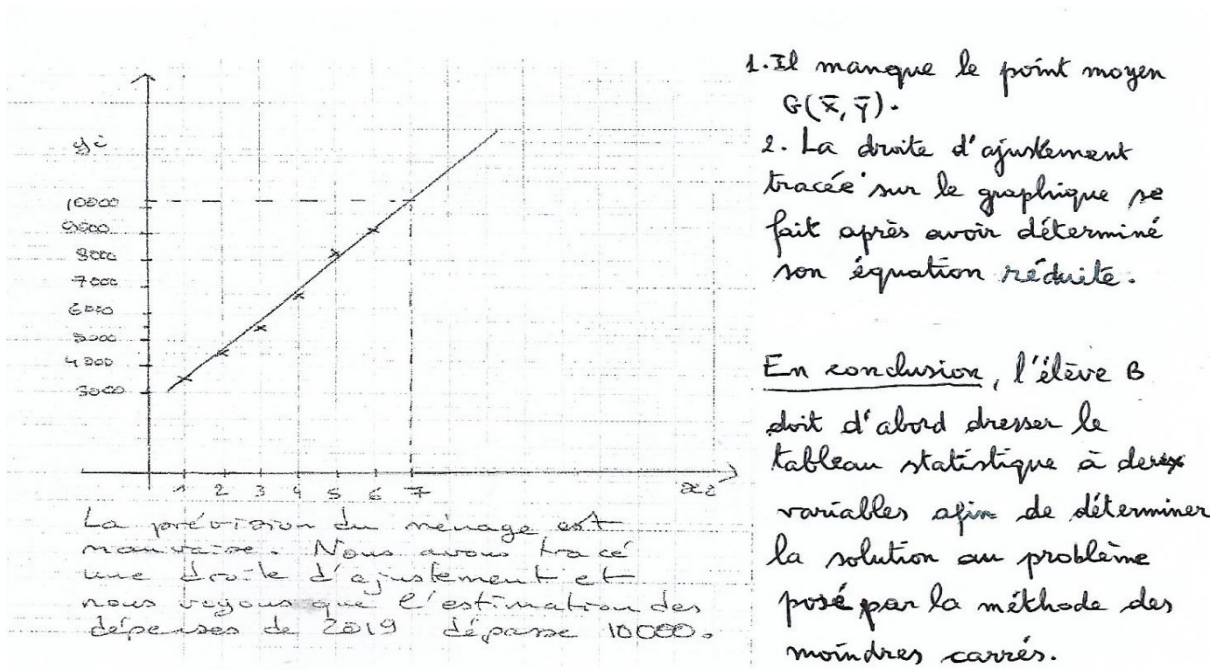


Figure 14: Extrait de la réponse de P4

Donc les enseignants (P₁, P₃, P₄, P₆, P₁₀) considèrent la méthode graphique comme une méthode qui manque de précision, ils privilégient la méthode de Mayer et celle des moindres carrés. Pour eux, la méthode graphique ne donne pas une bonne estimation. Il faut calculer l'équation de la droite pour faire l'estimation.

La pratique de l'enseignant centrée sur les séries statistiques à deux variables.

La onzième question concerne l'importance de quelques aspects en statistique. Pour cela nous avons trois items.

Le premier item porte sur l'avis des enseignants sur l'importance de l'apprentissage des méthodes d'ajustement. Cinq enseignants (P₁, P₂, P₃, P₄, P₆) donnent l'avis favorable. Pour eux, l'important est que l'élève apprenne des méthodes d'ajustement, c'est en maîtrisant les méthodes d'ajustement que l'élève peut parvenir à faire des estimations et à mieux interpréter ses résultats. Il est indispensable que l'élève apprenne ces méthodes. Car cela est prévu dans le programme officiel.

Thème 3 : Questions centrées sur la pratique de l'enseignant pour les statistiques à deux variables en général.

11) A propos du thème statistiques à deux variables, est-ce vous diriez que :

a) L'important est que l'élève apprenne des méthodes d'ajustement : moindre carrés, Méthode de Mayer :

Oui, car c'est en maîtrisant ces méthodes qu'il parviendra à faire de bonnes interprétations et une bonne estimation

b) L'important est que l'élève...

Figure 15: Extrait de la réponse de P3

Les six enseignants ($P_1, P_2, P_4, P_5, P_6, P_7$) disent que l'important est que l'élève sache interpréter les résultats de ses calculs, car c'est une suite logique de la modélisation.

Deux enseignants (P_1, P_6) disent que, proposer un modèle adéquat en statistique est le plus difficile à faire pour les élèves. Donc c'est important que les enseignants les accompagnent dans cette démarche, en donnant des indications sur les méthodes à utiliser. Par contre, selon l'enseignant (P_4), ce qui est important c'est que l'élève doit connaître les concepts et les méthodes apprises sur la statistique à deux variables.

De notre point de vue dans le cadre d'apprentissage de la statistique, les élèves ont souvent des difficultés à appréhender l'effet réel de la statistique, c'est-à-dire, à construire un modèle adéquat par une approche statistique qui permet d'interpréter de manière pertinente, la problématique posée.

La douzième question porte sur les difficultés que les élèves peuvent avoir en statistique.

Six enseignants ($P_1, P_2, P_3, P_4, P_5, P_6$) soutiennent l'idée que la principale difficulté des élèves provient du fait qu'ils ne connaissent pas bien les méthodes. Pour eux, cette difficulté provient d'une part de la mauvaise manière d'enseigner de certains enseignants. D'autre part, c'est parce que les élèves ne travaillent pas assez, c'est ce qui conduit à la non maîtrise de ces méthodes et qui cause des erreurs le plus souvent.

Les enseignants (P_1, P_2, P_4, P_6, P_7) pensent que la difficulté des élèves est qu'ils n'ont pas de calculatrice. Parfois on leur prête des calculatrices, parce qu'ils n'en n'ont pas, ils les utilisent occasionnellement, c'est évident qu'ils ont des difficultés à bien les manipuler. D'après eux, la véritable difficulté est cette non maîtrise de l'utilisation de la calculatrice.

Pour les enseignants (P_1, P_6, P_7) ils pensent que les élèves n'ont pas de difficultés particulières avec ce thème. Pour eux, si le cours est bien enseigné, alors les élèves n'auront pas des difficultés avec ce thème, c'est un thème qui semble assez simple à comprendre. Par contre, l'enseignant (P_4) pense qu'il y a toujours de difficultés, les élèves n'ont pas les mêmes niveaux de compréhension. Il y a ceux qui ont la capacité intellectuelle de comprendre plus vite que les autres.

Concernant les commentaires des enseignants sur l'enseignement de la statistique à deux variables.

L'enseignant (P_4) a mis l'accent sur les conditions de travail des élèves. Pour lui, c'est primordial que les élèves doivent avoir des calculatrices et les outils géométriques, permettant de faire des calculs et des représentations graphiques. Il a aussi mentionné que les élèves ont des difficultés à bien choisir les échelles.

Certains enseignants (P_1, P_6) recommandent aux enseignants de se rassurer des prérequis en statistique, c'est-à-dire qu'ils ont bien fait le cours de la série statistique à une variable en classe de première. Et les enseignants doivent accompagner les élèves pour mieux comprendre les méthodes d'ajustement qui est l'un des objectifs pour ce thème.

L'enseignant P_{10} , affirme que les difficultés mentionnées sont toutes importantes, il a cité d'autres difficultés des élèves en statistique :

- Non maîtrise des formules ;
- Compréhension de la langue française ;
- Absence de calculatrice (qui devrait faciliter la rapidité dans la résolution de l'exercice) ;

- L'une des difficultés des élèves avec ce thème est que ceux-ci évitent les calculs avec de grands nombres.

Pour lui, l'objectif de la statistique à deux variables, c'est la maîtrise des méthodes, des formules, l'interprétation des résultats dans le but de faire des prévisions.

En résumé, nous avons analysé un questionnaire qui était adressé aux enseignants de mathématiques de centrafricain qui interviennent au lycée. L'analyse du questionnaire nous a permis de recueillir leurs points de vue sur leurs expériences et leurs pratiques de l'enseignement de la statistique à deux variables. De façon générale, ils préfèrent proposer aux élèves un exercice articulé en plusieurs types de tâches, en utilisant la structuration usuellement présente dans les exercices.

Par ailleurs, ils privilégient les techniques des moindres carrés et Mayer plus que la technique graphique, car selon eux cette dernière manque de précision.

Ces réponses confirment selon nous le peu d'initiatives qui sera laissé aux élèves par ces professeurs. Il ne proposeront sans doute pas aux élèves de choisir une méthode pour réaliser une estimation. Ils n'ont pas non plus évoqué dans leurs réponses des compétences de modélisation. Ils attendent que les élèves sachent mettre en œuvre les méthodes apprises en cours, surtout la méthode de moindres carrés.

Il nous semble judicieux d'apporter un complément d'informations à nos données, afin d'explicitier les points de vue des enseignants établis à partir des questionnaires qu'ils avaient remplis. Pour cela, nous allons interviewer quelques enseignants qui ont rempli le questionnaire.

5.5 Analyse des entretiens avec les enseignants

Durant les entretiens que nous avons menés auprès des enseignants, nous nous sommes intéressé dans notre guide d'entretien plus particulièrement à l'enseignement et l'apprentissage de la série statistique à deux variables. Notre guide d'entretien comprend trois parties : la première partie fait un retour sur le questionnaire rempli par l'enseignant (§ 5.4), elle permet d'approfondir certaines questions abordées dans le questionnaire. La seconde partie est relative à la pratique d'enseignement de la statistique à deux variables. Et la troisième partie concerne les initiatives des élèves en statistique à deux variables. Nous avons interviewé trois enseignants, nommés P₂, P₄, au mois de Juillet 2020 et Aristide (P₁₀) au mois d'Août 2020.

Un retour sur le questionnaire rempli par l'enseignant

Concernant la première partie, la première question porte sur le choix d'utilisation d'autres manuels en dehors du manuel CIAM. Pour l'enseignant nommé P₂, c'est nécessaire d'utiliser d'autres manuels en dehors du manuel officiel CIAM. Selon lui, si les élèves sont en possession du manuel que l'enseignant utilise, ils ne vont pas être attentifs pour suivre les explications en classe. Pour eux (les élèves), le contenu du manuel est suffisant et tout ce que dira l'enseignant se trouve déjà dans leur manuel. D'où la nécessité d'utiliser d'autres manuels pour avoir des informations complémentaires. Pour l'enseignant P₄, il faut plusieurs manuels pour mieux préparer un cours et cet enseignant affirme qu'il ne fait jamais recours à l'internet pour préparer ses fiches de cours, faute de moyen financier pour payer un abonnement. Nous signalons que la connexion internet reste encore un luxe en Centrafrique, peu de gens ont accès à l'internet. Concernant l'enseignant P₁₀,

nommé Aristide, pour lui, c'est important d'utiliser plusieurs manuels à la fois, car ces manuels présentent différentes approches. Selon Aristide, si l'enseignant rencontre des difficultés sur une notion pendant la confection des fiches de cours, alors il peut utiliser d'autres manuels qui proposent d'autres approches.

La deuxième question concerne l'avis des enseignants sur le manuel CIAM utilisé en Centrafrique. Pour les enseignants P_2 et P_4 , le manuel CIAM est un bon manuel, il cadre bien avec le programme de mathématiques centrafricain. Du point de vue contexte, il cadre bien avec la réalité du pays. Pour Aristide, le manuel CIAM a été confectionné par les enseignants de l'Afrique francophone à base d'un programme prédéfini dans la zone francophone. « Il a certes ses atouts et certainement ses insuffisances, d'où la nécessité effectivement, comme je le disais tantôt d'utiliser plusieurs manuels à la fois » (entretien avec Aristide, 20 Août 2020).

La troisième question concerne l'avis des enseignants sur leurs réponses à la question 9 et 10 du questionnaire.

Concernant la production de l'élève A, d'après l'enseignant P_2 la procédure utilisée par l'élève A manque de clarté. P_2 suggère que l'élève A commence par faire un tableau statistique puis de déduire l'équation de la droite de régression. Nous rappelons que la procédure proposée par P_2 , est la procédure la plus travaillée dans la classe de terminale en Centrafrique. Pour déterminer les coefficients de l'équation de la droite de régression par la méthode des moindres carrés, il faut faire un tableau statistique qui permettra de calculer les moyennes, la variance, l'écart type et la covariance puis déduire les valeurs des coefficients de l'équation de la droite de régression. L'enseignant P_4 aussi répond dans ce sens que l'élève A doit commencer par faire un tableau statistique. Pour l'enseignant Aristide, l'élève A a bien déterminé l'équation de la droite par la méthode des moindres carrés, mais la réponse trouvée est fautive car il a donné une justification erronée.

Concernant l'élève B, pour l'enseignant P_2 , la méthode graphique utilisée par l'élève B est une bonne méthode, il souligne que l'élève n'a pas représenté le point moyen G ainsi que le nuage de points. Il nous semble que sur la représentation de cet élève, on voit bien qu'il a représenté le nuage de points et la droite de régression. Il signale le fait que la droite de régression doit passer forcément par le point moyen G. Pour P_4 , la résolution doit se faire à l'aide d'un tableau statistique pour aboutir à une équation de la droite et non par la méthode graphique. Selon Aristide, l'élève B a convenablement fait la représentation graphique, il a bien représenté le nuage de point certes, mais la méthode graphique est une conjecture, il doit expliciter sa réponse par une méthode analytique.

Selon Aristide, « il a convenablement construit le nuage des points par la lecture graphique, il a convenablement répondu à la question posée. Sauf que la lecture graphique n'est pas précise et donc, cette lecture graphique pouvait être considérée comme étant une conjecture qui devrait être démontrée par la suite » (entretien Aristide, 20 Août 2020).

Pour Aristide, la méthode graphique n'est pas précise. Il est bien vrai qu'elle permet de faire une interprétation graphique. Mais il faut déterminer l'équation de la droite de régression analytiquement, c'est la méthode des moindres carrés qui permet d'avoir une réponse exacte.

Pour remédier aux erreurs de ces élèves, Aristide suggère de leur proposer plus d'exercices à faire à la maison et qu'il y ait un suivi sur ces travaux à faire à la maison.

La pratique d'enseignement de la statistique à deux variables

Concernant la deuxième partie de l'entretien, elle est relative à la pratique d'enseignement de la statistique à deux variables. Pour l'introduction de la notion de la série statistique à deux variables en classe, l'enseignant P_2 commence par les prérequis sur les fondamentaux en statistique. Pour P_4 , il introduit son cours de statistique par un exemple d'application dans la vie courante, pour amener les élèves à comprendre l'utilité de la statistique dans la vie quotidienne. Et quant à l'enseignant Aristide, il nous rappelle que la série statistique à deux variables est introduite en classe de première SE mais pas en première C, et la série statistique à une variable est introduite depuis le collège. Pour Aristide, y a les acquis des classes antérieures sur la série à une variable qu'il faut compléter avec les séries à deux variables. Dons les trois enseignants commencent leur cours de la série statistique à deux variables par vérifier et consolider les acquis antérieurs en statistique.

En ce qui concerne l'objectif de l'enseignement de la statistique à deux variables en terminale, P_2 ne nous a pas dit avec précision l'objectif de l'enseignement de la statistique prescrit dans le curriculum. Il fait juste allusion à l'utilité de la statistique dans la vie quotidienne. Nous signalons que le programme de mathématiques du second cycle de 2002 n'est pas très détaillé, on met juste les intitulés (les titres des chapitres) mais on ne précise pas les savoirs, savoir faire, commentaires et instructions dans le programme. Certains enseignants se réfèrent au détail de l'ancien programme ou au contenu du manuel CIAM, pour préparer leur cours sur la statistique à deux variables en terminale SE. Pour P_4 , l'objectif est de se servir d'un tableau croisé au départ, pour aboutir à la question demandée, il précise aussi que l'un des objectifs est de représenter le nuage des points puis d'interpréter la forme du nuage, s'il est allongé, dispersé. Quant à Aristide, l'objectif de l'enseignement de la statistique à deux variables en terminale SE est selon lui de faire comprendre d'abord aux élèves, le lien qui peut y avoir entre deux grandeurs statistiques ensuite de pouvoir faire une interprétation dans le but d'aboutir à une prévision. Selon Aristide, le rôle de l'enseignement de la statistique est de faire aussi l'application avec les données réelles. Donc pour Aristide, il s'agit d'une part de trouver une relation liant deux variables. Ensuite on pourra utiliser cette relation pour estimer ou prévoir une des deux variables connaissant l'autre. Et d'autre part de faire l'interprétation. Nous voyons ici que Aristide a une prise de recul sur le sens de son enseignement des séries statistiques à deux variables.

Concernant les compétences importantes en statistique à deux variables à construire chez les élèves. Pour l'enseignant P_2 les compétences à consolider c'est la notion de la droite de régression. Pour lui, il faut mettre l'accent sur la droite de régression qui permet d'avoir un meilleur ajustement qui approche la réalité. Pour connaître le point de vue de P_2 sur l'interprétation des résultats en statistique au niveau de la classe de terminale SE, il fait référence aux élèves en ce sens qu'ils ne savent pas pourquoi ils étudient la statistique. Pour P_4 la compétence à développer chez les élèves, c'est l'interprétation du lien entre deux grandeurs statistiques. En donnant des exemples dans la vie quotidienne, comme par exemple, interpréter le lien des grandeurs âges et poids. Aristide cite deux objectifs à construire et à consolider, la maîtrise des formules et l'interprétation des résultats en vue de faire la prévision. On se rend compte que les trois enseignants mettent l'accent sur l'interprétation lorsqu'il s'agit de compétences importantes à construire chez les élèves.

Concernant la méthodologie utilisée par les enseignants pendant leur cours, P₂ déclare utiliser la méthode dite active, elle consiste à faire responsabiliser les élèves pendant le cours, c'est-à-dire, faire contribuer les élèves à l'établissement des prérequis par un jeu de questions-réponses. Il commence par les prérequis, en posant des questions aux élèves. En participant à la leçon, les élèves vont comprendre et apprendre la notion visée. Pour P₄, chaque cours a des objectifs à atteindre, donc il fait le cours suivi des exemples pour faciliter la compréhension, ensuite il fait des travaux dirigés à la fin du chapitre et enfin il propose des exercices d'approfondissement à faire à la maison. Selon lui, les élèves ont des problèmes de documentation. La plupart des élèves n'ont pas de manuels faute de moyens financiers, ils se contentent du contenu de cours et des exercices proposés par l'enseignant. Nous signalons que dans la majorité des lycées en Centrafrique les bibliothèques ne sont pas riches en documents et dans certains lycées la bibliothèque n'existe quasiment pas. En plus de cela, un bon nombre des élèves ne fréquentent pas la bibliothèque, du fait qu'elle n'est pas riche en manuel scientifiques en général. En ce qui concerne Aristide, il commence un chapitre ou une séquence avec des activités dirigées où il responsabilise les élèves par rapport à des savoirs nouveaux dans la découverte de la leçon visée. Pour faire responsabiliser les élèves, il les fait travailler en contribuant à l'établissement des savoirs nouveaux. En statistique, il propose des activités où apparaissent les formules des concepts en statistique, la représentation du nuage des points, un problème d'interprétation.

Les trois enseignants disent qu'ils organisent des travaux dirigés à la fin de chaque chapitre. Selon eux, c'est obligatoire de faire des travaux dirigés à la fin de chaque chapitre, car cela permet de savoir si les élèves ont compris ou pas la leçon. Concernant Aristide, il s'organise de la manière suivante : il propose de plus une série d'exercices aux élèves à la fin de chaque chapitre, où les élèves vont d'abord essayer de faire à la maison avant de passer les corrigés en classe, sous la direction de l'enseignant. Il demande aux élèves de commencer les exercices à la maison en fonction des difficultés de ces exercices. Dans son organisation (Aristide), il fait le cours suivi des exercices d'application puis d'approfondissement et des problèmes. Son objectif est de responsabiliser les élèves et de les amener à prendre des initiatives devant un exercice. Les enseignants P₂ et P₄ procèdent de la même manière que Aristide, ils proposent les exercices, et ils laissent quelques jours aux élèves de traiter les exercices à la maison. Pendant les séances des travaux dirigés les élèves passeront au tableau pour corriger. S'ils sont bloqués l'enseignant peut intervenir. D'après P₂ parfois il les met en petits groupes pour faire les travaux dirigés.

Pour évaluer les élèves, P₂ fait soit un devoir individuel ou un devoir de groupe sur table, une fois qu'il a fini de faire les travaux dirigés. Quant à P₄, il informe les élèves sur l'évaluation avant de la faire. Concernant Aristide, après avoir épuisé deux ou trois chapitres, il programme une évaluation et à la fin du semestre il évalue le contenu du programme déjà fait.

Pour P₂, les difficultés des élèves en statistique à deux variables, concernent le remplissage du tableau statistique, ils ne savent pas se servir du tableau, en plus leur deuxième difficulté se situe au niveau du choix des échelles d'un repère. P₄ évoque deux difficultés, selon lui les élèves ont des difficultés avec l'usage des calculatrices et la non maîtrise des notions mathématiques. Quant à Aristide, les difficultés des élèves sont selon lui la non maîtrise des formules, des difficultés par rapport à la compréhension de la langue et aussi des difficultés d'usage des calculatrices scientifiques. D'après ce dernier, quand les élèves se retrouvent devant un problème de statistique qui utilise de grands nombres, ceux qui n'ont pas de calculatrice scientifique ou qui ne savent pas

bien manipuler seront en difficulté de traiter l'exercice. Il souligne que, un bon nombre des enseignants de mathématiques ont aussi des difficultés à l'utilisation des calculatrices scientifiques. Un bon nombre des élèves de terminale SE n'ont pas de calculatrice scientifique et c'est ce qui explique la difficulté d'usage des calculatrices scientifiques.

Les initiatives des élèves en statistique à deux variables

Concernant la partie trois sur la prise d'initiatives des élèves pendant l'enseignement ordinaire en classe, d'après P₂ certains élèves ont l'habitude de consulter d'autres manuels pour chercher à comprendre d'autres méthodes de la statistique, il qualifie cela comme une initiative à la recherche qu'il faut développer « parce qu'il a l'initiative de la recherche et l'essentiel c'est de voir est ce que y a la logique mathématique dans ce qu'il fait » (entretien P₂, 07 juillet 2020). Pour le choix des méthodes d'ajustement, P₄ préconise de préciser toujours la méthode à utiliser dans les exercices. Pour l'interprétation, il reste sur des questions fermées, comme pour savoir si l'ajustement trouvé est bon ou mauvais. Nous remarquons que les méthodes décrites par P₄ peuvent laisser penser que la prise d'initiatives des élèves est réduite dans son cours, l'enseignant P₄ ne met pas les élèves en position de prise d'initiative, « on parle juste d'ajustement, est-ce que c'est un bon ajustement et c'est tout on arrête. Parce qu'on ne peut pas demander tout à l'élève » (entretien P₄, 20 juillet 2020). D'après Aristide, tous les élèves ont des initiatives à prendre devant une situation. Il rappelle la production des élèves A et B dans le questionnaire comme une prise d'initiative, « être capable d'interpréter les résultats, mais interpréter un résultat, c'est prendre des initiatives, réfléchir et voire même se poser des questions et tirer une conclusion » (entretien Aristide, 20 Août 2020). Aristide se réfère à un exemple dans le questionnaire où l'élève interprète son résultat, c'est un bon exemple où on voit une prise d'initiative mais il n'a pas fait référence à son propre cours que nous avons pourtant observé.

Pour P₄, si on doit tenter de développer les initiatives des élèves, cela doit rester dans le cadre du programme de mathématiques pour le niveau terminale. Aristide considère l'interprétation des résultats statistiques qui fait partie de l'objectif de l'étude de statistique à deux variables, comme une prise d'initiative qu'il faut développer. Selon lui, l'interprétation des résultats amène l'élève à réfléchir à poser des questions et tirer des conclusions. Donc l'accent doit être beaucoup mis sur l'aspect interprétation.

5.6 Synthèse du chapitre 5

Dans ce chapitre nous avons apporté des éléments de réponse à notre première question de recherche (qui comporte trois sous-questions).

QR.1 : Quel est le savoir à enseigner au sujet des séries statistiques doubles et de l'ajustement affine en Centrafrique (terminale SE), et quelles sont les prises d'initiatives possibles des élèves relativement à ce savoir ?

QR.2 : Quel est le savoir effectivement enseigné à ce sujet dans l'enseignement ordinaire ?

L'analyse historique, la revue de la littérature et l'analyse du programme de terminale SE et du manuel de BTS nous a permis de concevoir un modèle praxéologique de référence (MPR) pour les séries statistiques doubles et l'ajustement affine.

Nous résumons ici nos principales observations.

Nous avons noté que les types de tâches : T3 : choisir la méthode d'ajustement affine, T7 : recueillir des données et T8 : nettoyer des données réelles qui figurent dans le MPR ne se retrouvent pas dans le programme et le manuel. Ni le programme de Centrafrique, ni le manuel de BTS ne mentionnent ces types de tâches, qui pourtant peuvent donner du sens à un enseignement de la statistique (Fine, 2013 ; Régnier, 2005 ; Zendrera et al, 2017). Les élèves n'auront donc pas l'initiative de recueillir eux-mêmes des données, ou de travailler avec des données réelles, ce qui nécessite de supprimer les valeurs aberrantes par exemple.

Les prises d'initiatives possibles des élèves que nous avons notées lors de notre analyse du manuel de BTS de France apparaissent dans les types de tâches suivants :

- T_Rep_Nuage. Il a la responsabilité de choisir l'échelle pour représenter le nuage de points ;
- T_Tra_DroiteD. Pour la représentation graphique de la droite (D) dans la méthode d'ajustement à la règle, l'élève a à sa charge de trouver deux points de la droite, et de tracer à la règle la droite (D) passant par ces deux points ;
- Pour T5 donner une estimation, T6 interpréter l'estimation obtenue par le modèle.

Nous avons ensuite analysé le manuel CIAM utilisé en Centrafrique, et nous l'avons comparé au MPR. Cette analyse a permis aussi de savoir où se situent des prises d'initiatives potentielles des élèves dans l'institution terminale SE en Centrafrique.

A travers l'analyse du manuel CIAM, nous avons aussi noté l'absence des types de tâches T3, T7 et T8. Nous avons de plus observé que dans le type de tâches T4.3 est toujours précisé dans l'énoncé des exercices, laquelle des deux droites de régression qu'il faut utiliser, entre la droite de régression de y en x ou la droite de régression de x en y . Pour ce type de tâches les élèves n'auront pas l'initiative de choisir la droite de régression qu'il faut utiliser.

Nous avons aussi observé que dans le manuel CIAM que la démonstration de la propriété de l'équation de la droite des moindres carrés est donnée : il y a là un élément technologique important, qui justifie la méthode des moindres carrés

Nous avons noté l'ajout de trois nouveaux types de tâches en faisant le MPR qui correspond au manuel CIAM : T_PartagerG1G2, T_PartagerG1G2 et T_Etab_TabS.

Les prises d'initiatives possibles des élèves observées dans le manuel CIAM, il y en a moins que dans le manuel de BTS. Nous avons :

- Le type de tâches T_Rep_Nuage laisse place à l'initiative de l'élève lorsque le choix de l'échelle est à sa responsabilité.
- Dans T5, la prise d'initiative de l'élève se situe dans la transformation de données de départ selon ce qui a été choisi pour y (par exemple transformer une année en rang de

l'année) pour trouver y . Pour T6, elle correspond à l'interprétation du résultat de l'estimation obtenu.

D'un autre côté, nous constatons à travers l'analyse d'un enseignement observé en Centrafrique que l'enseignant privilégie plus un travail sur la technique que le discours technologique relatif à la technique. Le discours technologique de l'enseignant est réduit, par rapport à ce qui est dans le manuel CIAM. L'enseignant utilise un manuel français (ancien) mais c'est un livre d'exercices.

Nous résumons ci-dessous les principaux types de tâches présents dans le cours observé :

- T1 : Faire la représentation graphique de la série double et de son point moyen ;
- T2 : Étudier si un ajustement affine est pertinent ;
- T4.2 : Réaliser un ajustement affine avec la méthode de Mayer ;
- T4.3 : Réaliser un ajustement affine avec la méthode des moindres carrés. Cependant ici, la technique n'est pas présentée ;
- T5 : Donner une estimation.

Nous avons vu durant le cours que les prises d'initiatives des élèves étaient très limitées. L'enseignant débute par une activité, mais il n'apporte pas aux élèves le savoir nécessaire, et ne leur laisse pas suffisamment de temps pour faire les tâches qui sont à leur portée. Certains élèves prennent toutefois des initiatives en posant des questions, mais les réponses de l'enseignant sont succinctes et il ne profite pas des questions pour engager un échange avec les élèves. Les conditions des classes surchargées en RCA ne se prêtent pas beaucoup aux échanges.

Nous discuterons plus en détail les prises d'initiatives effectives des élèves dans le chapitre suivant (chapitre 6).

L'analyse du questionnaire et des entretiens aux enseignants confirme que les enseignants préfèrent proposer aux élèves un exercice en utilisant la structuration usuellement présente dans les manuels, ils ne proposent pas aux élèves de choisir une méthode pour réaliser une estimation. Cela laisse peu d'initiatives aux élèves.

Nous avons également noté à travers l'analyse du questionnaire et des entretiens aux enseignants que, ces derniers n'acceptent pas une méthode qui sort de ce qu'ils font habituellement. En exemple, le cas de la méthode graphique utilisée par l'élève B (fictif). Cela confirme le peu d'initiative laissée aux élèves dans le choix de la méthode pour faire l'ajustement.

6 MISE EN PLACE ET ANALYSE D'EXPERIMENTATIONS

Alors que le chapitre 5 est consacré à l'analyse de l'enseignement ordinaire des séries statistiques doubles en RCA, nous entrons dans le chapitre 6 dans la partie de notre thèse dédiée à des expérimentations.

Nous avons constaté au chapitre 5 que dans l'enseignement ordinaire en Centrafrique les prises d'initiatives possibles pour les élèves semblaient limitées. Nous souhaitons faire des propositions d'enseignements qui ouvrent plus de possibilités. Il s'agit dans cette partie de répondre à la question de recherche suivante :

QR.3 : Quelles sont les prises d'initiatives effectives des élèves de Terminale en Centrafrique, lorsqu'on leur propose un exercice sur le thème de l'ajustement affine qui permettrait une telle initiative ? A quelles conditions peut-on mettre en œuvre en Centrafrique un enseignement qui laisse plus d'initiative aux élèves que ce qui est pratiqué dans l'enseignement ordinaire ?

Nous avons mené pour répondre à cette question trois expérimentations successives, au cours des années 2019 ; 2020 et 2021. Chaque expérimentation a été construite à partir des résultats de la précédente. De plus ces expérimentations s'articulent avec les questionnaires et entretiens avec les enseignants présentés au chapitre 5. Le questionnaire a été construit à partir du résultat de l'expérimentation de 2020 pour le soumettre aux enseignants. Et l'entretien fait un retour sur le questionnaire rempli par les enseignants.

Nous présentons et analysons ces trois expérimentations ci-dessous.

6.1 Analyse de la première expérimentation en classe « Prévission du ménage 1 » (2019)

Dans cette section nous présentons et nous analysons la première expérimentation que nous avons menée. Cette première expérimentation consiste en une activité soumise à des élèves de Terminale SE. L'objectif de cette expérimentation est d'abord de laisser le choix de la méthode d'ajustement affine à la charge des élèves dans un exercice, et ensuite d'observer à travers leurs productions où se situent leurs prises d'initiatives effectives. Ci-dessous nous faisons l'analyse de l'exercice proposé aux élèves, ensuite nous faisons l'analyse de productions des élèves.

6.1.1 *Présentation de l'exercice*

Ce problème est une adaptation d'un exercice du manuel BTS (page 251) figure 16. Nous avons introduit un changement important en modifiant les questions pour laisser le choix de la méthode d'ajustement affine à la charge des élèves.

Exercice

Le tableau suivant indique en euros les montants des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense qu'en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette prévision ? Vous expliquerez en détail la méthode que vous avez choisie. Puis justifier pourquoi un ajustement affine est correct (ou approprié).

Bon travail

Figure 16: Énoncé de l'exercice « prévision du ménage 1 » soumis à des élèves de Terminale SE en 2019

Nous présentons ci-dessous l'analyse a priori de cet exercice.

6.1.2 Analyse a priori de l'exercice

Pour cette analyse a priori, nous allons identifier les types de tâches qui peuvent intervenir dans la résolution de cet exercice. Nous notons également les changements de registres nécessaires, et où peuvent se situer les prises d'initiatives de ces élèves,

Le type de tâches pour cette question est de: « faire une estimation en l'an 2019 », qui relève du type de tâches que nous avons nommé T5. Ici il n'y a aucune indication de méthode, sur la manière de réaliser cette estimation. Ainsi nous considérons que les élèves prennent une initiative en choisissant cette méthode. Ce choix peut être décrit de la manière suivante :

- Pour faire une estimation, on doit réaliser un ajustement ;
- on va essayer avec un ajustement affine ;
- donc il faut trouver l'équation d'une droite qui ajuste le nuage de points.

Cependant il faut souligner que les élèves ont récemment travaillé ce chapitre, ce qui réduit la prise d'initiative nécessaire.

Plusieurs méthodes pourraient être utilisées pour trouver l'équation de la droite. Nous avons entre autres la méthode des moindres carrés (qui relève de T4.3) et la méthode de Mayer (qui relève de T4.2). Ces deux méthodes sont enseignées dans l'institution « classe de terminale SE en Centrafrique ». Ainsi les types de tâches pour les élèves sont : T_Dét_EquatMCyx, T_Cal_Cov,

T_Etab pour ceux qui utilisent T4.3 et T_Dét_EquatG1G2, T_Cal_CoordG1G2, T_PartagerG1G2 pour ceux qui utilisent T4.2.

Nous donnons ci-dessous le détail de chaque méthode.

Première méthode : Méthode des moindres carrés

Pour déterminer les coefficients a et b de l'équation de la droite de la forme : $y = ax + b$ par la méthode des moindres carrés. La technique consiste à calculer en premier lieu le coefficient directeur a . Pour cela, calculons d'abord la moyenne des variables x_i qui représentent le rang de l'année et celle des variables y_i représentant le montant des remboursements.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^6 x_i = \frac{1}{6} (1+2+3+4+5+6) = \frac{21}{6} = 3,5 \text{ et}$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^6 y_i = \frac{1}{6} (3568+4561+5502+6693+8241+9120) = \frac{37685}{6} = 6295,83$$

Ensuite calculons leur covariance :

$$\begin{aligned} \sigma_{xy} &= \text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^6 x_i y_i - \bar{x} \bar{y} \\ &= \frac{1}{6} (1 \times 3568 + 2 \times 4561 + 3 \times 5502 + 4 \times 6693 + 5 \times 8241 + 6 \times 9120) - 3,5 \times 6295,83 \\ &= \frac{1}{6} (151983) - 22035,405 = 3295,08 \end{aligned}$$

Et enfin la variance des variables x_i :

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^6 x_i^2 - \bar{x}^2 = \frac{1}{n} \sum_{i=1}^6 x_i^2 = \frac{1}{6} (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - (3,5)^2 = \frac{91}{6} - 12,25 = 2,916$$

Puis déduisons le coefficient directeur a par la relation suivante :

$$a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{3295,08}{2,916} = 1129,5$$

Les éléments technologiques qui permettent de faire ces calculs sont : la définition de la moyenne arithmétique permet de calculer \bar{x} et \bar{y} . Puis la définition de la covariance pour le calcul σ_{xy} et ensuite la définition de la variance permet de faire le calcul de σ_x^2 . Le coefficient directeur est obtenu par le quotient de la covariance de x et y avec la variance de x .

Le rôle de la calculatrice est très important dans la réalisation de ce type de tâches. Les élèves peuvent utiliser la calculatrice de différentes manières.

Pour calculer les moyennes \bar{x} et \bar{y} , ils peuvent procéder ainsi, ils utilisent la parenthèse pour le calcul, c'est-à-dire ils ouvrent une parenthèse et effectuent le calcul, puis ils divisent le résultat par

6

$$\bar{x} = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = 3,5 ;$$

$$\bar{y} = \frac{1}{6}(3568+4561+5502+6693+8241+9120) = \frac{37685}{6} = 6295,83$$

Pour le calcul de la variance et la covariance, les élèves qui n'ont pas de calculatrice utilisent leur téléphone qui n'a pas de touche moyenne pour faire des calculs, peuvent utiliser la parenthèse pour effectuer le premier calcul, moins le produit du deuxième terme. Ou bien ils peuvent recalculer à chaque fois jusqu'à ce qu'ils aboutissent au résultat, c'est-à-dire ils vont recalculer chaque terme au carré pour la variance ou le produit pour la covariance qui est dans la parenthèse, puis recalculer le deuxième avant de faire la différence des résultats.

La technique qui permet de déterminer le second membre de l'équation de la droite b , consiste à utiliser la relation $\bar{y} = a\bar{x} + b \rightarrow b = \bar{y} - a\bar{x}$.

En remplaçant \bar{x} et \bar{y} par leur valeur dans l'expression ci-dessus, on a :

$$b = 6295,83 - 1129,5 \times 3,5 = 2342,6$$

D'où on a l'équation de la droite de régression : $y = 1129,5x + 2342,6$, qui nous permettra de faire l'estimation.

Les registres utilisés dans la méthode des moindres carrés sont le registre algébrique des variables et le registre numérique des variables quantitatives discrètes pour la détermination l'équation de la droite de régression de y en x .

La deuxième technique qui est beaucoup plus attendue, pour la détermination des coefficients a et b , est d'établir un tableau statistique à double entrée.

							Totaux
x_i	1	2	3	4	5	6	21
y_i	3568	4561	5502	6693	8241	9120	37685
x_i^2	1	4	9	16	25	36	91
$x_i y_i$	3568	9122	16506	26772	41205	54720	151893

Tableau 10: Tableau statistique pour l'exercice « prévision du ménage 1 »

Pour cette technique, les élèves utilisent nécessairement la calculatrice pour remplir le tableau. Pour avoir l'effectif de la variable x_i et y_i dans le tableau, ils font la somme de chaque valeur de ces variables. Pour calculer les x_i^2 , ils élèvent chaque valeur des x_i au carré ou qu'ils multiplient chaque valeur de x_i par elle-même. Et pour le calcul du produit $x_i y_i$, ils recalculent à chaque fois les valeurs des variables correspondantes, puis ils recalculent aussi les résultats trouvés.

Les résultats du tableau 10, permettent de calculer, les moyennes arithmétiques, la variance et la covariance.

La moyenne arithmétique des x_i est :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{6}(21) = 3,5$$

La moyenne arithmétique des y_i est :

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{6}(37685) = 6295,83$$

La variance des x_i est :

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \frac{1}{6}(91) - (3,5)^2 = 2,916$$

La covariance des x_i et des y_i est :

$$\sigma_{xy} = \text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^6 x_i y_i - \bar{x} \bar{y} = \frac{1}{6}(151893) - 3,5 \times 6295,83 = 3295,08$$

L'utilisation de la calculatrice pour effectuer le calcul de la covariance est de diviser 151 893 par 6 puis $3,5 \times 6295,83$ et recalculer les résultats, ou bien, ils peuvent entrer toutes les données dans la calculatrice, c'est-à-dire $151893 \div 6 - 3,5 \times 6295,83 = 3295,08$

Puis on calcule le coefficient directeur a par la relation suivante :

$$a = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{3295,08}{2,916} = 1129,5 \text{ et } b \text{ est déterminé par la relation : } b = \bar{y} - a \bar{x}$$

$$b = 6295,83 - 1129,5 \times 3,5 = 2342,6$$

D'où on a l'équation de la droite de régression : $y = 1129,5x + 2342,6$

La technologie relative à cette technique est la définition de : la moyenne arithmétique, la variance, la covariance, la notion du coefficient directeur et la calculatrice.

En plus du registre numérique des variables quantitatives discrètes et du registre algébrique des variables, nous avons aussi le registre des tableaux des variables quantitatives qui sont utilisés pour la réalisation de la tâche demandée.

Pour les deux techniques utilisées pour déterminer l'équation de la droite de régression de y en x , il n'y a pas de prise d'initiatives, car les élèves appliquent les techniques travaillées en classe.

Deuxième méthode : Méthode de Mayer

Pour la méthode de Mayer, on note G_1 le point du sous nuage formé des trois premiers points et G_2 celui du sous-nuage formé des trois derniers points. Cela peut être représenté par le registre des tableaux des variables quantitatives discrètes suivants (tableaux 11 et 12):

x_i	1	2	3
y_i	3568	4561	5502

Tableau 11: Sous- série 1

x_i	4	5	6
y_i	6693	8241	9120

Tableau 12: Sous- série 2

La technique consiste à calculer les coordonnées de ces deux points moyens $G_1(\bar{x}_1, \bar{y}_1)$ et $G_2(\bar{x}_2, \bar{y}_2)$, représentant respectivement les sous-séries 1 et 2.

$$\bar{x}_1 = \frac{1}{n} \sum_{i=1}^3 x_i = \frac{1}{3}(1+2+3) = \frac{6}{3} = 2$$

$$\bar{y}_1 = \frac{1}{n} \sum_{i=1}^3 y_i = \frac{1}{3}(3568+4561+5502) = \frac{13631}{3} = 4543,66$$

$$G_1(2; 4543,66)$$

$$\bar{x}_2 = \frac{1}{n} \sum_{i=1}^3 x_i = \frac{1}{3}(4+5+6) = \frac{15}{3} = 5$$

$$\bar{y}_2 = \frac{1}{n} \sum_{i=1}^3 y_i = \frac{1}{3}(6693+8241+9120) = \frac{24054}{3} = 8018$$

$$G_2(5; 8018)$$

L'élément technologique qui justifie cette technique est la définition de la moyenne arithmétique et du point moyen.

Il y a plusieurs techniques possibles pour déterminer les coefficients a et b . Pour cela, nous verrons la technique la plus attendue d'être utilisée par les élèves.

La technique pour la détermination du coefficient directeur a est donnée par la formule suivante :

$$a = \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1} = \frac{8018 - 4543,66}{5 - 2} = \frac{3474,34}{3} = 1158,11$$

Pour déterminer b par une résolution d'équation du premier degré, nous utilisons l'un des deux points moyens qui appartient à la droite d'équation : $y = ax + b$

Soit $\bar{y}_2 = a\bar{x}_2 + b$ on a : $b = \bar{y}_2 - a\bar{x}_2 = 8018 - 1158,11 \times 5 = 2227,43$

D'où $y = 1158,11x + 2227,43$

La méthode de Mayer ne nécessite pas une prise d'initiative, c'est une technique apprise et travaillée en classe.

En 2019, le rang vaut $x=7$. La technique consiste à déterminer la valeur de y correspondant à $x=7$. Remplaçant cette valeur de la variable x dans l'équation de la droite $y = 1129,5x + 2342,6$

$$y = 1129,5 \times 7 + 2342,6 = 10249,1 \text{ euros}$$

En 2019 le ménage va dépenser plus de 10000 euros, pour le remboursement.

La prise d'initiative de l'élève se situe dans la conversion de l'année 2019 au rang $x=7$ pour trouver y et l'interprétation du résultat.

Pour justifier la pertinence de l'ajustement affine, cela relève du type de tâches T2, on peut faire T_Cal_Coef, c'est-à-dire calculer le coefficient de corrélation entre les variables x et y . $r = 0,99 \approx 1$ r est proche de 1, on peut envisager un ajustement affine. En générale, le contrat didactique prévoit qu'on répond toujours oui, pour les exercices où il n'y a pas de changement de variable.

On peut aussi faire T_Méth_graph, c'est-à-dire faire la représentation graphique du nuage de points. Si les points semblent alignés, alors une droite d'ajustement affine est appropriée.

Nous notons cependant que dans cet exercice, seulement six valeurs sont données pour chaque série. L'estimation obtenue pour la valeur dépensée durant l'année 7 est proche de 10000 ; mais on ne peut pas affirmer avec certitude qu'elle dépassera 10000. On peut attendre qu'un élève fasse cette remarque sur la qualité de l'estimation.

En se référant à Robert (2018), cet exercice correspond au deuxième niveau de mise en fonctionnement des connaissances des élèves, c'est-à-dire, des connaissances mobilisables.

Liste des critères d'analyse des productions d'élèves en 2019

Nous sommes parti des critères, présentant une liste d'items relatifs à la démarche entreprise, comme outil méthodologique pour l'analyse de chaque copie. Les items se présentent comme suit :

- Valeur trouvée pour l'estimation en 2019 : la valeur est acceptable / non acceptable ;
- Erreur dans la procédure de résolution ;
- Méthode choisie :

A propos de la mise en œuvre de la méthode : application d'une technique apprise en cours/présence d'explication sur la méthode ;

- Présence d'un graphique : oui/non ;
- Si présence d'un graphique, à propos du graphique : choix de l'échelle pertinent oui/non ;
- légende sur les axes oui/non ;

- Représentation du nuage de points/de la droite d'ajustement/plus autres rubriques pertinentes ;
- Présence d'une justification : cette justification est- elle correcte, suffisante ?

6.1.3 Analyse de productions des élèves

Nous avons soumis cet exercice le 25 avril 2019 à une classe de Terminale SE dans laquelle il y avait 84 élèves.

Nous avons choisi une situation de classe où les élèves travaillent en petits groupes, et interagissent entre eux pour faciliter la résolution du problème.

Dans un premier temps, nous décrivons l'analyse de productions des élèves d'une manière globale, puis par la suite nous ferons l'analyse détaillée des réponses formulées par les élèves dans les différents groupes. Nous cherchons aussi à indiquer où se situe leur prise initiative dans la résolution de l'exercice.

6.1.3.1 Analyse globale

Les élèves interagissent en groupes pour résoudre cet exercice, en temps limité de 1 heure.

Onze groupes sur seize ont choisi de réaliser un ajustement affine. Ils ont pris des initiatives, pour répondre à la question : en effet ils ont d'eux même déduit que pour répondre, il fallait estimer la dépense de 2019. Pour chercher l'estimation, ils ont décidé de faire un ajustement affine. Et ces onze groupes ont choisi la méthode des moindres carrés, pour trouver l'équation de la droite.

Les cinq autres groupes n'ont pas réussi à proposer une modélisation adéquate. Ils ont représenté le nuage de points et calculé le coefficient de corrélation linéaire. Ils ont donné des réponses incomplètes et n'ont pas proposé d'estimation.

Aucun groupe n'a choisi la méthode de Mayer. Ceci s'explique par le fait que, la méthode n'est pas beaucoup travaillée en classe.

En ce qui concerne la détermination des coefficients a et b , de l'équation de la droite de la forme : $y=ax+b$, par la méthode des moindres carrés, parmi les onze groupes qui ont choisi la méthode des moindres carrés, deux groupes ont utilisé la technique qui consiste à appliquer directement la formule pour déterminer les coefficients a et b . Et les neuf autres groupes ont utilisé la technique du tableau statistique. Cette technique consiste à faire ou à compléter un tableau statistique, puis utiliser les résultats du tableau pour déduire les coefficients a et b de l'équation de la droite. six groupes sur onze, ont pu déterminer l'équation de la droite.

Concernant les six groupes qui ont déterminé l'équation de la droite, seuls cinq groupes ont réussi à faire une estimation.

Pour justifier la pertinence de l'ajustement affine, cinq groupes ont utilisé comme technique, le coefficient de corrélation linéaire entre les variables x et y . trois groupes ont utilisé le raisonnement graphique du nuage de points comme technique, pour justifier l'ajustement affine.

Sept groupes n'ont pas abordé la question. Et un groupe a fait une justification théorique par rapport à la notion de coefficient de corrélation linéaire sans effectuer le calcul.

Groupe	A fait un ajustement	A trouvé l'équation de la droite	A proposé une estimation pour 2019
1	Oui	Non	Non
2	Non	Non	Non
3	Oui	Non	Non
4	Non	Non	Non
5	Oui	Oui	Oui
6	Oui	Non	Non
7	Oui	Oui	Oui
8	Oui	Oui	Oui
9	Oui	Oui	Oui
10	Oui	Non	Non
11	Oui	Oui	Oui
12	Non	Non	Non
13	Non	Non	Non
14	Oui	Non	Non
15	Non	Non	Non
16	Oui	Oui	Non

Tableau 13: Récapitulatif de l'analyse globale

6.1.3.2 Analyse détaillée

Dans cette section, les items qui guident notre analyse détaillée sont les suivants :

- La valeur de l'estimation en 2019 ;
- La méthode choisie ;
- Le graphique ;
- La justification de la pertinence de l'ajustement affine.

Nous faisons par ailleurs une analyse détaillée de certaines copies.

Concernant la valeur de l'estimation en 2019.

Concernant la valeur de l'estimation en 2019, c'est le type de tâches T5, aucun groupe n'a trouvé une valeur acceptable de l'estimation en 2019. Nous entendons par valeur acceptable, une valeur

qui est approximative à la valeur de l'estimation en 2019 avec la méthode des moindres carrés, avec un écart de 10^{-2} près. Par contre, six groupes d'élèves sur les onze ont trouvé des valeurs fausses. Et les cinq autres groupes ont donné des réponses incomplètes. Ils sont restés sur la détermination des coefficients de l'équation de la droite, donc ils ne proposent pas d'estimation.

Sur ces six groupes qui ont trouvé l'équation de la droite, les groupes 7, 9, 11 et 16 ont fait l'erreur dans la détermination de l'équation de la droite. Les valeurs des concepts statistiques (la variance et la covariance) mobilisés pour trouver l'équation de la droite sont erronés, donc la mise en œuvre de ces concepts fait que ces groupes n'ont pas trouvé des valeurs acceptables de l'estimation.

Les élèves de ces groupes ont fait des erreurs qui sont liées aux formules. Cependant les groupes 7, 9 et 11 ont des initiatives dans la réalisation de tâches qui leur sont demandées : c'est-à-dire, ils ont réussi à proposer un modèle pour faire l'estimation en l'an 2019, en proposant le rang de l'année 2019, $x=7$. Certes, ils ont fait des erreurs sur les formules de la variance et de la covariance, mais leur modèle de l'estimation est juste.

Pour le groupe 5, les élèves ont trouvé l'équation de la droite, mais ils n'ont pas trouvé une valeur juste de l'estimation. Ils n'ont pas utilisé la bonne valeur de la variable x . Nous constatons que dans ce groupe, les élèves ont appliqué les techniques apprises en classe pour déterminer l'équation de la droite. Mais il leur manque des prises d'initiatives dans la détermination de la valeur du rang de l'année en 2019. Ils ont mal interprété le x . Au lieu du rang de l'année, ils ont mis la valeur 10000 de la dépense maximum prévue en 2019. Et c'est une difficulté liée à la modélisation, ici à l'interprétation de l'énoncé.

La procédure utilisée par le groupe 8 est bonne, mais la valeur de l'estimation est fausse. Cela est dû à des erreurs de calculs. Ils ont des prises d'initiatives, pour trouver la valeur de la variable $x=7$, et de réaliser la tâche qui leur est demandée.

Il nous semble que l'erreur de calcul s'explique par le fait que, dans certains groupes il y a qu'un seul élève qui a la calculatrice. Et s'il se trompe dans la manipulation des calculs élémentaires, alors cela va jouer sur le résultat du travail du groupe. Mais dans le groupe 8, on voit une prise d'initiatives dans la réalisation de la tâche demandée. Ils ont réussi à proposer une modélisation adéquate.

Rappelons que, l'exercice proposé aux élèves est en langage naturel et sous forme de tableau. Le fait qu'ils choisissent ou trouvent la méthode appropriée et arrivent à proposer une modélisation adéquate pour résoudre l'exercice peut être considéré comme une prise d'initiative. Leur prise d'initiatives se situe dans le choix dans le choix du rang de l'année 2019 qui est $x=7$.

Pour la méthode choisie :

Onze groupes ont choisi de travailler avec la méthode des moindres carrés. Cette méthode a été beaucoup plus travaillée en classe pendant les exercices.

Le groupe 1 a utilisé la technique apprise en classe pour déterminer les valeurs des concepts statistiques comme la variance, les moyennes et la covariance. Pour ce faire, ils utilisent la calculatrice pour remplir le tableau statistique. Dans ce groupe, ils n'ont pas pu finir l'exercice,

faute de temps. Ils ont une mauvaise gestion du temps, ils passent plus de temps à interagir entre eux, sans pouvoir mettre par écrit leurs idées.

Le groupe 3 utilise aussi la technique du tableau statistique pour trouver les valeurs des résumés statistiques. Et on voit l'utilisation du registre algébrique pour trouver l'équation de la droite. On constate que la technique était travaillée en classe, mais les élèves de ce groupe n'ont pas fini l'exercice.

Pour le groupe 5, ils ont utilisé le registre tableau et algébrique pour trouver l'équation de la droite de régression. Et l'explication fournie par ce groupe sur la valeur de l'estimation est juste, mais la valeur de l'estimation trouvée n'est pas acceptable. Il nous semble que les élèves de ce groupe connaissent bien la technique apprise en classe sur la méthode des moindres carrés.

Le groupe 6 (figure 17), a mis en œuvre, la méthode des moindres carrés, en utilisant directement les formules et les propriétés sur la méthode des moindres carrés apprises en cours. Après avoir calculé les résumés statistiques, ils ont expliqué la méthode, permettant de trouver l'équation de la droite de régression. En plus, on voit l'utilisation du registre tableau pour calculer l'effectif total.

$$S_y^2 = \frac{1}{8} [(37685)^2 - (\frac{37685}{6})^2] = \frac{1}{6}$$

$$S_y^2 = 230118392,9$$

- Appliquons la méthode des moindres carrés qui permettront de déterminer les droites de régression de (D) et $(D)'$

on a: $y = ax + b = (D)$ $\left\{ \begin{array}{l} a = \frac{\text{Cov}(X,Y)}{S_x^2} \\ b = \end{array} \right.$

Figure 17: Extrait de copie du groupe 6

Les groupes 7, 8, 9 et 11, utilisent la même technique du tableau statistique pour le calcul des concepts statistiques et le registre algébrique pour avoir l'équation. Dans leur procédure, les groupes 7, 8 et 9 n'ont pas expliqué la méthode utilisée. Seul le groupe 11 explique la méthode utilisée pour trouver la prévision.

Pour $n = 7$

$\hat{y} = 1787,28$

On cherche la droite de régression de y en x , puis on remplace x par le rang de l'année 2017 qui est égal à 7 pour affirmer la prévision du ménage.

Figure 18: Extrait de copie du groupe 11

Et on voit une prise d'initiative dans le calcul de la valeur de prévision en année 2019, ils ont d'abord trouvé le rang de l'année, puis calculé la valeur de la prévision.

Pour les groupes 10, 14 et 16, la technique utilisée consiste à calculer les concepts statistiques sans utiliser le tableau statistique. Ils se sont arrêtés sur le calcul de la moyenne et la covariance. Seul le groupe 16 est arrivé au calcul de l'équation de droite. Ces groupes manquent d'initiatives dans leur façon de procéder pour trouver la prévision.

Concernant le graphique.

Ici, le type de tâches est T1. Le graphique n'a pas été explicitement demandé et les élèves utilisent la même procédure fait en classe, pendant les travaux dirigés.

Trois groupes sur seize n'ont pas utilisé la représentation graphique pour réaliser l'ajustement affine.

Pour le groupe 3, le choix de l'échelle n'est pas juste, de plus ils ont confondu l'axe des abscisses avec l'axe des ordonnées. Pourtant l'énoncé précisait bien quelles valeurs étaient considérées comme x et lesquelles comme y .

Les groupes 2, 4, 5, 8, 9, 10, 13, 15 et 16, ont fait une représentation graphique, sans pouvoir préciser l'échelle. Tous ces groupes ont fait une représentation du nuage de points. Et seul les groupes 5 et 15 ont représenté la droite d'ajustement.

Le groupe 6 a fait une représentation graphique sans présenter l'échelle. L'ensemble des points de coordonnées $M_i(x_i, y_i)$, qui forme le nuage de point est représenté sous forme de légende, puis de graphique.

Pour les groupes 1 et 7, ils ont fait un choix pertinent de l'échelle qui est à la charge de l'élève, c'est une prise d'initiative pour les élèves de ces deux groupes. Le nuage de points est aussi représenté.

Dans la plupart des représentations, nous constatons que les élèves ont centré leur représentation graphique sur le nuage de points, sans penser à la représentation d'une droite d'ajustement affine. On observe leur prise d'initiatives quand ils font un choix pertinent de l'échelle.

Concernant la justification de la pertinence de l'ajustement affine.

C'est le type de tâches T2. Dix groupes sur seize n'ont pas répondu à la question, et c'est un manque de prise d'initiative de la part des élèves de ces groupes.

Le groupe 12, a donné une justification qui n'est pas juste à cause de la valeur du coefficient de corrélation linéaire qui est faux.

Cinq groupes ont donné une justification correcte, dont deux groupes (5 et 9) ont utilisé la valeur du coefficient de corrélation linéaire, pour justifier la pertinence d'un ajustement affine. Par ailleurs le groupe 5 trouve $r > 1$, ce qui n'est pas juste mais ils commentent que « *l'ajustement affine est vérifié* ». Nous constatons que ces élèves ne voient que l'aspect calculatoire de la statistique, ils

n'ont pas la capacité attendue de savoir qu'on a toujours : $-1 \leq r \leq +1$. Le fait de trouver r en dehors de cet intervalle ne leur pose pas problème.

or

$$r = \frac{\text{Cov}(u_i, y_i)}{V_{u_i} V_{y_i}} = \frac{33\,32,6}{1,7 \times 1957,76} = \frac{3332,6}{3328,19}$$

$$r = 1,001$$

~~Après~~ D'après le résultat obtenu avec $r \gg 1$ alors on peut conclure l'ajustement affine est vérifié

Figure 19: Extrait de copie du groupe 5

Par contre, les autres groupes (7, 10 et 13), ont utilisé la représentation graphique du nuage de points, pour la justification de la pertinence de l'ajustement affine.

Calculons la corrélation

On sait que : $r = \frac{\text{Cov}(u, y)}{\sigma_u \times \sigma_y}$

Or d'après le graphique on voit que les nuages de points sont assimilable à une droite donc il peut passer une droite le plus près possible des points de nuage d'où un ajustement affine est approprié.

Figure 20: Extrait de copie du groupe 7

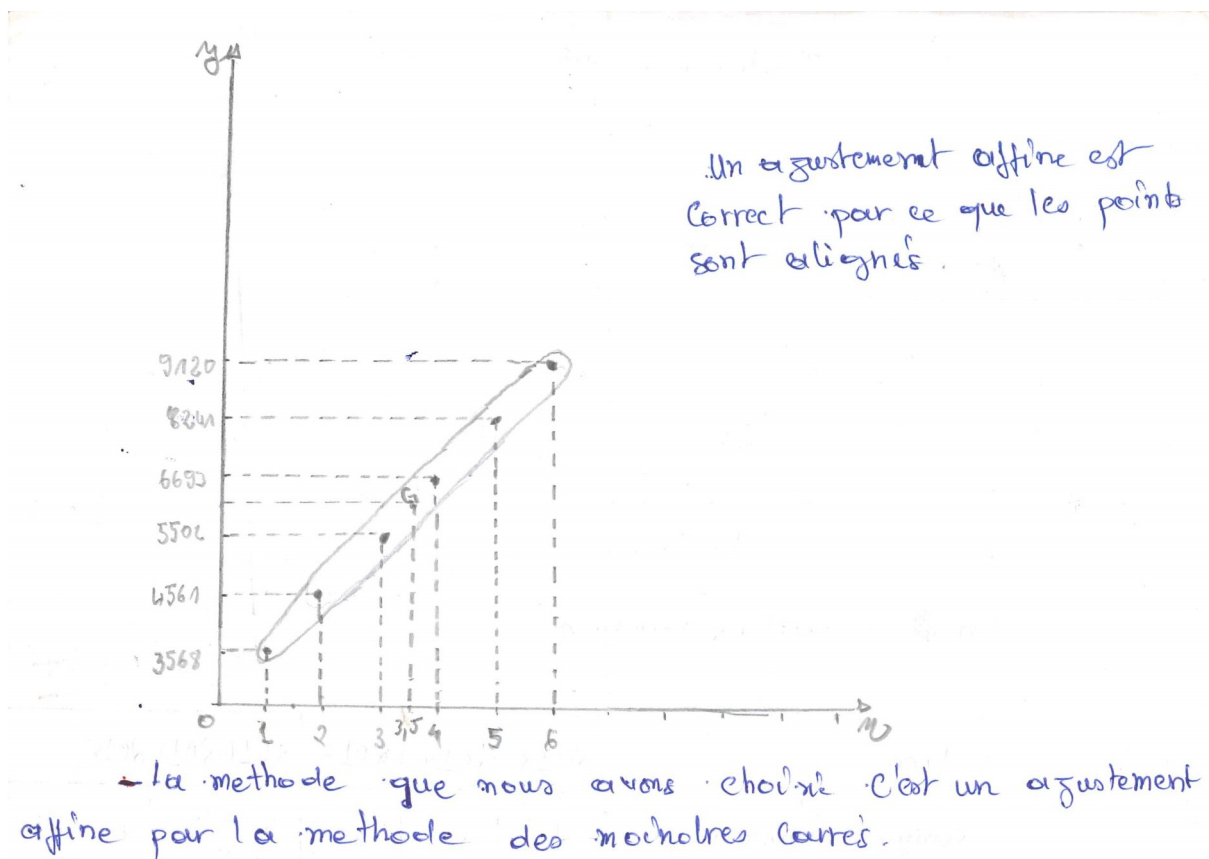


Figure 21: Extrait de copie du groupe 10

Les élèves de ces cinq groupes (5, 7, 9, 10 et 13) ont bien justifié la pertinence de l'ajustement affine.

Analyse de certaines copies : groupe 13, groupe 2 et groupe 8

Nous avons choisi ces productions pour les raisons suivantes. Pour les deux premières copies (groupe 13 et groupe 2), les élèves ont pris l'initiative de donner des modélisations qui ne relèvent pas de l'ajustement affine. Quant à la troisième copie (groupe 8), les élèves ont proposé une modélisation juste.

Pour le groupe 13. Les élèves de ce groupe ne semblent pas avoir pensé à un ajustement affine. Ils proposent une autre modélisation. Par ailleurs, les élèves de ce groupe se sont mis à faire la différence de montant des remboursements (y_i) des années successives (différence des dépenses de chaque année). Puis, ils ont fait la moyenne arithmétique, qui donne 925. Cette valeur est considérée comme la dépense en 2019. Pour obtenir le montant de remboursement en 2019, ils ont fait la somme du montant de remboursement en 2018 et la dépense en 2019 ($9120+925=10045$), ce qui permet d'avoir 10045 euro pour le remboursement.

- année
- Nous avons fait la somme des dépenses différences dépenses des années. C'est à dire de 2014 jus qu'au 2018, et nous avons obtenu 5552 et en suite nous avons divisé cette somme par 5 qui est le rang de l'année 2018, et nous avons obtenu la différence de l'année 2019 qui est: ~~773~~. 925
 - Et en fin nous avons obtenu la dépense de l'année 2019! $9120 + 925$ qui nous donne 10045.
 - Cette prévision est fautive car le ménage va dépenser plus de 10000 euros.
- L'allure sur le graphique est droite et montante alors l'ajustement affine est correct.

Figure 22: Extrait de copie du groupe 13

La méthode utilisée par ce groupe pour réaliser la tâche ne correspond pas à un ajustement affine. Toutefois elle donne une estimation dont l'ordre de grandeur est correct. Il serait possible d'accepter cet autre modèle et de le discuter dans la classe, pour réfléchir au sens de l'estimation.

Pour les élèves du groupe 2, ils ont fait la différence des montants de remboursement de deux années successives. Ils constatent que le montant décroît de 2014 à 2015 et croît de 2016 à 2017, puis décroît de nouveau en 2018. Et ils supposent que ce montant va décroître en 2019 aussi. Donc ils concluent que le ménage dépensera moins de 10000 euros en 2019. On observe ici, une prise d'initiatives de la part des élèves de ce groupe. Là encore, ils ne sont pas passés par un ajustement affine.

Nous pensons que ces deux copies montrent qu'il serait possible de mieux faire réfléchir les élèves au sens de ce qu'est une estimation, et à ce qui conduit à choisir une méthode ou une autre méthode. Il serait possible de travailler avec d'autres données, pour comparer les méthodes inventées par ces élèves et la méthode des moindres carrés.

Au rang des années 2013 à 2015, les taux des remboursements du ménage varient peu, de 993 en chute à 941 par contre de 2015 à 2017, les taux ajoutés sur les montants précédents sont importants de 191 à 1548 et en 2018 le taux ajouté sur le montant de 2017 est peu, ce qui veut dire qu'en 2019 aussi le taux qui va être ajouté aussi sera peu. Car le ménage alors le ménage dépensera moins de 10000 euros en 2019.

Figure 23: Extrait de copie du groupe 2

Concernant le groupe 8, les élèves ont associé « estimation » et « ajustement affine ». Ils ont choisi la méthode des moindres carrés pour trouver l'équation de la droite et proposé une estimation pour 2019. On observe une prise d'initiatives de la part des élèves du groupe 8.

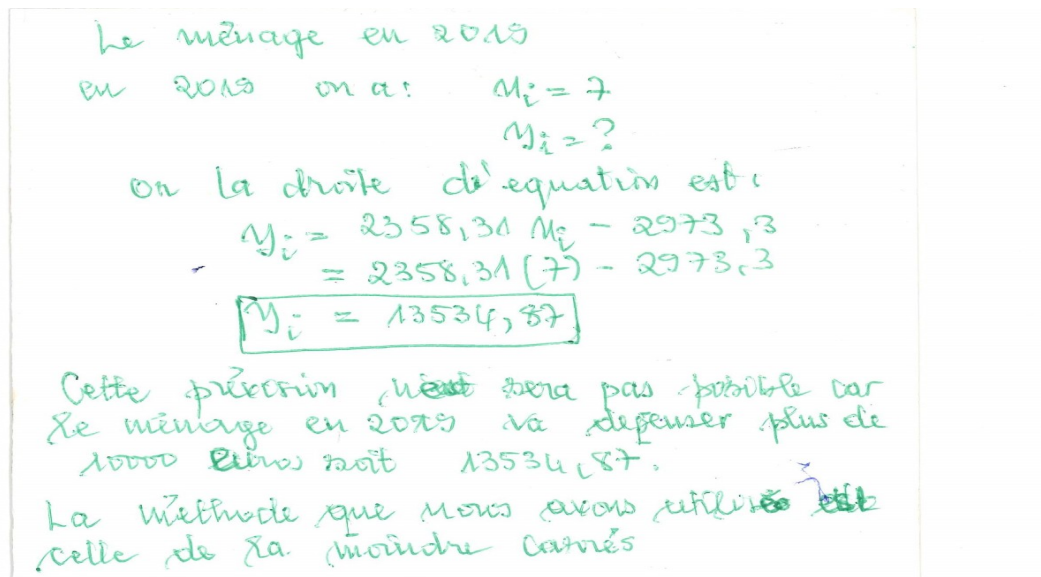


Figure 24: Extrait de copie du groupe 8

En ce qui concerne la production des élèves, nous avons observé des prises d'initiatives possibles des élèves pour les types de tâches T5, T4.3 et T_Rep_Nuage. Quand ils font l'estimation, leur prise d'initiative se situe d'abord dans le choix de méthode pour l'estimation. S'ils utilisent la méthode des moindres carrés (T4.3), ils ont une prise d'initiatives dans le choix du rang de l'année 2019 qui est $x=7$. On observe également leur prise d'initiatives quand ils arrivent à faire un choix pertinent de l'échelle du repère.

6.2 Analyse de la deuxième expérimentation en classe (2020)

A la suite de cette première expérimentation, nous avons élaboré un nouveau sujet à soumettre à des élèves. Nous nous sommes en particulier appuyé sur les résultats de la première expérimentation. Nous voulions en effet amener les élèves à réfléchir au choix de la méthode pour l'estimation. Nous sommes donc reparti de l'énoncé précédent en le modifiant, nous expliquons ci-dessous comment. Nous avons par ailleurs ajouté deux nouveaux problèmes.

Le but de cette partie du chapitre 6, consiste à observer où se situe la prise d'initiatives des élèves à travers ces trois problèmes qui leur sont proposés. Nous faisons l'analyse de ces trois problèmes proposés aux élèves, suivie de l'analyse des productions des élèves.

6.2.1 Analyse a priori des problèmes

Énoncé du problème 1 : « Prévion du ménage 2 »

L'exercice suivant a été proposé à des élèves de Terminale SE en Centrafrique.

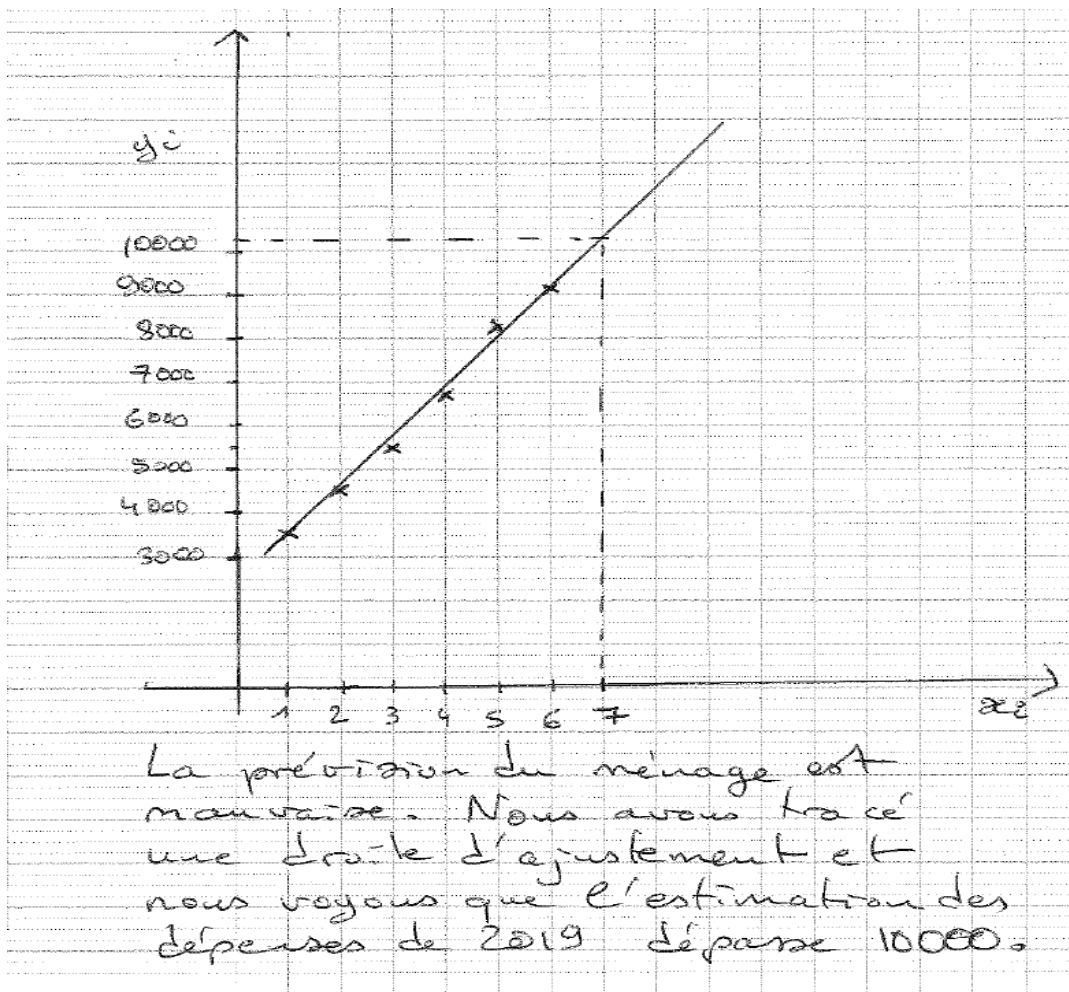
Énoncé de l'exercice donné à des élèves de Terminale Scientifique.

Le tableau ci-dessous indique le montant des remboursements annuels y_i , effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense que en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette affirmation ?

Ci-dessous est présentée la production d'un groupe d'élèves correspondant à cet énoncé.



Est-ce que vous jugez cette réponse juste? Justifier.

Analyse a priori du problème 1

A partir de l'énoncé du devoir précédent (« Prévion du ménage 1 », expérimentation 1, partie 6.1), nous avons fabriqué une production fictive nous-même. En effet, les élèves avaient pratiquement tous opté pour la même méthode utilisant les moindres carrés. Or, plusieurs avaient effectué une représentation graphique qui aurait pu servir de point d'appui à une procédure graphique. Nous souhaitons donc savoir ce que les élèves pensent d'une telle procédure, pourquoi ils ne l'utilisent pas. L'exercice 1 nécessite une importante prise d'initiative, parce qu'il est tout à fait atypique. Il demande aux élèves de prendre du recul pour évaluer la correction d'une production d'un autre élève (fictif).

Nous attendons de la part des élèves plusieurs types de réponses possibles :

1. Élèves qui considèrent que la réponse est juste.

Nous nous intéressons aux justifications utilisées par ces élèves. En effet, il est possible que certains élèves s'intéressent seulement à la conclusion de l'élève fictif, tandis que d'autres parleront de la méthode.

Pour ceux qui parlent de la méthode, nous nous attendons à des justifications du type suivant.

« En effet, les élèves ont bien représenté les données de l'énoncé par un nuage de points. Et ils ont tracé une droite qui approche bien le nuage de points, donc ils peuvent l'utiliser pour l'estimation. En plus ils ont bien vu que le rang de l'année est 7, et ils ont bien mis des pointillés pour indiquer où est la valeur correspondant à 7. »

Pour cet exercice, le type de tâches est de valider ou non la réponse proposée. Pour les élèves qui considèrent que la réponse fictive est juste, la technique pour accomplir ce type de tâches, consiste à observer si les points du nuage sont bien représentés. Ensuite de constater que la droite tracée approche bien le nuage de points. Alors pour faire l'estimation, On observe que le rang de l'année est $x=7$ indiqué en pointillé et la valeur correspondant à 7 dépasse 10000 euros. Ceci justifie que la méthode de l'élève fictif peut être considérée comme juste.

Cependant d'autres élèves pourront dire que la réponse est correcte, en se centrant seulement sur la conclusion : « les dépenses dépassent 10 000 ». Ils peuvent faire un calcul par différentes méthodes, et se baser sur l'estimation qu'ils ont obtenue. SI leur estimation dépasse 10000, ils diront que la production est juste, simplement parce que la conclusion est la même.

2. Élèves qui considèrent que la réponse n'est pas juste, ou pas correcte.

- Ceux qui pensent que la méthode graphique n'est pas correcte : On ne peut pas utiliser une méthode graphique, il faut appliquer la méthode des moindres carrés.
- Ceux qui pensent qu'on peut approcher par une droite, mais qu'il faut calculer son équation pour faire l'estimation : La réponse n'est pas juste, du fait que la méthode graphique ne donne pas une bonne estimation. Il aurait fallu calculer l'équation de la droite qu'ils ont tracée.

- Ceux qui pensent que la production manque de justifications : Ils auraient du dire comment ils ont tracé la droite. On ne peut pas vérifier numériquement l'affirmation de ces élèves.
- Ceux qui pensent que c'est imprécis : les points ne sont pas placés de manière exacte, donc cette méthode est trop approximative on ne peut pas répondre à la question.

Pour les élèves qui considèrent que la réponse est fausse, nous examinons la technique utilisée. Ces élèves peuvent s'appuyer sur l'imprécision du graphique : la technique pour la justification consiste à dire que comme on est proche de 10000, la méthode graphique ne suffit pas pour conclure qu'on est au-dessus de 10000 ou au-dessous de 10000.

D'autres élèves peuvent en revanche argumenter en disant simplement qu'il faut dans cette situation appliquer la méthode des moindres carrés. Ces élèves sont attachés au contrat habituel, et ne prennent pas l'initiative de sortir de ce contrat.

Par ailleurs, cet exercice correspond au troisième niveau de mise en fonctionnement des connaissances des élèves (Robert, 2018), autrement dit, le niveau des connaissances disponibles, qui consiste à utiliser des méthodes non prévues.

Énoncé du problème 2

Le tableau suivant représente des couples de notes en Mathématiques et en Comptabilité de 100 élèves que l'on a regroupés lorsqu'ils étaient égaux.

Note x_i en mathématiques	7	8	8	9	9	9	9	10	10	11	12	14	15
Note y_i en comptabilité	6	7	9	8	10	11	12	8	11	12	10	13	14
Effectif	5	6	8	10	12	14	8	11	7	9	3	5	2

1. Déterminer les coordonnées du point moyen G du nuage formé par les points $M_i(x_i, y_i)$
2. Calculer la covariance de la série statistique double (X, Y).
3. Déterminer une équation de la droite de régression de y en x, obtenue par la méthode des moindres carrés.
4. Un élève a obtenu la note de 13 en mathématiques. Peut-on estimer sa note en comptabilité ?

Justifier.

Analyse a priori du problème 2

Le problème 2 suivant est présenté dans le manuel de mathématiques comptabilité et gestion informatique de gestion (2000), utilisé en France en année de BTS, à la page 253. Nous avons ajouté une quatrième question d'interprétation, pour enrichir l'exercice.

L'erreur attendue pour cet exercice, est que les élèves ne tiennent pas compte de l'effectif associé aux notes de mathématiques et de comptabilité.

Nous observerons donc si les élèves font ou non cette erreur. Dans tous les cas, nous analysons en particulier la question 4 d'interprétation. En effet il est possible que les élèves fassent la bonne interprétation, même s'ils se trompent sur la bonne équation de la droite de régression.

1. Pour la première question dont le type de tâches est T_Cal_CoordG, la technique consiste à utiliser le type de tâches T_Cal_Moyxy, c'est-à-dire de calculer la moyenne \bar{x} de X et la moyenne \bar{y} de Y . La technologie qui justifie ce type de tâches est la définition de la moyenne pondérée de la série statistique. Et la procédure attendue de la part des élèves est d'utiliser la moyenne pondérée

pour calculer les coordonnées du point moyen G du nuage, c'est-à-dire : $\bar{x} = \frac{\sum_{i=1}^{13} n_i x_i}{100} = \frac{958}{100} = 9,58$

et $\bar{y} = \frac{\sum_{i=1}^{13} n_i y_i}{100} = \frac{990}{100} = 9,9$ avec $n = \sum_{i=1}^{13} n_i = (5+6+8+10+12+14+8+11+7+9+3+5+2) = 100$

on a : $G(9,58; 9,9)$

La difficulté prévisible pour cette question est l'interprétation du tableau statistique, en prenant en compte la pondération de chaque note du tableau. Pour ceux qui ne prennent pas en compte la pondération de la note, ils risquent de faire des erreurs de calcul. Le calcul des coordonnées du point moyen ne relève pas vraiment d'une prise d'initiative.

2. Pour le type de tâches T_Cal_Cov, la technique qui permet de réaliser le type de tâches est de se servir de T_Cal_Moyxy et T_Cal_Var, justifié par un discours technologique qui sont la définition de la variance et de la moyenne.

La procédure est d'utiliser la formule de la covariance :

$\sigma_{xy} = cov(x, y) = \frac{\sum_{i=1}^{13} n_i x_i y_i}{n} - \bar{x} \bar{y} = \frac{9700}{100} - 9,58 \times 9,9 = 2,158$. Soit d'utiliser un tableau statistique

pour calculer le numérateur d'abord et ensuite d'appliquer la formule de la covariance. L'erreur probable de certains élèves ici, est de faire le calcul sans tenir compte des pondérations des notes du tableau, pour ceux qui ne savent pas interpréter un tableau statistique. Et le calcul de la covariance ne relève pas d'une prise d'initiative.

3. Concernant la troisième question dont le type de tâches est : T_Dét_EquatMCyx, la procédure attendue est la méthode des moindres carrés, qui consiste à déterminer les coefficients a et b dans l'équation de la droite d'ajustement $y = ax + b$, on aura : $y = 0,88x + 1,383$ (1). Si la méthode des moindres carrés est déjà travaillée en classe, alors la question ne relève pas d'une prise d'initiatives. Les élèves vont utiliser la formule apprise en classe pour faire le calcul. En plus de cela, la covariance est déjà déterminée à la question précédente. Les élèves peuvent commettre des erreurs de calcul au niveau de la détermination des coefficients a et b .

Donc pour le types de tâches T_Dét_EquatMCyx, la technique consiste à déterminer les coefficient a et b de l'équation de la droite $y = ax + b$. La technologie relative à ce type de tâches est la propriété de la droite de régression et les définitions de la covariance, variance et la moyenne.

4. Pour la quatrième question dont les types de tâches sont :T5 et T6. La procédure est d'utiliser l'équation de la droite d'ajustement trouvée à la question précédente, pour déduire la note en comptabilité. Pour cela, ils doivent prendre $x=13$ pour trouver la valeur de y . $x=13$ correspond à la note de mathématiques, on remplace x par sa valeur dans l'équation de la droite de régression notée (1). On a $y=0,88 \times 13 + 1,383 = 12,94$. Pour avoir la note de la comptabilité de cet élève qui est 12,94.

La difficulté prévisible est que certains élèves peuvent confondre les deux variables, en donnant à y la valeur 13 et chercher la valeur de x . Si les élèves ont fait une erreur de calcul pour déterminer l'équation de la droite, leur interprétation peut tout de même être correcte dès qu'ils posent $x=13$ et remplacent dans l'équation trouvée.

Donc pour le type de tâches T5 : faire une estimation. La technique consiste à se servir de l'équation de la droite des moindres carrés pour déterminer la valeur de y . Ensuite, pour justifier l'estimation ou interpréter T5, on peut utiliser la notion du coefficient de corrélation linéaire.

Le problème 2 consiste à mettre en jeu de manières immédiates les théorèmes, des formules, des propriétés et des définitions. Il s'agit du niveau technique qui correspond au premier niveau de mise en fonctionnement des connaissances des élèves selon Robert (2018).

Énoncé du problème 3

Le tableau suivant donne, pour douze mois consécutifs, l'évolution des dépenses publicitaires (en milliers d'euros) d'une société commerciale.

Numéro du mois : x_i	1	2	3	4	5	6	7	8	9	10	11	12
Montant des dépenses : y_i	3	4,5	3,7	5,2	5,2	6	7,5	7,5	8,2	9,7	9,7	10

1. Représenter dans un repère orthogonal le nuage des points M_i de coordonnées (x_i, y_i) correspondant à cette série statistique.
2. Tracer la droite passant par les points A(1 ; 3) et B(9 ; 8,2). Déterminer l'équation réduite de cette droite.
3. On utilise cette droite pour réaliser un ajustement affine du nuage des points M_i .
 - a) Estimer le montant des dépenses durant le quatorzième mois.
 - b) Estimer le rang du mois au cours duquel le montant dépassera pour la première fois 13 milliers d'euros.

Analyse a priori du problème 3

Le problème proposé est un ancien sujet de CAPES de mathématiques (Dossier 8-4 de 2007). Pour éviter les erreurs de calcul de la part des élèves, nous avons modifié les valeurs du montant des dépenses y_i pour avoir des valeurs simples à manipuler.

1. Pour la première question dont le type de tâches est T_Rep_Nuage, la technique consiste à choisir l'échelle, et placer ensuite les points du nuage sur le repère orthogonal. Cette technique est justifiée par un discours technologique qui est le repérage des points dans un repère. La représentation du nuage de points et le choix de l'échelle relèvent d'une prise d'initiative, cependant celle-ci reste limitée car les valeurs du tableau statistique sont simples à manipuler.

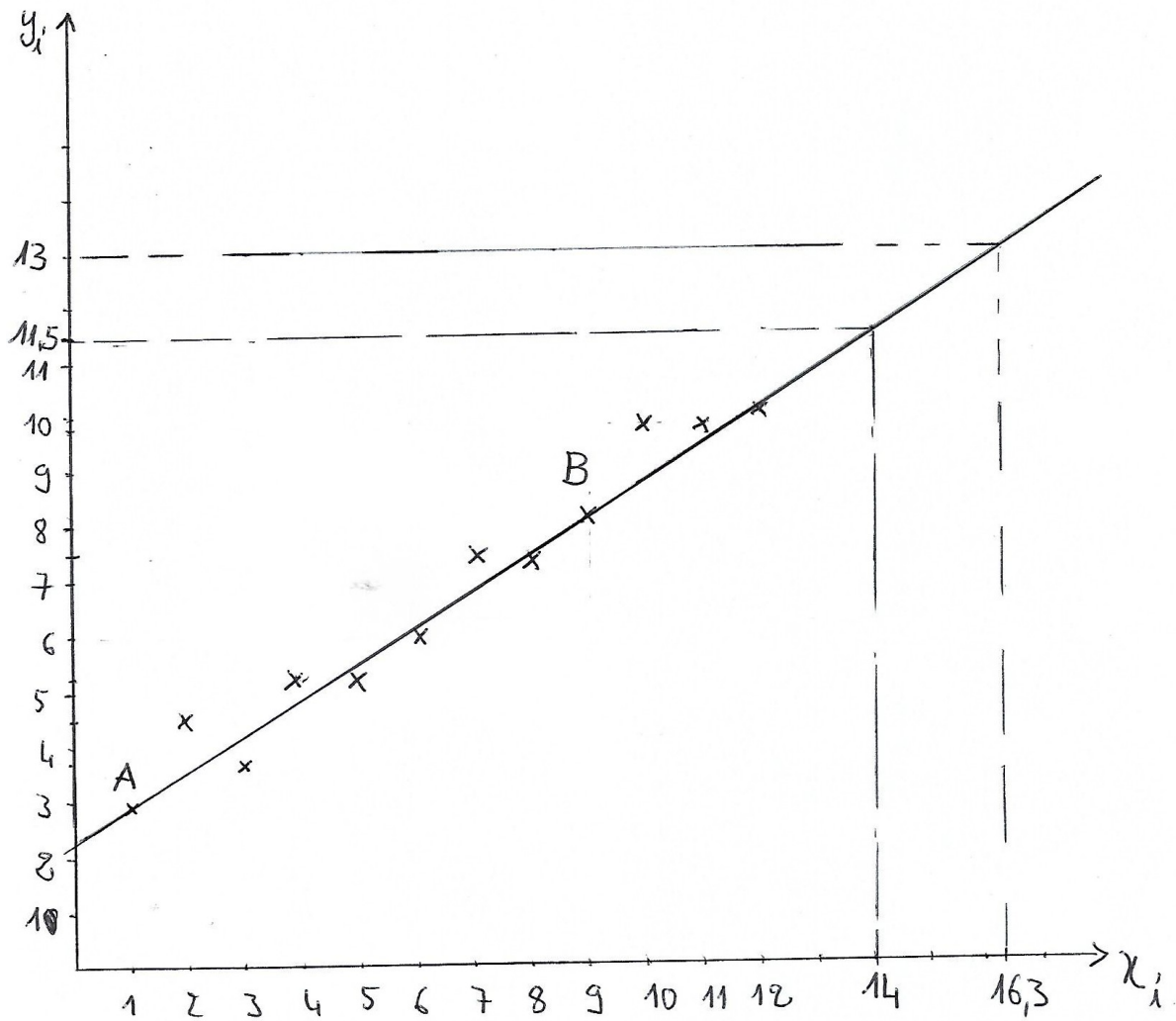


Figure 25: Représentation graphique de l'exercice 3

2. Pour tracer la droite (AB), on peut se servir du type de tâches T_Tra_DroiteG1G2. Les points A(1;3) correspond à G_1 et B(9;8,2) à G_2 . Pour le premier type de tâches qui est de tracer la

droite passant par les points $A(1;3)$ et $B(9;8,2)$. La technique consiste à placer les points A et B dans le repère puis de tracer une droite qui passe par ces deux points. Cette technique est justifiée par un discours technologique qui est le repérage des points dans le repère.

La méthode d'ajustement proposée n'est pas la méthode la plus habituelle. Pour déterminer l'équation de la droite $y=ax+b$, les élèves vont prendre des initiatives, pour trouver d'abord le coefficient directeur a en utilisant les coordonnées des points A et B: $a = \frac{y_B - y_A}{x_B - x_A} = \frac{8,2 - 3}{9 - 1} = 0,65$.

La procédure pour calculer b , consiste à considérer les coordonnées de l'un des deux points de la droite (A ou B). Puis de remplacer ces coordonnées dans l'expression de l'équation de la droite. Ces coefficients permettent d'avoir l'équation de la droite : $y=0,65x+2,35$. Pour les élèves qui vont se référer à la méthode de Mayer, peuvent se servir de T_Tra_DroiteG1G2 pour tracer la droite et T_Dét_EquatG1G2 pour déterminer l'équation réduite de cette droite, en considérant les points $A(1;3)$ comme G_1 et $B(9;8,2)$ comme G_2 . Pour obtenir l'équation d'une droite, on détermine son coefficient directeur en utilisant une propriété géométrique (deux points de la droite). Puis on détermine son ordonnée à l'origine en utilisant un point de la droite.

Donc pour le second type de tâches : déterminer l'équation réduite de la droite dont l'équation de la droite est de la forme $y=ax+b$. La technique pour déterminer le coefficient a consiste à utiliser les coordonnées des points A et B, $a = \frac{y_B - y_A}{x_B - x_A}$ et pour le coefficient, d'utiliser la relation $b = \bar{y}_B - a\bar{x}_B$ ou $b = \bar{y}_A - a\bar{x}_A$.

La troisième question correspond au types de tâches T5.

3.a- La procédure pour estimer le montant des dépenses durant le quatorzième mois, consiste à utiliser l'équation de la droite. 14 correspond au numéro du mois, on remplace $x=14$ dans l'équation de la droite : $y=0,65 \times 14 + 2,35 = 11,45$. Les élèves peuvent éventuellement confondre y et x . Ils peuvent aussi oublier d'écrire l'interprétation correspondante. Pour le premier type de tâches : estimer le montant des dépenses durant le quatorzième mois. La technique consiste à prendre $x=14$ ensuite de remplacer la valeur de x dans l'équation réduite déterminée à la question précédente, pour trouver la valeur de y . C'est cette valeur qui est le montant des dépenses.

b- Pour estimer le rang du mois, la procédure attendue, est de procéder de la même manière que la question précédente. Lorsque le montant dépassera pour la première fois 13 milliers d'euros. On prendra $y=13$, on remplace y par sa valeur dans l'équation de la droite $y=0,65x+2,35$, on a : $13=0,65x+2,35$, on a : $x=16,38$. Les élèves peuvent éventuellement confondre y et x . Ils peuvent aussi oublier d'écrire l'interprétation correspondante. Pour le deuxième type de tâches : estimer le rang du mois au cours duquel le montant dépassera pour la première fois 13 milliers d'euros. La technique est similaire à la précédente, elle consiste à prendre $y=13$ et de remplacer cette valeur dans l'équation réduite de la droite $13=0,65x+2,35$ pour trouver : $x=16,38$

Cette question 3 demande une prise d'initiative des élèves.

Le problème 3 correspond au niveau des connaissances mobilisables de Robert (2018).

Liste des critères d'analyse de productions d'élèves en 2020

Comme critères pour l'analyse des copies, nous avons les items suivants :

Problème 1

- Est-ce que les élèves disent « oui c'est correct » ou « non » ;
- La justification est- elle correcte, suffisante ?;
- Quel est le contenu de la justification : si oui « on peut faire un ajustement (à cause de la forme du nuage de points » « un graphique n'est pas suffisant » « il faut faire la méthode des moindres carrés ».

Problème 2

- Coordonnées du point moyen sont-elles justes ?/ représentation du point moyen dans le repère : oui/non On-t-il tenu compte de la pondération ?
- La formule de la covariance est-elle juste ? Si oui la valeur trouvée est correcte ?
- A propos de la mise en œuvre de la méthode des moindres carrés: application d'une technique apprise en cours / présence d'explication sur la méthode ;
- Quelle procédure, pour trouver l'estimation ?
- Valeur trouvée pour l'estimation de la note en comptabilité : la valeur est acceptable/ non acceptable ;
- Pour faire la justification, emploi du coefficient de corrélation ? (Noter si le calcul du coefficient de corrélation est juste) Emploi d'un graphique ? La justification est correcte ?

Problème 3

- Représentation du nuage de points: choix de l'échelle pertinent oui/non ;
- Représentation de la droite : acceptable/ non acceptable ;
- Équation de la droite : procédure apprise en classe oui/non ;
- Procédure pour trouver l'estimation ? Emploi de l'équation de la droite ; ou procédure graphique ;
- Valeur trouvée pour l'estimation au 14^{em} mois : la valeur est acceptable / non acceptable ;
- Valeur trouvée pour l'estimation, lorsque le montant dépassera 13 milliers d'euros: la valeur est acceptable / non acceptable ?
- Erreur dans la procédure de résolution ;
- Présence d'une justification : cette justification est- elle correcte, suffisante ?

6.2.2 Analyse des productions des élèves

Nous avons fait passer ce sujet composé de trois exercices durant l'année 2020 plus précisément le 31 Août 2020 dans une classe de terminale SE en Centrafrique. Il y avait 66 élèves présents, nous les avons organisés en 12 groupes.

6.2.2.1 Analyse globale

Problème 1

En ce qui concerne l'exercice 1, dix des douze groupes ont dit que la production de l'élève était correcte, par contre un groupe considère que la réponse n'est pas correcte et un groupe n'a pas répondu à la question. Cependant, comme nous l'avons noté dans l'analyse, certains groupes ont simplement calculé une estimation et répondu en fonction de la valeur de l'estimation trouvée, sans prendre en considération la méthode utilisée par l'élève fictif.

Nous récapitulons ces types de réponses dans le tableau ci-dessous.

Groupe N°	A pris en considération la méthode	A dit que la production était juste
Groupe 1	Oui	Oui
Groupe 2	Non	Oui
Groupe 3	Oui	Oui
Groupe 4	Oui	Oui
Groupe 5	Oui	Oui
Groupe 6	Non	Oui
Groupe 7	Non	Oui
Groupe 8	N'a pas fait l'exercice	N'a pas fait l'exercice
Groupe 9	Oui	Oui
Groupe 10	Non	Non
Groupe 11	Oui	Oui
Groupe 12	Non	Oui

Tableau 14 : Tableau récapitulatif des réponses des élèves à l'exercice 1

Nous reviendrons ci-dessous plus en détail sur les arguments utilisés par les élèves.

Problème 2

L'objectif principal du dépouillement des réponses des élèves aux différentes questions de l'exercice 2 dans cette partie, est d'analyser la nature des réponses qui peuvent être exactes, incomplètes, n'a pas répondu ou erronées, que nous avons résumé dans le tableau ci-dessous.

La première question consiste à déterminer les coordonnées du point moyen G du nuage formé par les points $M_i(x_i; y_i)$, c'est le type de tâches T_Cal_CoordG. Aucun groupe n'a trouvé la bonne réponse, ils n'ont pas pris en compte la pondération des variables x_i et y_i par des effectifs. Nous constatons que ces élèves ne maîtrisent pas les savoirs antérieurs sur la notion de la moyenne

pondérée. Dans le programme officiel en vigueur en Centrafrique (programme 2002), une section est consacrée à la statistique en classe de 4^{ème}. Dans la catégorie « *savoirs, savoir-faire* », c'est prescrit:

- Effectuer des calculs sur des pourcentages ;
- construire un graphique et calculer une moyenne pondérée par des effectifs avec un tableau ;
- calculer la moyenne et la moyenne pondérée par les effectifs avec une calculatrice ;
- déterminer le pourcentage relatif à un caractère d'un groupe constitué de la réunion de deux groupes dont les effectifs et les pourcentages relatifs à ce caractère sont connus.

Et dans la catégorie « *Commentaires et instructions* », on suggère que :

- *Les élèves sont confrontés à des situations familières où deux procédés de calcul différents de la moyenne sont mis en œuvre : somme des n données divisée par n , moyenne pondérée des valeurs par leurs effectifs.*
- Pour des petits effectifs, les élèves doivent savoir calculer une moyenne par la procédure de leur choix.
- Pour des grands effectifs, l'utilisation du tableur ou de la calculatrice est recommandée.

Les recommandations et les suggestions du programme de la classe de 4^{ème} insistent sur l'aspect calculatoire de la statistique descriptive à une variable. Par ailleurs, la notion de la moyenne pondérée est revue en classe de première. Cependant, c'est rare que les élèves revoient cette notion en terminale SE, c'est la pratique des enseignants de terminale. Force est de constater que ces élèves ont des difficultés avec la notion de la moyenne pondérée. Dans la pratique de cette notion dans l'exercice 2, les élèves n'ont pas pris en compte la pondération par des effectifs de cette série.

Questions	Réponse exacte	Réponse incomplète	Réponse erronée	N'a pas répondu
1	0	0	12	0
2	0	1	9	2
3	0	1	7	4
4	0	0	5	7

Tableau 15: Réponses des 12 groupes à l'exercice 2

Concernant la deuxième question qui consiste à calculer la covariance de la série double, T_Cal_Cov. Neuf groupes d'élèves ont trouvé des réponses erronées, deux groupes n'ont pas abordé la question, et un groupe a donné une réponse incomplète.

Pour la troisième question, qui consiste à déterminer l'équation de la droite de régression par la méthode des moindres carrés: T_Dét_EquatMCyx, six groupes ont trouvé des réponses erronées, deux groupes ont donné des réponses incomplètes et quatre groupes n'ont pas abordé la question 3. Pour ce type de tâches la méthode des moindres carrés est apprise en cours, les groupes qui ont réussi la procédure ont fait des erreurs de calculs et aucun groupe n'a trouvé la réponse exacte. Nous constatons que l'erreur de calcul est très récurrente, cela se justifie par le fait qu'un bon nombre d'élèves n'ont pas facilement accès à la calculatrice, faute de moyen pour s'en procurer. Nous faisons l'hypothèse que la cause de leur erreur de calcul pour les groupes qui ont au moins une calculatrice est qu'ils ne s'exercent pas assez avec la calculatrice.

La quatrième question consiste à faire une estimation : T5. Cinq groupes d'élèves ont trouvé une réponse erronée dont un seul groupe a proposé un modèle adéquat et sept groupes d'élèves n'ont pas répondu à la question. Nous rappelons que l'estimation est déduite à partir de l'équation de la droite, comme aucun groupe n'a pu trouver la bonne équation de la droite de régression de y en x , cela va de soi qu'aucun groupe n'a trouvé une estimation juste. Mais un groupe a appliqué le principe d'estimation, il a proposé une estimation qui est cohérente, avec l'équation trouvée.

Problème 3

Pour la première question, neuf groupes ont fait une représentation correcte du nuage des points et trois groupes ne l'ont pas fait. Nous constatons que la première question a été bien réussie par rapport aux autres questions de l'exercice 3. Nous faisons l'hypothèse que ces élèves ont acquis le sens de la représentation des coordonnées de points dans un repère au collège (programme 2017), et en plus les données de l'exercice 3 sont faciles à manipuler.

La deuxième question consiste à tracer une droite passant par deux points, puis de déterminer l'équation réduite de cette droite. Trois groupes ont tracé la droite qui passe par les points A et B. Mais l'équation de la droite trouvée est erronée. Par contre, trois groupes ont tracé la droite mais n'ont pas fini de déterminer l'équation de la droite. Donc ils ont donné une réponse incomplète. Et quatre autres groupes ont tracé la droite mais n'ont pas déterminé l'équation de la droite.

Questions	Réponse juste	Réponse fausse	Réponse incomplète	N'a pas répondu à la question
1	9	0	0	3
2	0	3	6	3
3.a	0	2	0	10
3.b	0	1	0	11

Tableau 16: Réponses des 12 groupes à l'exercice 3

Nous remarquons que quatre groupes ont tracé le segment [AB] plutôt que la droite (AB) (le trait tracé ne dépasse pas des points A et B).

Pour la troisième question, qui consiste à utiliser l'équation de la droite pour faire des estimations : T5. Pour 3.a) deux groupes ont fait l'estimation qui est erronée et leur justification est fausse. Par contre dix autres groupes n'ont pas répondu à la question, pour 3.b) un groupe a trouvé une réponse fausse et les onze autres groupes n'ont pas répondu à la question.

6.2.2.2 *Analyse détaillée*

Problème 1

Le groupe 1 considère que la réponse est correcte. Pour justifier leur réponse, ils parlent de la forme du nuage de points qui justifie un ajustement affine. Ils ne discutent pas la méthode utilisée par l'élève fictif pour tracer cette droite. En se basant sur la représentation graphique, ils ont donné un argument juste. Ils ont des prises d'initiative dans l'interprétation graphique.

Les élèves du groupe 2 affirment que la réponse est juste, mais ils ne tiennent pas compte de la méthode employée par l'élève fictif. Ils ont donné une justification qui n'est pas correcte, en se servant du tableau statistique de l'énoncé. Pour eux, le montant des remboursements croissent en fonction de l'année, alors au rang 7, c'est-à-dire en 2019 le montant de remboursements y_i dépassera 10000. Ces élèves n'ont pas fait allusion au graphique dans leur réponse, ils ne se sont pas intéressés à la méthode utilisée par l'élève fictif, mais seulement à la valeur numérique de l'estimation. Nous rappelons que cet exercice vise à interpréter la procédure graphique de la solution proposée. Force est de constater que ce groupe n'a pas tenu compte de ce qui leur a été demandé.

De même, les élèves du groupe 3 ne parlent pas de la méthode employée par l'élève fictif. Cependant ils se basent d'une part sur le tableau statistique et d'autre part, ils se servent du graphique, pour affirmer que la prévision du ménage est fausse. Comme ils se servent du graphique fourni, nous considérons qu'ils trouvent que la méthode de l'élève fictif est correcte : ils ont donc une prise d'initiative.

Pour le groupe 4, les élèves de ce groupe approuvent la réponse et proposent une justification juste. Ils constatent que d'après la droite d'ajustement tracée, l'estimation de la dépense va dépasser 10000. La justification de ces élèves est cohérente avec la réponse que nous avons prévue dans l'analyse a priori. Ils ont la prise d'initiative dans l'interprétation graphique des résultats : ils disent qu'il est justifié de faire un ajustement affine, et ne remettent pas en cause la méthode utilisée par l'élève fictif.

Le groupe 5 considère que la réponse est juste, les élèves de ce groupe ont fait une observation pertinente. Ils s'appuient sur la représentation graphique pour justifier leur réponse, ces élèves ont la capacité de faire l'analyse d'un graphique. Ils ont explicité leur réponse, en se basant sur la lecture et l'interprétation graphique, ils argumentent bien leur réponse. Ils ont utilisé le graphique, et la droite tracée par l'élève fictif. Ils ont posé $y=10000$ et ils ont cherché quelle date correspondait à la

valeur 10000. Ils constatent que, cette date est avant 2019. Ils ne remettent pas en cause la méthode employée par l'élève fictif, cependant ils éprouvent le besoin de faire une autre justification. De notre point de vue, nous pensons que ces élèves ont des initiatives pour lire et interpréter des résultats obtenus.

Exercice 11

Quel nous jugeons correcte cette affirmation.
Que la prévision du ménage est mauvaise parce
que si l'on fait la projection sur le graphique
(nous projetons le montant des remboursements
 $y_i = 10000$ sur la droite d'ajustement nous
constatons que ce dernier n'atteint pas 2019 sur
l'axe ou rang de l'année x_i .

Figure 26: Extrait de copie du groupe 5

Concernant le groupe 6. Les élèves de ce groupe disent que la réponse est correcte. Ils se justifient sur la base du tableau statistique de l'énoncé. Pour eux comme le montant des remboursements croissent et atteignent 9120 en l'an 2018. Ils déduisent qu'en 2019, en tenant compte de l'écart des montants y_i entre les années, le montant dépassera 10000. Ces élèves ont pris l'initiative de raisonner sur les données du tableau. Cependant ils ne se sont pas appuyés sur la représentation graphique, leur argument ne fait pas allusion au graphique. Ils cherchent à savoir si le résultat de l'élève fictif est juste, et ne parlent pas de la méthode employée.

Concernant le groupe 7, les élèves du groupe considèrent que la réponse est juste. Cependant leur justification n'est pas correcte, la réponse donnée par ces élèves n'est pas cohérente avec la question posée, ils n'ont pas de prise d'initiative. Nous supposons qu'ils n'ont pas compris l'exercice, ils n'ont même pas fait référence à la représentation graphique.

Le groupe 8, les élèves de ce groupe n'ont pas fait l'exercice.

Les élèves du groupe 9 considèrent que la réponse est juste. Ils constatent sur le graphique que la projection du point de la droite d'ajustement dont l'abscisse est l'année 7, sur l'axe des ordonnées dépassera 10000. Ces élèves ont donné une justification juste et cohérente, en analysant la représentation graphique.

Le groupe 10 considère que la réponse n'est pas juste. La justification proposée par ce groupe ne fait pas allusion à la représentation graphique, leur réponse manque de clarté.

Concernant le groupe 11, les élèves considèrent que la réponse est correcte. Ils justifient que la représentation graphique correspond au tableau et constatent que, graphiquement le point de l'année

7 dépassera le point y_i qui correspond au montant 10000. Ces élèves ont fait le lien entre les deux registres de représentation utilisés dans cet exercice, le registre de représentation de tableau et le registre de représentation graphique.

Le groupe 12 justifie que la réponse est juste, les élèves de ce groupe proposent une justification qui est insuffisante, en se basant sur la croissance des valeurs du tableau statistique, sans faire allusion à la représentation graphique. Or l'analyse des données du tableau n'est pas suffisante pour tirer une conclusion sur la prévision du ménage. Pour une bonne prévision, il faut changer de registre. Par exemple, passer du registre tableau au registre graphique pour pouvoir effectuer des estimations avec une droite d'ajustement.

Nous observons une prise d'initiatives dans les justifications des groupes 1 ; 3 ; 4 ; 5 ; 6 ; 9 ; 12. Ces groupes ont données des justifications basées soit sur la droite d'ajustement affine, soit le graphique. Par contre, les groupes 7 ; 10 ; 11 n'ont pas de prise d'initiative dans leur justifications. Ces derniers ont donnés des justifications en se basant, soit sur le tableau statistique, soit ils n'ont pas du tout répondu à la question.

Problème 2

Le type de tâches proposé aux élèves dans la question N°1 : T_Cal_CoordG, consiste à déterminer les coordonnées du point moyen G du nuage.

Dans l'analyse a priori nous avons expliqué que, pour déterminer les coordonnées du point moyen G du nuage des points, la procédure attendue est un calcul de la moyenne de chacune des deux séries statistiques, avec une prise en compte de la pondération des notes (x_i et y_i).

Le groupe 1, groupe 5, et groupe 8 ont trouvé la même réponse erronée qui est : $G(1,31;1,31)$. Ces élèves n'ont pas pris en compte la pondération des notes, mais ils ont bien fait référence à l'effectif total de 100. Ils ont utilisé la moyenne arithmétique plutôt que la moyenne pondérée, donc la procédure n'est pas juste ainsi que les coordonnées du point moyen trouvées.

Les groupes 1 et 8 n'ont pas fait une représentation graphique. Par contre, les élèves du groupe 5 ont fait la représentation graphique du nuage des points et placé le point moyen G. Nous constatons que le point moyen G est beaucoup plus à l'écart du nuage des points, et les élèves de ce groupe 5 n'ont pas pris l'initiative de vérifier leur calcul ou de faire un lien entre les données numériques et leur représentation graphique. Nous supposons que ces élèves n'ont pas acquis le sens de ce qu'était un point moyen. Par ailleurs dans le calcul des coordonnées du point moyen G, la technique consiste à calculer la moyenne pondérée de ces deux variables x_i et y_i . Par exemple, pour calculer la moyenne pondérée \bar{x} , les valeurs (note en mathématiques) de la série statistique x_i varient de 7 à 15, pondérées des effectifs. Alors on s'attendait à voir la valeur de la moyenne dans l'intervalle de 7 à 15. Mais ces élèves ont trouvé la valeur de la moyenne 1,31. Donc, nous remarquons que les élèves de ce groupe n'ont pas d'initiative dans le contrôle de l'ordre de grandeur du résultat. Ils n'ont pas de compétence en mathématique générale sur l'ordre de grandeur. Nous rappelons aussi que la notion d'ordre de grandeur est enseignée au collège en Centrafrique. Du point de vue du travail de groupe, nous constatons que, quand deux élèves du même groupe trouvent une même réponse de calcul, alors les autres élèves du groupe valident cette réponse. Nous l'avons observé

pendant la séance. Cette validation sur la base de deux résultats identiques trouvée est plus importante pour les élèves que la validation ou l'invalidation par le contrôle de l'ordre de grandeur du résultat.

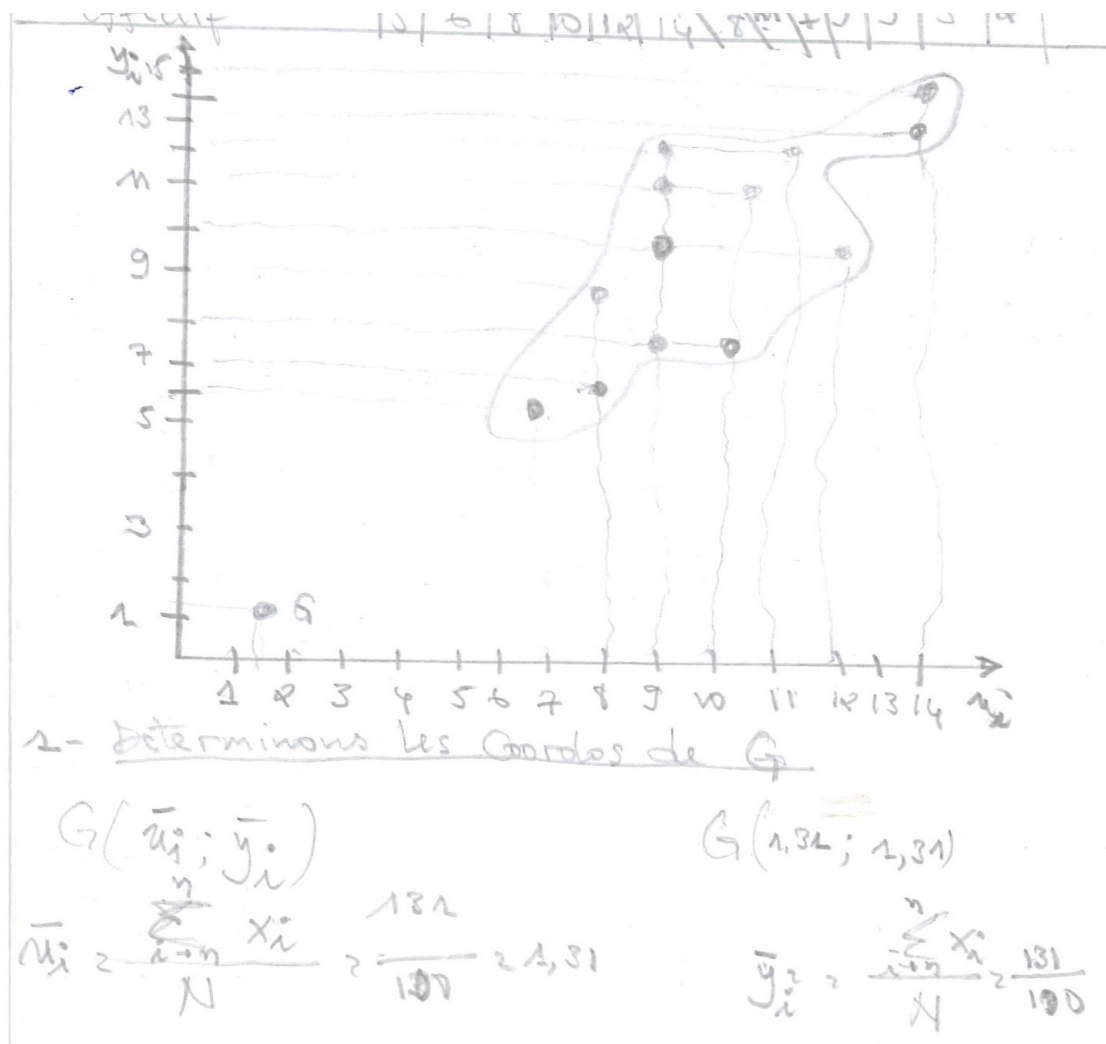


Figure 27: Extrait de copie du groupe 5

Pour les groupes, 2 ; 3 ; 4 ; 7 et 10. Ils ont trouvé la même réponse erronée : $G(10,07; 10,07)$ et $G(10; 10)$ pour le groupe 4. Du fait que ces groupes n'ont pas pris en compte la pondération, et de plus ils ont trouvé la valeur de l'effectif total $N=13$ au lieu de $N=100$. Pour ces élèves, l'effectif total est le nombre des couples de notes en Mathématiques et en Comptabilité que l'on a regroupé dans le tableau. Ils n'ont pas compris le sens du tableau statistique avec pondération. Nous faisons l'hypothèse que la difficulté de ces élèves réside dans le passage de la compréhension du tableau statistique à l'application du calcul à effectuer sur les données du tableau. En plus, ils n'ont pas placé le point moyen dans le repère.

Concernant les groupes 3 et 7, ils ont fait un tableau statistique sur les calculs élémentaires, sur les variables x_i , y_i et l'effectif. Dans ce tableau les élèves ont trouvé l'effectif total 100. Force est de constater que dans la détermination des coordonnées du point moyen G, ils ont utilisé le nombre des

couples de notes qui est $N=13$ sans tenir compte de la pondération. Ce qui confirme notre hypothèse que ces élèves ont une difficulté liée à la compréhension du tableau statistique avec pondération. C'est un effet du contrat didactique, ceci peut s'expliquer par le fait que, dans le contrat didactique habituel en terminale, les élèves ne rencontrent presque aucun tableau avec pondération. Donc ils ne peuvent pas faire l'analogie d'un exemple traité en classe avec cet exercice.

Exercice 2

x_i	7	8	8	9	9	9	9	10	10	11	12	14	15	131
y_i	6	7	9	8	10	11	12	8	11	12	10	13	14	131
Eff	5	6	8	10	12	14	8	11	7	9	3	5	2	100
x_i^2	49	64	64	81	81	81	81	100	100	121	144	196	225	1387
y_i^2	36	49	81	64	100	121	144	64	121	144	100	169	196	1389
$x_i y_i$	42	56	72	72	90	99	108	80	110	132	120	182	220	1378

1) Déterminons le point moyen du nuage : $G(x_i, y_i)$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \rightarrow \frac{1}{13} (131) \rightarrow \bar{x} = \frac{131}{13} = 10,07$$

$$\bar{x} = 10,07$$

$$\bar{y} = \frac{1}{13} (131) = 10,07 \rightarrow \bar{y} = 10,07$$

le point $G(10,07; 10,07)$

Figure 28: Extrait de copie du groupe 3

Quant aux élèves du groupe 6, leur procédure est juste, ils ont utilisé la formule de la moyenne pondérée pour déterminer les coordonnées du point moyen mais ils ont trouvé une réponse erronée, car ils ont fait une erreur de calcul au niveau de l'effectif total. Par ailleurs les élèves de ce groupe n'ont pas représenté le point moyen G sur un repère, cela s'explique par les valeurs des coordonnées du point moyen trouvées. Vu les valeurs trouvées : $\bar{x}=56,35$ et $\bar{y}=58,23$, et l'écart qu'il y a entre ces coordonnées et les valeurs des variables x_i et y_i , ils auront des difficultés à choisir l'échelle convenable pour ces données. Là aussi, il n'y a pas de contrôle de l'ordre de grandeur du résultat.

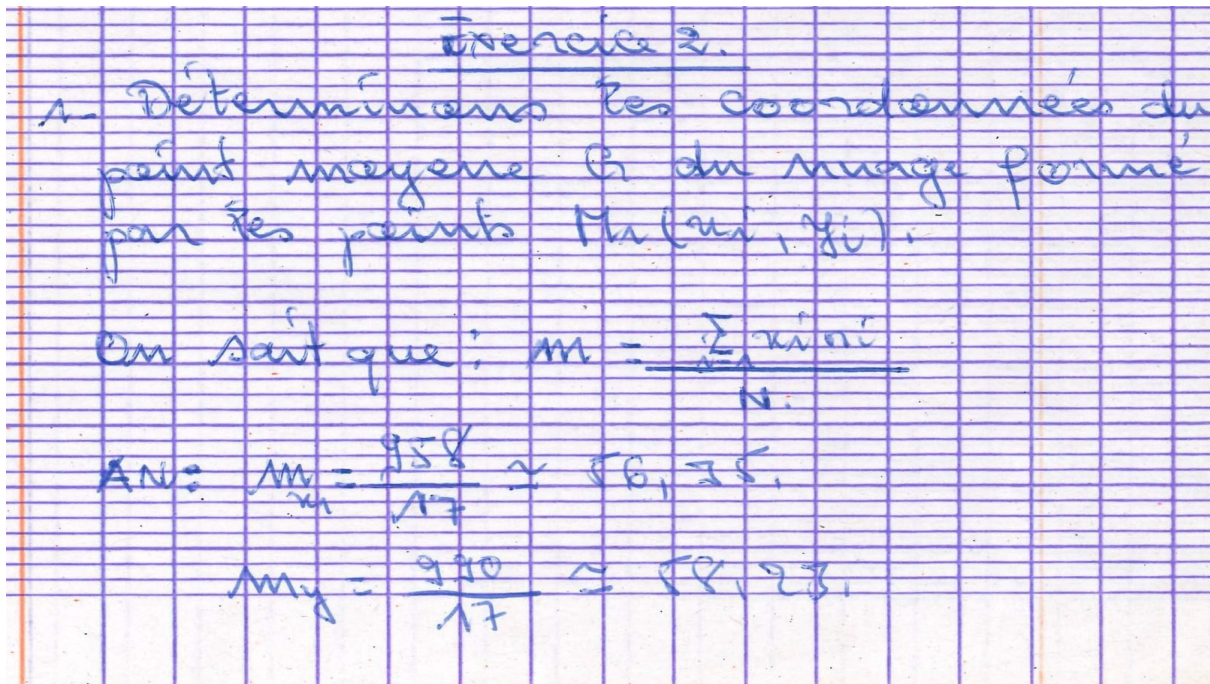


Figure 29: Extrait de copie du groupe 6

Le groupe 11 a proposé la formule de la moyenne pondérée mais dans leur calcul, ils n'ont pas pris en compte la pondération. Et l'effectif total trouvé n'est pas juste. Du point de vue du savoir-faire, la mise en application de la formule de moyenne pondérée pose problème avec ces élèves. Ils ont fait la représentation graphique du nuage de points et du point moyen G, c'est-à-dire, qu'ils ont réalisé le type de tâches que nous avons nommé T1.

Pour le groupe 9, les élèves ont trouvé une réponse erronée, ils n'ont pas donné le détail de leur procédure de calcul. Ils proposent juste la réponse finale qui n'est pas correcte.

Quant au dernier groupe, le groupe 12 ils ont scindé le tableau des notes de mathématiques et comptabilité en deux tableaux séparés. Le premier tableau représente les notes en mathématiques et le deuxième tableau les notes en comptabilité, suivi de leurs effectifs. De notre point de vue, ce groupe n'a pas compris la question posée .

Nous retenons que l'analyse des réponses des élèves relatives à la question N°1, montre que ces groupes d'élèves n'ont pas de prise d'initiatives dans l'interprétation d'un tableau statistique. La détermination des coordonnées du point moyen G du nuage formé par les couples de points relève d'une prise d'initiative par rapport au contrat habituel, pour noter qu'il y a des effectifs et qu'il faut en tenir compte. Mais ces élèves ont tous commis des erreurs dans la réalisation de cette tâche. Ces erreurs sont liées soit à la compréhension du tableau statistique, ou soit aux calculs.

Le type de tâches de la question N°2, consiste à calculer la covariance de la série.

Les élèves du groupe 5 et du groupe 11 ont procédé par un tableau pour calculer la covariance. Mais ces derniers n'ont pas pris en compte la pondération de chaque note dans le tableau statistique qu'ils ont établi. Par ailleurs la formule de la covariance utilisée est juste, mais le résultat trouvé est erroné. Nous constatons que ces élèves ne s'exercent pas assez sur l'aspect calculatoire. Ils arrivent

à poser les bonnes formules, mais leur aptitude de faire des calculs sur la notion de somme des termes finis pose problème.

Concernant les groupes 4 ; 7 et 10. Les élèves de ces groupes ont proposé des formules erronées, ils n'ont pas tenu compte de la pondération dans leur formule. Le groupe 4 a utilisé la formule explicite de la covariance (formule sous la forme développée) et les groupes 7 et 10 ont utilisé la formule implicite (formule sous la forme factorisée). Au niveau de l'application numérique ils ont trouvé des résultats erronés.

Les groupes 1 et 2 ont proposé des formules erronées et leurs résultats sont faux. En analysant leur copie, il nous semble que ces élèves n'ont pas compris le sens de l'expression de somme de termes consécutifs, ils confondent la somme des produits et le produit des sommes. Ils expriment toutes ces variables en fonction du symbole de la somme.

$$\begin{aligned} \text{Cov}(x, y) &= \frac{1}{n} \sum_{i=1}^n (x_i y_i - \bar{x} \bar{y}) \\ &= \frac{1}{100} [(131 \times 131) - (1,31 \times 1,31)] = \\ &= \frac{1}{100} [(131)^2 - (1,31)^2] \\ \text{Cov}(x, y) &= 171,6 \end{aligned}$$

Figure 30: Extrait de copie du groupe 1

Le groupe 6 a donné une réponse incomplète, il a juste proposé la formule qui est fautive.

Pour le groupe 9, les élèves de ce groupe ont donné leur réponse sans avoir présenté le détail de leur calcul. Et cette réponse n'est pas juste.

Pour le groupe 3, les élèves de ce groupe ont fait une autre tâche à la place de la tâche demandée. Ils ont déterminé la variance de la variable x , au lieu de déterminer la covariance de la série double.

Quant aux groupes 8 et 12, les élèves de ces deux groupes n'ont pas répondu à la question.

La question N°2 ne relève pas d'une prise d'initiatives. Le constat est qu'aucun groupe n'a pu réussir à trouver une réponse juste.

Pour la question N°3, le type de tâches est de déterminer une équation de la droite de régression de y en x , par la méthode des moindres carrés.

La procédure utilisée par les groupes 1 ; 2 ; 4 ; 5 ; 7 ; 9 et 11 est juste, c'est une méthode déjà travaillée en classe et qui fait partie du contrat. Cela ne relève pas d'une prise d'initiative, les élèves peuvent répondre par effet du contrat didactique sans initiative, c'est-à-dire, cette solution peut être obtenue par répétition de la méthode vue en classe. Certes, les élèves ont commis des erreurs de

calcul au niveau de la détermination des coefficients a et b . Seul, le groupe 4 a donné une explication sur son équation de la droite trouvée. Ce groupe a trouvé l'équation de la droite $y=x$ et précise que c'est la première bissectrice, c'est-à-dire la droite qui partage les premier et troisième quadrants en deux dans un repère cartésien. Pour les autres groupes, ils n'ont pas donné des explications sur la méthode utilisée.

Le groupe 3 a donné une réponse incomplète. Les élèves de ce groupe ont utilisé une procédure qui est juste, ils ont déterminé la valeur des coefficients a et b qui sont des réponses erronées. Mais ces derniers n'ont pas écrit l'expression de l'équation de la droite de régression de y en x .

Les groupes 6 ; 8 ; 10 et 12 n'ont pas déterminé l'équation de la droite de régression de y en x .

Concernant la question N°4, dont les types de tâches consistent à faire une estimation et de justifier la réponse trouvée. Nous rappelons qu'aucun des groupes n'a trouvé l'équation de la droite de régression de y en x .

Les groupes 1 ; 2 ; 3 ; 4 et 7 ont répondu à la question. Le groupe 1 a proposé une estimation qui est cohérente, les élèves de ce groupe ont utilisé une procédure qui est juste, en proposant un modèle adéquat pour faire l'estimation même si l'équation de la droite de régression de y en x trouvée n'est pas correcte. En prenant $x=13$ comme note de mathématique pour déterminer la valeur de y comme note de comptabilité. Cela relève d'une prise d'initiative de la part de ces élèves. Concernant le groupe 3, les élèves affirment qu'on ne peut pas estimer la note de comptabilité. Pour justifier leur propos, ils se basent sur le tableau statistique sans tenir compte de l'équation de la droite. Pour eux, il n'y a pas de note $x=13$ dans le tableau donc on ne peut pas estimer la note de comptabilité. Ils n'ont pas pu faire le lien entre la variable y et la variable x en utilisant l'équation de la droite de régression, c'est-à-dire que la connaissance de la valeur de x (note de mathématique) permet de prévoir la valeur de y (note de comptabilité). Nous rappelons aussi que l'une des difficultés de ce groupe est qu'ils ont déterminé la valeur des coefficients a et b sans établir l'expression de l'équation de la droite de régression de y en x . Donc sans l'équation de la droite, ils ne peuvent pas faire l'estimation algébriquement. Pour le groupe 7, les élèves affirment que ce n'est pas possible d'estimer la note de comptabilité. Pour eux, c'est incertain de savoir réellement si l'élève est aussi bon en mathématique qu'en comptabilité. Ces élèves ont donné une justification fautive du fait qu'ils n'ont pas de prise d'initiatives de se servir de leur équation de la droite pour estimer la note de comptabilité.

Le groupe 4 affirme qu'en se servant de la droite de régression trouvée, on peut estimer la note de comptabilité. Cependant, ils n'ont pas utilisé la droite trouvée pour faire l'estimation demandée.

Pour le groupe 2, les élèves se basent sur le tableau statistique pour justifier leur estimation. Pour eux, la note de mathématique se trouve dans l'intervalle $[12 ; 14[$ donc cela implique que la note de comptabilité se trouve dans l'intervalle $[10 ; 13]$. De notre point de vue, cette justification n'est pas suffisante, la méthode ne donne pas une bonne estimation.

Les autres groupes n'ont pas abordé la question.

Problème 3

La question N°1 relève du type de tâches T_Rep_Nuage: représenter le nuage de points dans un repère. Pour ce type de tâches, le choix de l'échelle ne pose pas de difficulté, du fait que les valeurs du tableau statistique sont simples à représenter. Les groupes 1 ; 2 ; 3 ; 4 ; 5 ; 6 ; 7 ; 8 et 9 ont bien représenté le nuage des points. Les valeurs du tableau sont simples et faciles à manipuler, donc leur représentation dans le repère orthogonal ne relève pas d'une prise d'initiatives. Les groupes 10 ; 11 et 12 n'ont pas fait l'exercice.

Pour la question N°2, le premier type de tâches est de tracer une droite passant par les points $A(1;3)$ et $B(9;8,2)$. Les élèves des groupes 3 ; 5 ; 6 ; 8 et 9 ont fait une représentation fautive. Nous remarquons que ces derniers n'ont pas pu faire la différence entre une droite (AB) et un segment $[AB]$. Dans leur représentation graphique, nous constatons une confusion entre ces deux notions. Par définition : « la droite (AB) est la ligne droite qui passe par les points A et B, elle est illimitée de chaque côté » et « le segment $[AB]$ est une partie de la droite (AB) limitée par deux extrémités, les points A et B ». Les représentations graphiques associées à chacune de ces deux notions font partie du contrat didactique établi depuis le collège, où ils ont appris à identifier une droite, une demi-droite, un segment et à tracer à l'aide d'une règle une droite, une demi-droite et un segment d'extrémités données (Programme 2017). Dans la pratique d'enseignement en Centrafrique, quand on demande aux élèves par exemple de tracer une droite passant par deux points sur une feuille de cahier, cela consiste à tracer la droite de sorte qu'elle dépasse un peu ces deux points.

Quant aux groupes 1 ; 2 ; 4 et 7, les élèves de ces groupes ont tracé une droite qui est juste.

Dans cet exercice, les élèves qui ont tracé le segment $[AB]$ ne peuvent pas utiliser ce segment pour faire une estimation en 14. Ils ne peuvent pas faire une estimation graphique, du fait que le segment qu'ils ont tracé ne dépasse pas le nuage des points. Par ailleurs, ces élèves ont bien représenté les coordonnées des points A et B dans le repère orthogonal. Dans le choix de l'origine du repère, le groupe 2 a choisi comme origine de leur repère les coordonnées $(1;3)$.

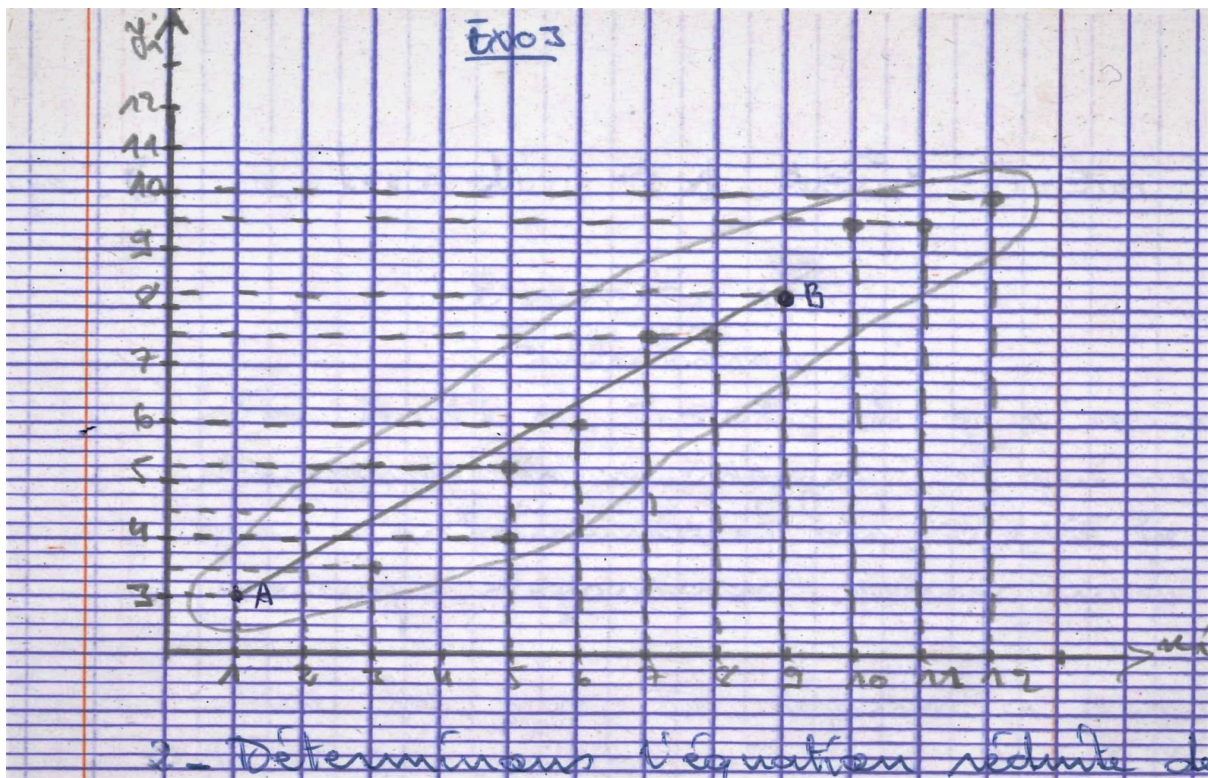


Figure 31: Extrait de copie du groupe 6

Pour le deuxième type de tâches de la question N°2, qui consiste à déterminer l'équation de cette droite (AB), la procédure attendue est de passer par des calculs intermédiaires pour déterminer les coefficients a et b . Cependant, les groupes 5 et 7 ont donné des réponses incomplètes, en utilisant la méthode des moindres carrés. Or la procédure attendue est d'utiliser les coordonnées des points A et B pour déterminer les coefficients a et b de l'équation de la droite. Le groupe 7 se base sur la méthode (moindres carrés) apprise en classe pour déterminer l'équation de la droite. Or on donne deux points A et B et on demande l'équation de la droite (AB). Ce groupe ne répond pas à cette question. Il détermine l'équation de la droite des moindres carrés à la place. C'est un effet de contrat typique. Les élèves du groupe 7 n'ont pas mis en équation sous la forme $y=ax+b$.

Concernant les groupes 2 et 4, ils ont déterminé l'équation de la droite, l'équation trouvée n'est pas juste. Ils ne se sont pas servis des coordonnées des points A et B pour déterminer l'équation de la droite, ils ont utilisé plutôt la méthode des moindres carrés apprise en classe. Comme le groupe 7, le contrat didactique habituel pour déterminer l'équation de la droite a été rompu, alors ces groupes ont des difficultés à progresser dans la construction du savoir. Ils n'ont pas de prise d'initiatives.

Les groupes 1 ; 3 ; 6 ; 8 ; 9 ; 10 ; 11 ; 12 n'ont pas répondu à ce type de tâche.

Concernant la question N°3, qui consiste à faire des estimations, seuls deux groupes ont proposé une estimation. Pour le premier type de tâches T5 (estimer le montant des dépenses durant le quatorzième mois) de la question N°3. Le groupe 1, propose une estimation qui n'est pas juste, les élèves de ce groupe affirment que le montant estimé est 11,7 millions. Nous rappelons que ce

groupe n'a pas trouvé l'équation de la droite qui permettra d'estimer le montant des dépenses. Mais

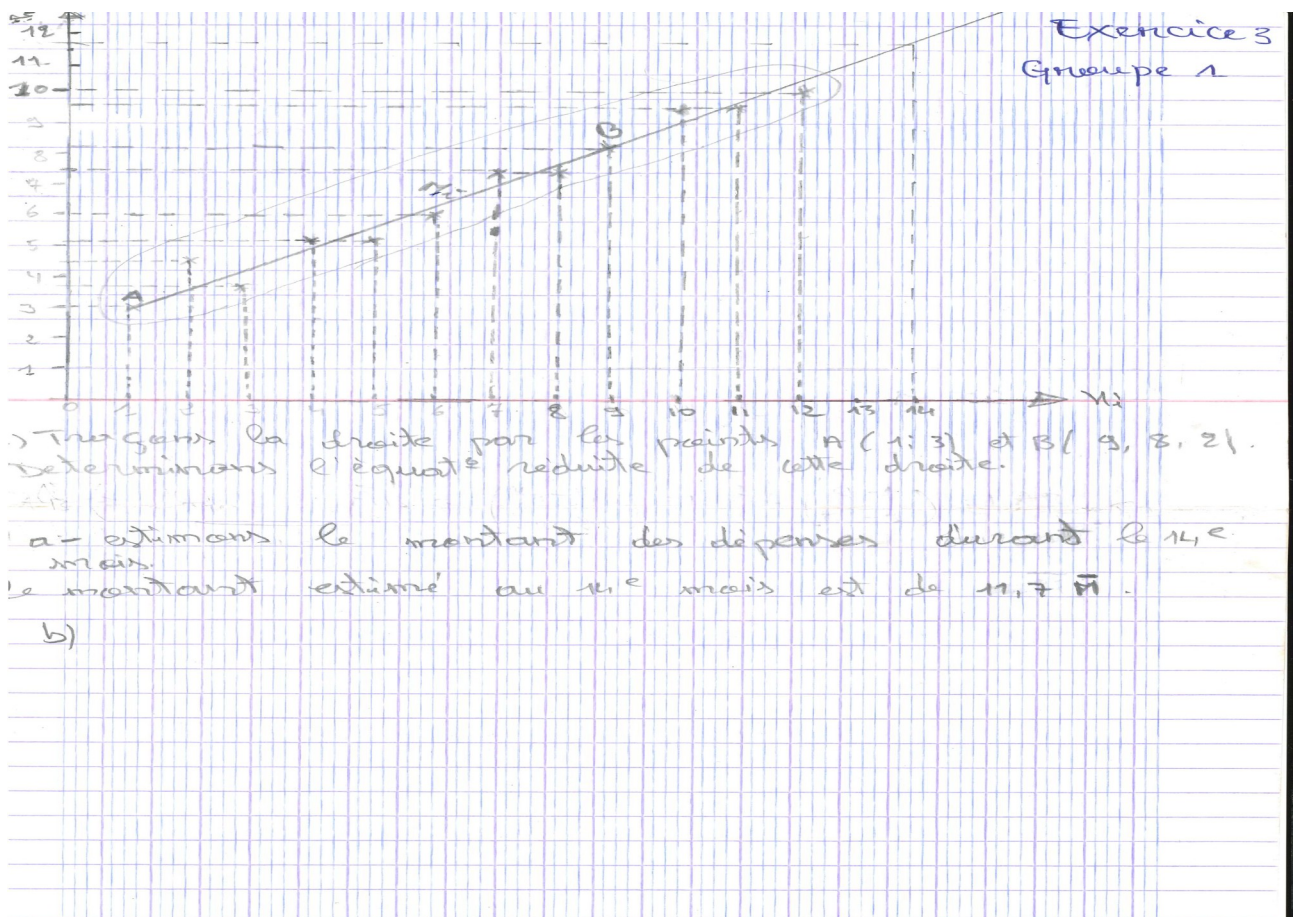


Figure 32: Extrait de copie du groupe 1

ils ont procédé en utilisant la droite (AB), avec une lecture graphique. On voit qu'ils ont représenté la valeur 14 sur l'axe des x , et ils ont bien mis des pointillés pour indiquer où est la valeur correspondant à 14. Cette procédure graphique nécessite une prise d'initiatives importante.

Quant au groupe 2, ils ont procédé de la même manière que le groupe précédent. Pour eux, le montant des dépenses dans le tableau croît et atteint 10 au douzième mois, donc au quatorzième mois le montant atteindra 11,7. On ne sait pas d'où provient cette valeur pour ce groupe, car ils n'ont pas placé 14 sur leur graphique.

Pour le deuxième type de tâches de la question 3. Seul le groupe 2 qui a proposé une réponse. Mais cette réponse est fautive. Pour ces élèves le rang estimé sera dans l'intervalle [14 ; 15]. Et cet intervalle ne figure pas sur la représentation graphique.

Concernant la procédure pour faire une estimation, nous dirons que la plupart de ces élèves n'ont pas de prise d'initiatives dans la manière de faire. Les élèves du groupe 1 se sont appuyés sur la méthode graphique utilisée dans l'exercice 1 pour répondre à la question.

Nous avons relevé quelques difficultés des élèves à travers l'analyse de leurs copies. Nous avons remarqué que dans la résolution de l'exercice 1, aucun groupe n'a remis en cause la méthode de l'élève fictif ; mais également peu de groupes ont spontanément parlé du fait qu'un ajustement affine est justifié (2 groupes). Concernant l'exercice 2, la majorité des élèves n'ont pas tenu compte de la pondération . Nous faisons l'hypothèse que l'erreur des élèves sur la notion de la moyenne

pondérée provient du choix didactique de l'enseignant, le fait de ne pas travailler régulièrement les moyennes pondérées.

Certains choix didactiques de l'enseignant peuvent avoir des conséquences sur les conceptions des élèves et notamment renforcer les obstacles culturels et didactiques... (Berté et al, 2013, p.13).

A titre de rappel la notion de la moyenne arithmétique et de la moyenne pondérée ont été apprises en classe de première. Nous constatons que les élèves confondent ces deux notions, si l'enseignant avait fait le choix didactique de rappeler ces deux notions pendant son cours de la série statistique à deux variables en terminale SE, ces erreurs pourraient être évitées.

Par ailleurs, ce qui attire aussi notre attention c'est l'erreur de calculs. Les élèves commettent très souvent des erreurs de calculs. Ils ne prennent pas du recul pour repérer leurs erreurs de calcul et les corriger. Le mode de validation du travail de certains groupes, c'est de voir si ils sont plusieurs à avoir le même résultat. Ils ne s'appuient pas sur les ordres de grandeur.

L'enseignement pourrait remédier à cela, en demandant régulièrement aux élèves de vérifier les ordres de grandeur des résultats pour qu'ils en prennent l'habitude.

Par ailleurs au moins certains élèves n'ont pas accès au sens du concept de point moyen, puisqu'il placent le point moyen en dehors du nuage de points (suite à une erreur de calcul) sans que cela leur pose problème.

Nous avons conçu ces exercices de manière à laisser aux élèves des possibilités de prise d'initiatives. Notre analyse des productions d'élèves montre que ces prises d'initiatives restent très limitées. Cependant, nous avons observé une prise d'initiatives d'une part, de certains groupes (groupes 1 et 4) dans le problème 1. Ces groupes s'appuient sur la représentation et la lecture graphique pour justifier et interpréter le résultat. D'autre part dans le problème 3, le groupe 1 s'appuie sur la méthode graphique pour faire le type de tâches T5, c'est une prise d'initiatives importante de la part de ce groupe.

Les difficultés évoquées ci-dessus limitent les prises d'initiatives. Les élèves ont appris à appliquer une méthode. Il leur est difficile de procéder autrement, surtout parce qu'ils n'ont pas d'éléments leur permettant de contrôler leurs procédures : contrôle des ordres de grandeur, compréhension du sens des concepts d'ajustement et de point moyen notamment.

6.3 Analyse de la troisième expérimentation en classe (2021)

Dans cette partie nous analysons notre troisième expérimentation. Nous avons nous-même conçu et implémenté l'enseignement du chapitre qui nous intéresse dans une classe que nous nommons classe 1 dans ce qui suit. Dans cette classe, nous avons d'abord soumis les élèves à un pré-test avant de commencer le cours, et à la fin de notre expérimentation nous avons soumis les élèves à un post-test. Nous avons aussi passé le même post-test dans une autre classe dans un établissement privé dans une classe que nous nommons classe 2 ; la classe 2 avait suivi un cours « ordinaire ». Les élèves de la classe 1 étaient informés par leur enseignant que cet enseignement expérimental se situe dans le cadre d'une recherche.

Nous allons d'abord présenter dans la classe 1, l'analyse du pré-test, ensuite l'analyse de l'enseignement conçu et son implémentation suivie de l'analyse du post-test. Enfin nous terminons par l'analyse du post-test dans la classe 2. Les deux points suivants seront vus avec attention : les erreurs de calcul et le contrôle de l'ordre de grandeur.

6.3.1 Analyse du pré-test dans la classe 1

Cette partie est consacrée à l'analyse des réponses des élèves au pré-test dans la classe 1. Nous avons fait passer le test le 12 avril 2021. Il y avait 39 élèves présents. Ils ont effectué un travail en 10 groupes, qui a duré environ 45 minutes. Nous présentons d'abord l'analyse a priori ensuite l'analyse des réponses des élèves.

Énoncé du problème

Deux groupes d'élèves de Terminale, dont le niveau en mathématiques est exactement le même, composent sur un même sujet, mais n'ont pas le même correcteur.

Voici les notes lues dans l'ordre alphabétique des copies.

- Correcteur A :

15 ; 11 ; 05 ; 05 ; 08 ; 11 ; 03 ; 11 ; 05 ; 08 ; 08 ; 03 ; 14 ; 08 ; 16 ; 11 ; 14 ; 11 ; 08 ; 15.

- Correcteur B :

03 ; 13 ; 01 ; 07 ; 11 ; 13 ; 02 ; 07 ; 11 ; 07 ; 14 ; 03 ; 03 ; 02 ; 07 ; 03 ; 14 ; 02 ; 03 ; 14.

- 1) Calculer la moyenne et l'écart type des notes pour chaque correcteur.
- 2) Que pensez vous de la manière de noter de chaque correcteur ? Justifiez votre réponse.
- 3) Comment modifier les notes du correcteur B pour que leur moyenne se rapproche de celle du correcteur A ? Expliquer votre proposition de manière détaillée.

Analyse a priori

Ce problème est dans un contexte familier aux élèves, il aborde la question des notes de mathématiques. L'exercice est centré sur la série statistique à une variable, il permet de vérifier concrètement si l'élève maîtrise les connaissances nécessaires (les concepts statistiques) pour mieux suivre la série statistique double.

- 1) Concernant la première question qui consiste à calculer la moyenne qui relève du type de tâches T_Cal_MoyX et l'écart type des notes de chaque correcteur. Deux procédures possibles que l'élève peut mettre en œuvre dans la situation qui lui est proposée. La première procédure consiste à utiliser la moyenne arithmétique, c'est-à-dire diviser la somme de toutes les notes par leur nombre :

Correcteur A

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{20} x_i = \frac{1}{20} (15 + 11 + 05 + 05 + 08 + 11 + 03 + 11 + 05 + 08 + 08 + 03 + 14 + 08 + 16 + 11 + 14 + 11 + 08 + 15)$$

$$\bar{x} = \frac{190}{20} = 9,5$$

Correcteur B

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{20} x_i = \frac{1}{20} (03+13+01+07+11+13+02+07+11+07+14+03+03+02+07+03+14+02+03+14)$$

$$\bar{x} = \frac{140}{20} = 7$$

Dans cette procédure, l'erreur que l'élève peut commettre est l'erreur de calcul. Sa prise d'initiative est moyenne, les valeurs sont simples à calculer et le calcul de la moyenne des notes est familier aux élèves.

La deuxième procédure consiste à faire un tableau statistique, en regroupant les mêmes notes et en tenant compte de leur effectif. Cette procédure consiste à utiliser la moyenne pondérée :

Correcteur A

x_i	03	05	08	11	14	15	16
n_i	2	3	5	5	2	2	1

$$\bar{x} = \frac{1}{n} \sum_{i=1}^7 n_i x_i = \frac{1}{20} (2 \times 03 + 3 \times 05 + 5 \times 08 + 5 \times 11 + 2 \times 14 + 2 \times 15 + 1 \times 16) = \frac{190}{20} = 9,5$$

Correcteur B

x_i	01	02	03	07	11	13	14
n_i	1	3	5	4	2	2	3

$$\bar{x} = \frac{1}{n} \sum_{i=1}^7 n_i x_i = \frac{1}{20} (1 \times 01 + 3 \times 02 + 5 \times 03 + 4 \times 07 + 2 \times 11 + 2 \times 13 + 3 \times 14) = \frac{140}{20} = 7$$

Si l'élève n'arrive pas à investir les connaissances nécessaires vues en classe de première, sur la série statistique à une variable dans le cas de regroupement des valeurs, l'éventuelle difficulté qu'il aura pour cette procédure est la conception du tableau statistique pondéré par des effectifs des notes. Et l'erreur éventuelle est la non prise en compte de la pondération de ces notes dans le calcul. Pour la deuxième procédure, la potentialité en terme de prise d'initiative consiste à interpréter ces

données, c'est-à-dire, si l'élève se rend compte que certaines notes se répètent, alors il va procéder par regroupement en tenant compte de leur nombre. Après il fera recours au tableau statistique pondéré pour calculer la moyenne des notes.

Concernant le calcul de l'écart type, la procédure consiste à calculer d'abord la variance des notes, qui relève du type de tâches T_Cal_Var. Si on utilise la formule développée de la variance (Théorème de Koenig) on a :

Correcteur A

$$V(X) = \frac{1}{n} \sum_{i=1}^7 n_i x_i^2 - \bar{x}^2 = \frac{1}{20} (2 \times 9 + 3 \times 25 + 5 \times 64 + 5 \times 121 + 2 \times 196 + 2 \times 225 + 1 \times 256) - 90,25 = 15,55$$

L'écart type est : $\sigma_x = \sqrt{V(X)} = \sqrt{15,55} = 3,94$

Correcteur B

$$V(X) = \frac{1}{n} \sum_{i=1}^7 n_i x_i^2 - \bar{x}^2 = \frac{1}{20} (1 \times 01 + 3 \times 04 + 5 \times 09 + 4 \times 49 + 2 \times 121 + 2 \times 169 + 3 \times 196) - 49 = 22,1$$

L'écart type est : $\sigma_x = \sqrt{V(X)} = \sqrt{22,1} = 4,70$

2) La deuxième question nécessite une prise d'initiative dans l'interprétation des résultats statistiques.

Pour ceux qui vont comparer les deux moyennes des notes : $\bar{x} = 9,5$ pour le correcteur A et $\bar{x} = 7$ pour le correcteur B, diront aussi que le correcteur B est plus sévère que le correcteur A, car la moyenne des notes du correcteur B est inférieure à celle du correcteur A.

Pour ceux qui vont comparer l'écart des notes : 3,94 pour le correcteur A et 4,70 pour le correcteur B, diront que le correcteur B est plus sévère que le correcteur A.

En effet, les notes du correcteurs B sont plus dispersées que celles du correcteur A. Et cette dispersion plus importante est due à la présence de notes très faibles chez le correcteur B.

D'autres élèves vont regarder les notes elles-mêmes et non les indicateurs statistiques calculés et diront que, pour le correcteur A 10 élèves ont eu la moyenne tandis que pour le correcteur B, 7 élèves seulement ont eu la moyenne donc le correcteur B est plus sévère que le correcteur A.

D'autres diront qu'il y a plus de notes sous la moyenne chez le correcteur B que chez le correcteur A, et la plus faible note c'est chez le correcteur B, tandis que la meilleure note c'est chez le correcteur A.

3) La troisième question demande une large prise d'initiatives de la part des élèves.

Pour ceux qui vont comparer l'écart des deux moyennes qui est 2,5 :

- soit ils vont proposer d'ajouter un même nombre à chaque note du correcteur B, puis recalculer la moyenne ;

- soit ils vont proposer de multiplier par le même nombre supérieure à 1 de chaque note du correcteur B ;
- soit ils vont faire des modifications portant sur seulement une partie des notes du correcteur B.

Ceux qui vont modifier les notes, en tenant compte de l'écart des deux moyennes, sont des élèves qui ont des prises d'initiatives.

D'autres élèves, vont diminuer les notes du correcteur A de 2,5 pour que la moyenne soit égale à celle du correcteur B.

Certains vont comparer les variances ou les écarts types et ajouter un nombre proche de l'écart sur les notes du correcteur B.

C'est un exercice qui correspond à deux niveaux de mise en fonctionnement des connaissances. Nous avons d'abord le niveau technique, qui consiste à appliquer les formules et les propriétés, et ensuite nous avons le niveau des connaissances disponibles qui permet d'utiliser des méthodes non prévues.

Analyse des réponses des élèves

D'une manière générale, quatre groupes (groupes 1 ; 3 ; 7 et 10) ont donné une réponse correcte à la première question, quatre groupes (groupes 4 ; 5 ; 6 et 8) ont trouvé la réponse juste du calcul des moyennes par contre leur calcul des écart-types sont faux. Un groupe (groupe 9) a trouvé la bonne réponse du calcul des moyennes mais n'a pas calculé les écart-types et un groupe (groupe 2) n'a pas trouvé la bonne réponse à la première question. Quant à la deuxième question six groupes (groupes 1 ; 5 ; 6 ; 7 ; 8 et 9) ont proposé des justifications justes, trois groupes (groupes 3 ; 4 et 10) ont proposé des justifications fausses et un groupe (groupe 2) n'a pas répondu à la question. Pour la dernière question, deux groupes (groupes 5 et 9) ont fait des propositions justes, ils ont fait des modifications sur une parties des notes. Trois groupes (groupes 1 ; 3 et 6) ont fait des propositions fausses et cinq groupes (groupes 2 ; 4 ; 7 ; 8 et 10) n'ont pas répondu à la troisième question.

Nous récapitulons ces réponses dans le tableau ci-dessous.

Groupe N°	A trouvé la moyenne juste	A trouvé l'écart type juste	A bien justifié la manière de noter de chaque correcteur	A proposé des modifications justes sur les notes du correcteur B
Groupe 1	Oui	Oui	Oui	Non
Groupe 2	Non	Non	N'a pas fait	N'a pas fait
Groupe 3	Oui	Oui	Non	Non
Groupe 4	Oui	Non	Non	N'a pas fait
Groupe 5	Oui	Non	Oui	Oui
Groupe 6	Oui	Non	Oui	Non

Groupe 7	Oui	Oui	Oui	N'a pas fait
Groupe 8	Oui	Non	Oui	N'a pas fait
Groupe 9	Oui	N'a pas fait	Oui	Oui
Groupe 10	Oui	Oui	Non	N'a pas fait

Tableau 17 : Tableau récapitulatif des réponses des élèves de la classe 1 au pré-test

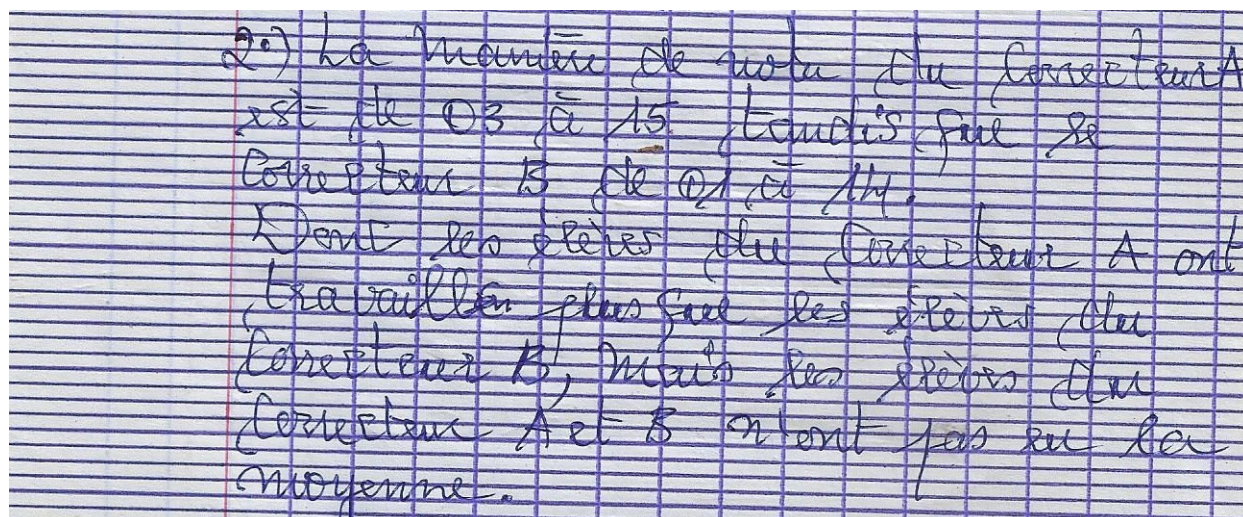
Dans ce qui suit nous présentons les réponses des élèves d'une manière détaillée.

Pour la première question, sept groupes (groupes 1 ; 2 ; 3 ; 4 ; 5 ; 8 et 10) ont proposé une formule juste de la moyenne (moyenne arithmétique ou moyenne pondérée) et les trois autres groupes (groupes 6 ; 7 et 9) ont fait directement l'application numérique sans écrire la formule. Tous les groupes ont fait le tableau statistique, ils ont tenu compte de la pondération en regroupant les mêmes notes et en tenant compte de leur effectif, ceci relève d'une prise d'initiative de la part de ces groupes. Par ailleurs, le groupe 1 a écrit la formule de la moyenne arithmétique et ne s'en est pas servi pour calculer la moyenne, il a utilisé la moyenne pondérée dans l'application numérique. Concernant la valeur des moyennes les groupes (groupes 1 ; 3 ; 4 ; 5 ; 6 ; 7 ; 8 ; 9 et 10) ont trouvé la réponse juste par contre le groupe 2 a trouvé une réponse erronée, il n'a pas trouvé la bonne valeur de l'effectif total.

Concernant le calcul de l'écart type, les groupes 1 ; 3 ; 7 ; et 10 ont trouvé la réponse correcte, ils ont calculé la variance et déduit l'écart type. Les groupes 2 ; 4 ; 5 ; 6 et 8 ont trouvé des réponses erronées, ils ont proposé la bonne formule de la variance mais ils ont fait des erreurs de calcul. Et le groupe 9 n'a pas répondu à la question.

La deuxième question consiste à justifier la manière de noter de chaque correcteur. Six groupes (groupes 1 ; 5 ; 6 ; 7 ; 8 et 9) ont proposé des justifications justes. La justification du groupe 1 est basée sur l'écart des notes mais il n'ont pas détaillé sa réponse. Le fait de comparer l'écart des notes fait allusion à la dispersion des notes. Les groupes 5 et 8 ne raisonnent pas sur les indicateurs statistiques, ils comparent le nombre des élèves qui ont eu des notes supérieures à 10. Quant aux groupes 6 et 7, les élèves ont comparé les deux moyennes, en plus les élèves du groupe 6 résumant les trois indicateurs statistiques (moyenne, variance et l'écart type) dans un tableau. Et pour le groupe 9, les élèves ont comparé les plus faibles notes, le nombre de notes sous la moyenne et les meilleures notes. Au regard de toutes ces justifications nous dirons que ces groupes (groupes 1 ; 5 ; 6 ; 7 ; 8 et 9) d'élèves ont des prises d'initiatives dans leur raisonnement. Les groupes (groupes 3 ; 4 et 10) ont donné des justifications fausses. Le raisonnement du groupe 3 n'est pas correct, car ce

groupe dit que les correcteurs ont attribué des notes en fonction du niveau



2.) La manière de nota du correcteur A est de 03 à 15 tandis que le correcteur B de 01 à 14. Dont les élèves du correcteur A ont travaillé plus que les élèves du correcteur B, mais les élèves du correcteur A et B n'ont pas eu la moyenne.

Figure 33: Extrait de copie du groupe 10

des élèves. Le groupe 4 a calculé l'étendue des notes. Nous rappelons que l'étendue d'une série statistique est la différence entre sa valeur la plus élevée et sa valeur la plus basse. Ce groupe a calculé l'étendue des notes chez le correcteur A ainsi que chez le correcteur B, il trouve 13 dans les deux cas et conclut que leur manière de noter est la même. Cette justification basée sur l'indicateur étendue n'est pas suffisante pour conclure. Quant au groupe 10, il compare l'intervalle des notes, c'est-à-dire, chez le correcteur A les notes varient de 03 à 16 tandis que chez le correcteur B les notes varient de 01 à 14, donc les élèves du correcteur A travaillent mieux que ceux du correcteur B. On observe une prise d'initiative dans le raisonnement du groupe 10, le fait de comparer les intervalles de notes mais cela n'est pas suffisant pour conclure que le correcteur B est plus sévère que le correcteur A. Et enfin le groupe 2 n'a pas répondu à la question.

La dernière question consiste à proposer de manière détaillée comment modifier les notes du correcteur B pour que la moyenne des notes qu'il a mises se rapproche de celle du correcteur A. Ce type de tâches nécessite une large prise d'initiatives de la part des élèves, car le fait de modifier par exemple une partie de notes du correcteur B relève d'une prise d'initiatives. Les groupes 5 et 9 ont fait des propositions correctes, leur proposition portant sur la modification d'une partie de notes du correcteur B pour se rapprocher de la moyenne du correcteur A. Les élèves du groupe 5 ont modifié une partie des notes du correcteur B de telle sorte qu'ils ont eu 11 notes sous la moyenne et 9 notes supérieures à 10. Et le groupe 9 modifie les notes du correcteur B de telle sorte qu'il a eu le même intervalle de notes. Nous constatons que l'effectif total trouvé par le groupe 9 est supérieur à 20. Nous avons observé une large prise d'initiatives de la part de ces deux groupes, ils ont su modifier une partie des notes du correcteur B dans l'optique que la nouvelle moyenne se rapproche de celle du correcteur A. Les groupes 1 ; 3 et 6 ont proposé des explications qui ne sont pas correctes. Concernant le groupe 1, il propose de modifier les notes du correcteur B pour se rapprocher des notes du correcteur A, pour qu'il y ait égalité entre les écart types des notes. Le groupe 3 propose d'ajouter 50 sur la somme des notes du correcteur B pour qu'il ait la même moyenne de notes que celle du correcteur A. Quant au groupe 6, il propose de faire la moyenne des trois indicateurs statistique (moyenne, variance et écart type) du correcteur A et B. Cette proposition, elle est fautive

car elle ne permet pas de rapprocher la moyenne du correcteur A de celle du correcteur B. Et enfin les groupes 2 ; 4 ; 7 ; 8 et 10 n'ont pas répondu à la question.

Nous retenons de ce pré-test plusieurs choses.

D'une part, les élèves ont des connaissances à propos des séries statistiques à une variable, qui seront utiles pour notre enseignement des séries statistiques doubles :

- Ils maîtrisent la formule de la moyenne,
- ils possèdent la capacité d'écrire les données statistiques sous la forme d'un tableau statistique en tenant compte de la pondération des données.
- Ils savent se servir des indicateurs statistiques pour justifier ou interpréter leur résultat.

D'autre part, ils rencontrent tout de même certaines difficultés, auxquelles nous devons être attentif :

- Ils font des erreurs de calcul quand ils calculent la variance,
- ils ont des difficultés à commencer une question qui nécessite une large prise d'initiatives.

Finalement, en ce qui concerne les prises d'initiatives :

- les élèves ont des prises d'initiatives dans les types de tâches où on leur demande de justifier ou d'interpréter le résultat trouvé. Par contre, ils n'ont pas de prises d'initiatives dans les types de tâches atypiques qui nécessitent de prendre du recul pour répondre (question 3).

6.3.2 Présentation de l'enseignement conçu et analyse de son implémentation

Nous présentons et analysons, dans cette partie, l'enseignement expérimental mené en 2021 par le chercheur lui-même dans une classe d'un établissement public que nous nommons classe 1. Cette classe comporte 67 élèves, son enseignant avait accepté de nous la confier pour le chapitre sur les séries statistiques doubles. Nous avons filmé les quatre séances du cours, chacune a duré entre une et deux heures.

La préparation d'une leçon nécessite le passage à l'écrit, cela permet de prévoir ce qui peut se produire pendant la situation d'apprentissage et permet aussi d'anticiper certaines réactions d'élèves et d'orienter certains choix de l'enseignant. Pour bâtir ce cours, nous nous sommes inspiré de trois manuels : le manuel CIAM, le manuel de BTS, et le manuel Transmath, et le programme officiel.

Nous présentons ci-dessous l'ensemble de l'enseignement. Pour l'analyse de l'implémentation, nous avons choisi des extraits qui diffèrent sensiblement de ce qui est pratiqué dans l'enseignement ordinaire, selon les analyses menées au chapitre 6.

TABLEAU DE PRÉSENTATION

Étapes	Dates	Durées	Thèmes	Différence avec l'enseignement ordinaire
--------	-------	--------	--------	--

Pré-test	12/04/2021	45 min	Contrôle sur la statistique à une variable.	
Première séance de cours	15/04/2021	1 h 35 min	Rappels sur la statistique à une variable et introduction de la série statistique double .	Faire le lien entre la statistique à une variable et la statistique à deux variables
Deuxième séance de cours	19/04/2021	1 h 52 min	La série statistique double : représentation graphique, nuage de points, point moyen du nuage, variance et covariance.	Positionnement du point moyen dans le nuage de points, et l'influence des effectifs sur la position du point moyen. Contrôle de l'ordre de grandeur.
Troisième séance de cours	22/04/2021	2 h	Coefficient de corrélation linéaire. Ajustement affine : méthode à la règle et méthode de Mayer.	Le choix d'aborder la notion du coefficient de corrélation avant la méthode des moindres carrés. La mise en valeur de la méthode « à la règle », qui souligne le sens de l'ajustement affine.
Quatrième séance de cours	26/04/2021	1 h 46 min	Méthode de Mayer et méthode des moindres carrés .	
Post-test	29/04/2021	2 h	Contrôle sur la statistique à deux variables .	

Tableau 18 : Tableau récapitulatif de la troisième expérimentation dans la classe 1

Les éléments importants qui diffèrent dans cet enseignement expérimental par rapport à l'enseignement ordinaire sont les suivants :

- Nous avons pris soin de faire un lien entre la statistique à deux variables (savoir nouveau) et ce que les élèves savent déjà sur la statistique à une variable ;
- Nous sommes revenu sur la question de la pondération, qui nous a permis notamment de discuter le positionnement du point moyen dans le nuage de points, et l'influence des effectifs sur la position du point moyen ;
- Nous avons introduit la notion de coefficient de corrélation avant les méthodes des moindres carrés et méthode de Mayer, ainsi les élèves ont tout de suite deux moyens pour examiner si un ajustement affine est justifié ;
- Nous avons mis en valeur l'ajustement « à la règle », qui permet de souligner le sens de l'ajustement affine : on fait passer une droite « près » du nuage de points ;
- Nous avons illustré chaque notion ou partie abordée par un exemple ou un exercice ;

- Les techniques utilisées sont justifiées par des technologies ;
- Chaque méthode d'ajustement affine est illustrée par un exemple ;

Nous avons ainsi produit et transcrit les synopsis des épisodes dans trois séances que nous jugeons intéressants. Le choix de ces épisodes a été fait pour mettre en évidence des aspects de cet enseignement expérimental qui diffèrent de ce que nous avons vu dans l'enseignement ordinaire.

Les épisodes choisis sont les suivants :

- Épisode 1, séance 1 : Faire le lien entre ce que les élèves savent de la statistique à une variable, et le savoir nouveau sur la statistique à deux variables. Introduire l'idée de lien entre deux variables.
- Épisode 2, séance 2 : Discuter la question de la pondération et la représentation du nuage de points.
- Épisode 3, séance 3 : Détermination graphique d'une droite d'ajustement affine et de son équation par la méthode à la règle.

Pour ce faire, nous analysons ces épisodes pour indiquer où se situent des prises d'initiatives des élèves.

6.3.2.1 Présentation et analyse du premier épisode (séance 1)

Nous avons débuté le cours par quelques rappels sur la série statistique à une variable, en particulier sur la notion de moyenne arithmétique et la moyenne pondérée, illustrés d'un exemple. Ce rappel permet de vérifier concrètement si l'élève maîtrise les connaissances de la série statistique à une variable, pour mieux suivre le cours sur la série statistique à deux variables et de faire les ajustements nécessaires en fonction des programmes et de ce qui a été traité en classe.

Nous présentons ci-dessous le synopsis du premier épisode, sous forme de tableau.

Minutage	Phases	Savoirs en jeu	Descriptions
00:02	Rappels sur la série statistique à une variable	La moyenne arithmétique	L'enseignant écrit la définition de la moyenne arithmétique, puis l'énoncé de l'exemple 1 au tableau, et il explique aux élèves comment on écrit la formule développée de la moyenne arithmétique. Pendant la correction de l'exercice proposé en exemple 1, l'enseignant interagit avec les élèves en posant des questions sur la résolution.
<u>Commentaires</u> : Un élève volontaire est passé au tableau faire l'exemple 1. Il n'arrive pas à commencer l'exercice, donc l'enseignant intervient pour lui expliquer que l'objectif de l'exemple 1, est d'utiliser la formule de la moyenne arithmétique pour calculer la moyenne des notes.			
16:08	Rappels sur la série statistique à une variable	La moyenne pondérée.	L'enseignant interroge les élèves sur la formule de la moyenne pondérée avant d'écrire la définition de la moyenne pondérée au tableau. Il écrit ensuite l'énoncé de l'exemple 2 au tableau. Un élève passe au tableau corrigé l'exemple 2.
<u>Commentaires</u> : Le deuxième élève qui est passé au tableau corriger l'exemple 2, arrive à commencer la résolution sans l'intervention de l'enseignant. L'enseignant intervient a un moment de la résolution pour donner des explication sur l'utilisation du tableau d'effectifs.			
33:32	Introduction de la série à deux variables	Présentation de la série statistique à deux variables.	L'enseignant commence par dicter le cours puis il écrit au tableau. Il interagit avec les élèves sur les exemples.
<u>Commentaires</u> : Trois élèves interrogés, pour donner un exemple de deux variables qui ont un lien entre elles.			
47:57	Présentation de l'énoncé l'exercice 1. Réponse des élèves à la première question	Définition de la série statistique à deux variables.	L'enseignant écrit l'énoncé de l'exercice au tableau, et pose oralement la première question aux élèves . Quatre élèves ont répondu à la question mais leur réponse était fausse.
<u>Commentaires</u> : L'enseignant pose une question où les élèves n'ont pas encore les moyens de répondre correctement, il voulait voir leur point de vue sur la notion d'estimation. Les élèves ont pris de l'initiative de proposer des réponses à cette question.			
58:45	Définition et représentation graphique de la série statistique à deux variables. Correction de la deuxième question	Représentation graphique. Repère, origine coordonnées.	Un élève est passé au tableau faire la représentation graphique de l'ensemble de couples (x_i, y_i) .

Commentaires : Pour représenter les coordonnées de points sur le graphique, il (l'élève) a fait une erreur, il a confondu les coordonnées de x à celles de y. Après l'erreur de l'élève, l'enseignant intervient pour lui expliquer ce qu'il faut faire. Après la représentation graphique, l'enseignant réexplique aux élèves comment on place les points $M_i(x_i, y_i)$.

Tableau 19: Synopsis du premier épisode

Au début de la première séance, l'enseignant avait commencé son cours par rappeler la notion de la moyenne arithmétique et la moyenne pondérée avec des exemples à l'appui, et avec des références au pré-test. Après ce rappel, il introduit la série statistique à deux variables en interagissant avec les élèves. Trois élèves sont interrogés, il leur demande de donner un exemple de série statistique à deux variables qui aurait un lien entre elles. Cette demande a été faite pour susciter une prise d'initiative des élèves, sur l'idée de lien entre deux variables qui est très importante dans ce cours.

Un seul élève a proposé un exemple juste. Ensuite, il aborde la définition et la représentation graphique de la série statistique à deux variables. Par la suite, l'enseignant aborde un exercice (Exercice 1) qui est transcrit ci-dessous.

Exercice 1

Le tableau suivant représente les couples de notes en Mathématiques et en Physique de 13 de ces élèves ; Ils ont eu des couples de notes tous différents.

Note x_i en mathématiques	7	8	8	9	9	9	9	10	10	11	12	14	15
Note y_i en physique	6	7	9	8	10	11	12	8	11	12	10	13	14

Un élève a obtenu la note de 13 en mathématiques. A votre avis, est-ce qu'on peut estimer sa note de physique ?

Proposer de faire le graphique après les réponses des élèves (un élève passe au tableau faire le graphique).

Figure 34: Exercice introductif

L'exercice 1 comporte deux types de tâches. Le premier type de tâche porte sur la notion des estimations. Ce type de tâche donne à voir ce que pensent les élèves sur la notion de l'estimation. De voir de quelle façon, ils vont procéder. Est-ce qu'ils vont estimer la note de mathématiques en fonction de la note de physique ?

L'un des objectifs de l'enseignement de la série statistique à deux variables est d'utiliser une droite d'ajustement pour faire l'estimation. Et donc, ce type de tâche introduit au début du cours sera répondu à la fin du cours par l'enseignant.

Plusieurs élèves ont répondu « oui », qu'en connaissant la note de mathématique qui correspond à la variable x on peut estimer la note de physique qui correspond à la variable y .

Deux élèves répondent :

« Si un élève a 13 en mathématique et y a un lien entre les mathématiques et la physique donc il peut mieux faire encore en physique. » (Épisode 1, 54:53)

« On peut prendre la note de la mathématique divisée par l'effectif total. » (Épisode 1, 56:58)

L'enseignant répond à la deuxième réponse :

« La note en math est 13 et l'effectif total aussi est 13. Quand tu divise 13 par 13, tu auras 1, donc pour toi la note de physique est 1 ? » (Épisode 1, 57:28)

Les élèves n'ont pas encore étudié la notion d'estimation. Le professeur veut amener cette idée intuitivement, avant de commencer le cours sur la notion d'ajustement. Il demande donc si on peut estimer.

La première réponse traduit une connaissance culturelle. Elle n'est pas une estimation, mais exprime l'idée de grandeurs liées. Nous faisons l'hypothèse que cet élève raisonne en se basant sur la logique que, si un élève travaille bien en mathématique ou comprend bien la mathématique alors il peut aussi mieux faire en physique. Le lien dont il parle, nous semble que, c'est la complémentarité qui existe entre certains chapitres en mathématiques et en physique. Concernant la deuxième réponse proposée par un autre élève, nous disons que cette réponse n'a rien à voir avec la question posée. Comme la note en mathématique est 13 et l'effectif total est donné (13) alors pour répondre à la question posée, l'élève veut utiliser toutes ces données, par rapport à la réponse de cet élève nous faisons référence à un problème proposé aux élèves de CE1 à l'école élémentaire sur « l'âge du capitaine » : il s'agit d'un effet de contrat didactique.

Ici le professeur pose une question pour laquelle les élèves n'ont justement pas encore le moyen de faire une réponse correcte. Les deux élèves prennent l'initiative de répondre. Le premier s'appuie sur des connaissances culturelles, et ne propose pas une estimation chiffrée. Le second effectue un calcul à partir des données de l'énoncé, par un effet de contrat didactique. Le professeur corrige le deuxième élève, mais il ne saisit pas l'opportunité de rebondir sur les propos du premier. Il aurait pu demander : « est-ce que tu peux estimer une valeur » ? Le premier élève a bien compris l'idée de dépendance, et le professeur pourrait en profiter pour discuter ce qu'on entend par « estimation ».

Le deuxième type de tâches consiste à faire la représentation graphique du nuage de points. Pour répondre, un élève passe au tableau, il a des prises d'initiative dans le choix et la représentation du repère. La représentation graphique est vue dans les classes antérieures, mais l'objectif premier de cet apprentissage est la réactivation et/ou la mobilisation, la consolidation de cette connaissance ancienne. Par ailleurs pour représenter l'ensemble des points du nuage, l'élève a confondu les coordonnées de x à celles de y . Cette erreur est prévisible ici : en effet les deux données de la série sont de même nature.

L'enseignant intervient : « Est-ce que la représentation qu'il est en train de faire est juste ? » (Épisode 1, 01:08:08)

« Non Monsieur » (Épisode 1, 01:08:09)

Vue l'erreur commise par cet élève, l'enseignant se trouve forcé, en vue d'apporter une aide indirecte à l'élève ou à toute la classe, en expliquant que les notes de mathématiques correspondent aux valeurs de x sur l'axe des abscisses et les notes de physique correspondent aux valeurs de y sur l'axe des ordonnées.

« Qu'est-ce qui ne va pas avec la représentation ?

Regardez la position des points, tu es entrain de confondre x et y , c'est lorsque $x=9,9,9$, y varie [L'enseignant montre du doigt les valeurs dans le tableau statistique]

Lorsque $x=9, y=8$, $x=9, y=10$ ainsi de suite. [Il a confondu l'abscisse et l'ordonnée, il efface et refait la représentation]

$x=7, y=6$; $x=8, y=7$; $x=8, y=9$; $x=9, y=8$; $x=9, y=10$; $x=9, y=11$; $x=9, y=12$; $x=10, y=8$; $x=10, y=11$; $x=11, y=12$; $x=12, y=10$; $x=14, y=13$ et $x=14, y=15$
 » (Épisode 1, 01:08:13)

L'intervention de l'enseignant réduit la prise d'initiative de l'élève qui est au tableau. Cependant, les autres élèves ont pris l'initiative de dire non, que la représentation du nuage de points n'est pas juste. Avec l'intervention de l'enseignant et la participation des autres élèves, l'élève désigné arrive à représenter le nuage de points.

6.3.2.2 Présentation et analyse du deuxième épisode (séance 2)

Dans cette section, nous présentons et analysons un extrait de la deuxième séance. Cet épisode concerne la correction d'un exemple portant sur les notions (représentation graphique, nuage de points, point moyen du nuage) abordées pendant la première séance.

Minutage	Phases	Savoirs en jeu	Descriptions
00:06	Correction de l'exemple 2	Point moyen	L'enseignant réécrit le tableau statistique au tableau et pose la question aux élèves de savoir la valeur trouvée. Il vérifie sur leur cahier s'ils ont bien représenté les coordonnées du point moyen sur leur cahier. Il écrit au tableau la solution proposée par les élèves.
<u>Commentaires</u> : A la fin de la dernière séance, l'enseignant a demandé aux élèves de calculer les coordonnées du point moyen à la maison. Il a vérifié la réponse des élèves qui ont fait l'exercice proposé. Ceux qui ont fait l'exercice ont trouvé la bonne réponse.			
07:18	Présentation de l'énoncé	Pondération par l'effectif. Représentation graphique.	L'enseignant écrit l'intégralité de l'énoncé au tableau. Il pose la première question sur la représentation graphique oralement. L'élève désigné trouve la bonne réponse.
<u>Commentaires</u> : Après avoir présenté l'énoncé. Un élève répond à la question posée oralement, la question consiste à savoir si on ajoute l'effectif aux notes des élèves présentées précédemment dans le tableau, il faut changer la représentation graphique ou pas ? L'élève qui a répondu à la question, dit qu'on n'a pas besoin de changer la représentation du nuage			

car les effectifs n'influencent pas sur la représentation de x_i et y_i .			
12:08	Détermination du point moyen G' du nuage.	Coordonnées du point moyen G'	Un élève volontaire passe au tableau déterminer les coordonnées du point moyen G' du nuage. Il écrit la bonne formule des coordonnées du point moyen G' . En interagissant avec les autres élèves ils ont trouvé les coordonnées du point moyen G' . Ensuite l'enseignant a commenté le résultat trouvé.
<u>Commentaires</u> : Pour calculer les coordonnées de G' , les autres élèves ont participé en comparant les calculs. Par la suite l'enseignant intervient pour commenter le résultat du calcul.			

Tableau 20: Synopsis du deuxième épisode

Après avoir défini le nuage de points et point moyen du nuage pendant la première séance. L'enseignant modifie les données de l'exercice 1 abordé pendant le premier épisode, en ajoutant les effectifs aux couples des notes. Dans cet exemple l'objectif de l'enseignant n'est pas de voir les élèves calculer les coordonnées du point moyen G et de le placer sur le graphique qui relèvent respectivement des types de tâches T_Cal_CoordG et T_Pla_MoyG mais surtout de faire réfléchir les élèves aux effectifs et à leur influence sur la position du point moyen. Nous rappelons l'exemple ci-dessous.

Exemple 2 :

- Reprenons l'exercice 1 précédent. Calculer les coordonnées du point moyen G et le placer sur le graphique.
- En fait, pour les 100 élèves de la classe, les couples de notes correspondent aux effectifs donnés dans le tableau ci-dessous (ce sont exactement les couples de notes précédents, mais il y a plusieurs élèves pour chaque couple de notes).

Note x_i en mathématiques	7	8	8	9	9	9	9	10	10	11	12	14	15
Note y_i en physique	6	7	9	8	10	11	12	8	11	12	10	13	14
Effectif n_i	5	6	8	10	12	14	8	11	7	9	3	5	2

Est-ce que selon vous, il faut changer la représentation graphique ?

Déterminer le point moyen G' du nuage.

Cet exemple 2 a été donné aux élèves à la fin de la première séance. L'enseignant commence la séance 2 par la correction de cet exemple 2. Dans cette situation, la question posée aux élèves oralement est :

« Selon vous, est ce qu'il faut changer la représentation graphique qu'on avait fait la fois passée, si on ajoute l'effectif ? » (Épisode 2, 10:34)

« oui Monsieur, non Monsieur » (Épisode 2, 10:34)

« Ceux qui disent oui, donnez les raisons et ceux qui disent non aussi. Si on ajoute l'effectif qu'est ce qui va changer ici ? » (Épisode 2, 10:47)

Il s'agit ici d'évaluer le point de vue des élèves en terme de prise d'initiatives sur deux registres de représentations, le registre tableau et le registre graphique et sur le passage de l'un à l'autre.

Après avoir représenté le nuage de points, l'enseignant a ajouté les effectifs au tableau statistique. Il y a une large prise d'initiatives ici pour cette question, l'élève doit prendre en compte les deux registres de représentation. Un élève qui a dit « oui », justifie sa réponse : « On a représenté le nuage de points en fonction de x et y » (Épisode 2, 11:11)

L'initiative attendue de la part des élèves est de constater qu'en ajoutant l'effectif les valeurs de x et y ne changent pas dans le registre tableau. En plus, pour représenter le nuage de points, on a utilisé les valeurs de x et y , donc l'effectif n'impacte pas sur le nuage de points. Cet élève semble avoir bien analysé les deux registres et proposer la réponse attendue. Par ailleurs les autres élèves qui ont dit « oui » ou « non » se sont retenus de parler après l'intervention de leur collègue. Nous faisons l'hypothèse qu'ils sont convaincus de la réponse de leur collègue et décident de ne pas parler. Alors l'enseignant conclut ainsi :

« Ok, d'après votre collègue, la représentation du nuage de points, on l'avait représenté en fonction de x et y , les valeurs de x et y ne changent pas donc l'effectif ne joue pas sur la représentation du nuage de points. » (Épisode 2, 11:21).

Pour répondre à la deuxième question qui relève du type de tâches T_Cal_CoordG', un élève volontaire est passé au tableau. Il a proposé la bonne formule des coordonnées du point moyen G', en tenant compte de la pondération des notes. Il arrive à faire la différence avec la formule des coordonnées du point moyen G calculé précédemment. Le résultat du calcul trouvé est juste.

Par la suite l'enseignant intervient pour commenter le résultat trouvé :

« on observe bien que les effectifs modifient le point moyen. On s'éloigne par exemple du point 15 car l'effectif n'est que 2. Par contre, on se rapproche du point 9 dont l'effectif est de 14. Vous pouvez constater cela graphiquement aussi. » (Épisode 2, 27:43).

Ici, nous voyons que l'enseignant essaie de faire comprendre aux élèves que la pondération influence la position du point moyen. Ceci permet de travailler sur la position du point moyen : le professeur espère que de cette manière, les élèves ne feront pas l'erreur de mettre le point moyen en dehors du nuage. Si ils font une erreur de calcul, le fait de savoir que le point moyen est dans le nuage leur permettra de se corriger : c'est le contrôle des ordres de grandeur. De plus, lors de cet épisode on travaille aussi sur la notion de pondération, pour des séries à deux variables. Même si le nuage ne change pas d'apparence, chaque point n'a pas le même « poids », ceci va jouer sur le point moyen, et ensuite sur la droite d'ajustement.

6.3.2.3 Présentation et analyse du troisième épisode (séance 3)

Le troisième épisode aborde la deuxième partie du cours qui porte sur la notion d'ajustement affine. Il a commencé d'abord par le principe et les caractéristiques d'un ajustement affine, par la suite il a abordé la méthode à la règle suivie d'un exemple qui est présenté ci-dessous. Notre troisième épisode est centré sur un exemple portant sur la méthode à la règle.

Minutage	Phases	Savoirs en jeu	Descriptions
30:10	Ajustement affine : méthode à la règle	Principe et caractéristique d'un ajustement affine : justifier la pertinence de faire un ajustement affine.	L'enseignant écrit le principe et caractéristique, ensuite il écrit la méthode à la règle au tableau. Il vérifie les travaux des élèves. Ensuite il représente la droite d'équation au tableau.
<u>Commentaires</u> : Après la méthode à la règle présentée par l'enseignant, il demande aux élèves de tracer la droite sur leur cahier en se servant des données de l'exemple 2b, ensuite l'enseignant passe dans le couloir vérifier les cahiers des élèves. Après avoir répondu à la question d'un élève, il trace la droite d'ajustement affine.			
01:03:53	Exemple sur la méthode à la règle	Équation des droites	L'enseignant interroge d'abord les élèves, et ensuite il écrit l'équation de la droite au tableau et il donne des explications aux élèves.
<u>Commentaires</u> : Les élèves ont oublié comment on détermine le coefficient directeur dans l'équation d'une droite. L'enseignant leur explique comment on détermine les coefficients a et b . Ensemble avec les élèves il établit l'équation de la droite.			

Tableau 21: Synopsis du troisième épisode

Dans son exposé sur l'ajustement affine, l'enseignant commence par présenter le principe et caractéristique d'un ajustement affine, où il y a le fait de justifier que c'est pertinent de faire un ajustement affine. Ensuite il demande aux élèves de déterminer graphiquement une droite d'ajustement à la règle, en se servant des données de l'exemple 1). Cette question relève du type de tâche T_Rea_AjReg.

L'énoncé est présenté ci-dessous.

Reprenons les données de l'exemple 1). Déterminer graphiquement une droite d'ajustement à la règle. Calculer l'équation de la droite que vous avez tracée.

Exemple 1 :

Note x_i en mathématiques	7	8	8	9	9	9	9	10	10	11	12	14	15
Note y_i en physique	6	7	9	8	10	11	12	8	11	12	10	13	14

Le choix de cet exercice permet de mettre en évidence plusieurs aspects de la série statistique double tels que : l'application de la méthode à la règle, la représentation graphique, faire une droite qui s'approche du nuage de points, et la détermination de l'équation de la droite sous forme $y=ax+b$.

Pour déterminer graphiquement la droite d'ajustement ils commencent par représenter le nuage de points, travail qui avait déjà été fait précédemment.

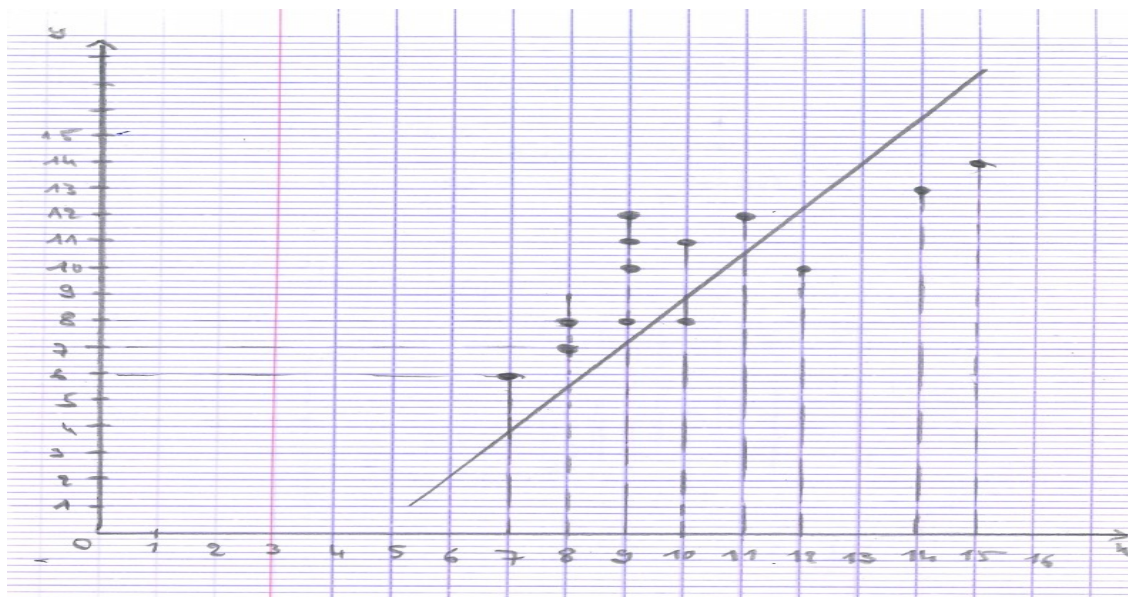


Figure 35: Extrait du cahier d'un élève

Ensuite il s'engage des échanges sur la manière de placer la droite.

Pour tracer la droite un élève a posé la question de savoir si la droite doit forcément passer sur la majorité des points du nuage :

« S'il vous plaît Monsieur, est-ce que la droite doit passer obligatoire sur la majorité des points ? » (Épisode 3, 41:07)

L'enseignant explique qu'il s'agit plutôt de passer « au milieu » du nuage :

« Relis bien le résumé de la méthode, c'est bien écrit que la droite passe le plus près des points du nuage et on essaie d'équilibrer le nombre de points situés au dessus et au dessous de cette droite.

Donc pour répondre à ta question, la droite ne doit pas forcément passer par la majorité des points. » (Épisode 3, 41:52)

Cette question de l'élève est intéressante et elle renvoie au différents choix possibles pour un ajustement de type graphique (Casey & Wasserman 2015). En effet, dans le nuage, on peut observer certains groupes de 4 points qui semblent alignés. Si la droite choisie passe par ces quatre points, les sommes des distances entre chaque point et la droite peut être réduite (puisque 4 valeurs sont nulles). Mais il y aurait pour ce nuage deux droites assez différentes ayant cette propriété (passer par 4 points).

Le fait que l'enseignant intervient pour répondre à la question posée par cet élève, réduit leur prise d'initiative. Car ce type de tâches est l'occasion de voir les élèves prendre l'initiative de faire une droite qui s'approche du nuage de points. Force est de constater que avec l'intervention de l'enseignant, ils ont réussi à tracer une droite qui ajuste le nuage des points. Par ailleurs, nous avons constaté que les élèves ont orienté leur droite dans le même sens que le nuage de points, cela relève de la prise d'initiative de leur part.

La deuxième question consiste à calculer l'équation de la droite tracée. Ça relève du type de tâches T_Dét_EquatC.

L'enseignant pose une question aux élèves pour savoir comment ils vont procéder pour déterminer l'équation de la droite. Comme les élèves ne réagissent pas, il pose autrement sa question :

« Regardez, vous avez vu au collège qu'avec deux points vous pouvez tracer une droite qui passe par ces deux points. Ensuite on vous a montré au collège comment déterminer l'équation de la droite qui passe par deux points de cette droite. Qui peut me rappeler la formule de l'équation de cette droite? » (Épisode 3, 01:04:21)

Un élève répond : « On ne sait pas Monsieur, on a déjà oublié » (Épisode 3, 01:05:03)

L'enseignant intervient pour prendre en charge ce type de tâches, il écrit la formule au tableau ensuite il demande à un élève de lui proposer les coordonnées de deux points de sa droite tracée.

L'enseignant : « Voici la formule qui permet de déterminer l'équation de la droite $y=ax+b$,

$a = \frac{y_B - y_A}{x_B - x_A}$ et $b = y_B - ax_B$ ou $b = y_A - ax_A$, les x_A, y_A, x_B et y_B sont les coordonnées des points A et B.

Qui peut me donner les coordonnées des deux points de la droite que vous avez tracée? » (Épisode 3, 01:07:00)

Une élève répond : « A a pour coordonnées 8 ; 9 et B a pour coordonnées 14 et 13 » (Épisode 3, 01:07:38).

On se rend compte que les coordonnées proposées par cet élève coïncident avec des points du nuage, l'enseignant valide le choix de ces deux points.

L'unique responsabilité confiée à l'élève dans la détermination de l'équation de cette droite est le choix des deux points de la droite tracée, ça relève d'une prise d'initiative mais qui reste limitée.

En interagissant avec les élèves sur les calculs, ils trouvent l'équation de la droite $y=0,6x+4,2$

A la suite de cet épisode, l'enseignant est passé à l'étude des méthodes « classiques » de Mayer et des moindres carrés. Il a souligné que pour la méthode à la règle, on a le choix de plusieurs droites, ce qui peut poser problème. De plus quand il y a une pondération on ne peut pas en tenir compte. Ceci justifie le fait de prendre d'autres méthodes. Cependant, il a souligné que la méthode à la règle reste pertinente et qu'ils peuvent l'utiliser.

6.3.3 Analyse du post-test dans la classe 1

Nous avons soumis de nouveau le même sujet de la deuxième expérimentation (Partie 6.2) dans deux classes de terminales SE en Centrafrique dans deux établissements différents. La première soumission a eu lieu le 29 avril 2021 dans un établissement public (classe 1). Il y avait 48 élèves présents. Et la deuxième soumission a eu lieu le 14 août 2021 dans un établissement privé (classe 2) dont 10 élèves étaient présents.

Dans la classe 1, les élèves ont encore travaillé en groupe. Nous avons 12 groupes, ces groupes ne sont pas les mêmes que lors du pré-test.

6.3.3.1 Analyse globale des réponses des élèves de la classe 1

Problème 1

Concernant le problème 1 deux groupes d'élèves (groupes 3 et 8) n'ont pas fait l'exercice. Par contre les huit autres groupes ont fait l'exercice dont cinq groupes d'élèves (groupes 1, 4, 5, 6 et 10) ont donné une réponse correcte et trois groupes d'élèves (groupes 2, 7 et 9) n'ont pas donné une réponse juste. Malgré la consigne trois groupes ont déterminé la droite des moindres carrés et deux groupes ont utilisé la méthode au jugé.

Nous récapitulons ces types de réponses dans le tableau ci-dessous.

Groupe N°	A pris en considération la méthode	A dit que la production était juste
Groupe 1	Oui	Oui
Groupe 2	Non	Non
Groupe 3	N'a pas fait l'exercice	N'a pas fait l'exercice
Groupe 4	Non	Oui
Groupe 5	Non	Oui
Groupe 6	Oui	Oui
Groupe 7	Non	Non
Groupe 8	N'a pas fait l'exercice	N'a pas fait l'exercice
Groupe 9	Non	Non
Groupe 10	Non	Oui
Groupe 11	Oui	Oui
Groupe 12	Non	Oui

Tableau 22 : Tableau récapitulatif des réponses des élèves de la classe 1 au problème 1

Nous reviendrons ci-dessous plus en détail sur les arguments utilisés par les élèves.

Problème 2

Le problème 2 consiste à faire une représentation graphique du nuage de points, puis d'utiliser deux points d'une droite pour déterminer son équation dans le but de faire des estimations.

La première question consiste à représenter le nuage de points dans un repère orthogonal. Les dix groupes ont bien représenté le nuage de points. Pour la deuxième question, elle consiste dans un premier temps à placer deux points dans le repère précédent et tracer une droite qui passe par ces deux points. Et nous constatons que les dix groupes ont bien répondu à cette tâche. Dans un second temps on demande aux élèves de déterminer l'équation de cette droite, six groupes (groupes 2 ; 3 ; 5 ; 7 ; 9 et 10) sur les dix ont donné une réponse juste, un groupe (groupe 6) a donné une réponse incomplète et trois groupes (groupes 1 ; 4 et 8) ont donné des réponses erronées. Concernant la troisième question, elle consiste à faire des estimations. Pour la première estimation cinq groupes (groupes 3 ; 5 ; 7 ; 9 et 10) ont proposé une réponse juste, deux groupes (groupes 1 et 4) ont donné des réponses erronées et trois groupes (groupes 2 ; 6 et 8) n'a pas répondu à la question. Quant à la deuxième estimation, trois groupes (groupes 5 ; 9 et 10) ont proposé une réponse juste, trois groupes (groupes 1 ; 3 et 7) ont donné des réponses erronées et les quatre autres groupes (groupes 2 ; 4 ; 6 et 8) n'ont pas répondu à la question.

Questions	Réponse exacte	Réponse incomplète	Réponse erronée	N'a pas répondu
1	10	0	0	0
2	10	0	0	0
	6	1	3	0
3.a	5	0	2	3
3.b	3	0	3	4

Tableau 23: Réponses des 10 groupes à l'exercice 2

Problème 3

Ce problème a pour but de faire une estimation de la tension théorique d'une personne et de comparer avec la tension réelle. Pour ce faire, on demande à l'élève d'utiliser la méthode des moindres carrés pour déterminer l'équation de la droite qui servira à faire l'estimation.

La première question consiste à faire une représentation graphique du nuage de points dans un repère. Deux groupes ont bien réussi la question, deux groupes ont proposé des réponses incomplètes, trois groupes ont proposé des réponses erronées et trois groupes n'ont pas répondu à la question. Pour la deuxième question 2.a, qui consiste à déterminer l'équation de la droite par la méthode des moindres carrés, un seul groupe a trouvé la bonne réponse, deux groupes ont proposé des réponses incomplètes et sept groupes n'ont pas répondu à cette question. Concernant la question 2.b, un groupe a trouvé la bonne réponse, un autre groupe a trouvé une réponse erronée et huit groupes n'ont pas répondu à la question. Quant à la dernière question qui consiste à faire une estimation et de la comparer avec une autre valeur, un seul groupe a trouvé la bonne réponse et les neuf autres groupes n'ont pas répondu à la question, cela est dû au fait que ces groupes n'ont pas trouvé l'équation de la droite qui permet de faire l'estimation. Et le groupe qui a trouvé la bonne réponse a comparé de manière pertinente la valeur de la tension théorique et celle de la tension réelle.

Questions	Réponse juste	Réponse incomplète	Réponse erronée	N'a pas répondu
1	3	2	2	3
2.a	1	2	0	7
2.b	1	0	1	8
3	1	0	0	9

Tableau 24: Réponses des 10 groupes au problème 3

6.3.3.2 Analyse détaillée des réponses des élèves de la classe 1

Problème 1

En ce qui concerne ce problème, pour justifier leur réponse, un bon nombre de groupes d'élèves a d'abord vérifié la réponse en établissant l'équation d'une droite qui leur permet de faire l'estimation, avant de proposer leur justification. Ainsi ces élèves n'ont pas fait ce qui était attendu, c'est-à-dire s'intéresser à la méthode utilisée par l'élève fictif.

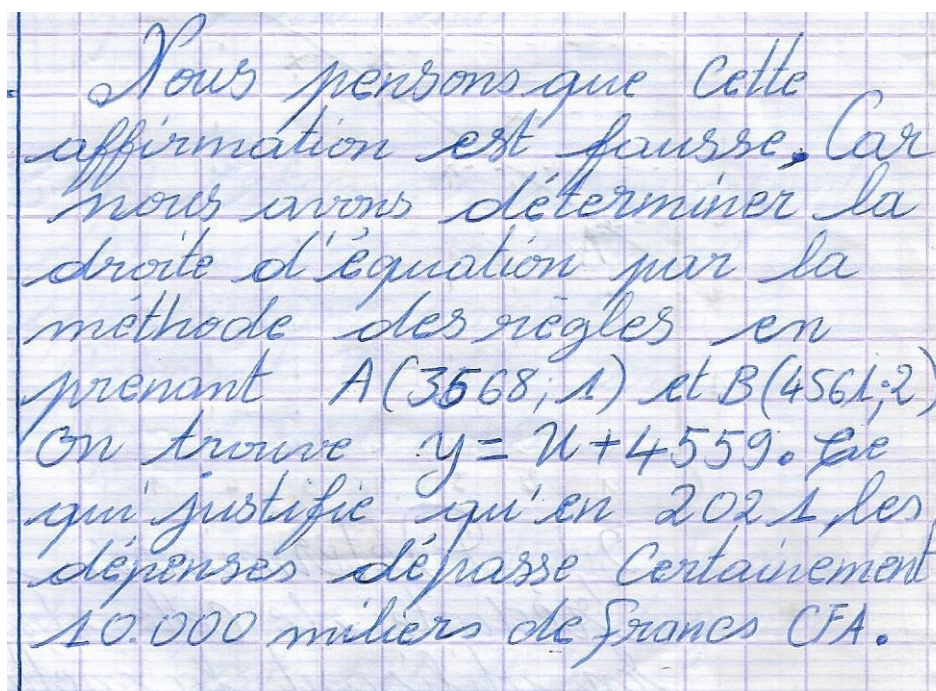
Pour le groupe 1, les élèves disent oui la réponse est correcte. Malgré l'instruction de ne pas faire l'exercice, ils ont fait une estimation au rang de l'année $x=7$ après avoir déterminé l'équation de la droite des moindres carrés. Et nous constatons que l'estimation trouvée dépasse 10000. La justification est correcte, ils ont bien vu que le rang de l'année est 7, et ils ont bien mis des pointillés pour indiquer où est la valeur correspondant à 7. Et ils ont constaté que cette valeur dépasse 10000. D'après les critères donnés dans la partie méthodologique, nous observons une prise d'initiative dans leur justification, car bien qu'ils aient utilisé la méthode des moindres carrés, leur justification est basée sur la lecture du graphique proposée dans l'énoncé (mais parlent par erreur de mois au lieu d'année).

En 2021 soit disant 7^e mois le ménage dépensera plus de 10000 F. L'affirmation du ménage n'est pas la bonne. Donc les élèves ont raison puisque sur le graphique au 7^e mois, on constate que la projection de $x=7$ sur y_7 dépasse 10000 F.

Figure 36: Extrait de copie du groupe 1

Pour le groupe 2, les élèves de ce groupe considèrent que la réponse n'est pas correcte. Pour eux, en 2021 le rang de l'année sera toujours le même rang qu'en 2020, c'est-à-dire $x=6$ et la dépense en 2021 sera aussi 9120. Le fait de dire que le rang de l'année en 2021 est le même qu'en 2020, nous disons que ces élèves n'ont pas compris le tableau statistique donné.

Concernant le groupe 4, les élèves disent oui et pour justifier leur réponse, les élèves de ce groupe ont déterminé l'équation de la droite par la méthode à la règle. Ils ont pris les deux premiers points dans le tableau (années 1 et 2). Ceci est une prise d'initiative, qui peut résulter de l'enseignement expérimental. Cependant ils se sont trompés dans l'équation car ils ont trouvé $y=x+4559$. Ensuite, ils confirment que les dépenses dépassent 10000 en 2021, sans effectuer des calculs. Mais si nous faisons l'estimation en l'an 2021 avec l'équation de la droite qu'ils ont trouvée, nous aurons $y=4566$. Ce montant sera inférieur à 10000, ce qui contredit la justification des élèves de ce groupe. Bien que nous ayons, dans l'enseignement expérimental, insisté sur l'importance de contrôler les ordres de grandeur, ce groupe n'a pas appliqué ce contrôle à son résultat. La prise d'initiative sur le contrôle de l'ordre de grandeur n'est pas présente.



Nous pensons que cette affirmation est fautive, car nous avons déterminé la droite d'équation par la méthode des règles en prenant $A(3568; 1)$ et $B(4561; 2)$. On trouve $y = x + 4559$. Ce qui justifie qu'en 2021, les dépenses dépassent certainement 10.000 milliers de Francs CFA.

Figure 37: Extrait de copie du groupe 4

Pour le groupe 5, les élèves de ce groupe disent oui la réponse est correcte. Pour la justification, ils ont trouvé l'équation de la droite d'ajustement par la méthode à la règle qui est : $y=1139,75x+2281,5$. Ces élèves ont eu une prise d'initiative dans leur méthode, en choisissant deux points qui sont bien situés, par rapport à leur représentation du nuage de points. En revanche, leur calcul d'équation est faux. Pour l'estimation des dépenses en 2021 au rang $x=7$, ils ont trouvé $y=10259,75 > 10000$. Donc la justification de ces élèves se base sur la méthode à la règle. Nous constatons que ces élèves ne se contentent pas seulement de donner leur avis sur la méthode graphique proposée par cet élève mais, ils utilisent la méthode d'ajustement vue en classe pour vérifier la réponse proposée par cet élève.

En ce qui concerne le groupe 6, les élèves disent oui, la réponse est correcte l'estimation des dépenses en 2021 dépasse 10000. Pour la justification, les élèves de ce groupe font allusion au graphique sans donner des explications.

Pour le groupe 7, les élèves considèrent que la réponse n'est pas juste, car le nuage de points ne semble pas aligné mais ils affirment que l'estimation dépasse 10000. Nous constatons que ce groupe n'a pas bien explicité sa réponse. Ces élèves ont soulevé deux objections, le nuage de points ne semble pas aligné mais si nous observons bien le graphique établi par cet élève, nous constatons qu'effectivement le nuage de points semble aligné et la droite tracée approche bien le nuage de points. Par ailleurs ces élèves confirment que l'estimation dépasse 10000.

Pour le groupe 9, les élèves disent non, la réponse n'est pas juste, ils ont appliqué la méthode des moindres carrés pour déterminer l'équation de la droite pour faire l'estimation. L'équation trouvée est $y=354,72x+5039,28$ et l'estimation est : $y=7522,32$. Nous disons que l'estimation proposée par les élèves du groupe 9 n'est pas juste, cependant ils ont une prise d'initiatives dans la méthode utilisée pour la vérification de la réponse proposée dans cet exercice.

Pour le groupe 10, les élèves disent oui, la réponse est correcte, ils appliquent la méthode des moindres carrés pour déterminer la droite qui leur permet de faire l'estimation. Pour l'estimation, ils ont trouvé $y=10357,24$. Et ils confirment que l'estimation des dépenses dépasse 10000 en l'an 2021. Pour leur justification, le groupe 10 aussi utilise la méthode des moindres carrés.

The image shows a student's handwritten work on a grid background. At the top, the equation $y = 1163,26x + 2194,42$ is written and boxed. Below it, the student writes "En 2021 $x=7$ ". Then, the equation $y = 1163,26 \times 7 + 2194,42$ is written. This is followed by another boxed equation: $y = 10357,24$. At the bottom, there is a handwritten justification in French: "En 2021 l'estimation des dépenses dépasse est de $y = 10357,24$ Donc la prévision du ménage est mauvaise, car l'estimation de 2021 dépasse 10000."

Figure 38: Extrait de copie du groupe 10

Quant aux groupes 3 et 8, ils n'ont pas répondu à l'exercice 1.

En résumé, nous retenons que les cinq groupes qui considèrent que la réponse est juste (groupes 1, 4, 5, 6 et 10), ont utilisé soit la méthode des moindres carrés (groupes 1 et 10), soit la méthode au jugé (groupes 4 et 5) pour justifier leur réponse, et seul le groupe 6 a utilisé la méthode graphique. Quant aux trois groupes (groupes 2, 7 et 9) qui considèrent que la réponse n'est pas juste, ou pas correcte, ils (groupes 2 et 7) ont proposé des justifications fausses ou erronées (groupe 9).

Pour rappel, l'important pour cet exercice n'est pas qu'ils utilisent la méthode graphique pour déterminer une estimation. C'est plutôt qu'ils se fondent sur le graphique donné dans l'énoncé : en effet, on leur demande de se prononcer sur la méthode employée par de groupe fictif. Force est de constater que cinq groupes sur les dix ont fait l'exercice et conclu sur leur résultat plutôt que de se prononcer sur la méthode employée par le groupe fictif. Nous considérons ceci comme un effet du contrat didactique : en effet, cet exercice est inhabituel, et plutôt que d'effectuer la tâche demandée, les élèves ont effectué la tâche habituelle qui consiste à déterminer une estimation.

Problème 2

La première question consiste à faire la représentation du nuage de points dans un repère orthogonal, les valeurs sont simples à manipuler donc le choix de l'échelle n'est pas difficile. Tous les groupes d'élèves ont réussi ce type de tâches. Le choix de l'échelle relève d'une prise d'initiative, mais celle-ci reste limitée car les valeurs du tableau statistique sont simples à manipuler.

La deuxième question consiste à placer deux points A et B dans le repère et de tracer une droite passant par ces deux points. Pour cette tâche, tous les groupes ont proposé une réponse juste. Nous constatons que les élèves maîtrisent bien les savoirs préalables sur les droites, ces savoirs consistent à tracer à l'aide d'une règle une droite passant par deux points. Par la suite on leur demande de déterminer l'équation de cette droite. Pour déterminer l'équation de la droite, les élèves doivent se servir des coordonnées de ces deux points pour trouver les coefficients de l'équation de cette droite. Nous rappelons que cette procédure n'est pas travaillée en classe de terminale, cependant elle fait partie des savoirs travaillés à la fin du collège. Six groupes (groupes 2, 3, 5, 7, 9, 10) d'élèves ont trouvé la bonne réponse. Les groupes 2, 3, 5 et 9, se sont servi des coordonnées des points A et B pour déterminer le coefficient directeur ensuite ils ont déduit la valeur du deuxième coefficient.

$$a = \frac{y_B - y_A}{x_B - x_A} = \frac{8,9 - 3}{9 - 1} = \frac{5,9}{8} = 0,65$$

$$a = \underline{\underline{0,65}}$$

$$b = y_A - ax_A = 3 - 0,65 \times 1 = 2,35$$

$$b = \underline{\underline{2,35}}$$

$$\text{d'où : } \boxed{y = 0,65x + 2,35}$$

Figure 39: Extrait de copie du groupe 5

Quant aux groupes 7 et 10, ils ont utilisé un système d'équations de deux équations à deux inconnues pour déterminer l'équation de la droite. Mais dans la mise en équation le groupe 7 a utilisé la méthode de Mayer pour déterminer l'équation, donc ce groupe n'a pas utilisé les coordonnées des points A et B. Ces groupes (groupes 2, 3, 5 et 9) ont des potentialités en terme de

prise d'initiative, ils ont utilisé une procédure non apprise en classe pour trouver l'équation de la droite.

La troisième question consiste à utiliser l'équation de la droite trouvée précédemment pour faire des estimations. Pour le premier type de tâche qui consiste à estimer le montant des dépenses durant le quatorzième mois. En d'autres termes, on demande aux élèves de trouver y lorsque $x=14$, en utilisant soit l'équation de la droite ou soit le graphique. Pour ce type de tâche, les groupes 3, 5, 7, 9 et 10 ont proposé une estimation juste. Ils ont tous utilisé l'équation de la droite pour faire l'estimation et les groupes 3, 5, 7 et 9 ont proposé des justifications correctes sur leur résultats. En terme de prise d'initiative, nous disons que ces groupes ont proposé un modèle qui est juste, ils ont des potentialités en terme de prise d'initiatives. Les groupes 1 et 4, ont trouvé des réponses erronées du fait que leur équation de la droite est fautive. Cependant le modèle d'estimation proposé par ces groupes est juste. De notre point de vue, le fait de proposer un modèle d'estimation juste relève d'une prise d'initiative de la part des élèves. Quant aux groupes 2, 6 et 8, ils n'ont pas répondu à la question.

Pour le deuxième type de tâche qui consiste à estimer le rang du mois au cours duquel le montant dépassera pour la première fois 13 millions de francs CFA. Autrement dit, on demande aux élèves de trouver la valeur de x lorsque $y=13$. Les groupes 5, 9 et 10 ont proposé un modèle juste, ils ont pris $y=13$ et ils ont trouvé la bonne réponse. Ils ont utilisé l'équation de la droite pour trouver l'estimation. Cela relève d'une prise d'initiative de la part de ces élèves. Les groupes 1, 3 et 7 ont proposé des réponses erronées, ils ont confondu les valeurs de x et y . Nous avons prévu cette erreur dans l'analyse a priori. Cependant le modèle d'estimation proposé est juste.

3-b) Estimons le rang du mois, au cours duquel le montant dépassera pour la 1^{ère} fois 13 M.
pour $n=13$
 $y = 1,02(13) + 0,05 = 13,31$.

Donc c'est au 14^e mois que le montant dépassera 13 millions de francs CFA puisque $y = 14,33$ M.

Figure 40: Extrait de copie du groupe 1

Quant aux autres groupes (groupes 2, 4, 6 et 8), ils n'ont pas répondu à la question. Nous constatons qu'aucun groupe n'a utilisé la procédure graphique pour faire l'estimation. Ceux qui ont proposé un modèle juste pour faire l'estimation, cela relève d'une prise d'initiative.

Problème 3

La première question concerne la représentation du nuage de points dans un repère dont l'origine est le point de coordonnées (30, 10). Pour cette question l'échelle est proposée, et ceci ne relève pas d'une prise d'initiative. Cependant la représentation des points du nuage dont l'origine du repère est (30, 10) relève d'une forme de prise d'initiative. Trois groupes (groupes 3, 5 et 10) ont proposé une représentation juste en tenant compte de l'origine. Les groupes 7 et 8 ont bien représenté le nuage de points mais ces groupes n'ont pas tenu compte que l'origine du repère est (30, 10). Donc ils ont proposé une réponse erronée.

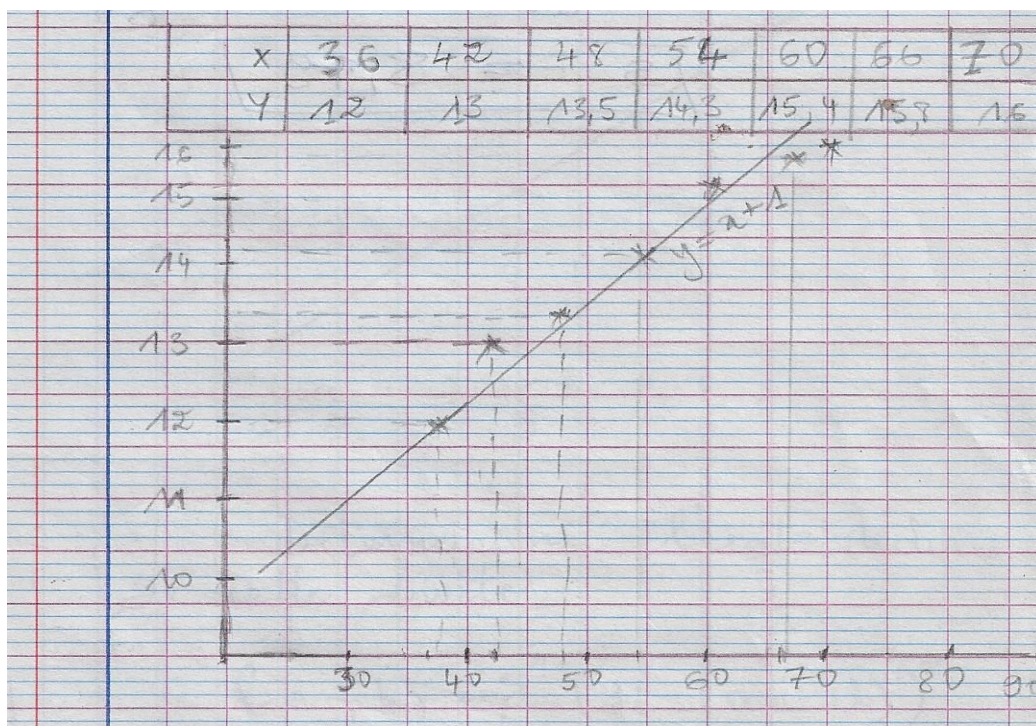


Figure 41: Extrait de copie du groupe 8

Deux groupes (groupes 1 et 2) ont proposé des réponses incomplètes. Quant aux groupes 4, 6 et 9, ils n'ont pas répondu à la question.

La deuxième question consiste à déterminer l'équation de la droite, en utilisant une méthode apprise en classe. Puis de représenter cette droite sur le repère précédent. Un seul groupe (groupe 10) a trouvé une réponse juste, les élèves de ce groupe ont trouvé la bonne équation de la droite par la méthode des moindres carrés. Ils ont justifié leur réponse, en disant que les points du nuage sont presque alignés. En ce qui concerne la représentation de cette droite, ils ont fait un tableau de valeurs de x et y qui permet de tracer la droite. De notre point de vue, ces élèves ont bien compris la question et ils ont utilisé une méthode apprise en classe pour déterminer l'équation de la droite, et ensuite ils ont représenté cette droite dans le repère précédent. Ceci relève d'une prise d'initiative de la part de ces élèves. Les groupes 5 et 8 ont proposé des réponses incomplètes. Quant aux groupes 1, 2, 3, 4, 6, 7 et 9, ils n'ont pas répondu à la question.

Concernant la dernière question qui consiste à faire une estimation et de la comparer avec une autre valeur. Seul le groupe 10 a proposé un modèle juste et les élèves de ce groupe ont trouvé la bonne réponse $y=16,83$.

3) Pour $\alpha = 75$

$$y = 0,12\alpha + 7,83$$

$$y = 0,12(75) + 7,83$$

$$y = 16,83$$

$$16,83 < 17,5$$

On constate que la tension théorique est inférieure à la tension réelle

Figure 42: Extrait de copie du groupe 10

La procédure utilisée pour trouver l'estimation est correcte. Et même la procédure pour comparer la tension théorique et la tension réelle est juste. Donc la procédure utilisée par ces élèves relève d'une prise d'initiative. Toute fois, il faut tout de même noter ici que quand on dit «comparer», on n'est pas dans le contrat didactique habituel qui consiste à mettre deux nombres en regard et dire lequel est le plus grand. Il s'agit de dire si l'estimation finalement est bonne ou pas, mesurer l'erreur d'estimation. Les élèves qui ont répondu à la question ne se sont pas placés dans cette perspective de mesure d'une erreur. Quant aux autres groupes (groupes 1, 2, 3, 4, 5, 6, 7, 8 et 9) ils n'ont pas répondu à la question.

Nous rappelons que nous avons précédemment fait passer ce test à d'autres élèves, qui avaient suivi un enseignement « traditionnel ». Nous pouvons ainsi comparer les résultats à ce test suite à un enseignement traditionnel, et suite à cet enseignement spécifique que nous avons conçu et mis en œuvre. Nous observons certaines différences, qui nous conduisent à faire les hypothèses suivantes :

- Les élèves semblent faire de meilleures représentations, en terme de représentations du nuage des points et de la droite d'équation ;
- La plupart des élèves maîtrisent le sens de la démarche d'estimation. Pour faire l'estimation, les élèves déterminent d'abord l'équation de la droite, soit par la méthode des moindres carrés ou soit par la méthode de Mayer. Ensuite ils se servent de cette équation pour estimer la variable demandée.
- Les élèves semblent pour la plupart faire plus attention aux ordres de grandeur, ils font moins d'erreurs de calcul. La plupart de ces élèves commencent à avoir le sens du contrôle des ordres de grandeur : cet aspect avait particulièrement été souligné dans le cours.

6.3.4 Analyse du post-test dans la classe 2 (classe témoin)

6.3.4.1 Analyse globale

Nous avons fait passer ce post-test dans un établissement privé (Lycée Saint Philippe) car nous connaissons le collègue qui enseigne dans cet établissement et qui était volontaire pour le faire passer auprès de ses élèves de la classe de terminale SE. L'effectif des élèves est de dix. Nous avons fait cinq groupes de deux élèves, chaque groupe est équipé au moins d'une calculatrice. Et le test a duré deux heures. Nous désignons les groupes d'élèves par G1, G2, G3, G4 et G5.

Nous soulignons la différence de conditions de travail, entre le lycée public où se trouvait la classe 1, avec un effectif de 69 élèves, et ce lycée privé où les effectifs sont réduits.

Problème 1

Pour ce problème, les cinq groupes d'élèves ont répondu à la question. Les réponses proposées par ces cinq groupes sont justes. Cependant sur les cinq groupes, trois groupes (G1, G2, G3) ont commenté la réponse de l'élève (fictif), par contre les autres groupes (G4 et G5) ont utilisé la technique de Mayer pour justifier leur réponse.

Nous récapitulons ces types de réponses dans le tableau ci-dessous.

Groupe N°	A pris en considération la méthode	A dit que la production était juste
G1	Oui	Oui
G2	Oui	Oui
G3	Oui	Oui
G4	Oui	Oui
G5	Non	Oui

Tableau 25: Tableau récapitulatif des réponses des élèves de la classe 2 au problème 1

Nous reviendrons ci-dessous plus en détail sur les arguments utilisés par les élèves.

Problème 2

La première question consiste à représenter le nuage des points dans un repère orthogonal. Les cinq groupes ont bien représenté le nuage des points mais le groupe G1 joint les points du nuage par des segments de droite. La deuxième question consiste à tracer une droite passant par deux points A et B ensuite, de déterminer l'équation de cette droite. Pour le traçage de la droite, les cinq groupes ont bien représenté la droite et ils ont trouvé l'équation de la droite. Cependant Les groupes G2 et G3 ont utilisé la méthode des moindres carrés pour déterminer l'équation de la droite. Concernant la troisième question, elle consiste à faire des estimations. Le 3.a) consiste à estimer la valeur de y à partir de la valeur de x , quatre groupes (G2, G3, G4 et G5) ont donné des réponses justes tandis que un groupe (G1) a donné une réponse erronée. Pour le 3.b) on demande aux élèves de prédire la valeur de x à partir de la valeur de y . Les groupes G1 et G3 ont donné des réponses erronées, Les groupes G4 et G5 ont donné des réponses justes et le groupe G2 n'a pas répondu à la question.

Problème 3

La première question est de représenter les points du nuage dans un repère dont l'origine est le point de coordonnées (30 ;10). Le groupe G2 a donné une représentation juste du nuage des points ainsi

que l'origine du repère est bien représentée, quant aux groupes G1, G3, G4 et G5 ces groupes ont fait une représentation erronée, soit du nuage des points (G3) ou soit de l'origine du repère (G1, G4 et G5). Concernant la deuxième question, elle consiste à déterminer la droite d'ajustement par la technique des moindres carrés ensuite de représenter cette droite. Les groupes G1, G2 et G4 ont donné des réponses justes par contre les groupes G3 et G5 ont donné des réponses erronées. Quant à la dernière question qui consiste à prédire la tension théorique et de la comparer avec la tension réelle. Les groupes G1 et G4 ont donné la bonne estimation de la tension par contre les groupes G2, G3 et G5 ont donné des réponses erronées, ces groupes ont fait des erreurs de calcul.

6.3.4.2 Analyse détaillée

Problème 1

Concernant cet exercice, les réponses proposées par les cinq groupes sont correctes. Les groupes G1, G2 et G3 ont proposé des justifications correctes. Ils s'appuient sur la justification graphique pour répondre à la question. En ce qui concerne les groupes G1 et G2, les élèves de ces groupes ont constaté d'après la représentation de la droite d'ajustement que l'estimation des dépenses dépassera 10 000 milliers de francs CFA. Par ailleurs ces deux groupes ont reproduit le graphique de l'élève fictif. Par contre le groupe G3 s'est focalisé sur la justification. Selon des critères donnés dans la partie méthodologique, les justifications proposées par ces trois groupes relèvent d'une prise d'initiatives.

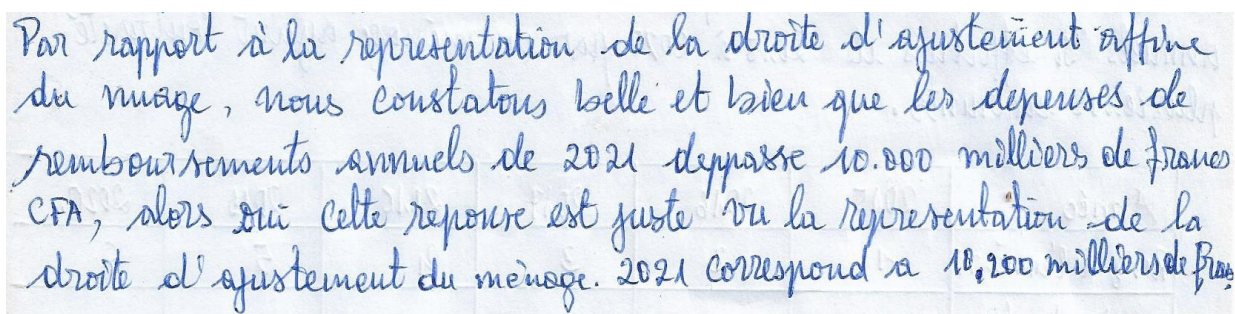


Figure 43: Extrait de copie du groupe G1

Quant aux groupes G4 et G5, les élèves ont proposé des réponses correctes. Mais, pour eux la technique graphique n'est pas suffisante. Ils ont utilisé la technique de Mayer pour montrer que le ménage dépensera plus de 10000 milliers de francs cfa. Par ailleurs le groupe G4 se base sur la représentation de la droite d'ajustement pour conclure que la réponse de l'élève fictif est juste.

Problème 2

Concernant la première question qui consiste à faire la représentation du nuage de points dans un repère orthogonal. Les cinq groupes d'élèves ont bien représenté les points du nuage dans le repère. Par ailleurs, le groupe G1 relie les points du nuage par des segments de droites et il obtient un polygone qui ne fait pas partie du type de tâches demandé. Pour la deuxième question, le type de

tâches demandé aux élèves consiste à tracer une droite passant par deux points et ensuite de déterminer l'équation de cette droite. Les cinq groupes ont réussi à tracer une droite qui passe par les deux points A et B. Ce type de tâche ne relève pas d'une prise d'initiative. Quant à la détermination de l'équation de la droite, tous les groupes ont trouvé la bonne réponse. Les groupes G1, G4 et G5 ont utilisé les coordonnées des points A et B pour déterminer le coefficient directeur a et ils se sont servis des coordonnées de l'un des deux points pour déterminer le coefficient b.

Déterminons l'équation réduite de cette droite. on a : A(1; 3) et B(9; 8,2)

On soit (AB) : $y = ax + b$

* Déterminons a

$$a = \frac{y_B - y_A}{x_B - x_A} = \frac{8,2 - 3}{9 - 1} = \frac{5,2}{8} = \frac{1,3}{2}$$

$a = \frac{1,3}{2}$

* Déterminons b

On sait que A ∈ (AB) d'où (AB) : $y_A = \frac{1,3}{2}x_A + b \Leftrightarrow$

$$3 = \frac{1,3}{2} \times 1 + b$$

$$b = 3 - \frac{1,3}{2} = \frac{6 - 1,3}{2} = 2,35$$

$b = 2,35$ ③

Figure 44: Extrait de copie du groupe G1

Pour la troisième question 3.a, on demande aux élèves d'estimer le montant des dépenses durant le quatorzième mois en se servant de l'équation de la droite déterminée précédemment. Les groupes (G2, G3, G4 et G5) ont proposé des estimations justes. Le modèle utilisé par ces groupes relève d'une prise d'initiatives. Nous rappelons que c'est la démarche que nous avons prévue dans l'analyse a priori. Quant au groupe G1, les élèves de ce groupe ont proposé une estimation fautive, ils ont utilisé la méthode graphique pour avoir le montant correspondant au quatorzième mois. Ensuite ils expliquent leur réponse littéralement sans faire de calcul. Même si la valeur trouvée est fautive mais ce groupe d'élèves a une prise d'initiative importante pour ce type de tâches.

usage des premiers mois.

a) Estimons le montant des dépenses durant le quatorzième mois.

Par rapport à la droite d'ajustement, nous estimons le montant des dépenses durant le quatorzième mois à 10,6 millions de francs en

Figure 45: Extrait de copie du groupe G1

Pour la question 3.b, elle consiste à estimer le rang du mois, lorsque le montant dépassera pour la première fois 13 millions de francs CFA. Les groupes G4 et G5 ont proposé une estimation juste, ils ont remplacé y par 13 dans l'équation de la droite précédente pour trouver le rang du mois. Cela relève d'une prise d'initiative de la part de ces élèves. D'autre part le groupe G4 a proposé une justification correcte. Quant au groupe G5 il a proposé une justification fautive, ce groupe considère x comme le montant en millions de francs CFA.

$R = 16,38$
 Donc on estime que le montant dépassera la première fois 13 millions de fcfa à 16,38 millions de fcfa

Figure 46: Extrait de copie du groupe G5

Les élèves du groupe G1 ont proposé une réponse fautive, ils ont utilisé une fois de plus encore la méthode graphique pour l'estimation. Ils ont trouvé l'estimation du rang du mois $y=18$. Ce groupe a une prise d'initiative importante. Le fait d'utiliser la méthode graphique pour faire l'estimation nécessite aux élèves de prendre de recul pour utiliser, car c'est une méthode qui est différente de types habituellement travaillé en classe.

Le groupe G3 a proposé une réponse fautive. Ces élèves $y=14,27$ n'ont pas expliqué leur technique, ils ont trouvé l'estimation du rang du mois sans présenter les détails de leur calcul.

Quant au groupe G2, les élèves de ce groupe n'ont pas répondu à la question.

Problème 3

La première question est de faire une représentation graphique des points du nuage dans un repère non classique. Nous entendons par repère non classique un repère dont l'origine n'est pas le point de coordonnées $(0, 0)$. Dans notre cas, l'origine du repère sur l'axe des abscisses débute à 30 et sur l'axe des ordonnées débute à 10. Un seul groupe (G2) a bien représenté l'origine de coordonnées $(30 ; 10)$. Et les points du nuage représentés par ce groupe sont justes. Ce type de tâche relève d'une prise d'initiative même s'il est limité. Quant aux groupes G1, G4 et G5 ils ont bien représenté le nuage de points, par contre ils n'ont pas pris comme l'origine du repère le point de coordonnées $(30 ; 10)$. Pour le groupe G3, les élèves ont bien représenté l'origine du repère mais ils n'ont pas bien représenté les points du nuage.

Concernant la deuxième question, qui consiste à déterminer l'équation de la droite par la méthode des moindres carrés, puis de la placer sur le graphique précédent. Pour la méthode des moindres carrés, cela est supposé être travaillé en classe déjà. Les groupes G1, G2 et G4 ont utilisé la procédure apprise en classe, en faisant un tableau à double entrée pour déterminer les coefficients de l'équation de la droite de régression de y en x . Ces trois groupes ont trouvé une équation de la droite juste. Par contre les groupes G3 et G5 ont trouvé des réponses erronées. Ces deux groupes n'ont pas utilisé le tableau statistique à double entrée pour calculer la covariance et la variance. Et ils ont fait l'erreur de calcul. En ce qui concerne la représentation de la droite de l'équation, les

groupes G1, G2 et G4 ont fait une représentation de la droite qui est correcte. Par contre les groupes G3 et G5 ont fait une représentation qui n'est pas juste.

La dernière question consiste à utiliser la droite des moindres carrés pour trouver la tension théorique et de comparer cette tension avec une tension réelle. Pour cela les groupes G1 et G4 ont utilisé une procédure qui est correcte. Ils ont remplacé l'âge qui correspondait à $x=75$ dans l'équation pour trouver la tension théorique. Par ailleurs, la procédure utilisée pour comparer ou discuter la tension théorique avec la tension réelle est correcte.

3- Calcul de la tension théorique par rapport à la droite.

* A 75 ans, $x = 75$.

$$y = 0,12 \times 75 + 7,83$$

$y = 16,83$

* $17,5 > 16,83$, la tension réelle est plus élevée que la tension théorique.

Figure 47: Extrait de copie du groupe G4

Quant aux groupes G2, G3 et G5 ces groupes ont proposé des réponses erronées, ils ont proposé des modèles justes mais ils ont fait des erreurs de calcul.

En ce qui concerne l'interprétation du résultat, les cinq groupes ont comparé leur résultat avec la tension réelle. Les groupes G1, G2 et G5 ont trouvé la tension théorique inférieure à la tension réelle tandis que les groupes G3 et G4 ont trouvé la tension théorique supérieure à la tension réelle.

6.3.5 Conclusion à propos de la troisième expérimentation

Concernant l'analyse portant sur l'implémentation de l'enseignant expérimental, les élèves font moins d'erreurs de calcul, ils maîtrisent la démarche d'estimation. L'enseignant privilégie l'interaction pendant le cours mais le fait de beaucoup intervenir quand un élève est au tableau réduit leur prise d'initiative.

En ce qui concerne le pré-test et post-test dans la classe 1, les élèves font plus d'erreurs lors du calcul de la variance mais pour justifier leur réponse, ils font des propositions pertinentes. Pendant le post-test les élèves maîtrisent la représentation graphique et la démarche d'estimation. Ils font plus attention aux ordres de grandeur et font moins d'erreurs de calcul.

Concernant la classe 1 et la classe 2. Nous rappelons que dans cette partie, nous comparons la production des élèves dans les deux classes. La première classe nommée « classe 1 » est la classe où nous avons enseigné le cours portant sur la notion de la série statistique double nous même au lycée d'application de l'école normale supérieure (LAENS) et la classe nommée « classe 2 » est une classe de terminale dont l'enseignement de la série statistique à deux variables est fait par un

collègue, sans intervention du chercheur. Nous avons récupéré le contenu de son cours à la fin du post-test.

Concernant le problème 1 du post-test, nous avons constaté que six élèves des deux classes (trois de la classe 1 et trois de la classe 2) ont justifié leur réponse sans faire l'exercice. Dans la classe 1 trois élèves ont utilisé la technique des moindres carrés et deux élèves la technique au jugé, par contre deux élèves ont utilisé la technique de Mayer dans la classe 2, pour justifier leur réponse. Nous signalons que la technique au jugé ou à la règle n'était pas travaillée en classe avec les élèves de la classe 2. Ainsi l'enseignement expérimental, en soulignant l'existence de différentes techniques, a permis aux élèves de la classe 1 de développer une plus grande prise d'initiative quant au choix des méthodes d'ajustement.

En ce qui concerne le problème 2, nous déclinons la comparaison de la production des élèves des deux classes en trois points :

- Le premier point concerne la représentation graphique, les élèves de deux classes ont bien repéré et représenté respectivement les points et le nuage de points dans le repère. Ce savoir-faire fait partie des acquis du collège et du début du lycée.
- Le deuxième point est relatif à la détermination de l'équation de la droite de régression. Pour la classe 1 six groupes ont trouvé une réponse juste, trois groupes ont trouvé des réponses erronées et un groupe a donné une réponse incomplète tandis que les cinq groupes de la classe 2 ont trouvé des réponses justes. Nous déduisons que les élèves de la classe 2 ont plus d'initiatives que ceux de la classe 1 dans la technique permettant de déterminer une équation de la droite.
- Le troisième point concerne l'estimation. Pour la première estimation, dans la classe 1 cinq groupes ont donné une réponse juste, deux groupes ont proposé des réponses erronées et trois groupes n'ont pas répondu. Par contre dans la classe 2 quatre groupes ont bien répondu et un groupe a proposé une réponse erronée. Quant à la deuxième estimation, dans la classe 1 trois groupes ont donné une réponse juste, trois groupes ont proposé des réponses erronées et quatre groupes n'ont pas répondu. Tandis que dans la classe 2 deux groupes ont donné une réponse juste deux groupes ont trouvé des réponses erronées et un groupe n'a pas répondu. A la lumière de cette analyse, nous dirons que les non réponses des élèves peuvent être liées à l'absence d'une technique permettant de trouver l'estimation, car ils ont laissé certaines questions pour aborder le problème suivant. Par ailleurs, dans la classe 2, un groupe a utilisé la méthode graphique pour faire les estimations.

Pour le problème 3, un seul groupe de la classe 1 a répondu à toutes les questions et trois groupes ont répondu seulement à la première question, qui consiste à faire une représentation graphique du nuage de points. Par contre tous les groupes de la classe 2 ont abordé la totalité des questions. Nous faisons l'hypothèse que cela est dû à la mauvaise gestion du temps de la part des élèves de la classe 1. Nous rappelons que chaque groupe de la classe 1 est constitué de cinq ou six élèves, s'ils passaient plus de temps à interagir entre eux sans mettre leur idées sur papier, ils ne pouvaient pas finir les trois exercices. Par contre dans la classe 2, les groupes sont constitués de deux élèves par groupes.

6.4 Synthèse du chapitre 6

Dans ce chapitre, nous avons analysé trois expérimentations mises en œuvre dans la classe de terminale SE en centrafricaine. L'analyse de ces expérimentations nous a permis d'apporter des éléments de réponse à notre troisième question de recherche.

QR3 : Quelles sont les prises d'initiatives effectives des élèves de Terminale en centrafricaine, lorsqu'on leur propose un exercice sur le thème de l'ajustement affine qui permettrait une telle initiative ? A quelles conditions peut-on mettre en œuvre en centrafricaine un enseignement qui laisse plus d'initiatives aux élèves que ce qui est pratiqué dans l'enseignement ordinaire ?

La première expérimentation nous a permis de dégager les éléments suivants :

Nous rappelons que l'objectif de première expérimentation est d'abord de laisser le choix de la méthode d'ajustement affine à la charge des élèves à travers un problème, et ensuite d'observer à travers leurs productions où se situent leurs prises d'initiatives effectives. Concernant le choix de la méthode d'ajustement affine, les élèves ont privilégié la méthode des moindres carrés, c'est la méthode qui est beaucoup travaillée en classe. Ainsi bien que le problème sort du contrat didactique habituel, les élèves sont restés dans ce contrat.

L'analyse relative à la première expérimentation nous a permis de relever les prises d'initiatives effectives des élèves pour les types de tâches suivants : T5, T4.3 et T_Rep_Nuage. Ces prises d'initiatives sont en fait aussi celles que nous avons repérées comme prises d'initiatives potentielles dans le chapitre 5. Le problème ne sortait pas suffisamment du contrat didactique habituel.

Dans la deuxième expérimentation, l'objectif consistait à indiquer la prise d'initiatives des élèves à travers trois problèmes qui leur sont proposés.

Nous résumons ainsi nos principales observations.

Nous avons proposé un problème 1 modifié à partir du problème de l'année précédente. En demandant aux élèves de se prononcer sur la validité d'une production d'un autre élève, nous sommes vraiment sorti des formats habituels d'exercice. Cependant nous n'avons pas complètement atteint nos objectifs :

- Les élèves ne discutent pas la méthode graphique utilisée par l'élève fictif (exercice 1) pour tracer la droite d'ajustement affine. Ils cherchent à savoir si le résultat de l'élève fictif est juste ou non. Ainsi, certains ont appliqué la méthode des moindres carrés, et ont jugé la validité de la production de l'élève non pas à partir de la méthode employée, mais du résultat obtenu.
- La validation des réponses au sein d'un même groupe se fait, lorsque deux élèves trouvent deux résultats identiques. Cette méthode de validation est plus importante pour eux que la validation par le contrôle de l'ordre de grandeur du résultat ;

Concernant le problème 2 où il y avait une pondération, les élèves n'ont pas compris le sens du tableau statistique avec pondération, la majorité n'a pas tenu compte de la pondération.

Nous avons constaté à travers l'analyse des productions d'élèves que les prises d'initiatives des élèves restent limitées. Ils ont appris à appliquer une méthode. Ils n'ont pas d'éléments leur permettant de contrôler leurs procédures : contrôle des ordres de grandeur, compréhension du sens

des concepts d'ajustement et de point moyen. Cependant, nous avons observé une prise d'initiative de la part de certains groupes (groupes 1 et 4) dans l'exercice 1. Ces groupes se basent sur la représentation et la lecture graphique pour justifier et interpréter le résultat. D'autre part dans le problème 3, le groupe 1 utilise la méthode graphique pour faire le type de tâches T5, c'est une prise d'initiative importante de la part de ce groupe.

Concernant la troisième expérimentation en classe, nous avons soumis les élèves d'une classe à un pré-test portant sur la série statistique à une variable, ensuite nous avons implémenté l'enseignement expérimental portant sur la série statistique à deux variables. Et à la fin ces mêmes élèves étaient soumis à un post-test. Parallèlement à cela, nous avons passé le même post-test dans une autre classe de terminale SE dans un autre établissement.

Nous résumons nos principales observations de la troisième expérimentation dans ce qui suit.

Nous retenons du pré-test :

D'une part, les élèves ont des connaissances à propos des séries statistiques à une variable qui seront utiles pour notre enseignement des séries statistiques doubles.

D'autre part, ils rencontrent tout de même certaines difficultés ; ils font des erreurs de calcul sur la variance ; ils ont des difficultés à commencer une tâche qui nécessite une large prise d'initiative.

Concernant les prises d'initiatives : les élèves ont des prises d'initiatives dans les types de tâches qui consistent à justifier ou à interpréter le résultat. Mais ils n'ont pas de prise d'initiatives dans les types de tâches atypiques qui nécessitent de prendre du recul pour répondre.

Nous avons observé où se situent des prises d'initiatives possibles dans l'enseignement conçu, mais pendant la mise en œuvre de cet enseignement expérimental les prises d'initiative des élèves sont réduites. Lorsque l'enseignant intervient pour répondre à une question ou lorsque l'enseignant est au tableau, il réduit la prise d'initiatives des élèves. Par ailleurs, nous avons aussi observé des prises d'initiative de la part des élèves, comme le choix des deux points de la droite tracée qui permet de déterminer les coefficients dans l'équation de la droite d'ajustement affine.

Concernant le post-test. Nous avons constaté dans la classe 1 que les élèves font moins d'erreurs de calcul par rapport au pré-test. Ils commencent à avoir le sens du contrôle des ordres de grandeur.

L'enseignement expérimental a permis aux élèves de la classe 1 de développer une prise d'initiative quant au choix des méthodes d'ajustement, ils se permettent de faire des ajustements à la règle.

Nous avons relevé dans les deux classes des prises d'initiatives effectives des élèves pour les types de tâches suivants : T5, T4.3 et T_Rep_Nuage, qui sont les mêmes cités dans la première expérimentation. En plus il y a des prises d'initiatives pour l'ajustement à la règle, puisque l'enseignant ne donne pas a priori de critère pour placer la droite. C'est aux élèves de faire le choix de critères, par exemple ils peuvent choisir deux points du nuage et les joindre. Ils peuvent placer le point moyen G du nuage, et tracer une droite qui passe par G et qui équilibre le nombre de points au-dessus et au-dessous de la droite, mais ne passe par aucun point du nuage. Le temps nous a manqué pour faire débattre les élèves sur ces choix, mais nous pourrions le faire dans la suite de nos travaux.

Nous n'avons pas tenté de mettre en place un enseignement où les élèves effectuent par eux mêmes une investigation, comme un Parcours d'Études et de Recherche (Markulin et al., 2021). En effet, en Centrafrique les conditions ne sont pas réunies pour un tel enseignement. Les élèves sont très nombreux dans les classes de l'enseignement public ; ils n'ont pas de calculatrice pour la plupart ; ils n'ont pas accès à Internet pour faire des recherches. C'est pourquoi nous avons choisi modestement de faire quelques modifications par rapport à l'enseignement habituel. Nous pourrions après notre thèse prévoir encore d'autres modifications ; nous y reviendrons dans la partie dédiée aux perspectives.

CONCLUSION GÉNÉRALE ET PERSPECTIVES

Dans cette partie, nous présentons une synthèse des résultats auxquels nous sommes parvenu et nous présentons quelques perspectives que nous envisageons à la suite de ce travail, en termes de recherche mais aussi de formation d'enseignants.

Synthèse d'éléments de réponse aux questions de recherche

Nous commençons par rappeler nos questions de recherche :

- QR.1 : Quel est le savoir à enseigner au sujet des séries statistiques doubles et de l'ajustement affine en centrafricaine (terminale SE), et quelles sont les prises d'initiatives possibles des élèves relativement à ce savoir ?
- QR.2 : Quel est le savoir effectivement enseigné à ce sujet dans l'enseignement ordinaire ?
- QR.3 : Quelles sont les prises d'initiatives effectives des élèves de Terminale en centrafricaine, lorsqu'on leur propose un exercice sur le thème de l'ajustement affine qui permettrait une telle initiative ? A quelles conditions peut-on mettre en œuvre en centrafricaine un enseignement qui laisse plus d'initiatives aux élèves que ce qui est pratiqué dans l'enseignement ordinaire ?

Pour répondre les question QR.1 et QR.2, nous avons construit un modèle praxéologique de référence (MPR) pour les séries statistiques doubles et l'ajustement affine grâce à l'analyse historique de l'ajustement affine, à notre synthèse de travaux, et à l'analyse du programme de terminale SE et du manuel de BTS de France. Dans ce MPR, les principaux types de tâches sont :

- T1 : faire la représentation graphique de la série double et de son point moyen ;
- T2 : Étudier si un ajustement affine est pertinent ;
- T3 : Choisir une méthode d'ajustement affine ;
- T4.1 : réaliser un ajustement affine avec la méthode « à la règle » ;
- T4.2 : réaliser un ajustement affine avec la méthode de Mayer ;
- T4.3 : réaliser un ajustement affine avec la méthode des moindres carrés ;
- T5 : donner une estimation ;
- T6 : Interpréter l'estimation ;
- T7 : Recueillir des données ;
- T8 : Nettoyer des données réelles.

Nous avons pu observer que les types de tâches T3 « choisir la méthode », T7 « recueillir des données » et T8 « Nettoyer des données réelles » ne se retrouvaient pas dans le programme de Centrafricaine, ni dans le manuel de BTS. Nous les avons mentionnées dans notre MPR en nous appuyant sur l'analyse historique et sur les travaux de recherche en didactique de la statistique.

L'analyse du manuel CIAM que nous avons menée à l'aide de ce MPR montre d'ailleurs que ces types de tâches, qui laissent une large part aux prises d'initiatives des élèves n'y sont pas non plus présents. Les élèves en Centrafrique n'ont pas à leur charge la collecte de données, ni la préparation de données réelles pour l'analyse. Ils n'ont pas non plus la liberté de choisir une méthode pour l'ajustement affine, dans le manuel la méthode est donnée. Nous notons par ailleurs qu'il leur serait difficile de comparer des méthodes d'ajustement, en l'absence d'outil informatique adapté.

Cette analyse permet en outre d'observer que dans le type de tâches T4.3, on précise toujours dans l'énoncé des exercices laquelle des deux droites de régression il faut utiliser, entre la droite de régression de y en x ou la droite de régression de x en y . Pour la méthode des moindres carrés, nous avons observé que dans le manuel CIAM la démonstration de la propriété de l'équation de la droite des moindres carrés est donnée (cependant cette démonstration n'est pas nécessairement présentée aux élèves lors du cours). Nous avons aussi relevé l'ajout de trois nouveaux types de tâches en faisant le MPR qui correspond au manuel CIAM : T_PartagerG1G2 et T_Etab_TabS, et la prise au compte du type de tâches T4.1.

Concernant l'interprétation du résultat, lorsqu'on propose aux élèves de commenter le résultat du coefficient de corrélation entre deux variables, ou de commenter la forme d'un nuage des points, ou encore d'interpréter l'estimation obtenue par le modèle, ces types de tâches nécessitent une prise d'initiatives.

Au regard de l'analyse d'un enseignement observé en Centrafrique, nous avons noté que l'introduction des savoirs nouveaux sur les séries statistiques et l'ajustement affine comme la définition de certains concepts, la procédure pour réaliser un ajustement affine, étaient particulièrement rapides. Nous avons aussi constaté que durant le cours, les prises d'initiatives des élèves étaient très limitées. L'enseignant débute son enseignement par une activité. Il semble vouloir mettre en avant le sens de l'ajustement : il peut y avoir un lien entre deux variables, et une manière d'exprimer ce lien pourrait être une équation de droite. Il veut se reposer sur le savoir des élèves à propos des équations de droites. Il pourrait mieux exploiter ce savoir, et présenter la méthode d'ajustement « au jugé » (ou « à la règle ») qui serait cohérente avec cette démarche de développement progressif du sens de l'ajustement. Mais il n'aborde pas du tout cette méthode. Il souhaite introduire d'emblée la méthode des moindres carrés. Mais il n'apporte pas aux élèves le savoir nécessaire, et ne leur laisse pas suffisamment de temps pour faire les tâches qui sont à leur portée. Aucune démonstration n'est faite dans ce cours, les aspects technologico-théoriques se réduisent dans ce cours à cette tentative de mise en avant du sens de l'ajustement. L'essentiel est centré sur le bloc pratico-technique de la praxéologie. Certains élèves prennent toutefois des initiatives en posant des questions, mais les réponses de l'enseignant sont succinctes et ne profitent pas des questions pour engager un échange avec les élèves. Les conditions des classes surchargées en RCA ne se prêtent pas beaucoup aux échanges.

L'analyse du questionnaire et des entretiens auprès des enseignants confirme que les professeurs accordent peu de place aux prises d'initiatives des élèves. Ils préfèrent proposer aux élèves un exercice en utilisant la structuration usuellement présente dans les manuels. Leur objectif semble être avant tout que les élèves apprennent à appliquer une méthode, surtout la méthode des moindres carrés ou la méthode de Mayer. La méthode de l'ajustement « au jugé » ou « à la règle » n'est pas pratiquée par les professeurs. Or elle pourrait être utile pour mettre en évidence le sens de

l'ajustement. Les professeurs responsabilisent les élèves par un jeu de questions-réponses pendant les cours. Cette pratique permet à certains élèves de s'exprimer. Mais elle peut renforcer les inégalités entre les établissements publics dans lesquels les effectifs sont pléthoriques, et les établissements privés dans lesquels les effectifs sont plus réduits.

L'analyse menée au chapitre 6 permet de répondre à la question notée QR.3. En effet, nous avons mené trois expérimentations dans la classe de terminale SE en Centrafrique. Lors de la première expérimentation, où on propose un exercice aux élèves dont le choix de la méthode d'ajustement affine est laissé à la responsabilité de ces derniers. Force est de constater que, même si cet exercice sort du contrat didactique habituel, les élèves ont privilégié la méthode des moindres carrés qui a été travaillée en classe. Cependant, nous avons observé les mêmes prises d'initiatives que celles repérées dans le chapitre 5 pour l'enseignement « ordinaire ». Concernant la deuxième expérimentation, dont le but était d'analyser la prise d'initiative effective des élèves à travers trois exercices qui leur étaient proposés. Pour l'exercice 1, l'objectif n'était pas atteint, car on s'attendait de la part des élèves à la validation ou non de la méthode graphique (« ajustement affine 'au jugé' ») proposée par un élève fictif pour faire l'ajustement affine. Or les élèves se sont plutôt centrés sur le résultat numérique obtenu que sur la méthode employée. La réflexion sur la méthode ne fait pas partie du contrat didactique usuel en Centrafrique. Par ailleurs, lors d'un travail en groupe sur des exercices de recherche d'une estimation, la majorité des élèves se contentent de valider leur résultats, lorsque deux élèves trouvent deux résultats identiques. Ils ne valident pas leur résultats par le contrôle de l'ordre de grandeur du résultat. Mais deux groupes se basent sur la représentation et la lecture graphique pour justifier et interpréter le résultat et c'est une prise d'initiative importante de la part de ces deux groupes. Pour l'exercice 2, où il y avait un tableau statistique avec pondération. Le constat est que la majorité des élèves n'ont pas tenu compte de la pondération dans leur calcul, ils n'ont pas compris le sens du tableau avec pondération. Concernant leur prise d'initiatives effectives dans l'exercice 2, celles-ci restent limitées. Ils se contentent d'appliquer la méthode, sans comprendre le sens des concepts d'ajustement et de point moyen, ni le contrôle des ordres de grandeur. Dans l'exercice 3, un groupe utilise la méthode graphique pour faire l'estimation, c'est une prise d'initiative importante de la part de ce groupe. Dans la troisième expérimentation en classe, les élèves ont été soumis à un pré-test portant sur la série statistique à une variable, ensuite ils ont suivi un enseignement sur la série statistique double, avant d'être soumis de nouveau à un post-test. En plus de cela, les élèves d'une autre classe de terminale SE ont été soumis à ce même post-test. Concernant le pré-test sur les séries statistiques à une variable, les élèves font des erreurs de calcul sur la notion de la variance et ils ont des difficultés à commencer une tâche qui nécessite une large prise d'initiative. Mais ils ont des prises d'initiatives quand-il s'agit de justifier ou interpréter le résultat. Lors de la conception de l'enseignement expérimental, nous avons observé des prises d'initiatives possibles, mais lors de la mise en œuvre la prise d'initiative des élèves est réduite. A chaque fois que l'enseignant intervient soit pour répondre à une question ou soit au tableau, il réduit la prise d'initiative des élèves. Mais pendant l'enseignement expérimental, l'enseignant avait pris en compte certaines difficultés des élèves relevées dans le pré-test. Il met en évidence la position du point moyen dans le cas de la pondération, qui permettra aux élèves d'avoir le sens du contrôle des ordres de grandeur. Pour le post-test, dans la classe 1, nous avons constaté que les élèves commencent à avoir le sens du contrôle des ordres de grandeur, ils font moins d'erreurs que dans le pré-test. Ils ont développé la prise d'initiatives quant au choix des

méthodes d'ajustement affine, ils se permettent de faire des ajustements à la règle. En plus de cela, nous avons relevé des prises d'initiatives pour l'ajustement à la règle concernant le choix des deux points de la droite et le traçage de la droite à la règle. Et enfin, nous avons constaté dans les deux classes, des prises d'initiatives effectives des élèves pour les types de tâches qui sont les mêmes que ceux cités dans la première expérimentation.

A travers ce bilan, nous retenons que, les expérimentations permettent de voir que même si on propose aux élèves des exercices où une prise d'initiative est possible, ceux-ci ne développent pas toutes les initiatives possibles. Il est difficile de faire sortir les élèves du contrat didactique habituel, avec un enseignement expérimental étalé sur quelques heures. Certains élèves sont capables de prise d'initiatives, ces derniers arrivent à trouver des modèles adéquats qui permettent de résoudre l'exercice. Même si la réponse est erronée, il peut y avoir des prises d'initiatives dans leur procédure. D'autre part, une majeure partie des élèves ne développent pas de prise d'initiatives pour proposer une modélisation adéquate. La plupart des élèves n'ont pas été habitués à avoir des initiatives, et de plus ils n'ont pas les outils informatiques qui leur permettraient de contrôler la validité d'une démarche originale.

Cependant il nous semble important d'avoir pu attirer l'attention des élèves sur quelques éléments : le sens de l'ajustement affine, avec la mise en avant de l'ajustement « à la règle » avant l'introduction systématique de la méthode des moindres carrés ; la maîtrise des ordres de grandeur, pour contrôler la validité d'un résultat ; l'interprétation des résultats numériques obtenus, dans le contexte réel de la situation de départ.

Perspectives à la suite de cette thèse

Perspectives pour la recherche

Nous souhaitons poursuivre notre recherche dans le domaine de la didactique de la statistique.

Nous pourrions étendre le type de recherche que nous avons fait dans cette thèse à d'autres aspects de la statistique, comme l'étude des séries à une seule variable et des indicateurs de tendance centrale (moyenne, médiane) comme des indicateurs de dispersion.

En ce qui concerne les séries statistiques doubles et l'ajustement affine nous souhaiterions mettre à nouveau en œuvre un enseignement expérimental. Le contexte de la Centrafrique fait qu'on n'essaie pas de faire quelque chose de révolutionnaire avec un enseignement comme un « parcours d'étude et de recherche » (Chevallard, 2009) où on pose une question initiale et on laisse les élèves réaliser l'investigation où il y aura beaucoup d'autonomie. Mais il serait possible de collecter des données dans la classe sur des thèmes qui intéressent les élèves, et qui leur permettent d'être plus sensibles au sens de l'estimation. Par ailleurs, un tel type d'enseignement de la statistique, dans lequel les élèves collectent eux-mêmes des données pourrait être mis en œuvre dès le collège, pour les enseignements de la statistique descriptive et pourrait contribuer à donner du sens aux indicateurs statistiques, et éviter que les élèves se centrent sur le calcul uniquement.

Le sujet de la statistique nous semble très important pour la formation des élèves en particulier en Centrafrique. Il est vrai que le manque d'accès à des calculatrices et des logiciels constitue un vrai obstacle pour cette étude ; mais nous espérons que dans l'avenir cette situation pourra s'améliorer.

On pourrait en particulier souhaiter que les établissements publics soient équipés avec une salle informatique, et un lot de calculatrices qui pourraient être partagées entre les différentes classes. Naturellement si ces équipements évoluent dans le sens où nous le souhaitons, nos recherches s'adapteront à ces nouvelles conditions.

Perspectives pour la formation

En tant que formateur à l'École Normale Supérieure de Bangui, nous allons transférer les travaux de cette thèse vers la formation de plusieurs manières.

Premièrement, en ce qui concerne directement la statistique, notamment l'ajustement affine pour le cas des séries statistiques doubles. Ce thème est important pour la formation des enseignants du second degré. Il nous semble nécessaire comme premier objectif, la formation continue des enseignants sur le thème de la statistique en mettant en avant l'ajustement affine pour le cas des séries statistiques doubles. Et veiller à ce que l'enseignement de la statistique soit vraiment effectif dans les collèges et les lycées. Nous rappelons qu'à la fin de la formation initiale à l'ENS de Bangui, les futurs enseignants sortent avec une base en mathématiques. Cependant certaines promotions des élèves professeurs n'ont pas eu la chance de recevoir une formation initiale en statistique, d'où la nécessité de la formation continue en statistique. Pour cela, nous préconisons de mettre en avant dans une telle formation des contenus sur l'histoire de la statistique, qui permet de réfléchir sur le sens des méthodes. Ils doivent savoir l'origine des concepts statistiques et l'évolution des méthodes statistiques qui sont aux programmes. De plus une formation à la didactique de la statistique nous semble indispensable. En nous appuyant notamment sur les travaux de Régnier (2002, 2005, 2012), nous pensons que nous pouvons utilement faire découvrir aux enseignants en Centrafrique le concept d'esprit statistique. De même nous pouvons mettre en avant l'importance de la modélisation statistique et notamment celle de l'interprétation statistique. Il nous semble nécessaire que les enseignants prennent en compte les activités de modélisation dans l'enseignement de la statistique auprès des élèves.

Deuxièmement, au-delà du thème de la statistique, nous envisageons de transférer à la formation des enseignants différents outils issus de notre formation en didactique des mathématiques. Les élèves professeurs auraient intérêt à apprendre à faire des analyses praxéologiques, par exemple pendant le cours de didactique des mathématiques, nous proposerons des situations en statistique ou dans d'autres domaines mathématiques où nous demanderons aux élèves professeurs de faire des analyses praxéologiques. Nous envisageons aussi de les former à enseigner la modélisation, en leur proposant par exemple des situations-problèmes.

Par ailleurs, nous envisageons aussi d'initier les enseignants du second degré aux logiciels statistiques comme Excel pour leur culture mathématique, dans le contexte Centrafricain où l'accès aux outils informatiques est difficile.

BIBLIOGRAPHIE

- Armatte, M. (2005). Une statistique mathématique sans probabilité ? *Journ@l'Électronique d'Histoire des Probabilités et de la Statistique* 1(1), 1-19.
- Batanero, C., Fine, J., Raoult, J-P & Vermandele, C. (2013). Le curriculum statistique dans le secondaire - comparaisons internationales. *Statistique et enseignement*, 4(1), 1-4.
- Batanero, C., Godino, J.D., Vallecillos, A., Green, D.R., & Holmes, H. (1994). Erros and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.
- Berté, A., et al (2013). L'erreur dans l'apprentissage des mathématiques. *Petit x*, 93, 07-28.
- Block, D., Nikolantonakis, K., & Vivier, L. (2012). Registres et praxis numérique en fin de première année de primaire dans trois pays. *Annales de didactique et de sciences cognitives*, 17, 59-86.
- Blum, W., & Leiß, D. (2005). Filling up - the problem of independence-preserving teacher interventions in lessons with demanding modelling tasks. In M. Bosch (Ed.), *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education (CERME 4)* (pp. 1623–1633). URLL and ERME.
- Blum, W. & Ferri, R. B. (2009). Mathematical modelling : can it be taught and learnt ? *Journal of Mathematical Modeling and Application*, 1(1), 45-58
- Bouvier, A. (Éd.). (2005). *Dictionnaire des mathématiques* (7. éd., mise à jour). Presses Univ. de France.
- Brousseau, G. (1983). Les « effets » du « contrat didactique ». *Actes de la 2ième école d'été de didactique des mathématiques*, 1-16, IREM Orléans.
- Brousseau, G. (1986). La relation didactique : le milieu, *Actes de la IVème École d'Été de didactique des mathématiques*, 54 - 68, IREM Paris 7.
- Brousseau, G. (1998). Glossaire de quelques concepts de la théorie des situations didactiques en mathématiques. <http://guy-brousseau.com/wp-content/uploads/2010/09/GlossaireV5.pdf>.
- Brousseau, G. (2004). *Théorie des situations didactiques : Didactique des mathématiques, 1970-1990*. La Pensée sauvage.

- Brousseau, G. (2009). *Alternatives en Didactique de la Statistique*. <http://guy-brousseau.com/wp-content/uploads/2011/03/09-Alternatives-en-didactique-de-la-statistique.pdf>
- Burrill, G. & Biehler, R. (2013). Les idées statistiques fondamentales dans le curriculum scolaire. *Statistique et Enseignement*, 4(1), 5-24.
- Casey, S. A., & Wasserman, N. H. (2015). Teachers' knowledge about informal line of best fit. *Statistics Education Research Journal*, 14(1), 8-35.
- Castro-Sotos, A. E., Van den noortgate, W., & Onghena, P. (2009). The transitivity misconception of Pearson's correlation coefficient. *Statistics Education Research Journal*, 8(2), 33-55.
- Chaachoua, H., Ferraton, G., & Desmoulins, C. (2017). Utilisation d'un modèle praxéologique de référence dans un EIAH. In G. Cirade et al. (Eds.), *Évolutions contemporaines du rapport aux mathématiques et aux autres savoirs à l'école et dans la société* (pp. 301-324). <https://citad4.sciencesconf.org>.
- 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*. IREM de Paris – Université Paris Diderot.
- Chevallard, Y. (1998). Analyse des pratiques enseignantes et didactique des mathématiques : L'approche anthropologique. *Actes de l'Université d'été de la Rochelle*, 91-120. IREM de Clermont-Ferrand.
- Chevallard, Y. (2001). Organiser l'étude 1. Structures et fonctions. In J.-L. Dorier, M. Artaud, M. Artigue, R. Berthelot, & R. Floris, (Eds.) *Actes de la XIème École d'été de didactique des mathématiques, Corps*, 3- 32. Pensée Sauvage.
- Chevallard, Y. & Wozniak, F. (2003). Enseigner la statistique au secondaire. Entre genre prochain et différence spécifique. In Mercier, A. & Margolinas, C. (Eds.), *Balises en didactique des mathématiques: cours de la XIIe école d'été de didactique des mathématiques, corps* (Isère) (p.195-218). La pensée sauvage.
- Coppé, S., & Houdement. C. (2009). Résolution de problèmes à l'école primaire française : perspectives curriculaire et didactique. Colloque de la COPIRELEM, Juin 2009, France. <https://halshs.archives-ouvertes.fr/halshs-00959613>

- Coppé, S. (2012). Démarche d'investigation et aspects temporel des processus d'apprentissage/enseignement. In J-L. Dorier, & S. Coutat, (Eds.). *Enseignement des mathématiques et contrat social: enjeux et défis pour le 21^e siècle. Actes du colloque EMF2012*, pp.1306-1318. Genève Suisse.
- Coulangue, L. (1997). Les Problèmes concrets à mettre en équations dans l'enseignement. *Petit x*, 47, 33-58. IREM de Grenoble.
- Doukhan, C. (2021). *Modèles praxéologiques dans la transition secondaire-supérieure : le cas des probabilités en filière biologie*. [Thèse de doctorat]. Université de Bretagne Occidentale.
- Dorier, J.-L., Gueudet, G., & Robert, A. (2018). Les difficultés de la modélisation mathématique : comment faire dépasser des intuitions erronées ? In J-L. Dorier, G. Gueudet, M.-L. Peltier, A. Robert, & E. Roditi, (Eds.). *Enseigner les mathématiques Didactique et enjeux de l'apprentissage*, pp.291-303. Belin: Éducation.
- Droesbeke, J.-J., & Vermandele, C. (2018). Faciliter l'enseignement de la moyenne arithmétique en se servant de son histoire. *Statistique et Enseignement*, 9(1), 99-120.
- Duval, R. (1993). Registres de représentation sémiotique et fonctionnement cognitif de la pensée. *Annales de Didactique et de Sciences Cognitives* 5, 37- 65.
- Eichler, A., & Zapata-Cardona, L. (2016). *Empirical Research in Statistics Education*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-38968-4>
- Estepa, A. & Sanchez-Cobo, F.T. (2003) Evaluacion de la comprension de la correlacion y regresion a partir de la resolucion de problemas. *Statistics Education Research Journal* 2(1), 54-68.
- Fine, J. (2013). Le rapport Gaise (U.S.) cadre d'un curriculum statistique de la maternelle à la terminale. *Statistique et Enseignement*, 4(1), 25-54.
- Gattuso, L., & Pannone, M.A. (2000). Une expérimentation d'enseignement des statistiques et les enseignants qui l'ont vécue. <http://cirdis.stat.unipg.it/files/Sperimentazione/Giornate%20di%20Studio/Relazioni/Gattuso-Pannone.pdf>
- Gattuso, L. (2011). L'enseignement de la statistique : où, quand, comment, pourquoi pas ? *Statistique et Enseignement*, 2(1), 5-30.

- Gattuso, L. & Mary, C. (1997). La moyenne, un concept évident? Bulletin Association Mathématique du Québec, 37 (3),10-19.
- Haddad-Ouédraogo, G., Binaté, M., Niamen, A., & Ouéhi, D. (1999). Mathématiques, terminale sciences expérimentales. Edicéf.
- Hahn, C. (2015). La recherche internationale en éducation statistique : état des lieux et questions vives. *Statistique et Enseignement*, 6 (2), 25- 39.
- Henry, M. (2010). Evolution de l'enseignement secondaire français en statistique et probabilités. *Statistique et Enseignement*, 1(1), 35-45 <http://www.statistique-et-enseignement.fr>
- Hersant, M. (2013). Le contrat didactique et l'organisation de la rencontre des élèves avec le savoir. In M. Hersant, & C. Morin, (Eds.), *Pratiques enseignantes en mathématiques : expérience, savoir et normes* pp. 89-110. Presses Universitaires de Bordeaux.
- Hersant, M. (2014). Facette épistémologique et facette sociale du contrat didactique : une distinction pour mieux caractériser la relation contrat didactique milieu, l'action de l'enseignant et l'activité potentielle des élèves. *Recherches en didactique des mathématiques*, 34(1), 9-31. La pensée Sauvage.
- Jamie, D. M. (2002). Using Computer Simulation Methods to Teach Statistics: A Review of the Literature. *Journal of Statistics Education*, 10(1), DOI: 10.1080/10691898.2002.11910548.
- Koparan, T. & Güven, B. (2012). The effect on the 8th grade students' attitude towards statistics of project based learning. *European Journal of Educational Research*, 3(2), 73-85.
- Kuzniak, A. (2021). *Enseignement de la modélisation mathématique et construction du travail mathématique : une dynamique problématique*. Communication présentée à l'École d'été de didactique des mathématiques, Île de Ré, France.
- Marasinghe, M. G., Duckworth, B., & Tae-sung Shin. (2004). Tools for Teaching Regression Concepts using Dynamic Graphics. *The Journal of Statistical Education*, 12(2) <http://www.amstat.org/publications/jse/v12n2/marasinghe.html>
- Markulin, K., Bosch, M., & Ferrando, I. F. (2021). Project-based learning in statistics : a critical analysis. 11(1), 199-220.

- Mary, C. & Theis, L. (2007). Les élèves à risque dans des situations problèmes statistiques : stratégies de résolution et obstacles cognitifs. *Revue des sciences de l'éducation*, 33(3), 579 - 599.
- Monod, R., et al. (2010). L'investigation en MI-SVT : un chemin vers l'autonomie des élèves ? In C. Loisy, J. Trgalova, & R. Monod, (Ed.), *Ressources et travail collectif dans la mise en place des démarches d'investigation dans l'enseignement des sciences. Actes des journées scientifiques DIES 2010 Lyon, 24 et 25 novembre 2010*, pp.87-97. Institut National de Recherche Pédagogique.
- Munier, V., Chesnais, A., & Molvinger, K. (2018). La mesure en mathématiques et en physique: enjeux épistémologiques et didactiques. In M. Bächtold, V. Durand-Guerrier, & V. Munier (Eds.), *Epistémologie & Didactique* (pp. 95-111). Presses Universitaires de Franche-Comté.
- Parzysz, B. (2011). Quelques questions didactiques de la statistique et des probabilités. *Annales de didactique et de sciences cognitives* 16. 127-147.
- Piednoir, J.-L. (2004). Modélisation et enseignement des mathématiques au lycée. In J.-P Raoult, (Ed.), *La modélisation*, pp. 31-33. IREM de Paris VII.
- Régner, J. C. (2003). A propos de la formation en statistique. Approches praxéologiques et épistémologiques de questions du champ de la didactique de la statistique. <https://www.researchgate.net/publication/32221515>
- Régner, J-C. (2005). Formation de l'esprit statistique et raisonnement statistique. Que peut-on attendre de la didactique de la statistique? *Séminaire National de Didactique des Mathématiques*, 13-38. Paris France.
- Régner, J-C. (2012). Enseignement et apprentissage de la statistique : entre un art pédagogique et une didactique scientifique. *Statistique et Enseignement*, 3(1), 19-36.
- Robert, A. (1998) Outils d'analyse des contenus mathématiques à enseigner au lycée et à l'Université. *Recherches en Didactique de Mathématiques*, 18 (2), 139-190.
- Roche, F. & Barny, F. (2000). *Mathématiques comptabilité et gestion Informatique de gestion*. Hachette Éducation.

- Rodoumdje, F. (2019). Initiatives des élèves dans l'apprentissage de la statistique descriptive : étude de manuels. Communication par poster, École d'été de didactique des mathématiques, Autrans.
- Rolland, A. (2020). 2009-2019 : dix ans de publications sur l'enseignement de la statistique en France. *Statistique et Société*, 8(1), 55-71.
- Samueli, J-J. (2010). Legendre et la méthode des moindres carrés. *OpenEdition journals*. <http://journals.openedition.org/bibnum/580>.
- Sotura, B. (2017). Un stage de l'IREM Paris Diderot autour des enjeux de l'enseignement des probabilités et de la statistique au lycée. In L. Vivier (Ed.), *Les probabilités et la statistique au lycée* (pp. 139-202). Presses Universitaires de Franche-Comté.
- 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 21^e siècle - Actes du colloque EMF 2012* (pp. 742–753). Genève. <http://emf.unige.ch/actes-emf-2012/>
- Stigler, S. M. (1990). The development of mathematical statistics in astronomy and geodesy before 1927. In S. M. Stigler (Ed.), *The history of statistics : the measurement of uncertainty before 1900* (pp. 9 – 61).
- Visnovska, J. & Cobb, P. (2019). Supporting shifts in teachers' views of a classroom statistical activity : problem context in teaching statistics. *Mathematical thinking and learning*, 21(4), 285-305.
- Zendreras, N., Dubreil-Frémond, V., Marion, J-M., & Bihan-Poudec, A. (2017). Les données et leur production : Réflexions sur une lacune paradoxale en éducation statistique. *Statistique et Enseignement*, 8 (2), 59 – 78.

ANNEXES

A. TRANSCRIPTION D'UN ENSEIGNEMENT ORDINAIRE EN CENTRAFRIQUE

Transcription de la première séance de cours

Lignes	Actants	Minutage	Productions verbales
1		00:01	Silence
2	P	00:07	<p>[Le professeur écrit au tableau, application] Le tableau suivant donne la moyenne grand Y des maxima de la tension artérielle, de la tension artérielle en fonction de l'âge X d'une population donnée.</p> <p>[Le Professeur fait un tableau statistique]</p> <p>L'âge est représenté par la variable X . Nous avons la tension qui est représenté par la variable Y .</p> <p>A 36 correspond 12 ; à 42 correspond 13 ; à 48 correspond 13,5 ; à 54 correspond 14,3 ; à 60 correspond 15,4 ; à 66 correspond 15,8 ; et enfin 70 correspond 16.</p> <p>Et donc voila ce tableau là, qui définit la tension en fonction de l'âge.</p> <p>L'individu qui à 36 ans doit avoir le maximum de la tension qui est de 12. Un individu qui à 42 ans doit avoir un maximum de tension qui est de 13, un individu qui à 48 ans doit avoir un maximum de tension qui est de 13,5 et ainsi de suite.</p> <p>Donc première question, c'est de représenter graphiquement le nuage de points $M(x, y)$. On prendra pour unités graphique, on prendra pour unités graphique 0,5cm pour 1 an en abscisse et, euh et 3cm pour l'unité de tension artérielle en ordonnée. Et donc l'origine, et donc l'origine sera le point de coordonnées 30 et 10.</p> <p>2.a) Vous constatez que le nuage de points permet d'effectuer, permet d'effectuer un ajustement affine et donnez une équation de la droite des moindres carrés.</p> <p>b) Représenter cette droite.</p> <p>3) une personne de 75 ans, à une tension artérielle de 17,5. Quelle serait sa tension théorique en utilisant, en utilisant la droite des moindres carrés. Comparer avec la tension réelle.</p>
3	P	09:50	<p>Voilà le petit travail, que vous avez à faire. J'espère que déjà à 45 minutes, on peut le faire en 45 minutes.</p> <p>On va chercher rapidement les éléments de réponse. Silence !</p> <p>Et après 45 minutes quelqu'un passera au tableau pour la correction. Pas quelqu'un mais des gens passeront au tableau pour la correction.</p> <p>En résolvant ici, nous avons affaire à des nombres plus, donc notre repère doit être compter de cette manière.</p> <p>Qu'est-ce-qu'on dit ?</p> <p>On suit !</p>
4	P	15:01	<p>L'origine du repère n'est pas le point de coordonnées 0 ; 0. On dit ici que l'origine du repère sera le point de coordonnées $x=30$, $y=10$.</p>

			<p>$x=30$, $y=10$. L'origine de l'axe des abscisses est passée par 30, l'origine de l'axe des ordonnées est passée par 30, et donc l'origine de l'axe des ordonnées est passée par ?</p> <p>10 en allant. [Le Professeur montre l'orientation de l'axe des ordonnées par la main puis celle des abscisses.]</p> <p>30 ; 31 ; 32 ; ainsi de suite ça commence à 10 ; 11 ; 12 ainsi de suite. Voilà ce que ça veut dire ces contraintes, c'est une précision.</p>
5	P	16:16	De 36 à 70 ça nous fait une distance de 30 à ?
6	E		34
7	P		<p>34. De 36 à 70, il y a une distance de 34</p> <p>Suivez ! c'est-à-dire que si, 36 se trouve quelque part ici, [le Professeur montre un endroit avec sa main sur l'axe des abscisses] le 36 à 70 ça doit nous faire une distance de ? De 34.</p> <p>Si on prend une unité comme étant le carreau, l'unité est égale à un carreau. Si on arrive à placer 36 quelque part ici, 36 à 70, ça va nous faire 34 et on voit que 34 carreaux, c'est trop grand, faut réduire. Nous disons que une unité va correspondre à demi carreau.</p> <p>Si une unité correspond à demi carreau, alors un carreau va correspondre à deux unités. Un carreau sur l'axe des abscisses va correspondre à deux unités, parce qu'on prendra pour unité graphique 0,5cm pour un an ou pour une unité sur l'axe des abscisses. En ordonnée, on prendra 3cm puisque nous sommes entrain de convertir, les centimètres en carreau puisque, nous travaillons euh ___ votre cahier.</p> <p>C'est aussi utile de convertir ___ unité . Considérez une demi carreau qui va correspondre à un an et 3 carreaux qui va correspondre à une unité de tension artérielle, une unité de tension artérielle.</p> <p>Un carreau, deux carreaux, trois carreaux, ici trois carreaux.</p> <p>Nous sommes ici là, à ?</p>
8	E		A 11
9	P		<p>Trois carreaux après.</p> <p>Un carreau, ici deux carreaux, ici trois carreaux, nous sommes ici à 12.</p> <p>Voici un carreau, deux carreaux, trois carreaux, nous sommes ici à 13,</p> <p>Voici un carreau, deux carreaux, trois carreaux, nous sommes ici à 14,</p> <p>Voici un carreau, deux carreaux, trois carreaux, nous sommes ici à 15,</p> <p>3 carreaux correspondent à une unité sur l'axe des ordonnées. Et l'unité sur l'axe sur l'axe des ordonnées, une unité correspond à ?</p>
10	E		A 1
11	P		A 1, c'est des données ici en tension. Sur l'axe des abscisses, un demi carreau correspond à deux unités. Et lorsqu'un demi carreau correspond à deux unités, alors un carreau va correspondre à ?

12	P	20:20	Un demi carreau correspond à une unité par conséquent, un carreau va correspondre à ?
13	E		Deux unités
14	P		Deux unités, là ça fait un carreau, voilà un carreau ici, voici deux carreaux, voici le carreau suivant, c'est gradué en carreau et donc chaque carreau correspond à deux unités. 30 ici, nous aurons 32, nous aurons 34, nous aurons 36. Voilà le point de départ par rapport à notre tableau : 36 après on a 38, on a 40, ainsi de suite. Entre 34 et 36, nous avons au milieu 30... ? La distance entre 32 et 34 correspond à un carreau. La distance entre 38 et 40 correspond à ? À un carreau. La distance 10 et 11 correspond à 3 carreaux, ainsi de suite.
15	P	21:53	Si X est égale à 36, le nombre moyen Y est égale à ?
16	E		A 12
17	P		X est égale à 36, voilà Y est égale à 12, voici le premier point, c'est le point M_1 de coordonnées $x=36$ et $y=12$, ainsi de suite. On va placer les points $M_2; M_3; M_4; M_5; M_6; M_7$ Y a t-il des questions ? Ces contraintes sont ainsi expliquées. On ne vous laisse pas la latitude de choisir les unités, on vous les impose, les unités sur l'axe des abscisses et sur l'axe des ordonnées. C'est ce qu'on a _ _ _
18	P	23:37	Monsieur, ceux qui travaillent sur les feuilles non graduées, du moins non graduelle peuvent utiliser ce que nous avons ainsi comme contraintes, c'est-à-dire 1cm qui correspond à 1 unité sur l'axe des abscisses et 3cm va correspondre à 1 unité sur l'axe des ordonnées. Et si on travaille avec le quadrillage d'un cahier, sur le cahier c'est bien de travailler avec les carreaux. La question un, c'est de représenter le nuage des points. J'ai besoin de quelqu'un au tableau pour la deuxième question. Passer au tableau _ _ _
19	E		_ _ _
20	P	26:13	Les ? Mais si on ne vous donne pas les unités, c'est à vous de, vous débrouillez pour faire le choix des unités. Si aucune unité a été donnée, a été donnée vous avez _ _ _ pour choisir vos unités.
21	P	26:48	Un ajustement affine est justifié, c'est un ajustement possible. Ajustement affine justifié, c'est-à-dire qu'on peut déterminer la droite des moindres carrés.[A ce stade, l'unique méthode introduit par le professeur est la méthode des moindres carrés] L'affirmation de l'ajustement affine permet de déterminer la droite de régression _ _ _
22	P	27:34	Et donc la deuxième question, on demande petit a, de constater que le nuage de points permet d'effectuer un ajustement affine. De constater à partir de la figure, mais la figure nous permet de

			constater effectivement que, un ajustement affine est justifié.
23	P	28:36	La question b, c'est de représenter cette droite. Pour la représenter, il faut d'abord déterminer son équation. Pour déterminer l'équation de la droite de régression, vous complétez le tableau par les variables x_i^2 , y_i^2 , $x_i y_i$ et aussi les totaux. De calculer des moyennes \bar{x} , \bar{y} , de calculer la covariance de x , y , de calculer la variance de x .
24	P	31:42	Complétez le tableau. Nous avons ici les totaux. Nous avons les x_i^2 , les y_i^2 , les $x_i y_i$ Complétez le tableau. Allez y ! A quoi est égale au total de x ? [Une élève est au tableau, elle utilise sa calculatrice pour faire le calcul.]
25	P	34:06	Je t'ai demandé de commencer par le total de x . [Elle écrit dans la case des totaux, la somme des x , $x=376$, puis elle écrit la somme des y : 100 dans la case correspondante à la somme des y]. Total de $x=376$, total de $y=100$ Les autres grandeurs. Allons y rapidement, dépêche madame . Ok, vous avez une calculatrice en main, on vous demande pas de réfléchir. Qu'est-ce que tu es entrain de calculer ? J'entends rien, 36 ? Mais tu écris. [Elle écrit dans la ligne correspondante aux valeurs des x_i^2 : 1296 correspond au carré de 36 ; 1764 correspond au carré de 42 ; 2304 correspond au carré de 48 ; 2916 correspond au carré de 54 ; 3600 correspond au carré de 60 ; 4356 correspond au carré de 66 ; 4900 correspond au carré de 70. Pour la somme des x_i^2 , elle écrit 21136] Donc la somme des $x_i^2=21136$ La somme des y_i^2 . Allons y
26	P	43:14	A ta place, laisse la machine là bas. [L'enseignant désigne un autre élève pour passer au tableau] Vous voyez l'intérêt de la tenue ? [Elle écrit dans la ligne correspondante aux valeurs des y_i^2 : 144 correspond au carré de 12 ; 169 correspond au carré de 13 ; 182,25 correspond au carré de 13,5 ; 204,49 correspond au carré de 14,3 ; 237,16 correspond au carré de 15,4 ; 249,64 correspond au carré de 15,8 ; 256 correspond au carré de 16. Pour la somme des y_i^2 , elle écrit 1,44254] Bien !

27	P	49:02	Ayez ce qu'on appelle le bon sens. Déjà tu me dis que ceux-ci égale à un virgule, un virgule alors que ici tu as 256. [L'enseignant montre la dernière colonne]. Ayez au moins ce qu'on appelle le bon sens.
28	P	49:54	La somme des y_i^2 est égale à ? [L'enseignant écrit au tableau.]
29	E		5485
30	P		5485, ?
31	E		Il n'y a pas de virgule
32	P		$36 \times 12 = ?$
33	Élèves		432
34	P		432. [L'enseignant complète le reste du tableau, en remplissant les cases.]
35	Élèves		546 ; 648 ; 772,
36	P		772, ?
37	Élèves		772,2 ; 924 ; 1042,8
38	P		1042,8
39	Élèves		1120
40	P	52:01	Avec ce tableau, on aura à calculer \bar{x} . On sait que \bar{x} est égale à la somme des x_i sur n , n est égale à ? c'est-à-dire 376 divisé par 7 est égale à ? 53, ?
41	Élèves		53,71
42	P		\bar{y} , \bar{y} serait égale à ?
43	Élèves		100
44	P		100 divisé par 7, ce qu'est égale à ?
45	Élèves		14,28
46	P		14, ?
47	Élèves		28
48	P		Et donc le point moyen G, c'est le point de coordonnées \bar{x} , \bar{y} , c'est-à-dire 53,71 et 14,28. On aura également à placer dans le repère le point G. On verra également que cette droite de régression, la droite des moindres carrés. Cette droite (D) passe par le point G, c'est-à-dire que le point G appartient à la droite (D). Le point G appartiendra à la droite (D'). La droite de régression de x en y. [pour la tâche, qui est de représenter la droite, le professeur n'a pas déterminer ni tracer la droite, il a expliqué oralement aux élèves] En suite _ _ _ c'est un repère.
49	P	53:25	Voilà ce que je dis, pour la troisième question. J'efface tout ce que j'ai ici. Et je reviens à l'énoncé. $x=36$, $y=12$; $x=66$, $y=15,8$; $x=70$, $y=16$;

			Si $x=75$, Monsieur à qui est égale à y ?
50	E		17,5
51	P		<p>Si $x=80$ à quoi est égale à y ? Voilà ce dont nous sommes à la recherche. Si $x=80$ à qui est égale à y ? Et donc, on ne voit pas facilement que la droite de régression, c'est le lien qu'il y a entre x et y. Si je connais x, alors je peux calculer la valeur de y. Si je connais y, je peux également calculer la valeur de x. Cette droite de régression, c'est le lien qu'il y a entre x et y [Discours technologique sur la covariance]. Et donc voilà l'intérêt de faire des prévisions. Si x est égale à ceci, à quoi est égale à y ? Si on connaît x, on peut prévoir la valeur de y</p>
52	P	55:17	<p>Une personne de 75 ans a une tension artérielle 17,5, une personne de 75 ans a une tension artérielle réelle de 17,5, c'est-à-dire une personne de 75 ans doit avoir réellement une tension de 17,5. Mais théoriquement à partir de cette droite de régression, à partir de ce lien qu'il y a entre x et y. Quelle serait alors cette tension ? Est-ce que c'est la même chose ? Si je donne la valeur 75 à x, quelle serait alors la valeur de y ? Je calcule, est ce que ça va me donner effectivement 17,5 ? Quelle serait sa tension théorique, en utilisant la droite des moindres carrés ? Et après on vous demande de comparer avec la tension réelle. La tension réelle c'est celle qui est de 17,5. Ce qu'on va calculer et on va comparer avec la tension réelle. [On ne voit pas le professeur déterminer la droite d'équation des moindres carrés, qui devrait servir à déterminer la tension artérielle prévue pour un homme de 75 ans] Voilà les indications par rapport à la suite de l'exercice à la question trois. Vous allez les interpréter. Est-ce-que, y a t-il des questions ? Est-ce-que, y a t-il des questions ?</p>
53	E	57:17	___ ?
54	P		<p>___ y en fonction de x, personne n'a encore fait ? On peut utiliser la droite (D) ou la droite (D') pour calculer la valeur de y. Mais avec une droite, la droite (D) c'est le lien qu'il y a entre x et y. La droite (D') c'est également le lien qu'il y a entre x et y.</p>
55	P	58:26	<p>Par exemple l'équation cartésienne d'une droite. Suivez, c'est l'équation cartésienne d'une droite, mais lorsqu'on écrit y en fonction de x, c'est ce qu'on appelle une équation réduite de y en fonction de x. Alors ça, ça équivaut à $y = -2x + 1$ Si on donne une valeur à x, on calcul ____, ça équivaut également à $2x = -y - 1$ donc x sera égale à $-\frac{1}{2}y + \frac{1}{2}$ Donc c'est équivaut à $x = -\frac{1}{2}y + \frac{1}{2}$ c'est-à-dire que même si je connais une équation de t, à partir de l'équation de t. Je peux déduire une équation de t'.</p>

			<p>Il suffit tout simplement d'exprimer x en fonction de y. Si la droite (D) a pour équation $y = -2x + 1$ alors voici l'équation de (D') $x = \frac{-1}{2}y + \frac{1}{2}$.</p> <p>Alors la prochaine fois c'est ____, la droite de Mayer suivi de quelques exercices, vous aurez également des exemples. (01 :00:36)</p>
Fin de la première séance			

Transcription de la deuxième séance de cours

Lignes	Actants	Minutage	Productions verbales
01	P		<p>Début de la deuxième séance On a donné un exemple, c'est ce que nous avons donné comme des questions. Petit b avec l'ajustement affine, la définition a été donné, on a ____ ce qu'on attendait par covariance de x et y. On a donné la formule qui permettait de déterminer les droites de régression par la méthode des moindres carrés suivi d'un exemple. Voilà pour ce qui concerne la partie 1.</p>
02	P	00:49	<p>Petit 2, voilà on va parler du coefficient, le coefficient de corrélation linéaire, coefficient de corrélation linéaire. Petit a, c'est la définition. Soit grand X et grand Y deux variables, deux variables d'une série double, d'une série double, évidemment quantitatives, d'écart types, d'écart types respectifs σ_X et σ_Y. Et de covariance, de covariance de X et Y. Et voilà ce que nous avons comme hypothèse, on a covariance de X, Y ; on a l'écart type de X et l'écart type de Y. L'écart type c'est la racine carrée de la ?</p>
03	Élèves		variance.
04	P		<p>Variance On appelle coefficient de corrélation linéaire entre X et Y, entre X et Y, le nombre qu'on appelle r qui est égale à la covariance de X, Y sur l'écart type de X fois l'écart type de Y ($r = \frac{Cov(X, Y)}{\sigma_X \times \sigma_Y}$). Voilà, la formule qui nous permet de calculer le coefficient de corrélation linéaire entre les valeurs de X et Y. Propriété, r est toujours compris entre -1 et 1. Soit (D), c'est la droite qui a pour équation $y = ax + b$, c'est la droite de régression de x en y. Et (D') qui a pour équation $x = a'x + b'$, la droite de régression de x en y. On a : $r^2 = a \times a'$</p>
05	P	06:17	<p>Remarques, on considère généralement la corrélation entre X et Y est forte si $r \geq 0,87$ ou $r \leq -0,87$. Si r est proche de 0, est proche de 0, la corrélation est faible, la</p>

			corrélation est faible. Donc l'ajustement affine n'a pas lieu d'être. Si elle est forte, alors l'ajustement affine est justifié.
06	E	08:54	S'il vous plaît Monsieur, concernant les remarques que vous avez faites.
07	P		Oui ?
08	E		Le premier intervalle, vous, on considère généralement le carré de la corrélation entre X et Y si ?
09	P		Si $r \geq 0,87$ ou $r \leq -0,87$
10	E		Ce n'est pas le contraire, le contraire Monsieur ?
11	P	09:33	<p>Voilà tu sais que r est compris entre -1 et 1, c'est ce qu'on a dit, n'est ce pas ?</p> <p>r est compris entre -1 et 1, ici on a : ça fait 0,25 c'est le milieu de ce côté, ça fait 0,75 par ce que c'est le quart, et donc 0,87 est par ici. Puisque c'est le quart, on prend le un huitième.</p> <p>A quoi est égale à un huitième, un divisé par huit? 0,25 multiplié par ?</p> <p>$\frac{1}{8}$ c'est 0,25 divisé par 2, c'est $\frac{1}{4}$ divisé par 2 et donc $\frac{1}{4}$ divisé par 2, ça donne combien ? Euh 0,25 divisé par 2, multiplié par 7 .</p> <p>Vous suivez !</p> <p>Voici 0 est là, c'est le milieu de l'intervalle -1 et 1. On a 0 ici,, la moitié de l'intervalle 0 à 1, ça fait 0,5. On partage l'intervalle de 0 à 1 en 4 intervalles, on aura alors de 0,25, on partage l'intervalle de 0 à 1, en 8 intervalles. La moitié ici, la moitié là, la moitié là.</p> <p>A quoi est égale ? qu'est l'abscisse de ce nombre ?</p> <p>Ça fait $\frac{1}{8}$, $\frac{2}{8}$, $\frac{3}{8}$, $\frac{4}{8}$, $\frac{4}{8}$ c'est $\frac{1}{2}$, $\frac{5}{8}$, $\frac{6}{8}$ et $\frac{7}{8}$ donc $\frac{7}{8}$ c'est 0,87.</p> <p>ça se trouve quelque part ici. Voilà le nombre 0,87, 1 c'est là. Ici c'est -0,87 ça, c'est la valeur de c.</p> <p>Si $r \geq 0,87$, r se trouve entre 0,87 et 1. Si r se trouve entre 0,87 et 1 alors on peut dire que la corrélation de X et Y, elle est forte. Ou bien $r \leq -0,87$, elle se trouve entre -1 et -0,87. On dit également que la corrélation entre X et Y, elle est forte.</p> <p>Et si r, d'après ce qu'on a dit : si r est proche de 0 c'est-à-dire si r se trouve entre -0,87 et 0,87, on dit alors que la corrélation entre X et Y est faible.</p> <p>Voilà j'hachure ici cet intervalle et on voit ici que c'est l'intervalle sur lequel la corrélation entre X et Y est faible. On écrit écrit ici, corrélation faible. Et entre 0,87 et 1, on a également une corrélation forte. Entre -1 et -0,87, on a également une corrélation forte.</p> <p>Si r se trouve ici la corrélation est forte. Si r se trouve ici la corrélation est forte.</p> <p>Mais qu'est ce qu'on appelle par corrélation ?</p> <p>C'est le lien qu'il y a entre X et Y. Le coefficient de corrélation linéaire, c'est cet ordre là qui nous permet d'affirmer qu'un ajustement affine est un ajustement linéaire entre X et Y est justifié. Ça nous permet de l'affirmer par sa valeur, la valeur de r.</p>

			On a donné ici propriété, des propriétés, c'est-à-dire r est compris entre -1 et 1, il n'y aura jamais de r supérieure à 1, il n'y aura jamais de r inférieure à -1.
12	P	15:15	<p>Regardez, si on connaît une équation de la droite de régression de y en x. On suit !</p> <p>Si (D) a pour équation $y=ax+b$ c'est la droite de régression de y en fonction de x, $y=ax+b$. On a (D') qui se trouve être la droite de régression de x en y, de x en fonction de y, $x=a'y+b'$.</p> <p>Alors si je connais les équations des deux droites, on peut facilement calculer r. Le coefficient de corrélation linéaire entre X et Y, et aussi $r=a \times a'$, $r=a \times a'$</p> <p>Si je ne connais pas, si je ne connais pas des équations des deux droites de régression, alors je vais calculer r en utilisant la covariance de X et Y sur l'écart type de X multiplié par l'écart type de Y ($r = \frac{cov(x,y)}{\sigma_x \times \sigma_y}$)</p> <p>A quoi est égale à r ?</p> <p>Pour l'exemple qui était donné. A quoi était égale à la covariance de X, Y ?</p> <p>C'est 13,6.</p> <p>A quoi est égale à la variance de X ?</p> <p>La variance de X est égale à ? 13,8</p> <p>A quoi est égale à la variance de Y ?</p> <p>$(r = \frac{13,6}{3,79 \times 3,82})$</p> <p>A quoi est égale à l'écart type de x ?</p> <p>L'écart type de x est égale à la racine carrée de la variance de x.</p> <p>A quoi est égale à la racine carrée de 13,8 ?</p>
13	E		3,71.
14	P		<p>A quoi est égale à l'écart type de y ?</p> <p>C'est la racine carrée de sa variance, c'est 3,.. ?</p>
15	E		C'est 3,82
16	P		<p>Je connais la covariance de x, y. Je connais la variance de y, alors je peux calculer r. A quoi est égale à r ?</p> <p>r est égale, c'est-à-dire 13,6 sur 3,79 multiplié par 3,82</p> <p>$(r = \frac{13,6}{3,79 \times 3,82})$, ce qui est égale à combien ?</p> <p>Ceci est égale à 0,.. ?</p>
17	E		0,95
18	P		<p>0,95, on connaît le coefficient de corrélation linéaire. On vient de le calculer, alors on peut l'interpréter. r=0,95 ça signifie quoi ?</p> <p>r est supérieure à 0,87, par conséquent la corrélation linéaire entre x et y, elle est ?</p>
19	E		forte
20	P	19:04	Dire que la corrélation linéaire entre x et y est forte, signifie qu'un ajustement affine est possible, un ajustement affine est justifié, c'est-à-dire on peut trouver une droite qui passera le plus prêt possible des

			points du nuage. Donc voilà, ce que ça veut dire. Vous allez représenter le schéma.														
21	P	19:50	Exemple, qu'est ce qu'on a ? covariance de x, y c'est égale à combien ?														
22	Élèves		13,6														
23	P		<p>La variance de x, c'est 13,8 ça entraîne que l'écart type de x est égale à la racine carrée de 13,8 qui égale à 3,71 :</p> $(\sigma_x^2 = 13,8 \Rightarrow \sigma_x = \sqrt{13,8} = 3,71)$ <p>La variance de y, c'est 14,64 alors l'écart type de y sera égale à la racine carrée de 14,64, ce qui a pour valeur 3,82 :</p> $(\sigma_y^2 = 14,64 \Rightarrow \sigma_y = \sqrt{14,64} \approx 3,82)$ <p>Et donc r est égale covariance de x, y sur l'écart type de x multiplié par l'écart type de y : $r = \frac{cov(x, y)}{\sigma_x \times \sigma_y} = \frac{13,6}{3,71 \times 3,82} = 0,95$</p> <p>Qu'est ce qu'on a ? $r \geq 0,87$ 0,95 est supérieure à 0,87 par conséquent ? La corrélation entre x et y est forte. Dire que la corrélation entre x et y est forte signifie que l'ajustement affine est possible. D'où un ajustement entre x et y est justifié.</p>														
24	P	23:58	<p>Nous allons donner un deuxième exemple. Pour cela nous allons continuer avec la deuxième méthode qui est la droite de Mayer, droite de Mayer. On parle également de l'ajustement affine par la méthode de Mayer.</p> <p>Exemple : On donne la série suivante. On a les x_i, on a les y_i On a 1, on a 5, on a 7, on a 15, n a 22 et on a 25. Et les y_i on a 10, on a 20, on a 42, on a 60, on a 80 et on a 90. [L'enseignant fait le tableau statistique</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>x_i</td> <td>1</td> <td>5</td> <td>7</td> <td>15</td> <td>22</td> <td>25</td> </tr> <tr> <td>y_i</td> <td>10</td> <td>20</td> <td>42</td> <td>60</td> <td>80</td> <td>90</td> </tr> </table> <p>]</p> <p>On se propose de construire la droite de Mayer. Pour cela, partageons la série en deux séries de même effectif. Calculons les coordonnées des points moyens G_1 et G_2 des deux séries ainsi définies.</p> <p>Alors la droite de Mayer est la droite $(G_1 G_2)$. On vous demande de partager la série en deux séries. Si on avait une série d'effectif sept, on aura la possibilité de partager la série en deux séries,. Mais cette fois ci, ce n'est pas de même effectif : c'est 3, 4 ou bien 4 et 3. Partageons la série en deux séries, on a les séries : Première série, x_i, y_i ; deuxième série : x_i, y_i Première série, allons y quoi ? On a ?</p>	x_i	1	5	7	15	22	25	y_i	10	20	42	60	80	90
x_i	1	5	7	15	22	25											
y_i	10	20	42	60	80	90											
25	E	26:46	On a : 1 ; 5 ; 7 ; 10 ; 20 ; 42. [L'enseignant écrit sous forme de tableau.]														

26	P		<table border="1"> <tbody> <tr> <td>x_i</td> <td>1</td> <td>5</td> <td>7</td> </tr> <tr> <td>y_i</td> <td>10</td> <td>20</td> <td>42</td> </tr> </tbody> </table> <p>Deuxième série</p>	x_i	1	5	7	y_i	10	20	42
x_i	1	5	7								
y_i	10	20	42								
27	E		15 ; 22 ; 25 ; 60 ; 80 ; 90. [L'enseignant écrit sous forme de tableau.]								
28	P		<table border="1"> <tbody> <tr> <td>x_i</td> <td>15</td> <td>22</td> <td>25</td> </tr> <tr> <td>y_i</td> <td>60</td> <td>80</td> <td>90</td> </tr> </tbody> </table> <p>On a partagé la série en deux séries de même effectif. Là c'est une série à part et l'autre aussi c'est une série à part. Si on considère cette série [première série], on peut calculer les coordonnées de son point moyen, appelées G_1 n'est pas ? A quoi est égale à \bar{y}_1 ? A quoi est égale à \bar{x}_2 ? A quoi est égale à \bar{y}_2 ? \bar{x}_1 est égale à quoi ? C'est égale à un tier multiplié par la somme des x_i, multiplié par 13, ce qui est égale à ? on va donner sa valeur exacte $\bar{x}_1 = \frac{1}{13}(13) = \frac{13}{3}$ A quoi est égale à \bar{y}_1 ? C'est un sur trois multiplié par la somme des y_i, c'est 72, simplification, ça vous donne une valeur exacte qui est égale à ?</p>	x_i	15	22	25	y_i	60	80	90
x_i	15	22	25								
y_i	60	80	90								
29	E		24								
30	P		$\bar{y}_1 = \frac{1}{13}(72) = 24$. Et donc le point moyen G_1 , c'est le point de coordonnées $\frac{13}{3}$ et 24, $G_1(\frac{13}{3}; 24)$ A quoi est égale ? Pour la deuxième série. \bar{x}_2 est égale à $\frac{1}{13}$ multiplié par.. ? est égale à $\frac{1}{13}$ multiplié par.. ? La somme des x_i , c'est ?								
31	Élèves		62								
32	P		62, la somme des y_i								
33	Élèves		230								
34	P		<p>Ce qu'est égale à 62 divisé par 3, ici ça nous donne ? C'est $\frac{230}{3}$ $(\bar{x}_2 = \frac{1}{3}(62) = \frac{62}{3}; \bar{y}_2 = \frac{1}{3}(230) = \frac{230}{3})$ Et donc le point moyen G_2, est le point de coordonnées : $\frac{62}{3}$ et $\frac{230}{3}$, $G_2(\frac{62}{3}; \frac{230}{3})$</p>								

			On a partagé la série en deux séries de même effectif. On a calculé les coordonnées des points moyens G_1 et G_2 . On a commencé par les figures, on aura alors tout simplement à construire la droite G_1 et la droite G_2 , la droite (G_1G_2) . C'est la droite qui passe par les points G_1 et G_2 .
35	P	34:27	Je redis donc, graphiquement, graphiquement, la droite de Mayer, la droite de Mayer c'est la droite (G_1G_2) qui passe par les points G_1 et G_2 . Et il suffit tout simplement de les joindre. Raison pour laquelle, je dis graphiquement. Vous êtes assez grand, pour comprendre ce que je suis entrain de dire. Voilà ce qu'on voit graphiquement, on a construit la droite (G_1G_2) , mais qu'elle sera alors une équation de la droite (G_1G_2) ? Qu'elle est son équation? Quelle sera alors équation? Déterminons une équation de la droite (G_1G_2) Débout!
36	P	35:59	La droite (G_1G_2) passe par quels points? Par les points?
37	E		La droite (G_1G_2) passe par le point G_1
38	P		Par le point G_1 seul? La droite (G_1G_2) passe par le point G_2 ? La droite (AB) passe par le point A. La droite (AB) passe par le point B. Dire que la droite (AB) passe par le point A cela signifie que le point A appartient à la droite (AB). Dire que la droite (AB) passe par le point B, signifie que le point appartient à la droite (AB). Dire que la droite (G_1G_2) passe par le point G_1 , cela signifie que le point G_1 appartient à la droite (G_1G_2) et aussi le point G_2 appartient à la droite (G_1G_2) . Soit la droite (G_1G_2) , c'est la droite qui a pour équation $y=ax+b$. Je vais changer a et b. $y=mx+p$, on va déterminer les valeurs, la valeur de m et celle de p . C'est noté déjà là bas? La droite ce n'est pas finie? Est-ce qu'on peut effacer à droite?
39	Élèves		Non Monsieur
40	P		Alors on suit, le point $G_1 \in (G_1G_2)$ n'est ce pas? Si la droite passe par ce point, ça signifie quoi? Lorsqu'on dit que la droite passe par le point G_1 ? Lorsqu'on dit que la droite passe par le point G_1 ? La droite passe par le point G_1 , n'est ce pas? La droite passe par le point G_1 , lorsque $G_1 \in (G_1G_2)$. Ça veut dire quoi? Lorsqu'on dit que le point $G_1 \in (G_1G_2)$? Le point $G_1 \in (G_1G_2)$ lorsque la droite passe par, évidemment, il faut faire les va et vient entre...
41	P	41:27	Quel est le prénom de ton père? Monsieur?

42	E		Isaïe
43	P		Et toi c'est ?
44	E		Ephrem
45	P		On vérifie que tu es le fils d'Isaïe si et seulement si Isaïe est ton père. C'est pour aboutir à quoi finalement ? Est ce que y a une logique dans ce qu'on dit ? Soyez capable de convaincre On sait qu'un point appartient à une droite si et seulement si, ce point vérifie l'équation de la droite. Un point appartient à une courbe si et seulement si, les coordonnées de ce point vérifie l'équation de la courbe. Si et seulement si, x et y, x c'est quoi et y ? x c'est l'abscisse et y c'est ?
46	E		l'ordonnée
47	P		Et donc l'ordonnée de point G_1 est égale à m multiplié par l'abscisse du point G_1 plus p , c'est : $y_{G_1} = mx_{G_1} + p$ ça équivaut à quoi ? Je connais x_{G_1} , je connais y_{G_1} , donc je remplace, c'est combien ? C'est $24 = \frac{13}{3}m + p$ Le point $G_2 \in (G_1 G_2)$ si et seulement si $y_{G_2} = mx_{G_2} + p$, si et seulement si, on connaît y_{G_2} qui est égale à : $\frac{230}{3} = \frac{62}{3}m + p$ Finalement, on aura un système d'équation à deux inconnues. On a le système définie par ? On peut déjà réduire au même dénominateur, le dénominateur commun c'est ?
48	Élèves		c'est 3
49	P		Donc c'est $13m + 3p = 72$ $62m + 3p = 230$ En faisant deuxième équation moins la première équation, les p devient nul et donc : $62m - 13m$ ce qui est égale à ?
50	Élèves		$49m$
51	P		qui est égale à ?
52	Élèves		158
53	P		Et donc $m = \frac{158}{49}$ qui est égale à ? Est ce que ça c'est simplifiable
54	Élèves		Non
55	P		Si ça se simplifie, on devrait simplifier par combien ? Donc on connaît m , on remplace m par sa valeur pour trouver la valeur de p .

			<p>ça équivaut à $3p = 72 - \frac{2054}{49}$</p> <p>ça équivaut à $p = \frac{72}{3} - \frac{2054}{147}$, si et seulement si p ?</p> <p>$p = \frac{1474}{147}$ voici la valeur de p. Donc finalement cette droite (G_1G_2)</p> <p>D'où la droite (G_1G_2) a pour équation : $y = \frac{158}{49}x + \frac{1474}{147}$</p>
56	P	51:35	<p>Comment peut-on être rassurer ?</p> <p>J'ai encore le tableau ici : si $x=1$, $y=10$</p> <p>Si $x=1$, ça fait $\frac{158}{49} + \frac{1474}{147}$ réduisons au même dénominateur, qu'est ce que ça nous donne ? Le dénominateur commun c'est 147, et donc 49 doit être multiplier 3 ; 3×158, $\frac{474+1474}{147}$ est ce que le tout donne 10 ?</p>
57	Élèves		c'est 108
58	P		<p>Vous êtes sûr ?</p> <p>Si $x=15$ est ce que $y=60$?</p> <p>Voilà une idée comme ça qui me vient à l'esprit. Est ce que vous suivez ?</p> <p>Quelque part dans le cours, qu'est ce que nous avons dit ?</p> <p>Nous avons dit dans le cours que, voilà ce qu'on sait, que la droite (G_1G_2) passe par le point G_1 et ça passe par le point G_2.</p>
59	P	54:32	<p>Mais rien ne le garantie que la droite (G_1G_2) passe par le point de coordonnées $x=1$ et $y=10$, non. Parce que, qu'est ce que j'attends par nuage de point ? Nuage de point, ajustement affine ?</p> <p>Cela revient à déterminer une droite qui passe le plus près possible des points du nuage. La droite ne passe pas forcément par chacun de ces points là. La droite passe proche de ces points et donc on ne peut pas. Si on vérifie, en prenant $x=1$ on va trouve $y=10$, si on vérifie en prenant $x=20$, on va trouver y forcément égale à 80, non. Pour être rassurer, certes, la droite (G_1G_2) passe par le point G_1, on vérifie avec les coordonnées du point G_1, on aura aussi à vérifier avec les coordonnées du point G_2.</p> <p>Alors, pour cela, est ce que le point $G_1(\frac{13}{3}, 24)$?</p> <p>On va vérifier, $x = \frac{158}{49} \times \frac{13}{3} + \frac{1474}{147}$ on a déjà 147, donc 13×158 ça donne combien ? Plus 1474 le tout divisé par 147</p>
60	Élèves		24
61	P		<p>Le point G_2 appartient à cette droite, cela équivaut à dire que les coordonnées du point G_2 vérifie l'équation de cette droite.</p> <p>$y = \frac{230}{3}$, on a : $\frac{158}{49} \times \frac{62}{3} + \frac{1474}{147}$ on a déjà le dénominateur</p>

			commun qui est 147, $158 \times 62 + 1474 = 11270$ sur 147 Et si on doit simplifier, on va simplifier par ?
62	Élèves		Par 7
63	P		Par ce que 49 est un multiple de ?
64	E		7
65	P		Est ce qu'on peut simplifier 147 par 3 ? Par ce que 147 c'est $1+4+7=12$, et 12 est divisé par 3. Et 11270 c'est $1+1+2+7+0=11$, est ce que 11 est un multiple de 3 ?
66	E		non
67	P	58:19	Donc, on ne peut pas simplifier par 3. Soyez intelligent, qu'est ce qu'on veut avoir ? On veut avoir un nombre sur 3, on sera obligé de simplifier par 49. 11270 divisé par 49, qu'est ce que ça donne ? $\frac{11270}{147} = \frac{230}{3}$, non ? Ok --- On se propose de déterminer une équation de la droite (G_1G_2) , on connaît les coordonnées $M(x,y) \in (G_1G_2)$ si et seulement si quoi ? Pour $M(x,y)$, voici la droite (G_1G_2) et un point $M(x,y)$ de la droite si et seulement si quoi ?
68	E		les trois points sont alignés
69	P	01:01:10	Les trois points sont alignés, ça signifie quoi ? Ça signifie quoi ? Dire que ces trois points sont alignés ?
70	E		ça veut dire que les trois points appartient à la droite.
71	P		La droite passera par les trois points, les trois points appartient à la droite ?
72	P	01:02:03	Si et seulement si les deux vecteurs formés à partir de ces trois points sont ? Deux vecteurs formés à partir de ces trois points. Le vecteur $\overrightarrow{G_1M}$ et le vecteur $\overrightarrow{G_1G_2}$ sont ? Si et seulement si $\overrightarrow{G_1M}$ et $\overrightarrow{G_1G_2}$ sont colinéaires. Quand dit-on que deux vecteurs sont colinéaires ?
73	P	01:02:49	Monsieur, quand dit-on que deux vecteurs sont colinéaires ? Je dis, quand dit-on que deux vecteurs sont colinéaires ?
74	E		Si leur déterminant est égale à à zéro
75			Deux vecteurs sont colinéaires si et seulement si leur déterminant est égale à zéro, quelles sont les coordonnées du vecteur $\overrightarrow{G_1M}$? C'est l'abscisse M moins coordonnées G_1 , c'est $x - \text{quoi}$? $\frac{-13}{3}$ Coordonnées de ce vecteur c'est : $x - \frac{13}{3}$ et $y - 24$ et les coordonnées de $\overrightarrow{G_1G_2}$ c'est : $x_{G_1} - x_{G_2}$ et $y_{G_1} - y_{G_2}$ c'est

			$62 - \frac{13}{3} = \frac{49}{3} \frac{230}{3} - 24 = ?$
76	E		$\frac{158}{3}$
77	P		Le déterminant c'est : $\frac{158}{3}(x - \frac{13}{3}) - \frac{49}{3}(y - 24) = 0$ c'est $\frac{158}{3}x - \frac{49}{3}y - \frac{474}{9}$ nous sommes d'accord ?
78	E		c'est 2054 à la place de 474
79	P		Ici ?
80	E		Oui Monsieur
81	P		Sur 9 ?
82	E		Oui
83	P		Parce que c'est 159×13 ?
84	E		Oui Monsieur
85	P		$3 \times 24 + 49$ le tout sur 3, c'est égale à ?
86	Élèves		$\frac{1176}{3}$
87	P		voilà ce que ça nous donne : $\frac{158}{3}x - \frac{49}{3}y - \frac{2054}{9} + \frac{1176}{3} = 0$ D'accord ? C'est $\frac{158}{3}x - \frac{49}{3}y$, $1176 \times 3 - 2054$ sur 9, c'est ? Pourquoi ? Pour quelle valeur ?
88	E		2130-
89	P		J'ai commencé par $\frac{1176}{3}$, j'ai réduis au même dénominateur. Donc 1176×3 , qu'est ce que ça va vous donner ?
90	E		3528
91	P		$3528 - 2054$ le tout sur 9.
92	E		$\frac{1474}{9}$
93	P		$\frac{158}{3}x - \frac{49}{3}y + \frac{1474}{9} = 0$, OK ?
94	E		oui Monsieur
95	P	01:09:06	je voulais exprimer y en fonction de x. C'est $\frac{-49}{3}y = \frac{158}{3}x$, c'est ça qui passe de l'autre côté, c'est moins. Donc y serait égale à, moins divisé par moins ça, c'est donne plus et donc ça va donner

			$y = \frac{158}{3}$ multiplié par l'inverse de $\frac{49}{3}$, c'est $\frac{3}{49}$ ($y = \frac{158}{3} \times \frac{3}{49}$). On voit que les trois vont se simplifier. On va se retrouver avec $\frac{158}{49}$. Et ce moins divisé par ce moins, ça donne plus. ça va me faire ici, c'est quel nombre ?
96	E		___ ?
97	P		Si l'effectif est un nombre ___ Si c'est impair, c'est 3, 4 ou bien 4, 3. Pour que l'ajustement affine soit justifié, on fait intervenir, une autre grandeur, une autre variable. On aura l'occasion devoir ça en exercice. Je disais la fois dernière entre et autres intérêt de la statistique, c'est de faire des prévisions. Si je donne une valeur à x alors, je peux calculer la variable y. Si je donne une valeur à y, j'ai également la possibilité de déterminer la valeur de x. x c'est quoi ? C'est une autre grandeur. Si je prends l'exemple1, si je connais la note de Biologie alors je peux calculer la note de Physique. Si je connais la note de Physique alors je peux calculer la note de Biologie dans le premier exemple, qui avait été donné. Ça veut dire que y a un lien entre ces deux notes. (01:17:47)
Fin de la deuxième séance			

B LE QUESTIONNAIRE

Questionnaire

Thème1 : Questions centrées sur l'enseignant

- 1) Vous avez combien d'années d'expérience en enseignement ?
- 2) Vous enseignez actuellement :
- Au collège Précisez le niveau.....
- Au lycée Précisez le niveau.....
- A l'université Précisez le niveau.....
- 3) Est-ce que vous utilisez le manuel CIAM pour préparer vos enseignements ?
- Oui Non
- 4) Utilisez-vous d'autres ressources, en dehors du manuel CIAM ?
- Oui Non
- 5) Si oui, lesquelles ?
- 6) Avez-vous déjà enseigné la classe de terminale S (D) ?
- Oui Non
- 7) Si oui, avez vous déjà enseigné la statistique à deux variables en terminale S (D) ?

Si oui, nous vous proposons la suite du questionnaire.
Si non, nous vous remercions de vos réponses.

La suite du questionnaire n'est proposée que si les enseignants ont une expérience, même réduite, de l'enseignement de la statistique à deux variables.

Thème2 : Questions centrées sur la pratique de l'enseignant, exemple d'un exercice.

L'exercice suivant a été proposé à des élèves de Terminale scientifique en Centrafrique.

Le tableau ci-dessous indique le montant des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense que en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette affirmation ?

-
- 8) Est-ce que vous proposeriez cet exercice à vos élèves de terminale S ?:

Oui, tel quel Oui, mais en le modifiant Non

Expliquez les raisons de votre réponse.

Ci-dessous sont présentées deux productions d'élèves correspondant à cet énoncé (la correction est en annexe). Pour chacune des productions d'élève :

9) Analyser la réussite et les erreurs éventuelles.

10) En cas d'erreurs identifiées, comment pourriez-vous aider ces élèves ?

Élève A

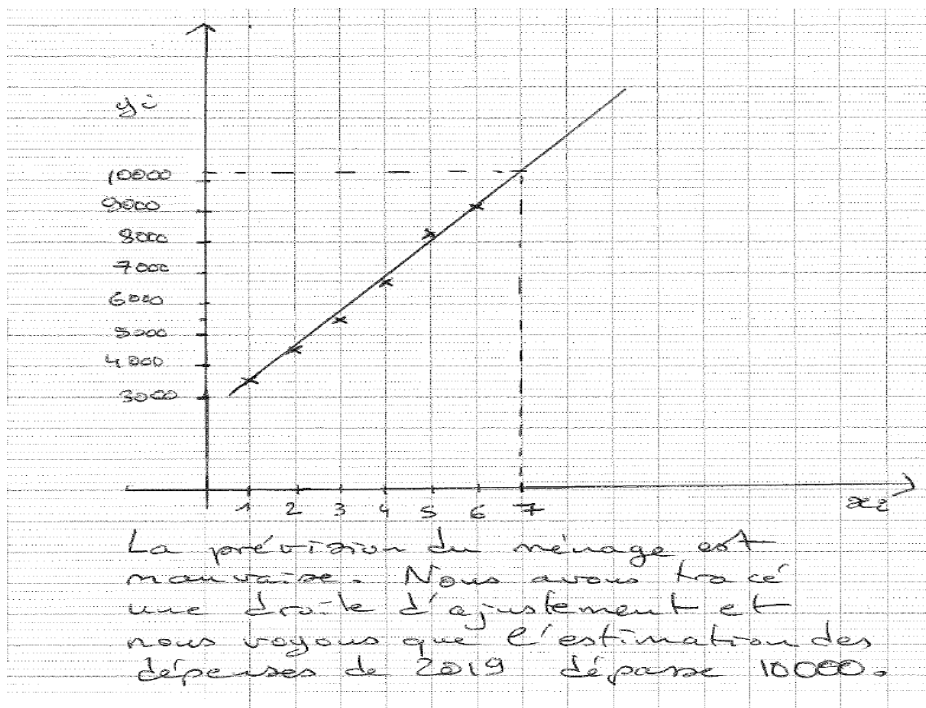
$y = ax + b \Leftrightarrow y = 10.000x + b$

Applique la formule.

$$a = \frac{\text{Cov}(x, y)}{V_x}$$
$$a = \frac{3338,6}{2,91}$$
$$b = \bar{y} - a\bar{x}$$
$$b = 6280,83 - 1145,22 \times 3,5$$
$$b = 2272,5$$
$$y = 1145,22x + 2272,5$$
$$y = 1145,22 \times 10000 + 2272,5$$
$$y = 11454472,5$$

le ménage pense que la prevision en 2019 est fause car le ménage a depense plus

Élève B



Thème 3 : Questions centrées sur la pratique de l'enseignant pour les statistiques à deux variables en général.

11) A propos du thème statistiques à deux variables, est-ce que vous diriez que :

a) L'important est que l'élève apprenne des méthodes d'ajustement : moindre carrés, Méthode de Mayer

b) L'important est que l'élève sache interpréter les résultats de ses calculs

c) L'important est que l'élève parvienne à proposer un modèle adéquat pour faire une estimation

12) A propos des difficultés des élèves avec ce thème, est-ce que vous diriez que :

a) La principale difficulté est que les élèves ne connaissent pas bien les méthodes

b) La principale difficulté est que les élèves n'ont pas de calculatrice pour la plupart

c) Les élèves n'ont pas particulièrement de difficultés avec ce thème.

Commentaires sur l'enseignement des statistiques à deux variables.

ANNEXE : une résolution correcte possible de l'exercice du thème 2

Méthode des moindres carrés.

Équation de la droite de régression de y en x est : $y=ax+b$ avec $a=\frac{\sigma_{xy}}{\sigma_x^2}$

et $b=\bar{y}-a\bar{x}$

- Calcul de \bar{x} et \bar{y}

$$\bar{x}=\frac{1}{n}\sum_{i=1}^6 x_i=\frac{1}{6}\cdot(1+2+3+4+5+6)=3,5$$

$$\bar{y}=\frac{1}{n}\sum_{i=1}^6 y_i=\frac{1}{6}\cdot(3568+4561+5502+6693+8241+9120)=6295,83$$

- Calcul de la variance de x : σ_x^2

$$\sigma_x^2=\frac{1}{n}\sum_{i=1}^6 x_i^2-\bar{x}^2=\frac{1}{6}(91)-(3,5)^2=2,916.$$

- Calcul de la covariance : σ_{xy}

$$\sigma_{xy}^2=\frac{1}{n}\sum_{i=1}^6 x_i y_i-\bar{x}\bar{y}=\frac{1}{6}(151893)-3,5*6295,83=3295,08$$

- $a=\frac{\sigma_{xy}}{\sigma_x^2}=1129,5$ et $b=\bar{y}-a\bar{x}=2342,6$

L'équation de la droite de régression de y en x est donc : $y=1129,5x+2342,6$

L'an 2019 correspond à $x=7$ d'où $y=10249,1$

En 2019 le ménage va dépenser plus de 10 000 euros, pour le remboursement.

Questionnaire Professeur P2

Questionnaire

Thème1 : Questions centrées sur l'enseignant

- 1) Vous avez combien d'années d'expérience en enseignement ? *17*
- 2) Vous enseignez actuellement :
 - Au collège Précisez le niveau.....
 - Au lycée Précisez le niveau... *Ter A, Ter B*
 - A l'université Précisez le niveau.....
- 3) Est-ce que vous utilisez le manuel CIAM pour préparer vos enseignements ?
 - Oui Non
- 4) Utilisez-vous d'autres ressources, en dehors du manuel CIAM ?
 - Oui Non
- 5) Si oui, lesquelles ?
Délic, Tercher
- 6) Avez-vous déjà enseigné la classe de terminale S (D) ?
 - Oui Non
- 7) Si oui, avez vous déjà enseigné la statistique à deux variables en terminale S (D) ?

- Si oui, nous vous proposons la suite du questionnaire.
Si non, nous vous remercions de vos réponses.

La suite du questionnaire n'est proposée que si les enseignants ont une expérience, même réduite, de l'enseignement de la statistique à deux variables.

Thème2 : Questions centrées sur la pratique de l'enseignant, exemple d'un exercice.

L'exercice suivant a été proposé à des élèves de Terminale scientifique en Centrafrique.

Le tableau ci-dessous indique le montant des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense que en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette affirmation ?

- 8) Est-ce que vous proposeriez cet exercice à vos élèves de terminale S ?:

Oui, tel quel Oui, mais en le modifiant Non

Expliquez les raisons de votre réponse.

- Dessiner le nuage des points*
- Compléter le tableau statistiques*
- Déterminer les moyennes \bar{x} et \bar{y} et préciser le point moyen G .*

- Déterminer les équations de la droite de régression.
- Déterminer la valeur de y à partir de la valeur x .

Ci-dessous sont présentées deux productions d'élèves correspondant à cet énoncé (la correction est en annexe). Pour chacune des productions d'élève :

9) Analyser la réussite et les erreurs éventuelles.

10) En cas d'erreurs identifiées, comment pourriez-vous aider ces élèves ?

Élève A

$y = ax + b \Leftrightarrow y = 10.000 \cdot x + b$

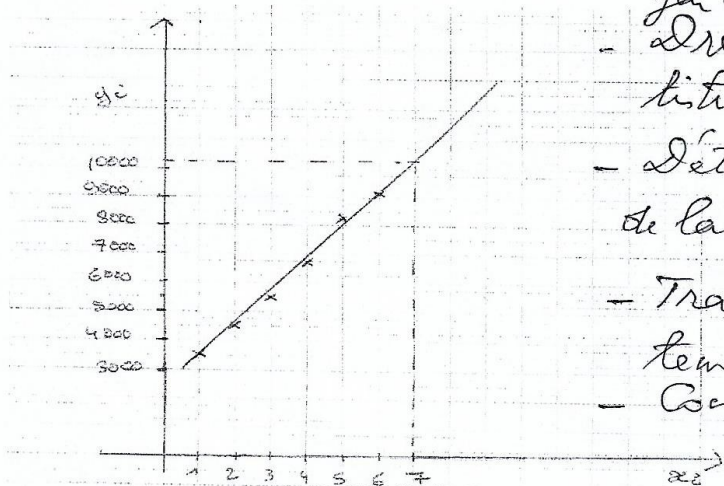
Après la formule.

$$a = \frac{\text{Cov}(x, y)}{V_x} \quad \Leftrightarrow \quad a = \frac{3332,6}{2,91}$$
$$b = \bar{y} - a \bar{x} \quad \Leftrightarrow \quad b = 6280,83 - 1145,22 \times 3,5$$
$$b = 2272,5$$
$$y = 1145,22x + 2272,5$$
$$y = 1145,22 \times 10000 + 2272,5$$
$$y = 11454472,5$$

le ménage pense que la précision en 2019 est faussée car le ménage a dépensé plus

- dresser le tableau statistique des moindres carrés
- Préciser l'équation de la droite de régression
- Déterminer les valeurs de a et b.

Élève B



- Représenter le nuage de points et placer le point moyen \bar{G} .
- Dresser le tableau statistique
- Déterminer l'équation de la droite
- Tracer la droite d'ajustement
- Conclusion.

La prévision du ménage est mauvaise. Nous avons tracé une droite d'ajustement et nous voyons que l'estimation des dépenses de 2019 dépasse 10000.

Thème 3 : Questions centrées sur la pratique de l'enseignant pour les statistiques à deux variables en général.

11) A propos du thème statistiques à deux variables, est-ce vous diriez que :

- a) L'important est que l'élève apprenne des méthodes d'ajustement : moindres carrés, Méthode de Mayer
- b) L'important est que l'élève sache interpréter les résultats de ses calculs
- c) L'important est que l'élève parvienne à proposer un modèle adéquat pour faire une estimation

12) A propos des difficultés des élèves avec ce thème, est-ce que vous diriez que :

- a) La principale difficulté est que les élèves ne connaissent pas bien les méthodes.
- b) La principale difficulté est que les élèves n'ont pas de calculatrice pour la plupart
- c) Les élèves n'ont pas particulièrement de difficultés avec ce thème.

Commentaires sur l'enseignement des statistiques à deux variables.

Questionnaire Professeur P4

Questionnaire

Thème1 : Questions centrées sur l'enseignant

- 1) Vous avez combien d'années d'expérience en enseignement ? **31**
- 2) Vous enseignez actuellement :
- Au collège Précisez le niveau.....
- Au lycée Précisez le niveau.. **Term. C. - Term. D - 2nde c**
- A l'université Précisez le niveau.....
- 3) Est-ce que vous utilisez le manuel CIAM pour préparer vos enseignements ?
- Oui Non
- 4) Utilisez-vous d'autres ressources, en dehors du manuel CIAM ?
- Oui Non
- 5) Si oui, lesquelles ?
- Dimathème , Délic**
- 6) Avez-vous déjà enseigné la classe de terminale S (D) ?
- Oui Non
- 7) Si oui, avez vous déjà enseigné la statistique à deux variables en terminale S (D) ?

Si oui, nous vous proposons la suite du questionnaire.
Si non, nous vous remercions de vos réponses.

La suite du questionnaire n'est proposée que si les enseignants ont une expérience, même réduite, de l'enseignement de la statistique à deux variables.

Thème2 : Questions centrées sur la pratique de l'enseignant, exemple d'un exercice.

L'exercice suivant a été proposé à des élèves de Terminale scientifique en Centrafrique.

Le tableau ci-dessous indique le montant des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense que en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette affirmation ?

- 8) Est-ce que vous proposeriez cet exercice à vos élèves de terminale S ?:

Oui, tel quel Oui, mais en le modifiant Non

Expliquez les raisons de votre réponse.

- En statistique, les étapes importantes sont :*
- 1- Dessiner le nuage des points,
 - 2- Compléter le tableau ci-dessus par des informations données pendant le cours,
 - 3- Déterminer l'équation de la droite de régression de y en x , puis la représenter graphiquement sur le nuage de points,
 - 4- Conclure soit analytiquement, soit géométriquement.

Ci-dessous sont présentées deux productions d'élèves correspondant à cet énoncé (la correction est en annexe). Pour chacune des productions d'élève :

9) Analyser la réussite et les erreurs éventuelles.

10) En cas d'erreurs identifiées, comment pourriez-vous aider ces élèves ?

Élève A

$y = ax + b \Leftrightarrow y = 10.000x + b$

Après la formule.

$$a = \frac{\text{Cov}(x, y)}{Vx} \quad a = \frac{3333,6}{2,91}$$
$$b = \bar{y} - a\bar{x} \quad b = 6280,83 - 1145,22 \times 3,5$$
$$b = 2272,5$$
$$y = 1145,22x + 2272,5$$
$$y = 1145,22 \times 10000 + 2272,5$$
$$y = 11454472,5$$

Le ménage pense que la prévision du coût est fautive car le ménage a dépensé plus

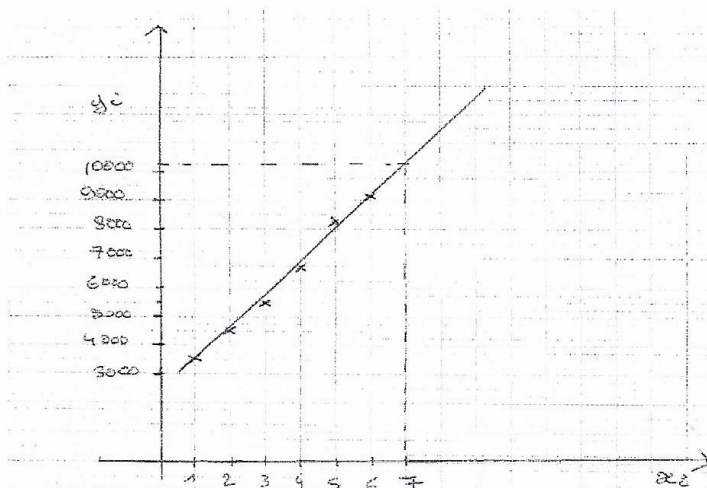
← C'est l'équation de quelle droite ?

← Dresser le tableau statistique donné au cours afin de déterminer la valeur de a , puis celle de b .

← Le résultat ne reflète pas la réalité.

En conclusion, l'élève A doit utiliser le tableau statistique des moindres carrés pour bien déterminer la solution au problème posé.

Élève B



La prévision du ménage est mauvaise. Nous avons tracé une droite d'ajustement et nous voyons que l'estimation des dépenses de 2019 dépasse 10000.

1. Il manque le point moyen $G(\bar{x}, \bar{y})$.
2. La droite d'ajustement tracée sur le graphique ne fait après avoir déterminé son équation réduite.

En conclusion, l'élève B doit d'abord dresser le tableau statistique à deux variables afin de déterminer la solution au problème posé par la méthode des moindres carrés.

Thème 3 : Questions centrées sur la pratique de l'enseignant pour les statistiques à deux variables en général.

11) A propos du thème statistiques à deux variables, est-ce vous diriez que :

a) L'important est que l'élève apprenne des méthodes d'ajustement : moindre carrés, Méthode de Mayer.

Oui, c'est prévu par le programme officiel.

b) L'important est que l'élève sache interpréter les résultats de ses calculs.

Oui.

c) L'important est que l'élève parvienne à proposer un modèle adéquat pour faire une estimation.

Non, l'élève doit connaître ce qu'on lui a appris sur la statistique à deux variables.

12) A propos des difficultés des élèves avec ce thème, est-ce que vous diriez que :

a) La principale difficulté est que les élèves ne connaissent pas bien les méthodes

C'est vrai pour les élèves qui ne travaillent pas.

b) La principale difficulté est que les élèves n'ont pas de calculatrice pour la plupart

Oui.

c) Les élèves n'ont pas particulièrement de difficultés avec ce thème.

Il y a toujours des difficultés car tous les élèves n'ont pas la même capacité intellectuelle de rétention.

Commentaires sur l'enseignement des statistiques à deux variables.

1. Les élèves doivent avoir une calculatrice et les outils de géométrie.

2. Ils ont des difficultés pour représenter graphiquement le nuage de points :

3. problème d'échelle sur les axes.

3. Pour les calculs dans le tableau statistique approprié, il vaut mieux préciser aux élèves les marges d'incertitude par exemple : $\pm 10^3$ près...

Questionnaire Professeur P10 (Aristide)

P10

Questionnaire

Thème1 : Questions centrées sur l'enseignant

- 1) Vous avez combien d'années d'expérience en enseignement ? **25**
- 2) Vous enseignez actuellement :
- Au collège Précisez le niveau.....
- Au lycée Précisez le niveau. **2^{me} C et T^{le} C**
- A l'université Précisez le niveau.....
- 3) Est-ce que vous utilisez le manuel CIAM pour préparer vos enseignements ?
- Oui Non
- 4) Utilisez-vous d'autres ressources, en dehors du manuel CIAM ?
- Oui Non
- 5) Si oui, lesquelles ? **Délic, Terracher, Fractale, Transmath, Annal BAC.**
- 6) Avez-vous déjà enseigné la classe de terminale S (D) ?
- Oui Non
- 7) Si oui, avez vous déjà enseigné la statistique à deux variables en terminale S (D) ?
- oui**

Si oui, nous vous proposons la suite du questionnaire.
Si non, nous vous remercions de vos réponses.

La suite du questionnaire n'est proposée que si les enseignants ont une expérience, même réduite, de l'enseignement de la statistique à deux variables.

Thème2 : Questions centrées sur la pratique de l'enseignant, exemple d'un exercice.

L'exercice suivant a été proposé à des élèves de Terminale scientifique en Centrafrique.

Le tableau ci-dessous indique le montant des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense que en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette affirmation ?

- 8) Est-ce que vous proposeriez cet exercice à vos élèves de terminale S ? :

Oui, tel quel Oui, mais en le modifiant Non

Expliquez les raisons de votre réponse.

1. On pouvait supposer que les variables étaient en corrélation linéaire.
2. On pouvait demander la représentation du nuage de points, cela permettrait une visibilité du lien entre les deux variables.
3. On pouvait préciser la méthode d'ajustement linéaire (Moins ou moindres carrés).

Ci-dessous sont présentées deux productions d'élèves correspondant à cet énoncé (la correction est en annexe). Pour chacune des productions d'élève :

9) Analyser la réussite et les erreurs éventuelles.

10) En cas d'erreurs identifiées, comment pourriez-vous aider ces élèves ?

Élève A

$y = ax + b \Leftrightarrow y = 10.000x + b$

Copier la formule.

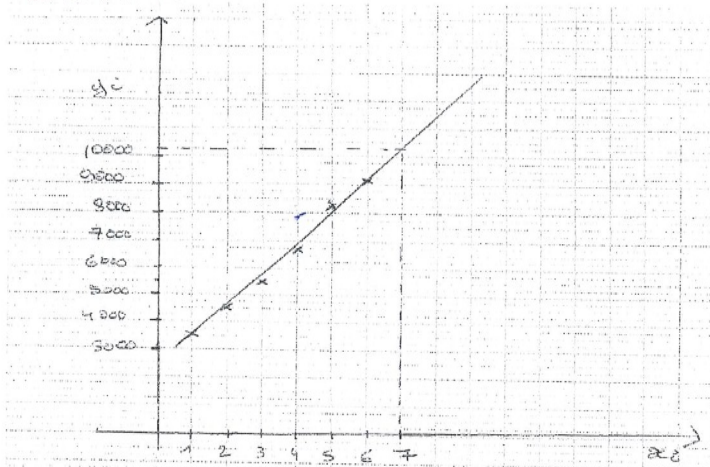
$$a = \frac{\text{Cov}(x, y)}{V_x} \Leftrightarrow a = \frac{3338,6}{4,91}$$
$$b = \bar{y} - a\bar{x}$$
$$b = 6280,83 - 1145,22 \times 5$$
$$b = 2278,5$$
$$y = 1145,22x + 2278,5$$
$$y = 1145,22 \times 10000 + 2278,5$$
$$y = 11454478,5$$

le mariage pose que la précision en 2019 est fautive car le mariage a dépensé plus

- l'élève A a convenablement déterminé une équation de la droite de régression de y en x par la méthode des moindres carrés. Sauf que la réponse à la question posée est fautive (voir corrigé en annexe), du fait que la justification est erronée.

- l'élève B a convenablement construit le nuage de points.
2 Par lecture graphique, il a convenablement répondu à la question posée. Sauf que la lecture graphique n'est pas toujours précise. la lecture graphique pourrait être une conjecture, qui devra être démontrée par la suite.

Elève B



La prévision du ménage est mauvaise. Nous avons tracé une droite d'ajustement et nous voyons que l'estimation des dépenses de 2019 dépasse 10000.

Thème 3 : Questions centrées sur la pratique de l'enseignant pour les statistiques à deux variables en général.

11) A propos du thème statistiques à deux variables, est-ce que vous diriez que :

a) L'important est que l'élève apprenne des méthodes d'ajustement : moindres carrés, Méthode de Mayer

b) L'important est que l'élève sache interpréter les résultats de ses calculs

c) L'important est que l'élève parvienne à proposer un modèle adéquat pour faire une estimation

12) A propos des difficultés des élèves avec ce thème, est-ce que vous diriez que :

a) La principale difficulté est que les élèves ne connaissent pas bien les méthodes

b) La principale difficulté est que les élèves n'ont pas de calculatrice pour la plupart

c) Les élèves n'ont pas particulièrement de difficultés avec ce thème.

Commentaires sur l'enseignement des statistiques à deux variables.

les difficultés mentionnées sont toutes importantes :

- 3
- Non maîtrise des formules
 - Compréhension de la langue (français)
 - Absence de calculatrice (qui devrait faciliter la rapidité dans la résolution de l'exercice)

La maîtrise des méthodes, des formules, l'interprétation
des résultats (faire des prévisions), sont des objectifs de ce
thème.

L'une des difficultés des élèves avec ce thème est que ceux-ci
évitent le calcul, avec de grands nombres

C GUIDE D'ENTRETIEN

Guide d'entretien

Première partie : retour sur le questionnaire rempli par l'enseignant

Vous avez rempli le questionnaire, merci, je vous propose de regarder ensemble vos réponses pour me donner quelques détails supplémentaires. (s'adapter ici aux réponses. Par exemple, si l'enseignant utilise un autre manuel que le CIAM, demander pourquoi etc.)

1. Pourquoi vous utilisez un autre manuel en dehors du manuel CIAM ? Pouvez-vous justifier votre choix ?
2. Que pensez-vous du manuel CIAM ?
3. Que-ce que vous pourriez me dire de plus sur vos réponses aux 9 et 10 ?

Deuxième partie : vos pratiques d'enseignement des séries statistiques doubles et de l'ajustement

1. Comment introduisiez-vous la série statistique à deux variables en classe ?
2. Est-ce que vous pourriez me dire, quels sont les objectifs de l'enseignement de la statistique à deux variables en terminale S (D) ?
3. D'après vous, quelles sont les compétences en statistique à deux variables qui vous paraissent importantes à construire chez les élèves ?
4. Est-ce que vous pourriez m'expliquer comment vous vous organisez avec vos élèves pendant votre cours ?
5. Est-ce que vous faites les travaux dirigés à la fin de chaque chapitre ?
6. Si oui, comment vous procédez ? Est-ce que, c'est les élèves qui passent au tableau corrigés les exercices ou bien ?
7. Est-ce que vous leur donniez des exercices à faire d'abord à la maison avant de corriger en classe ?
8. De quelle manière vous évaluez les élèves ?
9. D'après vous, quelles sont les difficultés des élèves en statistique à deux variables ?

Troisième partie : les initiatives des élèves en statistiques à deux variables

1. Selon vous, est-ce que dans vos cours habituels de statistiques à deux variables les élèves doivent prendre des initiatives (expliquer : les élèves peuvent choisir une méthode, une représentation graphique, une interprétation de leurs calculs) ?
2. Est-ce que vous pensez qu'il faudrait tenter de développer ces initiatives ?

Transcription d'entretien avec Professeur P2

07 juillet 2020

CH=Chercheur

P2=Professeur 2

Durée=15minutes

1	CH	<p>Merci de m'avoir accorder votre temps, euh notre entretien va se passer en trois parties. La première partie concerne le questionnaire que vous avez rempli la fois passée.</p> <p>Vous avez rempli le questionnaire, merci, je vous propose de regarder ensemble vos réponses pour me donner quelques détails supplémentaires.</p> <p>1. Pourquoi vous utilisez un autre manuel en dehors du manuel CIAM ? Pouvez-vous justifier votre choix ?</p>
2	P2	<p>Oui, merci d'avoir euh, poser cette question. Je crois que la réalité de la RCA et l'expérience avec nos élèves, si on utilise un seul manuel euh par exemple CIAM pendant le cours, si les élèves savent que c'est le livre CIAM qui est utilisé beaucoup plus, il y aura une négligence de la part des élèves, ils ne vont pas suivre ce cours là en tant que tel. Donc ils pensent que tout est déjà dans le livre CIAM, on fait le copier coller et donc ils ne vont pas bien suivre les explications en classe, d'où nécessité d'utiliser d'autres livres pour varier un peu les notions de telle manière que les élèves soient attentif dans la salle. Voilà les raisons pour lesquelles nous sommes obligés d'utiliser d'autres manuels.</p>
3	CH	<p>2. Que pensez-vous du manuel CIAM ?</p>
4	P2	<p>Je vois que le manuel CIAM, c'est un bon manuel qui répond à la réalité de la RCA, de la sous région. Et donc je crois que ceux qui ont travaillé sur le livre CIAM, ils ont bien vu l'importance et ils ont..., et ce livre là était conçu conformément à la réalité de la République Centrafricaine.</p>
5	CH	<p>Ok d'accord, maintenant on va regardé ensemble vos réponses à la question 9 et 10 du questionnaire.</p> <p>Je rappelle la question 9. On a proposé les réponses de deux élèves, on vous a demandé d'analyser leur réussite et leurs erreurs, et comment vous pouvez les aider.</p> <p>3. Que-ce que vous pourriez me dire de plus sur vos réponses aux 9 et 10 ? 4. Commençons par l'élève A. Pour l'élève A est-ce que vous avez des choses à me dire sur ce que vous avez donnez comme commentaire sur l'élève A ?</p>
6	P2	<p>Oui, ce que j'ai constaté sur la rédaction de l'élève A, il y a des valeurs qu'il a utilisé ici, mais on n'a pas vu réellement d'où sortent ces valeurs là. Donc ça veut dire qu'il n'y a pas le tableau statistique que l'enfant [élève] a établi en tant que tel, pour lui permettre d'avoir ces valeurs et ensuite il n'a pas précisé l'équation de la droite de régression, pour lui permettre de déterminer les valeurs de a et b avant de, de retrouver l'équation de la droite de régression. Voilà ce que j'ai constaté pour l'élève A.</p>
7	CH	<p>Donc d'après vous, pour trouver l'équation de la droite, vous suggérez à l'élève A de commencer par un tableau statistique ?</p>
8	P2	<p>Un tableau statistique afin de lui permettre de préciser l'équation de la droite.</p>
9	CH	<p>Ok d'accord, et concernant l'élève B, lui, il a utilisé une méthode graphique pour</p>

		répondre à la question. Qu'est-ce que vous pouvez me dire par rapport à l'élève B ?
10	P2	C'est une bonne méthode pour l'élève B, mais sauf qu'il n'a pas représenté le nuage de points, mais bon, il a représenté quelques points du nuage. Mais il n'a pas placé le point moyen G, il est obligé de tracer sa droite de régression. Et normalement le plus souvent, on sait que la droite de régression doit passer obligatoirement par le point moyen G, mais on n'a pas vu euh disant le point moyen G et il n'a pas déterminé l'équation de la droite de régression. Mais donc là, et il n'a pas pu conclure en tant que tel. Normalement l'importance de la droite de régression est de savoir l'estimation après et est ce que c'est à peu près de la réalité par rapport au nuage de points que l'enfant doit tracer.
11	CH	Ok d'accord merci, on va passer à la deuxième partie de notre échange, qui concerne vos pratiques en tant qu'enseignant sur la série statistique double. 4. Comment introduisiez-vous la série statistique à deux variables en classe de terminale? Bien sûr avec vos élèves
12	P2	Je crois que pour une bonne leçon, il faut d'abord commencer par les notions de prérequis, les notions fondamentales qui permettent à l'enfant d'assimiler vite, de maîtriser le chapitre. On introduit le cours en passant d'abord par les prérequis, les notions fondamentales que l'enfant doit d'abord connaître avant d'aborder ce chapitre.
13	CH	Ok, en parlant des prérequis, comme dans le programme, la statistique à une variable est enseignée en classe de première, vous considérez le chapitre sur la série à une variable, avant de commencer votre cours en terminale, vous considérez le chapitre sur la série à une variable comme les notions de prérequis que vous avez besoins pour introduire votre leçon sur la série à deux variables ?
14	P2	Évidemment, je crois que chez nous ici, maintenant on introduit même ce chapitre au niveau du premier cycle. Les notions fondamentales là, on peut dire que ce sont les intitulés qu'un élève de la terminale doit d'abord maîtriser avant de passer à la série statistique à deux variables.
15	CH	Ok 5. Est-ce que vous pourriez me dire, quels sont les objectifs de l'enseignement de la statistique à deux variables en terminale S (D) ?
16	P2	Oui, je crois que, à mon avis la statistique enseignée au niveau du fondamental 2 en général, ce sont les intitulés. Par ce qu'on sait qu'au niveau de la terminale, on a la statistique et pour permettre à ces enfants de maîtriser la statistique en tant que telle, il faut d'abord ces notions fondamentales, les intitulés en tant que telles pour ce chapitre là. Afin qu'à l'université, on sait que par exemple la statistique est beaucoup plus utilisée dans plusieurs départements. Si je prends par exemple, disant d'une manière générale la statistique son objectif c'est de proposer des solutions aux décideurs et donc là, ce que nous faisons au niveau du premier cycle ce sont les intitulés pour amener les enfants à comprendre et à étudier par exemple la population etc, l'environnement de la population et comme ça au niveau supérieure ils seront capable d'utiliser réellement la statistique pour avancer, pour permettre aux décideurs de prendre leur décision.
17	CH	Ok merci. 6. D'après vous, quelles sont les compétences en statistique à deux variables qui vous paraissent importantes à construire chez les élèves ?

18	P2	Je crois que, c'est beaucoup plus la droite de régression par ce que c'est avec..., c'est la représentation graphique, il faut représenter d'abord le nuage de points ensuite tracer la droite de régression, pour voir est ce qu'il ait un bon ajustement qui va nous permettre d'être proche de la réalité. Donc je crois beaucoup plus qu'on doit mettre l'accent sur la droite de régression.
19	CH	Mais que pensez-vous de l'interprétation des résultats en statistique au niveau de la classe de terminale ?
20	P2	Je crois que comme je viens de vous le dire. Je crois que les enfants n'ont pas encore bien compris, euh disant l'objectif de ce chapitre. Et ça fait que parfois il y a d'autres élèves qui font ça, parce que c'est une leçon qu'on doit enseigner en salle et ils cherchent à comprendre mais ils ne comprennent pas pourquoi on fait la statistique. Et c'est là que je réitère encore que la statistique, normalement ce qu'on fait c'est pour proposer une solution aux décideurs mais les enfants n'ont pas bien compris mais comme nous sommes là c'est notre travail donc, on doit mettre toujours de l'accent de telle manière que ces enfants là comprennent pourquoi on fait la statistique.
21	CH	OK, 7. Est-ce que vous pourriez m'expliquer comment vous vous organisez avec vos élèves pendant votre cours ?
22	P2	Oui, c'est ce qu'on a l'habitude de faire même pour toute les autres leçons. Euh y a les notions de prérequis et je crois que la méthodologie la plus utilisée, c'est ce la méthodologie, c'est la méthode active qu'il faut faire participer les élèves et donc euh, lorsque vous introduisez, vous posez les questions, c'est aux élèves de reprendre, de chercher à découvrir, de réfléchir et faire passer les élèves aussi au tableau pour que, eux même ils réalisent, ils comprennent la leçon en tant que telle.
23	CH	8. Est-ce que vous faites les travaux dirigés à la fin de chaque chapitre ?
24	P2	Oui, c'est une obligation mais si on ne fait pas des travaux dirigés, comment on peut évaluer ce qu'on a enseigné ? Parce que y a ce qu'on appelle l'auto..., je peux dire l'autocontrôle euh, donc lorsqu'on fait une leçon, à la fin il faut évaluer et pour évaluer il faut donner des travaux dirigés pour permettre de savoir si les enfants ont compris ou ils n'ont pas compris, et ça va vous permettre de revenir sur les parties qu'ils n'ont pas compris pour réexpliquer encore d'avantage.
25	CH	OK et puis, 9. Comment vous procédez avec vos élèves? Est-ce que, c'est les élèves qui passent au tableau corrigés les exercices ou bien ?
26	P2	Je crois qu'on dit bien travaux dirigés, donc, on doit donner l'exercice et donner d'abord le temps à ces enfants là de réfléchir, d'utiliser les notions qu'ils ont dans leur cahier et ensuite réagir. Et c'est à ce moment là que, si vous voyez les failles vous pouvez maintenant réexpliquez la partie du cours là, c'est-à-dire les enfants n'ont pas bien compris s'ils n'arrivent pas à reprendre correctement.
27	CH	10. Est-ce que vous leur donniez des exercices à faire d'abord à la maison avant de corriger en classe ?
28	P2	OUI, je crois que lorsqu'on fait un cours. Pour permettre une bonne compréhension aux élèves, il y a ce qu'on appelle les applications. Les applications on les fait ensemble en salle mais y a des exercices qu'on considère comme travaux dirigés. Il faut donner aux enfants, leur accordés à peu près trois jours ou une semaine pour

		faire de recherche, chercher à comprendre le cours pour appliquer à ces exercices là. Et par après, on revient en classe, eux mêmes ils vont corriger , et si la première personne passe au tableau et si ça ne marche pas une deuxième personne passe, eux mêmes ils vont essayer de se corriger entre eux. Si ça ne va pas que l'enseignant peut intervenir à la fin . En plus de ça, y a ce qu'on appelle le travail de groupe, parfois on fait les travaux dirigés en deux ou trois groupes, il faut pas que le nombre exagère sinon il y aura des bavardages inutiles. Le travaille de groupe permet aux élèves d'échanger entre eux.
29	CH	11. De quelle manière vous évaluer les élèves ?
30	P2	Lorsqu'on fait un cours après on fait les travaux dirigés, donc on évalue les élèves en fonction du cours et des travaux dirigés, pour savoir si réellement le message est passé.
31	CH	Et souvent c'est des évaluations sur table ou bien c'est des évaluations qu'ils font à la maison et vous remettent ?
32	P2	Oui, pour moi le plus souvent ce sont des évaluation sur table, je vais parfois de deux manières, soit c'est individuel ou ils peuvent faire à deux. Deux personnes sur une copie, ils vont réfléchir, échanger et rédiger leur devoir.
33	CH	12. D'après vous, quelles sont les difficultés des élèves en statistique à deux variables ?
34	P2	Oui, y a des élèves qui ne savent pas remplir le tableau statistique, parce que pour la statistique à deux variables ce qui est trop cruciale c'est le tableau statistique. Lorsqu'on arrive à remplir le tableau statistique correctement, je crois que l'application des formules est facile. De deux les difficultés des élèves c'est aussi au niveau, disant du repère, parce que le plus souvent la statistique à deux variables, on utilise un repère orthogonal donc y a un problème d'échelle qui se pose le plus souvent. Donc voilà à peu près là où les élèves ont des difficultés le plus souvent.
35	CH	Ok d'accord, on va aborder la troisième partie de notre échange, qui est portée sur les initiatives des élèves. Ce que je considère ici comme l'initiative des élèves lorsque pendant un exercice, les élèves peuvent choisir eux même les méthodes, faire des représentations graphique, interpréter les résultats des calculs, donc ils peuvent eux même trouver des stratégies, c'est ça que je considère comme des initiatives. Donc, 13. Selon vous, est-ce que dans vos cours habituels de statistiques à deux variables les élèves doivent prendre des initiatives ?
36	P2	Ok, si nous enseignons je crois que y a d'autres élèves qui ont cette capacité de faire des recherches, si par exemple moi j'utilise le livre CIAM et et par exemple Déclic ou Terracher mais un enfant qui peut avoir des documents sur la statistique et utilisées d'autres méthodes qu'il a vu dans d'autres livres. Je crois que pour l'enseignant, il s'agit d'encourager cet enfant là, parce qu'il a l'initiative de la recherche et l'essentiel c'est de voir est ce que y a la logique mathématique dans ce qu'il fait. Si y a la logique donc c'est appréciable.
37	CH	14. Est-ce que vous pensez qu'il faudrait tenter de développer ces initiatives ?
38	P2	Oui, il faut toujours encourager, nous sommes dans le monde de recherche donc y a des enfants qui ont la possibilité de l'accès à internet donc il faut les encourager à faire la recherche, parce que, si on fait un cours par rapport à l'objectif fixé mais y a d'autres enfants qui sont capables de faire leur recherche, ils peuvent aller au-delà et avoir des connaissances adéquates par rapport à la leçon. Moi je crois que c'est

		toujours à encourager.
39	CH	Ok, je vous remercie d'avoir répondu à mes questions.

Transcription d'entretien avec Professeur P4

20 juillet 2020

CH=Chercheur

P4=Professeur4

Durée=17minutes 39secondes

1	CH	<p>Bonjour, merci de m'avoir accorder votre temps pour cet échange. Notre échange va se passer en trois parties. La première, on va échanger sur le questionnaire que vous avez rempli la fois passée, la deuxième partie sera sur vos pratiques en tant qu'enseignant et la troisième partie, on va échanger sur les initiatives des élèves, en ce qui concerne la statistique descriptive.</p> <p>Vous avez rempli le questionnaire, merci, je vous propose de regarder ensemble vos réponses pour me donner quelques détails supplémentaires.</p> <p>1. Pourquoi vous utilisez un autre manuel en dehors du manuel CIAM ? Pouvez-vous justifier votre choix ?</p>
2	P4	<p>Oui, merci de m'avoir donner la parole. Pour préparer un cours, on ne peut pas seulement se borner sur un seul document donc il faut,</p> <p>C'est un professeur, il prépare son cours, il peut y avoir un, deux, trois, plusieurs documents pour que le cours soit solide.</p>
3	CH	<p>Ok, est-ce que vous faites parfois recours à l'internet pour préparer vos cours ? Et sinon pourquoi ?</p>
4	P4	<p>Pour moi, je n'ai pas recours à l'internet, y a un problème de moyen [moyen financier] qui se pose.</p>
5	CH	<p>2. Que pensez-vous du manuel CIAM ?</p>
6	P4	<p>Le manuel CIAM, c'est vraiment conforme au programme pour les pays francophone Africain, donc c'est un document très très convenable, donc ça sciait.</p>
7	CH	<p>3. Que-ce que vous pourriez me dire de plus sur vos réponses aux 9 et 10 ?</p> <p>Voilà pour l'élève A, ce que vous avez donnez comme commentaire. Vous avez dit que l'élève A doit utiliser le tableau statistique pour bien déterminer la solution au problème. Qu'est-ce que vous pouvez me dire de plus sur l'élève A ?</p>
8	P4	<p>Ça c'est le programme, le programme de l'enseignement de la statistique, parce que en statistique y a deux choses qui sont très importantes, y a le graphique et le tableau. Donc le tableau d'abord ensuite le graphique et sont deux choses très importantes et sont mêmes liées. Donc si un est là, l'autre doit en principe apparaître. C'est pourquoi j'avais demandé à ce que cet élève là fasse d'abord le tableau statistique.</p>

9	CH	Ok concernant l'élève B, lui il a utilisé la méthode graphique pour résoudre l'exercice, qu'est-ce que vous pouvez me dire par rapport à l'élève B ?
10	P4	Il a seulement utilisé le graphique mais le tableau, c'est-à-dire, c'est vice versa, l'élève A, il manquait le tableau et l'élève B aussi il manque le tableau. Parce que pour chercher le coefficient de l'équation de la droite de régression, c'est à partir du tableau, parce que vous avez le tableau statistique initialement au sujet là bas donc le sujet qui amène l'élève à avoir la droite de régression. Mais si déjà dans le tableau, l'enfant n'a pas le tableau, comment il va répondre ? Ça c'est imaginé, c'est purement mathématique, mais c'est pas la statistique qui est demandé là.
11	CH	Ok, on va passer à la deuxième de notre entretien qui concerne vos pratique en tant qu'enseignant. Vous pouvez me dire, 4 Comment introduisiez-vous la série statistique à deux variables en classe ?
12	P4	Bon pour introduire la leçon de statistique à deux variables en classe, l'enfant doit d'abord appris en premier cycle la série statistique à une variable et maintenant au second cycle précisément à partir de seconde et première, on doit maintenant introduire la notion de statistique à deux variables et quant il sera à l'université, ça sera à plusieurs variables. Et c'est pourquoi en classe de première par exemple, je parle d'un exemple simple, par exemple du sport. Le sport on peut lancer le poids et aussi on peut parler de taille et âge. Et pour l'amener l'enfant à comprendre si y a t-il des relations entre ces termes là. Et c'est pourquoi ça abouti même, en biologie y a ce qu'on appelle le tableau de, en génétique là de Mendel. Le tableau de Mendel, le tableau à double entré, dès que vous parlez de tableau à double entré automatiquement les enfants réagissent. Donc ce sont les applications de la série statistique à deux variables.
13	CH	5 Est-ce que vous pourriez me dire, quels sont les objectifs de l'enseignement de la statistique à deux variables en terminale S (D) ?
14	P4	L'objectif effectivement, comme je viens de le dire, y a beaucoup d'objectifs c'est-à-dire là, en biologie l'enfant se retrouve avec le tableau croisé, l'enfant se retrouve. Donc là ça prépare aussi l'élève après son bac à amorcer les études. Donc il devra aussi essayer la statistique à plusieurs variables. L'idée ici c'est de préparer le candidat après son bac à aborder les études supérieures.
15	CH	Ok c'est bien, mais concernant les objectifs, si on reste rien qu'en terminale. Est-ce que, à la fin du cours (chapitre) qu'est ce que les élèves doivent retenir comme les grandes lignes ?
16	P4	A la fin du cours l'enfant, comme tout à l'heure, je l'avais signalé, on est en statistique donc les applications en mathématiques. Donc là l'enfant doit savoir dès qu'on parle de la statistique niveau terminale, il doit y avoir un tableau, les données comme points de départ, faut qu'il y ait l'information et maintenant de ces données là, il peut exploiter, vue que les questions sont posées pour aboutir à ce qu'on lui a demandé.
17	CH	Comme vous parlez d'exploitation, est-ce que vous parlez en terme d'interprétation des résultats ou bien ?
18	P4	Oui, de fois dans le tableau, on peut lui demander d'interpréter des résultats, admettant qu'il a le nuage des points, si il arrive à dessiner, à construire le nuage des points. On voit la forme, est-ce que la forme est allongée, qu'est-ce qu'il dit sur la forme ainsi de suite. Il doit l'interpréter et cela lui donne l'idée en

		statistique, comment ça se passe en statistique donc c'est quelque chose de concret qu'on doit exploiter.
19	CH	6. D'après vous, quelles sont les compétences en statistique à deux variables qui vous paraissent importantes à construire chez les élèves ?
20	P4	Comme tout à l'heure je le dis, l'enfant sache interpréter des données qui peuvent se lier, comme par exemple : âge, poids. L'exemple encore simple à la santé. A la maternité par exemple quand les enfants sont nées, la maman va à l'hôpital, on lui donne un carnet. Vous voyez, y a des courbes là, poids de l'enfant, l'enfant qui a deux mois doit avoir tel poids ainsi de suite. Et si y a un moment donné, ça joue dans la vie quotidienne et donc la statistique joue un rôle très fondamentale dans la vie de l'enfant, l'élève.
21	CH	7. Est-ce que vous pourriez m'expliquer comment vous vous organisez avec vos élèves pendant votre cours ?
22	P4	Comme c'est moi le Professeur mais malheureusement n'y a pas assez des documents donc c'est comme un cours qui se prépare, y a des objectifs, donc on commence le cours, on donne des exemples avec des appui mathématiques, faut pas oublier qu'à la fin du cours on donne des exercices simples, qu'on traite et pour continuer avec des approfondissements à domicile, pour le candidat ou pour l'élève.
23	CH	Vous avez soulevé un aspect ici, vous dites qu'il n'y a pas assez des documents. Est-ce que chaque élève à au moins un manuel pendant le cours ou comment ça se passe ?
24	P4	Vu notre pays, l'état dans laquelle nous trouvons, y avait des documents à l'époque mais avec les événements qui sont passés tout est détruit, les parents des enfants qui ont les reins solides (c'est-à-dire les parents qui ont les moyens financiers) payent les documents à leurs enfants, auquel cas c'est le professeurs seul qui endosse la responsabilité, donc s'il prépare bien son cours pour ne pas que les élèves qui sont démunis ne soient pas pénaliser.
25	CH	Donc ça veut dire que, si les élèves qui n'ont pas de document, ils se contentent des cours et des exercices que l'enseignant leur donne ?
26	P4	Justement, l'enseignant doit tout faire pour que le candidat qui doit faire l'examen, ne soit pas dans des mauvais draps.
27	CH	8. Est-ce que vous faites les travaux dirigés à la fin de chaque chapitre ?
28	P4	Comme je l'ai tout à l'heure dit, tout cours qui fini, y a des exercices directs, des exercices d'approfondissement qu'on donne aux enfants à faire à la maison et le lendemain on corrige.
29	CH	9. Est-ce que vous leur donniez des exercices à faire d'abord à la maison avant de corriger en classe ?
30	P4	Oui, y a ce qu'on appelle exercice d'approfondissement donc ça amène l'enfant à se cultiver, et si y a des documents quelques part, il peut les utilisées pour résoudre ce problème là.
31	CH	10. Est-ce que, c'est les élèves qui passent au tableau corrigés les exercices ou bien ?
32	P4	Dans le cas de fois normalement dit, en classe d'examen, c'est l'élève qui est au

		tableau, parce qu'on le prépare à affronter l'université, les études supérieures donc il doit être au tableau, toi le professeur tu es là en tant que guide, tu es entrain de le guider pour ne pas que, si y a erreur, tu intervies. Mais dans les corrections c'est comme les TD, on prépare le candidat déjà, dès qu'il sera à l'université y a des TD qui se font et c'est l'étudiant qui est au tableau lors des TD, ce n'est pas le professeur.
33	CH	11. De quelle manière vous évaluer les élèves ?
34	P4	On évalue les élèves par rapport au programme, c'est-à-dire que là, en principe y a plusieurs sortent d'évaluation mais ici, vue la nouvelle méthode d'enseignement, on évalue par devoir de table. On informe d'abord et ils préparent le devoir.
35	CH	12. D'après vous, quelles sont les difficultés des élèves en statistique à deux variables ?
36	P4	Les difficultés c'est que, comme ils ont des problèmes en mathématiques, heureusement, ils sont des élèves de terminale C et D mais y a toujours encore des problèmes donc la plus grosse difficulté, ce sont les mathématiques et ensuite l'usage des calculatrices fait défaut. Par ce que l'élève qui n'a pas de calculatrice vous voyez ça fait mal. Là y a aussi du coté matériel, la calculatrice, l'enfant doit avoir la calculatrice pour mieux travailler. Mais les notions mathématiques sont beaucoup utilisées en statistiques, si l'enfant ne maîtrise pas bien ça fait un problème.
37	CH	D'accord, selon vous, les difficultés que les élèves ont beaucoup plus en statistique, c'est qu'ils ne maîtrisent pas bien les notions mathématiques qu'ils vont utiliser en statistique, en plus de ça, y a aussi le problème de calculatrice qui se pose ?
38	P4	Justement
39	CH	13. Selon vous, est-ce que dans vos cours habituels de statistiques à deux variables les élèves doivent prendre des initiatives? J'attends par initiatives des élèves, le choix de méthodes pour faire une prévision par exemple. Dans le cas des interprétations des résultats qu'ils ont trouvé, c'est ça que je qualifie des initiatives. Est-ce que pendant vos cours les élèves ont cet habitudes ?
40	P4	Non ce n'est pas refusé mais le cours suit un programme, le programme dans les instructions officielles euh donc l'élève et non seulement ça, nous sommes dans le domaines des mathématiques, c'est le sujet qui va diriger l'élève a utilisé la méthode qu'il faut. Mais si jamais y a pas une méthode précisée et que l'élève propose ça, mais toi le correcteur tu vois est-ce que ça cadre bien et si ça ne cadre pas, c'est-à-dire la méthode ne convient pas, sinon ça va bouleverser l'élève. Donc s'il propose une méthode qui ne cadre pas à l'énoncé mais, on laisse tomber.
41	CH	Mais il peut arriver que vous donnez par exemple un exercice et vous demandez aux élèves de déterminer la droite d'équation sans préciser la méthode par exemple, ni préciser si c'est la méthode de Mayer qu'ils vont utiliser ou la méthode des moindres carrés. Est-ce que ça arrive que vous dites simplement de déterminer la droite d'équation sans préciser laquelle des deux méthodes utiliser ?
42	P4	Y a un danger là, si on ne précise pas, parce que y a une méthode là, y a des méthodes qui font défaut. Mais ce sont les programmes, programmes qui précisent

		<p>. C'est la méthode des moindres carrés qui est beaucoup plus utilisée et c'est cette méthode là qui est la plus sûre, mais Mayer là, y a des inconvénient, c'est pourquoi on n'a pas parlé beaucoup de ça. Mais c'est plutôt avec la méthode des moindres carrés, c'est plus sûre, et le programme recommande ça. Et c'était d'abord dans le sujet, la méthode des moindres carrés, on signale mais si on ne signale pas dans ce cas, soit le sujet celui qui a conçu le sujet là n'a pas vue d'erreur si on utilise d'autres méthodes. Mais au quel cas surtout avec les parties décimales. Si avec la méthode des moindres carrés vraiment ça confine bien les erreurs, mais sans ça avec l'autre méthode de fois y a des erreurs et de fois le coefficients de corrélation ça dépasse 5. Voilà ça ne marche pas, or le coefficient de corrélation est toujours compris entre -1 et 1. Et c'est grâce à la méthode des moindres carrés que le carré qui confine très bien les erreurs de la partie décimale là.</p> <p>D'abord on signale ça, c'est ça qui fera que l'enfant va trouver la réponse entre -1 et 1, si ça dépasse ça, ça ne marche pas.</p>
43	CH	In peut arriver que vous demandez aux élèves d'interpréter les résultats trouvés ? Quant ils font des calculs par exemples, le cas de coefficient de corrélation. Est-ce que vous avez l'habitude de leur demander d'interpréter ?
44	P4	Justement c'est par rapport au programme, on parle juste d'ajustement, est-ce que c'est un bon ajustement et c'est tout on arrête. Par ce qu'on ne peut pas demander tout à l'élève. Parfois aussi, on demande de chercher une inconnue mais on reste au niveau des ajustements, est-ce que c'est un bon ajustement ou pas.
45	CH	14. Est-ce que vous pensez qu'il faudrait tenter de développer ces initiatives ?
46	P4	Non mais, comme je l'ai dit, on ne peut pas sortir du programme, l'enseignement de la statistique, l'enseignant reste dedans, par ce que, il sait que une fois le bac obtenu, l'enfant verra ça, donc il faut le strict respect du programme.
47	CH	Ok, si je comprends bien, selon vous, les initiatives à développer, c'est des initiatives qui reste dans le cadre du programme fixé pour le niveau des élèves de terminales ?
48	P4	Justement, par ce que si on amène l'élève à sortir hors du programme, y a le bac devant lui et le programme de mathématique au bac doit être respecté, c'est parce que le programme est très très indispensable pour un enseignant qui dispense le cours.

Transcription d'entretien avec P10 (Aristide)

20 Août 2020

CH=Chercheur

P10=Professeur10 (Aristide)

Durée=25minutes 45secondes

1	CH	<p>Merci de m'avoir accorder votre temps pour cette échange. On faire un entretien qui est répartie en trois parties. La première partie concerne le questionnaire que vous avez rempli.</p> <p>Vous avez rempli le questionnaire, merci, je vous propose de regarder ensemble vos réponses pour me donner quelques détails supplémentaires.</p> <p style="padding-left: 40px;">1. Pourquoi vous utilisez un autre manuel en dehors du manuel CIAM ? Pouvez-vous justifier votre choix ?</p>
2	P10	<p>Merci de m'avoir donner la parole. Vous savez que c'est très important d'utiliser plusieurs manuels à la fois, car ces manuels là utilisent plusieurs approches différentes. Au lieu d'utiliser seulement un seul manuel qui, lorsqu'on se confronte à un certains nombres de difficultés, on ne peut pas avancer dans la confection de ces fiches de cours. Et donc il est important d'avoir à sa disposition quatre, cinq manuels de tendance différentes du manuel.</p> <p>Chez nous ici, on a les manuels qui viennent du Cameroun, le manuel CIAM, on a aussi les manuels qui nous viennent de la France donc que, j'ai aussi remarqué dans les manuels là, y a des approches qui sont vues d'une autre manière que CIAM.</p>
3	CH	<p>OK merci.</p> <p style="padding-left: 40px;">2. Que pensez-vous du manuel CIAM utilisé en terminale D?</p>
4	P10	<p>Le manuel CIAM est un manuel comme tout les autres manuels qui a certes ces atouts et certainement ces insuffisances, d'où la nécessité effectivement, comme</p>

		je le disais tantôt d'utiliser plusieurs manuels à la fois, euh, vous êtes sans ignorés que le manuel CIAM a été confectionné grâce à la collaboration de certains professeurs de mathématique de la sous région mais pratiquement de l'Afrique. Parce que, on utilise à une certaine époque, on a utilisé en Afrique de l'ouest. Donc le manuel CIAM comme je le disais tantôt, il a ces atouts certes également ces insuffisances. C'est un bon livre euh n'empêche qu'on utilise d'autres manuels. Et peut être sans oublier le manuel CIAM a été conçu par rapport au programme qui a été défini dans la collection, donc par rapport au programme arrêté dans l'Afrique francophone. Le manuel CIAM reprend à tout ces attentes, je parle du programme.
5	CH	3. Que-ce que vous pourriez me dire de plus sur vos réponses aux 9 et 10 ?
6	P10	Les questions 9 et 10, il était question je crois d'analyser la réussite et les erreurs éventuelles, voici, ce qui concerne la question 9. Pour cette question, je crois avoir constaté que l'élève A a convenablement déterminer une équation de la droite de régression de y en x par la méthode des moindres carrés. Sauf que la réponse à la question posée est fausse. Le corrigé se trouve à l'annexe. La réponse est fausse pour la simple raison que cet élève a donné une justification erronée, voilà l'erreur commise par l'élève A. L'élève B quant à lui, il a convenablement construit le nuage des points par la lecture graphique, il a convenablement répondu à la question posée. Sauf que la lecture graphique n'est pas précise et donc, cette lecture graphique pouvait être considérée comme étant une conjecture qui devrait être démontrée par la suite. On donne l'occasion à l'élève de faire une interprétation graphique mais en réalité, cette interprétation graphique devrait être suivie d'une démonstration, parce que la démonstration c'est rassurante, ça prouve effectivement à l'élève que, voilà ce que j'ai eu à lire graphiquement, les idées que je me suis faites tout à l'heure se sont vérifiées à partir de la démonstration.
7	CH	Ok pour vous, concernant l'élève B, en plus de la résolution graphique, il devrait déterminer l'équation de la droite proprement dite ?
8	P10	Déterminer effectivement l'équation de la droite, mais puisque la question n'était pas posée, je ne sais pas. Est-ce que la question était posée à l'élève ? Oui, je crois ça été posée. Et donc l'élève B s'était contenté seulement du graphique pour répondre à la question, or on sait que cette réponse pouvait être fausse ou bien c'est de l'à peu près. Avec la droite de régression, on devrait s'attendre à une réponse exacte, précise.
9	CH	On va passer à la deuxième partie de de notre
10	P10	Mais aussi la question, excusé moi la question 10. Il est dit que, en cas d'erreur identifiée, et comment pourrez vous aider ces élèves, en cas d'erreur ? Comment les aidés ? C'est de multiplier les exercices, de leur proposer beaucoup plus d'exercices à faire à la maison. Et si on leur propose beaucoup d'exercices à faire à la maison, il faudrait que l'œil soit regardant là dessus. Est-ce que ces exercices sont traités ? Ici chez nous, nous sommes obligés d'exercer une certaine forme de pression sur les élèves afin qu'ils se mettent au travail. L'habitude est donc, on est obligé d'utiliser parfois un bâton comme on le dit, mais le bâton ce n'est pas véritablement le bâton mais c'est d'exercer une forme de pression afin que

		nos élèves puissent se mettre au travail. Les erreurs identifiées se trouvent sur l'annexe et donc les aidés, c'est de leur faire, faire des exercices à la maison. Ces exercices qui seront corrigés de manière, pas tout les exercices mais choisir certains pour corriger afin qu'ils se rendent compte effectivement de leur erreurs et se mettent véritablement au travail.
11	CH	On va passer à la deuxième partie de notre entretien qui concerne vos pratique d'enseignement de la statistique double. 4. Comment introduisiez-vous la série statistique à deux variables en classe ?
12	P10	Je crois que la série statistique à deux variables, c'est un programme chez nous ici, c'est un programme de la première. On le fait déjà en première D, pas en série C, en le fait déjà en première D et on le poursuit en terminale D. Si vous parlez d'introduction, on avait déjà introduit en première mais si on doit l'introduire en terminale D, nous avons les acquis du collège, les acquis de seconde basés sur euh la statistique à une variable. Alors on se dit avec la statistique à une seule variable, certaines choses ne pourraient pas être réaliser, je crois que à mon avis par rapport à ça que les mathématiciens ont vu bon d'inventer cette autre approche qui consiste à étudier deux variables et l'étude de ces deux variables, c'est juste une manière de mettre en lien deux grandeurs alors qu'au collège, jusqu'au niveau de la seconde, on étudie seulement une seule variable qui n'était pas tellement intéressante, alors qu'avec les deux variables, l'objectif c'est de faire des prévisions, on retrouve en économie, on retrouve en agriculture, on le retrouve un peu partout.
13	CH	5. Est-ce que vous pourriez me dire, quels sont les objectifs de l'enseignement de la statistique à deux variables en terminale S (D) ?
14	P10	Comme je le disais tantôt, l'objectif c'est d'amener nos élèves à comprendre un certains nombre de lien qui peut y exister entre deux grandeurs et pouvoir interpréter ces liens là. Après avoir interpréter ces liens, on peut faire des prévisions et je répète ce que j'avais dis, si j'ose le faire répéter, en disant que, on le retrouve en statistique, je dis en statistique en économie, on le retrouve euh, bon dans le système éducatif, on le retrouve euh, en programmation. Les plans quinquennaux ou bien, donc la statistique à également son rôle, c'est un rôle capital.
15	CH	6. D'après vous, quelles sont les compétences en statistique à deux variables qui vous paraissent importantes à construire chez les élèves ?
16	P10	Les objectifs à consolider ce sont la maîtrise des formules euh, l'interprétation des résultats qui conduisent certainement à des prévisions, euh je crois que ce sont ces deux objectifs que je pense retenir par rapport à l'objectif de la statistique à deux variables en terminale.
17	CH	Merci, 7. Est-ce que vous pourriez m'expliquer comment vous vous organisez avec vos élèves pendant votre cours ? Quand vous faites un cours, comment vous procédez avec vos élèves ?
18	P10	Alors pour ce qui me concerne, comme je le disais tantôt euh, le système éducatif Centrafricain, il est différent du système éducatif Français par exemple. Nous enseignants, nous avons l'habitude de, de tout donner aux élèves, lorsqu'on donne tout aux élèves parfois, finalement on attend rien de ces élèves et donc il est important, c'est ce que moi même j'ai l'habitude de le faire, de

		commencer une séquence ou un chapitre à travers des activités, des activités contrôlées bien évidemment, contrôlées ou chacun doit se mettre au travail afin que les élèves eux même qui puissent découvrir, découvrir ce dont en parle, les formules qui apparaissent, l'interprétation graphique, si c'est, on va représenter le nuage des points euh aussi l'interprétation en terme de problème, si c'est un problème d'interprétation, parce que au final, l'objectif de statistique, c'est l'interprétation, c'est-à-dire la prévision.
19	CH	Et, 8. Est-ce que vous faites les travaux dirigés à la fin de chaque chapitre ?
20	P10	A la fin de chaque chapitre, je propose des exercices. Vous savez qu'un enseignant ici, dans notre pays, vous le savez, nous avons des difficultés de produire des supports et moi souvent, je propose des exercices sur une feuille frappe, une ou deux feuilles de frappe que je mets à la disposition des élèves qui eux même vont photocopier afin que, après le cours un maximum de ces exercices soient traités et corrigés en classe. Les exercices ne doivent pas être corrigés seulement par l'enseignant. Les élèves doivent participer à la correction de ces exercices et le plus souvent, c'est les élèves qui passent au tableau, si y a des difficultés moi même personnellement j'interviens pour décanter la situation et on évolue ainsi de suite.
21	CH	10. Est-ce que vous leur donniez des exercices à faire d'abord à la maison avant de corriger en classe ?
22	P10	C'est ce que je disais tantôt, les exercices sont proposés sous forme de polycopie, c'est proposé sur une ou deux feuilles de frappe pour les enfants euh s'en appropriés, ils photocopient. Je leur demande souvent de commencer par réfléchir sur certains de ces exercices en les énumérant, commencer par tel exercice exercice, tel autre suivant les difficultés de ces exercices, suivant les difficultés de ces exercices et c'est après qu'on se retrouve en classe pour traiter ces exercices. Un cours sans travaux dirigés n'est pas un cours, n'est pas consistant, c'est-à-dire on aura fait tout simplement de cours mais les enfants ne savent pas à quoi sert finalement le cours, donc moi je sais que, un cours doit être suivi d'exercice d'application direct, d'exercice approfondis et aussi pourquoi pas de problème. C'est ça l'objectif, amené nos élèves à réfléchir et à prendre des initiatives.
23	CH	D'accord et, 11. De quelle manière vous évaluer les élèves ?
24	P10	L'évaluation des élèves se fait souvent à travers les enseignements à travers les chapitres, à travers les séquences. J'ai l'habitude pour moi après avoir épuisé deux ou trois chapitres, j'évalue, j'évalue par rapport à deux ou trois chapitres. Mais entre temps, les travaux dirigés ont été suffisamment fait en classe, donc je les évalue après avoir épuisé deux ou trois chapitres ainsi de suite. Mais je sais que pendant les compositions du premier, deuxième ou troisième trimestre, certaines choses qu'on a déjà vu peuvent revenir pendant les compositions.
25	CH	D'accord, 12. D'après vous, quelles sont les difficultés des élèves en statistique à deux variables ?
26	P10	Les difficultés des élèves ici chez nous, c'est la non maîtrise des formules euh, ils ont des difficultés, des difficultés par rapport à la compréhension de la langue. Vous savez que sincèrement dit, nous Centrafricain, nous avons quelques lacunes par rapport à certains de nos frères d'autres pays, nous avons des difficultés des lacunes euh par rapport à la langue. D'où il est très important

		<p>que le système éducatif met l'accent déjà au niveau de l'école primaire sur la lecture, c'est à travers ça que ces lacunes pourraient être endiguées. Et donc, les difficultés comme je le disais tantôt, difficulté de maîtrise des formules, difficulté par rapport à la compréhension de la langue, et aussi difficulté d'outils. Quand je parle d'outils, je pense à la calculatrice, la calculatrice scientifique, la calculatrice programmable. Vous allez vous retrouver avec un problème de statistique qui utilise de grand nombre mais quand l'enfant, l'élève se lance dans ces calculs, ça le lince automatiquement, très rapidement il est lincé, il se dit mais ça sert à rien de continuer à calculer ces grands nombres là. Mais y avait une calculatrice qui était déjà programmée, enseigné aux élèves que, voilà ça doit se faire de cette manière, je pense que ça pourrait être un outil de motivation, un outil de motivation, donc voilà les quelques difficultés et je crois même aussi que beaucoup d'enseignants de mathématiques ne savent pas utiliser les calculatrices programmables. Moi même je le dis sincèrement j'ai des difficultés par rapport à cet outil de travail, alors que ça ne devrait pas être une difficulté, c'est une machine qui coûte un peu cher voilà.</p>
27	CH	<p>Ok merci, on va passer à la troisième partie de notre entretien qui concerne les initiatives des élèves en statistique à deux variables.</p> <p>13. Selon vous, est-ce que dans vos cours habituels de statistiques à deux variables les élèves doivent prendre des initiatives. Bon, j'attends par initiatives des élèves par exemple, le choix des méthodes pour faire la prévision, ou faire l'estimation, faire l'ajustement et la représentation graphique, c'est tout ces trucs la que je considère comme initiative de la part des élèves.</p>
28	P10	<p>Effectivement dans notre questionnaire que j'ai eu à analyser par rapport aux élève A et l'élève B comme je le disais tantôt, il a rapidement construit le nuage de points et donc par lecture graphique il a répondu à la question, ça, là c'est aussi une initiative. Mais je disais tantôt que cette initiative là devrait être confirmé, parce que c'est sous forme de conjecture ça devrait être confirmé par un résultat à partir des calculs. L'élève A aussi il avait pris des initiatives, chaque élève à toujours des initiatives à prendre, tout les élèves, ils ont des initiatives à prendre. Donc voilà la réponse que je peux donner à cette question.</p>
29	CH	<p>Ok, si je vous comprends bien, il faudrait tenter de développer ces initiatives des élèves ?</p> <p>Selon vous, est-ce que vous pensez qu'il faudrait tenter de développer ces initiatives ?</p>
30	P10	<p>Évidement , ça fait même partir de l'objectif de l'étude de statistique à deux variables, être capable d'interpréter les résultats, mais interpréter un résultats, c'est prendre des initiatives, réfléchir et voir même se poser des questions et tirer une conclusion.</p>
31	CH	<p>Ok, bon surtout dans l'enseignement de la statistique à deux variables, sur quelle notion ou sur quelle partie, on se rend compte que y a des prises d'initiative de la part de l'élève ?</p> <p>Est-ce que c'est vraiment pendant les calculs ou bien pendant les interprétations ou bien sur quelle partie, on rencontre souvent ce genre de prise d'initiatives des élèves ?</p>
32	P10	<p>D'après mon expérience, je crois que nos élèves veulent souvent calculer, calculer ceci, calculer, calculer cela, calculer ceci, calculer cela, mais pour</p>

		<p>repandre à la question, à une question par rapport à un problème posé, c'est ça la difficulté parce que, comme je le disais tantôt les élèves ont un problème de la compréhension de la langue et quant à la question posée par rapport à un problème et qu'ils n'arrivent pas à cerner la question et parfois la réponse est donnée comme ça et parfois la réponse est erronée. Les initiatives on les retrouve souvent dans les calculs, je crois que sur ce point là, il n'a pas tellement des difficultés donc l'accent devrait être mit sur l'approche interprétation des résultats, parce que c'est l'objectif final. Je disais même l'objectif principal de l'étude statistique à deux variables. Vous allez voir que un peu partout on voit la statistique, la construction même des histogramme, un élève il est capable de quantifier à partir des statistique ses résultats, ses résultats en mathématiques, ses résultats en français, en anglais et en histoire géographie sur son portable. Si y a n'a et voir l'allure de la courbe statistique de son travail, si ça là, ça devrait être, si ça pouvait être inculquer à nos enfants ça serait une bonne chose, dommage que nous même, est-ce que nous avons l'habitude de faire ce travail ? Je ne crois pas.</p>
33	CH	Ok d'accord, on est à la fin de notre entretien, je vous remercie.
34	P10	C'est moi qui vous remercie beaucoup.

D PREMIÈRE EXPÉRIMENTATION EN CLASSE

Énoncé de la première expérimentation en classe

Exercice

Le tableau suivant indique en euros les montants des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense qu'en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette prévision ? Vous expliquerez en détail la méthode que vous avez choisie. Puis justifier pourquoi un ajustement affine est correct (ou approprié).

Bon travail

Copie des élèves du groupe 5

Exercice

Groupe n° 5.

u_i	1	2	3	4	5	6	21
y_i	3568	4561	5502	6693	8241	9120	37685
u_i^2	1	4	9	16	25	36	91
y_i^2	12730624	20802721	30272004	44796249	67914081	83174400	259690079
$u_i y_i$	3568	9122	16506	26772	41205	54720	151893

$n = 6$

Calculons \bar{u}_i et \bar{y}_i

$$\bar{u}_i = \frac{1}{n} \sum_{i=1}^n u_i$$

$$\bar{u}_i = \frac{1}{6} (21) = 3,5$$

$\bar{u}_i = 3,5$

$$\bar{y}_i = \frac{1}{n} \sum_{i=1}^n y_i$$

$$= \frac{1}{6} (37685) = 6280,83$$

$\bar{y}_i = 6280,83$

$G(3,5 ; 6280,83)$

Calculons le coefficient de corrélation linéaire.

$$r = \frac{\text{cov}(u_i, y_i)}{\sqrt{u_i} \sqrt{y_i}}$$

$$\begin{aligned} \text{or } \text{Cov}(u_i y_i) &= \frac{1}{n} \sum_{i=1}^n u_i y_i - \bar{u} \bar{y}_i \\ &= \frac{1}{6} [151893] - (3,5 \times 6280,83) \\ &= \frac{1}{6} (151893 - 21982,905) \end{aligned}$$

$$\text{Cov}(u_i y_i) = 25135,5 - 21982,905$$

$$\boxed{\text{Cov}(u_i y_i) = 3332,6}$$

Calcul de la variance

$$\begin{aligned} V_{u_i} &= \frac{1}{n} \sum_{i=1}^n u_i^2 - \bar{u}^2 \\ &= \frac{1}{6} (91 - 12,25) \\ &= 15,16 - 12,25 \end{aligned}$$

$$\boxed{V_{u_i} = 2,91}$$

$$\boxed{V_{u_i} = 1,97}$$

$$\begin{aligned} V_{y_i} &= \frac{1}{n} \sum_{i=1}^n y_i^2 - \bar{y}_i^2 \\ &= \frac{1}{6} (259690079) - (6280,83)^2 \end{aligned}$$

$$V_{y_i} = (43281679,83 - 39449825,19)$$

$$V_{y_i} = \sqrt{3832854,34}$$

$$V_{y_i} = 1957,76$$

or

$$r = \frac{\text{Cov}(x_i, y_i)}{V_{x_i} V_{y_i}} = \frac{3332,6}{1,7 \times 1957,76} = \frac{3332,6}{3328,19}$$

$$r = 1,001$$

~~Après~~ D'après le résultat obtenue avec $r \approx 1$
alors on peut conclure l'ajustement affine est vérifié

Gp : 5

$$y = ax + b \Leftrightarrow y = 10.000x + b$$

Après la formule

$$a = \frac{\text{Cov}(x_i, y_i)}{y_{ii}}$$

$$\Leftrightarrow a = \frac{3332,6}{2,91}$$

$$b = \bar{y} - a\bar{x}$$

$$b = 6280,83 - 1145,22 \times 3,5$$

$$a = 1145,22$$

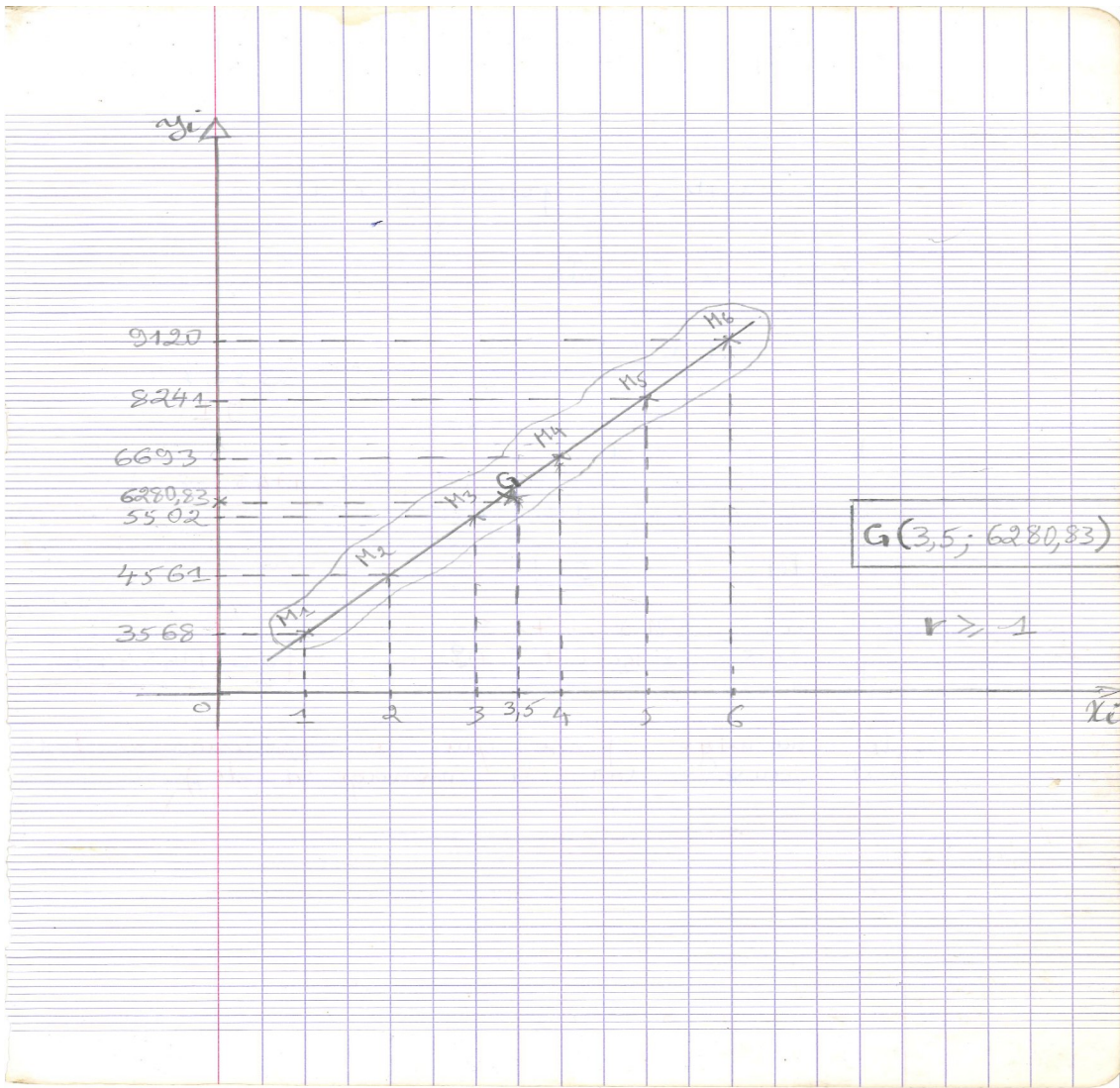
$$b = 2272,5$$

$$y = 1145,22x + 2272,5$$

$$y = 1145,22 \times 10000 + 2272,5$$

$$y = 11454472,5$$

le ménage pense que la prévision en 2019 est fautive car le ménage a dépense plus



Copie des élèves du groupe 6

Jeudi le 25 avril 2019

Groupe 6

Travaux dirigés

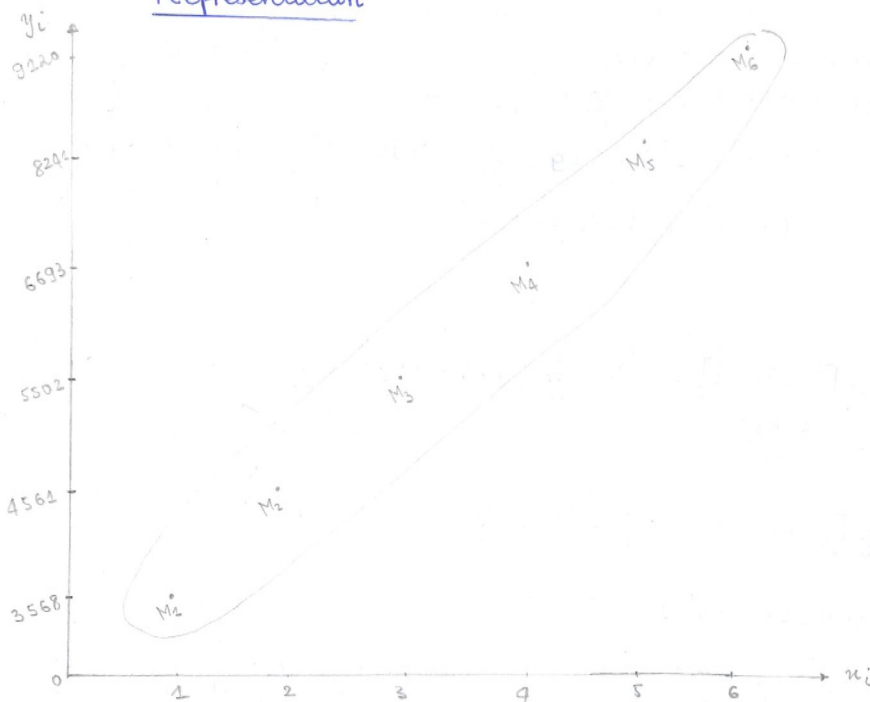
Exercice

Année	2013	2014	2015	2016	2017	2018
Rang de l'annexi	1	2	3	4	5	6
Montant de remboursements y_i	3 568	4 561	5 502	6 693	8 241	9 120

- Désignons par : M_1, M_2, M_3, M_4, M_5 et M_6 l'ensemble des points de nuage :

on a : $M_1 (1; 3568)$; $M_2 (2; 4561)$; $M_3 (3; 5502)$; $M_4 (4; 6693)$; $M_5 (5; 8241)$
et $M_6 (6; 9120)$

Représentation



- Un ajustement de nuage de points permet de faire cette prévision.

Calculons \bar{x} et \bar{y}

on sait que: $\frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$

AN: $\bar{x} = \frac{1}{6} (21) = \frac{21}{6} = \bar{x}$

$\frac{1}{n} \sum_{i=1}^n y_i = \bar{y}$

AN: $\bar{y} = \frac{1}{6} (37685) = \frac{37685}{6} = \bar{y}$

$\bar{x} = \frac{21}{6}$ et $\bar{y} = \frac{37685}{6}$

Complétons le tableau

Année	2013	2014	2015	2016	2017	2018	Totaux
Rang de l'année x_i	1	2	3	4	5	6	21
Montant de remboursement y_i	3568	4561	5502	6693	8241	9120	37685
							791385 $x_i y_i$

Calculons la Cov (X, Y)

on sait que: $Cov(x, y) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}$

AN: $Cov(x, y) = \frac{1}{6} [(791385) - \frac{791385}{36}] = \frac{276998475}{216}$

$Cov(x, y) = 1281233,6$

Variance:

$\sigma_x^2 = \frac{1}{6} [(21)^2 - (\frac{21}{6})^2] = \frac{1}{6} (441 - \frac{441}{36}) = \frac{1}{6} (\frac{15435}{36}) = \frac{15435}{216}$

$\sigma_x^2 = \frac{15435}{216}$

$\sigma_y^2 = \frac{1}{6} [(37685)^2 - (\frac{37685}{6})^2] = \frac{1}{6}$

$\sigma_y^2 = 230148392,9$

- Appliquons la méthode des moindres carrés qui permettront de déterminer les droites de regression de (\hat{y}) et (\hat{x})
 on a: $y = ax + b$ $\left\{ \begin{array}{l} a = \frac{Cov(x, y)}{\sigma_x^2} \\ b = \end{array} \right.$

Copie des élèves du groupe 7

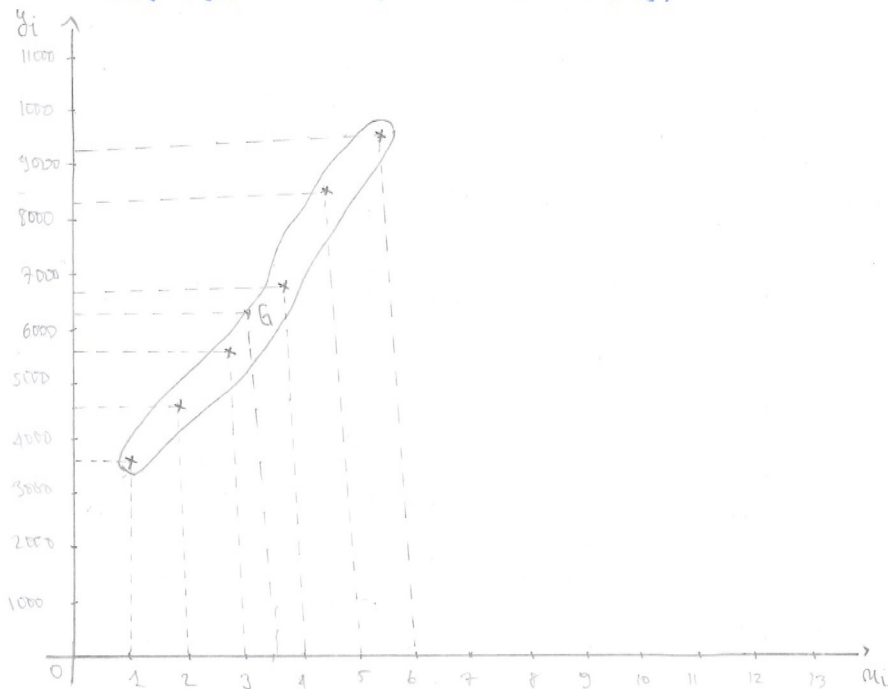
Groupe N° 7

EXERCICE DE STATISTIQUE

Par la méthode de Moindres Carrées

							Σ
Année	2013	2014	2015	2016	2017	2018	X
n_i	1	2	3	4	5	6	21
y_i	3568	4561	5502	6693	8242	9120	37685
n_i^2	1	4	9	16	25	36	91
y_i^2	12730624	20803721	30272004	44796249	6794064	83177400	259690073
$n_i y_i$	3568	9122	16506	26772	41205	54720	152893

1. Représentons le nuage de points: Echelle: 1cm \rightarrow 1année sur (0,4)
 et (n_i, y_i) dans le repère orthonormé (O, i, j) 1cm \rightarrow 10.000 \$ sur (0, y)



Calculons les coordonnées du point G (\bar{n}_i, \bar{y}_i)

$$\bar{n}_i = \frac{\Sigma n_i}{N} \Rightarrow \bar{n}_i = \frac{21}{6} \Rightarrow \bar{n}_i = 3,5$$

$$\bar{y}_i = \frac{1}{N} (\Sigma y_i) = \frac{37685}{6} \Rightarrow \bar{y}_i \approx 6,28 \times 10^3 \text{ près} \quad \text{donc } \boxed{G(3,5 ; 6,28)}$$

Déterminons la droite de régression de y en x .

$$(a) : y = a(x - \bar{x}) + \bar{y} \quad \text{or} \quad a = \frac{\text{Cov}(x, y)}{\sigma_x^2}$$

$$\text{Cov}(x, y) = \overline{xy} - \bar{x} \times \bar{y} = \frac{152893}{6} - (3,5 \times 6,28)$$

$$\text{Cov}(x, y) = 25293,52$$

$$\sigma_x^2 = \overline{x^2} - (\bar{x})^2 = \frac{94}{6} - (3,5)^2 = 2,92$$

$$a = \frac{25293,52}{2,92} = 8662,16 \quad a = 8662,16$$

$$y = 8662,16(x - 3,5) + 6,28 \\ = 8662,16x - 8662,16 \times 3,5 + 6,28$$

$$(b) : y = 8662,16x - 30311,28$$

Calculons la dépense du ménage en 2019

$$\text{En 2019, } x = 7 \quad \text{donc } y = 8662,16(7) - 30311,28$$

$$\text{d'où } y = 30323,84$$

Conclusion: la prévision qui prévoyait le ménage en 2019 n'est pas vérifiable car en 2019 il devrait dépenser 30.323,84 Euros.

Calculons la corrélation

$$\text{On sait que: } r = \frac{\text{Cov}(x, y)}{\sigma_x \times \sigma_y}$$

Or d'après le graphique on voit que les nuages de points sont assimilables à une droite donc il peut passer une droite la plus près possible des points de nuage d'où un ajustement affine est approprié.

Copie des élèves du groupe 10

Jeudi, le 25 Avril 2019

Groupe N° 10

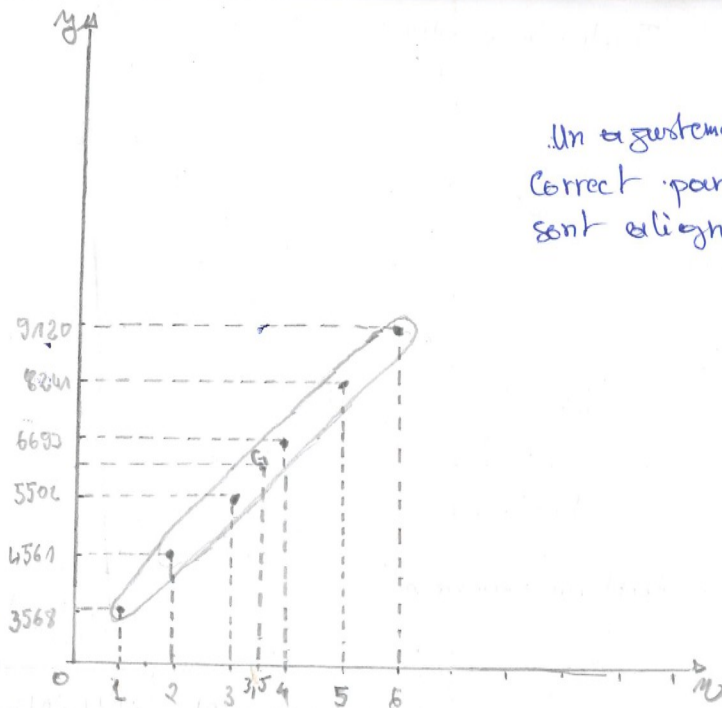
Exercice

On a le tableau suivant :

	Σ						
Année	2013	2014	2015	2016	2017	2018	
Rang de l'année x_i	1	2	3	4	5	6	21
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120	37685
x_i	1	4	9	16	25	36	91
y_i	12730	20802	30272	44796	67914	83174	256631
x_i, y_i	3568	9122	16506	26772	41205	54720	151858

* La courbe

Un ajustement affine est correct car ce que les points sont alignés.



La méthode que nous avons choisie c'est un ajustement affine par la méthode des moindres carrés.

* Calculons \bar{x}

$$\bar{x} = \frac{1}{n} \cdot \sum x_i \Leftrightarrow \frac{1}{6} \cdot (21) \Leftrightarrow \bar{x} = \frac{21}{6} = 3,5$$

$$\boxed{\bar{x} = 3,5}$$

$$\bar{y} = \frac{1}{n} \cdot \sum y_i \Leftrightarrow \frac{1}{6} \cdot (37685) \Leftrightarrow \bar{y} = \frac{37685}{6} = 6280$$

$$\boxed{\bar{y} = 6280}$$

* Calculons $\text{Cov}(x, y)$

$$\text{Cov} = \frac{\sum x_i \cdot y_i}{n} - \bar{x} \bar{y}$$

$$\text{Cov} = \frac{21 \times 37685}{6} - (3,5 \times 6280)$$

$$= (131 \cdot 857 - 21980)$$

$$= 131897 - 21980 = 109.917$$

$$\boxed{\text{Cov}(x, y) = 109.917}$$

Copie des élèves du groupe 12

Jeudi 25 Avril 2019

G.R. (12)

Solution d'exercice

							Total
Année	2013	2014	2015	2016	2017	2018	6 ans
Rang de l'année x_i	1	2	3	4	5	6	21
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120	37685
x_i^2	1	4	9	16	25	36	91
y_i^2	12730624	20702721	30272004	44782249	67914021	83174400	259690079
$x_i \cdot y_i$	3568	9122	16506	26772	41205	54720	151883

Appliquons la relation de CORRELATION linéaire

D'après la formule de Corrélation linéaire :

$$r = \frac{\text{Cov}(x, y)}{\sqrt{V(x) \cdot V(y)}} \quad , \quad V(x) = \frac{1}{6} \sum_{i=1}^6 x_i^2 - \bar{x}^2$$

$$= \frac{1}{6} (91) - (3,5)^2$$

$$= 15,16 - 12,25$$

$$V(x) = 2,91$$

$$V(y) = \frac{1}{6} \sum_{i=1}^6 y_i^2 - \bar{y}^2$$

$$= \frac{1}{6} (259690079) - 37685^2$$

$$= 43281679,8 - 37685^2$$

$$V(y) = 43243994,8$$

$$\begin{aligned} & \text{Cov}(x, y) \\ &= \frac{1}{6} \sum_{i=1}^6 x_i y_i - \bar{x} \bar{y} \\ &= \frac{1}{6} (151893) - 3,5 (6280,8) \\ &= 25315,5 - 21982,8 \end{aligned}$$

$$= 3332,7$$

$$\text{Donc } r = \frac{\text{Cov}(x, y)}{\sqrt{v(x) \cdot v(y)}} = \frac{3332,7}{\sqrt{(2,91)(3332,7)}} = \frac{3332,7}{\sqrt{9698,157}}$$

$$= \frac{3332,7}{98,45} = 33,95$$

$$r = 33,95$$

Justifions donc la Corrélation n'est pas affine car elle n'est pas proche de 1.

* Nous réalisons que la dépense de 2019 est :

$$37685 - 10000 = 27685$$

E. DEUXIÈME EXPÉRIMENTATION EN CLASSE

Énoncé de la deuxième expérimentation en classe

Exercice 1

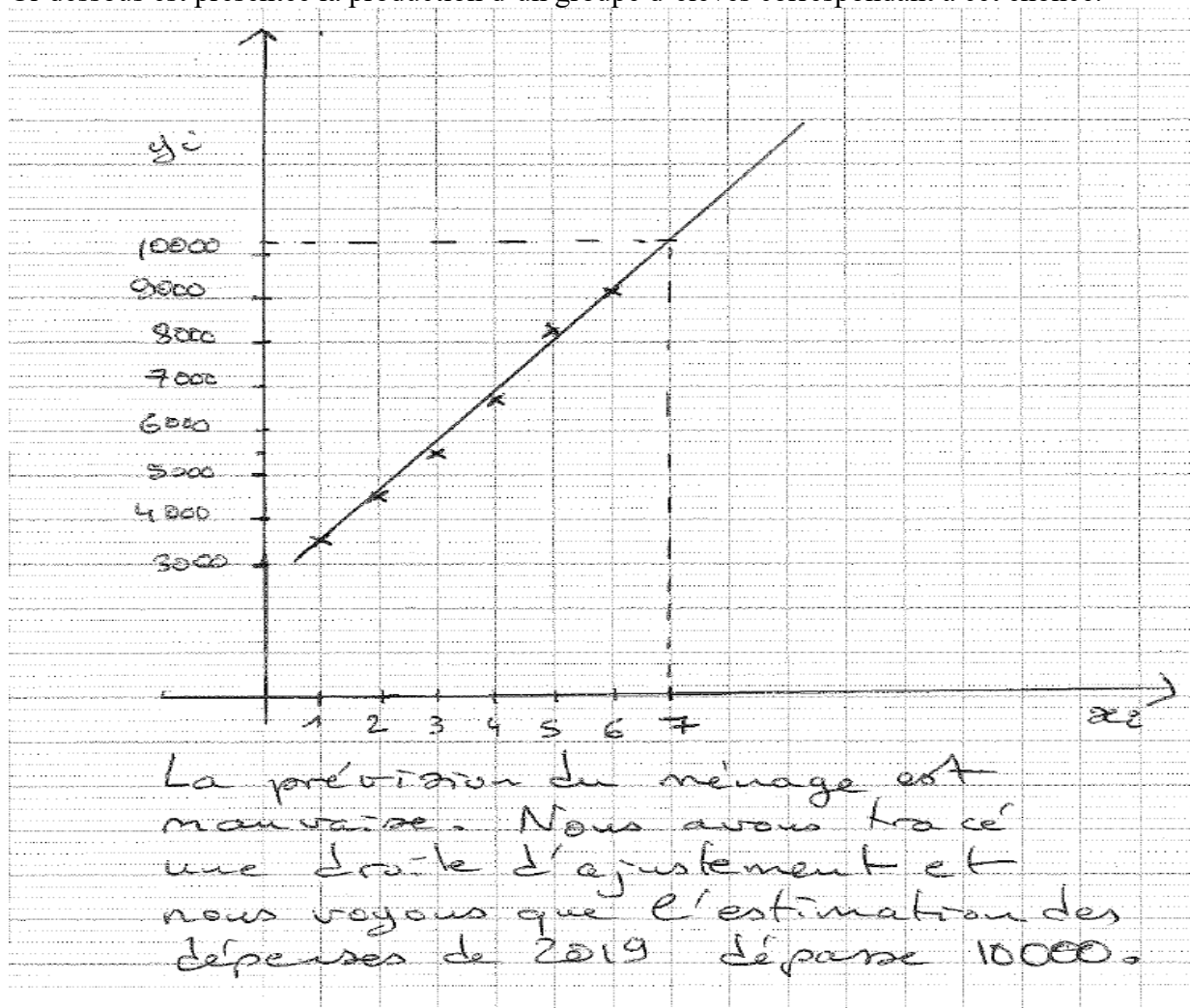
L'exercice suivant a été proposé à des élèves de Terminale scientifique en Centrafrique.

Le tableau ci-dessous indique le montant des remboursements annuels y_i effectués de 2013 à 2018 par un ménage ayant contracté plusieurs emprunts.

Année	2013	2014	2015	2016	2017	2018
Rang de l'année x_i	1	2	3	4	5	6
Montant des remboursements y_i	3568	4561	5502	6693	8241	9120

Le ménage pense que en 2019, il dépensera moins de 10000 euros. Que pensez-vous de cette affirmation ?

Ci-dessous est présentée la production d'un groupe d'élèves correspondant à cet énoncé.



Est-ce que vous jugez cette réponse juste? Justifier.

Exercice2

Le tableau suivant représente des couples de notes en Mathématiques et en Comptabilité de 100 élèves que l'on a regroupés lorsqu'ils étaient égaux.

Note xi en mathématiques	7	8	8	9	9	9	9	10	10	11	12	14	15
Note yi en comptabilité	6	7	9	8	10	11	12	8	11	12	10	13	14
Effectif	5	6	8	10	12	14	8	11	7	9	3	5	2

1. Déterminer les coordonnées du point moyen G du nuage formé par les points $M_i(x_i, y_i)$
2. Calculer la covariance de la série statistique double (X, Y) .
3. Déterminer une équation de la droite de régression de y en x, obtenue par la méthode des moindres carrés.
4. Un élève a obtenu la note de 13 en mathématiques. Peut-on estimer sa note en comptabilité ?

Justifier.

Exercice 3

Le tableau suivant donne, pour douze mois consécutifs, l'évolution des dépenses publicitaires (en milliers d'euros) d'une société commerciale.

Numéro du mois : xi	1	2	3	4	5	6	7	8	9	10	11	12
Montant des dépenses : yi	3	4,5	3,7	5,2	5,2	6	7,5	7,5	8,2	9,7	9,7	10

1. Représenter dans un repère orthogonal le nuage des points M_i de coordonnées (x_i, y_i) correspondant à cette série statistique.
2. Tracer la droite passant par les points A(1 ; 3) et B(9 ; 8,2). Déterminer l'équation réduite de cette droite.
3. On utilise cette droite pour réaliser un ajustement affine du nuage des points M_i .
 - a) Estimer le montant des dépenses durant le quatorzième mois.
 - b) Estimer le rang du mois au cours duquel le montant dépassera pour la première fois 13 milliers d'euros.

Copie des élèves du groupe 1

Mardi, le 31 Août 2020

Groupe 1

Exercice 1

Reponses aux questions

Oui la reponse donnée est juste car le nuage des points constitue une droite donc l'ajustement affine est possible. Or la droite d'ajustement affine permet de faire une prévision. Mais en faisant la prévision, on voit que l'affirmation du ménage n'est pas juste.

Exercice 2

Note x_i en Mathématiques	7	8	8	9	9	9	9	10	10	11	12	14	15	131
Note y_i en comptabilité	6	7	9	8	10	11	12	8	11	12	10	13	14	131
Effectif	5	6	8	10	12	14	8	11	7	9	3	5	2	100
														Σ

1- Déterminons les coordonnées du point moyen G .
 $G(\bar{x}, \bar{y})$: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{100} (131) = 1,31$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{100} (131) = 1,31$$

donc G a pour coordonnées $G(1,31; 1,31)$

2) Calculons la covariance: $Cov(x, y)$

$$\begin{aligned} \text{Cov}(x, y) &= \frac{1}{n} \sum_{i=1}^n (x_i y_i - \bar{x} \bar{y}) \\ &= \frac{1}{100} [(131 \times 131) - (1,31 \times 1,31)] \\ &= \frac{1}{100} [(131)^2 - (1,31)^2] \end{aligned}$$

$$\boxed{\text{Cov}(x, y) = 171,6}$$

3- Déterminons une équation de la droite de régression de y en x .

$$(D): y = ax + b$$

$$\text{Avec } a = \frac{\text{Cov}(x, y)}{\text{Cov} \sigma_x^2} ; \quad b = \bar{y} - a\bar{x}$$

$$\begin{aligned} \sigma_x^2 &= \frac{1}{n} \sum_{i=1}^n (x_i^2 - \bar{x}^2) \\ &= \frac{1}{100} [(131)^2 - (1,31)^2] \end{aligned}$$

$$\boxed{\sigma_x^2 = 171,6}$$

$$\text{d'où } a = \frac{171,6}{171,6} = 1$$

$$b = 1,31 - (1)(1,31) = 0$$

$$\text{Donc } (D): y = (1)x - 0 \Rightarrow \boxed{(D): y = x}$$

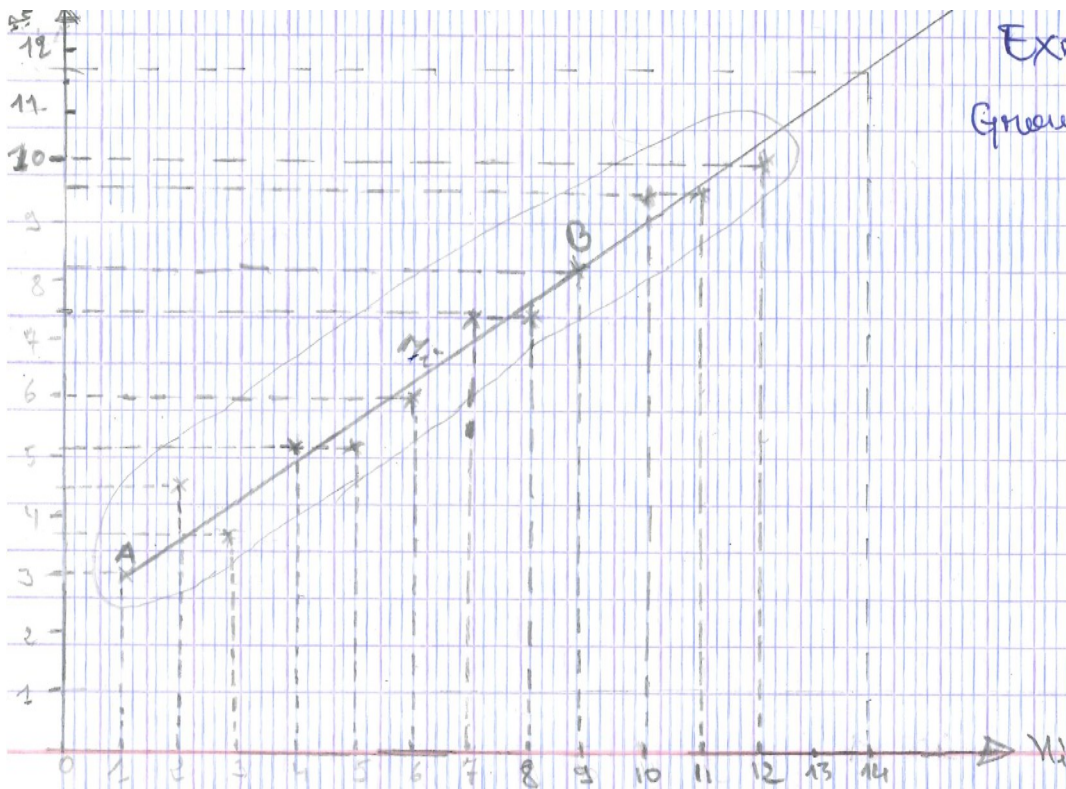
4) Estimons la note de la comptabilité

$x_i = 13$ or on sait que: $y = x$ donc

$$\boxed{y = 13} \quad \text{La note estimée en comptabilité est } y = 13$$

Exercice 3

Groupe 1



a) Trouver la droite par les points A (1; 3) et B (9; 8, 2).
 Déterminer l'équation réduite de cette droite.

a- estimons le montant des dépenses durant le 14^e mois.

le montant estimé au 14^e mois est de 11,7 M.

b)

Copie des élèves du groupe 4

Groupe (4)

Exercice 1

Année	2013	2014	2015	2016	2017	2018	2019
Rang de l'année	1	2	3	4	5	6	7
Montant des remboursements	3568	4561	5502	6693	8241	9120	10000

① L'affirmation du ménage est fautive, car ce que de 2013 à 2018 les montants dépensés s'accroît.

② Premièrement l'ajustement affine est simple. Qui nous apporte cette réponse, car d'après le tracé nous constatons que l'estimation des dépenses est bien belle et bien supérieur à 10000.

Exercice 2

Note x, en mathématiques	7	8	8	9	9	9	9	10	10	11	12	14	15
Note y _i en Comptabilité	6	7	8	8	10	11	12	8	11	12	10	13	14
Effectifs	5	6	8	10	12	14	8	11	7	9	3	5	2

1) Déterminons les coordonnées du point moyen G du nuage formé par les points M_i (x_i, y_i)

• Calcul de \bar{x} et \bar{y}

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n n_i x_i = \frac{\sum_{i=1}^n n_i x_i}{N} = \frac{131}{13} = 10$$

$$\bar{y} = \frac{131}{13} = 10$$

Point G (10, 10)

2) Calculons la covariance de la série statistique

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n n_i x_i y_i - \bar{x} \bar{y}$$

$$\text{AN: cov}(x, y) = \frac{1}{13} (1373 - 100)$$

$$= \frac{1273}{13} = 98$$

$$\text{cov}(x, y) = 98$$

3- déterminons une équation de la droite de régression de y en x

$$(\mathcal{D}) : y = ax + b \text{ d'où}$$

$$a = \frac{\text{cov}(x, y)}{V(x)}$$

$$b = \bar{y} - a\bar{x}$$

$$\text{on cov}(x, y) = 98$$

$$\bar{x} = 10$$

$$\bar{y} = 10$$

$$\text{donc } V(x) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

$$\text{AN: } V(x) = \frac{1}{13} (1367 - 100)$$

$$V(x) = 97,4$$

$$a = \frac{98}{97,4} = 1$$

$$b = 10 - 10 = 0$$

donc $(\mathcal{D}) : y = x$ première bissectrice

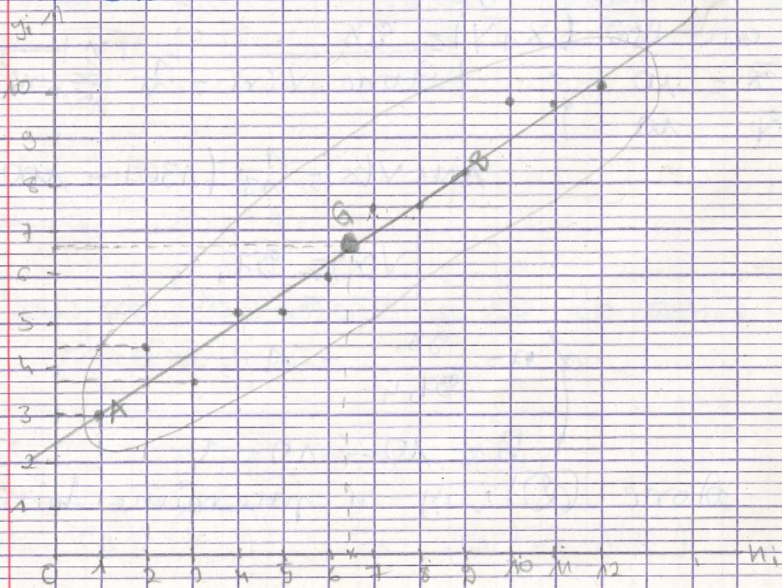
4. Justification

- Qui on peut estimer sa note en comparabilité ~~est~~. Par ce que la droite de régression nous donne; $y = a + b(x)$
 $(x_i = y_i)$

Exercice 3

Numéro du mois	1	2	3	4	5	6	7	8	9	10	11	12
Montant des dépenses	3	4,5	3,7	5,2	5,2	6	7,5	7,5	8,2	9,7	9,7	10

1°) Tracé



Suite exercice 3
Calculons \bar{x} et \bar{y}

$$\bar{x} = \frac{78}{12} = 6,5$$

$$\bar{y} = \frac{80,2}{12} = 6,6$$

G (6,5; 6,6)

2- Trouver la droite passant les points
A(1; 3) et B(9; 8,2)

- déterminons l'équation réduite de cette
droite

* utilisant la méthode de moindres carrés
on a: (D) de régression de y en x

$$a = \frac{\text{Cov}(x, y)}{V(x)}$$

$$b = \bar{y} - a\bar{x}$$

on sait que:

$$\text{Cov}(x, y) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}$$

$$\text{Cov}(x, y) = \frac{1}{12} (615,9 - 42,9) = \frac{573}{12}$$

$$\text{Cov}(x, y) = 47,75$$

$$V(x) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

$$= \frac{1}{12} (630 - 42,25)$$

$$= \frac{587,75}{12}$$

$$V(x) = 49$$

$$\rho = \frac{47,75}{49} = 1$$

$$b = 6,6 - 6,5 = 0,1$$

$$\hat{y} = x + 0,1$$

Copie des élèves du groupe 7

Lundi 31 août 2020

Groupe n° 7

Exercice 2

Le tableau suivant représente -les couples de notes en Mathématiques et en comptabilité de 100 élèves que l'on a regroupés lorsqu'ils étaient égaux

Note en Mathématiques	7	8	8	9	9	9	9	10	10	11	12	14	15	Σ
Note en comptabilité	6	7	9	8	10	11	12	8	11	12	10	13	14	131
Effectif	5	6	8	10	12	14	8	11	7	9	3	5	2	100

1°) Déterminons les coordonnées du point moyen G du nuage formé par les points $M_i(x_i, y_i)$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} = \frac{131}{13} = 10,07$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{N} = \frac{131}{13} = 10,07$$

$M_G = (10,07; 10,07)$

2°) Calculons la covariance double de (x, y)

$$\begin{aligned} \text{Cov}(x, y) &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\ &= \frac{1}{13} (131 - 10,07)(131 - 10,07) \\ &= \frac{(131 - 10,07)(131 - 10,07)}{13} \\ &= \frac{14641}{13} \end{aligned}$$

$\text{Cov}(x, y) = 1126,23$

3°) Déterminons l'équation de la droite

(d); $y = ax + b$ d'où $b = -a\bar{x} + \bar{y}$

Calculons a

$$a = \frac{\text{cov}(x, y)}{V(x)} = \frac{\frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum (x_i - \bar{x})^2}$$
$$= \frac{1126,23}{9,30} = 121,1$$

$$\underline{a = 121,1}$$

En a ; (d) : $y = 121,1x + b$

Calculons b

$$b = -a\bar{x} + \bar{y} = -121,1 \times 10,07 + 10,07$$

$$b = -1209,40$$

d' où : $y = 121,1x + 1209,40$

4) On ne peut estimer sa note en comptabilité.

Parce qu'on ne peut déterminer réellement s'il est autant bon en Mathématique qu'en comptabilité.

Exercice 1

Nous pensons que cette réponse est juste.

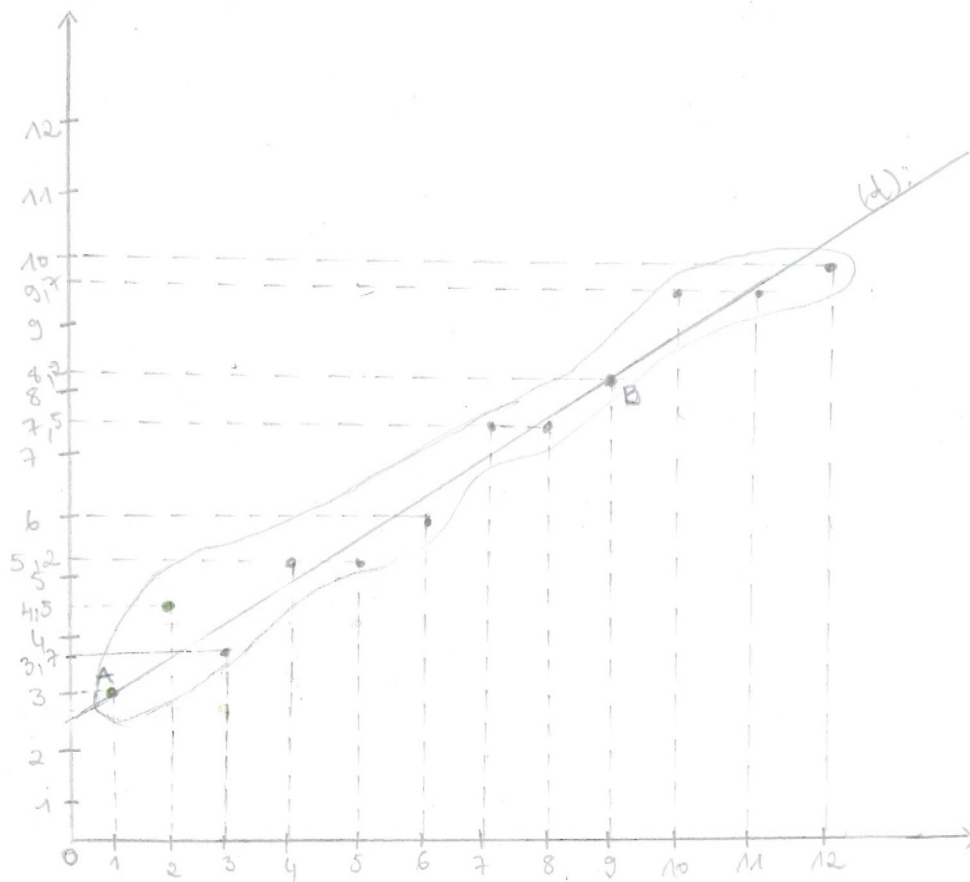
Parce que la gestion de ~~ce~~ ménage est mauvaise, ~~il~~ ne pourra pas gérer ; afin de dépenser moins de 10.000 euros en 2019

Exercice 3

Le tableau suivant donne, pour douze mois consécutifs, l'évolution des dépenses publicitaires (en milliers d'euros) d'une société commerciale.

Numéro du mois n_i	1	2	3	4	5	6	7	8	9	10	11	12
Montant de dépenses: y_i	3	4,5	3,7	5,2	5,2	6	7,5	7,5	8,2	9,7	9,7	10

1°) Représentons dans un repère orthogonal le nuage des points M_i de coordonnées (n_i, y_i) correspondant à cette série statistique.



2°) Traçons la droite passant par les points A(1, 3) et B(9, 8,2).
Déterminons l'équation réduite de cette droite.

$$(d): y - \bar{y} = a(x - \bar{x})$$

Calculons a :

$$a = \frac{\text{Cov}(x, y)}{V(x)} = \frac{\frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum (x_i - \bar{x})^2}$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} = \frac{78}{12} = 6,5$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{N} = \frac{80,2}{12} = 6,68$$

Variance :

$$V(x) = \frac{1}{n} \sum (x_i - \bar{x})^2$$

$$= \frac{(78 - 6,5)^2}{12} = \frac{71,5^2}{12} = 5,95$$

$$\text{Cov} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{71,5 \times 73,52}{12} = \frac{5256,68}{12}$$

$$\text{Cov} = 438,05$$

$$a = \frac{438,05}{5,95} = 73,62$$

L'équation réduite est : $y - \bar{y} = 73,62x - 6,5$

F. TROISIÈME EXPÉRIMENTATION EN CLASSE

Concernant la troisième expérimentation, nous présentons un exemplaire de copie de pré-test et post-test des élèves de la classe 1 et 2. Notre choix porte sur les copies dont les extraits figurent dans le texte.

Copie des élèves du groupe 1 de la classe 1 : pré-test

Travail de groupe N° 1

Données : 9, 25, 64, 121, 146, 225, 256

exercice

M_i	3	5	8	11	14	15	16
m_i	2	3	5	5	2	2	1

1) Calculons la moyenne et l'écart type de la note

on sait que
$$\bar{M} = \frac{1}{N} \sum_{i=1}^n M_i$$

$M_i = 20$
 $3 \times 2 + 5 \times 3 + 8 \times 5 + 11 \times 5 + 14 \times 2 + 15 \times 2 + 16 \times 1$
 $190 \div 20 =$

$$\bar{M} = 9,5$$

Calculons la variance

$1/20 (2 \times 3^2 + 3 \times 5^2 + 5 \times 8^2 + 5 \times 11^2 + 2 \times 14^2 + 2 \times 15^2 + 1 \times 16^2)$
 $= 105,8$
 $105,8 - 9,5^2 = 15,55$

$$|V_{Fx}| = 15,55$$

écart-type = $\sqrt{15,55}$
 $= 3,94$

correcteur B

M_i	1	2	3	7	11	13	14
n_i	1	3	5	4	2	2	2

1) Calculons \bar{M}
 on sait que
$$\bar{M} = \frac{1}{N} \sum_{i=1}^n M_i$$

$\bar{M} = 20$
 $\frac{1}{20} (1 \times 1 + 3 \times 2 + 3 \times 5 + 4 \times 7 + 2 \times 11 + 2 \times 13 + 2 \times 14)$

$$\bar{m} = 7$$

Calculons la variance

$$V(x) = 71,1 - 49$$

$$V(x) = 22,1$$

Calculons l'écart type

$$\sqrt{V(x)}$$
$$\Rightarrow \sqrt{22,1} = 4,70$$

~~1) Nous constatons que la moyenne du correcteur A est inférieure à celle du correcteur B.~~

2) on constate que $\sqrt{V(x)_A} < \sqrt{V(x)_B}$

donc les élèves du correcteur A ont plus de moyenne que ~~seul~~ celui du correcteur B. car $A > B$.

3) Il faut juste multiplier les notes du correcteur B pour que les notes se rapprochent donc $\sqrt{V(x)_A} = \sqrt{V(x)_B}$

n	FA	FB	FC	FD	FE	FF	FG
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7

$$\bar{m} = \frac{\sum_{i=1}^n x_i}{n} = \bar{m}$$

Trame de cours implémenté dans la classe 1

Étapes	Titre et contenu des étapes	Initiatives des élèves	Durée
1	Rappels - Moyenne arithmétique et exemple 1 - Moyenne pondérée et exemple 2	- Proposition de définitions de la moyenne arithmétique et la moyenne pondérée. - Résolution des exemple 1 et 2.	25 minutes
2	Introduction de la série statistique double et Exercice 1	Exprimer leur avis : est-ce qu'on peut estimer la note de physique en fonction de la note de mathématique ?	10 minutes
3	Découverte de la notion de nuage de points, point moyen du nuage suivi d'un exemple.	- Contrôler de l'ordre notion de grandeur. - Faire une interprétation graphique.	30 minutes
4	Notion de variance, covariance et coefficient de corrélation linéaire et des exemples	- Contrôle de l'ordre de grandeur.	35 minutes
5	Découverte de la notion de l'ajustement par la règle et exemple	Faire une représentation graphique.	20 minutes
6	Découverte de la notion de l'ajustement par la méthode de Mayer et exemple	Résolution de l'exemple.	45 minutes
7	Découverte de la notion de l'ajustement par la méthode des moindres carrés et exemple	Résolution de l'exemple.	45 minutes

Copie des élèves du groupe 4 de la classe 1 : post-test

Groupe 4

GROUPE 4

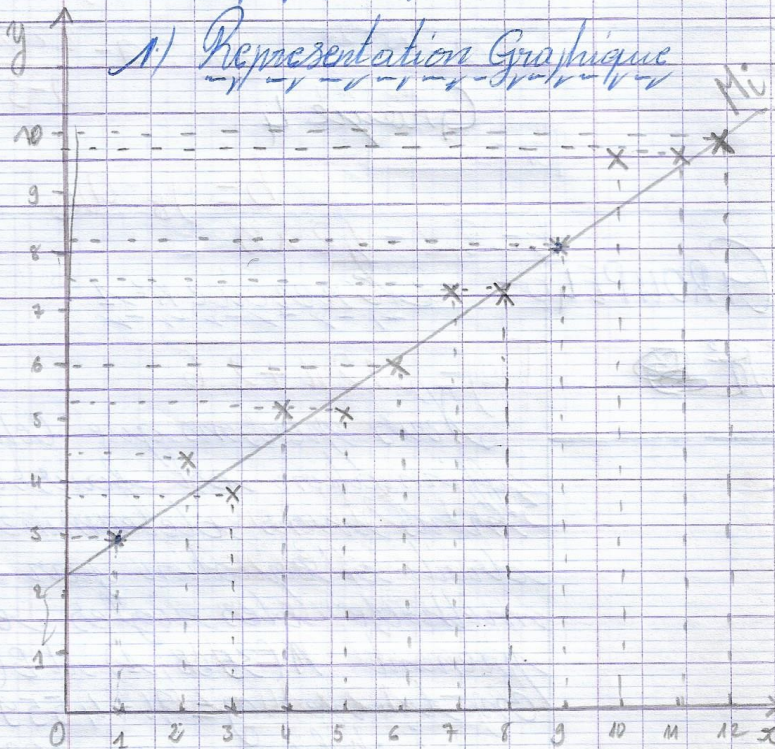
N° 9

Exercice n° 1

Nous pensons que cette affirmation est fautive, car nous avons déterminé la droite d'équation par la méthode des moindres carrés en prenant $A(3568; 1)$ et $B(4561; 2)$. On trouve $y = x + 4559$. Ce qui justifie qu'en 2021, les dépenses dépassent certainement 10.000 milliers de Francs CFA.

Exercice n°2

1) Représentation Graphique



2) Déterminons l'équation réduite de cette droite:
En utilisant la méthode des règles on a:

$$A(1,3), B(9,8,2) \Rightarrow y = ax + b$$

$$\text{et } a = \frac{y_B - y_A}{x_B - x_A} \quad \text{A.N.: } a = \frac{8,2 - 1}{9 - 3} = \frac{7,2}{6}$$

$$a = 1,2$$

$$b = y_B - ax_B$$

$$b = 8,2 - 1,2 \times 9$$

$$b = -2,6$$

$$y = 1,2x + 2,6$$

3-a.) Estimons le montant des dépenses durant le quatorzième mois:

Présumant $x = 14$ on a:

$$y = 1,2 \times 14 + 2,6$$

$$y = 19,2$$

donc le montant de dépense au quatorzième mois est 19,2.

Copie des élèves du groupe 1 de la classe 2 : post-test

Exercice 1

G1

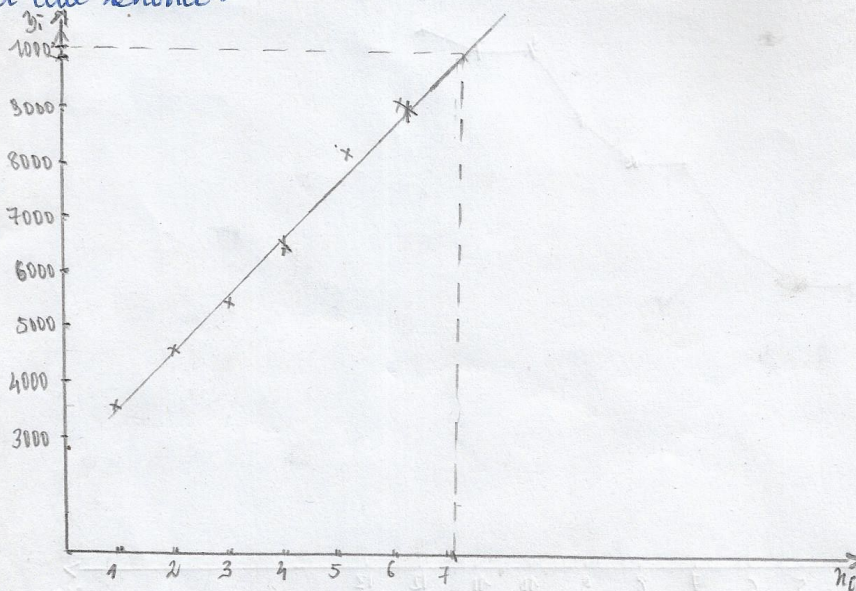
Le tableau ci-dessous indique le montant des remboursements y_i annuels effectués de 2015 à 2020 par un ménage ayant contracté plusieurs emprunts.

Année	2015	2016	2017	2018	2019	2020
Rang de l'année n_i	1	2	3	4	5	6
Montant des remboursements y_i (en milliers de francs CFA)	3568	4564	5502	6683	8241	9120

Le ménage pense que en 2021, il dépensera moins de 10.000 milliers de francs CFA. Que pensez-vous de cette affirmation?

Selon la différence du montant des remboursements y_i qui croît de plus en plus chaque année, alors nous pensons que cette affirmation est fautive et non vérifiée et que le ménage pourra bien dépenser plus de 10.000 milliers de francs CFA en 2021.

Ci-dessous est présentée la production d'un groupe d'élève correspondant à cette énoncé.



La prévision du ménage est mauvais. Nous avons tracé une droite d'ajustement et nous voyons que l'estimé des dépenses de 2021. (1)

depassera 10.000 milliers de francs CFA.

Est-ce que vous jugez cette réponse juste? Justifier.

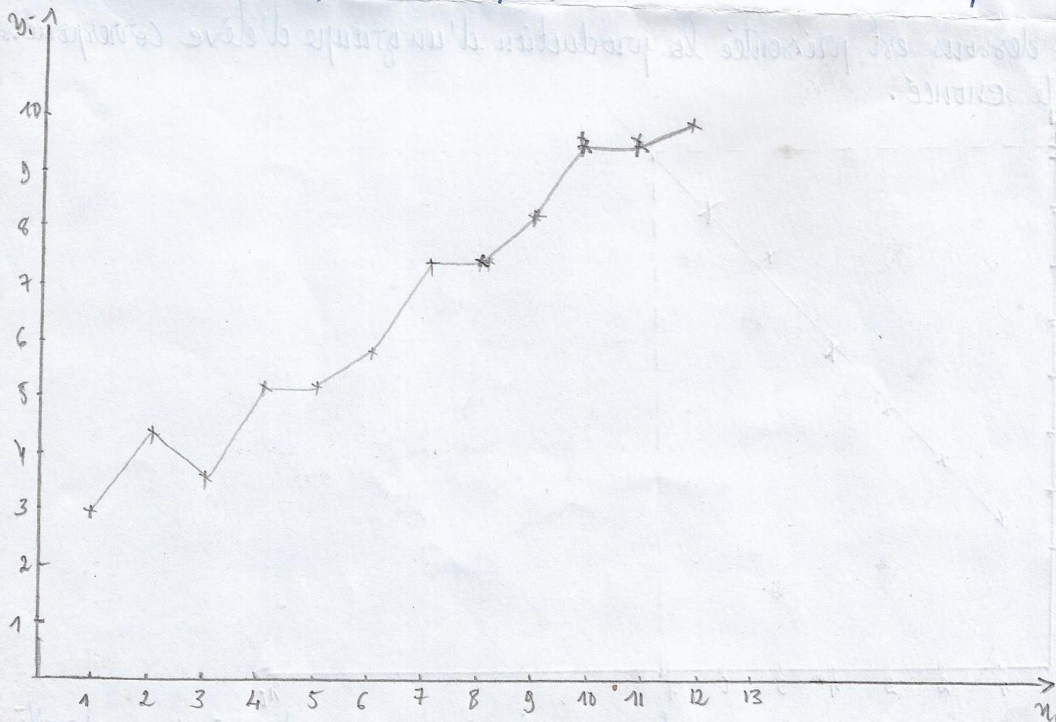
Par rapport à la représentation de la droite d'ajustement affine du nuage, nous constatons bien et bien que les dépenses de remboursements annuels de 2021 dépassent 10.000 milliers de francs CFA, alors oui cette réponse est juste vu la représentation de la droite d'ajustement du nuage. 2021 correspond à 10.200 milliers de francs.

Exercice 2

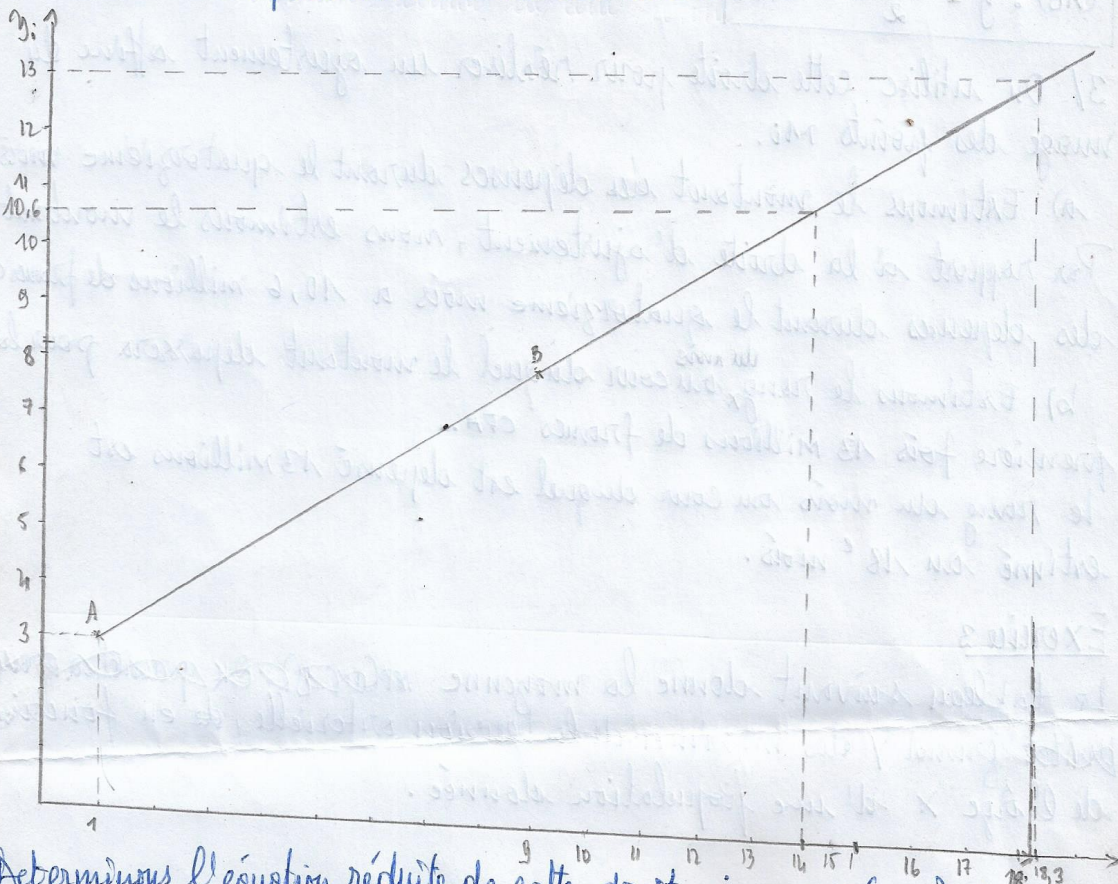
Le tableau suivant donne pour douze mois consécutifs, l'évaluation des dépenses publicitaires (en milliers de francs CFA) d'une société commerciale.

Numéro du mois: n_i	1	2	3	4	5	6	7	8	9	10	11	12
Montant des dépenses: y_i	3	4,5	3,7	5,2	5,2	7,6	7,5	7,5	8,2	9,7	9,7	10

1. Représentons dans un repère orthogonal le nuage des points M_i de coordonnées $(n_i; y_i)$ correspondant à cette série statistique.



2/ Tracons la droite passant par les points $A(1; 3)$ et $B(9; 8,2)$. Et déterminons l'équation réduite de cette droite.



Déterminons l'équation réduite de cette droite. on a : $A(1; 3)$ et $B(9; 8,2)$

On soit $(AB) : y = ax + b$

* Déterminons a

$$a = \frac{y_B - y_A}{x_B - x_A} = \frac{8,2 - 3}{9 - 1} = \frac{5,2}{8} = \frac{1,3}{2}$$

$$a = \frac{1,3}{2}$$

* Déterminons b

On sait que $A \in (AB)$ d'où $(AB) : y_A = \frac{1,3}{2} x_A + b \Leftrightarrow$

$$3 = \frac{1,3}{2} \times 1 + b$$

$$b = 3 - \frac{1,3}{2} = \frac{6 - 1,3}{2} = 2,35$$

$$b = 2,35$$

③

Donc l'équation réduite de cette droite est :

$$(AB) : y = \frac{1,3}{2}x + 2,35$$

3/ On utilise cette droite pour réaliser un ajustement affine du nuage des points M_i .

a) Estimons le montant des dépenses durant le quatorzième mois. Par rapport à la droite d'ajustement, nous estimons le montant des dépenses durant le quatorzième mois à 10,6 millions de francs CFA.

b) Estimons le rang ^{du mois} ou cour duquel le montant dépassera pour la première fois 13 millions de francs CFA. le rang du mois ou cour duquel est dépensé 13 millions est estimé au 18^e mois.

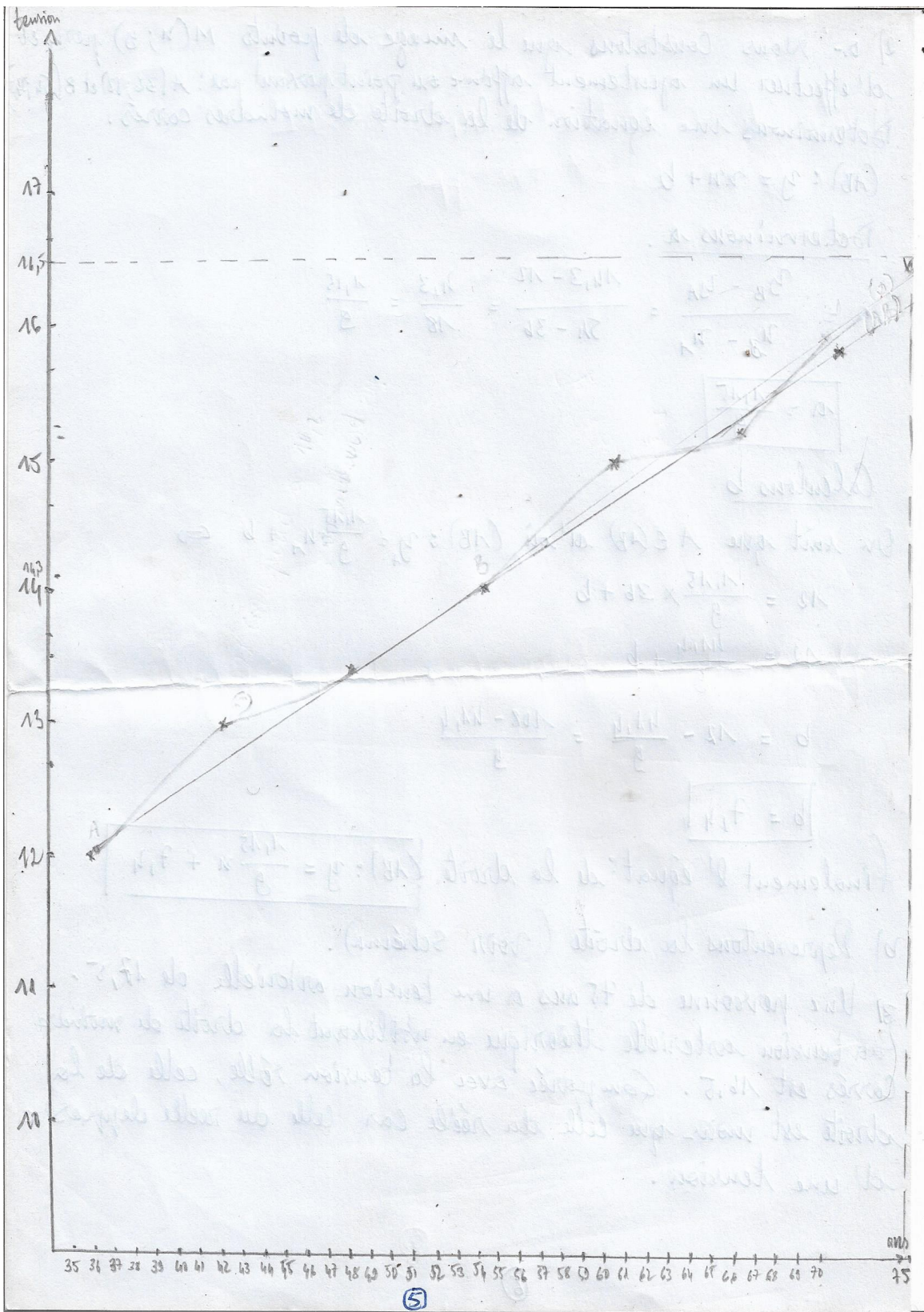
Exercice 3

Le tableau suivant donne la moyenne ~~max~~ ~~de~~ ~~la~~ ~~tension~~ ~~artérielle~~ ~~de~~ ~~en~~ ~~fonction~~ ~~de~~ l'âge x et d'une population donnée.

Âge x	36	42	48	54	60	66	70
Tension y	12	13	13,5	14,3	15,4	15,8	16

1/ Représentons graphiquement le nuage de points $M(x; y)$. On prendra pour unités graphiques 0,5 cm pour 1 an en abscisse et 3 cm pour l'unité de tension artérielle en ordonnée. L'origine sera le point de coordonnées (30; 10).

②



2) a) Nous constatons que le nuage de points $M(n; y)$ permet d'effectuer un ajustement affine au point passant par: $A(36; 12)$ et $B(54; 23)$.
 Determinons une equation de la droite de moindres carrés.

$$(AB): y = an + b$$

Determinons a .

$$a = \frac{y_B - y_A}{n_B - n_A} = \frac{23 - 12}{54 - 36} = \frac{11}{18} = \frac{1,15}{9}$$

$$\boxed{a = \frac{1,15}{9}}$$

Calculons b

On sait que $A \in (AB)$ d'où $(AB): y = \frac{1,15}{9}n + b \Leftrightarrow$

$$12 = \frac{1,15}{9} \times 36 + b$$

$$12 = 4,4 + b$$

$$b = 12 - \frac{44,4}{9} = \frac{108 - 44,4}{9}$$

$$\boxed{b = 7,4}$$

Finallement l'équation de la droite $(AB): y = \frac{1,15}{9}n + 7,4$

b) Représentons la droite (voir schéma).

3) Une personne de 75 ans a une tension artérielle de 17,5.
 Sa tension artérielle théorique en utilisant la droite de moindres carrés est 16,5. Comparée avec la tension réelle, celle de la droite est moins que celle du réel car celle du réel dépasse d'une tension.

⑥

Titre : Les initiatives des élèves dans l'enseignement et l'apprentissage des séries statistiques doubles : exemple de la classe de Terminale SE en Centrafrique

Mots clés : Séries statistiques doubles, ajustement affine, initiatives des élèves, modélisation.

Résumé : Notre étude concerne les prises d'initiatives des élèves en mathématiques, dans le contexte de la République Centrafricaine, en particulier, pour des élèves de terminale Scientifique et Expérimentale (SE). Nous nous centrons sur les séries statistiques doubles et l'ajustement affine. Nous nous référons à la théorie anthropologique du didactique (TAD) de Chevallard (1998), en considérant que le savoir mathématique est façonné par les institutions sous forme de praxéologies mathématiques, tout en empruntant à d'autres approches théoriques (la théorie des situations didactiques (TSD), la notion de registre de représentation de Duval). Nous construisons un modèle praxéologique de référence (MPR) pour les séries statistiques doubles et l'ajustement affine en appui sur une analyse historique, sur une revue de la littérature, sur les programmes de Centrafrique et sur un manuel français. A l'aide de ce MPR, nous étudions tout d'abord l'enseignement ordinaire en Centrafrique,

et la place que celui-ci laisse à l'initiative des élèves. Pour ce faire, nous analysons le manuel officiel de Terminale SE, ainsi qu'un cours que nous avons observé et filmé. De plus nous avons proposé un questionnaire à des enseignants et mené des entretiens. Ensuite nous construisons, mettons en oeuvre et analysons trois expérimentations. Il s'agit de s'intéresser à l'activité des élèves, lorsqu'on leur propose un énoncé ou un cours sortant de la pratique ordinaire, et destiné à favoriser leurs prises d'initiatives. Nos analyses montrent que certains types de tâches importants du MPR n'apparaissent pas dans l'enseignement ordinaire : notamment le fait de choisir une méthode d'ajustement. Les initiatives des élèves sont réduites, les enseignants visent l'apprentissage de techniques. Il est toutefois possible de mettre en place un enseignement qui amène les élèves à mettre plus de sens dans l'ajustement affine.

Title : The initiative of the students in the teaching and learning of double statistics series : The example of experimental and scientific terminal class in Centre African Republic

Keywords : double statistics series, affine adjustment, initiative of the students, modelisation

Abstract : Our study concerns the state initiative of students in mathematics in the context of Central African Republic for the experimental and scientific terminal students. We are stating on double statistic series and the affine adjustment. Our reference is on of the didactic antropologie theory of Chevallard (1998), knowing that the mathematic knowldge is done by the institutions under praxeologic mathematics taken from some approaches of theories of didactic situations, the notion of the registration of Duval. Our of praxeologic reference model (PRM) for the double statistics series and the affine adjustment with an historic analysis on the review of literature on the program in Central African Republic and on a french book. With the aid of the PRM, we are studying first the usual teaching in Central African Republic and the place given to the initiative of the students.

For this, we are analysing the official manual of scientific experimental of terminale class, together with a course that we have observed and taken pictures. We have also proposed some questionnaires to the teachers with some interviews. After we have constructed and put into practice and analysed some experimentations. We have to take into account the activities of the students by given them a topic or a course due to the ordinary way to help them in their own initiatives. Our analyses show that some important tasks of PRM do not appear in the ordinary teaching, in order to choose an adjustment method. The initiatve of the students are reduced, the teachers draw their attention on the learning and the techniques. It is necessary to bring out a teaching which helps the students to pay more attention on the affine adjustment.