UNIVERSITÉ DU BURUNDI
GRENIER DU SAVOIR DU BURUNDI
Dépôt institutionnel officiel
https://repository.ub.edu.bi

Impact des cours disciplinaires sur la réussite en première session pour les étudiants des troisièmes années de Baccalauréat à l'École Normale Supérieure : Cas du Département des Sciences Naturelles

Mbazumutima, Louis; Sous la direction de : Dr Nkurunziza Jean de Dieu

2024
UB, ENS
https://repository.ub.edu.bi/handle/123456789/1980

Document téléchargé depuis le **Grenier du Savoir du Burundi** — https://repository.ub.edu.bi

UNIVERSITÉ DU BURUNDI **ÉCOLE NORMALE SUPÉRIEURE**

MASTER EN DIDACTIQUE DES SCIENCES

OPTION : MATHÉMATIQUE

Impact des cours disciplinaires sur la réussite en première session pour les étudiants des troisièmes années de Baccalauréat à l'École Normale Supérieure : Cas du Département des Sciences Naturelles

Par

MBAZUMUTIMA Louis

Mémoire présenté et défendu publiquement en vue de l'obtention du Diplôme de Master en Didactique des Sciences, option Mathématique

Sous la direction de :

Dr NKURUNZIZA Jean de Dieu

Bujumbura, juillet 2024

Membres du jury

Président du jury : Pr BARAHINDUKA Etienne

Directeur de Mémoire : Dr NKURUNZIZA Jean de Dieu

Secrétaire du jury : Pr BARANKANIRA Emmanuel

Dédicace

À ma femme ;

À mes enfants ;

À mes parents ;

Je dédie ce mémoire.

Remerciements

Avant tout, nous aimerions signaler que ce travail n'est pas le fruit de nos seuls efforts raison pour laquelle nous voudrions présenter nos sentiments de gratitude à toutes les personnes qui ont contribué à sa réalisation.

Nous tenons à adresser nos vifs remerciements au directeur de mémoire Dr NKURUNZIZA Jean de Dieu pour nous avoir dirigé et initié à la rédaction d'un article lequel article a été soumis, accepté et publié dans le journal Iga de l'École Normale Supérieure (ENS); qu'il trouve ici l'expression de notre profonde gratitude.

Nous exprimons également nos vifs remerciements à l'endroit de tous les enseignants et éducateurs pour la formation tant humaine que scientifique dont ils nous ont fait bénéficier. Nous ne pouvons pas trouver l'expression adéquate pour les remercier mais nous leur disons merci infiniment.

Nous tenons aussi à remercier tous les membres du jury pour le temps qu'ils ont bien voulu consacrer pour lire ce travail; leurs précieux commentaires pour l'amélioration de ce travail nous seront d'une importance capitale.

Nos remerciements s'adressent à nos parents qui se sont sacrifiés en accomplissant la noble mission éducative dès le bas âge jusqu'à présent. Sans eux, nous ne serions pas ce que nous sommes aujourd'hui ; qu'ils trouvent ici l'expression de notre profonde gratitude.

Nous terminons à exprimer toute notre gratitude à tous ceux qui ont contribué, de près ou de loin, à la réalisation de ce travail.

MBAZUMUTIMA Louis

Résumé

L'objectif de ce travail était de montrer l'impact des cours disciplinaires sur la réussite en première session pour les étudiants des troisièmes années du Département des Sciences Naturelles. L'École Normale Supérieure n'a jamais réalisé une étude pour se rendre compte pourquoi les étudiants échouent en première session. C'est pour cela que notre étude montre l'impact des cours disciplinaires sur la réussite en première session pour les étudiants des troisièmes années de Baccalauréat du Département des Sciences Naturelles. Il ressort des résultats que quatre cours disciplinaires sont significativement associés à la réussite dans la Section de Physique-Technologie; deux en Section des Sciences Géographiques et un seul cours en Section Mathématique et en Section Biologie-Chimie. Le modèle logistique a été utilisé pour modéliser et analyser les données.

Mots-clés : analyse descriptive, analyse bivariée, analyse multivariée, rapport des cotes.

Abstract

The aim of this work was to show the impact of disciplinary courses on first-term success for third-year students in the Natural Sciences Department. The Ecole Normale Supérieure has never carried out a study to find out why students fail in the first term. That is why our study shows the impact of subject courses on first-term success for third-year Baccalauréat students in the Natural Sciences Department. The results show that three subject courses are significantly associated with success in the Physics-Technology Section; two in the Geographical Sciences section and one course in the Mathematics section and the Biology-Chemistry Section. The logistic model was used to model and analysis the data.

Keywords: descriptive analysis, bivariate analysis, multivariate analysis, odds ratio.

Table des matières **Page**

Membres du jury ..... **i**
Dédicace ..... **ii**
Remerciements ..... **iii**
Résumé ..... **iv**
Abstract ..... **iv**
Table des matières ..... **v**
Liste des illustrations ..... **viii**
Liste des tableaux viii
Liste des figures ix
Sigles et abréviations ..... **x**
Avant-propos ..... **xi**

Introduction générale **1**
Objectifs de l'étude 2
Questions de recherche 2
Hypothèses de recherche 2
Intérêt du sujet 3

Chapitre 1 : Méthodologie **4**
1.1. Introduction 4
1.2. Population cible 4
1.3. Échantillon 4
1.4. Collecte des données 5
1.5. Variables à l'étude 5
1.6. Méthodes d'analyse des données 8
1.7. Cadre conceptuel 9
1.8. Conclusion 9

Chapitre 2 : Théorie du modèle logistique 10
2.1. Introduction 10
2.2. Quelques notions sur le modèle logistique 11
2.2.1. Modèle logistique binaire 11
2.2.2. Cote 12
2.2.3. Rapport de cotes 12
2.3. Tests des paramètres et intervalle de confiance 14
2.3.1. Test de Wald 14
2.3.2. Test de Hosmer et Lemeshow 15
2.3.3. Courbe ROC 15
2.3.4. Intervalle de confiance 16

Chapitre 3 : Impact des cours disciplinaires comme déterminants de la réussite 17
3.1. Introduction 17
3.2. Évolution des taux de réussite 17
3.3. Comparaison des taux de réussite des étudiants des Baccalauréats III 19
3.4. Section Physique-Technologie 21
3.4.1. Description de la réussite dans chaque cours disciplinaires 21
3.4.2. Influence des cours disciplinaires à la réussite en bivarié 22
3.4.3. Analyse multivariée 23
3.4.4. Diagnostique du modèle 24
3.5. Section des Sciences Géographiques 25
3.5.1. Analyse descriptive des variables explicatives 25
3.5.2. Analyse bivariée 26
3.5.3. Validation du modèle 30

3.6. Section de Biologie-Chimie 31
3.6.1. Effectif de réussites par cours disciplinaire en section Biologie-Chimie 31
3.6.2. Analyse bivariée 32
3.7. Section Mathématique 33
3.7.1. Description de la réussite pour chaque cours disciplinaire 33
3.7.2. Association des cours disciplinaires avec la réussite en Section Mathématique 34
3.8. Discussion des résultats 34
3.9. Conclusion 35

Conclusion générale, recommandations et perspective 36
Conclusion générale 36
Recommandations 36
Perspective 37

Références bibliographiques 38

Annexes 40

Liste des illustrations

Liste des tableaux Page

Tableau 1 : Sélection de l'échantillon 5
Tableau 2 : Description des cours disciplinaires pour la Section Physique-Technologie 6
Tableau 3 : Description des cours disciplinaires pour la Section des Sciences Géographiques 6
Tableau 4 : Description des cours disciplinaires pour la Section Mathématique 7
Tableau 5 : Description des cours disciplinaires pour la Section Biologie-Chimie 7
Tableau 6 : Évaluation de l'AUC 8
Tableau 7 : Comparaison des taux de réussite 19
Tableau 8 : Effectif de réussite pour chaque cours en Physique-Technologie 21
Tableau 9 : Association entre la réussite et les cours disciplinaires en Physique-Technologie 22
Tableau 10 : Modèle complet pour la Section de Physique-Technologie 23
Tableau 11 : Modèle saturé pour la Section de Physique-Technologie 24
Tableau 12 : Effectif de réussites pour la Section des Sciences Géographiques 26
Tableau 13 : Association entre cours disciplinaires et réussite pour la Section des Sciences Géographiques 27
Tableau 14 : Modèle complet pour la Section des Sciences Géographiques 28
Tableau 15 : Choix du meilleur modèle 29
Tableau 16 : Modèle sélectionné par le critère d'Akaike (modèle saturé) 29
Tableau 17 : Diagnostique du modèle (Sciences Géographiques) 30
Tableau 18 : Effectif de réussites et pourcentage pour la Section Biologie-Chimie 31
Tableau 19 : Association des cours disciplinaires avec la réussite en bivarié 32
Tableau 20 : Effectif des réussites pour chaque cours disciplinaire 33
Tableau 21 : Facteurs associés à la réussite pour la Section Mathématique 34

Liste des figures **Page**

Figure 1 : Cadre conceptuel 9
Figure 2 : Évolution des taux de réussite pour la Section Physique-Technologie 17
Figure 3 : Évolution des taux de réussite pour la Section des Sciences Géographiques..... 18
Figure 4 : Évolution des taux de réussite pour la Section Mathématique 18
Figure 5 : Évolution des taux de réussite pour la Section Biologie-Chimie 19
Figure 6 : Taux de réussite des étudiants des Baccalauréats III 20
Figure 7 : Courbe ROC 25

Sigles et abréviations

AA **:** Année académique

AIC : Akaike information criterion (Critère d'Information d'Akaike)

AUC : Area under curve (Aire en dessous de la courbe)

BC : Biologie-Chimie

ddl : Degré de liberté

DSN : Département des Sciences Naturelles

E : Échec

ENS : École Normale Supérieure

GEO : Géographie

HL : Hosmer-Lemeshow

IC : Intervalle de confiance

MATH: Mathématiques

NP : Nombre de participants

OR : Odds ratio

PT : Physique-Technologie

R : Réussite

ROC : Receiver operating characteristic

TP : Travail pratique

TR : Taux de réussite

Avant-propos

Ce mémoire traite de l'influence des cours disciplinaires à la réussite tout en utilisant le modèle logistique qui est l'une des méthodes statistiques utilisées pour identifier la relation qui existe entre les variables explicatives X_i (quantitatives, qualitatives ou quantitatives et qualitatives) et une variable qualitative y appelée aussi variable dépendante (ou variable d'intérêt). La variable dépendante était la réussite en première session et cours disciplinaires vus en troisième année étaient considérés comme variables explicatives.

Depuis la création de l'École Normale Supérieure du Burundi, aucune étude n'a été faite pour se rendre compte pourquoi les étudiants échouent lamentablement en première session. De ce fait, l'objectif de ce mémoire est de montrer l'impact des cours disciplinaires sur la réussite en première, dans chaque section, des étudiants de Bac III du Département des Sciences Naturelles en 2021-2022.

Ce mémoire montre à l'École Normale Supérieure en général et au Département des Sciences Naturelles en particulier que les grilles de délibération renferment des informations importantes pour améliorer la science qualité auprès de ses étudiants.

La partie méthodologique de ce mémoire décrit toutes les étapes de recherche en statistique : elle parle de population cible, du calcul de l'échantillon, du logiciel statistique utilisé, des variables à l'étude et des méthodes d'analyse envisagées. Avant de faire l'analyse et présentation des résultats, quelques théories sur le modèle logistique ont été conduites dans le but de montrer théoriquement les formules statistiques utilisées pour trouver les résultats.

Introduction générale

Certains auteurs se sont intéressés aux facteurs de réussite à l'Université : les étudiants qui n'ont pas repris l'année et qui avaient de bonnes notes durant leur scolarité ont en effet plus de chance de réussir à l'université (Degrave et Martou,1996 ; Galand, Frenay et Bourgeois, 2004). Les étudiants qui obtiennent de bons résultats à des tests d'intelligence ou de performance standardisés ont également de meilleure chance de réussite, mais cet effet n'est bien entendu pas indépendant de celui du parcours scolaire (Farsides et Woodfield, 2003). La section suivie au secondaire semble également liée au succès au supérieur : par exemple, en communauté française, les étudiants provenant des sections « latin-math » ont de meilleures chances de réussite que les étudiants ayant suivi des sections « sciences humaines » (Droesbeke, Hecquet et Wattelaer, 2001). Un des résultats les plus consistants dans la littérature scientifique est que plus un étudiant croit en ses capacités à réussir et plus il se fixe des objectifs élevés, plus il a de chances de réussir (Chemers, Hu et Garcia, 2001 ; Multon, Brown et Lent, 1991 ; Torres et Solbeg ,2001). La confiance en ses facultés d'apprentissage, mais aussi en ses capacités de gestion du temps et des exigences de la formation, jouerait donc un rôle crucial (Bandura, 1997). La valeur accordée aux cours, c'est-à-dire à la fois l'intérêt et l'importance accordée à leur contenu et à leur réussite, est un autre élément qui se révèle déterminant pour la performance (Eccles et Wigfied, 2002). L'efficacité personnelle et l'attitude proactive sont toutes deux liées à la réussite scolaire. La confiance en soi aide les étudiants à gérer leur environnement d'apprentissage, quel que soit le programme d'études. Les étudiants qui possèdent ces traits d'autorégulation réussissent également bien dans les cours où ils doivent travailler en équipe, car ils doivent collaborer avec d'autres pour mener à bien la tâche (Burton et Dowling, 2005).

À l'École Normale Supérieure, depuis sa création, aucune étude n'a été réalisée pour savoir les facteurs de réussite en troisième année de Baccalauréat. C'est dans cette optique que cette étude a été conduite. Elle veut montrer si les cous disciplinaires vus en troisième année sont des facteurs explicatifs de la réussite académique pour les étudiants des troisièmes années de baccalauréat au Département des Sciences Naturelles de l'École Normale Supérieure. Dans notre étude, nous nous focaliserons sur l'année précédente (2021-2022).

Notre étude se basera sur les cours disciplinaires et non sur la confiance de facultés, ni la bonne gestion du temps et des exigences de la formation, ni l'obtention de bons résultats dans les tests de performance ou d'intelligence (concours de sélection).

Objectifs de l'étude

L'objectif général de notre recherche est de montrer l'impact des cours disciplinaires, dans chaque section à la réussite des étudiants de Bac III du Département des Sciences Naturelles en 2021-2022.

Les objectifs spécifiques sont de :

1. Examiner les facteurs qui influencent le taux de réussite au Département des Sciences Naturelles pour l'année académique 2021-2022;
2. Comparer les taux de réussite pour l'année académique 2021-2022 selon les sections.

Questions de recherche

Les questions de recherche de notre étude sont :

1. Les cours disciplinaires optique physique, Physique de l'état solide et mécanique quantique sont-ils associés à la réussite des étudiants de bac III de la section Physique-technologie en première session ?
2. Le cours de Géographie de l'Europe et celui de Géographie de l'Océanie influencent-ils la réussite des étudiants de Bac III dans la section des Sciences Géographiques en première session ?
3. En section mathématique, le cours d'Algèbre commutative est-il associé à la réussite des étudiants de bac III en première session ?
4. En section Biologie-Chimie, le cours de Chimie clinique est-il associé à la réussite des étudiants de bac III en première session ?
5. Les taux de réussite selon les sections sont-ils égaux pour l'année académique 2012-2022 ?

Hypothèses de recherche

Nous pouvons formuler les hypothèses de cette étude comme suit :

1. Les cours comme optique physique, Physique de l'état solide et mécanique quantique influencent la réussite des étudiants de Bac III en première pour la section Physique-Technologie(H_1) ;
2. Le cours de Géographie de l'Europe et celui de Géographie de l'Océanie influencent la réussite des étudiants de Bac III dans la section des Sciences Géographiques en première session (H_2) ;
3. En section Mathématique, le cours d'Algèbre commutative est associé à la réussite des étudiants de bac III en première session (H_3) ;
4. En section Biologie-Chimie, le cours de chimie clinique associé à la réussite des étudiants de bac III en première session (H_4) ;
5. Les taux de réussite des étudiants de Bac III sont différents au cours de l'année académique 2021-2022 (H_5).

Intérêt du sujet

Ce sujet montre à l'ENS en général et au Département des Sciences Naturelles en particulier l'information que renferment les grilles de délibération qui sont stockées dans ses tiroirs et rend disponible les résultats pouvant servir comme référence lors des planifications afin d'améliorer le service rendu à la nation.

Notre travail est subdivisé en trois chapitres. Le premier chapitre parle de la description des variables, population cible et échantillon, collecte des données et les méthodes d'analyse des données. Le second chapitre montre quelques théories du modèle logistique à savoir : rapport de cotes, test de Wald, intervalle de confiance, etc.et le troisième chapitre parle de l'analyse statistique de la réussite par le modèle logistique binaire tout en utilisant les cours disciplinaires comme variables explicatives. Signalons que ce mémoire commence par une introduction générale et se termine par conclusion générale, recommandations et perspective.

Chapitre 1 : Méthodologie

1.1. Introduction

Si nous voulons construire une maison, nous nous demandons là où nous trouverons les matériaux de construction. De même, pour la rédaction d'un travail scientifique, nous devons clarifier les matériels et les méthodes à utiliser pour le faire. D'où, dans ce chapitre, nous allons montrer les descriptions des variables, la population, l'échantillon et les méthodes d'analyse statistique à utiliser.

1.2. Population cible

La population est constituée de l'ensemble d'étudiants de Baccalauréat III qui ont fait les épreuves de la première session, dans chaque section du Département des Sciences Naturelles, pour l'année académique 2021-2022. En conséquence, 51 étudiants en section Physique-Technologie, 56 en section Biologie-Chimie, 75 en section Mathématiques et 58 en section Géographique ont été considérés. Cette étude a porté sur une population de 240 étudiants de quatre classes de 3èmeannées de Baccalauréat du Département des Sciences Naturelles.

1.3. Échantillon

La détermination de la taille de l'échantillon a été trouvée en appliquant la formule de Krecje et Morgan (REA L.M.et al., 1997).

$$n = \frac{t^2 \times p(1-p) \times N}{t^2 \times p(1-p) + (N-1) \times y^2} \tag{1.1}$$

avec : • n : taille de l'échantillon ;

- N : taille de la population;
- p : proportion attendue d'une réponse de la population ou proportion réelle;
- t : valeur associée à l'intervalle de confiance d'échantillonnage ;
- y : marge d'erreur d'échantillonnage.

Pour notre cas, $N = 240; p = 0,5; t = 1,96; y = 0,03$. L'application de la formule (1.1) a abouti à la taille de l'échantillon $n = 196$.

En appliquant la formule (1.1), nous avons fait la stratification de notre population par section. Les effectifs des étudiants dans chaque section ont été tirés par l'échantillonnage aléatoire simple.

Tableau 1 : Sélection de l'échantillon

DSN	Section	N	n
	Biologie chimie	56	46
	Sciences géographiques	58	47
	Mathématique	75	61
	Physique technologie	51	42
	Total	240	**196**

1.4. Collecte des données

Les données ont été des grilles de délibérations des résultats de la première session et fiches d'identification des étudiants des troisièmes années de Baccalauréat en 2021-2022 via le service des étudiants. La base des données a été constituée suivant le nombre d'étudiants qui ont fait les épreuves de la première session dans chaque section. La saisie des données a été effectuée en Excel et la sortie des résultats a été faite avec le logiciel stata version 15.

1.5. Variables à l'étude

La variable dépendante désigne la réussite en première session pour les étudiants des troisièmes années au département des sciences naturelles. Cette variable est dichotomique et prend deux modalités à savoir 1 si l'étudiant a réussi en première session et 0 sinon. Les cours disciplinaires vus en troisième ont été considérés comme les variables indépendantes suivant la section. Les tableaux suivant illustrent toutes les variables.

Tableau 2 : Description des cours disciplinaires pour la Section Physique-Technologie

Description	Codage	Variable
Cristallographie	1=Oui ; 0=Non	CRISTA
Physique de l'état solide	1=Oui ; 0=Non	PHES
Mécanique quantique	1=Oui ; 0=Non	MEQUA
Mécanique analytique III	1=Oui ; 0=Non	MEANIII
Optique physique	1=Oui ; 0=Non	OPPH
Matériel et installation électrique	1=Oui ; 0=Non	MAIE
Technique d'assemblage	1=Oui ; 0=Non	TEES
TP spéciaux de la physique + excursions	1=Oui ; 0=Non	TPE
Dessin industriel	1=Oui ; 0=Non	DEIN
Dessin du bâtiment	1=Oui ; 0=Non	DEBA
Géométrie descriptive	1=Oui ; 0=Non	GEDE
Histoire des sciences physiques	1=Oui ; 0=Non	HIPH
Relativité générale	1=Oui ; 0=Non	REGE

Tableau 3 : Description des cours disciplinaires pour la Section des Sciences Géographiques

Description	Codage	Variable
Géographique du Burundi	1=Oui ; 0=Non	GEOBU
Géographie générale d'Afrique	1=Oui ; 0=Non	GEOGEA
Analyse démographique	1=Oui ; 0=Non	ANDEMO
Géographie des transports	1=Oui ; 0=Non	GEOTRA
Environnement et développement durable	1=Oui ; 0=Non	ENVDEDU
Géographie des Amériques	1=Oui ; 0=Non	GEOAM
Tourisme et environnement	1=Oui ; 0=Non	TOUENV
Géographie de l'Europe	1=Oui ; 0=Non	GEOEURO
Géographie de l'Asie	1=Oui ; 0=Non	GEOASI
Géographie de l'Océanie	1=Oui ; 0=Non	GEOOCE
Géographie régionale d'Afrique	1=Oui ; 0=Non	GEOREA
Visite guidée sur les milieux au Burundi	1=Oui ; 0=Non	VMB

Tableau 4 : Description des cours disciplinaires pour la Section Mathématique

Description	Codage	Variable
Analyse complexe	1=Oui ; 0=Non	ANCO
Équations aux dérivées partielles	1=Oui ; 0=Non	EDP
Astronomie	1=Oui ; 0=Non	ASTRO
Analyse numérique et Programmation	1=Oui ; 0=Non	ANNU
Géométrie différentielle	1=Oui ; 0=Non	GEDI
Géométrie projective	1=Oui ; 0=Non	GEPR
Calcul statistique sur l'ordinateur	1=Oui ; 0=Non	CASTO
Langage de programmation	1=Oui ; 0=Non	LAPRO
Systèmes dynamiques	1=Oui ; 0=Non	SYDY
Algèbre commutative	1=Oui ; 0=Non	ALCO
Algèbre de Boole	1=Oui ; 0=Non	ALBO

Tableau 5 : Description des cours disciplinaires pour la Section Biologie-Chimie

Description	Codage	Variable
Génétique et reproduction	1=Oui ; 0=Non	GERE
Embryologie animale	1=Oui ; 0=Non	EMAN
Introduction aux sciences environnementale	1=Oui ; 0=Non	INSCEN
Alimentation et nutrition	1=Oui ; 0=Non	ALNU
Parasitologie et maladies infectieuses	1=Oui ; 0=Non	PMI
Microbiologie	1=Oui ; 0=Non	MICROB
Chimie clinique	1=Oui ; 0=Non	CHICL
Chimie industrielle	1=Oui ; 0=Non	CHIND
Biochimie métallique	1=Oui ; 0=Non	BIME
Notion de toxicologie et d'Ecotoxicologie	1=Oui ; 0=Non	TOXEC

1.6. Méthodes d'analyse des données

L'analyse bivariée des données a été la base de la description de la variable d'intérêt sur toutes les sections ainsi que l'analyse multivariée. Pour dégager les facteurs pris en compte dans le cas de l'analyse multivariée, le test du chi-deux d'indépendance de Pearson au seuil de $0,2$ a été utilisé. Notons que le modèle logistique multivariable

$$\log it\left(p\right)=\log it\left[\frac{p\left(y=1/x_1,....,x_k\right)}{p\left(y=0/x_1,...,x_k\right)}\right]=\beta_0+\beta_1x_1+\cdots+\beta_kx_k+\varepsilon$$

a été considéré; p étant une probabilité de subir l'événement (réussir à la première session), β_0 l'intercept, $\beta_1,...,\beta_k$ les paramètres correspondant aux k variables indépendantes $x_1,...,x_k$ respectivement, y la variable dépendante qui représente la réussite en première session et ε un terme d'erreur. Les rapports de cotes ont été rapportés et leur significativité a été évaluée au seuil de $0,05$. De plus, une sélection manuelle pas-à-pas descendante au seuil de $0,05$ a été faite pour sélectionner les variables à faire figurer dans le modèle saturé et le choix du meilleur modèle a été fait par le Critère d'Information d'Akaike (AIC). La discrimination du modèle a été évaluée suivant les valeurs de l'aire sous la courbe (AUC *pour Area Under Curve* en anglais).Voici en bas, le tableau qui montre les différentes valeurs de l'aire sous la courbe.

Tableau 6 : Évaluation de l'AUC

Valeurs de l'AUC	Discrimination
0,5 ≤ AUC < 0,6	Pas de discrimination
0,6 ≤ AUC < 0,7	Faible
0,7 ≤ AUC < 0,8	Acceptable
0,8 ≤ AUC < 0,9	Bonne
0,9≤AUC ≤ 1	Excellente

(Source : Almquist, Åkesson et Brannström, 2021)

1.7. Cadre conceptuel

Le schéma suivant montre la variable d'intérêt (réussite) et les variables explicatives (cours disciplinaires pour chaque section) pouvant influencer la réussite des étudiants de baccalauréat III.

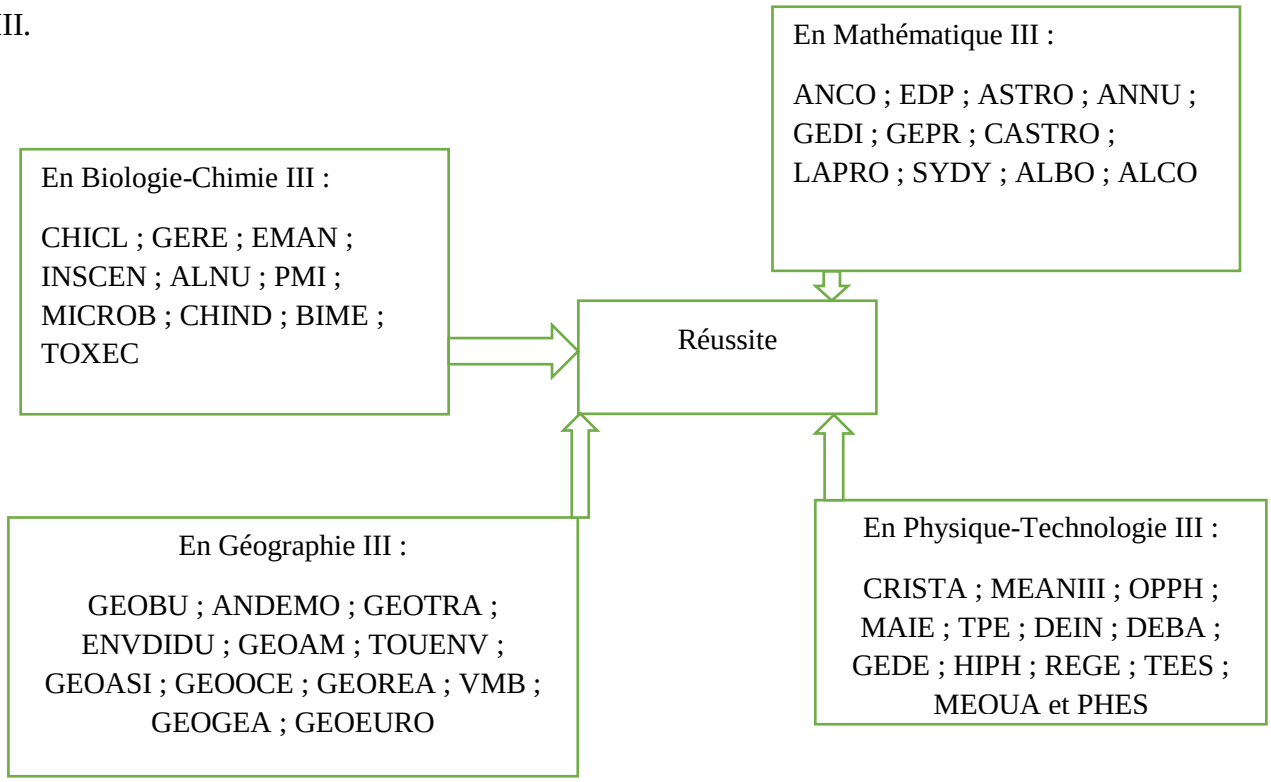


Figure 1 : Cadre conceptuel

1.8. Conclusion

Ce chapitre était de montrer les méthodes mises en avant pour pouvoir analyser l'impact des cours disciplinaires sur la réussite des étudiants du Département des Sciences Naturelles : modèle logistique a été conduite.

Dans le chapitre suivant, nous développerons quelques notions théoriques du modèle logistique.

Chapitre 2 : Théorie du modèle logistique

2.1. Introduction

La statistique est une science définie comme un ensemble de méthodes de recueil des données, d'observation, d'analyse des informations relatives au sujet d'étude et qui permet de conclure et de prendre une décision (Dagnelie, 2006). En se basant à l'analyse des informations relatives à notre sujet d'étude, nous utiliserons la régression logique qui est l'un des modèles statistiques car la variable dépendante (Réussite) est qualitative.

La régression logistique est un modèle mathématique d'analyse multivariée permettant d'identifier la relation entre les variables explicatives X_i (quantitatives, qualitatives ou mélange de deux) et une variable qualitative y appelée aussi variable dépendante (variable à expliquer). Dans le cas où la variable dépendante a deux modalités, on parle de la régression logistique binaire (c'est-à-dire de type vrai ou faux, oui ou non) et il existe la régression logistique binaire simple (à une seule variable explicative) et la régression logistique binaire multiple (à plusieurs variables explicatives). Dans le cas où la variable d'intérêt a plus de deux modalités, on parle de régression logistique polytomique et ce type de régression se scinde en deux : régression logistique ordinale si les modalités de la variable d'intérêt sont ordonnées et régression logistique nominale si ses modalités sont nominales (Stéphane, 2010).

Dans la pratique, la régression logistique est largement répandue dans de nombreux domaines notamment :

- En médecine, elle permet par exemple de trouver les facteurs qui caractérisent un groupe de sujets malades par rapport à des sujets sains ;
- Dans le domaine des assurances, elle permet de cibler une fraction de la clientèle qui sera sensible à une police d'assurances sur un tel ou tel risque particulier
- Dans le domaine bancaire, pour détecter les groupes à risque lors de la souscription d'un crédit ;
- En économétrie, pour expliquer une variable discrète par exemple, les intentions de vote aux élections ;
- En éducation, par exemple, elle permet d'identifier les facteurs de réussite ou de l'échec.

Dans ce chapitre, nous développerons la théorie liée à la régression logistique tout en montrant les notions essentielles utilisées pour faire l'analyse de notre base de données. Pour notre étude, la variable d'intérêt prendra deux modalités à savoir 1 s'il y a la survenue de l'évènement (si

l'étudiant a réussi en première session) et 0 sinon. Nous utiliserons donc le modèle logistique binaire.

2.2. Quelques notions sur le modèle logistique

2.2.1. Modèle logistique binaire

Soient $\beta' = (\beta_1, \beta_2, \beta_3, ..., \beta_k)$ le vecteur des paramètres à estimer, et $x = (x_1, x_2, x_3, ..., x_k)$ le vecteur des variables explicatives. La probabilité conditionnelle de chaque valeur de y (1 ou 0), étant donné le vecteur x des variables explicatives, est modélisée suivant la formule :

$$P(y = 1/x) = \frac{e^{\gamma' z}}{1 + e^{\gamma' z}} = \pi(x) \qquad (2.1)$$

d'où $$P(y = 0/x) = 1 - \pi(x) \qquad (2.2)$$

avec $$\gamma' = (\beta_0, \beta_1, ..., \beta_k) = (\beta_0, \beta'); z = (1, x_1, x_2, ..., x_k) = (1, x).$$

Pour notre étude, selon la section considérée, k désigne le nombre de cours disciplinaires et y la réussite ayant deux modalités à savoir 1 si l'étudiant a réussi en premier session et 0 sinon. En prenant l'exemple de la section physique technologie, nous constatons que :

- x_1 = Cristallographie (CRISTA) ;
- x_2 = Physique de l'état solide (PHES) ;
- x_3 =Mécanique quantique (MEQUA);
- x_4 = Mécanique analytique III (MEANIII) ;
- x_5 =Optique physique (OPPH) ;
- x_6 = Matériel et installation électrique (MAIE) ;
- x_7 = Technique d'assemblage (TEES) ;
- x_8 =TP spéciaux de la physique + excursions (TPE) ;
- x_9 =Dessin industriel (DEIN) ;
- x_{10} =Dessin du bâtiment (DEBA) ;
- x_{11} =Géométrie descriptive (GEDE) ;
- x_{12} =Histoire des sciences physiques (HIPH) ;
- x_{13} =Relativité générale (REGE).

Le modèle logistique est souvent exprimé par le $\log it$ dont l'expression est la suivante :

$$\log it(p) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \log it\left[\frac{p(y=1/x_1,...,x_k)}{p(y=0/x_1,...,x_k)}\right] = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \varepsilon \quad (2.3)$$

où log est le logarithme népérien $(\ln)$, p une probabilité de subir l'événement (réussir en première session), y la variable dépendante qui représente la réussite à la première session et ε un terme d'erreur (Jean-Pierre et Josiane, 2003).

Signalons en passant que la régression logistique a été introduite en 1944 par Berkson en biostatistique, puis en 1973 par McFadden en économétrie (Gilbert, 2006).

2.2.2. Cote

En statistique et probabilités, la cote d'un évènement (odds en anglais) est le ratio entre la probabilité que l'évènement se produise et la probabilité qu'il ne se produise pas .Dans le cas où la variable dépendante et la variable explicative sont binaires, la cote est donnée par :

$$\cot e = odds = \frac{P(y=1/x=1)}{1-P(y=1/x=1)} \quad (2.4)$$

La cote permet à son tour de déterminer le rapport de cotes.

2.2.3. Rapport de cotes

Le rapport de cotes (RC) ou odds ratio (OR) est une mesure statistique communément appelée rapport de chances et rapport de cotes. Il est donné par la formule suivante :

$$OR = \frac{\dfrac{P(y=1/x=1)}{1-P(y=1/x=1)}}{\dfrac{P(y=1/x=0)}{1-P(y=1/x=0)}} \quad (2.5)$$

Par la notation exponentielle, la formule ci-haut mentionnée devient :

$$OR = \frac{e^{\beta_0+\beta_1}}{e^{\beta_0}} \quad (2.6)$$

Ainsi, par simplification, elle devient :

$$OR = e^{\beta_1} \quad (2.7)$$

L'estimation du paramètre β_1 est donnée par l'expression suivante :

$$p\left(y=1/x=0\right)=\frac{e^{\beta_0}}{1+e^{\beta_0}} \tag{2.8}$$

Le rapport de cotes communément appelé rapport de chances s'interprète comme suit :

- Si $OR=1$: Absence d'effet c'est -à- dire la survenue de la modalité considérée pour la variable à expliquer y est fortement libre avec la survenue de la modalité considérée pour la variable explicative x.
- Si $OR<1$: la survenue de la modalité considérée pour la variable à expliquer y est fortement liée avec la survenue de la modalité considérée pour la variable explicative x
- Si $OR>1$: la survenue de la modalité considérée pour la variable à expliquer y est fortement nuisible (néfaste) à la survenue de la modalité considérée pour la variable explicative x.

Pour mettre en pratique la théorie d'odds et d'odds ratio, considérons, pour notre étude, une variable indépendante binaire (par exemple un des cours disciplinaires) qui prend la valeur 1 $\left(x=1\right)$ si un étudiant a eu une note supérieure ou égale à 10 et 0 sinon $(x=0)$ et une variable à expliquer (Réussite) qui prend la valeur 1 $\left(y=1\right)$ si l'étudiant a réussi en première session et 0 sinon. Dans ce cas, la probabilité de réussir en première session si l'étudiant a eu une note supérieure ou égale 10 dans le cours disciplinaire sera donnée par l'expression suivante (Gilbert, 2006).

$$p\left(y=1/x=1\right)=\frac{e^{\beta_0+\beta_1}}{1+e^{\beta_0+\beta_1}} \tag{2.9}$$

Cette probabilité est comparable à la probabilité de ne pas avoir réussi en première session si l'étudiant a eu une note supérieure ou égale à 10 dans le cours disciplinaire. Elle est donnée par l'expression suivante:

$$P\left(y=0/x=1\right)=1-P\left(y=1/x=1\right)=\frac{1}{1+e^{\beta_0+\beta_1}} \tag{2.10}$$

La cote chez les étudiants qui ont eu une note supérieure ou égale à 10 dans le cours disciplinaire est le rapport de ces deux probabilités définies précédemment:

$$odds=\frac{P\left(y=1/x=1\right)}{P\left(y=0/x=1\right)}=e^{\beta_0+\beta_1} \tag{2.11}$$

Nous faisons les mêmes opérations pour trouver l'odds chez les étudiants qui n'ont pas eu une note supérieure à 10 mais qui ont réussi en première session.

Donc, la probabilité de réussir en première session sans toutefois avoir une note supérieure à 10 dans le cours disciplinaire est donnée par :

$$P\left(y=1/x=0\right)=\frac{e^{\beta_0}}{1+e^{\beta_0}} \tag{2.12}$$

De même, cette probabilité est comparable à celle de ne pas avoir réussi quand on n'a pas eu une note supérieure ou égale à10. Nous avons:

$$P\left(y=0/x=0\right)=\frac{1}{1+e^{\beta_0}} \tag{2.13}$$

L'odds chez les étudiants n'ayant pas eu une note supérieure ou égale à 10 est:

$$P\left(y=0/x=0\right)=e^{\beta_0} \tag{2.14}$$

Finalement, nous trouvons l'odds ratio

$$OR=e^{\beta_1} \tag{2 .15}$$

2.3. Tests des paramètres et intervalle de confiance

2.3.1. Test de Wald

Le test de Wald une méthode utilisée pour tester l'apport d'une variable explicative au modèle. Ce test suit asymptotiquement un khi-deux à un degré de liberté sous l'hypothèse de nullité du coefficient théorique (Gilbert, 2006).

Le test de Wald est un test fondé sur la normalité asymptotique des coefficients et consiste à tester la nullité du paramètre β_1.

Les hypothèses du test sont:

$$H_0:\beta_1=0 \text{ et } H_1:\beta_1\neq 0$$

La statistique suivante est utilisée pour tester ces hypothèses

$$W_1=\frac{\hat{\beta}_1^2}{s_1^2}$$

où s_1 est la variance du coefficient $\hat{\beta}_1$, lue sur la diagonale principale de la matrice de variances-covariances. Sous H_0, la statistique W_1 du test de Wald suit une loi du khi-deux à un degré de liberté. Au risque, l'hypothèse H_0 sera rejetée $W_1 > \chi_1^2$.

2.3.2. Test de Hosmer et Lemeshow

Le test de Hosmer-Lemeshow (test HL) est un test d'adéquation pour la régression logistique, en particulier pour les modèles de prédiction des risques. Un test d'adéquation indique dans quelle mesure les données correspondent au modèle. Plus précisément, le test HL calcule si les taux d'événements observés correspondent aux taux d'événements attendus dans les sous-groupes de population.

La statistique du test de Hosmer-Lemeshow est calculée à l'aide de la formule suivante :

$$G_{HL}^2 = \sum_{j=1}^{n} \frac{\left(O_j - E_j\right)}{E_j\left(1 - \frac{j}{n_j}\right)} \sim \chi^2 \qquad (2.16)$$

où:

χ^2 : Chi-deux ;

n_j : Nombre d'observations dans le j^{ème} groupe;

O_j : Nombre de cas observés dans le j^{ème} groupe;

E_j : Nombre de cas attendus dans le j^{ème} groupe.

2.3.3. Courbe ROC

La courbe ROC correspond à la représentation graphique du couple (1 – spécificité ; sensibilité) pour les différentes valeurs seuil. Son allure est en soit en escalier soit en droites par morceaux. L'aire sous la courbe (ou Area Under the Curve-AUC) est un indice synthétique calculé pour les courbes ROC. Pour notre étude, la sensibilité est la capacité du test à bien détecter ceux qui ont réussi en première session et la spécificité est la capacité du test à bien détecter ceux qui n'ont pas réussi.

L'aire sous la courbe ROC (ou Area Under the Curve, AUC) peut être interprétée comme la probabilité que, parmi deux sujets choisis au hasard, réussi et un non réussi, la valeur du

marqueur soit plus élevée pour celui qui a réussi que pour celui qui n'a pas réussi. Par conséquent, une AUC de 0,5 (50%) indique que le marqueur est non-informatif. Une augmentation de l'AUC indique une amélioration des capacités discriminatoires, avec un maximum de 1,0 (100%). La courbe ROC est plus utilisée dans le domaine médical (Delacour, Servonnet, Perrot, Vigezzi et Ramirez, 2005).

Signalons que nous l'avons utilisée pour voir si les variables explicatives sont en quantité suffisante pour la prise de décision en rapport avec la réussite.

2.3.4. Intervalle de confiance

L'intervalle de confiance IC (95%) permet de voir s'il y a une relation entre la variable d'intérêt et la variable explicative. Il est donné par l'expression suivante:

$$IC = \exp\left[\hat{\beta}_i \pm u_i.s\left(\hat{\beta}_i\right)\right] \tag{2.17}$$

Les théories émises dans ce chapitre seront utilisées pour la mise en œuvre dans l'analyse de l'impact des cours disciplinaires comme déterminant de réussite en première session.

Chapitre 3 : Impact des cours disciplinaires comme déterminants de la réussite

3.1. Introduction

Dès la sortie de la première promotion des étudiants de baccalauréat III à l'École Normale Supérieure en général et au Département des Sciences Naturelles en particulier, l'évolution du taux de réussite en première session n'a jamais connu une allure croissante. Dans ce chapitre, nous commençons à montrer graphiquement l'évolution du taux de réussite section par section pour se rendre compte des lacunes de cette réussite. Puis, nous passerons à l'association entre cours disciplinaire et le taux de réussite. Ce dernier varie successivement au sens contraire (fluctuation) comme le montrent les figures suivantes.

3.2. Évolution des taux de réussite

Les figures suivantes décrivent l'évolution des taux de réussite des étudiants de baccalauréat III, dans toutes les sections du Département des Sciences Naturelles, pour une période allant de 2014 à 2022.

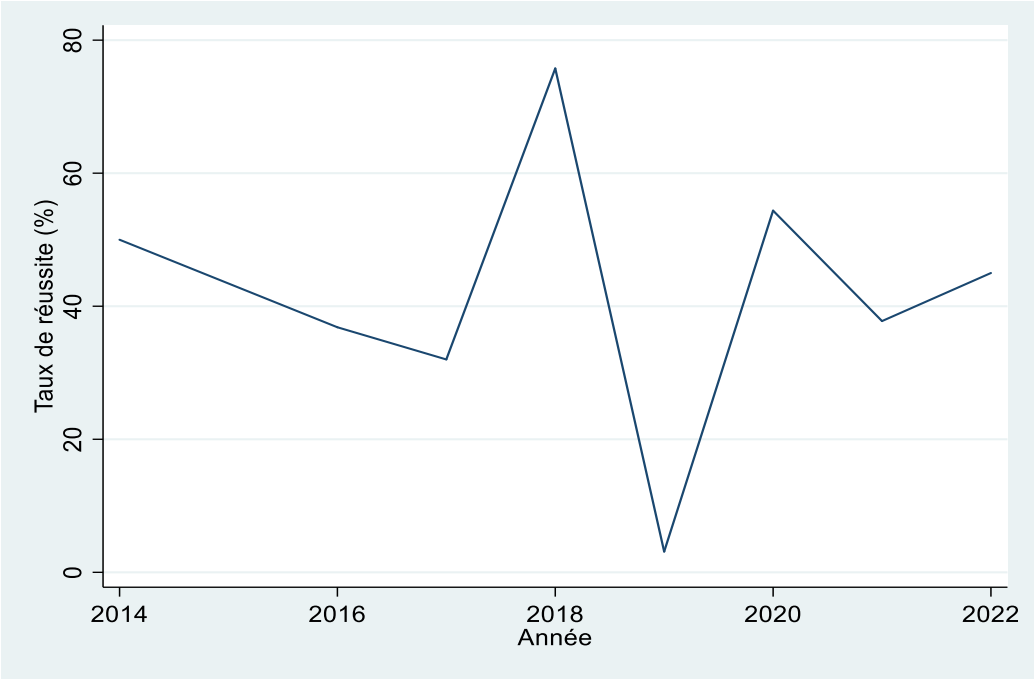


Figure 2 : Évolution des taux de réussite pour la Section Physique-Technologie

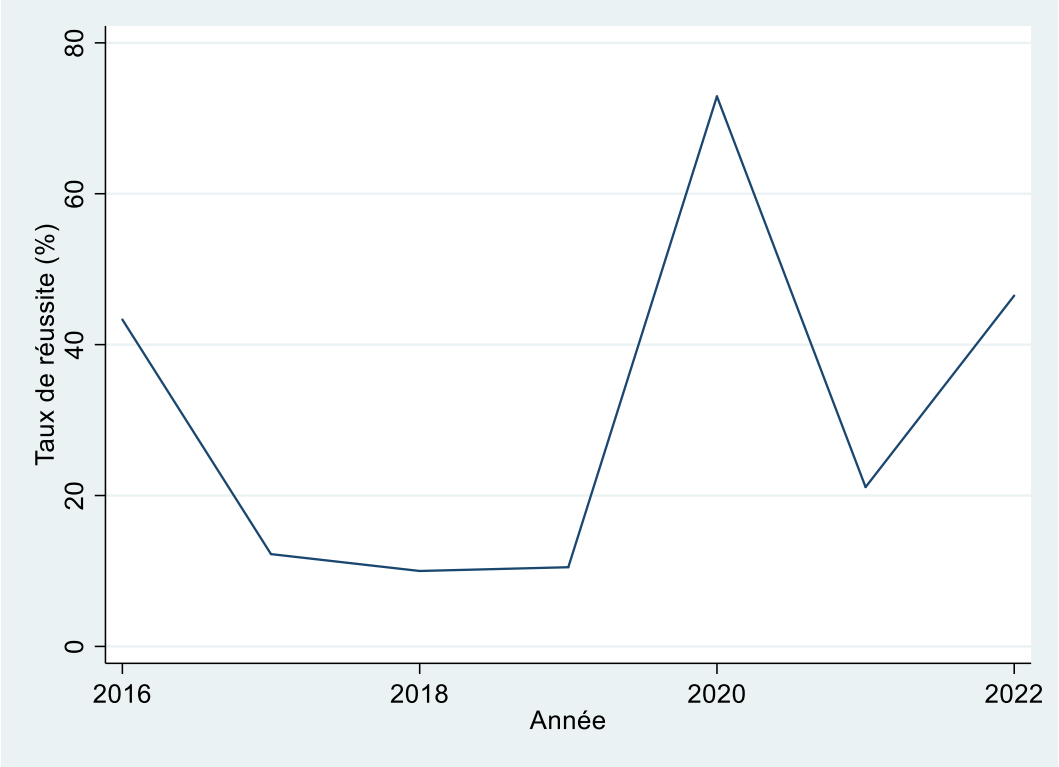


Figure 3 : Évolution des taux de réussite pour la Section des Sciences Géographiques

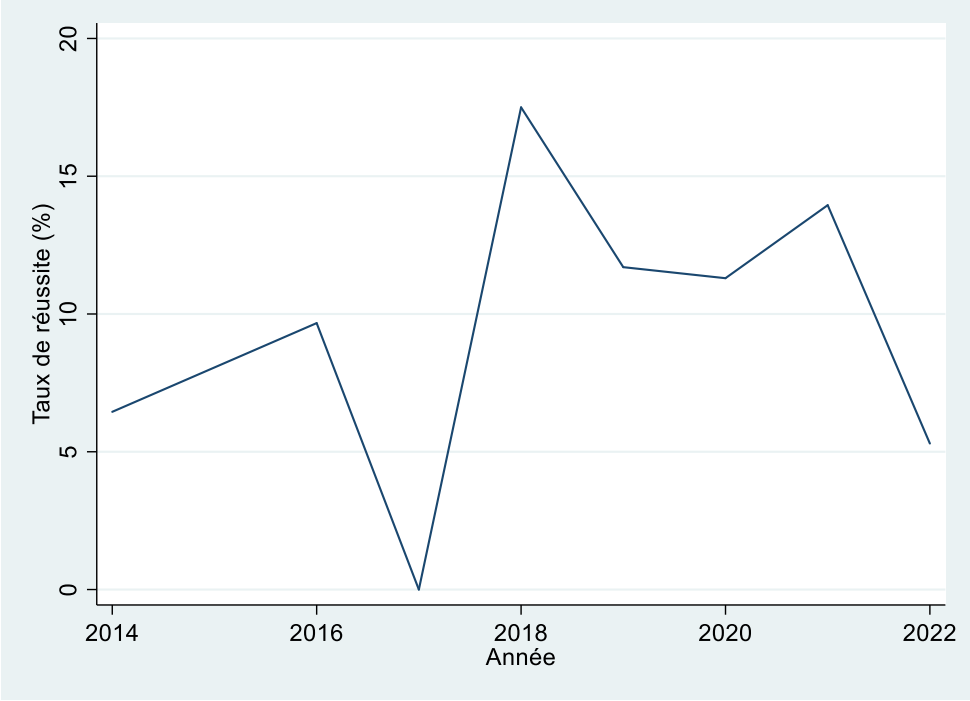


Figure 4 : Évolution des taux de réussite pour la Section Mathématique

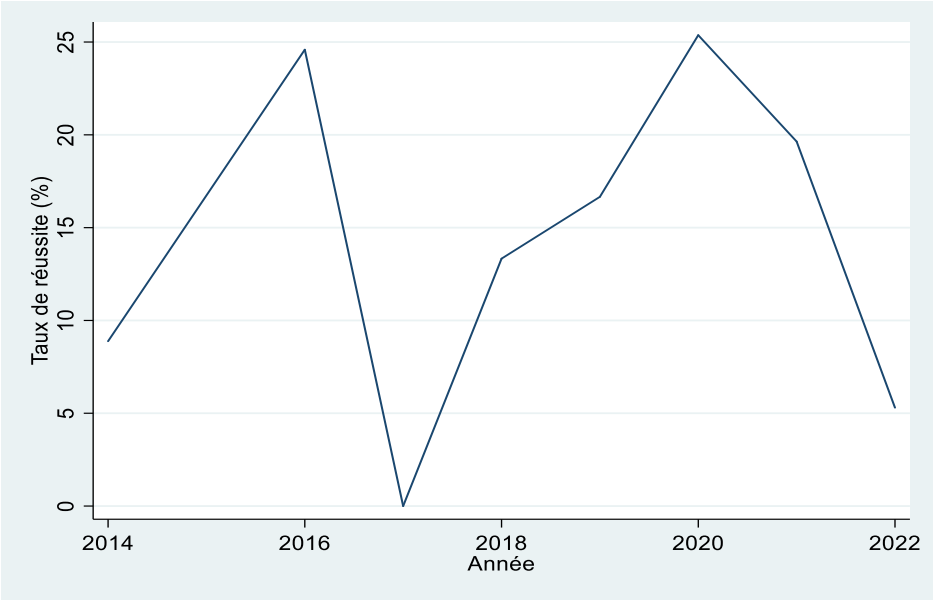


Figure 5 : Évolution des taux de réussite pour la Section Biologie-Chimie

Ces quatre figures montrent que les taux de réussite connaissent des variations successives au sens contraire dans toutes les sections: tantôt l'augmentation tantôt la chute du taux de réussite. Donc, nous pouvons nous poser des questions: Quelle est la contribution des cours disciplinaires à cette fluctuation? Est-il possible de mener une étude scientifique afin de se rendre compte des cours disciplinaires qui sont associés à la réussite ? Pour tenter de répondre à ces questions, voici, dans la suite, les résultats trouvés lors de l'analyse des données par le modèle logistique.

3.3. Comparaison des taux de réussite des étudiants des Baccalauréats III

Selon le test de Student, la comparaison des taux de réussites se présente comme suit:

Tableau 7 : Comparaison des taux de réussite

Sections	t^{obs}	ddl	P-value
Géo-PT	-7,3e+02	46	<0,001
Géo-Maths	-81,262	46	<0,001
Géo-BC	-81,262	46	<0,001
PT-Maths	-76,7	41	<0,001
PT-BC	-76,7	41	<0,001
BC-Maths	11,72	45	1

Le taux de réussite de la section des Géographie (p-value =0,000) est statistiquement différent de ceux des autres sections. De même, le taux de réussite de la section Physique-Technologie est différent de celui de la section Mathématique et Biologie-Chimie. En revanche, Les sections Mathématique et Biologie-Chimie ont statistiquement un même taux de Réussite (p-value=1). Donc, nous confirmons partiellement notre hypothèse (H_5). L'illustration de cette comparaison se voit dans la figure suivante où R désigne la réussite et E désigne l'Échec.

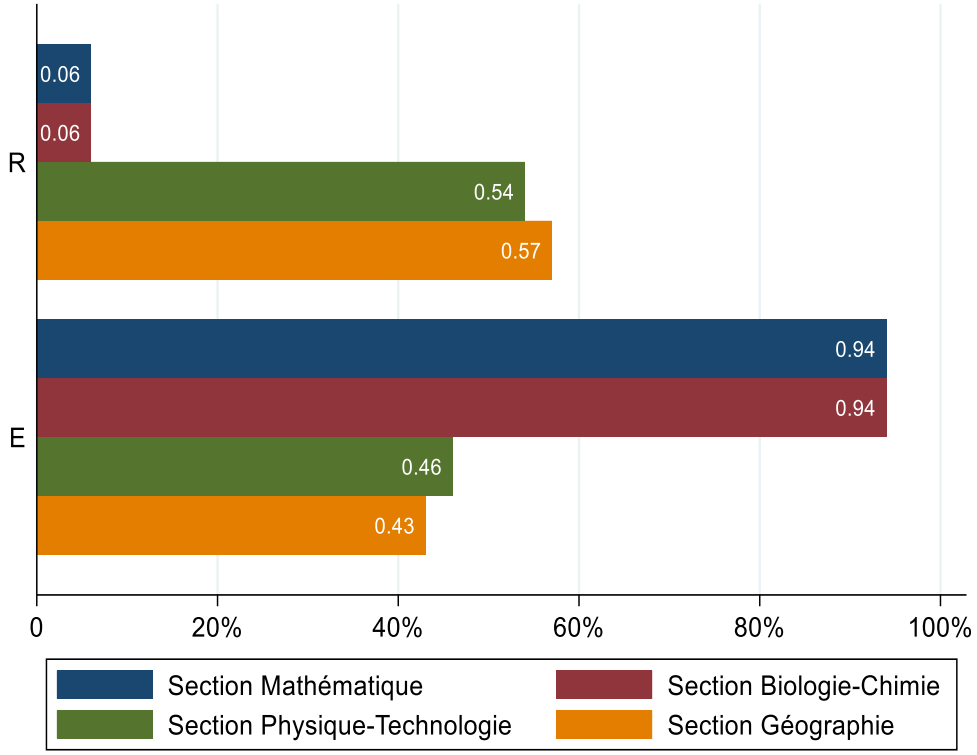


Figure 6 : Taux de réussite des étudiants des Baccalauréats III

3.4. Section Physique-Technologie

3.4.1. Description de la réussite dans chaque cours disciplinaires

Le tableau suivant montre le nombre des étudiants qui ont eu une note supérieure ou égale à dix sur vingt dans chaque cours considéré comme variable explicative.

Tableau 8 : Effectif de réussite pour chaque cours en Physique-Technologie

Variables	Effectif	Pourcentage
CRISTA		
Oui	38	90,48
PHES		
Oui	15	35,71
MEQUA		
Oui	15	35,71
MEANIII		
Oui	41	97,62
OPPH		
Oui	7	16,67
MAIE		
Oui	36	85,71
TEES		
Oui	22	52,38
TPE		
Oui	42	100
DEIN		
Oui	25	59,52
DEBA		
Oui	42	100
GEDE		
Oui	42	100
HIPH		
Oui	41	97,62
REGE		
Oui	21	50

Ce tableau montre que tous les étudiants de Physique-Technologie baccalauréat III ont réussi en première session dans les cours de travaux pratiques spéciaux de la physique et excursions (TPE), dessin du bâtiment (DEBA) et géométrie descriptive (GEDE). Dans les cours de cristallographie, mécanique analytique III (MEANIII) et histoire des sciences physiques (HIPH), matériel et installation électrique (MAIE), technique d'assemblage (TEES) et dessin industriel (DEIN), les étudiants ont réussi à plus de 50% c'est-à-dire plus de la moitié ont eu une note supérieure ou égale à 10 sur 20.Et, nous voyons que les étudiants ont réussi à 50% dans le cours de relativité générale (REGE). Cependant, dans les autres cours, le nombre d'échecs est supérieur au nombre de réussite.

3.4.2. Influence des cours disciplinaires à la réussite en bivarié

Ce tableau montre l'association entre chaque cours disciplinaire avec la réussite des étudiants de baccalauréat III.

Tableau 9 : Association entre la réussite et les cours disciplinaires en Physique-Technologie

Variable	Modalités	OR brut	IC (95%)	P-value
CRISTA				
	Non	1,00		
	Oui	0,41	[0,03; 4,52]	0,453
PHES				
	Non	1,00		
	Oui	0,21	[0,04; 0,91]	**0,021**
MEQUA				
	Non	1,00		
	Oui	0,12	[0,02 ; 0,65]	**0,003**
OPPH				
	Non	1,00		
	Oui	0,50	[0,09 ; 2,67]	0,408
MAIE				
	Non	1,00		
	Oui	3,14	[0,47 ; 20,67]	0,208
TEES				
	Non	1,00		
	Oui	0,14	[0,02 ; 0,69]	**0,004**
DEIN				
	Non	1,00		
	Oui	1,33	[0,37 ; 4,71]	0,653
REGE				
	Non	1,00		
	Oui	0,30	[0,07 ; 1,16]	0,064

Les cours de physique de l'état solide (PHES), mécanique quantique (MEQUA) et technique d'assemblage (TEES) sont significatifs au seuil de 0,05 et au vu de l'intervalle de confiance à 0,95. En effet, les étudiants qui ont eu des points supérieurs ou égaux à dix sur vingt, dans ces cours respectifs (OR=0,41, IC= [0,04 ; 0,91] et P-value=0,021; OR=0,12, IC= [0,02 ; 0,65] et P-value=0,003; OR=0,14, IC= [0,02 ; 0,69] et P-value= 0,004) sont moins susceptibles d'avoir échoué que ceux qui n'en ont pas. En revanche, les cours de cristallographie (CRISTA), matériels et installation électrique (MAIE), dessin industriel (DEIN), optique physique (OPPH) et relativité générale (REGE) n'influencent pas la réussite en première session. Les variables mécaniques analytiques III (MEANIII) et histoire des sciences physiques (HIPH) n'ont pas été utilisées dans les modèles à cause de faibles effectifs dans les proportions. Il en est de même pour les variables physique et excursions (TPE), dessin du bâtiment (DEBA) et géométrie descriptive (GEDE) qui expliquent à 100% la réussite en première session.

Le tableau suivant est constitué de facteurs significatifs à 0,05 et à 0,2 via le modèle logistique binaire multiple.

3.4.3. Analyse multivariée

Le tableau ici-bas montre le modèle complet pour la Section Physique-Technologie.

Tableau 10 : Modèle complet pour la Section de Physique-Technologie

Variable	Modalités	OR	IC (95%)	P-value
PHES				
	Non	1,00		
	Oui	0,10	[0,01 ; 0,76]	**0,026**
MEQUA				
	Non	1,00		
	Oui	0,09	[0,01 ; 0,63]	**0,016**
TEES				
	Non	1,00		
	Oui	0,08	[0,01 ; 0 ,61]	**0,015**
REGE				
	Non	1,00		
	Oui	0,15	[0,02 ; 1,08]	0,060

Ainsi, les cours significativement associés à la réussite à la première session des étudiants de Physique-Technologie, Bac III sont physique de l'état solide (OR=0,10; p-value=0,026**),** la mécanique quantique (OR=0,09; p-value =0,016) et Technique d'assemblage (OR=0,08; p-

value=0,015). Plus précisément, les étudiants ayant une note supérieure ou égale à dix sur vingt sont :

- 0,10 fois plus susceptibles d'avoir réussi que ceux qui n'en ont pas dans le cours de physique de l'état solide ;
- 0,09 fois plus susceptibles de réussir en première session que ceux qui n'en pas ont dans le cours de la mécanique quantique ;
- 0,08 fois plus susceptibles de réussir en première session par rapport à ceux qui n'ont pas une note supérieure ou égale à 10 sur 20 dans le cours de technique d'assemblage.

Après la sélection pas-pas descendante, nous obtenons le modèle saturé suivant :

Tableau 11 : Modèle saturé pour la Section de Physique-Technologie

Variable	Modalités	OR	IC (95%)	P-value
PHES				
	Non	1,00		
	Oui	0,14	[0,02 ; 0,82]	**0,030**
MEQUA				
	Non	1,00		
	Oui	0,10	[0,01 ; 0,60]	**0,012**
TEES				
	Non	1,00		
	Oui	0,13	[0,02 ; 0,73]	**0,021**

En effet, les cours de Physique de l'état solide, mécanique quantique et techniques d'assemblage sont associés à la réussite des étudiants de Bac III en section physique-Technologie. Donc, nous confirmons partiellement l'hypothèse (H_2).

L'équation du modèle s'écrit donc comme suit:

$$\log it(p) = 3,03 - 2,03 \times TEES - 2,265 \times MEQUA - 1,957 \times PHES + \varepsilon$$

3.4.4. Diagnostique du modèle

Au seuil de 0,05, le test d'adéquation de Wald rejette l'hypothèse nulle qui stipule que le modèle n'est pas adéquat ou que tous les paramètres du modèle sont nuls (p-value =0,009<0,05). Le modèle est donc adéquat, ce qui signifie qu'il existe au moins une variable indépendante qui explique le modèle. Selon le test de Hosmer-Lemeshow, l'ajustement est bon (p-value=0,517< 0,05), ce qui signifie que les valeurs observées et les valeurs prédites de réussite sont proches.

La figure suivante montre la courbe ROC (*Receiver Operating Characteristic* en anglais) et l'aire sous la courbe (AUC *pour Area Under Curve* en anglais) estimée à partir des résultats du modèle logistique saturé.

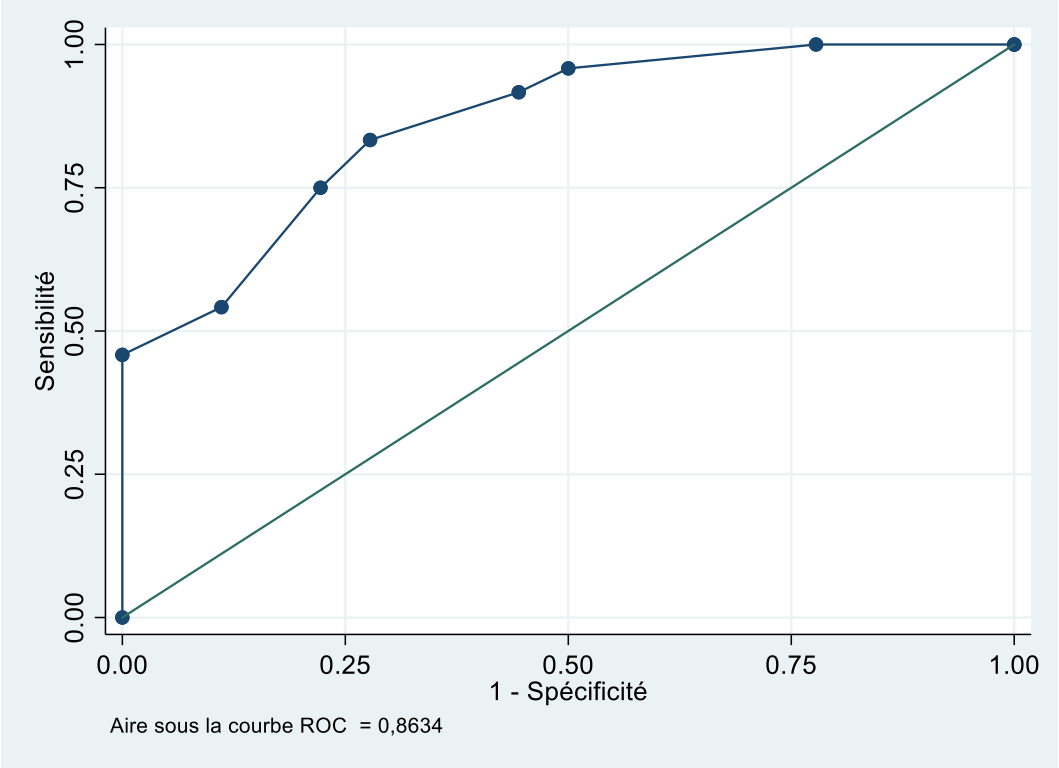


Figure 7 : Courbe ROC

L'aire sous la courbe ROC est de 0,8634, ce qui nous garantit à dire que le modèle a une discrimination bonne. Donc, les variables explicatives considérées dans notre modèle sont suffisantes et permettent de prendre une décision adéquate en rapport avec la réussite en première session des étudiants finalistes de la Section Physique-Technologie.

3.5. Section des Sciences Géographiques

3.5.1. Analyse descriptive des variables explicatives

L'analyse descriptive montre que tous les étudiants ont une note supérieure ou égale dix sur vingt dans les cours de géographie du Burundi (GEOBU), analyse démographique (ANDEMO) et géographie d'Asie (GEOASI). Nous voyons que les étudiants ont échoué dans les cours de géographie des transports (GEOTRA), géographie générale d'Afrique (GEOGEA) et visite guidée sur les milieux au Burundi (VMB). En revanche, une bonne réussite se voit dans les autres cours disciplinaires.

Tableau 12 : Effectif de réussites pour la Section des Sciences Géographiques

Variables	Effectif	Pourcentage
GEOBU		
Oui	47	100
GEOGEA		
Oui	16	34,04
ANDEMO		
Oui	47	100
GEOTRA		
Oui	15	31,91
ENVDEDU		
Oui	26	55,32
GEOAM		
Oui	25	53,19
TOUENV		
Oui	44	93,62
GEOEURO		
Oui	35	74,47
GEOASI		
Oui	47	100
GEOOCE		
Oui	39	82,98
GEOREA		
Oui	43	91,49
VMB		
Oui	16	34,04

3.5.2. Analyse bivariée

Le tableau ci-dessous indique l'effet de chaque variable explicative à la réussite. Toutes les variables explicatives statistiquement significatives à 0,2 vont faire partie du modèle logistique multivariable afin de trouver les facteurs significatifs au seuil de 0,05 .Toutefois, les cours de géographie d'Asie (GEOASI), Analyse démographique (ANDEMO) et géographie du Burundi

(GEOBU) n'ont pas été utilisés dans les modèles car ils sont colinéaires. Il en est de même pour le cours de tourisme et environnement (TOUANV) qui présente un faible effectif dans les proportions. Le modèle logistique saturé sera trouvé en rejetant les variables explicatives les moins significatives par la méthode de sélection manuelle descendante. La validation du modèle utilisé va être faite par le critère de l'information d'Akaike.

Tableau 13 : Association entre cours disciplinaires et réussite pour la Section des Sciences Géographiques

Variable	Modalités	OR brut	IC (95%)	P-value
GEOGEA				
	Non	1,00		
	Oui	0,10	[0,02 ; 0,58]	**0,001**
GEOTRA				
	Non	1,00		
	Oui	0,51	[0,14 ; 1,85]	0,304
ENVDEDU				
	Non	1,00		
	Oui	0,13	[0,03 ; 0,60]	**0,002**
GEOAM				
	Non	1 ,00		
	Oui	0,26	[0,07 ; 0,95]	**0,029**
GEOEURO				
	Non	1,00		
	Oui	0,05	[0,004 ; 0,64]	**0,001**
GEOOCE				
	Non	1,00		
	Oui	1,05	[0,223 ; 4,90]	0,947
GEOREA				
	Non	1.00		
	Oui	1,04	[0,13 ; 8,31]	0,964
VMB	Non	1,00		
	Oui	0,28	[0,07 ; 1,10]	0,053

Au seuil de 0,05 et au vu de l'intervalle de confiance à 0,95 du rapport de cotes, les cours comme géographie générale d'Afrique (GEOGEA), environnement et développement durable (ENVDEDU), géographie des Amériques (GEOAM) et géographie de l'Europe (GEOEURO) sont statistiquement significatifs. Plus précisément, les étudiants qui ont une note supérieure ou égale à 10 sur 20 points dans ses cours disciplinaires sont moins susceptibles d'avoir échoué que ceux qui n'en ont pas. En revanche, d'autres facteurs ne sont pas significatifs. A ces cours statistiquement significatifs à 0,05 s'ajoute le cours de visite guidée sur les milieux au Burundi (VMB), significatif à 0,2, pour former le modèle complet en faisant la régression logistique binaire multiple.

Tableau 14 : Modèle complet pour la Section des Sciences Géographiques

Variable	Modalités	OR brut	IC (95%)	P-value
GEOGEA				
	Non	1,00		
	Oui	0,07	[0,006 ; 0,74]	**0,028**
ENVDEDU				
	Non	1,00		
	Oui	0,42	[0,07 ; 2,37]	0,333
GEOAM				
	Non	1,00		
	Oui	0,48	[0,08 ; 2,66]	0,409
GEOEURO				
	Non	1,00		
	Oui	0,04	[0,002 ; 0,62]	**0,022**
VMB				
	Non			
	Oui	0,42	[0,07 ; 2,61]	0,359

Le tableau montre que le cours de géographie générale de l'Afrique (GEOGEA) et celui de géographie de l'Europe (GEOEURO) sont statistiquement significatifs au seuil de tolérance de 0,05. En effet, les étudiants qui ont une note supérieure ou égale à 10 sur 20 dans ces cours (0R=0,07; IC= [0,006 ; 0,74]; P-value=0,028 et OR=0,04 ; IC= [0,002 ; 0,62]; P-value=0,022 respectivement) sont moins susceptibles d'avoir échoué que ceux qui n'en ont pas.

Cependant, les cours d'environnement et développement durable (ENVDEDU), de géographie des Amériques (GEOAM) et de visite guidé sur les milieux au Burundi (VMB) ne sont pas significatifs : plus précisément, les étudiant ayant une note supérieure ou égale à dix sur vingt dans le cours d'environnement et développement durable, géographie des Amériques et visite guidé sur les milieux au Burundi (OR=0,42 , IC=[0,07 ; 2,37], P-value=0,333 ; OR=0,48, IC=[0,08 ; 2,66], P-value=0,409; OR=0,42 , IC=[0,07 ; 2,61], P-value=0,359 respectivement) sont moins susceptibles d'avoir réussi que ceux qui n'en ont pas.

À partir de ces résultats, nous allons procéder à la recherche du meilleur modèle par la méthode de pas- à- pas descendante et le critère de l'information d'Akaike.

Tableau 15 : Choix du meilleur modèle

Modèle	AIC	Variables supprimées
Complet	49,36	-
En étape 1	48,04	GEOAM
En étape 2	**47,12**	**VMB**
En étape 3	47,64	ENVDEDU

Tableau 16 : Modèle sélectionné par le critère d'Akaike (modèle saturé)

Variables	**OR (95%IC)**	**p-value**
GEOGEA	0,06 [0,007 ; 0,616]	**0,017**
ENVDEDU	0,28 [0,058 ; 1,375]	0,118
GEOEURO	0,037 [0,002 ; 0,610]	**0,021**

En conséquence, les cours disciplinaires associés statistiquement à la réussite à la première session des étudiants de Géographie, Bac III sont géographie générale de l'Afrique (OR=0,06 ; p-value=0,017) et géographie de l'Europe (OR=0,037, p-value =0,021). Pour ce, nous confirmons partiellement notre hypothèse (H_2). Cette insuffisance des cours disciplinaires significativement associés à la réussite en première session des étudiants de Géographie révèle la non-importance de la diagnostique de ce modèle. Donc, l'équation du modèle s'écrit comme suit :

$$\log it(p) = 4,21 - 3,28 \times GEOEURO - 2,70 \times GEOGEA - 1,26 \times ENVDEDU + \varepsilon$$

3.5.3. Validation du modèle

Pour la validation du modèle, Certaines hypothèses du modèle logistique binaire multiple ont été vérifiées dans le but de diagnostiquer le bon choix du meilleur modèle. Pour ce faire, le test de Wald, test d'ajustement du modèle (test d'ajustement d'Homsmer-Lemeshow) ont été notre outils de travail.

- **Test d'adéquation du modèle** (Test de Wald): Au seuil de 0,05, le test d'adéquation de Wald rejette l'hypothèse nulle qui stipule que le modèle n'est pas adéquat ou que tous les paramètres du modèle sont nuls (p-value=0,010 <0,05). Le modèle est donc adéquat, ce qui signifie qu'il existe au moins une variable indépendante qui explique le modèle.
- **Test d'ajustement du modèle** : Le test d'ajustement d'Hosmer-Lemeshow (p-value= 0,592) n'a pas rejeté l'hypothèse nulle qui stipule que les valeurs théoriques sont statistiquement égales aux valeurs observées. Voici, en bas, les résultats de ces tests :

Tableau 17 : Diagnostique du modèle (Sciences Géographiques)

Diagnostique	Type de test	P-value
Adéquation du modèle	Test de Wald	0,010
Ajustement du modèle	Test de Hosmer-Lemeshow	0,592

3.6. Section de Biologie-Chimie

3.6.1. Effectif de réussites par cours disciplinaire en section Biologie-Chimie

Les étudiants qui ont eu une note supérieure ou égale à dix sur vingt, dans chaque cours disciplinaire, sont conférés dans le tableau suivant.

Tableau 18 : Effectif de réussites et pourcentage pour la Section Biologie-Chimie

Variables	Effectif	Pourcentage
GERE		
Oui	18	39,13
EMAN		
Oui	41	89,13
INSCEN		
Oui	33	71,74
ALNU		
Oui	30	53,57
PMI		
Oui	27	58,70
MICROB		
Oui	46	100
CHICL		
Oui	7	15,22
CHIND		
Oui	11	23,91
BIME		
Oui	40	86,96
TOXEC		
Oui	35	76,09

Notre étude montre que la majorité des étudiants ont eu une note supérieure ou égale à dix sur vingt dans les cours disciplinaires suivants: embryologie animale (89,13%), introduction aux sciences environnementale (71,74%), alimentation et nutrition (53,57%) parasitologie et maladies

infectieuses (58,70%), microbiologie (100%), Biochimie métallique (86,96%) et notion de toxicologie et d'écotoxicologie (76,09%).

3.6.2. Analyse bivariée

Le tableau ici-bas montre l'influence de chaque cours disciplinaire à la réussite.

Tableau 19 : Association des cours disciplinaires avec la réussite en bivarié

Variable	Modalités	OR brut	IC (95%)	P-value
GERE				
	Non	1,00		
	Oui	0,62	[0,03 ; 11,12]	0,750
CHICL				
	Non	1,00		
	Oui	0,15	[0,007 ; 3,16]	**0,016**

Le tableau 19 (en bivarié) montre qu'un seul cours disciplinaire «Chimie clinique » (CHICL) est statistiquement significatif au seuil de 0,05.Alors, l'hypothèse (H_4) est vérifiée. Cependant, d'autres facteurs ne sont pas pris en considération à cause de non convergence statistique. Le taux de réussite des étudiants de Baccalauréat de cette section est très faible (0,06) de sorte que cette proportion soit gratuitement assimilée au taux de non-réussite. Ceci revient à dire que la variable réussite pour cette section se comporte comme une constante. En conséquence, l'analyse multivariable n'a plus de sens. Cependant, le modèle retenu peut donc s'écrire de la façon suivante:

$$\log it(p) = 3,44 - 2,72 \times CHICL + \varepsilon$$

3.7. Section Mathématique

3.7.1. Description de la réussite pour chaque cours disciplinaire

Les effectifs des étudiants qui ont eu une note supérieure ou égale à dix sur vingt dans chaque cours disciplinaires sont inscrits dans le tableau suivant.

Tableau 20 : Effectif des réussites pour chaque cours disciplinaire

Variables	Effectif	Pourcentage
ANCO		
Oui	35	57,38
EDP		
Oui	35	57,38
ASTRO		
Oui	53	86,89
ANNU		
Oui	3	4,92
GEDI		
Oui	42	68,85
GEPR		
Oui	49	80,33
CASTO		
Oui	33	54,10
LAPRO		
Oui	34	55,74
SYDY		
Oui	35	57,38
ALCO		
Oui	32	52,46
ALBO		
Oui	61	100

La description de la réussite des étudiants de baccalauréat III montre que le nombre de réussites, pour chaque cours disciplinaire, dépasse celui d'échecs. Cependant, le cours d'analyse

numérique et programmation (ANNU) se distingue en nombre d'échecs. Nous voyons que tous les étudiants ont réussi en première session dans le cours d'algèbre de Boole (ALBO).

3.7.2. Association des cours disciplinaires avec la réussite en Section Mathématique

Le tableau suivant montre l'influence des cours disciplinaires à la réussite annuelle des étudiants en première session.

Tableau 21 : Facteurs associés à la réussite pour la Section Mathématique

Variable	Modalités	OR brut	IC (95%)	P-value
CASTO				0,656
	Non	1,00		
	Oui	0,57	[0,04 ; 6,85]	
ALCO				**0,043**
	Non	1,00		
	Oui	0,23	[0,03 ; 4,07]	

Un seul cours (Algèbre commutative) est statistiquement associé à la réussite (p-value < 0.05). Nous confirmons notre hypothèse (H_4).L'équation du modèle est donnée par :

$$\log it = 2,27 - 0,12 \times ALCO + \varepsilon$$

3.8. Discussion des résultats

Les résultats de cette étude montrent que la réussite des étudiants de baccalauréat III, dans le département des sciences naturelles, en première session, est influencée par certains cours disciplinaires vus en troisième année. Ces résultats sont similaires aux résultats trouvés dans l'étude menée par Neuville et ses coéquipiers où ils ont prouvé que la valeur accordée aux cours, c'est-à-dire à la fois l'intérêt et l'importance accordée à leur contenu et à leur réussite, est un élément qui se révèle déterminant pour la performance (Neuville, Bourgeois et Frenay, 2007; Eccles, Jacquelyne et Allan, 2002).

Les résultats montrent que 6% des étudiants de la section Mathématique, 6% des étudiants de la section Biologie-Chimie, 54% des étudiants de la section Physique-Technologie et 53% de ceux de la section des Sciences Géographiques ont eu un diplôme après les épreuves de la première session. Pour ce, ces résultats renseigne qu'il est difficile de réussir en première session au département des sciences naturelles en général et ils sont aussi semblables à ceux de l'étude effectuée en Allemagne (Danilowicz-Gösele, Lerche, Meya et Schwager, 2017). Ces taux sont déplorables particulièrement dans la section Mathématique et Biologie-Chimie, nous

pourrions dire que ces résultats sont dus du fait que les étudiants ne travaillent presque pas en groupe. Or, pour bien mener la tâche d'étudier, ça demande, à part des explications de l'enseignant, des explications complémentaires émanant des pairs (Burton et Dowling, 2005). Toutefois, ils sont différents de ceux de l'étude effectuée par Droesbeke et ses coéquipiers dans laquelle on a trouvé que l'option (section) suivie au secondaire semble liée au risque d'échec au supérieur (Droesbeke, Hecquet et Wattelaer, 2001).

3.9. Conclusion

Ce chapitre a été introduit pour montrer l'impact des cours disciplinaires à la réussite, en première session, des étudiants de baccalauréats III du département des sciences naturelles. Avant de faire l'analyse des données par la régression logistique binaire, certaines notions de la régression logistique ont été développées théoriquement: entre autres, rapport de cotes, estimations des paramètres, tests des paramètres et tests d'adéquation du modèle. Certains cours disciplinaires vus en troisième année sont liés à la réussite: entre autre, en Physique-Technologie, les cours comme physique de l'état solide, mécanique quantique et technique d'assemblage sont liés à la réussite; en Sciences géographiques, les facteurs comme géographie générale de l'Afrique et géographie de l'Europe sont associés à la réussite; en section Mathématique, le cours d'algèbre commutative se distingue en se montrant qu'il est le seul à influencer l'obtention du diplôme. Il en est de même qu'en section Biologie-Chimie, le cours de chimie clinique est statistiquement, en bivarié, significatif au seuil de 0,05.

Conclusion générale, recommandations et perspective

Conclusion générale

Dans ce travail, nous avons construit des modèles mathématiques permettant d'expliquer la réussite en première session des étudiants de Baccalauréat III du Département des Sciences Naturelles, en 2022 à l'aide des variables explicatives comme les cours disciplinaires vus en troisième année selon la section.

Ce travail était subdivisé en trois chapitres. Le premier parlait de la méthodologie utilisée et la description des variables extraites dans les grilles de délibération et maquettes de cours. Le second chapitre portait sur quelques théories du modèle logistique: entre autres, le modèle logistique binaire, notion de rapport de cotes, tests des paramètres et intervalle de confiance. Ces notions ont été utilisées dans le troisième chapitre afin d'étudier l'impact des cours disciplinaires comme déterminant de la réussite en première session. En fait, le troisième chapitre concernait l'explication de la réussite par les cours disciplinaires vus en troisième année pour l'année académique 2021-2022. Signalons que l'extrait du chapitre trois a été un article soumis, accepté et publié dans le journal Iga de l'ENS.

Les cours disciplinaires qui expliquent cette réussite selon la section sont: en Physique-Technologie, les cours comme physique de l'état solide, mécanique quantique et technique d'assemblage; en Sciences géographiques, les facteurs comme géographie générale de l'Afrique et géographie de l'Europe sont associées à la réussite; en section Mathématique, le cours d'algèbre commutative se distingue en se montrant qu'il est le seul à influencer l'obtention du diplôme et il en est de même pour le cours de chimie clinique en section Biologie Chimie. Nous signalons que ces résultats sont valables si rien n'est fait.

Recommandations

Aux étudiants, nous recommandons ce qui suit :

- Multiplier les efforts afin d'augmenter le taux de réussite en première session ;
- D'étudier sérieusement tous les cours disciplinaires car c'est dans ces cours qu'ils tireront des notions importantes pour le travail qui leur est attendu au secondaire ;
- Mettre une attention particulière à l'obtention du diplôme après les épreuves de la première session.

Aux enseignants :

- Donner, en première session, les questions semblables à celles des travaux dirigés et/ou pratiques;
- Corriger les travaux dirigés et/ou pratiques avant de choisir les questions d'examen pour la première session.

L'ENS;

- De mener une étude sur les autres Départements;
- De faire une autoévaluation, dans tous les départements, chaque fois que 5 ans s'écoulent, pour se rendre compte d'évolution du taux de réussite en première session pour les étudiants de Baccalauréat III.

Perspective

Nous ne pouvons pas admettre que le contour de la théorie et pratiques du modèle logistique a été épuisé. Nous nous sommes bornés sur le modèle logistique binaire et sa mise en pratique pour la réussite des étudiants de Baccalauréat III du Département de Sciences Naturelles de l'ENS. Dans la même optique, nous invitons tout autre chercheur désirant étendre cette recherche sur les autres départements de l'ENS, qu'il soit le bienvenu.

Références bibliographiques

Almquist, Y., Åkesson, C. et Brannström, L. (2021). An applied guide to quantitative methods with stata. *Research Reports in Public Health Science*, Stockholm University, n°: 2021:1.

Bandura, A. (1997). *Self- efficacy: the exercises of control*: New York: W. H. Freeman and company.

Burton, L.J. and Dowling, D.G. (2005). In search of the key factors that influence student success at university. *Preceeding of the 28th Higher Education Research and development society of Australasia Annual conference (DERDSA);* Sydney, Australia 03-06 juillet 2005: PP68.

Chemers, M .M., Hu , L. and Garcia, B.F.(2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Education psychology, 93,* 55-64.

Dagnelie P. *Statistique théorique et appliquée* .Tome 2 ; Bruxelles : DE boeck ; 2006

Danilowicz-Gösele, K., Lerche, K., Meya, J. and Schwager, R. (2017). Determinants of students'*success* at university. *Education Economics, 25(5), 513–532.* doi:10.1080/09645292.2017.1305329.

Degrave, F. et Martou, F. (1996). *Efficacité de l'enseignement : Une question de mesure à propos de l'enseignement en communauté Française de Belgique. In Education et Formation, facteurs de compétitivité (12ème congrès des économistes belges de la langue française, commission 4 pp. 143-167)*. Charlera : Centre Interuniversitaire de Formation Permanente.

Delacour, H., Servonnet,A. ,Perrot ,A., Vigezzi, J.F.et Ramirez, J.M. (2005). La courbe ROC (Receiver operating characteristic) : *principes et principales applications en biologie clinique,* 63(2) : 145-54.

Droesbeke, J.J., Hecquet, I. et Wattelaer, C. (2001). *La population étudiante: description, évolution, perspective*s. Bruxelles : Edition de l'Université de Bruxelles et Edition Ellipses.

Eccles, S.J, and Wigfield, A. Motivational beliefs, values, and goals. *Annual Review Psychology, 2002, 53: 109-32.*

Farsides, T. and Woodfield, R. (2003). Individual differences and undergraduate academic success: the roles of personality, intelligence and application. Personality and individualdifferences, 34, 1225-1243.

Galand, Frenay et Bourgeois, (2004). *Facteurs de réussite en 1ère candidature de la faculté des sciences et à l'Institut de l'Education physique et de Réadaptation*. Communication orale à la journée d'étude de la chaire de pédagogie universitaire, Louvain-la- Neuve, Belgique.

Gilbert, S. (2006). *Probabilités, analyses des données et statistique (*2ème édition révisée et augmentée), Paris: TECHNIP.

Jean-Pierre et Josiane (2003) .*Statistique explicative appliquée.* Paris, Editions TECHNIP.

Multon, K.D., Brown, S.D. and Lent, R.W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of counseling psychology, 38, 30-38.*

Neuville, S., Bourgeois, E. and Frenay, M.(2007). Task value, self-efficacy and goal orientations: impact on self-regulated learning, choice and performance among university students. *European journal of educational psychology, 47-1/2, 95-17.*

Rea, L.M. (1997). Calcul de la taille d'un échantillon pour une enquête. https://memento-assainissement.gret.org/IMG/pdf/memento-assainissement-fiche4.pdf (consulté le 01/08/2014 à 13h40 minutes.

Stéphane Tufféry .*Data Mining et statistique décisionnelle* : L'intelligence des données *;* Editions TECHNIP, Paris, 2010, P 439.

Torres, J. and Solbeg, V. (2001). Role of self-efficacy, stress, social integration, and family support in latino college student persistence and health. *Journal of vocational Behavior, 59*, 53-63.

Annexes

1. Analyse de l'impact des cours disciplinaires sur la réussite des étudiants

1.1. Taux de réussite des étudiants des Baccalauréats III

****

1. Générer la variable réussite

****

encode Reussite, generate(REUSSITE)

encode Maths, generate(Mathématique)

encode Bio, generate(Biologie)

encode PT, generate(Physique_Technologie)

encode Geo, generate(Géographie)

**********2. Graphique**********

input Fonc

1

2

end

gen Mod1=Fonc in 1/2

recode Mod1(1=1)(2=2)

lab def Mod1 1"R" 2"NR"

lab val Mod1 Mod1

tab Mod1

****

gen Mo=Fonc in 1/2

recode Mo(1=1)(2=2)

lab def Mo 1"a Réussi" 2"n'a pas Réussi"

lab val Mo Mod

tab Mod

input Mathématiques

0.06

0.94

end

input Biologie_Chimie

0.06

0.94

end

input Physiques_Technologie

0.54

0.46

end

input Géographique

0.57

0.43

end

****

graph hbar Mathématiques Biologie_Chimie Physiques_Technologie Géographique, over(Mod1) graphregion(color(white)) legend(on) ylabel(0 "0" .2 "20%" .4 "40%" .6 "60%" .8 "80%" 1 "100%") ///

graphregion (color(white)) blabel(bar, color(white) position(inside) format(%10.2f))

2. Commandes pour l'analyse de l'influence des cours disciplinaires à la réussite

2. 1. Section Physique-Technologie

******1.1.Catégorisation des variables*****

gen RESS= REUSSITE

recode RESS(0=1)(1=0)

lab def RESS 0"Réussi" 1"échoué"

lab val RESS RESS

tab RESS

lab def CRISTA 1"OUI" 0"NON"

lab val CRISTA CRISTA

codebook CRISTA

lab def PHES 1"OUI" 0"NON"

lab val PHES PHES

codebook PHES

lab def MEQUA 1"OUI" 0"NON"

lab val MEQUA MEQUA

codebook MEQUA

tab MEQUA

lab def MEANIII 1"OUI" 0"NON"

lab val MEANIII MEANIII

codebook MEANIII

tab MEANIII

lab def OPPH 1"OUI" 0"NON"

lab val OPPH OPPH

codebook OPPH

tab OPPH

lab def MAIE 1"OUI" 0"NON"

lab val MAIE MAIE

codebook MAIE

tab MAIE

lab def TEES 1"OUI" 0"NON"

lab val TEES TEES

codebook TEES

tab TEES

tab TEES

lab def TPE 1"OUI" 0"NON"

lab val TPE TPE

codebook TPE

tab TPE

lab def DEIN 1"OUI" 0"NON"

lab val DEIN DEIN

codebook DEIN

lab def DEBA 1"OUI" 0"NON"

lab val DEBA DEBA

codebook DEBA

tab DEBA

lab def GEDE 1"OUI" 0"NON"

lab val GEDE GEDE

codebook GEDE

tab GEDE

lab def HIPH 1"OUI" 0"NON"

lab val HIPH HIPH

codebook HIPH

tab HIPH

lab def REGE 1"OUI" 0"NON"

lab val REGE REGE

codebook REGE

tab REGE

1.2.**Test du chi-carré**

tab RESS CRISTA, chi2

tab RESS PHES, chi2

tab RESS MEQUA, chi2

tab RESS MEANIII, chi2

tab RESS OPPH, chi2

tab RESS MAIE, chi2

tab RESS TEES, chi2

tab RESS TPE , chi2

tab RESS DEIN, chi2

tab RESS DEBA, chi2

tab RESS GEDE, chi2

tab RESS HIPH, chi2

tab RESS REGE, chi2

*****1.3.Odds ratio******

tabodds RESS CRISTA,or

tabodds RESS PHES, or

tabodds RESS MEQUA, or

tabodds RESS MEANIII,or

tabodds RESS OPPH, or

tabodds RESS MAIE, or

tabodds RESS TEES, or

tabodds RESS TPE , or

tabodds RESS DEIN, or

tabodds RESS DEBA , or

tabodds RESS GEDE, or

tabodds RESS HIPH, or

tabodds RESS REGE, or

*****1.4. Etude du modèle******

logit RESS PHES MEQUA TEES REGE

logit RESS PHES MEQUA TEES

logit RESS ib(0).PHES ib(0).MEQUA ib(0).TEES

logit $ylist $xlist ,or

quietly logit $ylist $xlist

estat ic

******1.4.1.Modèle retenu , test de wald et ROC******

logit RESS ib(0).PHES ib(0).MEQUA ib(0).TEES

logit $ylist $xlist ,or

quietly logit $ylist $xlist ,or

test PHES MEQUA TEES

lroc

*****1.4.2.The Hosmer and Lemeshow test (adéquation)******

estat gof

2.2. Section des sciences géographiques

*****2.1.Catégorisation des variables*****

gen RESS= REUSSITE

recode RESS(0=1)(1=0)

lab def RESS 0"Réussi" 1"échoué"

lab val RESS RESS

tab RESS

lab def GEOGEA 1"OUI" 0"NON"

lab val GEOGEA GEOGEA

codebook GEOGEA

tab GEOGEA

lab def ANDEMO 1"OUI" 0"NON"

lab val ANDEMO ANDEMO

codebook ANDEMO

tab ANDEMO

lab def GEOBU 1"OUI" 0"NON"

lab val GEOBU GEOBU

codebook GEOBU

tab GEOBU

lab def GEOTRA 1"OUI" 0"NON"

lab val GEOTRA GEOTRA

codebook GEOTRA

tab GEOTRA

lab def ENVDEDU 1"OUI" 0"NON"

lab val ENVDEDU ENVDEDU

codebook ENVDEDU

tab ENVDEDU

lab def GEOAM 1"OUI" 0"NON"

lab val GEOAM GEOAM

codebook GEOAM

tab GEOAM

lab def GEOEURO 1"OUI" 0"NON"

lab val GEOEURO GEOEURO

codebook GEOEURO

tab GEOEURO

lab def GEOASI 1"OUI" 0"NON"

lab val GEOASI GEOASI

codebook GEOASI

tab GEOASI

lab def GEOREA 1"OUI" 0"NON"

lab val GEOREA GEOREA

codebook GEOREA

tab GEOREA

lab def VMB 1"OUI" 0"NON"

lab val VMB VMB

codebook VMB

tab VMB

lab def GEOOCE 1"OUI" 0"NON"

lab val GEOOCE GEOOCE

codebook GEOOCE

tab GEOOCE

lab def TOUANV 1"OUI" 0"NON"

lab val TOUANV TOUANV

codebook TOUANV

tab TOUANV

*****2.2.Test du chi-carré******

tab RESS GEOBU,chi2

tab RESS GEOGEA,chi2

tab RESS ANDEMO,chi2

tab RESS ENVDIDU, chi2

tab RESS TOUANV, chi2

tab RESS GEOAM,chi2

tab RESS GEOEURO,chi2

tab RESS GEOASI,chi2

tab RESS GEOOCE ,chi2

tab RESS GEOREA,chi2

tab RESS VMB, chi2

tab RESS GEOTRA,chi2

*****2.3. 0dds ratio*****

tabodds RESS GEOBU,or

tabodds RESS GEOGEA,or

tabodds RESS ANDEMO,or

tabodds RESS ENVDIDU,or

tabodds RESS TOUANV, or

tabodds RESS GEOAM,or

tabodds RESS GEOEURO,or

tabodds RESS GEOASI,or

tabodds RESS GEOOCE ,or

tabodds RESS GEOREA,or

tabodds RESS VMB, or

tabodds RESS GEOTRA,or

*****2.4. Etude du modèle*****

logit RESS GEOGEA ENVDIDU GEOAM GEOEURO VMB

logit $ylist $xlist ,or

estat ic

logit RESS ib(0).GEOGEA ib(0).ENVDIDU ib(0).GEOEURO ib(0).VMB

logit $ylist $xlist ,or

estat ic

logit RESS ib(0).GEOGEA ib(0).GEOEURO ib(0).VMB ib(0).ENVDIDU

logit $ylist $xlist ,or

estat ic

logit RESS ib(0).GEOGEA ib(0).GEOEURO ib(0).ENVDIDU

logit $ylist $xlist ,or

estat ic

logit RESS ib(0).GEOGEA ib(0).GEOEURO ib(0).ENVDIDU

logit $ylist $xlist ,or

estat ic

*****2 .4.1.Modèle retenu , test de wald et courbe ROC*****

logit RESS GEOGEA GEOEURO ENVDIDU

logit RESS GEOGEA GEOEURO ENVDIDU

logit $ylist $xlist ,or

quietly logit $ylist $xlist ,or

test GEOGEA GEOEURO ENVDIDU

lroc

*****2.4.2.The Hosmer and Lemeshow test (adéquation)*****

estat gof

2. 3. Section biologie chimie

*****3.1. Catégorisation des variables*****

gen RESS= REUSSITE

recode RESS(0=1)(1=0)

lab def RESS 0"Réussi" 1"échoué"

lab val RESS RESS

tab RESS

lab def GERE 1"OUI" 0"NON"

lab val GERE GERE

codebook GERE

tab GERE

lab def EMAN 1"OUI" 0"NON"

lab val EMAN EMAN

codebook EMAN

tab EMAN

lab def INSCEN 1"OUI" 0"NON"

lab val INSCEN INSCEN

codebook INSCEN

tab INSCEN

ab def ALNU 1"OUI" 0"NON"

lab val ALNU ALNU

codebook ALNU

tab ALNU

lab def PMI 1"OUI" 0"NON"

lab val PMI PMI

codebook PMI

tab PMI

lab def MICROB 1"OUI" 0"NON"

lab val MICROB MICROB

codebook MICROB

tab MICROB

lab def CHICH 1"OUI" 0"NON"

lab val CHICH CHICL

codebook CHICL

tab CHICL

lab def CHIND 1"OUI" 0"NON"

lab val CHIND CHIND

codebook CHIND

tab CHIND

lab def BIME 1"OUI" 0"NON"

lab val BIME BIME

codebook BIME

tab BIME

lab def TOXEC 1"OUI" 0"NON"

lab val TOXEC TOXEC

codebook TOXEC

tab TOXEC

*****3.2.Test du chi-carré*****

tab RESS GERE,chi2

tab RESS EMAN, chi2

tab RESS INSCEN,chi2

tab RESS ALNU, chi2

tab RESS PMI , chi2

tab RESS MICROB,chi2

tab RESS CHICL, chi2

tab RESS CHIND, chi2

tab RESS BIME, chi2

tab RESS TOXEC, chi2

*****3.3. Odds ratio*****

tabodds RESS GERE,or

tabodds RESS EMAN, or

tabodds RESS INSCEN,or

tabodds RESS ALNU, or

tabodds RESS PMI , or

tabodds RESS MICROB,or

tabodds RESS CHICL, or

tabodds RESS CHIND, or

tabodds RESS BIME, or

tabodds RESS TOXEC,or

2.4. Section Mathématique

*****4.1.Catégorisation des variables*****

gen RESS2= REUSSITE

recode RESS2(0=1)(1=0)

lab def RESS2 0"Réussi" 1"échoué"

lab val RESS2 RESS2

tab RESS2

lab def ANCO 1"OUI" 0"NON"

lab val ANCO ANCO

codebook ANCO

tab ANCO

lab def EDP 1"OUI" 0"NON"

lab val EDP EDP

codebook EDP

tab EDP

lab def ASTRO 1"OUI" 0"NON"

lab val ASTRO ASTRO

codebook ASTRO

tab ASTRO

lab def ANNU 1"OUI" 0"NON"

lab val ANNU ANNU

codebook ANNU

tab ANNU

lab def GEDI 1"OUI" 0"NON"

lab val GEDI GEDI

codebook GEDI

tab GEDI

lab def GEPR 1"OUI" 0"NON"

lab val GEPR GEPR

codebook GEPR

tab GEPR

lab def CASTO 1"OUI" 0"NON"

lab val CASTO CASTO

codebook CASTO

tab CASTO

lab def LAPRO 1"OUI" 0"NON"

lab val LAPRO LAPRO

codebook LAPRO

tab LAPRO

lab def SYDY 1"OUI" 0"NON"

lab val SYDY SYDY

codebook SYDY

tab SYDY

lab def ALCO 1"OUI" 0"NON"

lab val ALCO ALCO

codebook ALCO

tab ALCO

lab def ALBO 1"OUI" 0"NON"

lab val ALBO ALBO

codebook ALBO

tab ALBO

*****4.2.Test du chi2*****

tab RESS GEDI,chi2

tab RESS CASTO, chi2

tab RESS GEPR, chi2

tab RESS LAPRO, chi2

tab RESS SYDY, chi2

tab RESS ALCO, chi2

tab RESS ALBO,chi2

tab RESS ANCO, chi2

tab RESS EDP , chi2

tab RESS ASTRO ,chi2

tab RESS ANNU, chi2

*****4.3. 0dds ratio*****

tabodds RESS GEDI,or

tabodds RESS CASTO, or

tabodds RESS GEPR, or

tabodds RESS LAPRO, or

tabodds RESS SYDY, or

tabodds RESS ALCO, or

tabodds RESS ALBO,or

tabodds RESS ANCO, or

tabodds RESS EDP ,or

tabodds RESS ASTRO ,or

tabodds RESS ANNU,or