

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA
FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la
VIE
DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : **Statistique**

Par

RIGUET Lamia

Titre :

Sur Quelques Tests Statistiques Et Applications

Membres du Comité d'Examen :

Pr.	BRAHIMI BRAHIM	UMKB	Président
Pr.	SAYAH ABDALLAH	UMKB	Encadreur
Dr.	ZOUAOUI NOUR EL HOUDA	UMKB	Examinatrice

27 Juin 2022

Dédicace

Je dédie ce mémoire à :

À mes chers parents

Mon père qui m'a beaucoup aidé pour accéder à ce degré de savoir, j'exprime toute ma gratitude pour le soutien et l'encouragement qu'ils m'ont apportée pour mener à bien mon travail.

Ma mère qui je ne pourrai jamais décrire sa tendresse et je la considère mon exemple de sacrifices et ma source de l'espoir durant toutes mes années d'étude.

À mes chères sœurs.

À mes chers frères.

REMERCIEMENTS

J'exprime d'abord mes profonds remerciements à "*ALLAH*" qui m'a donné le courage et la volonté d'achever ce travail.

Un grand merci au Dr **SAYAH Abdallah**, le directeur de mon mémoire, qui a dirigé et suivi mes travaux. Je remercie également tous les autres membres du jury qui ont acceptés de juger mon travail.

Je tiens à remercier toute l'équipe pédagogique d'UNIVERSITÉ MOHAMED KHIDER, BISKRA et les participants professionnels responsables de cette formation.

En fin, je remercie mes très chers parents et toutes ma famille et mes amis (Maroua Saad, Hana Massa, Sofia Drihem, Hind Sabegh, Fatima Zohra Settouti, Ahlam Dendouga, Lobna Kara, Sadika Yakhlef, Chihab eddine Menacer, younes Boukamoune, ...) qui ont toujours été là pour moi leurs soutien inconditionnel et leurs encouragements ont été d'une grande aide à tous ces intervenants, je présente mes remerciements, mon respect et ma gratitude.

L'étudiante : **RIGUET Lamia**

Table des matières

Remerciements	ii
Table des matières	iii
Table des figures	vi
Liste des tables	vii
Introduction	1
1 Généralités sur les tests	3
1.1 Caractéristique d'un échantillon	3
1.2 La statistique et Test statistique	5
1.2.1 La statistique	5
1.2.2 Test statistique	5
1.2.3 Le but des tests statistiques	6
1.3 Tests d'hypothèses, généralités	6
1.3.1 Hypothèses statistiques	6
1.3.2 Les deux espèces d'erreur	7
1.3.3 Région de rejet et niveau de signification	8

1.3.4	Variable de décision (la statistique de test)	9
1.3.5	P-valeur	9
1.3.6	Règle de décision	10
1.3.7	Fonction puissance	10
1.3.8	Les tests unilatéral et bilatéral	10
1.3.9	Catégories des tests	12
1.3.10	Démarche pour faire un test statistique	12
1.4	Test paramétrique et non paramétrique	13
1.4.1	Test paramétrique	13
1.4.2	Test non paramétrique	13
2	Tests paramétriques	14
2.1	Test de la moyenne	14
2.1.1	Variance connue	14
2.1.2	Variance inconnue	18
2.2	Test de la variance	19
2.2.1	Moyenne connue	20
2.2.2	Moyenne inconnue	21
2.3	Test de la fréquence (proportion)	23
2.4	Test de comparaison de deux moyennes	26
2.4.1	Variance connue	26
2.4.2	Variance inconnue et $\sigma_1 = \sigma_2 = \sigma$	27
2.4.3	Variance inconnue et $\sigma_1 \neq \sigma_2$	31
2.5	Test de comparaison de deux variances	32
2.6	Test de comparaison de deux proportions	34

3 Application sous R	37
3.1 Test de la moyenne	37
3.2 Test de la fréquence (proportion)	38
3.3 Test de comparaison de deux moyennes	39
3.4 Test de comparaison de deux variances	40
3.5 Test de comparaison de deux proportions	41
Conclusion	43
Bibliographie	44
Annexe A : Logiciel R	46
Annexe B : Notations et symbols	47

Table des figures

1.1	Test bilatéral et unilatéral (à gauche et à droite).	11
-----	--	----

Liste des tableaux

1.1	Table de décision d'un test statistique	8
2.1	Table de Test de la moyenne.	19
2.2	Table des prix des salles de mariage.	22
2.3	Table de Test de la variance.	23
2.4	Table de test de proportion.	25
2.5	Table de la solidité d'une tasse du premier lot.	28
2.6	Table de la solidité d'une tasse du deuxième lot.	28
2.7	pression artérielle (sains/hémoglobinopathie).	30
2.8	Table de test de comparaison deux moyennes.	32
2.9	Table de poid de cacao en poudre de machine 1.	33
2.10	Table de poid de cacao en poudre de machine 2.	33
2.11	Table de test de comparaison deux variances et deux proportions.	36
3.1	Table des commandes R des tests.	42

Introduction

En statistique, les tests statistiques sont des méthodes de la statistique qui permettent d'analyser des données obtenues, ils consistent à généraliser les propriétés constatées sur des observations d'un phénomène aléatoire de loi de probabilité inconnue.

Il existe deux types de tests : test paramétrique et test non paramétrique .

Il y' a beaucoup des chercheurs qui ont étudié les tests statistiques par exemple : mémoire Pierre Simon de Laplace (1749 – 1827) publié en 1773 et réédité en 1891, une étude heuristique des conditions sous lesquelles l'hypothèse H_0 devrait être rejetée est donnée dans le livre d'Émile Borel (1871 – 1956) Le Hasard 1914. À cette époque les idées concernant les tests statistiques étaient encore floues, Borel insistait sur l'existence d'un test statistique « *remarquable* » dont les propriétés pourraient être démontrées rigoureusement, il disait aussi que le choix du test devait être fait avant toute expérience.

Les notions d'erreurs de première et de seconde espèces, sous les noms de risque de producteur et de risque de consommateur, étaient introduites en 1929 dans les recherches appliquées consacrées au contrôle de qualité de Harold French Dodge (1893 – 1976) et Harry Gutelius Romig (1900 – 1985). Plus tard, ces notions furent réinventées par des théoriciens et sont devenues les notions de base.

Aujourd'hui, les tests statistiques sont utilisés dans tous les domaines par exemple : mathématiques, la médecine, la pharmacie, la psychologie, la sociologie, l'économie,

l'éducation,... etc. Nous verrons certaines d'entre elles appliquées dans les exemples de cette mémoire, donc la question que je peux poser : comment faire un test statistique dans la réalité et quel résultat pouvons-nous tirer de cette expérience ?

L'objectif principal de ce travail est de savoir comment certains tests statistiques sont effectués et appliqués dans la réalité, tout en connaissant les résultats que nous obtenons de chaque expérience.

Ce mémoire se compose de trois chapitres :

- **Premier chapitre** : Dans ce chapitre, j'ai donné un rappel de généralité des tests statistiques, qu'est-ce qu'un test statistique ?, son objectif, les étapes pour faire un test statistique, ... etc.

- **Deuxième chapitre** : Dans ce chapitre, j'ai étudié certains des tests statistiques avec des détails dans les cas existants, j'ai donné quelques exemples, et j'ai terminé l'étude de chaque test statistique par un résumé sous forme de tableau pour faciliter une meilleure compréhension.

- **Troisième chapitre** : Je termine cette mémoire par ce chapitre (simulation), dans chaque test statistique, j'ai choisi un exemple (étudié au deuxième chapitre) et je l'ai appliqué dans le logiciel **R**, j'ai pris les résultats et tiré une conclusion, j'ai terminé ce chapitre par un tableau résumé les commande R des tests.

Chapitre 1

Généralités sur les tests

Ce chapitre est consacré à un rappel sur les tests statistiques : l'échantillon, la statistique, test statistique, tests d'hypothèses, définition des tests paramétriques et non paramétriques.

1.1 Caractéristique d'un échantillon

- On appelle échantillon aléatoire de taille n une suite de n variables aléatoires indépendantes et de même loi, noté : X_1, X_2, \dots, X_n .
- La moyenne de l'échantillon ou la moyenne empirique est \bar{X} définie par :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (1.1)$$

- La variance de l'échantillon est S^2 définie par :

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (1.2)$$

- La variance empirique est \tilde{S}^2 définie par :

$$\tilde{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (1.3)$$

- Le moment empirique d'ordre r est M_r définie par :

$$M_r = \frac{1}{n} \sum_{i=1}^n X_i^r \quad (1.4)$$

- Le moment centré empirique d'ordre r est M'_r définie par :

$$M'_r = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^r \quad (1.5)$$

Remarque 1.1.1

$$E(\bar{X}) = \mu \quad (1.6)$$

ou : μ est la moyenne.

$$Var(\bar{X}) = \frac{\sigma^2}{n} \quad (1.7)$$

ou : σ^2 est la variance.

$$E(\tilde{S}^2) = \sigma^2 \quad (1.8)$$

$$E(S^2) = \frac{n-1}{n} \sigma^2 \quad (1.9)$$

$$E(M_r) = \mu_r \quad (1.10)$$

ou : μ_r est la moyenne d'ordre r .(voir [17], pages : 69,70,71])

1.2 La statistique et Test statistique

1.2.1 La statistique

Définition 1.2.1 *La statistique est la discipline qui étudie des phénomènes à travers la collecte de données, un traitement leur analyse, l'interprétation des résultats et leur présentation afin de rendre ces données compréhensibles par tous. C'est à la fois une branche des mathématiques appliquées.*

Exemple 1.2.1 *Répartition par classe d'âge d'une population de poissons. Si on veut caractériser la population de saumons qui vivent en abondance dans les océans Atlantique et Pacifique et dans les grands lacs, on va prélever quelques poissons (ces quelques poissons vont constituer l'échantillon). Puis on va mesurer leur âge, leur poids, leur taille, ... on va enfin chercher à extrapoler ces résultats à toute la population.*

Remarque 1.2.1 *On trouve les applications de la statistique dans tous les domaines, on cite quelques domaine : médecine, pharmacie, économie, marketing, sport, ...etc.*

1.2.2 Test statistique

Définition 1.2.2 *Un test statistique est une procédure de décision entre deux hypothèses. Il s'agit d'une démarche consistant à rejeter ou à ne pas rejeter une hypothèse statistique, appelée hypothèse nulle, en fonction d'un jeu de données (échantillon). On distingue deux types de test : test paramétrique et non paramétrique.*

Exemple 1.2.2 *Dans le domaine de la commercialisation : tester un produit marketing par exemple cosmétique, aide à comprendre son efficacité à atteindre l'objectif marketing et les tendances des clients.*

Exemple 1.2.3 *Dans le domaine pharmaceutique : il est obligatoire de tester l'efficacité d'un traitement avant de le mettre sur le marché. On procède alors de la façon suivante : on sélectionne deux groupes de patients. L'un reçoit le médicament, l'autre un placebo, il faut alors vérifier que le groupe qui reçoit le médicament voit bien ses symptômes diminuer en moyenne.*

1.2.3 Le but des tests statistiques

- ▶ Aider à la validation d'hypothèses.
- ▶ Permettre de tirer des conclusions claires, mathématiquement rigoureuses à partir des données.
- ▶ Ils permettent de réduire la subjectivité, en rendant les choix plus objectifs et plus transparents pour pouvoir les critiquer ...etc.

1.3 Tests d'hypothèses, généralités

1.3.1 Hypothèses statistiques

- Une **hypothèse statistique** est un énoncé (une affirmation) concernant les caractéristiques (valeurs des paramètres, forme de la distribution des observations) d'une population.
- Types d'hypothèses : On distingue deux types d'hypothèses :
 - ▶ **L'hypothèse nulle** : est celle dont le paramètre qu'on veut tester ne change pas, ce qui est symbolisé le plus souvent par une égalité. On la note H_0 .
 - ▶ **L'hypothèse alternative** : est la contre-hypothèse de l'hypothèse nulle. On la note H_1 .
- À partir des observations sur un échantillon, le test d'hypothèse permet donc de

déterminer s'il faut rejeter ou non l'hypothèse nulle.

Exemple 1.3.1 $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$

$H_0 : \mu = \mu_0$ contre $H_1 : \mu > \mu_0$

1.3.2 Les deux espèces d'erreur

- Dans un problème de décision qui permet à choisir H_0 ou H_1 , quatre situations sont possibles :

1. Décider que H_0 est vraie alors que H_0 est vraie.
2. Décider que H_1 est vraie alors que H_1 est vraie.
3. Décider que H_1 est vraie alors que H_0 est vraie.
4. Décider que H_0 est vraie alors que H_1 est vraie.

- Dans les deux premiers cas, la décision prise est bonne mais les deux derniers cas elle est erronée.

- L'erreur qui consiste à rejeter une hypothèse vraie appelée **erreur de première espèce** et sa probabilité appelée **risque de première espèce**. On le note : α :

$$\alpha = P(\text{rejeter } H_0 | H_0 \text{ est vraie}) = P(H_1 | H_0) \quad (1.11)$$

- L'erreur commise en acceptant une hypothèse fausse est appelée **erreur de deuxième espèce** et sa probabilité appelée **risque de deuxième espèce**. On le note : β :

$$\beta = \mathbb{P}(\text{accepter } H_0 | H_0 \text{ est fausse}) = \mathbb{P}(H_0 | H_1) \quad (1.12)$$

- On peut résumer ceci dans le tableau suivant :

Décision \ Vérité	H_0	H_1
H_0	$1 - \alpha$	β
H_1	α	$1 - \beta$

TAB. 1.1 – Table de décision d’un test statistique

Remarque 1.3.1 *La puissance du test est la probabilité de rejeter H_0 si H_0 est effectivement fausse, c’est-à-dire $1 - \beta$.*

1.3.3 Région de rejet et niveau de signification

Région de rejet

La région de rejet d’un test est l’ensemble des points (X_1, \dots, X_n) de \mathbb{R}^n pour lequel l’hypothèse nulle H_0 est écartée au profit de l’hypothèse alternative H_1 . On appelle aussi région critique du test et on la note généralement par W . Elle est définie par la relation :

$$\mathbb{P}(W|H_0) = \alpha \quad (1.13)$$

Le complémentaire de la région critique est appelée **région d’acceptation** du test. Elle est notée par \bar{W} et est définie par :

$$\mathbb{P}(\bar{W}|H_0) = 1 - \alpha \quad (1.14)$$

Niveau de signification

Le niveau de signification d’un test (ou seuil du test) est égal au risque de première espèce maximum. On le note par α où :

$$\alpha = \sup_{\theta \in \Theta_0} \alpha(\theta) \quad (1.15)$$

Dans la pratique des tests statistiques on fixe α .

Les valeurs usuelles de α sont 10%, 5%, 1%.

Remarque 1.3.2 - *Le test le plus puissant est le test qui fournit l'erreur β la plus petite, pour une même valeur de α .*

- *On dit qu'un test est sans biais si sa puissance est supérieure ou égale à son risque α .*

- *On dit qu'un test est convergent si et seulement si sa puissance tends vers 1 ($1 - \beta \rightsquigarrow 1$).*

1.3.4 Variable de décision (la statistique de test)

La statistique de test est une variable aléatoire construite à partir d'un échantillon aléatoire permettant de formuler une règle de décision pour le test. Elle n'est pas unique, ce qui permet de construire différentes règles de décision et de les comparer à l'aide de la notion de puissance. On optera pour une statistique de test contenant le maximum d'informations sur l'échantillon observé, de lois différentes sous H_0 ou H_1 .

1.3.5 P-valeur

En statistique la p-valeur est la probabilité pour un modèle statistique donne sous l'hypothèse nulle d'obtenir la même valeur ou une valeur encore plus extrême que celle observée sur l'échantillon. On va comparer un seuil de signification α et p-valeur en vue de savoir accepter ou rejeter H_0 :

-Si **p – valeur** $\leq \alpha$ on va rejeter l'hypothèse H_0 .

-Si **p – valeur** $> \alpha$ on va accepter l'hypothèse H_0 .

1.3.6 Règle de décision

Pour un niveau de signification α fixé, il existe deux façons de prendre la décision :

-Règle de décision 1 : si la statistique de test appartient à la région critique alors l'hypothèse H_0 est rejetée au risque d'erreur, sinon elle ne peut être rejetée.

-Règle de décision 2 : on compare entre p-valeur et niveau de signification α et on décide.

1.3.7 Fonction puissance

La fonction puissance d'un test est la fonction $\pi : \Theta_1 \rightarrow [0; 1]$ pour tout $\theta \in \Theta_1$ définie par :

$$\begin{aligned}\pi(\theta) &= \mathbb{P}(\text{rejeter } H_0 / H_0 \text{ fausse}) \\ &= 1 - \mathbb{P}(\text{accepter } H_0 / H_0 \text{ fausse}) \\ &= 1 - \beta(\theta)\end{aligned}\tag{1.16}$$

1.3.8 Les tests unilatéral et bilatéral

D'une manière générale, si nous définissons θ comme étant un paramètre inconnu d'une population sur laquelle nous voulons faire des tests d'hypothèses, nous disposons des trois possibilités suivantes :

1. Le **test bilatéral** : qui sert à tester si le paramètre inconnu est égal à une valeur fixe θ_0 (valeur initialement connue) ou s'il est différent de cette valeur. On l'appelle ainsi, car si jamais ce paramètre n'est pas égal à la valeur avec laquelle on le compare, il peut être soit supérieur, soit inférieur à cette valeur. On le symbolise par $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$.
2. Le **test unilatéral à droite** : qui sert à tester si le paramètre inconnu est égal

à une valeur fixe ou s'il est supérieur à cette valeur. On l'appelle ainsi, car si jamais ce paramètre n'est pas égal à valeur avec laquelle on le compare, il ne peut être que plus grand, donc à la droite de cette valeur. On le symbolise par $H_0 : \theta = \theta_0$ contre $H_1 : \theta > \theta_0$.

3. Le **test unilatéral à gauche** : qui sert à tester si le paramètre inconnu est égal à une valeur fixe ou s'il est inférieur à cette valeur. On l'appelle ainsi, car si jamais ce paramètre n'est pas égal à valeur avec laquelle on le compare, il ne peut être que plus petit, donc à la gauche de cette valeur. on le symbolise par $H_0 : \theta = \theta_0$ contre $H_1 : \theta < \theta_0$.

Les figures suivantes résument cela :

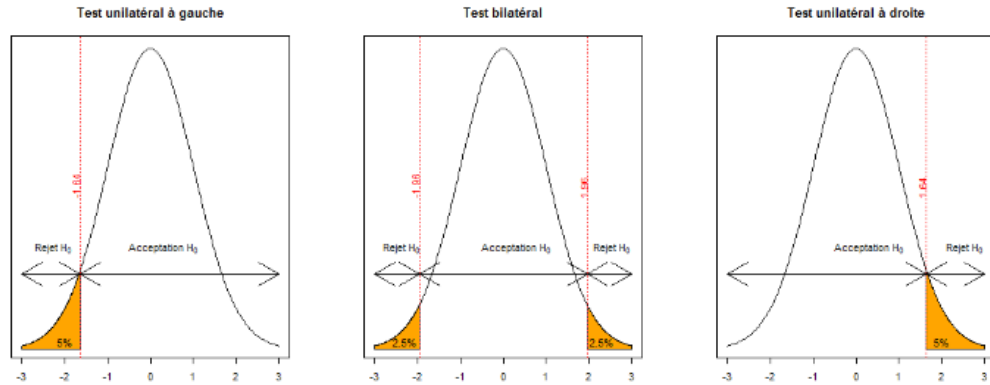


FIG. 1.1 – Test bilatéral et unilatéral (à gauche et à droite).

1.3.9 Catégories des tests

Nous avons quatre catégories principales de tests :

1. **Tests de conformité** : qui sont destinés à comparer entre un paramètre (moyenne, variance,..) d'une population de distribution et une valeur hypothétique du paramètre.
2. **Tests d'ajustements** : qui sont destinés à vérifier si un échantillon provient d'une population de distribution donnée.
3. **Tests d'indépendances** : qui sont pour le but de vérifier si deux ou plusieurs population sont indépendantes.
4. **Tests d'homogénéités** (comparaison) : Ils servent à vérifier si les paramètres de deux (ou plusieurs) populations de distributions connues sont égaux.

1.3.10 Démarche pour faire un test statistique

Il est possible de définir 6 étapes distinctes :

1. Choix de H_0 et H_1 .
2. Choix de la variable de décision (statistique du test).
3. Choix de α .
4. Calcul de la région critique.
5. Calcul de la valeur observée de la variable de décision.
6. Comparaison et conclusion : rejet de l'hypothèse si la valeur calculée \in région critique.

1.4 Test paramétrique et non paramétrique

1.4.1 Test paramétrique

Définition 1.4.1 *Un test paramétrique est un test pour lequel on fait une hypothèse sur la forme des données sous H_0 (normale, Poisson,...). Les hypothèses du test concernant alors les paramètres gouvernant cette loi.*

1.4.2 Test non paramétrique

Définition 1.4.2 *Un test non paramétrique est un test ne nécessitant pas d'hypothèse sur la forme des données. Les données sont alors remplacées par des statistiques ne dépendant pas des moyennes, variances des données initiales (tables de contingence, statistique d'ordre...).*

Chapitre 2

Tests paramétriques

Dans ce chapitre on étudie quelques tests paramétriques avec des détails sur les cas existants (moyenne et variance connue où inconnue et le test est bilatéral où unilatéral) dans chaque test et des exemples.

2.1 Test de la moyenne

On suppose que l'on a un échantillon X_1, X_2, \dots, X_n suit une loi normale $\mathcal{N}(\mu, \sigma^2)$. La moyenne de l'échantillon suit la loi $\mathcal{N}\left(\mu_0, \frac{\sigma^2}{n}\right)$ est :

$$\boxed{\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i} \quad (2.1)$$

2.1.1 Variance connue

La statistique de test est :

$$\boxed{Z = \frac{(\bar{X} - \mu)}{\sqrt{\frac{\sigma^2}{n}}} \rightsquigarrow \mathcal{N}(0, 1)} \quad (2.2)$$

(suit une loi normale centrée réduite).

a) Pour une hypothèse bilatérale $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$

- On rejette H_0 au risque α si : $|Z| \geq q_{1-\frac{\alpha}{2}}$ tq $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale, donc la région de rejet est : $]-\infty; -q_{1-\frac{\alpha}{2}}[\cup]q_{1-\frac{\alpha}{2}}; +\infty[$.

- Pour un risque d'erreur α fixé on a :

$$\mathbb{P}(|Z| \leq q_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$$

$$\Leftrightarrow \mathbb{P}(Z \leq q_{1-\frac{\alpha}{2}} \text{ où } Z \geq -q_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$$

$$\Leftrightarrow \mathbb{P}(Z \leq q_{1-\frac{\alpha}{2}}) + P(Z \geq -q_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$$

$$\Leftrightarrow \mathbb{P}(Z \leq q_{1-\frac{\alpha}{2}}) = P(Z \geq -q_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} \text{ (par la symétrie de la loi normale de moyenne 0 et de variance 1)}$$

$$\Leftrightarrow \Phi(q_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} \text{ (ou } \Phi \text{ est la fonction de répartition de la loi normale centrée réduite)}.$$

$$\Leftrightarrow q_{1-\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2}) \text{ (déduire d'après la table de la fonction de répartition inverse de la loi normale)}.$$

Exemple 2.1.1 Une entreprise qui possède plusieurs parcs de stationnement dans différentes parties d'une ville avec des revenus quotidiens moyens estimée à 60000 DA, la société a décidé d'augmenter le prix pour déterminer si les rendements quotidiens moyens changeraient. On s'attendait à ce que les rendements moyens augmentent, mais le manque à gagner était probablement dû à la perte de certains clients. Après que la demande se soit stabilisée sous l'augmentation des prix, un échantillon de 30 jours a été sélectionné montrant que le rendement quotidien moyen était de 55200 DA avec un écart-type de 7900 DA, en supposant que les rendements quotidiens moyens des positions de la société suivaient la loi normale. Les données de l'échantillon suggèrent-elles qu'il y a eu un changement dans le rendement quotidien moyen (à un niveau de 5%) ?

Donc : $\mu = 60000$; $n = 30$; $\bar{X} = 55200$; $\sigma = 7900$.

$$1. \text{ L'hypothèse est : } \begin{cases} H_0 : \mu = 60000 \\ H_1 : \mu \neq 60000 \end{cases}$$

$$2. \text{ La statistique de test (2.2) : } Z = \frac{|\bar{X} - \mu|}{\sqrt{\frac{\sigma^2}{n}}} = \frac{|55200 - 60000|}{\sqrt{\frac{(7900)^2}{30}}} = 3.3279 >$$

$$q_{1-\frac{\alpha}{2}} = q_{0.975} \simeq 1.88$$

$$3. \text{ Région de rejet : }]-\infty; -1.88[\cup]1.88; +\infty[$$

4. Règle de décision : Z se trouve dans la région de rejet. Donc on rejette l'hypothèse qui dire il y a un changement.

$$\text{b) Pour une hypothèse unilatérale à droite } \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$$

- On rejette H_0 au risque α si : $Z \geq q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi normale, donc la région de rejet est : $]q_{1-\alpha}; +\infty[$.

- Pour α fixé on a :

$$\mathbb{P}(Z \leq q_{1-\alpha}) = 1 - \alpha$$

$$\Leftrightarrow \Phi(q_{1-\alpha}) = 1 - \alpha$$

$$\Leftrightarrow q_{1-\alpha} = \Phi^{-1}(1 - \alpha).$$

Exemple 2.1.2 Dans un département Algérien, nous avons fait une enquête sur la durée de sommeil par nuit de 9.7 heures dans un groupe de 60 enfants. L'écart-type est 3 heures, la moyenne attendue de sommeil est de 12.1 heures chez les enfants de cet âge. Les enfants examinés dorment autant que ceux de la population générale ? (avec $\alpha = 0.05$).

Solution 2.1.1 $\mu = 9.7$; $n = 60$; $\sigma = 2$; $\bar{X} = 12.1$.

$$1. \text{ L'hypothèse est : } \begin{cases} H_0 : \mu = 9.7 \\ H_1 : \mu > 9.7 \end{cases}.$$

2. La statistique de test : $Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{12.1 - 9.7}{\sqrt{\frac{(3)^2}{60}}} = 6.1968 > q_{1-\alpha} = q_{0.95} \simeq 1.64$

3. La région de rejet : $]1.64; +\infty[=]1.64; +\infty[$

4. Règle de décision : On rejette l'hypothèse (car Z se trouve dans la région de rejet). Donc les enfants examinés présentent un temps de sommeil plus court que les enfants de la population générale.

c) Pour une hypothèse unilatérale à gauche $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$

- On rejette H_0 au risque α si : $Z \leq -q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi normale, donc la région de rejet est : $] -\infty; -q_{1-\alpha}[$.

- Pour un risque d'erreur α fixé on a :

$$\mathbb{P}(Z \geq -q_{1-\alpha}) = 1 - \alpha$$

$$\Leftrightarrow 1 - \mathbb{P}(Z < -q_{1-\alpha}) = 1 - \alpha$$

$$\Leftrightarrow \mathbb{P}(Z < -q_{1-\alpha}) = \alpha$$

$$\Leftrightarrow \Phi(-q_{1-\alpha}) = \alpha$$

$$\Leftrightarrow q_{1-\alpha} = -\Phi^{-1}(\alpha).$$

Exemple 2.1.3 Dans les petites villes où les élèves souffrent du manque de transport, un élève a besoin de 45 minutes à pied pour rejoindre l'école secondaire, un des responsables a voulu vérifier si le temps pris sur la route différait de 45 minutes, un échantillon aléatoire de 20 élèves a été tiré, il a été constaté que le temps moyen pris est de 39 minutes. L'écart type est de 18 minutes. Que concluons-nous ? (avec $\alpha = 0.05$).

On a : $\mu = 45$; $n = 20$; $\bar{X} = 39$; $\sigma = 18$.

1. L'hypothèse est : $\begin{cases} H_0 : \mu = 45 \\ H_1 : \mu < 45 \end{cases}$

2. La statistique de test : $Z = \frac{|\bar{X} - \mu|}{\sqrt{\frac{\sigma^2}{n}}} = \frac{|39 - 45|}{\sqrt{\frac{18^2}{20}}} = 1.4907 < q_{1-\alpha} = q_{0.95} \simeq 1.6449$

3. Région de rejet : $] -\infty; -1.6449[\cup]1.6449; +\infty[$

4. Règle de décision : on accepte l'hypothèse.

2.1.2 Variance inconnue

Comme la variance est inconnue on l'estime par la variance empirique :

$$\tilde{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (2.3)$$

alors la statistique de test est :

$$T = \frac{\bar{X} - \mu_0}{\sqrt{\frac{\tilde{S}^2}{n}}} \rightsquigarrow \mathcal{T}(n-1) \quad (2.4)$$

(suit la loi de student à $n - 1$ degrés de liberté).

a) Pour une hypothèse bilatérale $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$

On rejette H_0 au risque α si : $|T| \geq t_{1-\frac{\alpha}{2}}$ tq $t_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi de student, donc la région de rejet est : $] -\infty; -t_{1-\frac{\alpha}{2}}[\cup]t_{1-\frac{\alpha}{2}}; +\infty[$.

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$

On rejette H_0 au risque α si : $T \geq t_{1-\alpha}$ tq $t_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi de student, donc la région de rejet est : $]t_{1-\alpha}; +\infty[$.

c) Pour une hypothèse unilatérale à gauche $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$

On rejette H_0 au risque α si : $T \leq -t_{1-\alpha}$ tq $t_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la

loi de student, donc la région de rejet est : $] -\infty; -t_{1-\alpha}[$.

Remarque 2.1.1 Pour un risque d'erreur α fixé on a :

$$\mathbb{P}(|T| \leq t_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} \text{ (hypothèse bilatérale).}$$

$$\mathbb{P}(T \leq t_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à droite).}$$

$$\mathbb{P}(T \geq -t_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à gauche).}$$

Remarque 2.1.2 Lorsque le nombre d'observations n est grand (supérieur à 30), on peut utiliser le théorème de centrale limite pour approcher la statistique à la loi normale.

Ceci est résumé dans le tableau suivant :

	Test	Statistique de test	Région de rejet
σ^2 connue	bilatéral	$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$	$] -\infty; -q_{1-\frac{\alpha}{2}}[\cup] q_{1-\frac{\alpha}{2}}; +\infty [$
	unilatéral à droite		$] q_{1-\alpha}; +\infty [$
	unilatéral à gauche		$] -\infty; -q_{1-\alpha}[$
σ^2 inconnue	bilatéral	$T = \frac{\bar{X} - \mu_0}{\sqrt{\frac{\tilde{S}^2}{n}}}$	$] -\infty; -t_{1-\frac{\alpha}{2}}[\cup] t_{1-\frac{\alpha}{2}}; +\infty [$
	unilatéral à droite		$] t_{1-\alpha}; +\infty [$
	unilatéral à gauche		$] -\infty; -t_{1-\alpha}[$

TAB. 2.1 – Table de Test de la moyenne.

2.2 Test de la variance

On suppose que l'on a un échantillon qui suit une loi normale $\mathcal{N}(\mu, \sigma^2)$.

2.2.1 Moyenne connue

On définit S^2 comme suit :

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 \quad (2.5)$$

La statistique de test est :

$$V = \frac{nS^2}{\sigma_0^2} = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma_0} \right)^2 \rightsquigarrow \mathcal{X}_n^2 \quad (2.6)$$

où \mathcal{X}_n^2 est une loi de khi-deux à n degrés de libertés.

a) Pour une hypothèse bilatérale $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$

- On rejette H_0 au risque α si : $V \leq l_{\frac{\alpha}{2}}$ ou $V \geq l_{1-\frac{\alpha}{2}}$ tq $l_{\frac{\alpha}{2}}$ et $l_{1-\frac{\alpha}{2}}$ sont les quantiles d'ordre $\frac{\alpha}{2}$ et $1 - \frac{\alpha}{2}$ respectivement de la loi khi-deux \mathcal{X}_n^2 , donc la région de rejet est : $]0; l_{\frac{\alpha}{2}}[\cup]l_{1-\frac{\alpha}{2}}; +\infty[$.

- Pour α fixé on a :

$$\mathbb{P}(V \leq l_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$$

$$\Leftrightarrow F(l_{\frac{\alpha}{2}}) = \frac{\alpha}{2} \text{ (ou } F \text{ est la fonction de répartition de la loi } \mathcal{X}_n^2)$$

$$\Leftrightarrow l_{\frac{\alpha}{2}} = F^{-1}\left(\frac{\alpha}{2}\right).$$

$$\mathbb{P}(V \geq l_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$$

$$\Leftrightarrow \mathbb{P}(V < l_{1-\frac{\alpha}{2}}) = \frac{\alpha}{2}$$

$$\Leftrightarrow F(l_{1-\frac{\alpha}{2}}) = \frac{\alpha}{2}$$

$$\Leftrightarrow l_{1-\frac{\alpha}{2}} = F^{-1}\left(\frac{\alpha}{2}\right) \text{ (déduire d'après la table de la fonction de répartition inverse de la loi khi-deux).}$$

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 > \sigma_0^2 \end{cases}$

- On rejette H_0 au risque α si : $V \geq l_{1-\alpha}$ tq $l_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi khi-deux \mathcal{X}_n^2 , donc la région de rejet est : $]l_{1-\alpha}; +\infty[$.

- Pour α fixé on a :

$$\mathbb{P}(V \leq l_{1-\alpha}) = 1 - \alpha$$

$$\Leftrightarrow F(l_{1-\alpha}) = 1 - \alpha$$

$$\Leftrightarrow l_{1-\alpha} = F^{-1}(1 - \alpha).$$

c) Pour une hypothèse unilatérale à gauche $\left\{ \begin{array}{l} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 < \sigma_0^2 \end{array} \right.$

- On rejette H_0 au risque α si : $V \leq l_\alpha$ tq l_α est le quantile d'ordre α de la loi khi-deux \mathcal{X}_n^2 , donc la région de rejet est : $[0; l_\alpha[$.

- Pour α fixé on a :

$$\mathbb{P}(V \leq l_\alpha) = \alpha$$

$$\Leftrightarrow F(l_\alpha) = \alpha$$

$$\Leftrightarrow l_\alpha = F^{-1}(\alpha).$$

Remarque 2.2.1 Soit n le degrés de liberté alors, si $n > 30$ donc on a les deux approximations suivantes :

1. Approximation de Fisher : $\mathcal{X}_n^2 = \frac{1}{2} (q_{1-\alpha} + \sqrt{2n-1})^2$, où $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi normale.

La variable aléatoire $2\mathcal{X}_n^2 - (q_{1-\alpha} + \sqrt{2n-1})^2$ suit la loi $\mathcal{N}(0, 1)$. Par exemple

Si $\alpha = 0.05$ alors $q_{1-\alpha} = 1.64$

2. Approximation de Wilson Hilferty : $\mathcal{X}_n^2 = p \left(\sqrt{\frac{2}{9n}} q_{1-\alpha} - \frac{2}{9n} + 1 \right)^3$.

2.2.2 Moyenne inconnue

On va estimer la moyenne et la variance à l'aide des formules [2.1](#) et [2.3](#)

La statistique de test est :

$$U = \frac{(n-1)\tilde{S}^2}{\sigma_0^2} = \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma_0} \right)^2 \rightsquigarrow \mathcal{X}_{n-1}^2 \quad (2.7)$$

où \mathcal{X}_{n-1}^2 est une loi de khi-deux à $(n-1)$ degrés de libertés.

a) **Pour une hypothèse bilatérale** $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$

On rejette H_0 au risque α si : $U \leq k_{\frac{\alpha}{2}}$ ou $U \geq k_{1-\frac{\alpha}{2}}$ tq $k_{\frac{\alpha}{2}}$ et $k_{1-\frac{\alpha}{2}}$ sont les quantiles d'ordre $\frac{\alpha}{2}$ et $1 - \frac{\alpha}{2}$ respectivement de la loi \mathcal{X}_{n-1}^2 , donc la région de rejet est : $]0; k_{\frac{\alpha}{2}}[\cup]k_{1-\frac{\alpha}{2}}; +\infty[$.

Exemple 2.2.1 *Les hôteliers en Algérie affirment que l'écart type pour les prix des salles de mariage est de 4000 DA, un échantillon aléatoire de 14 hôtels a été retiré, le tableau suivant indique les prix :*

Hotel	1	2	3	4	5	6	7
Prix(DA)	88000	76000	150000	110000	166000	195000	200000
Hotel	8	9	10	11	12	13	14
Prix(DA)	90000	72000	150000	186000	145000	122000	130000

TAB. 2.2 – Table des prix des salles de mariage.

Soutenez-vous l'affirmation des gérants d'hôtel ? ($\alpha = 5\%$)

Solution 2.2.1 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = 1.2036 \times 10^5$

$$\tilde{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = 2.0883 \times 10^9$$

$$\sigma_0^2 = 4000$$

1. *Le test est :* $\begin{cases} H_0 : \sigma^2 = 4000 \\ H_1 : \sigma^2 \neq 4000 \end{cases}$

2. *La statistique de test est :* $U = \frac{(n-1)\tilde{S}^2}{\sigma_0^2} = \frac{13 \times 2.0883 \times 10^9}{4000} = 6.7870 \times 10^6$

3. $k_{1-\frac{\alpha}{2}} = k_{1-\frac{0.05}{2}} = k_{0.975} = 5.63$

4. Règle de décision : On rejette H_0 car $U > k_{1-\frac{\alpha}{2}}$.

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 > \sigma_0^2 \end{cases}$

On rejette H_0 au risque α si : $U \geq k_{1-\alpha}$ tq $k_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi χ_{n-1}^2 , donc la région de rejet est : $]k_{1-\alpha}; +\infty[$.

c) Pour une hypothèse unilatérale à gauche $\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 < \sigma_0^2 \end{cases}$

On rejette H_0 au risque α si : $U \leq k_\alpha$ tq k_α est le quantile d'ordre α de la loi khi-deux χ_{n-1}^2 , donc la région de rejet est : $[0; k_\alpha[$.

Remarque 2.2.2 pour un risque d'erreur α fixé on a :

$\mathbb{P}(k_{\frac{\alpha}{2}} \leq U) = \frac{\alpha}{2}$ et $P(U \leq k_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$ (hypothèse bilatérale).

$\mathbb{P}(U \leq k_{1-\alpha}) = 1 - \alpha$ (hypothèse unilatérale à droite).

$\mathbb{P}(U \leq k_\alpha) = \alpha$ (hypothèse unilatérale à gauche).

Ceci est résumé dans le tableau suivant :

	Test	Statistique de test	Région de rejet
μ connue	bilatéral	$V = \frac{nS^2}{\sigma_0^2}$	$]0; l_{\frac{\alpha}{2}}[\cup]l_{1-\frac{\alpha}{2}}; +\infty[$
	unilatéral à droite		$]l_{1-\alpha}; +\infty[$
	unilatéral à gauche		$[0; l_\alpha[$
μ inconnue	bilatéral	$U = \frac{(n-1)\tilde{S}^2}{\sigma_0^2}$	$]0; k_{\frac{\alpha}{2}}[\cup]k_{1-\frac{\alpha}{2}}; +\infty[$
	unilatéral à droite		$]k_{1-\alpha}; +\infty[$
	unilatéral à gauche		$[0; k_\alpha[$

TAB. 2.3 – Table de Test de la variance.

2.3 Test de la fréquence (proportion)

On dispose d'une population dans laquelle chaque individu présente ou non un certain caractère, la proportion d'individus présentant le caractère étant notée p , et

Y_n un échantillon aléatoire de taille n extrait de cette population. La proportion F calculée à partir de l'échantillon est considérée comme une réalisation d'une va de loi binomiale $B(n, p)$ qu'on peut assimiler si n est assez grand à une loi normale $\mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$.

La statistique de test est :

$$L = \frac{F - p}{\sqrt{\frac{p(1-p)}{n}}} \rightsquigarrow \mathcal{N}(0, 1) \quad (2.8)$$

(d'après le théorème de centrale limite).

a) Pour une hypothèse bilatérale $\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$

On la note par $F \left(F = \frac{Y_n}{n} \right)$ la fréquence empirique de p .

on rejette H_0 au risque α si : $|L| \geq q_{1-\frac{\alpha}{2}}$ tq $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}(0, 1)$, donc la région de rejet est : $] -\infty; -q_{1-\frac{\alpha}{2}} [\cup] q_{1-\frac{\alpha}{2}}; +\infty [$.

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : p = p_0 \\ H_1 : p > p_0 \end{cases}$

on rejette H_0 au risque α si : $L \geq q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{N}(0, 1)$, donc la région de rejet est : $] q_{1-\alpha}; +\infty [$.

Exemple 2.3.1 *Je pense que le taux de fraudeurs dans le programme d'allocation chômage en Algérie a augmenté ces derniers mois, il y a 3 mois le taux était de 15% du nombre de jeunes chômeurs. Pour vérifier l'hypothèse, 500 jeunes sont tirés au sort parmi ceux qui ont bénéficié de la bourse et un audit est réalisé. Il y a exactement 45 personnes qui ont triché. L'hypothèse est-elle validée ? (avec un niveau de $\alpha = 5\%$).*

Solution 2.3.1 1. Pour déterminer si l'hypothèse est vérifiée, on prend la v.a

$X : 1$ s'il ya un fraude, 0 sinon, donc $X \rightsquigarrow \text{Bin}(1, p_0)$ tel que p_0 es le taux de

fraudeurs.

$$2. \text{ le test est : } \begin{cases} H_0 : p = 0.15 \\ H_1 : p > 0.15 \end{cases}$$

$$3. \text{ La statistique de test : } L = \frac{|F - p|}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\left| \frac{45}{500} - 0.15 \right|}{\sqrt{\frac{0.15(1-0.15)}{500}}} = 3.7573 >$$

$$q_{1-\alpha} = q_{0.95} \simeq 1.64$$

4. La région de rejet est : $]1.64; +\infty[$.

5. On rejette l'hypothèse H_0 .

$$\text{c) Pour une hypothèse unilatérale à gauche } \begin{cases} H_0 : p = p_0 \\ H_1 : p < p_0 \end{cases}$$

on rejette H_0 au risque α si : $L \leq q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{N}(0, 1)$, donc la région de rejet est : $] -\infty; -q_{1-\alpha}[$.

Remarque 2.3.1 pour un risque d'erreur α fixé on a :

$$\mathbb{P}(|L| \leq q_{1-\frac{\alpha}{2}}) = 1 - \alpha \text{ (hypothèse bilatérale).}$$

$$\mathbb{P}(L \leq q_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à droite).}$$

$$\mathbb{P}(L \geq q_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à gauche).}$$

Ceci est résumé dans le tableau suivant :

Test	Statistique de test	Région de rejet
bilatéral	$L = \frac{F - p}{\sqrt{\frac{p(1-p)}{n}}}$	$] -\infty; -q_{1-\frac{\alpha}{2}}[\cup] q_{1-\frac{\alpha}{2}}; +\infty[$
unilatéral à droite		$] q_{1-\alpha}; +\infty[$
unilatéral à gauche		$] -\infty; -q_{1-\alpha}[$

TAB. 2.4 – Table de test de proportion.

2.4 Test de comparaison de deux moyennes

Soient $(X_1^{(1)}, X_2^{(1)}, \dots, X_{n_1}^{(1)})$ et $(X_1^{(2)}, X_2^{(2)}, \dots, X_{n_2}^{(2)})$ deux échantillons indépendants de tailles n_1 et n_2 respectivement, issus des variables aléatoire X_1 et X_2 tq : $X_1 \rightsquigarrow \mathcal{N}(\mu_1, \sigma_1)$ et $X_2 \rightsquigarrow \mathcal{N}(\mu_2, \sigma_2)$.

2.4.1 Variance connue

On a :

$$\bar{X}_{n_1} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i^{(1)} \rightsquigarrow \mathcal{N}\left(\mu_1, \frac{\sigma_1^2}{n_1}\right) \quad (2.9)$$

$$\bar{X}_{n_2} = \frac{1}{n_2} \sum_{i=1}^{n_2} X_i^{(2)} \rightsquigarrow \mathcal{N}\left(\mu_2, \frac{\sigma_2^2}{n_2}\right) \quad (2.10)$$

La statistique de test est :

$$Z^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \rightsquigarrow \mathcal{N}(0, 1) \quad (2.11)$$

a) Pour une hypothèse bilatérale $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$

On rejette H_0 au risque α si : $|Z^*| \geq q_{1-\frac{\alpha}{2}}$ tq $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale, donc la région de rejet est : $] -\infty; -q_{1-\frac{\alpha}{2}}[\cup] q_{1-\frac{\alpha}{2}}; +\infty[$.

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$

On rejette H_0 au risque α si : $Z^* \geq q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi normale, donc la région de rejet est : $] q_{1-\alpha}; +\infty[$.

c) **Pour une hypothèse unilatérale à gauche** $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$

On rejette H_0 au risque α si : $Z^* \leq -q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi normale, donc la région de rejet est : $] -\infty; -q_{1-\alpha}[$.

Remarque 2.4.1 Pour un risque d'erreur α fixé on a :

$$\mathbb{P}(|Z^*| \leq q_{1-\frac{\alpha}{2}}) = 1 - \alpha \text{ (hypothèse bilatérale).}$$

$$\mathbb{P}(Z^* \leq q_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à droite).}$$

$$\mathbb{P}(Z^* \geq -q_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à gauche).}$$

2.4.2 Variance inconnue et $\sigma_1 = \sigma_2 = \sigma$

On a :

$$\tilde{S}_{n_1}^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} \left(X_i^{(1)} - \bar{X}_{n_1} \right)^2 \quad (2.12)$$

$$\tilde{S}_{n_2}^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} \left(X_i^{(2)} - \bar{X}_{n_2} \right)^2 \quad (2.13)$$

i) $n_1 + n_2 - 2 < 30$

On pose

$$S^2 = \frac{n_1 \tilde{S}_{n_1}^2 + n_2 \tilde{S}_{n_2}^2}{n_1 + n_2 - 2} \quad (2.14)$$

La statistique de test est :

$$T^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightsquigarrow \mathcal{T}(n_1 + n_2 - 2) \quad (2.15)$$

(suit la loi de student à $(n_1 + n_2 - 2)$ degrés de liberté).

a) **Pour une hypothèse bilatérale** $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$

On rejette H_0 au risque α si : $|T^*| \geq t_{1-\frac{\alpha}{2}}$ tq $t_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{T}(n_1 + n_2 - 2)$, donc la région de rejet est : $]-\infty; -t_{1-\frac{\alpha}{2}}[\cup]t_{1-\frac{\alpha}{2}}; +\infty[$.

Exemple 2.4.1 On considère deux lots de tasses et on souhaite comparer la solidité de ces tasses. Pour chacun des deux lots, on dispose d'un échantillon de 8 tasses et on mesure la résistance de chacune d'entre eux. Les résultats sont :

- Pour le premier échantillon :

31.8	31.4	32.1	32.8	31.6	32.5	32.6	31.7
------	------	------	------	------	------	------	------

TAB. 2.5 – Table de la solidité d'une tasse du premier lot.

- pour le deuxième échantillon :

31.2	31.7	32.9	32.2	31.5	32.3	31.3	32.4
------	------	------	------	------	------	------	------

TAB. 2.6 – Table de la solidité d'une tasse du deuxième lot.

La solidité d'une tasse du premier lot peut être modélisée par une $\text{Var}(X_1)$, et celle du tasse du second lot peut être modélisée par une $\text{Var}(X_2)$. On suppose que X_1 et X_2 suivent des lois normales de variances égales.

Peut-on affirmer que ces deux échantillons ne proviennent pas de la même production ? (au niveau 20%) .

Solution 2.4.1 1. On a :

$$\left\{ \begin{array}{l} \bar{X}_{n_1} \simeq 32.06 \\ \bar{X}_{n_2} \simeq 31.93 \\ \tilde{S}_{n_1}^2 \simeq 0.35 \\ \tilde{S}_{n_2}^2 \simeq 0.36 \\ S \simeq 0.4 \end{array} \right.$$

$$2. T^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{32.06 - 31.93}{0.4 \times \sqrt{\left(\frac{2}{8}\right)}} = 0.65$$

$$3. t_{1-\frac{\alpha}{2}} = t_{0.9} \simeq 0.13$$

4. Règle de décision : $T^* > t_{1-\frac{\alpha}{2}}$ donc on rejette H_0 .

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$

On rejette H_0 au risque α si : $T^* \geq t_{1-\alpha}$ tq $t_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{T}(n_1 + n_2 - 2)$, donc la région de rejet est : $]t_{1-\alpha}; +\infty[$.

c) Pour une hypothèse unilatérale à gauche $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$

On rejette H_0 au risque α si : $T^* \leq -t_{1-\alpha}$ tq $t_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{T}(n_1 + n_2 - 2)$, donc la région de rejet est : $]-\infty; -t_{1-\alpha}[$.

Remarque 2.4.2 Pour un risque d'erreur α fixé on a :

$$\mathbb{P}(|T^*| \leq t_{1-\frac{\alpha}{2}}) = 1 - \alpha \text{ (hypothèse bilatérale).}$$

$$\mathbb{P}(T^* \leq t_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à droite).}$$

$$\mathbb{P}(T^* \geq -t_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à gauche).}$$

ii) $n_1 + n_2 - 2 > 30$

La statistique de test est :

$$T^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{\sqrt{\frac{\tilde{S}_{n_1}^2}{n_1} + \frac{\tilde{S}_{n_2}^2}{n_2}}} \rightsquigarrow \mathcal{N}(0, 1) \quad (2.16)$$

a) Pour une hypothèse bilatérale $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$

On rejette H_0 au risque α si : $|T^*| \geq q_{1-\frac{\alpha}{2}}$ tq $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi normale, donc la région de rejet est : $]-\infty; -q_{1-\frac{\alpha}{2}}[\cup]q_{1-\frac{\alpha}{2}}; +\infty[$.

Exemple 2.4.2 On désire comparer la pression artérielle diastolique d'un groupe de

sujets sains et d'un groupe de sujets atteints de drépanocytose (hémoglobinopathie).

Une étude donne les résultats suivants :

	n	Pression artérielle diastolique moyenne (mm Hg)	Variance (S_n^2)
Sains	62	60.1	8.8
Hémoglobinopathie	58	50.6	5

TAB. 2.7 – pression artérielle (sains/hémoglobinopathie).

La pression artérielle est différente chez les sujets drépanocytaires ? (au niveau 10%).

Solution 2.4.2 On a $n_1 + n_2 - 2 > 30$, σ_1 et σ_2 inconnue donc :

$$1. \text{ Le test : } \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

$$2. \text{ La statistique de test : } T^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{\sqrt{\frac{\tilde{S}_{n_1}^2}{n_1} + \frac{\tilde{S}_{n_2}^2}{n_2}}} = \frac{60.1 - 50.6}{\sqrt{\frac{8.8^2}{62} + \frac{5^2}{58}}} = 7.3293 > q_{1-\frac{\alpha}{2}} =$$

$$q_{0.95} = 1.6449$$

$$3. \text{ Région critique : }]-\infty; -1.6449[\cup]1.6449; +\infty[$$

4. Règle de décision : on rejette H_0 .

$$\text{b) Pour une hypothèse unilatérale à droite } \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$$

On rejette H_0 au risque α si : $T^* \geq q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi normale, donc la région de rejet est : $]q_{1-\alpha}; +\infty[$.

$$\text{c) Pour une hypothèse unilatérale à gauche } \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$$

On rejette H_0 au risque α si : $T^* \leq -q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi normale, donc la région de rejet est : $] -\infty; -q_{1-\alpha}[$.

Remarque 2.4.3 Pour un risque d'erreur α fixé on a :

$$\mathbb{P}(|T^*| \leq q_{1-\frac{\alpha}{2}}) = 1 - \alpha \text{ (hypothèse bilatérale).}$$

$\mathbb{P}(T^* \leq q_{1-\alpha}) = 1 - \alpha$ (*hypothèse unilatérale à droite*).

$\mathbb{P}(T^* \geq -q_{1-\alpha}) = 1 - \alpha$ (*hypothèse unilatérale à gauche*).

2.4.3 Variance inconnue et $\sigma_1 \neq \sigma_2$

$$n_1 + n_2 - 2 \leq 30$$

La statistique de test est :

$$T^* = \frac{X_{n_1} - X_{n_2}}{\sqrt{\frac{\tilde{S}_{n_1}^2}{n_1} + \frac{\tilde{S}_{n_2}^2}{n_2}}} \rightsquigarrow \mathcal{T}(n_1 + n_2 - 2) \quad (2.17)$$

a) Pour une hypothèse bilatérale $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$

On rejette H_0 au risque α si : $|T^*| \geq t_{1-\frac{\alpha}{2}}$ tq $t_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{T}(n_1 + n_2 - 2)$, donc la région de rejet est : $] -\infty; -t_{1-\frac{\alpha}{2}}[\cup] t_{1-\frac{\alpha}{2}}; +\infty[$.

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$

On rejette H_0 au risque α si : $T^* \geq t_{1-\alpha}$ tq $t_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{T}(n_1 + n_2 - 2)$, donc la région de rejet est : $] t_{1-\alpha}; +\infty[$.

c) Pour une hypothèse unilatérale à gauche $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$

On rejette H_0 au risque α si : $T^* \leq -t_{1-\alpha}$ tq $t_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{T}(n_1 + n_2 - 2)$, donc la région de rejet est : $] -\infty; -t_{1-\alpha}[$.

Ceci est résumé dans le tableau suivant :

cdt : condition

T : test

S-de-test : statistique de test

b : test bilatéral

uàd : test unilatéral à droite

uàg : unilatéral à gauche

cdt			T	S-de-test	Région critique
σ^2 est connue			b	$Z^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \rightsquigarrow \mathcal{N}(0, 1)$	$] -\infty; -q_{1-\frac{\alpha}{2}} [\cup] q_{1-\frac{\alpha}{2}}; +\infty [$
			uàd		$] q_{1-\frac{\alpha}{2}}; +\infty [$
			uàg		$] -\infty; -q_{1-\frac{\alpha}{2}} [$
σ^2 est inconnue	$\sigma_1 = \sigma_2$	$n_1 + n_2 - 2 < 30$	b	$T^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightsquigarrow \mathcal{T}(n_1 + n_2 - 2)$	$] -\infty; -t_{1-\frac{\alpha}{2}} [\cup] t_{1-\frac{\alpha}{2}}; +\infty [$
			uàd		$] t_{1-\alpha}; +\infty [$
			uàg		$] -\infty; -t_{1-\frac{\alpha}{2}} [$
		$n_1 + n_2 - 2 > 30$	b	$T^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{\sqrt{\frac{\tilde{S}_{n_1}^2}{n_1} + \frac{\tilde{S}_{n_2}^2}{n_2}}} \rightsquigarrow \mathcal{N}(0, 1)$	$] -\infty; -t_{1-\frac{\alpha}{2}} [\cup] t_{1-\frac{\alpha}{2}}; +\infty [$
			uàd		$] t_{1-\alpha}; +\infty [$
			uàg		$] -\infty; -t_{1-\frac{\alpha}{2}} [$
	$\sigma_1 \neq \sigma_2$	$n_1 + n_2 - 2 \leq 30$	b	$T^* = \frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{\sqrt{\frac{\tilde{S}_{n_1}^2}{n_1} + \frac{\tilde{S}_{n_2}^2}{n_2}}} \rightsquigarrow \mathcal{T}(n_1 + n_2 - 2)$	$] 0; -t_{1-\frac{\alpha}{2}} [\cup] t_{1-\frac{\alpha}{2}}; +\infty [$
			uàd		$] t_{1-\alpha}; +\infty [$
			uàg		$] -\infty; -t_{1-\frac{\alpha}{2}} [$

TAB. 2.8 – Table de test de comparaison deux moyennes.

2.5 Test de comparaison de deux variances

Soient $(X_1^{(1)}, X_2^{(1)}, \dots, X_{n_1}^{(1)})$ et $(X_1^{(2)}, X_2^{(2)}, \dots, X_{n_2}^{(2)})$ deux échantillons indépendants de tailles n_1 et n_2 respectivement, issus des variables aléatoire X_1 et X_2 tq : $X_1 \rightsquigarrow \mathcal{N}(\mu_1, \sigma_1)$ et $X_2 \rightsquigarrow \mathcal{N}(\mu_2, \sigma_2)$. (à l'aide des formules : [2.12](#) et [2.13](#))

La statistique de test est :

$$V^* = \frac{\tilde{S}_{n_1}^2 n_1}{(n_1 - 1) \sigma_1^2} \times \frac{(n_2 - 1) \sigma_2^2}{\tilde{S}_{n_2}^2 n_2} \rightsquigarrow \mathcal{F}(n_1 - 1, n_2 - 1) \quad (2.18)$$

où $\mathcal{F}(n_1 - 1, n_2 - 1)$ est loi de fisher à $(n_1 - 1, n_2 - 1)$ degrés de liberté.

et pour $i = 1, 2$ on a : $\frac{\tilde{S}_{n_i}^2}{(n_i - 1)\sigma_i^2} \rightsquigarrow \chi^2(n_i - 1)$

a) **Pour une hypothèse bilatérale** $\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 \neq \sigma_2^2 \end{cases}$

On rejette H_0 au risque α si : $V^* \leq c_{\frac{\alpha}{2}}$ ou $V^* \geq c_{1-\frac{\alpha}{2}}$ tq $c_{\frac{\alpha}{2}}$ et $c_{1-\frac{\alpha}{2}}$ sont les quantiles d'ordre $\frac{\alpha}{2}$ et $1 - \frac{\alpha}{2}$ respectivement de la loi $\mathcal{F}(n_1 - 1, n_2 - 1)$, donc la région de rejet est : $]0; c_{\frac{\alpha}{2}}[\cup]c_{1-\frac{\alpha}{2}}; +\infty[$.

Exemple 2.5.1 *La société de Monsieur Labrador utilise deux machines, machine 1 et machine 2, pour remplir automatiquement des paquets de cacao en poudre.*

On prélève un échantillon de 9 paquets remplis par la machine 1 et on les pèse, les résultats(en grammes) sont :

107	106.1	107.3	106.2	106.7	107.3	107.7	106.7	107.2
-----	-------	-------	-------	-------	-------	-------	-------	-------

TAB. 2.9 – Table de poid de cacao en poudre de machine 1.

On prélève un échantillon de 7 paquets remplis par la machine 2 et on les pèse, les résultats(en grammes) sont :

106.2	107.1	107.4	107.0	106.4	107.3	107.6
-------	-------	-------	-------	-------	-------	-------

TAB. 2.10 – Table de poid de cacao en poudre de machine 2.

On suppose que : $X_1 \rightsquigarrow \mathcal{N}(\mu_1, 1.2^2)$ (rempli à la machine 1) et $X_1 \rightsquigarrow \mathcal{N}(\mu_2, 0.7^2)$ (rempli à la machine 2). Peut-on affirmer que les machines sont réglées de manière différente ? (au niveau 5%).

Solution 2.5.1 1. On a : $\begin{cases} \bar{X}_{n_1} = 106.91 \\ \bar{X}_{n_2} \simeq 107 \\ \tilde{S}_{n_1}^2 \simeq 0.28 \\ \tilde{S}_{n_2}^2 \simeq 0.27 \end{cases}$

$$2. \text{ La statistique de test : } V^* = \frac{\tilde{S}_{n_1}^2 n_1}{(n_1 - 1) \sigma_1^2} \times \frac{(n_2 - 1) \sigma_2^2}{\tilde{S}_{n_2}^2 n_2} = \frac{0.28 \times 9}{(9 - 1) \times 1.2^2} \times \frac{(7 - 1) \times 0.7^2}{0.27 \times 7} = 0.34028$$

$$3. c_{1-\frac{\alpha}{2}} = c_{0.975} = 5.60$$

4. Règle de décision : $V^* < c_{1-\frac{\alpha}{2}}$ donc on accepte H_0 .

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 > \sigma_2^2 \end{cases}$

On rejette H_0 au risque α si : $V^* \geq c_{1-\alpha}$ tq $c_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{F}(n_1 - 1, n_2 - 1)$, donc la région de rejet est : $]c_{1-\alpha}; +\infty[$.

c) Pour une hypothèse unilatérale à gauche $\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 < \sigma_2^2 \end{cases}$

On rejette H_0 au risque α si : $V^* \leq c_\alpha$ tq c_α est le quantile d'ordre α de la loi $\mathcal{F}(n_1 - 1, n_2 - 1)$, donc la région de rejet est : $[0; c_\alpha[$.

Remarque 2.5.1 Pour un risque d'erreur α fixé on a :

$$\mathbb{P}(c_{\frac{\alpha}{2}} \leq V^* \leq c_{1-\frac{\alpha}{2}}) = 1 - \alpha \text{ (hypothèse bilatérale).}$$

$$\mathbb{P}(V^* \leq c_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à droite).}$$

$$\mathbb{P}(V^* \leq c_\alpha) = \alpha \text{ (hypothèse unilatérale à gauche).}$$

2.6 Test de comparaison de deux proportions

Soient $(X_1^1, X_2^1, \dots, X_{n_1}^1)$ et $(X_1^2, X_2^2, \dots, X_{n_2}^2)$ deux échantillons indépendants de tailles n_1 et n_2 respectivement, issus des variables aléatoire *iid* de loi de Bernoulli $\mathcal{B}(p_1)$ et $\mathcal{B}(p_2)$ respectivement.

On a :

$$\boxed{T_1 = \sum_{i=1}^{n_1} X_i^1 \rightsquigarrow \mathcal{N}(n_1 p_1, (1 - p_1) n_1 p_1)} \quad (2.19)$$

et

$$T_2 = \sum_{i=1}^{n_2} X_i^2 \rightsquigarrow \mathcal{N}(n_2 p_2, (1 - p_2) n_2 p_2) \quad (2.20)$$

Soient F_1 et F_2 deux fréquences empiriques associées à l'échantillon de p_1 et p_2 tq :

$$F_1 = \frac{T_1}{n_1} \rightsquigarrow \mathcal{N}\left(p_1, \frac{p_1(1-p_1)}{n_1}\right) \quad (2.21)$$

et

$$F_2 = \frac{T_2}{n_2} \rightsquigarrow \mathcal{N}\left(p_2, \frac{p_2(1-p_2)}{n_2}\right) \quad (2.22)$$

Donc :

$$F_1 - F_2 \rightsquigarrow \mathcal{N}\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} - \frac{p_2(1-p_2)}{n_2}\right) \quad (2.23)$$

La statistique de test (Sous l'hypothèse $H_0 : (p_1 = p_2 = p)$) est :

$$L^* = \frac{F_1 - F_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \rightsquigarrow \mathcal{N}(0, 1) \quad (2.24)$$

- Si p est inconnue on la remplace par son estimation (f_1 et f_2 sont les estimations de p_1 et p_2) : $\hat{p} = \frac{n_1 f_1 + n_2 f_2}{n_1 + n_2}$.

a) Pour une hypothèse bilatérale $\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{cases}$

on rejette H_0 au risque α si : $|L^*| \geq q_{1-\frac{\alpha}{2}}$ tq $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}(0, 1)$, donc la région de rejet est : $] -\infty; -q_{1-\frac{\alpha}{2}}[\cup] q_{1-\frac{\alpha}{2}}; +\infty[$.

b) Pour une hypothèse unilatérale à droite $\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 > p_2 \end{cases}$

on rejette H_0 au risque α si : $L^* \geq q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{N}(0, 1)$, donc la région de rejet est : $] q_{1-\alpha}; +\infty[$.

c) Pour une hypothèse unilatérale à gauche $\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 < p_2 \end{cases}$

on rejette H_0 au risque α si : $L^* \leq q_{1-\alpha}$ tq $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{N}(0, 1)$, donc la région de rejet est : $] -\infty; -q_{1-\alpha}[$.

Remarque 2.6.1 Pour un risque d'erreur α fixé on a :

$$\mathbb{P}(|L^*| \leq q_{1-\frac{\alpha}{2}}) = 1 - \alpha \text{ (hypothèse bilatérale).}$$

$$\mathbb{P}(L^* \leq q_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à droite).}$$

$$\mathbb{P}(L^* \geq q_{1-\alpha}) = 1 - \alpha \text{ (hypothèse unilatérale à gauche).}$$

Test de comparaison de deux variances et deux proportions résumé dans le tableau suivant :

	Test	Statistique de test	Région de rejet
deux variances	bilatéral	$V^* = \frac{\tilde{S}_{n_1}^2 n_1}{(n_1-1)\sigma_1^2} \times \frac{(n_2-1)\sigma_2^2}{\tilde{S}_{n_2}^2 n_2}$	$]0; c_{\frac{\alpha}{2}}[\cup]c_{1-\frac{\alpha}{2}}; +\infty[$
	unilatéral à droite		$]c_{1-\alpha}; +\infty[$
	unilatéral à gauche		$[0; c_{\alpha}[$
deux proportions	bilatéral	$L^* = \frac{F_1 - F_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	$] -\infty; -q_{1-\frac{\alpha}{2}}[\cup]q_{1-\frac{\alpha}{2}}; +\infty[$
	unilatéral à droite		$]q_{1-\alpha}; +\infty[$
	unilatéral à gauche		$] -\infty; -q_{1-\alpha}[$

TAB. 2.11 – Table de test de comparaison deux variances et deux proportions.

Chapitre 3

Application sous R

On termine ce mémoire par une application sur quelques tests qui ont déjà été étudiés dans le deuxième chapitre sous logiciel **R**.

3.1 Test de la moyenne

► **Instruction R** : on utilise la fonction `t.test()`

► Application de l'exemple 2.1.3 (cas unilatéral à gauche) :

Code R :

```
x=c(45, 40, 43, 35, 37, 20, 25, 15, 17, 45, 19, 40, 43, 18, 47, 27, 23, 38, 20, 37)
```

```
t.test(x, alternative = "less", mu = 45, conf.level = 0.05)
```

► One Sample t-test

data : x

t = -5.3503, df = 19, p-value = 1.829e - 05

alternative hypothesis : true mean is less than 45

5 percent confidence interval :

-Inf 27.40163

sample estimates :

mean of x

31.7

► p- value est inférieure à 0.05, donc on rejette l'hypothèse H_0 .

► Si le test est bilatéral ou unilatéral à droite l'application être comme suit : (respectivement)

```
t.test (x, alternative = "two.sided", mu = 45, conf.level = 0.05)
```

```
t.test (x, alternative = "greater", mu = 45, conf.level = 0.05)
```

3.2 Test de la fréquence (proportion)

► **Instruction R** : on utilise la fonction `prop.test()`

► Application de l'exemple 2.3.1 (cas unilatéral à droite) :

Code R :

```
► prop.test (45, 500, p = 0.15, alternative = "greater", conf.level = 0.05)
```

► 1-sample proportions test with continuity correction

data : 45 out of 500, null probability 0.15

X-squared = 13.651, df = 1, p-value = 0.9999

alternative hypothesis : true p is greater than 0.15

5 percent confidence interval :

0.1122181 1.0000000

sample estimates :

p

0.09

► p- value est supérieur à 0.05, donc on accepte l'hypothèse H_0 .

► Si le test est bilatéral où unilatéral à gauche l'application être comme suit : (respectivement)

```
prop.test (45, 500, p = 0.15, alternative = "two.sided", conf.level = 0.05)
```

```
prop.test (45, 500, p = 0.15, alternative = "less", conf.level = 0.05)
```

3.3 Test de comparaison de deux moyennes

► **Instruction R** : on utilise la fonction : `t.test()`

► Application de l'exemple 2.4.1 :

Code R :

```
S1=c(31.8, 31.4, 32.1, 32.8, 31.6, 32.5, 32.6, 31.7)
```

```
S2=c(31.2, 31.7, 32.9, 32.2, 31.5, 32.3, 31.3, 32.4)
```

```
t.test (S1, S2, var.equal = T)
```

► Two Sample t-test

data : S1 and S2

t = 0.44506, df = 14, p-value = 0.6631

alternative hypothesis : true difference in means is not equal to 0

95 percent confidence interval :

-0.4773909 0.7273909

sample estimates :

mean of x mean of y

32.0625 31.9375

► p- value est supérieur à 0.2, donc on accepte l'hypothèse H_0 .

3.4 Test de comparaison de deux variances

► **Instruction R** : on utilise la fonction `var.test()`

► Application de l'exemple 2.5.1 (cas bilatéral) :

Code R :

```
grammes1 =c(107, 106.1, 107.3, 106.2, 106.7, 107.3, 107.7, 106.7, 107.2)
```

```
grammes2 =c(106.2, 107.1, 107.4, 107.0, 106.4, 103.3, 107.6)
```

```
var.test (grammes1, grammes2, alternative = "two.sided", conf.level = 0.05)
```

► F test to compare two variances

data : grammes1 and grammes2

F = 0.13156, num df = 8, denom df = 6, p-value = 0.01155

alternative hypothesis : true ratio of variances is not equal to 1

5 percent confidence interval :

0.1215574 0.1342654

sample estimates :

ratio of variances

0.1315625

► p- value est inférieure à 0.05, donc on rejette l'hypothèse H_0 .

► Si le test est unilatéral à gauche ou à droite l'application être comme suit : (respectivement)

```
var.test ( grammes1, grammes2, alternative = "less", conf.level = 0.05)
```

```
var.test ( grammes1, grammes2, alternative = "greater", conf.level = 0.05)
```


3.5 Test de comparaison de deux proportions

► **Instruction R** : on utilise la fonction `prop.test`

Exemple 3.5.1 (*cas unilatéral à gauche*) : Un producteur de desserts lactés au caramel se trouve en concurrence avec d'autres marques. Au début de l'année 2010, il décide d'investir dans une nouvelle présentation de ses desserts. Avant d'avoir le bilan de l'année, il fait une rapide enquête auprès d'un certain nombre de magasins.

- Avant la nouvelle présentation, sur 200 desserts vendus, 44 étaient ceux du producteur.

- Après la nouvelle présentation, sur 290 desserts vendus, 88 étaient ceux du producteur.

Est-ce que le producteur peut affirmer que la nouvelle présentation a augmenté sa part de marché sur les desserts lactés au caramel ? (au niveau 0.05)

Code R :

► `prop.test(x = c(44, 88), n = c(200, 290), alternative = "less")`

► 2-sample test for equality of proportions with continuity correction

data : c(44, 88) out of c(200, 290)

X-squared = 3.7747, df = 1, p-value = 0.02602

alternative hypothesis : less

95 percent confidence interval :

-1.00000000 -0.01370086

sample estimates :

prop 1 prop 2

0.2200000 0.3034483

► p- value est inférieure à 0.05, donc on rejette l'hypothèse H_0 .

► Si le test est bilatéral où unilatéral à droite l'application être comme suit : (respectivement)

`prop.test (x = c(44, 88), n = c(200, 290), alternative = "two.sided")`

`prop.test (x = c(44, 88), n = c(200, 290), alternative = "greater")`

Tout est résumé dans le tableau suivant :

Test	L'hypothèse	commandes R
Moyenne	$\mu \neq \mu_0$	<code>t.test (x,alternative = "two.sided",mu = ,conf.level =)</code>
	$\mu < \mu_0$	<code>t.test (x,alternative = "less",mu = ,conf.level =)</code>
	$\mu > \mu_0$	<code>t.test (x,alternative = "greater",mu = ,conf.level =)</code>
Proportion	$p \neq p_0$	<code>prop.test (, ,p= ,alternative = "two.sided",conf.level =)</code>
	$p < p_0$	<code>prop.test (, ,p= ,alternative = "less",conf.level =)</code>
	$p > p_0$	<code>prop.test (, ,p= ,alternative = "greater",conf.level =)</code>
Comparison 2 moyennes	$\mu_1 \neq \mu_2$	<code>t.test (, ,var.equal = T)</code>
	$\mu_1 < \mu_2$	
	$\mu_1 > \mu_2$	
Comparison 2 variances	$\sigma_1^2 \neq \sigma_2^2$	<code>var.test (a,b,alternative = "two.sided",conf.level =)</code>
	$\sigma_1^2 < \sigma_2^2$	<code>var.test (a,b,alternative = "less",conf.level =)</code>
	$\sigma_1^2 > \sigma_2^2$	<code>var.test (a,b,alternative = "greater",conf.level =)</code>
Comparison 2 proportions	$p_1 \neq p_2$	<code>prop.test (x = c(,), n = c(,), alternative = "two.sided")</code>
	$p_1 < p_2$	<code>prop.test (x = c(,), n = c(,), alternative = "less")</code>
	$p_1 > p_2$	<code>prop.test (x = c(,), n = c(,), alternative = "greater")</code>

TAB. 3.1 – Table des commandes R des tests.

Conclusion

En conclusion, les tests statistiques jouent un rôle très important dans tous les domaines, donc ils sont très nécessaires.

Dans ce mémoire on a présenté quelques types de tests paramétriques usuels, et pour chaque test on a défini sa statistique ainsi que sa loi de probabilité, et on a aussi défini pour un seuil donné les régions critiques appropriées, à la fin on a utilisé le langage R, les différents packages et les commandes qui correspondent à chaque test.

Bibliographie

- [1] Arouche Faiza. 2016. Tests séquentiels. Mémoire. Université de Béjaia.
- [2] Achour Chams. 2021. Tests de normalité et applications. Mémoire de master. UMKB.
- [3] Benhissem Amina. 2013. Théorie des tests statistiques. Mémoire de master. Université Mohamed Khider Biskra.
- [4] Boubeche Nouredine. 2020. Tests d'hypothèses statistique. Mémoire de master. Université Mohamed Khider Biskra.
- [5] Bouden Ahlem. septembre 2020. Tests de comparaison de deux distributions. Mémoire de master. UMKB.
- [6] Chesneau, C. (2016). Introduction aux tests statistiques avec R.
- [7] Dib Khaled. Mémoire tests d'hypothèses dans un modèle de censure. Université de Mentouri constantine.
- [8] Etienne Birmele. Tests statistiques. M1 IMSV.
- [9] Houde Lois. Module 10 tests d'hypothèses. université du Québec à trois-Rivières
- [10] <https://dimension.usherbrooke.ca/pages/78>.
- [11] <https://www.universalis.fr/encyclopedie/tests-d-hypotheses-statistiques>.
- [12] <https://fr.m.wikipedia.org/wiki/statistique>.
- [13] Jean-Jacques Ruch. (2012.2013). Statistique : tests d'hypothèses.
- [14] Merghni Djamel, Necir Abdelhakim. Cours de tests statistique master1. UMKB.

- [15] Mountassir M'hammed. Probabilités et Statistique.
- [16] Monbet, V. (2009). Tests statistiques Notes de cours. Cité en, 133.
- [17] Lejeune, M. (2004). Statistique : La théorie et ses applications. Springer Science & Business Media.
- [18] Labarere José. 2011/2012. Chapitre 5 tests paramétriques de comparaison de 2 moyennes exercice commentés. Université Joseph Fourier de Grenoble.
- [19] Medouer Nawel. Rappel de cour (tests d'hypothèses). Université de Batna 2.
- [20] Menacer Chihab eddine. 2021. Mémoire de master. Estimation non paramétrique de la densité de probabilité. UMKB.
- [21] Saous cheikh. 2018. université Ahmed Draia Adrar : <https://www.researchgate.net/publication/341763603>.
- [22] Vitam. Tests statistiques. centre intégré universitaire de santé et de services sociaux et la capitale nationale.
- [23] Vincent Jalby. Octobre 2020. Une tres courte introduction au statistique avec R. Université de Limoges.
- [24] Vincent Rivoirard. Statistique mathématique. 2021/2022. université paris Dauphine-psl, licences3-mathématique appliquées département MLDO.

Annexe A : Logiciel R

► Le langage **R** est un langage de programmation et un environnement mathématique utilisés pour le traitement de données. Il permet de faire des analyses statistiques aussi bien simples que complexes comme des modèles linéaires ou non linéaires, des tests d'hypothèse, de la modélisation de séries chronologiques, de la classification, etc. Il dispose également de nombreuses fonctions graphiques très utiles et de qualité professionnelle.

► **R** a été créé par **Ross Ihaka** et **Robert Gentleman** en 1993 à l'Université d'Auckland, Nouvelle-Zélande, et est maintenant développé par la **R** Développement Corre Team. L'origine du nom du langage provient, d'une part, des initiales des prénoms des deux auteurs (Ross Ihaka et Robert Gentleman) et, d'autre part, d'un jeu de mots sur le nom du langage S auquel il est apparenté.

Annexe B : Notations et symbols

X	: Variable aléatoire.
X_1, X_2, \dots, X_n	: Échantillon.
n	: Taille de l'échantillon.
\bar{X}	: La moyenne de l'échantillon.
S^2	: La variance de l'échantillon.
\tilde{S}^2	: La variance empirique.
M_r	: Le moment empirique d'ordre r .
M'_r	: Le moment centré empirique d'ordre r .
E	: Espérance.
V	: Variance.
$E(\bar{X})$: Espérance de la moyenne.
$Var(\bar{X})$: Variance de la moyenne de l'échantillon.
$E(\tilde{S}^2)$: Espérance de la La variance empirique.
$E(M_r)$: Espérance de Le moment empirique d'ordre r .
H_0	: L'hypothèse nulle.
H_1	: L'hypothèse alternative.
α	: Risque de première espèce.
β	: Risque de deuxième espèce.
$1 - \beta$: La puissance du test.

W	: Région de rejet.
\bar{W}	: Région d'acceptation.
θ	: Paramètre inconnu.
Θ_0	: Ensemble des valeurs de θ .
π	: Fonction puissance.
μ	: Moyenne.
σ^2	: Variance.
$\mathcal{N}(\mu, \sigma^2)$: Loi normale.
$\mathcal{N}(0, 1)$: Loi normale centrée réduite.
$\Phi(\cdot)$: La fonction de répartition de loi normale centrée réduite.
$\mathcal{T}(n-1)$: Loi de Student à $(n-1)$ degrés de liberté.
χ_n^2	: Loi de khi-deux à n degrés de libertés.
χ_{n-1}^2	: Loi de khi-deux à $(n-1)$ degrés de libertés.
p	: Proportion d'individus.
Y_n	: Un échantillon aléatoire de taille n .
$\mathcal{B}(n, p)$: Loi Binomiale.
F	: La fréquence empirique de p .
c-à-d	: C'est t-à-dire.
$v.a$: Variable aléatoire.
tq	: Tel que.
$\mathcal{T}(n_1 + n_2 - 2)$: Loi de Student à $(n_1 + n_2 - 2)$ degrés de liberté.
$\mathcal{F}(n_1 - 1, n_2 - 1)$: Loi de fisher de $(n_1 - 1, n_2 - 1)$ degrés de liberté.
$\mathcal{B}(p)$: Loi de Bernoulli.
\mathbb{P}	: La probabilité.
\mathbb{R}	: Ensemble des valeurs réelles.
q, t, k, l, c	: Les quantiles.
\rightsquigarrow	: Suit approximativement.

Résumé

Un test statistique est une procédure mathématique qui se déroule entre deux hypothèses statistiques avec la sélection d'un échantillon, dans laquelle l'hypothèse statistique est rejetée ou non, puis les résultats obtenus sont extrapolés. L'objectif de ce mémoire est d'étudier quelques tests statistiques en raison de leur rôle important dans tous les domaines.

Mots clé : la statistique, test statistique, hypothèse, échantillon, paramétrique, risque, erreur, niveau de signification, région de rejet, statistique de test, accepter, rejeter.

المخلص

الاختبار الإحصائي هو إجراء رياضي يتم بين فرضيتين إحصائيتين مع اختيار عينة، يتم فيه رفض أو عدم رفض الفرضية الإحصائية، ثم يتم استقراء النتائج المتحصل عليها. الهدف من هذه المذكرة هو دراسة بعض الاختبارات الإحصائية نظرا لدورها المهم في جميع المجالات

الكلمات المفتاحية: الإحصاء، الاختبار الإحصائي، الفرضية، العينة، البارامترية، المخاطر، الخطأ، مستوى الأهمية، منطقة الرفض، اختبار الإحصاء، القبول، الرفض

Abstract

A statistical test is a mathematical procedure that takes place between two statistical hypotheses with the selection of a sample, in which the statistical hypothesis is rejected or not, and then the results obtained are extrapolated. The objective of this memory is to study some statistical tests because of their important role in all fields.

Keywords : statistics, statistical test, hypothesis, sample, parametric, risk, error, level of significance, region of rejection, test statistic, accept, reject.