



Université Sidi Mohamed Ben Abdellah
Faculté des Sciences et Techniques de Fès

MÉMOIRE DE FIN D'ÉTUDES

MATHÉMATIQUES APPLIQUÉS

Traitement statistique des résultats scolaires dans la province de Taza

Author:

Najlae SGHIOURI IDRISSE

Supervisor:

Pr. Azzeddine EL BARAKA

LIEU DE STAGE : DÉLÉGATION DU MINISTÈRE DE L'ÉDUCATION TAZA

SOUTENU LE 16 JUIN 2015 DEVANT LE JURY COMPOSÉ DE :

Pr. F Ezzaki

Pr. A El Baraka

Pr. M Akhmouch

Pr. O Sidki

Année universitaire : 2014/2015

Dédicace :

Je dédie ce mémoire à

Ma mère, qui depuis ma naissance, n'a cessé de me bercer avec des bons conseils et, sans doute, celle qui me réconforte et m'encourage dans les moments les plus difficiles de ma scolarité.

Mon père, celui qui ne se lassait jamais pour aménager tous ses efforts pour que je sois quelqu'un dans la vie. Pour lui, l'instruction est le socle fondamental de ses enfants.

Mes frères, qui n'ont jamais été absents dans les moments turbulents de ma scolarité, vos apports me sont inestimables.

Remerciements :

L'élaboration du présent travail a requis le soutien de nombreuses personnalités sans lesquelles l'accès aux ressources nécessaires n'aurait pas été possible.

Pour cette raison, je tiens à remercier tous ceux qui, de près ou de loin se sont impliqués dans la réalisation de ce mémoire.

Mes plus vifs remerciements s'adressent tout d'abord à mon encadrant , ***Mr Azzeddine El Baraka***, qui n'a cessé de m'apporter son soutien.

J'exprime mon extrême gratitude à ***Mme F.Ezzaki*** pour son encouragement, ses conseils, et son appui moral.

Je tiens à remercier vivement ***Mr Azouz Ziane***, chef du service de planification au sein de la délégation de Taza, ainsi que tout son staff.

Enfin, je ne peux pas achever cette page sans adresser tout particulièrement mes sincères remerciements à ***Mr M.Khihel***, pour son apport inestimable à la réalisation de ce mémoire.

Les personnes qui m'ont apporté de l'aide sont très nombreuses et je vous assure de ma profonde gratitude.

Contents

| | | |
|------------|---|-----------|
| I | Introduction | 4 |
| II | Présentation de la délégation de Taza | 6 |
| 1 | L'organigramme de la délégation | 7 |
| 2 | Cinq Bureaux Indépendants | 8 |
| 3 | Six Services Principaux | 8 |
| III | Éléments De Cours | 12 |
| 4 | L'échantillonnage Systématique | 13 |
| 5 | Les Statistiques Inférentielles | 14 |
| 6 | Définitions | 14 |
| 6.1 | Hypothèses nulle et alternative | 14 |
| 6.2 | les tests paramétriques | 15 |
| 6.2.1 | Comparaison de deux moyennes : | 15 |
| 6.2.2 | Comparaison de deux proportions | 18 |
| 6.2.3 | Comparaison de plusieurs proportions | 19 |
| IV | Collecte et traitement des données | 22 |
| 7 | Étude comparative des notes scolaires | 23 |
| 7.1 | Par milieu | 23 |
| 7.1.1 | Urbain | 26 |
| 7.1.2 | Rural | 29 |
| 7.2 | Par genre | 30 |
| 7.2.1 | Rural | 30 |
| 7.2.2 | Urbain | 35 |
| 7.3 | Par matière | 40 |
| 7.3.1 | Urbain | 41 |
| 7.3.2 | Rural | 43 |
| 8 | Évolution des taux de réussite du baccalauréat | 45 |
| 8.1 | Enseignement Général | 45 |
| 8.2 | Enseignement Originel | 47 |
| 8.3 | Enseignement Techniques Commerciales | 48 |
| 8.4 | Enseignement Techniques Industrielles | 50 |
| V | Conclusion | 52 |

Part I
Introduction

Le service de planification dont lequel j'ai été affectée au sein de la délégation du ministère de l'éducation nationale de Taza, et qui contient le bureau de recensement souhaitaient

savoir pas mal d'informations concernant les moyennes des élèves de la 3^{ème}

année collégiale ainsi que les taux de réussite du baccalauréat durant les 5 dernières années. Ils se posent les questions suivantes :

- Est ce que les moyennes du monde urbain sont supérieures à celles du monde rural ?
- Il est comment le rendement des filles par rapport à celui des garçons ?
- Est ce que également les élèves ont une tendance vers les matières scientifiques ou bien littéraires ?
- Le taux de réussite du baccalauréat a t-il évolué durant ces cinq dernières années ?

Le rapport en main essaiera de répondre à ces questions, en utilisant surtout les statistiques inférentielles.

La première partie de ce rapport présentera d'une façon générale la délégation du ministère de l'éducation nationale de Taza, avec tout ses bureaux et services principaux.

Dans la deuxième partie, j'ai ajouté quelques éléments de cours de statistique qui me seront utile pour bien traiter les données en main.

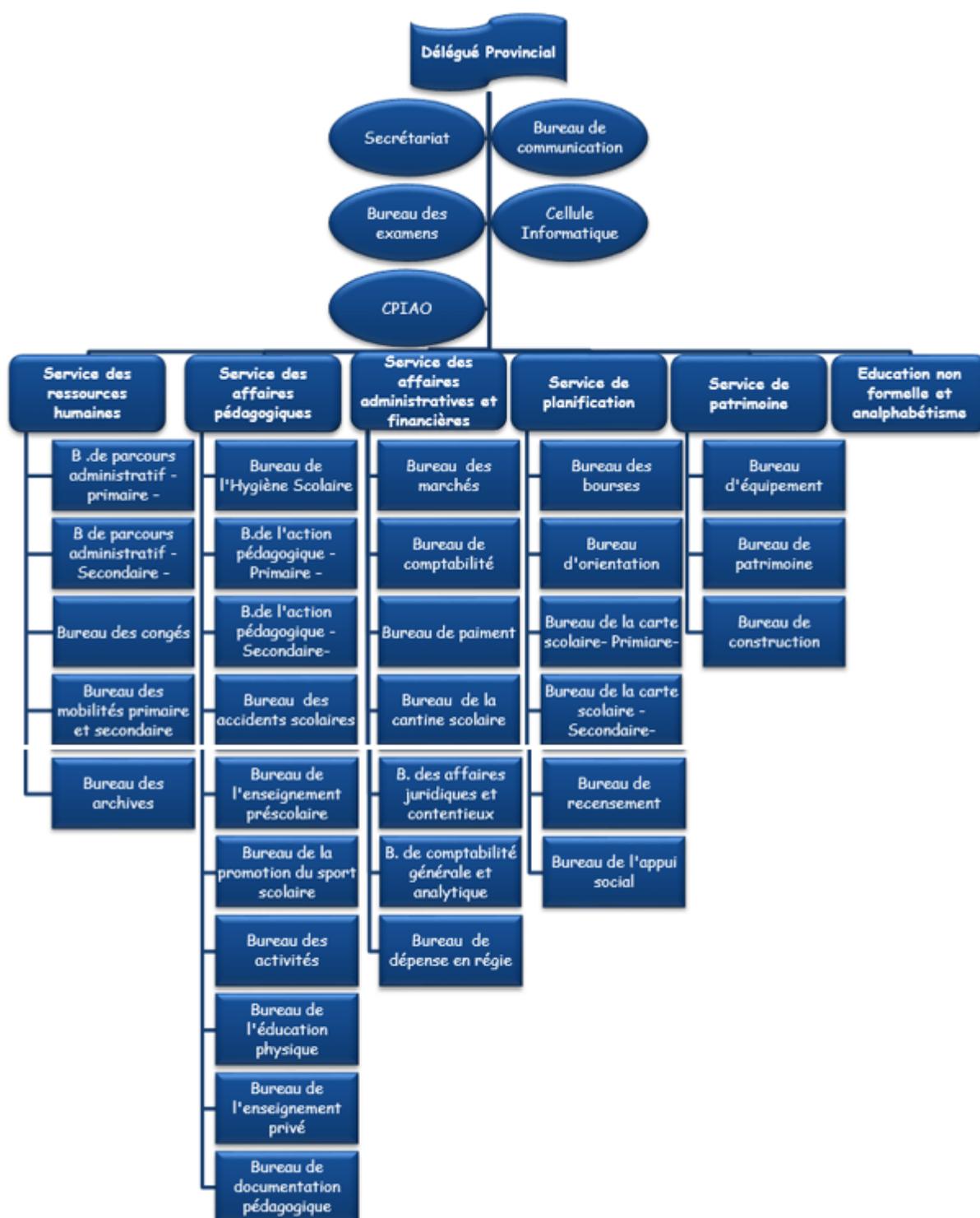
L'objectif de la troisième partie, est la collecte des données, leur analyse et ordonnancement.

Finalement, une petite conclusion qui résumera tout les résultats de ce traitement.

Part II

Présentation de la délégation de Taza

1 L'organigramme de la délégation



La structure administrative de la délégation est subdivisée comme suit:

2 Cinq Bureaux Indépendants

- Secrétariat.
- Cellule informatique.
- Bureau de communication.
- Bureau des examens.
- Centre Provincial de l'informatique et Aide à l'orientation.

3 Six Services Principaux

- Service des ressources humaines.
- Service des affaires pédagogique.
- Service de planification.
- Service des affaires administratives et financières.
- Service Bâtiment Et Équipement.
- Éducation Non Formelle et Analphabétisme

Les Bureaux Indépendants :

Le délégué provincial:

Le délégué provincial est le premier responsable de la délégation, c'est le représentant du ministère de l'éducation nationale dans la province.

Sa signature doit figurer dans tous les documents administratifs.

2- Secrétariat:

Ce bureau fait l'intermédiaire entre les services existants dans la délégation et le délégué, il s'occupe des relations externes entre la délégation et le ministère ou l'académie en cas d'absence du délégué. Parmi ses tâches :

- Classement des données administratives pédagogiques.

- Le stock des informations.
- Classement des dossiers.
- Gérer la correspondance du délégué.

3- La cellule informatique :

Cette cellule a été créée vers la fin des années 80 pour le but d'informatiser les différents services de la délégation.

La cellule informatique est équipée par différent matériel et logiciel qui lui permettent d'accomplir différentes tâches comme .

- Traitement des applications informatiques.
- Stockage des informations.
- Entretien des ordinateurs.

4- Bureau de communication et des partenariats :

Le rôle de ce bureau est la coordination interne (entre les différents services de la délégation) et externe (avec les services extérieurs d'état, la presse, les ONG ...), il veille aussi sur la conclusion de partenariats avec les différents acteurs.

6- Bureau des examens :

L'objectif de ce bureau est la gestion des examens et des concours qui se déroulent dans la province de TAZA, ainsi que l'authentification des relevés de notes du Baccalauréat.

Les services principaux :

1- Service des affaires pédagogique:

Via un travail synergique entre ses différents bureaux, il veille sur le bon déroulement de l'opération enseignement-apprentissage dans les différents établissements scolaires publiques et privés, il assure la diffusion et l'application des notes ministérielles en vigueur ainsi que la valorisation de la vie scolaire.

2- Service des ressources humains :

Le service des ressources humaines est rattaché à la direction des ressources humaines. Il gère les dossiers administratifs du personnel, élabore des plans d'action en vue d'améliorer les conditions de travail.

- Élaborer les stratégies d'emploi et des carrières en vue du développement des ressources humaines;
- Assurer la gestion administrative intégrée des personnels enseignants administratifs et techniques;
- Veiller à un meilleur déploiement des ressources humaines;
- Assurer la gestion de la base des données des ressources humaines et des carrières;
- Contribuer aux actions de promotion des œuvres sociales
- Élaborer la politique éducative nationale en matière de formation des cadres;
- Encourager et suivre la recherche pédagogique dans les établissements de formation des cadres;

3- Service de planification :

Le service de la planification a pour missions :

- Élaborer les études prospectives portant sur le système éducatif.

- Élaborer le plan provincial de développement de la scolarisation et de veiller à sa mise en œuvre;
- Réaliser les enquêtes et analyses statistiques;

4- Service des affaires administratives et financières :

Ce service prépare le budget de la gestion financière et matérielle de la délégation en coordination avec les chefs des autres services en présence du délégué.

Il détermine ainsi tous les besoins concernant les établissements primaires et secondaires en matériel didactique, fourniture de bureau, fournitures sanitaires, cantines scolaires, internats ...

5- Service de patrimoine :

Ce service s'occupe de la programmation des constructions prévus par le service de la planification.

- L'aménagement des établissements scolaires et des différents services de la délégation du matériel nécessaire pour la gestion éducative et administrative.
- La construction, la réparation et l'entretien des bâtiments scolaires.

6- Service de l'éducation non formelle et Analphabétisme :

Ce service veille sur la mise en œuvre des programmes du ministère visant la lutte contre l'analphabétisme et la vulgarisation de l'offre éducative.

Part III
Éléments De Cours

4 L'échantillonnage Systématique

L'échantillonnage systématique est une méthode qui exige aussi l'existence d'une liste de la population où chaque individu est numéroté de 1 jusqu'à N .

Avantages :

La sélection des individus se fait d'une façon aléatoire.

On peut obtenir une bonne précision parce que la méthode permet de répartir l'échantillon dans l'ensemble de la liste.

Inconvénients :

Les données peuvent être biaisées à cause de la périodicité.

Notons n , le nombre d'individus que doit comporter l'échantillon (la taille de l'échantillon).

L'entier voisin de $\frac{N}{n}$ sera noté r et appelé *raison de sondage* ou *pas de sondage*.

Pour constituer l'échantillon il faut :

Choisir au hasard un entier naturel k entre 1 et r (cet entier sera le point de départ).

L'individu dont le numéro correspond à k est le premier individu de l'échantillon.

Pour sélectionner les autres, il suffit d'ajouter à k la raison de sondage : les individus choisis seront alors ceux dont les numéros correspondent à

- $k + r$
- $k + 2r$
- $k + 3r$
- etc...

5 Les Statistiques Inférentielles

Confronté à des phénomènes complexes et aléatoires, la prise de décision est difficile et les outils adaptés de la théorie des tests ont pour objet de guider les choix entre différentes alternatives.

De façon générale, il s'agira de décider si des différences observées entre un modèle posé a priori et des observations sont significatives ou peuvent être considérées comme étant dues au simple effet du hasard consécutif aux aléas du tirage d'un échantillon.

Réaliser un test statistique consiste à mettre en oeuvre une procédure permettant :

- de confronter une hypothèse avec la réalité, ou plus exactement, avec ce que l'on perçoit de la réalité à travers les observations à disposition ;
- de prendre une décision à la suite de cette confrontation.

Si les problèmes traités par l'estimation (ponctuelle ou par intervalle de confiance) sont de type quantitatif, i.e. conduisent à un résultat numérique, ceux traités par les tests d'hypothèses sont d'ordre qualitatif, i.e. conduisent à une réponse du type rejet/acceptation de l'hypothèse statistique considérée.

6 Définitions

6.1 Hypothèses nulle et alternative

Un *test paramétrique* est une procédure de décision entre deux hypothèses concernant la valeur des paramètres d'un modèle probabiliste.

L'*hypothèse nulle* notée H_0 est celle que l'on considère vraie à priori. Le but du test est de décider si cet à priori est crédible.

L'*hypothèse alternative* notée H_1 est l'hypothèse complémentaire de H_0 .

- Pour un test bilatéral, nous pouvons émettre les hypothèses suivantes :

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

- Pour un test unilatéral, les hypothèses deviennent :

$$\begin{aligned}
H_0 &: \mu_1 = \mu_2 \\
H_1 &: \mu_1 < \mu_2 \quad \text{Test unilatéral à gauche} \\
H_1 &: \mu_1 > \mu_2 \quad \text{Test unilatéral à droite}
\end{aligned}$$

Remarques :

Les signes =, \neq , > et < ne correspondent pas à l'égalité ou aux inégalités au sens mathématique du terme.

Il s'agit d'une façon succincte d'écrire :

$$\begin{aligned}
H_0 &: \text{Il est crédible de penser que } \mu_1 = \mu_2 \\
H_1 &: \mu_1 \text{ est significativement différent de } \mu_2
\end{aligned}$$

H_0 et H_1 ne jouent pas le même rôle. En fait, on fait tout le raisonnement sous H_0 et H_1 n'est choisi que par défaut si H_0 ne convient pas.

6.2 les tests paramétriques

6.2.1 Comparaison de deux moyennes :

L'objectif dans cette section est de comparer des caractéristiques de deux populations à l'aide d'échantillons extraits de chacune de celle-ci.

Deux échantillons indépendants Cette section considère deux variables X et Y gaussiennes de moyennes $(\mu_1; \mu_2)$ et de variances $(\sigma_X^2; \sigma_Y^2)$ dont sont tirés des échantillons d'effectifs $(n_1; n_2)$; \bar{X} (resp. \bar{Y}) désigne toujours l'estimateur de l'espérance de la variable X (resp. Y) et S_1^2 et S_2^2 celui de la variance. Notons également :

$$S_p^2 = \frac{(n_1 - 1).S_1^2 + (n_2 - 1).S_2^2}{n_1 + n_2 - 2}$$

S_p^2 est appelée : Variance Commune.

Cas 1 : σ_1^2 et σ_2^2 Connues

Le test de l'hypothèse $H_0 : \mu_1 = \mu_2$ est basé sur la loi des différences des moyennes. Sachant que X et Y sont des variables aléatoires gaussiennes $X \rightsquigarrow \mathcal{N}(\mu_1, \frac{\sigma_1^2}{n_1})$

$$Y \rightsquigarrow \mathcal{N}(\mu_2, \frac{\sigma_2^2}{n_2})$$

La loi de la variable $(\bar{X} - \bar{Y})$ étant gaussienne on a $\bar{X} - \bar{Y} \rightsquigarrow \mathcal{N}(0, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

Tester $\mu_1 = \mu_2$ revient donc à tester la moyenne $\mu_1 - \mu_2 = 0$ d'une variable aléatoire $(\bar{X} - \bar{Y})$ Gaussienne et de variance connue comme dans le cas d'un seul échantillon.

Cas 1 : σ_1^2 et σ_2^2 Inconnues mais supposées égales

Cette hypothèse d'égalité des variances peut être vérifiée avec le test ci-dessous. Dans ce cas, la variable aléatoire $(\bar{X} - \bar{Y})$ est Gaussienne

$$\bar{X} - \bar{Y} \rightsquigarrow \mathcal{N}(0, \sigma^2 (\frac{1}{n_1} + \frac{1}{n_2}))$$

Dans ce cas σ^2 doit être estimée par S^2 et la statistique de test est le quotient d'une variable gaussienne par une variable du Khi-Deux ; c'est une variable de Student à $n_1 + n_2 - 2$ degrés de liberté.

$$\frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightsquigarrow \mathcal{T}_{n_1+n_2-2}$$

Tester $\mu_1 = \mu_2$ revient donc à tester la moyenne $\mu_1 - \mu_2 = 0$ d'une variable aléatoire $(\bar{X} - \bar{Y})$ Gaussienne et de variance connue comme dans le cas d'un seul échantillon.

Remarque :

$\sigma_1^2 \neq \sigma_2^2$ inconnues :

Lorsque les effectifs des deux échantillons sont élevés (supérieurs à 30), le test précédent peut être utilisé.

Test de student :

Ce test permet de comparer :

- Une moyenne d'un échantillon à une valeur donnée
- Les moyennes de deux échantillons indépendants

- Les moyennes de deux échantillons appariés.

L'emploi de ce test reste subordonné en général à deux conditions d'application importantes qui sont *la normalité et le caractère aléatoire et simple des échantillons*.

La première condition n'est toutefois pas essentielle lorsque les échantillons ont des effectifs suffisants (en pratique, la valeur de 30 est souvent retenue) pour assurer la quasi-normalité des distributions d'échantillonnage des moyennes. En plus, de ces deux conditions, nous devons supposer, dans certains tests relatifs aux moyennes, *l'égalité des variances des échantillons considérées*.

Conditions d'application du test de Student : Le caractère de l'échantillon étant supposé aléatoire, l'hypothèse de normalité de la variable X doit être vérifiée (par exemple) avec le test de Kolmogorov-Smirnov si $n < 30$.

Calcul :

Soit X une variable aléatoire distribuée selon une loi normale, la variable aléatoire définie ci-dessus suit une loi de Student avec $n - 1$ degrés de liberté.

Conditions d'application du test :

- Les deux échantillons sont indépendants entre eux, sont aléatoires et ont n_1 et n_2 unités indépendantes (cette condition est d'ordinaire satisfaite en utilisant une procédure de randomisation ; procédure pour laquelle on affecte au hasard chaque individu à un groupe expérimental).
- La variable aléatoire suit une loi normale ou elle a des effectifs supérieurs à 30.
- Il est aussi nécessaire de vérifier l'égalité des variances des échantillons (grâce au test de Fisher). **Cette condition est indispensable pour des effectifs inégaux.**

Remarques :

Plusieurs auteurs ont montré que **l'hypothèse de normalité est d'importance relativement secondaire dans le test d'égalité de deux moyennes lorsque $n > 30$** . En effet, dans certaines limites, la non-normalité des populations ne modifie pas sensiblement les risques d'erreur de première et deuxième espèce. Ceci est vrai surtout pour les distributions symétriques, même très différentes des distributions normales.

De même, l'hypothèse d'égalité des variances n'est pas fondamentale au point de vue pratique lorsque les effectifs des échantillons sont égaux.

En raison de cette faible sensibilité du test à la non-normalité et à l'inégalité des variances, on dira qu'il s'agit, pour des effectifs égaux, **d'un test robuste**. Par contre, **lorsque les effectifs des échantillons sont inégaux, il est absolument indispensable de s'assurer de l'égalité des variances** et, si cette hypothèse n'est pas vérifiée, il est indispensable d'utiliser une méthode adaptée à ces circonstances. On peut notamment procéder à une transformation de variable, destinée à stabiliser les variances, et utiliser ensuite le test de Student.

Comparaison de deux variances : Fisher

Il est utile de tester l'égalité de deux variances préalablement au test d'égalité des moyennes. La statistique de test est $\frac{S_1^2}{S_2^2}$ qui ne doit être ni trop petite ni trop grande afin d'accepter H_0 . Ce rapport de deux variables suivant des lois du Khi-deux est une variable aléatoire suivant une loi de Fisher à $(n_1 - 1)$ et $(n_2 - 1)$ degrés de liberté.

$$\frac{S_1^2}{S_2^2} \rightsquigarrow \mathcal{F}_{n_1-1, n_2-1}$$

En pratique, la plus grande quantité, disons S_1^2 est placée au numérateur et le test de niveau α conduit à rejeter H_0 si la réalisation de la statistique $\frac{S_1^2}{S_2^2}$ est plus grande que le quantile $\mathcal{F}_{n_1-1, n_2-1; \frac{\alpha}{2}}$ de la loi de Fisher.

6.2.2 Comparaison de deux proportions

Ici les variables X et Y ne sont plus supposées gaussiennes mais binomiales de paramètres $(n_1 ; p_1)$ et $(n_2 ; p_2)$. Le test qui suit n'est valide qu'approximativement pour des échantillons de tailles n_1 et n_2 suffisamment grandes car il est basée sur des approximations gaussiennes des lois binomiales.

- Hypothèses :

$$\begin{aligned} H_0 &: p_1 = p_2 \\ H_1 &: p_1 \neq p_2 \end{aligned}$$

Sous $H_0 : p_1 = p_2$ on obtient une estimation de la valeur p par :

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

avec $s(\hat{p}_1 - \hat{p}_2) = \sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

Sous H_0 et les conditions d'application du test : $n_1 \cdot \hat{p} \gtrsim 5 ; n_1 \cdot (1 - \hat{p}) \gtrsim 5$

et $n_2 \cdot \hat{p} \gtrsim 5 ; n_2 \cdot (1 - \hat{p}) \gtrsim 5$

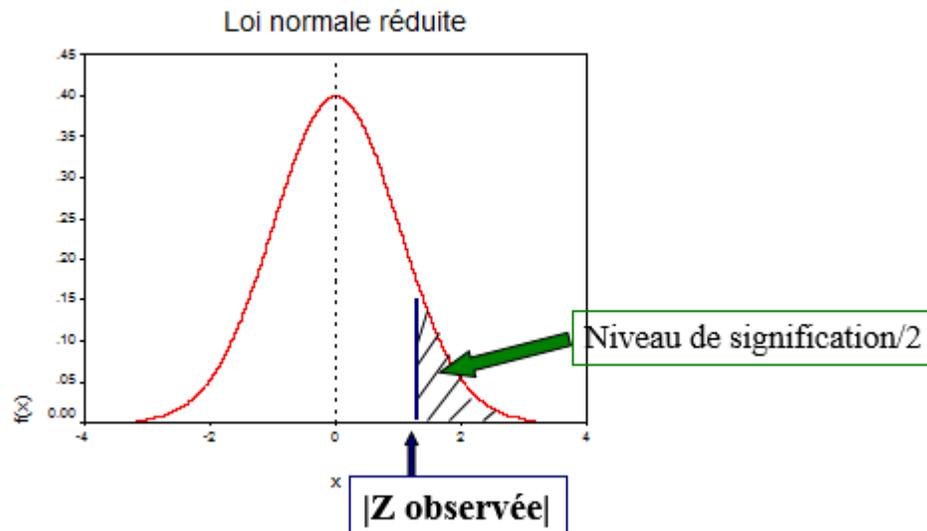
On obtient alors comme statistique du test :

$$Z_{exp} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \rightsquigarrow \mathcal{N}(0, 1)$$

Règle de décision :

Rejeter H_0 si $Z_{exp} \notin [-Z_{\alpha/2}; +Z_{\alpha/2}]$

Niveau de signification du Z observé



6.2.3 Comparaison de plusieurs proportions

Considérons p populations (colonnes). Pour chaque objet (élément), on a noté la valeur d'une variable qualitative comportant k modalités. Le tableau de contingence suivant contient les fréquences absolues observées f_{oij} :

| Modalités | Populations | | | | Σ |
|-----------|-------------|----------|-------|----------|----------|
| | 1 | 2 | j | p | |
| 1 | f_{11} | f_{12} | ... | f_{1p} | m_1 |
| 2 | f_{21} | f_{22} | ... | f_{2p} | m_2 |
| 3 | f_{31} | f_{32} | ... | f_{3p} | m_3 |
| i | f_{i1} | f_{i2} | ... | f_{ip} | m_i |
| k | f_{k1} | f_{k2} | ... | f_{kp} | m_k |
| Σ | n_1 | n_2 | n_j | n_g | n |

Hypothèses

- $H_0 : p_{i1} = p_{i2} = \dots = p_{ik} = p_i$ avec $i = 1, 2, \dots, k$
- H_1 : Les populations d'individus ne sont pas indépendantes dans les k populations pour au moins une modalité.

avec $\bar{p}_i = \sum_{j=1}^k \frac{f_{ojj}}{n}$

n : La taille totale du tableau des données.

f_{ojj} : La fréquence absolue associée au modalité i de la population j .

$\sum_{j=1}^k f_{ojj}$: La somme des fréquences absolues des k échantillons pour la modalité i .

$$f_{tij} = n_j \cdot \bar{p}_i$$

- Statistique : $\chi^2 = \sum_{i=1}^n \sum_{j=1}^k \frac{f_{ojj} - f_{tij}}{f_{tij}}$
- Calcul de la fréquence f_{tij} de chaque case :
La valeur attendue dans chaque case, sous l'hypothèse nulle, se calcule par
$$f_{tij} = \frac{n_j \cdot m_i}{n}$$
- Degré de liberté :
Le nombre de degrés de liberté associés à une statistique est le nombre de ses composantes indépendantes, i.e. le nombre des composantes de base utilisées dans le calcul de la statistique moins le nombre de relations (paramètres) qui lient celles-ci.
 $d.d.l = (k - 1)(p - 1)$

- Règle de décision :
Si $\mathcal{X}^2 > \mathcal{X}_\alpha^2$ on rejette H_0 avec un risque d'erreur de type I égal à α

Part IV

Collecte et traitement des données

7 Étude comparative des notes scolaires

7.1 Par milieu

NB : Les 3 parties sont indépendantes

Soit à comparer les moyennes des classes dans le monde rural et le monde urbain. Notons par \bar{x}_i la moyenne de chaque classe i dans le monde rural avec $i = 1, \dots, 117$ et par \bar{y}_j la moyenne de chaque classe j dans le monde urbain avec $j = 1, \dots, 173$

| | | | | | | | | | | | | | | | |
|----------------|--------|----------------|--------|----------------|--------|----------------|--------|----------------|--------|----------------|-------|-----------------|--------|-----------------|-------|
| \bar{x}_1 | 8,005 | \bar{x}_{17} | 7,910 | \bar{x}_{33} | 7,510 | \bar{x}_{49} | 6,680 | \bar{x}_{65} | 6,585 | \bar{x}_{82} | 9,120 | \bar{x}_{98} | 7,930 | \bar{x}_{114} | 7,360 |
| \bar{x}_2 | 7,170 | \bar{x}_{18} | 11,310 | \bar{x}_{34} | 9,975 | \bar{x}_{50} | 8,610 | \bar{x}_{67} | 9,755 | \bar{x}_{83} | 9,370 | \bar{x}_{99} | 9,890 | \bar{x}_{115} | 8,690 |
| \bar{x}_3 | 10,280 | \bar{x}_{19} | 11,000 | \bar{x}_{35} | 9,350 | \bar{x}_{51} | 9,340 | \bar{x}_{68} | 9,345 | \bar{x}_{84} | 9,360 | \bar{x}_{100} | 8,240 | \bar{x}_{116} | 9,250 |
| \bar{x}_4 | 8,920 | \bar{x}_{20} | 9,620 | \bar{x}_{36} | 8,825 | \bar{x}_{52} | 8,640 | \bar{x}_{69} | 7,825 | \bar{x}_{85} | 6,815 | \bar{x}_{101} | 9,460 | \bar{x}_{117} | 7,550 |
| \bar{x}_5 | 9,565 | \bar{x}_{21} | 9,970 | \bar{x}_{37} | 10,740 | \bar{x}_{53} | 9,215 | \bar{x}_{70} | 8,220 | \bar{x}_{86} | 9,045 | \bar{x}_{102} | 8,270 | | |
| \bar{x}_6 | 8,870 | \bar{x}_{22} | 8,855 | \bar{x}_{38} | 10,725 | \bar{x}_{54} | 9,00 | \bar{x}_{71} | 8,570 | \bar{x}_{87} | 8,560 | \bar{x}_{103} | 8,855 | | |
| \bar{x}_7 | 9,545 | \bar{x}_{23} | 9,465 | \bar{x}_{39} | 10,430 | \bar{x}_{55} | 7,170 | \bar{x}_{72} | 7,645 | \bar{x}_{88} | 7,490 | \bar{x}_{104} | 8,670 | | |
| \bar{x}_8 | 9,720 | \bar{x}_{24} | 8,080 | \bar{x}_{40} | 9,390 | \bar{x}_{56} | 8,275 | \bar{x}_{73} | 8,030 | \bar{x}_{89} | 8,325 | \bar{x}_{105} | 9,290 | | |
| \bar{x}_9 | 10,250 | \bar{x}_{25} | 10,020 | \bar{x}_{41} | 11,330 | \bar{x}_{57} | 8,740 | \bar{x}_{74} | 7,370 | \bar{x}_{90} | 8,075 | \bar{x}_{106} | 10,175 | | |
| \bar{x}_{10} | 9,200 | \bar{x}_{26} | 9,300 | \bar{x}_{42} | 9,840 | \bar{x}_{58} | 9,125 | \bar{x}_{75} | 8,450 | \bar{x}_{91} | 8,950 | \bar{x}_{107} | 10,260 | | |
| \bar{x}_{11} | 9,770 | \bar{x}_{27} | 9,480 | \bar{x}_{43} | 10,265 | \bar{x}_{59} | 8,860 | \bar{x}_{76} | 7,470 | \bar{x}_{92} | 8,780 | \bar{x}_{108} | 9,385 | | |
| \bar{x}_{12} | 9,635 | \bar{x}_{28} | 9,525 | \bar{x}_{44} | 9,330 | \bar{x}_{60} | 10,090 | \bar{x}_{77} | 10,615 | \bar{x}_{93} | 7,315 | \bar{x}_{109} | 10,525 | | |
| \bar{x}_{13} | 9,480 | \bar{x}_{29} | 10,645 | \bar{x}_{45} | 8,935 | \bar{x}_{61} | 10,010 | \bar{x}_{78} | 9,730 | \bar{x}_{94} | 9,790 | \bar{x}_{110} | 9,750 | | |
| \bar{x}_{14} | 10,480 | \bar{x}_{30} | 11,370 | \bar{x}_{46} | 7,880 | \bar{x}_{62} | 10,425 | \bar{x}_{79} | 10,630 | \bar{x}_{95} | 9,160 | \bar{x}_{111} | 10,355 | | |
| \bar{x}_{15} | 10,640 | \bar{x}_{31} | 10,680 | \bar{x}_{47} | 8,430 | \bar{x}_{63} | 7,215 | \bar{x}_{80} | 10,555 | \bar{x}_{96} | 8,110 | \bar{x}_{112} | 10,805 | | |
| \bar{x}_{16} | 10,560 | \bar{x}_{32} | 11,125 | \bar{x}_{48} | 9,650 | \bar{x}_{64} | 8,145 | \bar{x}_{81} | 10,680 | \bar{x}_{97} | 8,745 | \bar{x}_{113} | 10,560 | | |

La moyenne de la population est $m_1=9,204$

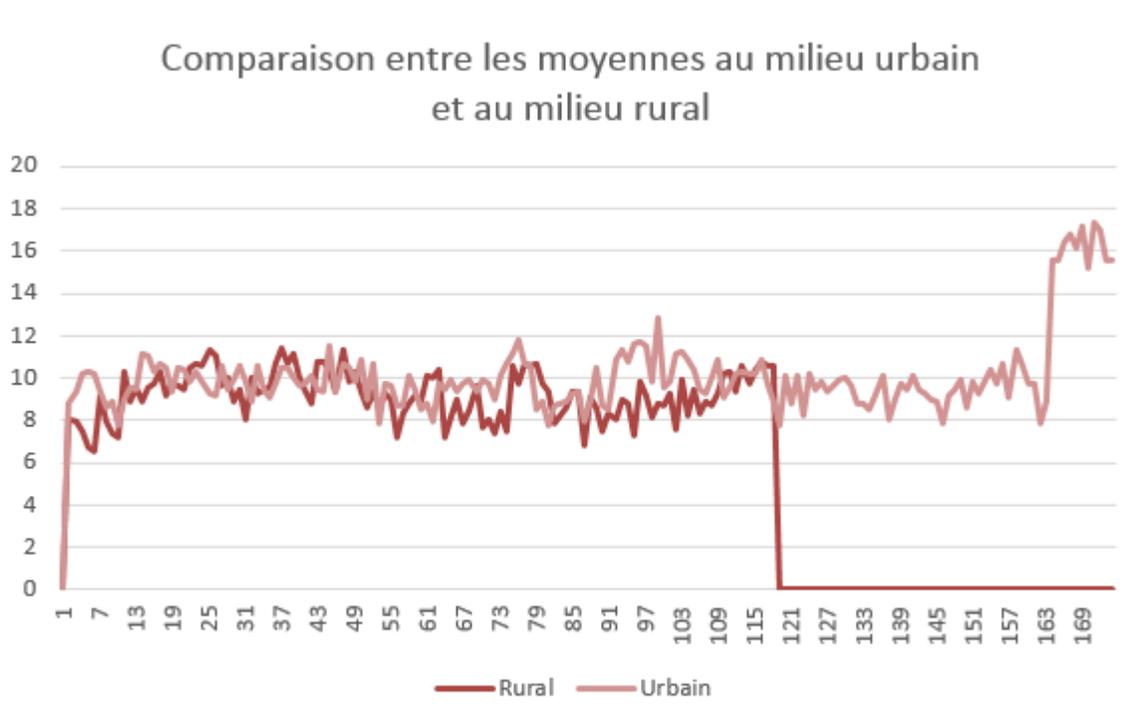
Le tableau ci-dessus regroupe les moyennes des classes des collèges du monde rural.

Le tableau ci-dessous regroupe les moyennes des classes des collèges du monde urbain.

| | | | | | | | | | | | | | | | | | | | | | |
|----------------|--------|----------------|--------|----------------|--------|----------------|--------|----------------|--------|----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|-------|
| \bar{y}_1 | 8,82 | \bar{y}_{17} | 10,47 | \bar{y}_{33} | 9,45 | \bar{y}_{49} | 10,81 | \bar{y}_{65} | 9,77 | \bar{y}_{82} | 8,75 | \bar{y}_{98} | 12,79 | \bar{y}_{114} | 10,24 | \bar{y}_{130} | 9,66 | \bar{y}_{146} | 9,20 | \bar{y}_{162} | 8,92 |
| \bar{y}_2 | 9,395 | \bar{y}_{18} | 9,375 | \bar{y}_{34} | 9,025 | \bar{y}_{50} | 8,970 | \bar{y}_{67} | 9,910 | \bar{y}_{83} | 8,905 | \bar{y}_{99} | 9,530 | \bar{y}_{115} | 10,810 | \bar{y}_{131} | 8,770 | \bar{y}_{147} | 9,400 | \bar{y}_{163} | 15,53 |
| \bar{y}_3 | 10,230 | \bar{y}_{19} | 10,490 | \bar{y}_{35} | 9,710 | \bar{y}_{51} | 10,690 | \bar{y}_{68} | 9,395 | \bar{y}_{84} | 9,210 | \bar{y}_{100} | 9,780 | \bar{y}_{116} | 9,745 | \bar{y}_{132} | 8,785 | \bar{y}_{148} | 9,930 | \bar{y}_{164} | 15,53 |
| \bar{y}_4 | 10,320 | \bar{y}_{20} | 10,340 | \bar{y}_{36} | 10,520 | \bar{y}_{52} | 7,880 | \bar{y}_{69} | 9,930 | \bar{y}_{85} | 9,310 | \bar{y}_{101} | 11,130 | \bar{y}_{117} | 8,870 | \bar{y}_{133} | 8,480 | \bar{y}_{149} | 8,620 | \bar{y}_{165} | 16,40 |
| \bar{y}_5 | 10,150 | \bar{y}_{21} | 9,805 | \bar{y}_{37} | 10,440 | \bar{y}_{53} | 9,770 | \bar{y}_{70} | 9,750 | \bar{y}_{86} | 7,960 | \bar{y}_{102} | 11,220 | \bar{y}_{118} | 7,775 | \bar{y}_{134} | 9,150 | \bar{y}_{150} | 9,800 | \bar{y}_{166} | 16,78 |
| \bar{y}_6 | 9,380 | \bar{y}_{22} | 10,300 | \bar{y}_{38} | 10,040 | \bar{y}_{54} | 9,670 | \bar{y}_{71} | 9,010 | \bar{y}_{87} | 8,880 | \bar{y}_{103} | 10,830 | \bar{y}_{119} | 10,055 | \bar{y}_{135} | 10,120 | \bar{y}_{151} | 9,280 | \bar{y}_{167} | 16,07 |
| \bar{y}_7 | 8,580 | \bar{y}_{23} | 9,805 | \bar{y}_{39} | 9,650 | \bar{y}_{55} | 8,675 | \bar{y}_{72} | 10,080 | \bar{y}_{88} | 10,470 | \bar{y}_{104} | 10,360 | \bar{y}_{120} | 8,830 | \bar{y}_{136} | 8,040 | \bar{y}_{152} | 9,820 | \bar{y}_{168} | 17,12 |
| \bar{y}_8 | 8,855 | \bar{y}_{24} | 9,285 | \bar{y}_{40} | 9,710 | \bar{y}_{56} | 8,700 | \bar{y}_{73} | 10,665 | \bar{y}_{89} | 8,710 | \bar{y}_{105} | 9,480 | \bar{y}_{121} | 10,055 | \bar{y}_{137} | 8,970 | \bar{y}_{153} | 10,405 | \bar{y}_{169} | 15,21 |
| \bar{y}_9 | 7,790 | \bar{y}_{25} | 9,205 | \bar{y}_{41} | 10,075 | \bar{y}_{57} | 10,070 | \bar{y}_{74} | 11,130 | \bar{y}_{90} | 8,505 | \bar{y}_{106} | 9,280 | \bar{y}_{122} | 8,175 | \bar{y}_{138} | 9,690 | \bar{y}_{154} | 9,725 | \bar{y}_{170} | 17,31 |
| \bar{y}_{10} | 8,965 | \bar{y}_{26} | 10,540 | \bar{y}_{42} | 9,454 | \bar{y}_{58} | 9,470 | \bar{y}_{75} | 11,770 | \bar{y}_{91} | 10,865 | \bar{y}_{107} | 10,020 | \bar{y}_{123} | 10,230 | \bar{y}_{139} | 9,470 | \bar{y}_{155} | 10,690 | \bar{y}_{171} | 16,97 |
| \bar{y}_{11} | 9,500 | \bar{y}_{27} | 9,435 | \bar{y}_{43} | 9,375 | \bar{y}_{59} | 8,530 | \bar{y}_{76} | 10,680 | \bar{y}_{92} | 11,310 | \bar{y}_{108} | 10,890 | \bar{y}_{124} | 9,415 | \bar{y}_{140} | 10,070 | \bar{y}_{156} | 9,040 | \bar{y}_{172} | 15,58 |
| \bar{y}_{12} | 9,470 | \bar{y}_{28} | 9,880 | \bar{y}_{44} | 11,530 | \bar{y}_{60} | 8,805 | \bar{y}_{77} | 10,540 | \bar{y}_{93} | 10,745 | \bar{y}_{109} | 9,030 | \bar{y}_{125} | 9,800 | \bar{y}_{141} | 9,470 | \bar{y}_{157} | 11,345 | \bar{y}_{173} | 15,53 |
| \bar{y}_{13} | 11,120 | \bar{y}_{29} | 10,570 | \bar{y}_{45} | 9,340 | \bar{y}_{61} | 7,985 | \bar{y}_{78} | 8,460 | \bar{y}_{94} | 11,560 | \bar{y}_{110} | 9,550 | \bar{y}_{126} | 9,375 | \bar{y}_{142} | 9,270 | \bar{y}_{158} | 10,58 | | |
| \bar{y}_{14} | 11,070 | \bar{y}_{30} | 9,850 | \bar{y}_{46} | 10,630 | \bar{y}_{62} | 9,760 | \bar{y}_{79} | 8,680 | \bar{y}_{95} | 11,690 | \bar{y}_{111} | 10,210 | \bar{y}_{127} | 9,660 | \bar{y}_{143} | 8,950 | \bar{y}_{159} | 9,70 | | |
| \bar{y}_{15} | 10,285 | \bar{y}_{31} | 8,850 | \bar{y}_{47} | 10,520 | \bar{y}_{63} | 9,490 | \bar{y}_{80} | 7,765 | \bar{y}_{96} | 11,490 | \bar{y}_{112} | 10,255 | \bar{y}_{128} | 9,920 | \bar{y}_{144} | 8,890 | \bar{y}_{160} | 9,77 | | |
| \bar{y}_{16} | 10,680 | \bar{y}_{32} | 10,575 | \bar{y}_{48} | 9,830 | \bar{y}_{64} | 9,870 | \bar{y}_{81} | 8,690 | \bar{y}_{97} | 9,850 | \bar{y}_{113} | 10,235 | \bar{y}_{129} | 9,975 | \bar{y}_{145} | 7,880 | \bar{y}_{161} | 7,88 | | |

Moyenne de la population $m_2=10,123$

Comparaison graphique entre les moyennes au milieu urbain et au milieu rural.



Pour comparer ces deux moyennes, on va procéder par le test T de Student.

- il s'agit ici d'une moyenne, donc le meilleur estimateur de ce paramètre est : \bar{x}
 Avant de commencer ce test, il faut d'abord vérifier si les conditions d'application de ce test mentionnées dans le chapitre précédent sont vérifiées.

- Les échantillons sont indépendants entre eux, sont aléatoires et ont n_1 et n_2

unités indépendantes.

- La variable aléatoire a un effectif supérieur à 30.
- L'égalité des variances : cette condition dans ce cas est indispensable car les effectifs sont inégaux.

Hypothèses Statistiques :

$$H_0 : m_1 = m_2$$

$$H_1 : m_1 < m_2$$

- Seuil de signification : $\alpha = 5$
- Conditions d'application du test sont vérifiées.

- Statistique du test : $T_{Exp} = \frac{(\bar{x} - \bar{y}) - (m_1 - m_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \rightsquigarrow \mathcal{T}_{n_1+n_2-2}$

- Règle de décision : Fractile de la loi de Student (cf. table) $t_\alpha = 1.650$

- Décision et conclusion :

- Calcul de la statistique sous H_0 c'est à dire avec $m_1 - m_2 = 0$

$$T_{Exp} = \frac{9.2038 - 10.1225}{\sqrt{\frac{1.276}{117} + \frac{3.3141}{173}}} = -5.2981$$

Le signe (-) est utilisé car la différence $\bar{x} - \bar{y}$ est négative.

T_{Exp} appartient à la zone de rejet de l'hypothèse nulle ($T_{Exp} = -5.2981 < t_\alpha = -1.650$)

Résultat :

On peut donc conclure, avec un risque $\alpha = 0.05$ que les deux moyennes diffèrent entre elles,
 i.e la moyenne des notes du monde urbain est significativement supérieure à celle du monde rural.

7.1.1 Urbain

- *Contrôle Continu / Examen Local* :

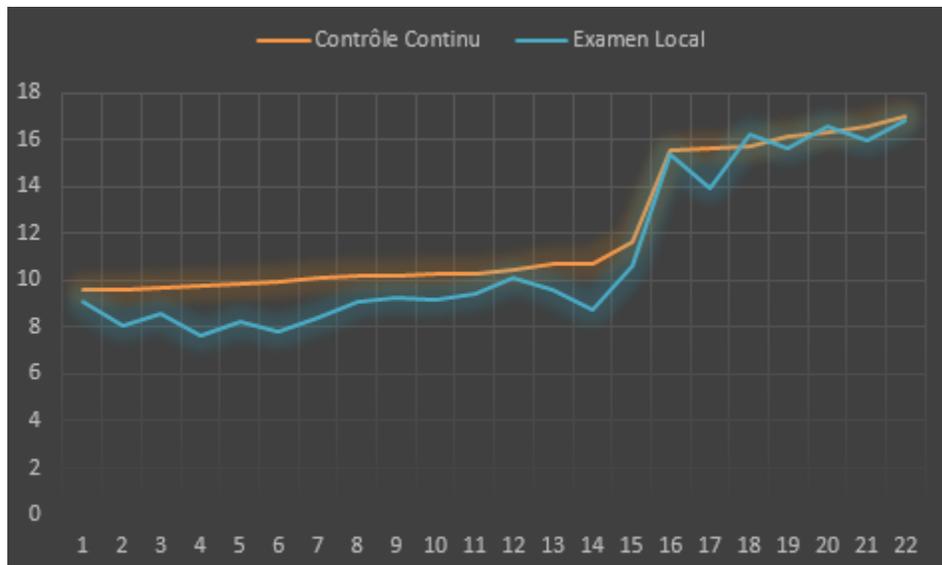
Les tableaux ci-dessous regroupent les moyennes des notes du contrôle continu et de l'examen local pour dans le monde urbain :

Contrôle Continu :

| Moyennes \ Secteur | Public | Privé |
|--------------------|--------|--------|
| 1 | 9,589 | 15,535 |
| 2 | 9,594 | 15,636 |
| 3 | 9,717 | 15,734 |
| 4 | 9,802 | 16,128 |
| 5 | 9,859 | 16,283 |
| 6 | 9,960 | 16,970 |
| 7 | 10,087 | |
| 8 | 10,165 | |
| 9 | 10,177 | |
| 10 | 10,289 | |
| 11 | 10,280 | |
| 12 | 10,455 | |
| 13 | 10,685 | |
| 14 | 10,693 | |
| 15 | 11,671 | |

Examen Local :

| Moyennes | | Secteur | |
|----------|--|---------|--------|
| | | Public | Privé |
| 1 | | 9,056 | 15,531 |
| 2 | | 8,115 | 13,954 |
| 3 | | 8,613 | 16,237 |
| 4 | | 7,647 | 15,625 |
| 5 | | 8,269 | 16,572 |
| 6 | | 7,846 | 15,938 |
| 7 | | 8,423 | |
| 8 | | 9,088 | |
| 9 | | 9,282 | |
| 10 | | 9,217 | |
| 11 | | 9,419 | |
| 12 | | 10,073 | |
| 13 | | 9,631 | |
| 14 | | 8,712 | |
| 15 | | 10,59 | |



D'après ce graphe on peut remarquer que les notes concernant, soit le contrôle continu ou l'examen local des élèves aux écoles privées sont plus élevées par rapport à celles aux écoles publics.

Et cela pourrait être due aux différents raisons à l'instar du gonflement des notes, l'élite qui préfère continuer sans cursus scolaire au sein d'un établissement privé pour assurer une bonne qualité d'enseignement.

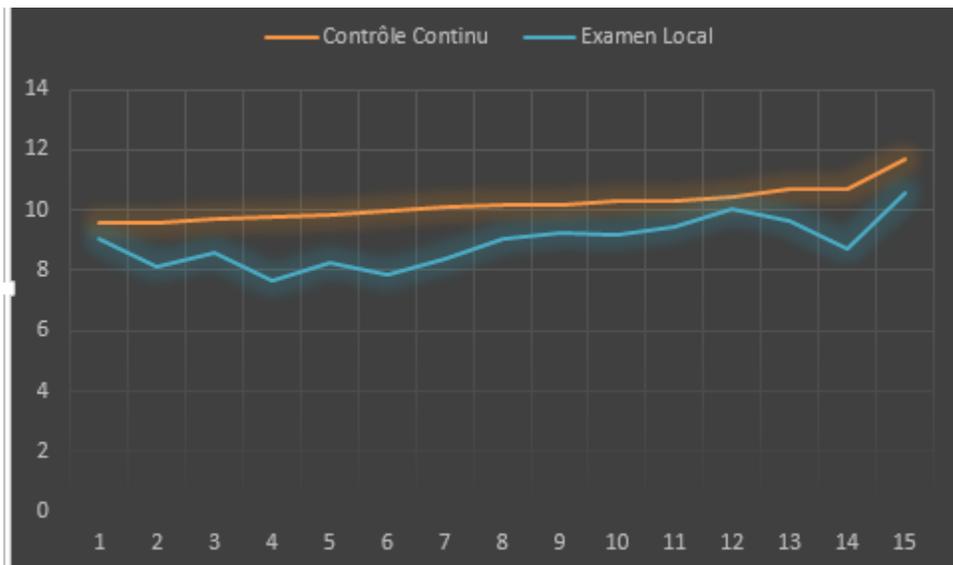
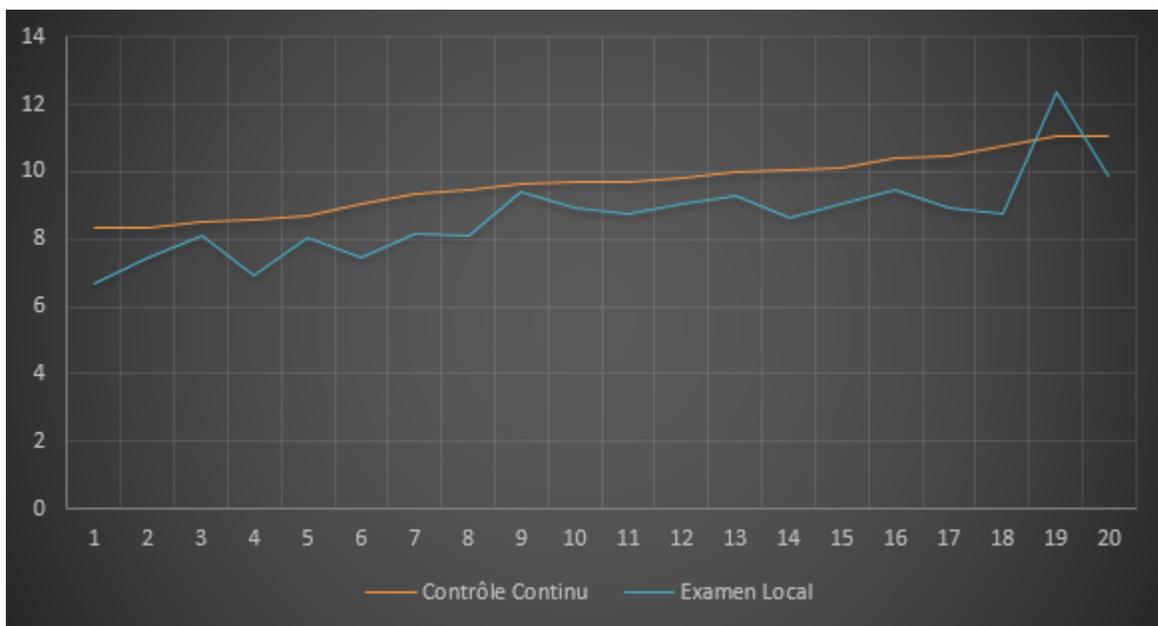


Figure 3 : Établissements privés exclus

Concernant ce graphe, on remarque que les notes du contrôle continu sont plus élevées par rapport à celles de l'examen local.

7.1.2 Rural

| | \bar{x}_1 | \bar{x}_2 | \bar{x}_3 | \bar{x}_4 | \bar{x}_5 | \bar{x}_6 | \bar{x}_7 | \bar{x}_8 | \bar{x}_9 | \bar{x}_{10} | \bar{x}_{11} | \bar{x}_{12} | \bar{x}_{13} | \bar{x}_{14} | \bar{x}_{15} | \bar{x}_{16} | \bar{x}_{17} | \bar{x}_{18} | \bar{x}_{19} | \bar{x}_{20} | |
|---------------------------|-----------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|--|
| <i>Contrôle Continu :</i> | 8.333 | 8.360 | 8.536 | 8.565 | 8.705 | 9.048 | 9.336 | 9.463 | 9.632 | 9.690 | 9.701 | 9.800 | 9.974 | 10.077 | 10.119 | 10.399 | 10.452 | 10.744 | 10.039 | 10.066 | |
| | 6.668 | 7.424 | 8.096 | 6.89 | 8.042 | 7.464 | 8.153 | 8.107 | 9.406 | 8.930 | 8.772 | 9.051 | 9.303 | 8.632 | 9.019 | 7.471 | 8.952 | 8.730 | 12.35 | 9.880 | |
| | <i>Examen Local :</i> | | | | | | | | | | | | | | | | | | | | |



La figure ci-dessus montre la corrélation entre les moyennes des contrôles continus et celles de l'examen local .

7.2 Par genre

Notons tout d'abord le nombre des filles et des garçons dans chaque milieu.

Le tableau ci-dessous regroupe le nombre des filles et des garçons dans les deux milieux :

| Genre \ Milieu | Urbain | Rural |
|----------------|--------|-------|
| Féminin | 2933 | 1428 |
| Masculin | 3762 | 3256 |

Remarque : A propos du tableau précédent, une remarque doit être faite concernant la différence entre le nombre des filles dans le monde urbain et dans le monde rural.

Dans cette partie, on va essayer de comparer le niveau d'étude par genre c'est à dire entre les garçons et les filles, soit dans le milieu rural ou dans le milieu urbain.

7.2.1 Rural

Féminin :

Avant de commencer, il nous faut un échantillon sur le quel on va faire le test.

Pour constituer un échantillon représentatif de la population, on va procéder par l'échantillonnage systématique décrit dans le cours.

Soit N la taille de la population

Notons n , le nombre d'individus que doit comporter l'échantillon (la taille de l'échantillon).

L'entier voisin de N/n sera noté r et appelé raison de sondage ou pas de sondage.

$$r = \frac{N}{n}$$

Dans ce cas $N = 1428$ et $n = 119$

$$r = \frac{1428}{119} = 12$$

On choisit aléatoirement un nombre entre 1 et $r = 12$

Notons ce nombre $a = 5$

Constituons cet échantillon avec a : le premier élément de l'échantillon

$$x_{nF}^- = a + (n - 1) \cdot r$$

| | | | | | | | | | | | | | | | |
|----------------|-------|----------------|-------|----------------|-------|----------------|-------|----------------|-------|----------------|-------|-----------------|-------|-----------------|-------|
| \bar{x}_1 | 7,73 | \bar{x}_{17} | 10,48 | \bar{x}_{33} | 6,73 | \bar{x}_{49} | 10,6 | \bar{x}_{65} | 6,24 | \bar{x}_{82} | 17,9 | \bar{x}_{98} | 10,98 | \bar{x}_{114} | 12,44 |
| \bar{x}_2 | 8,42 | \bar{x}_{18} | 9,85 | \bar{x}_{34} | 7,28 | \bar{x}_{50} | 15,06 | \bar{x}_{67} | 5,94 | \bar{x}_{83} | 15,29 | \bar{x}_{99} | 8,77 | \bar{x}_{115} | 12,53 |
| \bar{x}_3 | 7,94 | \bar{x}_{19} | 17,47 | \bar{x}_{35} | 15,11 | \bar{x}_{51} | 14,7 | \bar{x}_{68} | 15,54 | \bar{x}_{84} | 14,27 | \bar{x}_{100} | 11,6 | \bar{x}_{116} | 11,54 |
| \bar{x}_4 | 11,25 | \bar{x}_{20} | 11,13 | \bar{x}_{36} | 10,18 | \bar{x}_{52} | 7,56 | \bar{x}_{69} | 6,32 | \bar{x}_{85} | 7,28 | \bar{x}_{101} | 6,08 | \bar{x}_{117} | 10,97 |
| \bar{x}_5 | 15,35 | \bar{x}_{21} | 8,88 | \bar{x}_{37} | 11,85 | \bar{x}_{53} | 9,47 | \bar{x}_{70} | 6,98 | \bar{x}_{86} | 13,08 | \bar{x}_{102} | 18,55 | \bar{x}_{118} | 9,7 |
| \bar{x}_6 | 17,54 | \bar{x}_{22} | 10,38 | \bar{x}_{38} | 8,26 | \bar{x}_{54} | 11,32 | \bar{x}_{71} | 12,53 | \bar{x}_{87} | 5,52 | \bar{x}_{103} | 7,93 | \bar{x}_{119} | 13,63 |
| \bar{x}_7 | 6,01 | \bar{x}_{23} | 8,04 | \bar{x}_{39} | 9,88 | \bar{x}_{55} | 7,36 | \bar{x}_{72} | 6,05 | \bar{x}_{88} | 14,58 | \bar{x}_{104} | 11,32 | | |
| \bar{x}_8 | 6,69 | \bar{x}_{24} | 6,63 | \bar{x}_{40} | 10,5 | \bar{x}_{56} | 11,1 | \bar{x}_{73} | 3,39 | \bar{x}_{89} | 9,91 | \bar{x}_{105} | 8,24 | | |
| \bar{x}_9 | 10,96 | \bar{x}_{25} | 12,51 | \bar{x}_{41} | 14,46 | \bar{x}_{57} | 11,97 | \bar{x}_{74} | 11,76 | \bar{x}_{90} | 11,03 | \bar{x}_{106} | 11,56 | | |
| \bar{x}_{10} | 14,16 | \bar{x}_{26} | 11,85 | \bar{x}_{42} | 8,44 | \bar{x}_{58} | 13,36 | \bar{x}_{75} | 7,15 | \bar{x}_{91} | 8,82 | \bar{x}_{107} | 5,16 | | |
| \bar{x}_{11} | 14,34 | \bar{x}_{27} | 10,26 | \bar{x}_{43} | 12,7 | \bar{x}_{59} | 9,17 | \bar{x}_{76} | 9,33 | \bar{x}_{92} | 5,80 | \bar{x}_{108} | 9,99 | | |
| \bar{x}_{12} | 11,95 | \bar{x}_{28} | 7,76 | \bar{x}_{44} | 6,09 | \bar{x}_{60} | 8,87 | \bar{x}_{77} | 14,34 | \bar{x}_{93} | 9,54 | \bar{x}_{109} | 12,29 | | |
| \bar{x}_{13} | 9,57 | \bar{x}_{29} | 9,56 | \bar{x}_{45} | 10,03 | \bar{x}_{61} | 11,51 | \bar{x}_{78} | 15,1 | \bar{x}_{94} | 14,57 | \bar{x}_{110} | 13,81 | | |
| \bar{x}_{14} | 13,77 | \bar{x}_{30} | 3,45 | \bar{x}_{46} | 16,47 | \bar{x}_{62} | 12,46 | \bar{x}_{79} | 13,03 | \bar{x}_{95} | 11,40 | \bar{x}_{111} | 7,46 | | |
| \bar{x}_{15} | 9,97 | \bar{x}_{31} | 16,74 | \bar{x}_{47} | 11,7 | \bar{x}_{63} | 11,3 | \bar{x}_{80} | 11,72 | \bar{x}_{96} | 13,12 | \bar{x}_{112} | 8,15 | | |
| \bar{x}_{16} | 12,19 | \bar{x}_{32} | 11,42 | \bar{x}_{48} | 10,27 | \bar{x}_{64} | 14,36 | \bar{x}_{81} | 8,53 | \bar{x}_{97} | 8,24 | \bar{x}_{113} | 8,81 | | |

La moyenne de cet échantillon est de $m_F=10,66$.

Masculin :

De la même manière on constitue le deuxième échantillon

Dans ce cas $N = 3256$ et $n = 119$

Le pas de sondage dans ce cas sera $r' = \frac{3256}{119} = 27,36$

On prend $E(r') = 27$

On choisit aléatoirement un nombre entre 1 et 27

Notons ce nombre $b = 5$

avec

$$y_{1M} = 5$$

$$y_{nM} = b + (n - 1) \cdot r$$

| | | | | | | | | | | | | | | | |
|----------------|-------|----------------|-------|----------------|-------|----------------|-------|----------------|-------|----------------|-------|-----------------|-------|-----------------|-------|
| \bar{y}_1 | 0,28 | \bar{y}_{17} | 13,13 | \bar{y}_{33} | 6,79 | \bar{y}_{49} | 1,84 | \bar{y}_{65} | 8,00 | \bar{y}_{82} | 7,76 | \bar{y}_{98} | 11,27 | \bar{y}_{114} | 10,53 |
| \bar{y}_2 | 16,86 | \bar{y}_{18} | 8,92 | \bar{y}_{34} | 12,07 | \bar{y}_{50} | 8,18 | \bar{y}_{67} | 8,63 | \bar{y}_{83} | 10,49 | \bar{y}_{99} | 6,97 | \bar{y}_{115} | 6,76 |
| \bar{y}_3 | 11,25 | \bar{y}_{19} | 7,15 | \bar{y}_{35} | 6,6 | \bar{y}_{51} | 11,34 | \bar{y}_{68} | 7,79 | \bar{y}_{84} | 14,24 | \bar{y}_{100} | 11,38 | \bar{y}_{116} | 11,92 |
| \bar{y}_4 | 9,94 | \bar{y}_{20} | 9,93 | \bar{y}_{36} | 10,73 | \bar{y}_{52} | 4,72 | \bar{y}_{69} | 10,96 | \bar{y}_{85} | 5,64 | \bar{y}_{101} | 10,82 | \bar{y}_{117} | 9,40 |
| \bar{y}_5 | 9,94 | \bar{y}_{21} | 10,51 | \bar{y}_{37} | 5,70 | \bar{y}_{53} | 7,71 | \bar{y}_{70} | 5,16 | \bar{y}_{86} | 8,34 | \bar{y}_{102} | 7,00 | \bar{y}_{118} | 13,41 |
| \bar{y}_6 | 11,33 | \bar{y}_{22} | 6,02 | \bar{y}_{38} | 11,38 | \bar{y}_{54} | 13,85 | \bar{y}_{71} | 10,49 | \bar{y}_{87} | 10,09 | \bar{y}_{103} | 17,63 | \bar{y}_{119} | 11,95 |
| \bar{y}_7 | 9,13 | \bar{y}_{23} | 11,92 | \bar{y}_{39} | 19,13 | \bar{y}_{55} | 9,19 | \bar{y}_{72} | 9,31 | \bar{y}_{88} | 10,85 | \bar{y}_{104} | 7,21 | | |
| \bar{y}_8 | 9,20 | \bar{y}_{24} | 10,22 | \bar{y}_{40} | 1,97 | \bar{y}_{56} | 5,22 | \bar{y}_{73} | 7,79 | \bar{y}_{89} | 6,39 | \bar{y}_{105} | 11,37 | | |
| \bar{y}_9 | 4,64 | \bar{y}_{25} | 8,45 | \bar{y}_{41} | 13,23 | \bar{y}_{57} | 11,36 | \bar{y}_{74} | 15,20 | \bar{y}_{90} | 5,57 | \bar{y}_{106} | 11,55 | | |
| \bar{y}_{10} | 7,24 | \bar{y}_{26} | 11,61 | \bar{y}_{42} | 8,14 | \bar{y}_{58} | 11,89 | \bar{y}_{75} | 12,04 | \bar{y}_{91} | 6,21 | \bar{y}_{107} | 13,47 | | |
| \bar{y}_{11} | 10,61 | \bar{y}_{27} | 6,91 | \bar{y}_{43} | 12,37 | \bar{y}_{59} | 11,27 | \bar{x}_{76} | 9,51 | \bar{y}_{92} | 7,60 | \bar{y}_{108} | 12,18 | | |
| \bar{y}_{12} | 4,94 | \bar{y}_{28} | 4,73 | \bar{y}_{44} | 6,62 | \bar{y}_{60} | 12,00 | \bar{y}_{77} | 4,72 | \bar{y}_{93} | 10,45 | \bar{y}_{109} | 12,29 | | |
| \bar{y}_{13} | 4,36 | \bar{y}_{29} | 3,52 | \bar{y}_{45} | 9,08 | \bar{y}_{61} | 8,73 | \bar{y}_{78} | 6,46 | \bar{y}_{94} | 12,96 | \bar{y}_{110} | 13,52 | | |
| \bar{y}_{14} | 6,99 | \bar{y}_{30} | 8,39 | \bar{y}_{46} | 9,25 | \bar{y}_{62} | 6,58 | \bar{y}_{79} | 7,5 | \bar{y}_{95} | 11,49 | \bar{y}_{111} | 6,68 | | |
| \bar{y}_{15} | 8,24 | \bar{y}_{31} | 9,70 | \bar{y}_{47} | 8,65 | \bar{y}_{63} | 7,03 | \bar{y}_{80} | 6,31 | \bar{y}_{96} | 6,45 | \bar{y}_{112} | 9,31 | | |
| \bar{y}_{16} | 15,53 | \bar{y}_{32} | 6,9 | \bar{y}_{48} | 10,25 | \bar{y}_{64} | 8,99 | \bar{y}_{81} | 10,67 | \bar{y}_{97} | 10,91 | \bar{y}_{113} | 11,79 | | |

La moyenne de cet échantillon est de : $m_M = 9,25$

Test de *Fisher* pour vérifier l'égalité des variances.

Pour le premier échantillon, nous obtenons les statistiques suivantes : $\bar{x}_F=10,71$;
 $n_F = 119$; $S_F^2 = 10,227$

Pour le deuxième échantillon, nous obtenons les statistiques suivantes : $\bar{x}_M=9,30$;
 $n_M = 119$; $S_M^2 = 9,476$

Étape 1 :

Hypothèses :

$$H_0 : \sigma_F^2 = \sigma_M^2$$
$$H_1 : \sigma_F^2 > \sigma_M^2$$

Étape 2 :

Le risque α sera de 5%.

Étape 3 :

Déterminer le critère de rejet d'un test d'égalité des variances.

$$F_{exp} = \frac{\text{La plus grande variance}}{\text{La plus petite variance}} = \frac{10,2279861}{9,476985} = 1,0792446$$

Étape 4 :

Règle de décision :

Il faut maintenant trouver dans la table de Fisher la valeur critique $F_{0,05,118,118} = 1,48$ ou α est le risque d'erreur, ν_1 est le degré de liberté de l'échantillon ayant la plus grande variance et ν_2 est le degré de liberté de l'échantillon ayant la plus petite variance.

La règle de décision s'énoncera alors ainsi :

Si $F_{exp} > 1,48$ on doit rejeter H_0 avec un risque d'erreur de 5%

Puisque $1,07 < 1,48$; nous conservons H_0 .

Étape 5 :

Conclusion :

Avec un risque d'erreur de 5% on peut conclure que les variances sont égales.

Test t de Student :

Vérification des conditions d'application du test :

- Les échantillons sont indépendants entre eux, sont aléatoires et ont n_1 et n_2 unités indépendantes.
- La variable aléatoire a un effectif supérieur à 30.
- Les variances sont égales

Hypothèses Statistiques :

$$H_0 : m_M = m_F$$

$$H_1 : m_M < m_F$$

- Seuil de signification : $\alpha = 5$
- Conditions d'application du test sont vérifiées.
- Statistique du test :
$$T_{Exp} = \frac{(\bar{x}_F - \bar{x}_M) - (m_F - m_M)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \rightsquigarrow \mathcal{T}_{n_1+n_2-2}$$
- Règle de décision : Fractile de la loi de Student (cf. table) $t_\alpha = 1.650$
- Décision et conclusion :
- Calcul de la statistique sous H_0 c'est à dire avec $m_F - m_M = 0$

Dans ce cas, les effectifs sont égaux $n = n_1 = n_2 = 119$. La valeur de t devient :

$$T_{Exp} = \frac{(\bar{x}_F - \bar{x}_M) - (m_F - m_M)}{\sqrt{\frac{2S_p^2}{n}}} \quad (1)$$

$$\text{avecc : } S_p^2 = \frac{(n_1 - 1).S_F^2 + (n_2 - 1).S_M^2}{n_1 + n_2 - 2}$$

Calcul :

$$S_p^2 = \frac{(119 - 1).10,227 + (119 - 1).9,476}{119 + 119 - 2} = 9,936$$

On remplace dans (1) :

$$T_{Exp} = \frac{10,716 - 9,304}{\sqrt{\frac{2.9,936}{119}}} = 3,455$$

On a $T_{exp} > -1.650$ nous conservons alors H_0 .

Résultat :

On peut donc conclure, avec un risque $\alpha = 0.05$ de se tromper qu'il n'y a pas de différence significative entre les deux moyennes.

7.2.2 Urbain

Féminin :

Constituons le premier échantillon, soit $n = 120$ comme taille de l'échantillon.

Soit $N = 2933$ la taille de la population.

$$r = \frac{N}{n} = \frac{2933}{120} = 31.35$$

On prend $E(r) = 31$

Choisissons aléatoirement un nombre entre 1 et $r = 31$

Notons ce nombre $a' = 10$

Le premier élément de cet échantillon sera la valeur correspondante au nombre 10 (*dans excel*).

avec

$$\bar{y}'_{nF} = a' + (n - 1) \cdot k$$

L'échantillon alors sera constitué comme suit :

| | | | | | | | | | | | | | | | |
|-----------------|-------|-----------------|-------|-----------------|-------|-----------------|-------|-----------------|-------|-----------------|-------|------------------|-------|------------------|-------|
| \bar{y}'_1 | 12,58 | \bar{y}'_{17} | 10,84 | \bar{y}'_{33} | 8,30 | \bar{y}'_{49} | 10,85 | \bar{y}'_{65} | 7,09 | \bar{y}'_{82} | 8,69 | \bar{y}'_{98} | 15,02 | \bar{y}'_{114} | 13,37 |
| \bar{y}'_2 | 11,87 | \bar{y}'_{18} | 14,09 | \bar{y}'_{34} | 11,02 | \bar{y}'_{50} | 10,56 | \bar{y}'_{67} | 8,77 | \bar{y}'_{83} | 11,11 | \bar{y}'_{99} | 17,15 | \bar{y}'_{115} | 7,96 |
| \bar{y}'_3 | 11,29 | \bar{y}'_{19} | 11,39 | \bar{y}'_{35} | 14,75 | \bar{y}'_{51} | 14,76 | \bar{y}'_{68} | 6,79 | \bar{y}'_{84} | 9,28 | \bar{y}'_{100} | 8,90 | \bar{y}'_{116} | 16,29 |
| \bar{y}'_4 | 16,57 | \bar{y}'_{20} | 14,96 | \bar{y}'_{36} | 9,48 | \bar{y}'_{52} | 13,62 | \bar{y}'_{69} | 6,07 | \bar{y}'_{85} | 7,64 | \bar{y}'_{101} | 15,31 | \bar{y}'_{117} | 10,65 |
| \bar{y}'_5 | 10,82 | \bar{y}'_{21} | 11,32 | \bar{y}'_{37} | 9,05 | \bar{y}'_{53} | 8,76 | \bar{y}'_{70} | 14,90 | \bar{y}'_{86} | 11,69 | \bar{y}'_{102} | 6,18 | \bar{y}'_{118} | 15,28 |
| \bar{y}'_6 | 8,13 | \bar{y}'_{22} | 14,66 | \bar{y}'_{38} | 13,42 | \bar{y}'_{54} | 10,09 | \bar{y}'_{71} | 7,8 | \bar{y}'_{87} | 12,65 | \bar{y}'_{103} | 6,05 | \bar{y}'_{119} | 11,5 |
| \bar{y}'_7 | 10,27 | \bar{y}'_{23} | 7,690 | \bar{y}'_{39} | 9,98 | \bar{y}'_{55} | 11,77 | \bar{y}'_{72} | 17,20 | \bar{y}'_{88} | 7,54 | \bar{y}'_{104} | 12,75 | \bar{y}'_{120} | 17,07 |
| \bar{y}'_8 | 8,42 | \bar{y}'_{24} | 12,26 | \bar{y}'_{40} | 11,05 | \bar{y}'_{56} | 13,13 | \bar{y}'_{73} | 8,22 | \bar{y}'_{89} | 10,58 | \bar{y}'_{105} | 13,73 | | |
| \bar{y}'_9 | 9,73 | \bar{y}'_{25} | 9,44 | \bar{y}'_{41} | 9,13 | \bar{y}'_{57} | 13,77 | \bar{y}'_{74} | 15,75 | \bar{y}'_{90} | 17,53 | \bar{y}'_{106} | 7,88 | | |
| \bar{y}'_{10} | 9,67 | \bar{y}'_{26} | 8,78 | \bar{y}'_{42} | 11,28 | \bar{y}'_{58} | 9,09 | \bar{y}'_{75} | 9,53 | \bar{y}'_{91} | 11,62 | \bar{y}'_{107} | 8,83 | | |
| \bar{y}'_{11} | 10,78 | \bar{y}'_{27} | 11,69 | \bar{y}'_{43} | 11,71 | \bar{y}'_{59} | 9,73 | \bar{y}'_{76} | 13,32 | \bar{y}'_{92} | 14,21 | \bar{y}'_{108} | 16,17 | | |
| \bar{y}'_{12} | 19,46 | \bar{y}'_{28} | 6,51 | \bar{y}'_{44} | 14,75 | \bar{y}'_{60} | 10,01 | \bar{y}'_{77} | 9,59 | \bar{y}'_{93} | 9,41 | \bar{y}'_{109} | 12,23 | | |
| \bar{y}'_{13} | 5,17 | \bar{y}'_{29} | 8,49 | \bar{y}'_{45} | 17,51 | \bar{y}'_{61} | 16,02 | \bar{y}'_{78} | 14,07 | \bar{y}'_{94} | 12,24 | \bar{y}'_{110} | 17,92 | | |
| \bar{y}'_{14} | 15,58 | \bar{y}'_{30} | 9,46 | \bar{y}'_{46} | 17,90 | \bar{y}'_{62} | 14,28 | \bar{y}'_{79} | 10,77 | \bar{y}'_{95} | 15,29 | \bar{y}'_{111} | 8,28 | | |
| \bar{y}'_{15} | 18,47 | \bar{y}'_{31} | 13,59 | \bar{y}'_{47} | 13,55 | \bar{y}'_{63} | 9,45 | \bar{y}'_{80} | 15,85 | \bar{y}'_{96} | 11,68 | \bar{y}'_{112} | 10,61 | | |
| \bar{y}'_{16} | 14,64 | \bar{y}'_{32} | 19,50 | \bar{y}'_{48} | 11,04 | \bar{y}'_{64} | 13,31 | \bar{y}'_{81} | 16,69 | \bar{y}'_{97} | 9,5 | \bar{y}'_{113} | 10,98 | | |

La moyenne observée de cet échantillon est : $m'_F = 11,822$.

Masculin :

Le deuxième échantillon sera constitué de la même façon, soit $n = 120$ comme taille de l'échantillon.

Soit $N = 3762$ la taille de la population.

$$r' = \frac{N}{n} = \frac{3762}{120} = 24,44$$

On prend $E(k') = 24$

Choisissons aléatoirement un nombre entre 1 et $r' = 24$

Notons ce nombre $b' = 20$

Le premier élément de cet échantillon sera la valeur correspondante au nombre 20 (*dans excel*).

avec

$$\boxed{x'_{nM} = b' + (n - 1) \cdot k'}$$

L'échantillon constitué est le suivant :

| | | | | | | | | | | | | | | | |
|-----------------|-------|-----------------|-------|-----------------|-------|-----------------|-------|-----------------|-------|-----------------|-------|------------------|-------|------------------|-------|
| \bar{x}'_1 | 8,53 | \bar{x}'_{17} | 10,26 | \bar{x}'_{33} | 11,28 | \bar{x}'_{49} | 8,69 | \bar{x}'_{65} | 10,32 | \bar{x}'_{82} | 7,92 | \bar{x}'_{98} | 15,10 | \bar{x}'_{114} | 15,40 |
| \bar{x}'_2 | 12,68 | \bar{x}'_{18} | 9,19 | \bar{x}'_{34} | 8,82 | \bar{x}'_{50} | 5,01 | \bar{x}'_{67} | 10,64 | \bar{x}'_{83} | 14,05 | \bar{x}'_{99} | 8,71 | \bar{x}'_{115} | 11,81 |
| \bar{x}'_3 | 9,72 | \bar{x}'_{19} | 15,04 | \bar{x}'_{35} | 12,30 | \bar{x}'_{51} | 8,52 | \bar{x}'_{68} | 8,13 | \bar{x}'_{84} | 8,14 | \bar{x}'_{100} | 7,05 | \bar{x}'_{116} | 6,26 |
| \bar{x}'_4 | 15,05 | \bar{x}'_{20} | 10,77 | \bar{x}'_{36} | 9,71 | \bar{x}'_{52} | 10,76 | \bar{x}'_{69} | 7,13 | \bar{x}'_{85} | 14,36 | \bar{x}'_{101} | 8,33 | \bar{x}'_{117} | 5,94 |
| \bar{x}'_5 | 15,57 | \bar{x}'_{21} | 8,30 | \bar{x}'_{37} | 7,15 | \bar{x}'_{53} | 9,78 | \bar{x}'_{70} | 14,56 | \bar{x}'_{86} | 10,01 | \bar{x}'_{102} | 12,32 | \bar{x}'_{118} | 6,84 |
| \bar{x}'_6 | 10,14 | \bar{x}'_{22} | 6,91 | \bar{x}'_{38} | 9,03 | \bar{x}'_{54} | 8,15 | \bar{x}'_{71} | 10,02 | \bar{x}'_{87} | 16,79 | \bar{x}'_{103} | 10,95 | \bar{x}'_{119} | 11,82 |
| \bar{x}'_7 | 6,93 | \bar{x}'_{23} | 6,37 | \bar{x}'_{39} | 8,24 | \bar{x}'_{55} | 6,87 | \bar{x}'_{72} | 4,96 | \bar{x}'_{88} | 5,27 | \bar{x}'_{104} | 9,77 | \bar{x}'_{120} | 6,94 |
| \bar{x}'_8 | 11,25 | \bar{x}'_{24} | 3,90 | \bar{x}'_{40} | 8,67 | \bar{x}'_{56} | 17,71 | \bar{x}'_{73} | 4,72 | \bar{x}'_{89} | 12,74 | \bar{x}'_{105} | 5,17 | | |
| \bar{x}'_9 | 15,77 | \bar{x}'_{25} | 5,42 | \bar{x}'_{41} | 5,99 | \bar{x}'_{57} | 3,39 | \bar{x}'_{74} | 10,3 | \bar{x}'_{90} | 8,54 | \bar{x}'_{106} | 7,74 | | |
| \bar{x}'_{10} | 11,36 | \bar{x}'_{26} | 4,22 | \bar{x}'_{42} | 10,07 | \bar{x}'_{58} | 9,93 | \bar{x}'_{75} | 8,92 | \bar{x}'_{91} | 8,71 | \bar{x}'_{107} | 6,42 | | |
| \bar{x}'_{11} | 11,48 | \bar{x}'_{27} | 8,56 | \bar{x}'_{43} | 7,29 | \bar{x}'_{59} | 10,65 | \bar{x}'_{76} | 8,18 | \bar{x}'_{92} | 5,75 | \bar{x}'_{108} | 10,96 | | |
| \bar{x}'_{12} | 10,16 | \bar{x}'_{28} | 7,75 | \bar{x}'_{44} | 13,68 | \bar{x}'_{60} | 11,47 | \bar{x}'_{77} | 9,6 | \bar{x}'_{93} | 16,21 | \bar{x}'_{109} | 9,68 | | |
| \bar{x}'_{13} | 13,03 | \bar{x}'_{29} | 10,56 | \bar{x}'_{45} | 11,98 | \bar{x}'_{61} | 8,14 | \bar{x}'_{78} | 10,01 | \bar{x}'_{94} | 10,30 | \bar{x}'_{110} | 7,74 | | |
| \bar{x}'_{14} | 8,14 | \bar{x}'_{30} | 3,32 | \bar{x}'_{46} | 10,83 | \bar{x}'_{62} | 5,20 | \bar{x}'_{79} | 9,26 | \bar{x}'_{95} | 5,04 | \bar{x}'_{111} | 15,08 | | |
| \bar{x}'_{15} | 8,17 | \bar{x}'_{31} | 9,35 | \bar{x}'_{47} | 14,66 | \bar{x}'_{63} | 9,82 | \bar{x}'_{80} | 6,89 | \bar{x}'_{96} | 10,89 | \bar{x}'_{112} | 4,24 | | |
| \bar{x}'_{16} | 9,83 | \bar{x}'_{32} | 3,21 | \bar{x}'_{48} | 7,97 | \bar{x}'_{64} | 9,15 | \bar{x}'_{81} | 16,43 | \bar{x}'_{97} | 6,35 | \bar{x}'_{113} | 6,67 | | |

La moyenne observée de cet échantillon : $m'_M = 9,480$.

Test de *Fisher* pour vérifier l'égalité des variances.

Pour le deuxième échantillon, nous obtenons les statistiques suivantes : $\bar{x}'_M = 9,225$; $n'_M = 120$; $S'^2_M = 10,015$

Pour le premier échantillon, nous obtenons les statistiques suivantes : $\bar{y}'_F = 11,305$; $n'_F = 120$; $S'^2_F = 10,681$

Étape 1 :

Hypothèses :

$$\begin{aligned} H_0 : \sigma'^2_F &= \sigma'^2_M \\ H_1 : \sigma'^2_F &\neq \sigma'^2_M \end{aligned}$$

Étape 2 :

Le risque α sera de 5%.

Étape 3 :

Déterminer le critère de rejet d'un test d'égalité des variances.

$$F_{exp} = \frac{\text{La plus grande variance}}{\text{La plus petite variance}} = \frac{10,681132223}{10,01569276} = 1,066500$$

Étape 4 :

Règle de décision :

Il faut maintenant trouver dans la table de Fisher la valeur critique $F_{0,05,119,119} = 1,48$ où α est le risque d'erreur et ν'_1 le degré de liberté de l'échantillon ayant la plus grande variance et ν'_2 est le degré de liberté de l'échantillon ayant la plus petite variance.

La règle de décision s'énoncera alors ainsi :

Si $F_{exp} > 1,48$ on doit rejeter H_0 avec un risque d'erreur de 5%

Puisque $1,06650 < 1,48$; nous conservons H_0 .

Étape 5 :

Conclusion :

Avec un risque d'erreur de 5% on peut conclure que les variances sont égales.

Test t de Student :

Vérification des conditions d'application du test :

- Les échantillons sont indépendants entre eux, sont aléatoires et ont n_1 et n_2 unités indépendantes.
- La variable aléatoire a un effectif supérieur à 30.
- Les variances sont égales.

Hypothèses Statistiques :

$$\begin{aligned} H_0 : m'_M &= m'_F \\ H_1 : m'_M &< m'_F \end{aligned}$$

- Seuil de signification : $\alpha = 5$
- Conditions d'application du test sont vérifiées.

- Statistique du test :
$$T_{Exp} = \frac{(\bar{x}'_M - \bar{y}'_F) - (m'_M - m'_F)}{\sqrt{\frac{S'^2_F}{n'_F} + \frac{S'^2_M}{n'_M}}} \rightsquigarrow \mathcal{T}_{n_1+n_2-2}$$

- Règle de décision : Fractile de la loi de Student (cf. table) $t_\alpha = 1.650$

- Décision et conclusion :

- Calcul de la statistique sous H_0 c'est à dire avec $m'_M - m'_F = 0$

Dans ce cas, les effectifs sont égaux $n' = n'_M = n'_F = 120$. La valeur de t devient :

$$T_{Exp} = \frac{(\bar{x}'_M - \bar{y}'_F) - (m_F - m_M)}{\sqrt{\frac{2S_p'^2}{n'}}} \quad (1)$$

$$\text{avec : } S_p'^2 = \frac{(n'_M - 1) \cdot S_M'^2 + (n'_F - 1) \cdot S_F'^2}{n'_M + n'_F - 2}$$

Calcul :

$$S_p'^2 = \frac{(120 - 1) \cdot 10,015 + (120 - 1) \cdot 10,681}{120 + 120 - 2} = 10,348$$

On remplace dans (1) :

$$T_{Exp} = \frac{9,225 - 11,305}{\sqrt{\frac{2 \cdot 10,348}{120}}} = -5,0085$$

Le signe (-) est utilisé car la différence $\bar{x}'_M - \bar{y}'_F$ est négative.

T_{Exp} appartient à la zone de rejet de l'hypothèse nulle ($T_{Exp} = -5.0085 < t_\alpha = -1.650$)

Résultat :

On pourrait dire alors que les filles ont une moyenne significativement supérieure à celle des garçons.

7.3 Par matière

Dans cette partie on va comparer les moyennes des élèves par matières (matières scientifiques et littéraires), on dispose de deux populations, il est indispensable de présenter les notes de chaque individu, due à leur énorme nombre.

Désignons respectivement par \bar{x}_1 et \bar{y}_1 la moyenne des matières scientifiques et celle des matières littéraires.

Les tableaux ci-dessous regroupent toutes les informations nécessaires pour effectuer le test statistique utile pour comparer ces deux moyennes :

7.3.1 Urbain

| Statistiques \ Matières | Scientifiques | Littéraires |
|-------------------------|---------------|-------------|
| | Moyennes | 8,5703 |
| Variances | 14,8135 | 12,8911 |
| Écart-type | 3,8496 | 3,5918 |
| Observations | 6330 | 6330 |

Test de Fisher

- Hypothèses :

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

- Le risque α sera de 5%.

- Critère de rejet

$$F_{exp} = \frac{\text{La plus grande variance}}{\text{La plus petite variance}} = \frac{14,8135}{11,3198} = 1,3086$$

- Règle de décision :

Il faut maintenant trouver dans la table de Fisher la valeur critique $F_{0.05,6329,6329} = 1.0099$ ou α est le risque d'erreur et ν_1 le degré de liberté de l'échantillon ayant la plus grande variance et ν_2 est le degré de liberté de l'échantillon ayant la plus petite variance.

La règle de décision s'énoncera alors ainsi :

Si $F_{exp} > 1,0099$ on doit rejeter H_0 avec un risque d'erreur de 5%

Puisque $1,3086 > 1,0099$; nous rejettons H_0 .

• **Conclusion :**

Avec un risque d'erreur de 5% on peut conclure que les variances sont différentes

Choix du test, et vérification des conditions d'application :

Loi normale centrée et réduite

- Deux échantillons indépendants de taille $\gg 30$
- Les deux variables sont continues
- Les moyennes \bar{x}_1 et \bar{y}_1 des notes suivent une loi normale

Définition des hypothèses :

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Risque α :

Risque bilatéral : 5%

Règle de décision :

$$\text{Accepter } H_0 \text{ si } Z_{Exp} = \frac{(\bar{x}_1 - \bar{y}_1) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \in [-Z_{\alpha/2}; +Z_{\alpha/2}]$$

Calcul de la statistique

$$Z_{Exp} = \frac{(\bar{x}_1 - \bar{y}_1) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{9,7460 - 8,5703}{\sqrt{\frac{14,8135}{6330} + \frac{12,8911}{6330}}} = 17,7714$$

On a $Z_{Exp} = 17,7714 \notin [-1,960; 1,960]$

• **Conclusion :**

Avec un risque d'erreur de 5% on peut conclure que les moyennes sont significativement différentes

7.3.2 Rural

| Statistiques \ Matières | Scientifiques | Littéraires |
|-------------------------|---------------|-------------|
| Moyennes | 7,8794 | 9,2901 |
| Variances | 13,6805 | 11,3198 |
| Écart-type | 3,7002 | 3,3656 |
| Observations | 4229 | 4229 |

Test de Fisher.

- Hypothèses :

$$H_0 : \sigma_1'^2 = \sigma_2'^2$$
$$H_1 : \sigma_1'^2 \neq \sigma_2'^2$$

- Le risque α sera de 5%.

- Critère de rejet

$$F_{exp} = \frac{\text{La plus grande variance}}{\text{La plus petite variance}} = \frac{13,6805}{11,31198} = 1,3086$$

- Règle de décision :

Il faut maintenant trouver dans la table de Fisher la valeur critique $F_{0.05,4228,4228} = 1.0099$ ou α est le risque d'erreur et ν_1' le degré de liberté de l'échantillon ayant la plus grande variance et ν_2' est le degré de liberté de l'échantillon ayant la plus petite variance.

La règle de décision s'énoncera alors ainsi :

Si $F_{exp} > 1,00$ on doit rejeter H_0 avec un risque d'erreur de 5%

Puisque $1,3086 > 1,0099$, nous rejettons H_0 .

- Conclusion :

| |
|---|
| Avec un risque d'erreur de 5% on peut conclure que les variances sont différentes |
|---|

Choix du test, et vérification des conditions d'application:

Loi normale centrée et réduite

- Deux échantillons indépendants de taille $\gg 30$
- Les deux variables sont continues
- Les moyennes \bar{x}_2 et \bar{y}_2 des notes suivent une loi normale

Définition des hypothèses :

$$\begin{aligned} H_0 : \mu'_1 &= \mu'_2 \\ H_1 : \mu'_1 &\neq \mu'_2 \end{aligned}$$

Risque α :

Risque bilatéral : 5%

Règle de décision :

$$\text{Accepter } H_0 \text{ si } Z_{Exp} = \frac{(\bar{x}_2 - \bar{y}_2) - (\mu'_1 - \mu'_2)}{\sqrt{\frac{\sigma_1'^2}{n_1} + \frac{\sigma_2'^2}{n_2}}} \in [-Z_{\alpha/2}; +Z_{\alpha/2}]$$

Calcul de la statistique

$$Z_{Exp} = \frac{(\bar{x}_2 - \bar{y}_2) - (\mu'_1 - \mu'_2)}{\sqrt{\frac{S_1'^2}{n_1} + \frac{S_2'^2}{n_2}}} = \frac{9,2901 - 7,8794}{\sqrt{\frac{13,6805}{4229} + \frac{11,3198}{4229}}} = 18,3476$$

On a $Z_{Exp} = 18,3476 \notin [-1,960; 1,960]$

Conclusion :

Avec un risque d'erreur de 5% on peut conclure que les moyennes sont significativement différentes.

8 Évolution des taux de réussite du baccalauréat

Dans cette partie, on va essayer de poursuivre l'évolution du taux de réussite du baccalauréat pour 4 groupes :

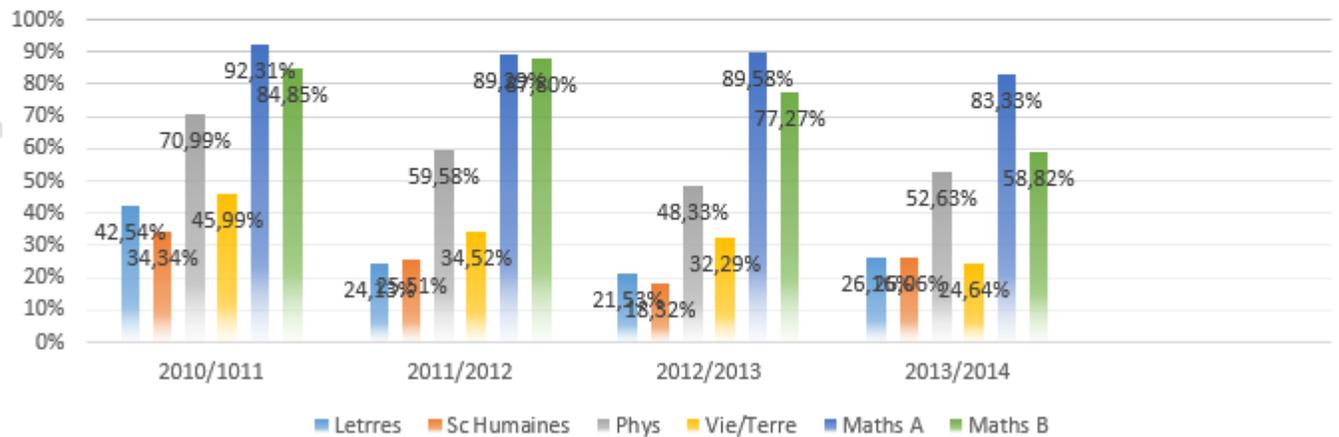
- Enseignement Général
- Enseignement Originel
- Techniques Commerciales
- Techniques Industrielles

8.1 Enseignement Général

Le tableau ci-dessous regroupe les taux de réussite de 5 années :

| Option \ Année | 2009/2010 | 2010/2011 | 2011/2012 | 2012/2013 | 2013/2014 |
|-----------------------------|-----------|-----------|-----------|-----------|-----------|
| Lettres | 31,32% | 42,45% | 24,13% | 21,53% | 26,16% |
| Sciences Humaines | 27,49% | 34,34% | 25,51% | 18,32% | 26,06% |
| Sciences Physiques | 64,47% | 70,99% | 59,58% | 48,33% | 52,63% |
| Sciences de la Vie et Terre | 39,21% | 45,99% | 34,52% | 32,29% | 24,64% |
| Sciences Mathématiques A | 85,71% | 92,31% | 89,29% | 89,58% | 83,33% |
| Sciences Mathématiques B | 85,19% | 84,85% | 87,80% | 77,27% | 58,82% |

EVOLUTION DDE TAUX DE RÉUSSITE DE CHAQUE FILIÈRE



Pour savoir comment évolue le taux de réussite pour ce groupe, on va procéder par le test d'indépendance de *Khi – deux* décrit dans la partie du cours.

Tableau de contingence :

| Filière \ Année | 2009/2010 | 2010/2011 | 2011/2012 | 2012/2013 | 2013/2014 | Σ |
|-----------------------------|---------------|---------------|---------------|---------------|---------------|----------------|
| Lettres | 31,32 | 42,45 | 24,13 | 21,53 | 26,16 | 146,62 |
| Sciences Humaines | 27,49 | 34,34 | 25,51 | 18,32 | 26,06 | 131,72 |
| Sciences Physiques | 64,47 | 70,99 | 59,58 | 48,33 | 52,63 | 296 |
| Sciences de la Vie et Terre | 39,21 | 45,99 | 34,52 | 32,29 | 24,64 | 176,65 |
| Sciences Mathématiques A | 85,71 | 92,31 | 89,29 | 89,58 | 83,33 | 440,22 |
| Sciences Mathématiques B | 85,19 | 84,85 | 87,80 | 77,27 | 58,82 | 393,32 |
| Σ | 333,39 | 370,93 | 320,83 | 287,32 | 271,64 | 1584,11 |

$$Stat = \frac{[f_{ij} - f_{tij}]^2}{f_{tij}}$$

La somme de ces valeurs $\langle Stat \rangle$ donne la statistique χ^2 globale.

Pour calculer cette statistique, il faut calculer les espérances f_{tij} de chaque case, les espérances se calculent de la façon suivante :

$$f_{tij} = \frac{n_j \cdot m_i}{n}$$

On calcule la première Stat de la case une :

$$f_{t11}) = \frac{n_1 \cdot m_1}{n} = \frac{333, 39.146, 62}{1584, 11} = 30, 857$$

$$\text{Alors } Stat = \frac{[f_{11} - f_{t11}]^2}{f_{t11}} = \frac{[31, 32 - 30, 875]^2}{30, 875} = 0, 0220$$

On continue le calcul de la même manière jusqu'à trouver la valeur de la Stat globale .

$$\chi^2 = \sum Stat$$

Après calcul, on a trouvé que $\chi^2 = 15, 8248$

Degré de liberté :

$$d.d.l = (k - 1)(p - 1) = (6 - 1)(5 - 1) = 20$$

Règle de décision :

La statistique $\chi^2 = 15, 8248$ est inférieure à la valeur critique $\chi_{0,05(\nu=20)}^2 = 31, 41$

On accepte l'hypothèse nulle.

Conclusion :

On en conclut alors que le taux de réussite de ces six filières n'a pas connu une différence significative ces cinq dernières années.

8.2 Enseignement Originel

| Option \ Année | 2009/2010 | 2010/2011 | 2011/2012 | 2012/2013 | 2013/2014 |
|-----------------------|-----------|-----------|-----------|-----------|-----------|
| Sciences de La CHARIA | 14,29% | 27,45% | 25,33% | 10,00% | 16,30% |
| Langue ARABE | 7,32% | 17,65% | 5,17% | 5,00% | 16,39% |

On procède de la même manière faite pour l'enseignement général .

$$\chi^2 = \sum Stat$$

On obtient comme statistique globale après calcul :

| Option \ Année | 2009/2010 | 2010/2011 | 2011/2012 | 2012/2013 | 2013/2014 | Σ |
|-----------------------|--------------|-------------|-------------|--------------|--------------|--------------|
| Sciences de La CHARIA | 14,29 | 27,45 | 25,33 | 10,00 | 16,30 | 93,37 |
| Langue ARABE | 7,32 | 17,65 | 5,17 | 5,00 | 16,39 | 51,53 |
| Σ | 21,61 | 45,1 | 30,5 | 15,00 | 32,69 | 144,9 |

$$\chi^2 = 8,3322$$

Degré de liberté :

$$d.d.l = (2 - 1)(5 - 1) = 4$$

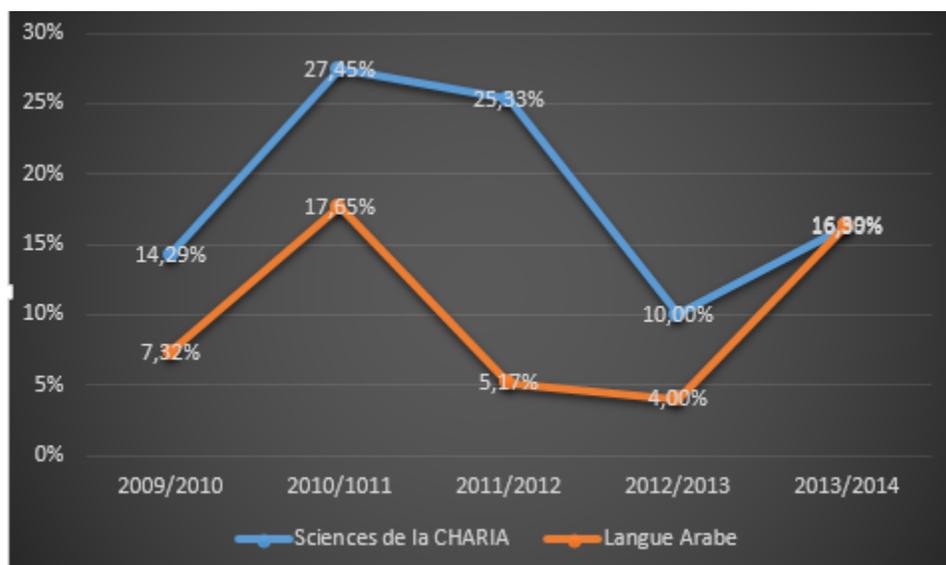
Règle de décision :

La statistique $\chi^2 = 8,3322$ est inférieure à la valeur critique $\chi^2_{0,05(\nu=4)} = 9,49$

On accepte l'hypothèse nulle.

Conclusion :

On en conclut alors que le taux de réussite des deux filières n'a pas connu une différence significative ces cinq dernières années.



8.3 Enseignement Techniques Commerciales

De la même façon,

| Option | Année | | | | | Σ |
|----------------------------|--------------|--------------|--------------|--------------|--------------|---------------|
| | 2009/2010 | 2010/2011 | 2011/2012 | 2012/2013 | 2013/2014 | |
| Sciences Économiques | 41,13 | 57,01 | 42,06 | 51,39 | 39,22 | 230,81 |
| Sciences Gestion Comptable | 49,15 | 32,00 | 49,15 | 39,58 | 51,22 | 221,12 |
| Σ | 90,28 | 89,01 | 91,23 | 90,97 | 90,44 | 451,93 |

$$\chi^2 = \sum Stat$$

On obtient comme statistique globale après calcul :

$$\chi^2 = 11,2128$$

Degré de liberté :

$$d.d.l = (2 - 1)(5 - 1) = 4$$

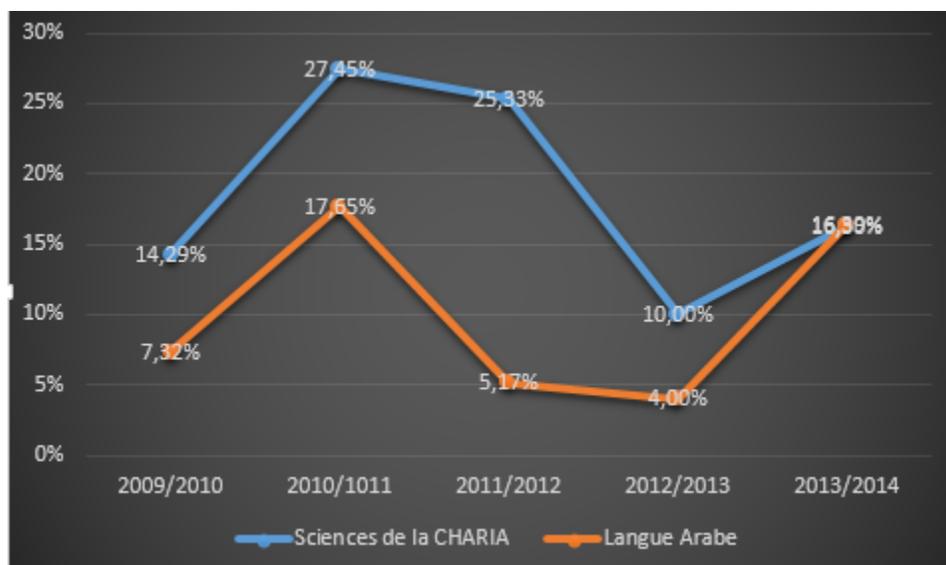
Règle de décision :

La statistique $\chi^2 = 11,2128$ est supérieure à la valeur critique $\chi^2_{0,05(\nu=4)} = 9,49$

On rejette l'hypothèse nulle au seuil de $\alpha = 0,05$. La probabilité d'obtenir de tels écarts sous H_0 est en fait de 0,024

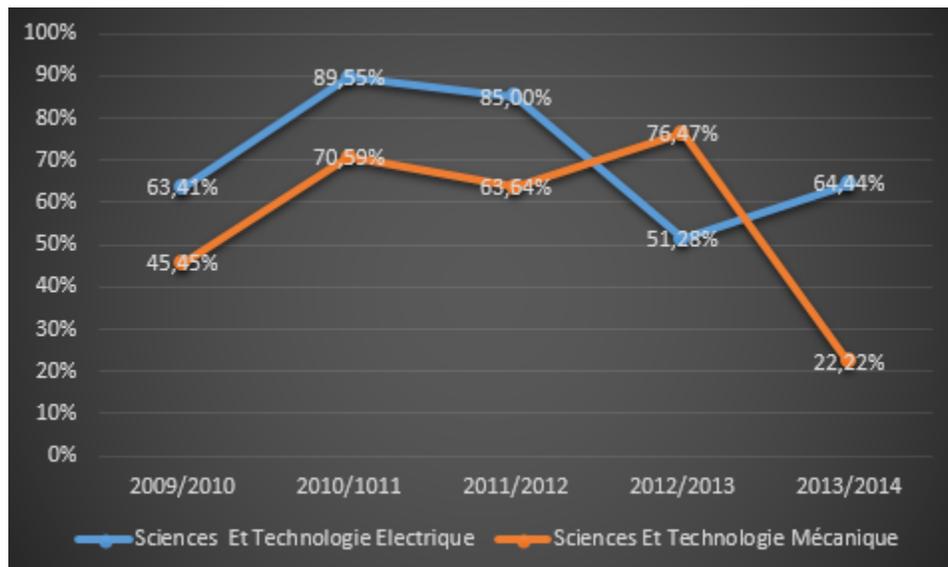
Conclusion :

On en conclut alors que le taux de réussite des deux filières a connu une différence significative ces cinq dernières années.



8.4 Enseignement Techniques Industrielles

| Option \ Année | 2009/2010 | 2010/2011 | 2011/2012 | 2012/2013 | 2013/2014 |
|-------------------------------------|-----------|-----------|-----------|-----------|-----------|
| Sciences Et Technologies Électrique | 63,41% | 89,55% | 85,00% | 51,28% | 64,44% |
| Sciences Et Technologie Mécanique | 45,45% | 70,59% | 63,64% | 76,47% | 22,22% |



$$\chi^2 = \sum Stat$$

| Option \ Année | 2009/2010 | 2010/2011 | 2011/2012 | 2012/2013 | 2013/2014 | Σ |
|-------------------------------------|---------------|---------------|---------------|---------------|--------------|---------------|
| Sciences Et Technologies Électrique | 63,41 | 89,55 | 85,00 | 51,28 | 64,44 | 353,68 |
| Sciences Et Technologie Mécanique | 45,45 | 70,59 | 63,64 | 76,47 | 22,22 | 278,37 |
| Σ | 108,86 | 160,14 | 148,64 | 127,75 | 86,66 | 632,05 |

On obtient comme statistique globale après calcul :

$$\chi^2 = 25,3618$$

Degré de liberté :

$$d.d.l = (2 - 1)(5 - 1) = 4$$

Règle de décision :

La statistique $\chi^2 = 25,3618$ est supérieure à la valeur critique $\chi^2_{0,05(\nu=4)} = 9,49$

On rejette l'hypothèse nulle au seuil de $\alpha = 0,05$. La probabilité d'obtenir de tels écarts sous H_0 est en fait de $0,000425$

Conclusion :

On en conclut alors que le taux de réussite des deux filières a connu une différence significative ces cinq dernières années.

Part V
Conclusion

Après avoir collecté et traité les données en main, on retient quelques conclusions :

Concernant les notes de la 3^{ème} année collégiale, et après avoir utilisé les tests nécessaires, on conclut que les moyennes du milieu urbain sont supérieures à celles du milieu rural et cela revient à plusieurs raisons, la majeure raison est la bonne qualité d'enseignement.

Par sexe, le rendement des filles est supérieur à celui des garçons.

Par matière, le niveau des élèves en matières scientifiques est plus bas par rapport à leur niveau en matières littéraires.

Concernant les taux de réussite du baccalauréat, pour l'enseignement général et originel, leur taux de réussite n'a pas connu une progression considérable.

Alors que, pour l'enseignement techniques commerciales et techniques industrielles, le taux de réussite a assez évolué.