

الجمهورية الجزائرية الديمقراطية الشعبية  
République Algérienne Démocratique et Populaire  
وزارة التعليم العالي والبحث العلمي  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



N° Réf :.....

Centre Universitaire  
Abd Elhafid Boussouf Mila

Institut des Sciences et Technologie

Département de Mathématiques et Informatique

## Mémoire préparé en vue de l'obtention du diplôme de Master

EN : Mathématiques

Spécialité : Mathématiques Appliquées

# L'analyse statistique dans les méthodes séquentielles

Préparé par : MALIM KHADIDJA

Devant le jury

Haj Ali Thuraya (MAA)

C.U.Abd Elhafid Boussouf

Président

Zerari Amel (MAA)

C.U.Abd Elhafid Boussouf

Rapporteur

Ahmed Ghezal (MCB)

C.U.Abd Elhafid Boussouf

Examineur

Année Universitaire : 2017/2018

# **REMERCIEMENTS**

*Je remercie d'abord et avant tout le bon Dieu qui m'a donné le courage et la patience pour réaliser ce travail.*

*Je tiens à remercier toutes les personnes qui ont participé à différents degrés pour réaliser ce modeste travail.*

*Je tiens tout d'abord à exprimer notre profonde gratitude à madame ZERARI AMEL qui a dirigé ce travail et avoir su nous apporter l'aide. Nécessaire à sa réalisation, pour son aide, sa patience et ses conseils tout au long de cette recherche.*

*Sans oublier aussi de remercier infiniment tous les enseignants du département de mathématique, sur leurs aides, efforts et conseils qui ont contribué à notre réussite.*

*Sans oublier bien-sûre remercier le Dr "Oudjih" professeur en épidémiologie et les enseignants "Frah Mouhamed" de Université de Batna et "M.Tallai" du centre Universitaire de Mila et tous les enseignants qui nous ont formé durant les années d'études.*

## DEDICACE

*Je dédie ce mémoire à mon père pour le soutien qu'il m'a toujours donné et  
qui m'aider beaucoup toute durée de mes études.*

*Et ma mère qui m'a encouragé et m'a donné tout son amour.*

*Je dédie à mes sœurs et mon frère et aussi ma grand-mère et toute ma familles  
et tous mes amis et les coulages sur tout*

*KENZA, MOUNIRA, YOUSRA, ASMA ,AMENI et AMEL*

# Table des matières

<b>Introduction Générale</b>	<b>3</b>
<b>1 Les essais cliniques</b>	<b>5</b>
1.1 Les définitions des essais cliniques . . . . .	6
1.1.1 Étude clinique, essai clinique et éthique médicale . . . . .	7
1.1.2 L'objectif des essais cliniques . . . . .	7
1.1.3 Effet placebo . . . . .	8
1.2 Les conceptions et les phases et les types des essais cliniques . . . . .	8
1.2.1 Conception des essais cliniques . . . . .	8
1.2.2 Les phases des essais cliniques . . . . .	9
1.2.3 Les types des essais cliniques . . . . .	10
1.3 Les avantages et les risques des essais cliniques . . . . .	10
1.3.1 Les avantages . . . . .	11
1.3.2 Les risques . . . . .	11
1.3.3 Les inconvénients des essais cliniques . . . . .	11
<b>2 L'analyse séquentielle</b>	<b>13</b>
2.1 Rappel mathématique . . . . .	14
2.1.1 La fonction de répartition et la fonction de densité . . . . .	14
2.1.2 Loi normale (loi de Laplace-Gauss) . . . . .	15
2.2 Le test d'hypothèses classiques . . . . .	17
2.2.1 Définitions générales . . . . .	17
2.2.2 Le test unilatéral et le test bilatéral . . . . .	18
2.2.3 La nature générale d'une procédure de test . . . . .	19
2.3 Le Test séquentiel . . . . .	25
2.3.1 Notion d'un test séquentiel . . . . .	25
2.3.2 Contrôle par échantillonnage . . . . .	27
2.3.3 Le test séquentiel du rapport de probabilité . . . . .	32
2.4 Les essais séquentiels par groupe et arrêt précoce . . . . .	36

<b>3 Les tests séquentiels dans les essais cliniques</b>	<b>37</b>
3.1 Les tests séquentiels de groupe . . . . .	38
3.1.1 Le test bilatéral . . . . .	38
3.1.2 Test d'échantillon fixe . . . . .	39
3.2 Test de Pocock . . . . .	40
3.2.1 Définition de test de Pocock . . . . .	41
3.2.2 L'exigence de puissance . . . . .	42
3.3 Test d'O'Brien et Fleming . . . . .	43
3.3.1 Définition de test d'O'Brien et Fleming . . . . .	44
3.3.2 L'exigence de puissance . . . . .	44
3.4 Comparaison des tests séquentiels . . . . .	46
3.4.1 Comparaison de test de Pocock et de test d'O'Brien et Fleming . . . . .	46
3.4.2 La taille maximale de l'échantillon . . . . .	46
3.4.3 Le nombre des groupes . . . . .	47
3.4.4 Discussion . . . . .	47
 <b>Annexe : Les tableaux statistiques</b>	 <b>49</b>
 <b>Conclusion Générale</b>	 <b>57</b>
 <b>Bibliographie</b>	 <b>59</b>

# Introduction Générale

La statistique est une branche mathématique d'un ensemble des techniques et des méthodes qui permet d'étudier et traduire l'information des données expérimentales, dans cette branche, l'analyse séquentielle est une technique où la taille d'échantillon n'est pas fixée à l'avance, mais elle nécessite l'évaluation risque. C'est une vieille idée qui est découverte par les scientifiques britanniques est eux qui réclament de l'avoir formulé en premier.

Historiquement, en 1939 et 1942, le groupe de recherche statistique fondé à l'université de Columbia, dirigé par W.A.Wallis a collaboré avec le département de la Défense des États-Unis est composé par des chercheurs suivants : H.A.Freeman, M.A Girshik, H.Hotelling et J.Wolfowitz.

Au début de 1943 M.Friedman et WA-Wallis ont énoncé le problème de réaliser un test d'hypothèse dans des séquences, il conçoit le test séquentiel du rapport de probabilité et rédige une théorie fondamentale. En 1944, M.Friedman a formulée la courbe d'efficacité du test de rapport de probabilité. En peu de temps, d'une manière similaire C.Stockman avait présenté une courbe d'efficacité et en même temps, AW a donné une fonction d'efficacité général pour le test de rapport de probabilité, quelques mois plus tard, il développa la théorie générale des sommes cumulatives, qui fournit non seulement la courbe d'efficacité de tous les tests de rapport de probabilité, en septembre 1944 il a publié les résultats de la fonction d'efficacité, dans « Sommes Cumulatives De Variable Aléatoire » et en juin 1945 dans les « Analyse de Mathématique Statistique ».

En 1975, Armitage a appliqué l'analyse séquentielle dans les essais cliniques, cette dernière est une étude sur des êtres humain utiliser pour expliquer l'efficacité de traitement expérimental est analyse les données accumulent avec des résultats après chaque observation et pour chaque participants sont remportés, et la décision est soit continue l'étude en nouvelle participation, et soit arrêter l'étude avec le résultat qu'il y a ou non un différent statistiquement entre les traitements.

Aux États-Unis, l'administration des aliments et des médicaments supervise la réglementation et l'approbation des médicaments, des produits biologiques et des dispositifs médicaux. Ses processus d'examen et d'approbation influent sur la conception et la conduite de la plupart des nouveaux essais cliniques. Le test préclinique d'une intervention expérimentale est effectué

avant que les investigateurs commencent un essai clinique. Ces études sont réalisées en laboratoire et dans des études avec des animaux pour fournir des preuves préliminaires que l'intervention expérimentale sera sûre et efficace pour les humains. L'administration des aliments et des médicaments exige des tests précliniques avant de pouvoir commencer les essais cliniques. Les informations de sécurité provenant des tests précliniques sont utilisées pour appuyer une demande à l'administration des aliments et des médicaments de commencer à tester l'intervention expérimentale dans des études avec des humains.[2]

Ce mémoire présente des applications mathématique, en particulier dans les essais cliniques, elle contient trois chapitres :

Dans le premier chapitre, nous donnons le domaine des essais cliniques comme les définitions des essais cliniques, les conceptions et les phases et les types, et nous expliquons les avantages et les risques des essais cliniques.

Le deuxième chapitre, nous donnons deux types d'hypothèses, les hypothèses classiques et les hypothèses séquentiels. Nous utilisons la loi normale dans les hypothèses classiques qui représentons le début de l'étude de règle décision qui décider si l'hypothèse a tester soit accepter ou rejeter, ils contiennent deux hypothèses l'un nul  $H_0$  n'existe pas une différence entre les variables comparés et l'autre alternative  $H_1$  c'est la négation de l'hypothèse nul, avec deux types des erreurs,  $\alpha$  1<sup>re</sup> espèce c'est la taille de région critique et  $\beta$  c'est la deuxième espèce et la puissance  $1 - \beta$ . Et les hypothèses séquentiels c'est le choix d'un des trois hypothèses accepté  $H_0$  ou accepté  $H_1$  ou en continuant l'expérience en faisant une observation supplémentaire, en utiliser le test séquentiel de rapport de probabilité pour choisir entre deux hypothèse simple dans l'analyse séquentielle, et nous expliquons la règle d'arrêtée précoce dans les essais séquentiels par groupe.

Le troisième chapitre, nous nous intéressons au test de Pocock et le test d'O'Brien et Fleming sont des tests bilatéral comme deux applications d'analyse séquentielle dans les essais cliniques, nous expliquons les tests séquentiels de groupe pour comparer les résultats de deux traitements et choisir le nombre de groupe convenable.

# Chapitre 1

## Les essais cliniques

## **Introduction**

Les essais thérapeutique est une étude scientifique des testes sur l'être humain pour évaluer l'efficacité de traitement et les effets secondaires ou la pharmacologie.

Les essais cliniques est un outil pour déterminer les possibilités et les efficacité de traitement préventive ou curative pour certaine maladie donné.

Tous les recherche qui affecte pour une étude prospective des sujets humains à des groupes d'intervention et de comparaison de la relation d'effet entre un acte médical et l'évolution d'un état de santé étudier.

Dans ce chapitre on expliquer quelque notion est présenter dans les définitions des essais cliniques, les phases des essais cliniques, les types des essais cliniques, et les avantages et les risques des essais cliniques

### **1.1 Les définitions des essais cliniques**

**Définition 1 (Cha. H. Evans. 2001)** *On a quelques chercheurs définissent les essais cliniques comme suit :*

1. **Friedman, Furberg et DeMets**, en 1996

*Un essai clinique est une étude prospective comparant l'effet et la valeur de l'intervention contre le contrôle chez les êtres humains.*

2. **Meinert**, en 1996

*Une expérience contrôlée ayant un événement clinique comme mesure de résultat et réalisée dans une clinique ou un milieu clinique et impliquant des personnes ayant une maladie ou un état de santé spécifique.*

3. **Piantadosi**, en 1997

*Une expérience est une série d'observations faites dans des conditions contrôlées par le scientifique. Un essai clinique est une expérience testant des traitements médicaux sur des participants humains.*

4. **Pocock**, en 1984

*Le terme « essais cliniques » peut s'appliquer à toute forme d'expérience planifiée impliquant des patients et visant à élucider le traitement le plus approprié pour de futur patients présentant un trouble médical donné.*

### **1.1.1 Étude clinique, essai clinique et éthique médicale**

#### **1.1.1.1 Étude clinique**

Est un domaine scientifique dont l'évaluation technique des résultats de traitement préventive d'une pathologie et en utilisant des données existantes à l'absence des sujets.

#### **1.1.1.2 Essai clinique**

L'essai clinique c'est une expérience conçue pour évaluer l'innocuité ou l'efficacité d'un traitement. Il est se réfèrent à des expériences planifiées qui impliquent des participants humains et qui sont conçues pour élucider les traitements les plus appropriés pour de futur patients atteints d'un trouble médical donné.

L'essai clinique utiliser les résultats d'un échantillon restreint des patients pour faire des inférences sur la façon dont le traitement devrait être administré à la population générale des patients qui nécessiteront un traitement futur.

#### **1.1.1.3 Éthique médicale**

C'est un mot grec «Ethos» qui explique une méthode d'être et de se comporter à la morale qui répond à plusieurs règles de bonne conduite c-à-d action humaine, et aussi elle est très importante à la médecine. Sans l'éthique la médecine perde son sens.

### **1.1.2 L'objectif des essais cliniques**

Les objectifs principaux des essais cliniques sont :  
Sécurité d'utilisation des essais cliniques, à visée pharmacocinétique, recherche d'efficacité, essais de pharmacologie clinique, évaluation des méthodes nouvelles, pharmacovigilance, études épidémiologique, détermination du mécanisme d'action, évaluation dans la population générale, études exploratoires.

#### **1.1.2.1 L'importance des essais cliniques**

Les essais cliniques sont des étapes nécessaires dans le domaine d'un traitement thérapeutique. Ils permettent de déterminer le traitement le plus efficace et plus protégé pour les patients d'après l'utilisation de résultat d'un échantillon des patients comme des réponses à des questions spécifiques et importantes se posent dans la recherche sur le traitement, ou comme un nouveau médicament comparer avec traitement existant.

Les paragraphes suivante sont relever de livre de **B.D. Bouhanick (2011-2012)**.

### 1.1.3 Effet placebo

#### 1.1.3.1 La définition de placebo

**Placebo** Est un mot latin «*placere*», (plaire), pour expliquer (je plairai), c'est un traitement simple sous forme des comprimés, ou liquides, ou injections, utilisé en pharmacologie comme une méthode pour contrôler l'activité d'un médicament, il présente un mécanisme de connaissance thérapeutique dans les essais cliniques et induit des effets psychologique chez l'humain. Il y a deux types de placebo :

**Placebo pur** : c'est une substance inerte, et évaluer les nouveaux médicaments, et en pratique c'est exceptionnel et peu recommandable.

**Placebo impur** : c'est un traitement contient un ou plusieurs principes actifs dont l'efficacité est variable. Utilisation insuffisante en médecine.

#### 1.1.3.2 L'effet placebo

L'effet placebo c'est l'effet thérapeutique déduit par l'administration d'une substance, comprimés, liquides, injections, ou c'est une réaction d'un médecin face au patient dans une position d'acceptation et sur la méthode de la guérison, il présente la réponse qui remplace à l'administration d'un placebo. L'interface d'effet placebo dans différentes sciences comme neurobiologie, psychologie, sociologie.

L'effet placebo dépend du patient, comportement des médecins, maladie, type de produit ordonné.

#### 1.1.3.3 L'utilisation du placebo dans les essais cliniques

En générale, les études d'un médicament différentes sur la comparaison de deux groupes de patient : le premier groupe traité par le médicament, et le deuxième groupe ne prend pas. Mais il est observé se voit administrer un placebo dans tous les points pareille au produit à tester comme présentation, aspect, mode d'administration. Il est impossible éthiquement comparer l'efficacité d'un traitement par placebo sachant que le traitement de référence existe déjà.

Les paragraphes suivants sont relevés de livre de “ *Cha. H. Evans. (2001).*”

## 1.2 Les conceptions et les phases et les types des essais cliniques

### 1.2.1 Conception des essais cliniques

La conception d'essai clinique nécessite 3 considérations suivantes :

1. L'étude basé sur des questions de recherche biomédicale importantes.
2. Baser sur une méthodologie stricte a une question de recherche spécifique posée.
3. Elle doit baser sur un groupe des règles éthiques, dont moins risques pour les participants à l'étude.

Le choix d'un plan d'étude approprié dépend d'un certain nombre de considérations, comme suivant :

1. Le plan d'étude capable à répondre à la question principale de recherche.
2. Si l'essai étudie un nouveau traitement potentiel pour une condition pour laquelle un traitement efficace existe déjà.
3. Si la maladie pour laquelle un nouveau traitement est recherché est grave.
4. La probabilité et l'ampleur du risque et des avantages probables pour les participants.
5. La population à étudier sa taille, sa disponibilité et son accessibilité.
6. comment utilisées les données seront (par exemple pour initier un traitement ou comme données préliminaires pour un essai plus important).

### 1.2.2 Les phases des essais cliniques

Il existe 4 phases classées en essais cliniques comme suit :

#### 1. Phase I : Sécurité et profils pharmacologiques

C'est l'introduction initiale d'un vaccin candidat ou d'un médicament dans une population humaine pour déterminer sa sécurité et son mode d'action ou de la pharmacogénétique. Cette phase peut inclure des études sur la toxicité et la dose maximale sûre d'un nouveau médicament. Et généralement testent moins de 100 volontaires sains.

#### 2. Phase II : Etudes d'efficacité clinique d'un médicament

Cet essai initial vise à examiner l'efficacité chez environ 200 à 500 volontaires. L'objectif de cette phase c'est l'immunogénicité tandis que avec les médicaments, l'accent est mis sur la démonstration de l'innocuité et de l'efficacité par rapport à celles des autres schémas thérapeutiques existants. Souvent les sujets sont répartis au hasard pour étudier et contrôler les groupes.

#### 3. Phase III : Essai clinique approfondi

Cette phase vise à compléter l'évaluation de la sécurité et de l'efficacité. Il implique un grand nombre, peut-être des milliers des volontaires d'un centre ou de plusieurs centres, généralement avec une allocation aléatoire aux groupes d'étude et de contrôle.

#### 4. Phase IV : La commercialisation d'un médicament

Cette phase est menée après que l'autorité nationale d'enregistrement des médicaments

a approuvé le médicament pour la distribution ou la commercialisation. L'essai est conçu pour déterminer un effet pharmacologique spécifique ou les effets d'une utilisation à long terme ou pour établir l'incidence d'effets indésirables.

### 1.2.3 Les types des essais cliniques

#### 1. Essais randomisé ou essais contrôlé :

Ce sont de différent groupe dans un essai thérapeutique qui ne sont pas contrôler par médecin et qui sont composés de façon comparable et aussi ont la même répartition de patient de base comme âge, sexe, taille, poids, et la gravité de la maladie.

#### 2. Essais ouverte :

Un essai clinique ouverte est contre essai contrôlé, c'est un essai en aveugle puisque le médecin et le sujet ou le patient inclus sont au courant du traitement administré.

#### 3. Essais croisé :

C'est un plan expérimental d'un essai clinique comparatif, les sujets ont les mêmes périodes de traitement. L'essai croisé est randomisé quand l'ordre des périodes de traitement est tiré au sort pour chaque sujet.

#### 4. Essais en groupe parallèles :

C'est l'essai clinique où le traitement étudié comparé à un traitement contrôlé avec des patients, les différents groupes des patients sont suivis dans une durée déterminé, chaque groupe a un traitement ou un placebo. Les essais en groupes parallèles beaucoup utilisés dans les essais de phase II et de phase III. Un sous type d'essai en groupe parallèle est plus fréquent par la technologie de l'information ce sont des étude séquentielles de groupe.

#### 5. Essais séquentielle :

On dit que l'essai séquentiel un groupe des patients qui reçoit un première traitement puis un deuxième c-à-d patients reçoit un traitement expérimental tandis que l'autre de traitement de référence. Les observations s'évaluent lorsque se produit et dépend des résultats accumulés.

Les paragraphes suivante sont relever de livre de “ *Société Canadienne Du Cancer. (2007)*”.

## 1.3 Les avantages et les risques des essais cliniques

Tous essais cliniques sont liés des avantages et des risques possibles, ils est important dans la recherche de médecine.

### **1.3.1 Les avantages**

Les avantages représentent comme suivant :

1. Les essais cliniques représentent un espoir et pour participer au progrès des connaissances scientifiques.
2. Accessibilité à un médicament ou traitement nouveaux pour les meilleures conditions de sécurité avant sa commercialisation.
3. Le nouveau traitement étudié dans les essais cliniques se révèle efficace dans la prévention du maladie.
4. L'occasion de venir en aide à d'autres humains et l'amélioration d'un traitement du la maladie.
5. dans la comparaison entre le traitement expérimentale et le traitement standard (placebo) ne fasse pas partie du groupe qui reçoit le meilleur et le nouveau traitement standard disponible pour lutter le maladie.

### **1.3.2 Les risques**

Les risques représentés comme suivant :

- (a) Les nouveaux traitements ou médicaments dans les essais cliniques se sont pas toujours meilleurs et ni aussi efficaces que le traitement standard (placebo).
- (b) Apparition les effets secondaires c'est plus grave de médicament.
- (c) Les nouveaux traitements ne sont pas nécessairement efficace pour tout participants.
- (d) Le traitement standard ne donne pas d'avantage que le nouveau traitement.
- (e) Les essais cliniques demande plus de temps ou impose des contraintes.

### **1.3.3 Les inconvénients des essais cliniques**

Les effets secondaire d'un essai clinique en générale et traitement en particulière dépend de la mesure thérapeutique et de l'état de malade, ils sont les plus prévisible de nature passagère et disparaissent après la fin du traitement.

Les traitements peuvent ne pas être efficace ou efficace indésirable voire dangereux.

### **Conclusion**

Dans ce chapitre on impliquons sur les sujets humains modale concerne un essai clinique pour savoir l'efficacité d'un traitement soit par intervention ou d'un schéma thérapeutique des patients atteints d'un maladie spécifique, en utilisent la taille d'échantillon nécessaire

pour conserver la puissance statistique et on observons le taux d'erreur de 1<sup>re</sup> espèce. Et ceci dans le deuxième chapitre.

# Chapitre 2

## L'analyse séquentielle

## Introduction

En 1941, H. Hotelling dans «Détermination Expérimentale du Maximum de Fonction», donne une représentation par séquentielle du design sur une expérience à grande échelle dans des stades successifs et ce cas est formulé par P.C.Mahalanobis en 1940.

W.A.Wallis dans son livre « Analyse séquentielle » a montré le test séquentiel du rapport de probabilité. Il présente un test pour lequel la taille d'échantillon n'est pas déterminée à l'avance. Le principe de l'analyse séquentielle c'est analyser les données accumulées pour arrêter l'étude bien que les résultats soient convainquants statistiquement.

Dans ce chapitre, nous donnons, au premier point, quelques rappels mathématiques, les hypothèses classiques et le test séquentiel. Nous expliquons le test séquentiel du rapport de probabilité. Et en fin les essais séquentiels par groupe et arrêt précoce.

## 2.1 Rappel mathématique

**Définition 2** *Variable aléatoire* (D. Foata, A. Fuchs. 1998)

Un espace probabilisé  $(\Omega, \mathcal{A}, \mathbf{P})$  étant donné avec  $\mathcal{A}$  tribu, on appelle variable aléatoire réelle (resp. variable aléatoire à valeurs dans  $\mathbb{R}^n$  ou à  $n$  dimensions) une fonction mesurable de  $(\Omega, \mathcal{A})$  dans  $(\mathbb{R}, \mathcal{B}^1)$  resp. dans  $(\mathbb{R}^n, \mathcal{B}^n)$ .

**Définition 3** *Une statistique* (R. Veysseyre. 2001-2006)

Soit  $X = (x_1, \dots, x_n)$  un échantillon de taille  $n$  avec  $X$  variable aléatoire muni de la fonction de répartition  $F$  et la densité  $f$ . On dit que la fonction mesurable  $T$  des variables aléatoires  $x_i$ ,  $i = 1, \dots, n$  **une statistique** si :  $T(x_1, \dots, x_n)$  à un échantillon, on peut associer différentes statistiques.

**Exemple 2.1** La fonction moyenne arithmétique des  $n$  observations  $\bar{X}$  comme une statistique :

$$\bar{X} = T(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i$$

### 2.1.1 La fonction de répartition et la fonction de densité

**Définition 4** *Fonction de répartition*

Notés  $F(x)$  la probabilité que  $X$  prenne une valeur inférieure à une valeur donnée  $x$  défini par :

$$F(x) = \mathbf{P}(X \leq x), \quad \forall x \in \mathbb{R}^n$$

Qui vérifie les propriétés suivantes :

1.  $0 \leq F(x) \leq 1$ .

2.  $x < y \Rightarrow F(x) \leq F(y)$ .

3.  $P(\{X, x \leq X < y\}) = F(y) - F(x), \quad \forall x, y \in \mathbb{R}$  et  $X$  variable aléatoire.

**Définition 5** *La fonction de densité*  $f(x)$

La dérivée de la fonction de répartition  $F(x)$  s'appelle la fonction de densité de la variable  $x$  et noté par  $f(x)$  comme suivant :

$$f(x) = \frac{dF(x)}{dx}$$

Dans le cas continu, si  $X$  variable continue alors la fonction de répartition est

$$F(x) = \int_{-\infty}^x f(u)du$$

et  $f$  est densité de probabilité de  $X$  vérifie les propriétés suivantes :

1.  $\int_{-\infty}^{+\infty} f(x)dx = 1$ .

2.  $\forall x, f(x) \geq 0$ .

3.  $P(X \in ]x_1, x_2]) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(u)du$ .

4.  $P(X = x_0) = \int_{x_0}^{x_0} f(u)du = 0$ .

**2.1.2 Loi normale (loi de Laplace-Gauss)**

La loi normale ou bien la loi de Laplace-Gauss c'est la loi la plus connue, et elle représenter comme une loi terminée de la loi binomiale.

**Définition 6** (*R. Veysseyre. 2001-2006*)

On dit qu'une variable aléatoire réelle  $X$ , prenant ses valeurs dans  $\mathbb{R}$ , suit une loi de Laplace-Gauss ou loi normale, de paramètres  $m$  c'est la moyenne et  $\sigma$  c'est la variance, si sa densité de probabilité est donnée par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

La fonction  $f$  définit une densité. En effet :

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

En général, on utilisent la notation  $N(\text{moyenne}; \text{variance}) = N(m; \sigma^2)$ .

Et la fonction de répartition :

$$P(X < a) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{(x-m)^2}{2\sigma^2}} dx \tag{2.1}$$

**Remarque 2.1** (R. Veysseyre. 2001-2006)

L'intégrale (2.1) n'ayant pas d'expression mathématique simple, il représente un tableau (3.1) (voir l'annexe) donnent les valeurs de la fonction de répartition.

L'espérance et la variance sont représentés pour  $n$  échantillon comme suivante :  $E(\bar{X}) = m$  et  $Var(\bar{X}) = \frac{\sigma^2}{n}$

Les paragraphes suivante sont relever de livre de “**R. Veysseyre. (2001-2006)**”.

### 2.1.2.1 Variable aléatoire centrée réduite

Pour exprimer la fonction de répartition mathématique simple on utiliser la formule de la variable centrée réduite associée à la variable aléatoire  $X$  comme suit :

$$Z = \frac{X - m}{\sigma}$$

pour  $m = 0$  et  $\sigma^2 = 1$ , la loi normale  $N(0, 1)$  est appelée centrée réduite muni par les relations suivant : Densité de probabilité de la variable  $Z$  :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Fonction de répartition :

$$\mathbf{P}(X < a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{z^2}{2}} dz$$

Le tableau (3.1) (voir l'annexe) contient les valeurs de la fonction de répartition pour  $z \geq 0$ , et pour les valeurs  $z \leq 0$ , on a la formule

$$\mathbf{P}(Z < -z) = 1 - \mathbf{P}(Z < z)$$

Et le tableau (3.2) (voir l'annexe) est contient les fractiles de la loi normale réduite, puisque elle ne dépend d'aucun paramètre, on utilisons les formules de changement de variables qui admet de transformer d'une variable à l'autre :

$$Z = \frac{X - m}{\sigma} \quad \text{et} \quad X = \sigma Z + m$$

### 2.1.2.2 Méthode d'utilisation des tableaux de la loi normale

Le fractile d'ordre  $p$  est défini par :  $\mathbf{P}(Z < z_p)$

Pour  $\mathbf{P}(Z < z_p) \leq 0.50$ , on choisissons la colonne de gauche et la ligne supérieure, et les fractiles sont négatifs.

Pour  $\mathbf{P}(Z < z_p) \geq 0.50$ , on choisissons la colonne de droite et la ligne inférieure, et les fractiles sont positifs. Pour simplifier on expliquons cette méthode dans exemple :

**Exemple 2.2** Soit  $X$  une variable aléatoire muni de la loi normale  $N(3; 4)$ , avec la moyenne  $m = 3$  et d'écart-type  $\sigma = 4$ . On va calculer les probabilités :

$$\mathbf{P}(X < 4), \mathbf{P}(X < -2), \mathbf{P}(X > 1).$$

Et les valeurs de  $c_i$  dans les cas suivantes :  $\mathbf{P}(X < c_1) = 0,75$ ,  $\mathbf{P}(X > c_2) = 0,85$ .

D'après la variable centrée réduite  $Z$  associée à la variable  $X$  on a :

$$Z = \frac{X - m}{4} \text{ et } X = 4Z + 3$$

1.  $\mathbf{P}(X < 4) = \mathbf{P}(4Z + 3 < 4) = \mathbf{P}(Z < 0,25) = 0.5987$
2.  $\mathbf{P}(X < -2) = \mathbf{P}(4Z + 3 < -2) = \mathbf{P}(Z < -1.25) = 1 - \mathbf{P}(Z < 1) = 1 - 0.8943 = 0.1057$
3.  $\mathbf{P}(X > 1) = \mathbf{P}(4Z + 3 > 1) = \mathbf{P}(Z > -0.5) = 1 - \mathbf{P}(Z > 0.5) = 1 - 0.6914 = 1.6916$
4.  $\mathbf{P}(X < c_1) = \mathbf{P}(4Z + 3 < c_1) = \mathbf{P}\left(Z < \frac{c_1 - 3}{4}\right) = 0,75.$

Puisque  $\mathbf{P}\left(Z < \frac{c_1 - 3}{4}\right) = 0,75 \geq 0.50$  et d'après le tableau (3.2),

on a  $\frac{c_1 - 3}{4} = 0.6745$  Donc  $\mathbf{P}(Z < 0,6745) = 0,75$  et  $c_1 = 5.70$

5.  $\mathbf{P}(X > c_2) = \mathbf{P}(4Z + 3 > c_2) = \mathbf{P}\left(Z < \frac{c_2 - 3}{4}\right) = 0,85 \geq 0.50$

D'après le tableau (3.2) il y a deux valeurs pour :  $\frac{c_2 - 3}{4} = -1,0364$  sont

$$\mathbf{P}(Z < -1,0364) = 0,15 \quad \text{et} \quad \mathbf{P}(Z > -1,0364) = 0,85$$

Donc :  $c_2 = -1,0364 \times 4 + 3 = -1.146$ .

## 2.2 Le test d'hypothèses classiques

Les tests d'hypothèse sont des outils qui permet de prendre la décision d'accepter ou rejeter des variables ou des paramètres par un testeur d'après une règle de décision qui admet le choix de variable de décision et de la région critique pour la méthode de Neyman-Pearson. Et il est composer de deux types différents, les tests unilatérales et les tests bilatérales.

### 2.2.1 Définitions générales

#### Définition 7 1. Un test

Est un outil qui admet de choisir entre deux hypothèses, ces hypothèses présenteront dans des paramètres (valeur d'un moment, égalité de variances, nature d'une loi de probabilité) ou trouvé des résultats d'un échantillon.

#### 2. Hypothèse classique

Est une confirmation relatif à des caractéristiques d'une population telle que la valeur d'un ou de plusieurs paramètres.

3. **Test statistique (test d'hypothèse)**

C'est le premier cas à étudier la règle de décision qui admet à choisir entre deux hypothèses, l'hypothèse nul  $H_0$  et l'hypothèse alternative  $H_1$ . Si elle détermine uniquement des valeurs pour les paramètres  $\theta_i, i = 1, \dots, k, k \in \mathbb{R}$ .

4. **L'hypothèse nul  $H_0$**

Est une hypothèse statistique telle que n'existe pas de différence entre une variable et autre variable, ou un ensemble de variable ou que deux ou plusieurs distributions de population comparés. Il est testé :

- une valeur particulière d'un paramètre :  $\theta = \theta_0$ ,
- les valeurs d'un paramètre sont égales sur deux populations différentes.

5. **L'hypothèse alternative  $H_1$**

C'est une hypothèse contre laquelle l'hypothèse nul est testée, c'est-à-dire la négation de  $H_0$ . Il peut tester :  $\theta = \theta_1, \theta > \theta_0, \theta < \theta_0, \theta \neq \theta_0$

6. **L'hypothèse simple** : C'est l'hypothèse de type  $H : \theta = \theta_0$  avec le paramètre  $\theta$  est une valeur unique.

2.2.2 Le test unilatéral et le test bilatéral

Avant d'exécuter tout test statistique, il faut définir le problème posé. Et en appliquent un permet deux tests, unilatéral ou bilatéral.

**Définition 8 Le test unilatéral**

C'est le test s'exécute pour une variable est supérieure ou inférieure à une autre, ou à une valeur donnée.

Le rejeter de l'hypothèse  $H_0$  est déterminée d'un seul côté de distribution de probabilité de référence. La première figure présente le teste unilatéral à droite  $H_0 : \theta = \theta_0$  et  $H_1 : \theta > \theta_0$

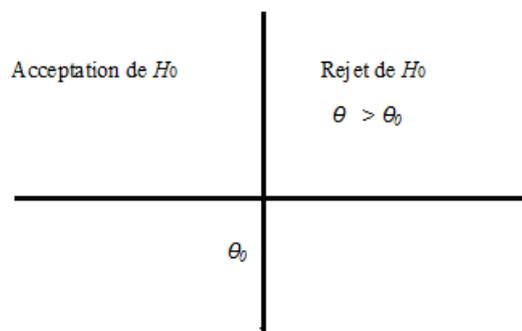


FIGURE 2.1 – Test unilatéral à droite

La deuxième figure présente le teste unilatéral à gauche  $H_0 : \theta = \theta_0$  et  $H_1 : \theta < \theta_0$

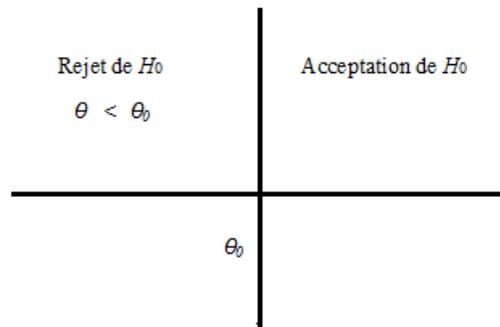


FIGURE 2.2 – Test unilatéral à gauche

**Définition 9 Le test bilatéral**

*C'est le test s'exécute pour deux variables différents, ou entre une variable et une valeur donnée sans se préoccuper du signe ou du sens de la différence.*

Le rejeter de l'hypothèse  $H_0$  est déterminée de deux côtés de distribution de référence. La figure suivante présente le teste bilatéral  $H_0 : \theta = \theta_0$  et  $H_1 : \theta \neq \theta_0$  :

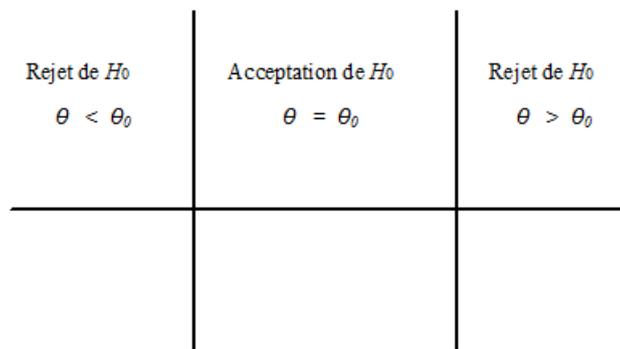


FIGURE 2.3 – Test bilatéral

Les paragraphes suivante sont relever de livre de “**A. Wald. (2013)**” et le livre de “**G. Saporta. (2006)**” .

**2.2.3 La nature générale d'une procédure de test**

La procédure de test conduisant à l'acceptation ou au rejet d'une hypothèse divisant le fonctionnement de la totalité de tous les échantillons possibles de taille  $n$  en deux parties : **Partie 1** : Est appelée **région critique**  $W$  c'est l'ensemble des valeurs de la variable aléatoire de décision et l'hypothèse soit rejeté.

**Partie 2** : C'est la **région complémentaire**  $\bar{W}$  est représenté la région d'acceptation déterminée de manière unique par la partie 1, et l'hypothèse soit acceptée.

Le choix d'un test procédure est équivalente à la détermination d'une **région critique**.

### 2.2.3.1 Les principes choix d'une région critique

1. Nous considérons deux types des erreurs :

Erreur de 1<sup>er</sup> espèce : la décision de rejet  $H_0$  si  $H_0$  vraie

Erreur de 2<sup>e</sup> espèce : l'acceptation de  $H_0$  si  $H_1$  est vraie ( $H_0$  fausse).

2. Après avoir choisi une région critique particulière, on déterminons la probabilité de commettre une erreur 1<sup>er</sup> espèce et de 2<sup>e</sup> espèce de façon unique comme suivant :

(a)  $\alpha$  : La probabilité commettre une erreur du 1<sup>er</sup> espèce est égale à la possibilité que l'échantillon  $(x_1, \dots, x_n)$ , contenu dans la région critique ( $H_0$  est rejetée) sous l'hypothèse que  $H_0$  est vrai :

$$\alpha = \mathbf{P}(\text{rejeter } H_0 | H_0 \text{ vraie}) = \mathbf{P}(\text{accepter } H_1 | H_0 \text{ vraie}) \quad (2.2)$$

(b)  $\beta$  : La probabilité de commettre une erreur de 2<sup>e</sup> espèce est égale à la possibilité que l'échantillon observé n'appartient pas la région critique ( $H_0$  est accepté), c'est la probabilité déterminée sur l'hypothèse que  $H_1$  est vraie :

$$\beta = \mathbf{P}(\text{accepter } H_0 | H_1 \text{ vraie}) = \mathbf{P}(\text{rejeter } H_1 | H_1 \text{ vraie}) \quad (2.3)$$

3.  $\alpha$  est appelée **la taille** de  $W$ , et  $1 - \beta$  est appelée **la puissance** de  $W$ . Donc en écrire **La région critique** comme suit :

$$\mathbf{P}(W | H_1) = 1 - \beta \quad \mathbf{P}(\bar{W} | H_0) = 1 - \alpha$$

#### 4. La règle de décision

Si la valeur de statistique considérée appartient à la région d'acceptation  $\bar{W}$ , on choisissons l'hypothèse nul  $H_0$ , si elle appartient à la région critique  $W$ , on choisissons l'hypothèse alternative  $H_1$ .

**Remarque 2.2** (*R. Veyseyre. 2001-2006*)

*L'utilisateur a choisie la probabilité  $\alpha$  entre les valeurs les plus utilisées sont 0.05 et 0.01 et 0.1. Pour calculer la probabilité  $\beta$ , on utilisons le tableau suivant :*

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

TABLE 2.1 – Décisions et probabilités

### 2.2.3.2 Théorème de Neyman-Pearson

Pour le choix de la variable de décision et de la région critique, Neyman-Pearson donne le cas simple de paramètre inconnu  $\theta$  prend deux valeurs  $\theta_0$  et  $\theta_1$ , et  $f(x, \theta)$  est la densité de  $X$  note par  $f_i(x) := f(x, \theta_i)$ ,  $i = 0, 1$ , avec l'hypothèse nul  $H_0$  et l'hypothèse alternative  $H_1$  sont formulent comme :

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \end{cases}$$

Le principe de Neyman-Pearson est de chercher ou choisir la plus puissant région critique  $W$  d'un taille  $\alpha$  fixe et  $\beta$  soit minimisé. Avec la taille d'échantillon  $n$  est fixée.

**théorème 2.1** *Théorème de Neyman-Pearson* (G. Saporta. 2006)

$\alpha$  étant donné,  $W$  la région critique optimale est définir par l'ensemble des points de  $\mathbb{R}^n$  et constituée de tous les échantillons  $(x_1, \dots, x_n)$  tels que :

$$\frac{f_1(x_1) \dots f_1(x_n)}{f_0(x_1) \dots f_0(x_n)} \geq k(\alpha)$$

**Preuve 2.1** Supposons que  $\alpha$  est connu, soit  $L(x, \theta)$  la densité de l'échantillon  $(x_1, \dots, x_n)$ , c-a-d  $L_i(x, \theta) = f_i(x_1) \dots f_i(x_n)$ ,  $i = 0, 1$ , avec  $f_i(x_1)$  est un variable aléatoire, telle que  $i = 0$ , l'hypothèse  $H_0$  et  $i = 1$ , l'hypothèse  $H_1$ .

– Il suffit de montrer qu' il existe une constante  $k(\alpha)$ , telle que  $W$  un région de  $\mathbb{R}^n$  et

$$\frac{L(x, \theta_1)}{L(x, \theta_0)} \geq k(\alpha),$$

avec  $\alpha = \mathbf{P}(W|H_0) = \int_W L(x, \theta_0) dx$  probabilité sous  $H_0$ , alors le région  $W$  réalise le maximum de  $1 - \beta$ . Telle que

$$1 - \beta = \mathbf{P}(W|H_1) = \int_W L(x, \theta_1) dx = \int_W \frac{L(x, \theta_1)}{L(x, \theta_0)} L(x, \theta_0) dx \quad (2.4)$$

Soit  $W'$  un autre région de  $\mathbb{R}^n$  différent de  $W$  en des points où

$$\frac{L(x, \theta_1)}{L(x, \theta_0)} < k(\alpha)$$

avec probabilité  $\alpha = \mathbf{P}(W'|H_0) = \int_{W'} L(x, \theta_0) dx$  et muni de la relation :

$$1 - \beta = \mathbf{P}(W'|H_1) = \int_{W'} L(x, \theta_1) = \int_{W'} \frac{L(x, \theta_1)}{L(x, \theta_0)} L(x, \theta_0) dx \quad (2.5)$$

Les relations (2.4) et (2.5) sont différents pour les parties non participé à  $W$  et  $W'$ .

$(W - W')$  c'est l'ensemble des points de  $W$  n'appartient pas à  $W'$ , et l'ensemble  $(W' - W)$

des points de  $W'$  n'appartenant pas à  $W$ , avec  $W$  et  $W'$  admet la même mesure  $\alpha$  sous  $H_0$ . Comme suit :

$$\int_{(W-W')} \frac{L(x, \theta_1)}{L(x, \theta_0)} L(x, \theta_0) dx \quad \text{et} \quad \int_{(W'-W)} \frac{L(x, \theta_1)}{L(x, \theta_0)} L(x, \theta_0) dx$$

Avec

$$\int_{(W-W')} \frac{L(x, \theta_1)}{L(x, \theta_0)} L(x, \theta_0) dx > \int_{(W'-W)} \frac{L(x, \theta_1)}{L(x, \theta_0)} L(x, \theta_0) dx$$

D'après le théorème de la moyenne et puisque la mesure  $L(x, \theta_0) dx$  sur un ensemble de même mesure, d'une fonction strictement supérieure on a :

$$\begin{aligned} \int_{W-W'} \frac{L(x, \theta_1)}{L(x, \theta_0)} L(x, \theta_0) dx &= \frac{L(\xi, \theta_1)}{L(\xi, \theta_0)} \mathbf{P}(W - W' | H_0) \quad \text{avec} \quad \xi \in (W - W') \\ \int_{W'-W} \frac{L(x, \theta_1)}{L(x, \theta_0)} L(x, \theta_0) dx &= \frac{L(\xi', \theta_1)}{L(\xi', \theta_0)} \mathbf{P}(W' - W | H_0) \quad \text{avec} \quad \xi' \in (W' - W) \end{aligned}$$

Puisque

$$\frac{L(\xi', \theta_1)}{L(\xi', \theta_0)} < k(\alpha) \leq \frac{L(\xi, \theta_1)}{L(\xi, \theta_0)}$$

Donc

$$\frac{L(\xi, \theta_1)}{L(\xi, \theta_0)} \geq k(\alpha)$$

– Maintenant nous montrons que  $k(\alpha)$  existe :

Soit  $A(K)$  la région de  $\mathbb{R}^n$  où  $L(x, \theta_1) > KL(x, \theta_0)$  et  $P(A(K)|H_0)$  est une fonction continue monotone de  $K$ , si  $X$  est à densité continue. Comme  $L(x, \theta_1)$  c'est une densité alors elle est toujours positive donc  $\mathbf{P}(A(0)|H_0) = 1$ .

D'autre part si  $K \rightarrow \infty$ , avec une densité bornée on a  $P(A(K)|H_n) \rightarrow 0$ . Il existe donc une valeur intermédiaire  $k_\alpha$  telle que  $\mathbf{P}(A(k(\alpha))) = \alpha$ .

### 2.2.3.3 Test de la moyenne d'une loi normale $N(m; \sigma)$ avec $\sigma$ connu

Cette paragraphe est relever de livre de “**R. Veyseyre. (2001-2006)**”.

Il y a 3 types différent pour l'hypothèse  $H_1$  avec la valeur de  $\sigma$  est connu pour le test sur la moyenne :

$$\begin{cases} H_0 = m_0 \\ H_1 = m_1 > m_0 \end{cases} \quad (2.6)$$

La statistique  $\bar{X}$  suit un loi normale de  $H_0$  et  $H_1$  comme suivant :  $N(m_0; \frac{\sigma}{\sqrt{n}})$  et  $N(m_1; \frac{\sigma}{\sqrt{n}})$

La région critique est définie par erreur de 1<sup>er</sup> espèce  $\alpha$  :  $\alpha = \mathbf{P}(\bar{X}|H_0) = \mathbf{P}\left(Z > \frac{k - m_0}{\frac{\sigma}{\sqrt{n}}}\right)$

On utilisons le tableau (3.2) (voir l'annexe) pour conclut la valeur de  $k$  dans la relation  $z = \frac{k - m_0}{\frac{\sigma}{\sqrt{n}}}$ .

Pour erreur de 2<sup>e</sup> espèce on a :  $\beta = \mathbf{P}(\bar{X} < k|H_1) = \mathbf{P}\left(Z < \frac{k - m_1}{\frac{\sigma}{\sqrt{n}}}\right)$

Avec  $\frac{k - m_1}{\frac{\sigma}{\sqrt{n}}}$  est dépendant de  $\beta$  dans le tableau (3.2) (voir l'annexe).

La même chose pour le cas suivant :

$$\begin{cases} H_0 = m_0 \\ H_1 = m_1 < m_0 \end{cases} \quad (2.7)$$

La région critique est définie par :  $\alpha = \mathbf{P}(\bar{X} < k|H_0) = \mathbf{P}\left(Z < \frac{k - m_0}{\frac{\sigma}{\sqrt{n}}}\right)$ , avec  $\frac{k - m_0}{\frac{\sigma}{\sqrt{n}}}$  est dépendant de  $\alpha$ .

pour la valeur de  $\beta$  on a :  $\beta = \mathbf{P}(\bar{X} > k|H_1) = \mathbf{P}\left(Z > \frac{k - m_1}{\frac{\sigma}{\sqrt{n}}}\right)$

Le dernier cas c'est :  $H_1 : m_1 > m_0$  ou  $H_1 : m_1 < m_0$

$$\begin{cases} H_0 = m_0 \\ H_1 = m_1 \neq m_0 \end{cases} \quad (2.8)$$

La région critique est déterminer avec la même variable de décision  $\bar{X}$  par :

$$\alpha = \mathbf{P}(|\bar{X}| > k|H_0) = \mathbf{P}\left(|Z| < \frac{k - m_0}{\frac{\sigma}{\sqrt{n}}}\right) = \mathbf{P}(|Z| < z)$$

La valeur de  $\beta$  nous utilisons le même calcule dans le cas précédent. Les valeurs de  $z$  sont dans le tableau (3.2) (voir l'annexe).

L'exemple suivant est relever de livre de "**R. Veysseyre. (2001-2006)**".

**Exemple 2.3** *Considérons  $X$  un variable aléatoire représenter un ensemble des lampes à luminescence dont la durée de vie, muni un loi normale  $N(m; \sigma)$ , avec  $m = 1000$  heures et  $\sigma = 100$  heures sous les deux hypothèses.*

*Un chercheur présente une procédé de fabrication moderne et améliorer. Il a affirmé la durée de vie moyenne de rendre doit être égale à 1075 heures, si l'amélioration est réelle, le procédé accepter puisque cette processus étant plus cher. Donc les deux hypothèses sont comme suite :*

*-  $m = 1000$  heures est une hypothèse vraie et le procédé moderne n'a pas défrent de façon*

significative la durée de vie des lampes.

–  $m = 1075$  heures et le procédé moderne a présenté une réelle amélioration.

Cette hypothèse est formalisée comme suite :

$$\begin{cases} H_0 : m = m_0 = 1000 \text{ heures} \\ H_0 : m = m_1 = 1075 \text{ heures} \end{cases}$$

Nous concluons as-que l'échantillon aléatoire de taille  $n = 25$  qui permet cette hypothèse rejeter ou accepter avec l'échantillon a donné pour la statistique  $\bar{X} = 1050$  heures.

On choisir  $\alpha = 0.05$  de rejeter l'hypothèse  $H_0$ , sous la dernière est vraie, et comme le risque d'accepter l'hypothèse  $H_1$ , si la conséquence obtenu sur un possibilité d'échantillon n'avait que 5 chances sur 100 de se produire.

La variable de décision  $\bar{X}$ , c'est la moyenne d'un échantillon de taille  $n = 25$  lampes muni de loi normale  $N(1000; \frac{100}{\sqrt{25}})$

1. Puisque  $\alpha = 0.05$ , et d'après la relation (2.2) la région critique d'accepter  $H_1$  (rejet  $H_0$ ) est :

$$\alpha = \mathbf{P}(\text{accepter } H_1 | H_0 \text{ vraie}) = \mathbf{P}(W | H_0) = 0.05$$

C'est-à-dire  $\mathbf{P}(\bar{X} > k) = 0.05$ .

Soit  $Z$  la variable aléatoire centrée réduite associée à  $X$ , alors :

$$\mathbf{P}(\bar{X} > k) = \mathbf{P}\left(Z = \frac{\bar{X} - 1000}{20} > \frac{k - 1000}{20}\right) = 0.05$$

D'après les valeurs dans le tableau (3.2) (voir l'annexe) on a

$$\frac{k - 1000}{20} = 1.6449 \Rightarrow k = 1032.898 \cong 1033 \text{ heures}$$

Si  $\bar{X}$  est supérieur à un seuil  $k$  qui n'a que 5 chances sur 100 d'être dépassé, en accepter  $H_1$  de probabilité 0.05. Cette formule est indiquée par la nature de  $H_0$ .

Si  $\bar{X}$  inférieur à  $k$ , on sauvegarder  $H_0$  faute de preuve suffisant.

2. La région critique ou région de rejet de  $H_0$  si  $[\bar{X} \geq 1033]$ . Et la région d'acceptation de  $H_0$  si  $[\bar{X} < 1033]$ .
3. D'après la relation (2.3) la valeur de  $\beta$  de deuxième espèce :

$$\beta = \mathbf{P}(\text{accepter } H_0 | H_1 \text{ vraie}) = \mathbf{P}(\bar{W} | H_1) = \mathbf{P}(\bar{X} < k | H_1)$$

$$\beta = \mathbf{P}(\bar{X} < 1033) = \mathbf{P}\left(Z = \frac{\bar{X} - 1075}{20} < \frac{1033 - 1075}{20} = -2.10\right)$$

D'après le tableau (3.1) (voir l'annexe), on a

$$\beta = 1 - P(Z > 2.10) = 1 - 0.982 = 0.018$$

Et la puissance du test,  $1 - \beta = 0.982$ . Donc pour l'échantillon de taille  $n = 25$ , l'hypothèse  $H_0$  est rejeter, et  $H_1$  est accepter.

## 2.3 Le Test séquentiel

L'analyse séquentielle est une méthode mathématique d'analyse permettant de quitter des conclusions statistiques au rapport, et à mesure qu'une application se déroule ou que des informations arrivent.

Le nombre d'observations (la taille de l'échantillon  $n$ ) est une caractéristique essentielle du test séquentiel et est une variable aléatoire, puisque il est dépende de leurs résultats successifs et il dénommé comme constante pour une procédure de test.

### 2.3.1 Notion d'un test séquentiel

#### Définition 10 *L'analyse séquentielle*

*C'est une méthode qui contient des étapes de l'expérience pour définir le nombre final d'observation d'un test, chaque étape de la décision d'expérience basée sur tous les résultats précédents, est faite si l'expérience est terminée ou sera reportée.*

#### Définition 11 (A. Wald. 2013) **Test séquentiel**

*Si  $H_0$  est l'hypothèse à tester, la méthode séquentielle est réalisée par étapes. nous prendrons l'une de trois décisions  $k = 1, 2, 3$  on pose :*

$$k = 1, \text{ en acceptant } H_0, \quad k = 2, \text{ en rejetant } H_0$$

$$k = 3, \text{ en continuant l'expérience en faisant une observation supplémentaire}$$

*Si la décision 1 ou 2 est satisfait, le processus est terminé, c'est le cas de l'hypothèse statistique. Si 3 est satisfait, un autre teste est effectué. Ce processus est continue jusqu'à ce que la décision 1, ou 2, soit satisfait.*

*$M_k$  est l'espace de dimension  $k$  contenant la totalité de tous les échantillons possibles  $(x_1, \dots, x_k)$  de taille  $k$ , et pour chaque valeur entière  $k$ ,  $M_k$  est divisé en trois sous-ensembles mutuellement exclusifs,  $\overline{W}_k^0$ ,  $\overline{W}_k^1$ , et  $\overline{W}_k$ .*

La procédure de test de  $H_0$  décrit sur la forme suivant :

$$(x_1, \dots, x_k) \in \begin{cases} \overline{W}_k^0 & \Rightarrow H_0 \text{ accepté, } n = k \\ \overline{W}_k^1 & \Rightarrow H_0 \text{ rejetée, } n = k \\ \overline{W}_i & \Rightarrow \text{test suivant } (k \rightarrow k + 1) \end{cases} \quad k = 1, 2, 3, \dots$$

**Définition 12 L'échantillon inefficace**

L'échantillon inefficace c'est l'échantillon  $(x_1, \dots, x_k)$  qui contient un segment initial  $(x_1, \dots, x_{k'})$ , où  $k' < k$  tel que  $(x_1, \dots, x_{k'}) \in \overline{W}_k^0 \cup \overline{W}_k^1$ . Sinon, un échantillon est appelé **efficace**.

Après un choix du  $\overline{W}_k^0, \overline{W}_k^1, \overline{W}_k$ ,  $k = 1, 2, \dots$  a été faite, la probabilité que le processus se termine avec l'acceptation de  $H_0$  dépend du test seulement sur la distribution de la variable aléatoire  $X$ , il est connue sauf pour les valeurs d'un nombre fini avec des paramètres  $\theta = (\theta_1, \dots, \theta_k)'$  inconnue, et notée la forme fonctionnelle  $f$  est connue par la forme suivant :  $f(x, \theta_1, \dots, \theta_k)$

**2.3.1.1 Degré de préférence pour l'acceptation ou le rejet**

Pour déterminer le degré de préférence pour l'acceptation ou le rejet en fonction de  $\theta$ , on notée  $\omega$  l'ensemble de tous les paramètres sont dépendant de  $H_0$ .

On peut dire que l'acceptation de  $H_0$  est préférée à chaque fois que  $\theta \in \omega$ , et que le rejet de  $H_0$  est préféré chaque fois que  $\theta \notin \omega$ , et en fonction de la proximité de la limite de  $\omega$ , c-à-d les paramètres  $\theta$  se trouvent dans  $\omega$  mais sont proche de la limite.

Nous obtenons une subdivision l'espace des paramètres en trois zones :

1. **zone d'acceptation** : est sous-ensemble de  $\omega$  contient tous les  $\theta \in \omega$  pour lesquels l'acceptation de  $H_0$  est fortement préférée.
2. **zone de rejet** : est une sous ensemble du complément de  $\omega$  et notée  $\bar{\omega}$  contient tous les paramètres  $\theta$  pour lesquels le rejet de  $H_0$  est fortement préféré.
3. **zone d'indifférence** : est une sous ensemble des points près de la limite de  $\omega$ , c-à-d les paramètres  $\theta$  pour lequel ni l'acceptation ni le rejet de  $H_0$  ne sont fortement préférés.

Une représentation raffinée du degré de préférence de décision donnée par deux fonctions :

1.  $\omega_0(\theta)$  formule de l'importance relative de l'erreur d'accepter  $H_0$  quand  $\theta$  est vraie.
2.  $\omega_1(\theta)$  formule de l'importance relative de l'erreur de rejeter  $H_0$  quand  $\theta$  est vraie.

$$\theta \in \begin{cases} \omega & \Rightarrow \text{resp. } \omega_1(\theta) > 0 \\ \bar{\omega} & \omega_0(\theta) > 0 \end{cases}$$

Pour la zone d'indifférence si on définit les constantes  $c_0 > 0$  et  $c_1 > 0$ , et d'après le figure (2.4) on a  $\theta$  dans la zone d'indifférence, si  $\omega_0(\theta) = 0, \text{ resp. } \omega_1(\theta) > c_1 > 0$  alors  $\theta$  appartient

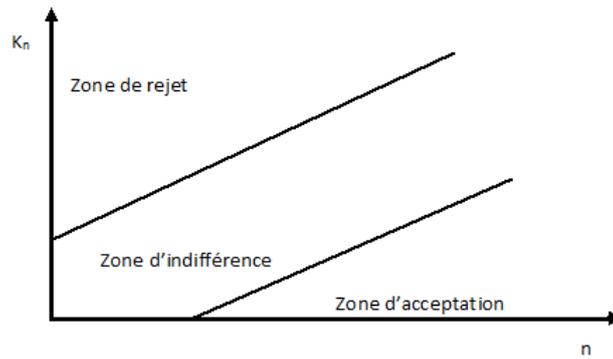


FIGURE 2.4 – Les zones de espace des paramètres

à la zone d'acceptation  $\omega$ , et si  $\omega_0(\theta) > c_0 > 0$ , resp.  $\omega_1(\theta) = 0$  alors  $\theta$  appartient à la zone de rejeter  $\bar{\omega}$ . Dans le cas pratiques, l'utilisation de la fonction de  $\omega_i(\theta)$ ,  $i = 0, 1$ , la subdivision de l'espace des paramètres en trois zones donnée une bonne approximation pour les applications pratiques.

### 2.3.1.2 Définition de la caractéristique de fonctionnement

**Définition 13** *La fonction d'efficacité* : est une fonction  $L$  de paramètre  $(\theta)$  représenter la probabilité d'accepter  $H_0$ .

$$L(\theta) = P(\text{accepter } H_0) \quad (2.9)$$

**Proposition 2.1** 1. Si  $\theta$  est compatible avec  $H_0$ , alors  $L(\theta)$  est la probabilité de faire une bonne décision correcte.

2. Si  $\theta$  n'est pas compatible avec  $H_0$ , la probabilité de prendre une décision correcte est égale à  $1 - L(\theta)$ .

3. Si  $L(\theta)$  est élevée pour  $\theta$  dépendant de  $H_0$ , et  $L(\theta)$  faible si pour  $\theta$  indépendant de  $H_0$ . La probabilité de rejeter  $H_0$  est égale à  $1 - L(\theta)$  ou égal à  $1 - \beta$ .

### Définition 14 Moyen de nombre d'échantillon

Est une fonction notée par  $E_\theta(n)$ , c'est la valeur attendue  $E(n)$  avec  $n$  nombre des observations avec distribution dépend de  $\theta$ .

Les ides des paragraphes suivante sont relever de livre de "P. Gabe. (1969)".

### 2.3.2 Contrôle par échantillonnage

En 1929, Dodge et Konig sont représenté un cas le plus nécessaire dans le procédure séquentiel, c'est le plan double d'échantillonnage qui permet dans la première étape en utilise

un échantillon de taille fixe et on s'arrête quand on a un résultat approprié avec une évidence complète. Sinon on prend un deuxième échantillon de taille fixe et on termine l'expérience.

### 2.3.2.1 Plan d'échantillonnage

#### Définition 15 *Plan d'échantillonnage*

*C'est un plan consiste à choisir la taille d'un telle analyse des lots, chaque lot contient un grand nombre d'élément donné, chacun élément pouvant être qualifié de bon ou défectueux, qui permet en fournir des règles pour l'acceptation ou le rejeter d'un lot.*

#### Définition 16 *Plan d'échantillonnage simple*

*Pour chaque lot on analyse un nombre donné  $N$  d'élément. Le lot est accepter lorsque nous ne dépasserons pas l'élément défectueux  $c$ . par contre le lot est rejeté, si l'on rencontre ou plus l'élément défectueux  $c$ .*

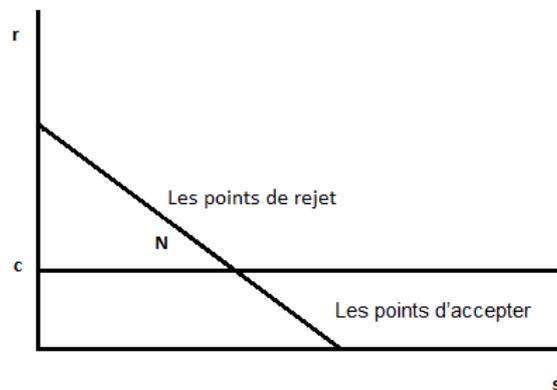


FIGURE 2.5 – Plan d'échantillonnage simple

*Pour le plan d'échantillonnage simple on a le point représentatif du résultat sur la droite  $r + s = N$ , elle se trouve plus haut, puisque on observe  $N$  élément dans chaque lot. Les points de rejet c'est les points dans cette droite qui vérifient  $r \geq c$ , et les points d'accepter c'est les points qui satisfait  $r \leq c - 1$ .*

### 2.3.2.2 Contrôle tronqué

Le plan tronqué contient les résultats dans deux axes, l'axe vertical représenter le nombre des éléments défectueux  $r$ , et l'axe horizontal représenter le nombre des éléments acceptable  $s$ , telle que  $r + s = N$ .

#### Définition 17 *Plan tronqué*

*C'est un plan d'échantillonnage qui contrôler les éléments de  $r \geq c$  et arrêter l'observation et rejeter le lot. Et l'élément de  $s \geq N - c + 1$  et arrêter l'observation et accepter le lot.*

**Exemple 2.4** *Pour  $N = 20$  et  $c = 3$  on le plan tronqué suivant :*

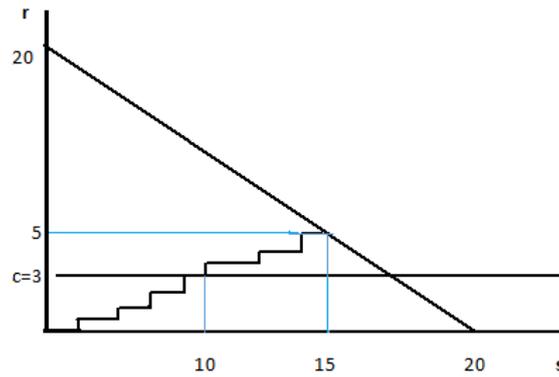


FIGURE 2.6 – Plan tronqué

Le graphique est défini par deux frontières, la première équation  $r = c$  c'est l'axe horizontal et la deuxième équation  $s = N - c + 1$  c'est l'axe vertical, et on arrête le contrôle si que l'une des frontières est remplie. Cette plan est séquentiel car pour chaque élément observé on arrête pour si l'un ou l'autre des conditions est satisfait et on continuerons l'observation, et en plus la taille d'échantillon n'est pas fixe elle est entre  $c$  et  $N$  avec  $c$  le première élément est défectueux.

Le plan séquentiel est plus compliqué puisque plan de taille fixe, et l'observation est plus lente. Les frontières de décision de graphique (2.6) sont modifiées et améliorer les propriétés de la taille moyenne de l'échantillon sans modifications à mener des plans plus en plus compliqué, en pratique il est difficile pour définir les condition pour remplir d'échantillonnage.

Pour simplifier les plans séquentiels et diminue le nombre de vérification par rapport aux frontières de décision on mettre les éléments par groupe.

### 2.3.2.3 Les propriétés du plan tronqués

1. Dans le plan tronqué on définit la fonction d'efficacité c'est la probabilité de lot qui accepter comme suit :

$$\mathbf{P}(\text{acceptation}|\theta) = \mathbf{P}(r \leq c - 1|\theta) = \sum_{r=0}^{c-1} \binom{N}{r} \theta^r (1 - \theta)^{N-r}$$

avec  $\theta$  le proportion de défectueux dans le lot, et  $c$  l'élément de défectueux et  $r$  les nombre des éléments défectueux. La courbe d'efficacité est représentée sur la figure suivante :



FIGURE 2.7 – Courbe d'efficacité

En théorie nécessaire d'inspecter la totalité de chaque lot pour tracer cette courbe au moins  $\theta \neq 0$  ou  $\theta \neq 1$  pour la façon suivant

$$\mathbf{P}(\textit{acceptation}|\theta) = \begin{cases} 1 & \text{si } \theta < \theta_0 \\ 0 & \text{si } \theta > \theta_1 \end{cases}$$

Et  $[\theta_0, \theta_1]$  c'est intervalle de région d'indifférence et on ne nécessite pas la décision qui on a utilisé, pour simplifier on a le figure :

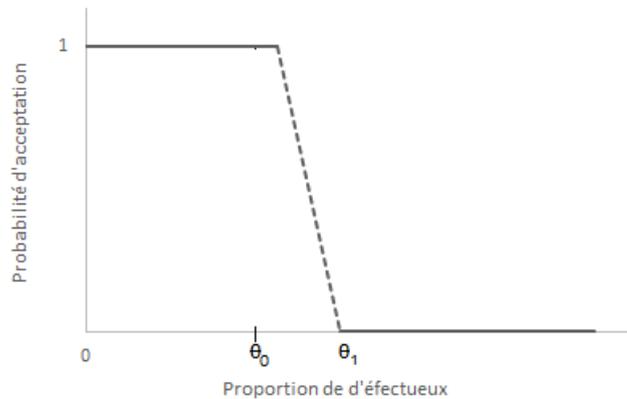


FIGURE 2.8 – Courbe d'efficacité idéale pour  $[\theta_0, \theta_1]$

Même dans le cas de moins désagréable. On n'éviter pas le contrôle complet de chaque lot, mais on a la modification suivant :

$$\mathbf{P}(\textit{acceptation}|\theta) = \begin{cases} 1 - \alpha & \text{si } \theta = \theta_0 \\ \beta & \text{si } \theta = \theta_1 \end{cases} \quad (2.10)$$

Avec la taille  $\alpha$  et  $\beta$  la puissance des tests d'hypothèse de Neymann-Pearson, et les paramètres  $N$  et  $c$  du plan tronqué choisir pour satisfait la relation 2.10.

2. Pour la distribution de la taille de l'échantillon. Un lot rejeté avec la taille de l'échantillon entre  $c$  et  $N$  on a :

Pour  $n = c$  :  $\mathbf{P}(n) = \theta^c$

Pour  $n = c + r$  :  $\mathbf{P}(n) = \binom{c+r-1}{c-1} \theta^c (1-\theta)^r$ , pour  $r = 0, 1, \dots, N - c$ .

Et pour un lot accepter avec la taille de l'échantillon entre  $N - c + 1$  et  $N$ , pour  $n = N - c + 1 + r$  et  $r = 0, 1, \dots, c - 1$  :  $\mathbf{P}(n) = \binom{N-c+r}{r} \theta^r (1-\theta)^{N-c+1}$

Dans le plan tronqué d'effectif moyen dépende de  $\theta$  présenter un maximum aux environs de  $\theta = \frac{c}{N}$ .

L'ordre de l'effectif moyen pour  $\theta$  à des valeurs faibles est  $\frac{N - c + 1}{1 - \theta}$ . Et  $\frac{c}{N}$  c'est l'ordre de  $\theta$  à des valeurs élevées sur la figure (2.8), on a l'axe horizontal indique la taille de l'échantillon fixe.

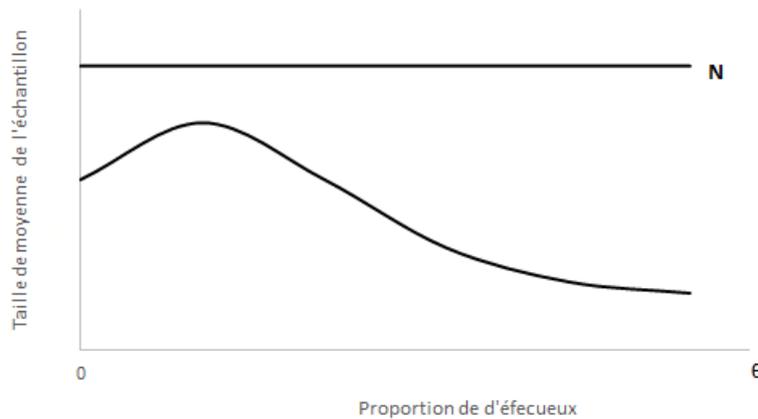


FIGURE 2.9 – La fonction de l'effectif de moyenne de l'échantillon dans le cas de plan tronqué

3. Plan de contrôle par échantillonnage ce sont des frontières pour l'acceptation ou rejeter un lot par rapport l'élément de défautueux  $c$ , tandis que la taille de l'échantillon est faible pour tous les éléments inférieure à  $c$  et la taille de l'échantillon est élevées pour les éléments supérieure à  $c$ .
4. Il y a deux type des frontières pour les valeurs faibles et élevées de  $\theta$  qui remplir les condition (2.10). Wald définir cette frontières comme des droites parallèles et une probabilité non nul de la taille de l'échantillon grande comme suivant :  
Les frontières fermées ont la taille de l'échantillon à une limite supérieure dans la figure suivant :

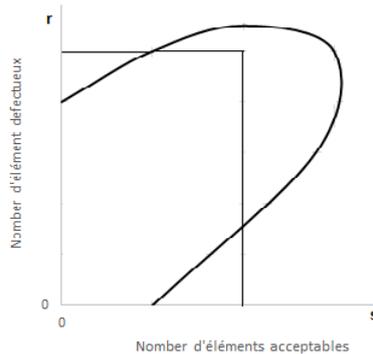


FIGURE 2.10 – Frontières fermées

Les frontières ouvertes ont la taille de l'échantillon infinie dans la figure suivant :

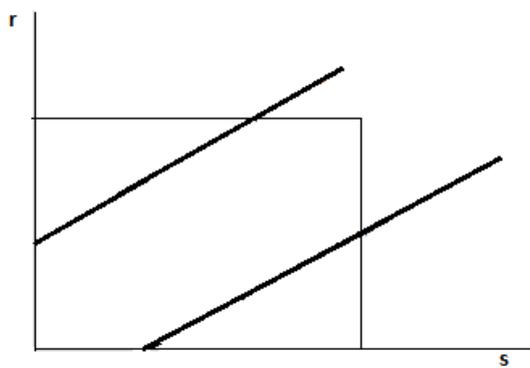


FIGURE 2.11 – Frontières ouvertes

### 2.3.3 Le test séquentiel du rapport de probabilité

Le test séquentiel du rapport de probabilité permet de choisir entre deux hypothèses simple et représenter un cas nécessaire dans la théorie séquentielle à préférer les propriétés qui existant dans le cette section.

#### 2.3.3.1 Le rapport de probabilité $p_{1l}/p_{0l}$

**Définition 18** (A. Wald. 2013)

$f(x, \theta_0)$  la densité de probabilité de variable aléatoire  $X$  pour l'hypothèse nul  $H_0 : \theta = \theta_0$  quand  $H_0$  est vrai, et  $f(x, \theta_1)$  par l'hypothèse alternative quand  $H_1$  est vrai. Supposons que les échantillons  $(x_1, \dots, x_l)$  indépendamment les uns des autres. Pour tout entier positif  $l$  et à chaque  $l$ -ème étape de l'expérience. **Le rapport de probabilité**  $p_{1l}/p_{0l}$  est calculé comme suit :  $p_{0l}$  pour  $H_0$  est vrai, et  $p_{1l}$  si  $H_1$  est vrai avec

$$p_{0l} = f(x_1, \theta_0), \dots, f(x_l, \theta_0) \quad \text{et} \quad p_{1l} = f(x_1, \theta_1), \dots, f(x_l, \theta_1)$$

$$p_{1l}/p_{0l} = \frac{f(x_1, \theta_1) \dots f(x_l, \theta_1)}{f(x_1, \theta_0) \dots f(x_l, \theta_0)} \quad \text{et} \quad B < p_{1j}/p_{0j} < A, \quad j = 1, \dots, l-1,$$

Pour deux constantes positives  $A, B$  telle que  $B < A$ , et le test est terminé si :

$$p_{1l}/p_{0l} \geq A \quad \text{et} \quad p_{1l}/p_{0l} \leq B$$

et le test considéré le rejet de  $H_0$  ou acceptation de  $H_1$  si  $p_{1l}/p_{0l} \geq A$  et si  $p_{1l}/p_{0l} \leq B$  rejet de  $H_1$  ou acceptation de  $H_0$

### 2.3.3.2 Les propriétés de test séquentiel de rapport

#### 1. Relations entre les paramètres $\alpha, \beta, A$ et $B$

Un échantillon  $(x_1, \dots, x_l)$  est appelé de **type 0** si

$$B < p_{1j}/p_{0j} < A \quad j = 1, \dots, l-1 \quad \text{et} \quad p_{1l}/p_{0l} \leq B,$$

et conduit à l'acceptation de  $H_0$ . Et de **type 1** si

$$B < p_{1j}/p_{0j} < A \quad j = 1, \dots, l-1 \quad \text{et} \quad p_{1l}/p_{0l} \geq A,$$

et conduit à l'acceptation de  $H_1$  ou rejet de  $H_0$ .

Supposons que le processus séquentiel toujours terminer. La probabilité d'un échantillon de type 1 est au moins  $A$  fois plus grande sous  $H_1$  que sous  $H_0$ , donc la probabilité de la totalité de tous les échantillons de type 1 est  $A$  fois plus grande sous  $H_1$  que sous  $H_0$ , et est par la relation (2.2) cette probabilité égale à  $\alpha$ , quand  $H_0$  est vrai, et égal à  $1 - \beta$  quand  $H_1$  est vrai.

Ceci conduit à la limite supérieure suivante pour  $A$  (inférieure pour  $B$ ) :

$$\frac{1 - \beta}{\alpha} \geq A \quad \text{et} \quad \frac{\beta}{1 - \alpha} \leq B$$

#### 2. Détermination des paramètres $A$ et $B$ en pratique

**théorème 2.2 (Wald)** (P. Gabe. 1969)

Les constantes  $A$  et  $B$  est défini dans le test séquentiel du rapport de probabilité comme suivant :

$$A \simeq \frac{1 - \beta}{\alpha} \quad \text{et} \quad B \simeq \frac{\beta}{1 - \alpha}$$

**Preuve 2.2** Nous avons étudié les erreurs de la 1<sup>er</sup> espèce et le 2<sup>e</sup> espèce c-à-d, nous obtenons les probabilités dans ce cas par :

(a) Si  $H_0$  est vrai et le test décide est  $H_0$  alors la probabilité requise c'est  $1 - \alpha$ .

- (b) Si  $H_0$  est vrai et le test décide est  $H_1$  alors la probabilité requise c'est  $\alpha$ .  
(c) Si  $H_1$  est vrai et le test décide est  $H_1$  alors la probabilité requise c'est  $1 - \beta$ .  
(d) Si  $H_1$  est vrai et le test décide est  $H_0$  alors la probabilité requise c'est  $\beta$ .

Soit  $p_i(x_1, x_2, \dots, x_l)$  la probabilité de l'ensemble d'observations  $(x_1, x_2, \dots, x_l)$  quand  $H_i, i = 0, 1$  est vrai. Donc, pour tous ensemble d'observation pour choisir  $H_0$  on a :

$$p_1(x_1, x_2, \dots, x_l) < Bp_0(x_1, x_2, \dots, x_l) \quad (2.11)$$

et pour choisir  $H_1$  on a :

$$Ap_1(x_1, x_2, \dots, x_l) < p_0(x_1, x_2, \dots, x_l) \quad (2.12)$$

- (a) Pour le cas de  $H_0$  est décider lorsque  $H_0$  est vrai, on a  $D_0$  l'ensemble de toutes les cas possibles  $(x_1, x_2, \dots, x_l)$  pour tout  $l$  :

$$\sum_{D_0} p_0(x_1, x_2, \dots, x_l) = 1 - \alpha \quad \text{et} \quad \sum_{D_0} p_1(x_1, x_2, \dots, x_l) = \beta$$

D'après la relation (2.11) on a

$$\sum_{D_0} p_1(x_1, x_2, \dots, x_l) \leq B \sum_{D_0} p_0(x_1, x_2, \dots, x_l)$$

Donc

$$\beta \leq B(1 - \alpha) \Rightarrow \frac{\beta}{1 - \alpha} \leq B$$

- (b) La même chose pour  $H_1$  est décider lorsque  $H_1$  est vrai :

$D_1$  l'ensemble de toutes les cas possibles  $(x_1, x_2, \dots, x_l)$  pour tout  $l$  :

$$\sum_{D_1} p_0(x_1, x_2, \dots, x_l) = 1 - \beta \quad \text{et} \quad \sum_{D_1} p_1(x_1, x_2, \dots, x_l) = \alpha$$

D'après la relation (2.12) on a

$$A \sum_{D_1} p_1(x_1, x_2, \dots, x_l) \leq \sum_{D_1} p_0(x_1, x_2, \dots, x_l)$$

Donc

$$A\alpha \leq 1 - \beta \Rightarrow \frac{1 - \beta}{\alpha} \geq A$$

### 3. La fonction d'efficacité de test séquentiel du rapport de probabilité

**théorème 2.3 (Wald)** (P. Gabe. 1969)

La fonction d'efficacité de test séquentiel du rapport de probabilité définir pour la relation

suivent :

$$L(\theta) = L(\theta|\alpha, \beta, \theta_0, \theta_1) \simeq \frac{A^{h(\theta)} - 1}{A^{h(\theta)} - B^{h(\theta)}}$$

Avec  $L(\theta)$  est la probabilité de choisir  $\theta_0$  et  $h(\theta)$  est solution de fonction dans le cas continu et cas discret :

$$\int_{-\infty}^{+\infty} \left( \frac{f(x, \theta_1)}{f(x, \theta_0)} \right)^{h(\theta)} f(x, \theta) dx = 1 \quad \text{resp.} \quad \sum_x \left( \frac{f(x, \theta_1)}{f(x, \theta_0)} \right)^{h(\theta)} f(x, \theta) = 1 \quad (2.13)$$

**Preuve 2.3** Supposons que une valeur  $h(\theta)$  telle que

$$\int_{-\infty}^{+\infty} \left( \frac{f(x, \theta_1)}{f(x, \theta_0)} \right)^{h(\theta)} f(x, \theta) dx = 1$$

Est satisfaite. Donc la fonction de distribution est donnée par suite :

$$f^*(x, \theta) = \left( \frac{f(x, \theta_1)}{f(x, \theta_0)} \right)^{h(\theta)} f(x, \theta)$$

Le test du rapport séquentiel de probabilité  $S^*$  comme un test  $H : f(x, \theta)$  est la distribution vraie, contre  $H^* : f^*(x, \theta)$  est la distribution vraie, pour  $\theta$  donner avec  $A^{h(\theta)}$  et  $B^{h(\theta)}$  de la forme

$$B^{h(\theta)} < \frac{f^*(x, \theta)}{f(x, \theta)} < A^{h(\theta)}$$

Le test du rapport séquentiel de probabilité  $S^*$  s'arrête et décide pour  $f(x, \theta)$  ou  $f^*(x, \theta)$ , et le test du rapport de probabilité séquentielle  $S$  : comme une test  $H_0 : \theta = \theta_0$  contre  $H_1 : \theta = \theta_1$ , avec  $A$  et  $B$  de la forme  $B < p_{1j}/p_{0j} < A$ .

Le test du rapport séquentiel de probabilité permet  $S$  s'arrête et décide pour  $\theta_0$  ou  $\theta_1$ . C'est-à-dire que la valeur de  $L(\theta)$  est exactement égale la probabilité que  $S^*$  à l'acceptation de  $H$ , donc  $H_0$  en raison de la probabilité pour que le test  $S$  décidé  $\theta_0$  est  $L(\theta)$  égala la probabilité pour que le test  $S^*$  décidé  $f(x, \theta)$ . Donc le test du rapport de probabilité séquentielle  $S^*$  est défini comme suite :

(a)  $\mathbf{P}(f(x, \theta) \text{ est vrai et le test } S^* \text{ décidé en } f(x, \theta)) = 1 - \alpha^*$ .

(b)  $\mathbf{P}(f(x, \theta) \text{ est vrai et le test } S^* \text{ décidé en } f^*(x, \theta)) = \alpha^*$ .

(c)  $\mathbf{P}(f^*(x, \theta) \text{ est vrai et le test } S^* \text{ décidé en } f^*(x, \theta)) = 1 - \beta^*$ .

(d)  $\mathbf{P}(f^*(x, \theta) \text{ est vrai et le test } S^* \text{ décidé en } f(x, \theta)) = \beta^*$ .

Donc

$$L(\theta) = 1 - \alpha^* \quad (2.14)$$

D'après le théorème (2.2), on a

$$A^{h(\theta)} \simeq \frac{1 - \beta^*}{\alpha^*} \quad \text{et} \quad B^{h(\theta)} \simeq \frac{\beta^*}{1 - \alpha^*} \quad (2.15)$$

En déduire la valeur de  $\alpha^*$ , d'après la relation (2.15) :

$$1 - A^{h(\theta)}\alpha^* = B^{h(\theta)}(1 - \alpha^*) \Rightarrow \alpha^* \approx \frac{1 - B^{h(\theta)}}{A^{h(\theta)} - B^{h(\theta)}}$$

En remplacer la valeur de  $\alpha^*$  dans (2.14) en conclue

$$L(\theta) \simeq \frac{A^{h(\theta)} - 1}{A^{h(\theta)} - B^{h(\theta)}}$$

#### 4. La fonction moyenne de nombre d'échantillon de test séquentiel du rapport de probabilité

La fonction moyenne de nombre d'échantillon  $\mathbf{E}_\theta(n)$  est la valeur attendue de  $n$  lorsque  $\theta$  est la vraie valeur du paramètre, pour  $n$  est le nombre d'observation.

La formule d'approximation pour la fonction moyenne de nombre d'échantillon du test séquentiel du rapport de probabilité, dans la fin de test séquentiel l'excès de  $p_{1m}/p_{0m}$  sur les frontières  $A$  et  $B$  est négligé. Avec :  $z := \ln \frac{f(x, \theta_1)}{f(x, \theta_0)}$

Et  $\mathbf{E}_\theta(z)$  comme valeur attendue de  $z$  quand  $\theta$  est la vraie valeur :

$$\mathbf{E}_\theta(n) \approx \frac{L(\theta) \ln B + (1 - L(\theta)) \ln A}{\mathbf{E}_\theta(z)}, \quad \mathbf{E}_\theta(z) \neq 0 \quad (2.16)$$

## 2.4 Les essais séquentiels par groupe et arrêt précoce

La conception séquentielle de groupe est une avance statistique dans la conception d'essais cliniques c-à-d la taille de l'échantillon de l'essai n'est pas fixée à l'avance et que les données successivement évaluées sont groupées et les données sont analyses a des point prédétermines dans l'essai quand les essais arrêté à ces points, cette règle d'arrêt est montre que l'essai soit arrêté quand on trouve les résultats positive, alors l'essai pour suivi, sinon le résultat est négative, il termine à tout moment.

### Conclusion

On remarque que le test est un utile nécessaire dans la décision pour accepter ou rejeter une hypothèse scientifique par la méthode séquentielle où la taille d'échantillon n'est pas fixe, et où il est évalué les données et l'échantillonnage est arrêté par une règle d'arrêt bien définie. Dans le troisième chapitre, nous expliquerons plus.

## Chapitre 3

# Les tests séquentiels dans les essais cliniques

## Introduction

Dans ce chapitre on va voir tous les tests présentés peuvent être observés comme un test de signification répétée car, chaque analyse l'hypothèse nul  $H_0$  est rejetée si le test non-séquentiel significatif à un certain niveau. Et nous présentent le test de Pocock et le test d'O'Bring et Fleming, et expliqué la comparaison entre des cette test et la fin la discussion.

### 3.1 Les tests séquentiels de groupe

Les tests séquentiels de groupe est un cas particulier d'un test séquentiel généralement, puisque les données accumulées sont analysées à intervalles plutôt qu'après chaque nouvelle observation. La surveillance continue des données peut constituer une lourde charge pratique, et l'introduction des tests séquentiels de groupe a conduit à une utilisation beaucoup plus large des méthodes séquentielles leur impact a été particulièrement évident dans les essais cliniques.

Dans la comparaison de deux traitements de base,  $K$  un maximum de nombre de groupes, et  $l$  une taille de groupe sont choisis, les sujets sont attribués aux traitements selon un schéma de randomisation contraint qui assure  $l$  sujets reçoivent chaque traitement dans chaque groupe, et les données accumulées sont analysé après chaque groupe de  $2l$  réponses pour chaque  $k = 1, \dots, K$ .

La statistique standardisé  $Z_k$ , est calculée à partir des  $k$  premiers groupes d'observations, si  $|Z_k| \geq c_k$  alors  $H_0$  est rejeté et le test est terminé, sinon le test continue à l'analyse  $K$ -ieme et  $|Z_k| < c_k$  alors  $H_0$  est accepté et s'arrête à ce point. La séquence des constants critiques,  $c_1, \dots, c_K$ , est choisi pour réaliser une erreur de 1<sup>re</sup> espèce spécifiée, et différents types de test séquentiel de groupe à différentes séquences. Et la taille du groupe  $l$ , est déterminée isolément par une puissance condition.

#### 3.1.1 Le test bilatéral

En dénotant  $\theta$  c'est la différence de moyen. L'hypothèse nul  $H_0 : \theta = 0$  représenter les réponses de la même distribution sous les deux traitements. L'hypothèse alternative  $H_1 : \theta \neq 0$ , contient deux cas,  $\theta < 0$  et  $\theta > 0$ , qui correspondent à un traitement étant supérieur (inférieur) à l'autre.  $Z$  la statistique de test standardisée, si  $|Z| > c$  avec  $c$  une constante, alors  $H_0$  est rejet. On définit la probabilité d'erreur de 1<sup>re</sup> espèce comme la probabilité de rejeter l'hypothèse nul comme :  $\mathbf{P}_{\theta=0}(|Z| > c)$

La puissance d'un test est la probabilité de rejeter l'hypothèse nul lorsqu'il ne le fait pas

tenir  $\mathbf{P}_\theta (|Z| > c)$ , pour les valeurs de  $\theta \neq 0$  avec  $\beta$  la probabilité d'erreur à  $\theta = \pm\delta$

$$\mathbf{P}_{\theta=\delta} (|Z| > c) = \mathbf{P}_{\theta=-\delta} (|Z| > c) = 1 - \beta \quad (3.1)$$

$\delta$  : représente la différence de traitement que les chercheurs espèrent détecter avec une forte probabilité.

**Remarque 3.1** (*Chr. Jennison. 2000*)

*En termes pratiques, il n'est pas souhaitable de rejeter  $H_0$  dans faveur de  $\theta < 0$  alors qu'en réalité  $\theta > 0$  car cela impliquerait une recommandation de l'inférieur des deux traitements. Dans ce cas, l'exigence de puissance ci-dessus peut être remplacé par*

$$\mathbf{P}_{\theta=\delta} (Z > c) = \mathbf{P}_{\theta=-\delta} (Z < -c) = 1 - \beta \quad (3.2)$$

*pour les valeurs de  $c$  et  $\delta$ , les probabilités  $\mathbf{P}_{\theta=\delta} (Z < -c)$  et  $\mathbf{P}_{\theta=-\delta} (Z > c)$  sont extrêmement petits et les deux les exigences (3.1) et (3.2) sont égales.*

### 3.1.2 Test d'échantillon fixe

Soit  $X_{Ai}$  c'est la réponse de sujet de traitement  $A$  et  $X_{Bi}$  c'est la réponse de sujet de traitement  $B$ , pour  $i = 1, 2, \dots$

Supposons que  $A$  et  $B$  sont normalement distribués avec la variance  $\sigma^2$  et la moyenne  $m_A, m_B$ , que nous écrivons

$$X_{Ai} \sim N(m_A, \sigma^2), \quad X_{Bi} \sim N(m_B, \sigma^2), \quad i = 1, 2, \dots$$

Et toutes les observations sont indépendantes.

Considérez le problème de teste l'hypothèse nul de différence de traitement contre le alternative bilatéral :

$$\begin{cases} H_0 : m_A = m_B \\ H_1 : m_A \neq m_B \end{cases} \quad (3.3)$$

Avec  $\alpha$  probabilité d'erreur de 1<sup>re</sup> espèce et puissance  $1 - \beta$  à  $m_A - m_B = \pm\delta$ .

Si  $n$  sujets sont attribués à chaque traitement,  $Z$  la statistique standardisée de la forme suivent :

$$Z = \frac{1}{\sqrt{2n\sigma^2}} \left( \sum_{i=1}^n X_{Ai} - \sum_{i=1}^n X_{Bi} \right) \sim N((m_A - m_B)\sqrt{(n/(2\sigma^2))}, 1)$$

Donc  $Z \sim N(0, 1)$  sous  $H_0$ ,  $H_0$  rejet si  $|Z| > F^{-1}(1 - \frac{\alpha}{2})$ , avec  $F$  la fonction de répartition. Pour remplir l'exigence de puissance, on a

$$1 - \beta = \mathbf{P} \left( |Z| > F^{-1}(1 - \frac{\alpha}{2}) \right)$$

Lorsque  $Z \sim N(\pm\delta\sqrt{(n/(2\sigma^2))}, 1)$ , les très petites probabilités sont ignorer que

$$Z < -F^{-1}(1 - \frac{\alpha}{2}) \text{ lorsque } m_A - m_B = \delta \text{ et } Z > F^{-1}(1 - \frac{\alpha}{2}) \text{ lorsque } m_A - m_B = -\delta$$

Donc nous avons besoin à l'alternative positive  $m_A - m_B = \delta$

$$\mathbf{E}(Z) = F^{-1}(1 - \frac{\alpha}{2}) + F^{-1}(1 - \beta)$$

Cette valeur attendue à  $\delta\sqrt{(n/(2\sigma^2))}$ . Et nous saisissons **la taille d'échantillon nécessaire  $n_f$  sujets par bras de traitement** :

$$n_f(\alpha, \beta, \delta, \sigma^2) = \frac{2\sigma^2}{\delta^2} \left( F^{-1}(1 - \frac{\alpha}{2}) + F^{-1}(1 - \beta) \right)^2 \quad (3.4)$$

En utilisons l'approximation de  $n_f$  pour un entier taille d'échantillon.

**Exemple 3.1** supposons  $\sigma^2 = 4$ , donc  $X_{Ai} \sim N(m_A, 4)$  et  $X_{Bi} \sim N(m_B, 4)$ ,  $i = 1, 2, \dots$ , et il tester  $H_0 : m_A = m_B$  avec erreur 1<sup>re</sup> espèce  $\alpha = 0,05$  et puissance  $1 - \beta = 0,9$  à  $m_A - m_B = \pm 1$ . D'après la relation (3.4) et le tableau (3.2) (voire l'annexe), donc le test de l'échantillon fixe

$$\begin{aligned} n_f(0.05, 0.1, 1, 4) &= 2 \times 4 \times \left( F^{-1}(1 - \frac{0.05}{2}) + F^{-1}(1 - 0.1) \right)^2 \\ &= 2 \times 4 \times (F^{-1}(0.975) + F^{-1}(0.9))^2 = 84.1 \end{aligned}$$

Que nous arrondissons à 85 sujets sur chaque traitement, si

$$\left| \left( \sum_{i=1}^{85} X_{Ai} - \sum_{i=1}^{85} X_{Bi} \right) \right| \geq 1.96\sqrt{(85 \times 2 \times 4)} = 51.1$$

Alors L'hypothèse nulle  $H_0$  est rejetée, sinon  $H_0$  est accepté.

## 3.2 Test de Pocock

En 1977, Pocock a développé l'idée d'un **test de signification répétée** à un constante niveau de signification nominale pour analyser les données accumulées à un relativement petit nombre de fois au cours d'une étude. L'entrée du patient est divisée en  $K$  groupes également

de taille contenant  $l$  sujets sur chaque traitement, et les données sont analysées après chaque nouveau groupe d'observations a été observé.

### 3.2.1 Définition de test de Pocock

Pour le problème d'essai (3.3), les observations  $X_{Ai} \sim N(m_A, \sigma^2)$  et  $X_{Bi} \sim N(m_B, \sigma^2)$ ,  $i = 1, 2, \dots$ , sont disponibles auprès des groupes de traitement  $A$  et  $B$ , respectivement. Le test de Pocock utilise la statistique standardisée après chaque groupe d'observation comme suivant :

$$Z = \frac{1}{\sqrt{2lk\sigma^2}} \left( \sum_{i=1}^{lk} X_{Ai} - \sum_{i=1}^{lk} X_{Bi} \right), \quad k = 1, \dots, K \quad (3.5)$$

Si, à la  $k$ -ème analyse,  $k = 1, \dots, K$ , la valeur absolue de  $Z_k$  est suffisamment grande. Le rejet de  $H_0$  est donnée arrête d'étude. Sinon  $H_0$  n'a pas été rejeté par la finale analyse  $K$ , il est accepté.

**Définition 19** (Chr. Jennison. 2000)

Le test de Pocock conserve la forme générale d'un test de signification appliqué de façon répétée,  $H_0$  rejeté à l'étape  $k$  si  $|Z_k| \geq C_P(K, \alpha)$ ,  $k = 1, \dots, K$ , sont des constante avec  $c_k = C_P(K, \alpha)$ , mais la valeur critique  $C_P(K, \alpha)$  est choisi dans le tableau (3.3) (voir l'annexe) avec  $\alpha = 0.01, 0.05, 0.1$ , pour donner une  $\alpha$  erreur globale de 1<sup>re</sup> espèce c-à-d :

$$P_{m_A - m_B = 0} (\text{Rejet } H_0 \text{ à l'analyse } k = 1, 2, \dots, \text{ ou } k = K) = \alpha$$

Donc, le test est de forme suivant :

$$\begin{aligned} &\text{Après le groupe } k = 1, \dots, K - 1 \\ &\text{Si } |Z_k| \geq C_P(K, \alpha) \quad \text{arrêter, rejeter } H_0 \\ &\text{Si } |Z_k| < C_P(K, \alpha) \quad \text{continuez à grouper } k + 1 \\ &\text{Après le groupe } K \\ &\text{Si } |Z_K| \geq C_P(K, \alpha) \quad \text{arrêter, rejeter } H_0 \\ &\text{Si } |Z_K| < C_P(K, \alpha) \quad \text{arrêter, accepté } H_0 \end{aligned} \quad (3.6)$$

**Remarque 3.2** (Chr. Jennison. 2000)

La valeur  $K = 1$  dans le tableau (3.3) (voir l'annexe) est appelle le cas non-séquentiel de  $H_0$ . Notez qu'à chaque analyse, le test rejette  $H_0$  si le niveau signification des deux côtés d'un test non-séquentiel de  $H_0$  en utilisant les données disponibles à ce moment est au-dessous

$$\alpha' = 2[1 - F(C_P(K, \alpha))]$$

Avec  $\alpha'$  c'est le niveau de signification nominale.

**Exemple 3.2** si  $K = 5$  et  $\alpha = 0,05$

Le niveau de signification nominale appliqué à chaque analyse est

$$\alpha' = 2(1 - F(2.413)) = 0,0158$$

On remarque que  $\alpha' = 0.0158 < \alpha = 0.05$ , c-à-d un niveau de signification nominal inférieur à la probabilité globale d'erreur de 1<sup>re</sup> espèce est nécessaire pour éviter le problème des "multiples visions".

### 3.2.2 L'exigence de puissance

Le test défini par la relation (3.6) a une probabilité d'erreur de 1<sup>re</sup> espèce  $\alpha$  pour toute taille de groupe  $l$ . C'est l'exigence de puissance

$$1 - \beta = \mathbf{P}_{m_A - m_B = \pm\delta}(\text{Rejet } H_0)$$

qui détermine la taille du groupe approprié. Avec la taille maximale de l'échantillon que le test séquentiel de groupe dépend de  $K, \alpha, \beta, \frac{\sigma^2}{\delta^2}$ . Et la même erreur de 1<sup>re</sup> espèce et la puissance  $1 - \beta$  à  $m_A - m_B = \pm\delta$  est également proportionnelle à  $\frac{\sigma^2}{\delta^2}$  pour la taille de l'échantillon d'un test d'échantillon fixe.

Il définir le rapport  $R_P(K, \alpha, \beta)$  de la taille maximale de l'échantillon du test séquentiel du groupe à la taille de l'échantillon fixe, et s'appliquera pour tout  $\delta$  et  $\sigma^2$ , avec les valeurs de  $R_P(K, \alpha, \beta)$  dans tableau (3.4) (voir l'annexe) qui déterminent les tailles de groupe pour les tests bilatéral avec  $K$  groupes d'observations, et la probabilité d'erreur de 1<sup>re</sup> espèce  $\alpha$  et puissance  $1 - \beta$ .

**Définition 20** (Chr. Jennison. 2000)

La taille de l'échantillon par bras de traitement,  $n_f(\alpha, \beta, \delta, \sigma^2)$ , requise par un test échantillon fixe avec erreur de 1<sup>re</sup> espèce  $\alpha$  et puissance  $1 - \beta$  à  $m_A - m_B = \pm\delta$  est donné par la relation (3.4). Un test séquentiel de groupe Pocock avec  $K$  groupes a besoin d'un maximum de  $R_P(K, \alpha, \beta)n_f(\alpha, \beta, \delta, \sigma^2)$  observations par bras de traitement et groupes de  $l$  à une valeur entière aura un léger effet sur la puissance réelle du test avec

$$l = \frac{R_P(K, \alpha, \beta)n_f(\alpha, \beta, \delta, \sigma^2)}{K}$$

observation sur chaque traitement.

**Exemple 3.3** D'après l'exemple (3.1) pour tester  $H_0 : m_A = m_B$  avec erreur de 1<sup>re</sup> espèce  $\alpha = 0.05$  et puissance  $1 - \beta = 0.9$  à  $m_A - m_B = \pm 1$  avec variance  $\sigma^2 = 4$ . De la relation

## CHAPITRE 03. LES TESTS SEQUENTIELLE DANS LES ESSAIS CLINIQUES

(3.4), la taille de l'échantillon observations par bras de traitement est  $n_f(0.05, 0.1, 1, 4) = 84.1$ . D'après le tableau (3.4) (voir l'annexe) on a  $R_P(5, 0.05, 0.1) = 1.207$ , par conséquent, un test de Pocock avec 5 groupes de les observations, la taille maximale de l'échantillon par bras de traitement

$$R_P(5, 0.05, 0.1)n_f(0.05, 0.1, 1, 4) = 1.207 \times 84.1 = 101.5$$

et la taille du groupe nécessaire  $l$  par traitement est  $l = \frac{101,5}{5} = 20,3 = 21$ , le test arrêter et l'hypothèse nul  $H_0$  est rejet à l'analyse  $k$ ,  $k = 1, \dots, 5$  si

$$\left| \left( \sum_{i=1}^{21k} X_{Ai} - \sum_{i=1}^{21k} X_{Bi} \right) \right| \geq 2.413 \sqrt{(21k \times 2 \times 4)} = 31.28\sqrt{k}$$

Si l'hypothèse nulle n'a pas été rejetée par l'analyse 5, elle est acceptée a ce point.

Pour simplifie la limite du test de Pocock par cet exemple qui à donnes la figure (3.1) comme suivant :

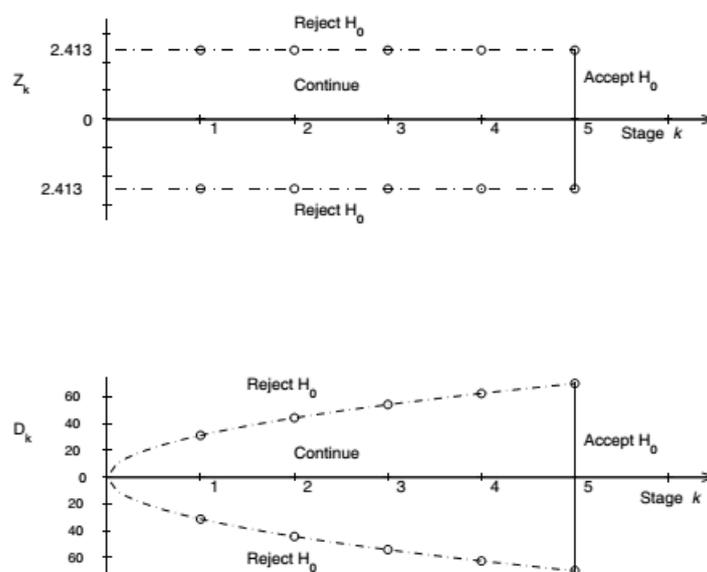


FIGURE 3.1 – Test de Pocock avec 5 groupe d'observation

On remarque que :

1. Le premier diagramme représente les valeurs critiques appliquées aux statistiques standardisées  $Z_k$ , sont des constants sur les 5 analyses pour le test de Pocock.
2. La deuxième diagramme, le même test est formulé en terme de différence non normalisées des sommes d'échantillon sur les traitements  $A$  et  $B$ ,

$$D_k = \sum_{i=1}^{21k} X_{Ai} - \sum_{i=1}^{21k} X_{Bi}, k = 1, \dots, 5,$$

Et dans cette représentation les points de limite de Pocock se trouvent sur une forme de courbe de racine carrée,  $C_P(K, \alpha)\sqrt{42k\sigma^2}$ .

3. Le test de Pocock utilise le nombre total maximum de sujets sur les deux bras de traitement nécessaire est 210. Ceci est un peu plus grand que le 170 nécessaire pour un test d'échantillon fixe.

### 3.3 Test d'O'Brien et Fleming

En 1979, O'Brien et Fleming ont donné un test dans lequel le niveau de signification nominale nécessaires pour rejeter  $H_0$  à chaque augmentation de l'analyse, Ainsi que une alternative au test de signification répétée avec constante nominale.

#### 3.3.1 Définition de test d'O'Brien et Fleming

Le test O'Brien et Fleming est défini comme suit :

**Définition 21** (*Chr. Jennison. 2000*)

*Les sujets sont attribués aux traitements en groupes, et les données accumulées sont analysées après terminé des groupes des données deviennent disponibles, et il est le terme de statistique standardisé  $Z_k$  de la forme (3.5), et  $H_0$  est rejeté après le groupe  $k$  si  $|Z_k| \geq c_k$  pour une séquence de valeurs critiques  $c_1, \dots, c_K$ .*

*Le test se la forme :*

$$\begin{aligned}
 &\text{Après le groupe } k = 1, \dots, K - 1 \\
 &\text{Si } |Z_k| \geq C_B(K, \alpha)\sqrt{\left(\frac{K}{k}\right)} \quad \text{arrêter, rejeter } H_0 \\
 &\text{Si } |Z_k| < C_B(K, \alpha)\sqrt{\left(\frac{K}{k}\right)} \quad \text{continuez à grouper } k + 1 \\
 &\text{Après le groupe } K \\
 &\text{Si } |Z_K| \geq C_B(K, \alpha) \quad \text{arrêter, rejeter } H_0 \\
 &\text{Si } |Z_K| < C_B(K, \alpha) \quad \text{arrêter, accepté } H_0
 \end{aligned} \tag{3.7}$$

Avec  $c_k = C_B(K, \alpha)\sqrt{\left(\frac{K}{k}\right)}$ ,  $k = 1, \dots, K$  et  $C_B(K, \alpha)$  des valeurs pour les tests bilatéraux avec  $K$  groupes des observations et probabilité globale de type 1<sup>re</sup> espèce  $\alpha$  est fournie dans le tableau (3.5) (voir l'annexe).

D'autre part au niveau de signification nominale, si le niveau de signification bilatéral de  $H_0$  est inférieur à  $\alpha'_k = 2\left(1 - F\left(C_B(K, \alpha)\sqrt{\left(\frac{K}{k}\right)}\right)\right)$ , alors  $H_0$  est rejeté à l'analyse  $k$ ,  $k = 1, \dots, K$ . Puisque  $C_B(K, \alpha)\left(\sqrt{\left(\frac{K}{k}\right)}\right)$  diminue avec l'augmentation de  $k$ , et  $\alpha'_k$  augmente.

**Exemple 3.4** Pour  $K = 5$  et  $\alpha = 0,05$  les niveaux de signification nominale appliqués aux analyses 1 à 5 sont

$$\alpha'_1 = 0,000005, \quad \alpha'_2 = 0,0013, \quad \alpha'_3 = 0,0084, \quad \alpha'_4 = 0,0225, \quad \alpha'_5 = 0,0413$$

### 3.3.2 L'exigence de puissance

Pour calculer la taille de l'échantillon on utilisons les constantes  $R_B(K, \alpha, \beta)$  de le même rôle que le  $R_P(K, \alpha, \beta)$  pour le test de Pocock. Avec les constantes  $R_B(K, \alpha, \beta)$  représenter dans le tableau (3.6) (voir l'annexe), pour déterminer la taille des groupes de tests bilatéraux avec  $K$  groupes d'observations, et probabilité d'erreur de 1<sup>re</sup> espèce  $\alpha$  et puissance  $1 - \beta$  La taille maximale de l'échantillon sur chaque bras de traitement est

$$R_B(K, \alpha, \beta)n_f(\alpha, \beta, \delta, \sigma^2)$$

Ou  $n_f(\alpha, \beta, \delta, \sigma^2)$  est le nombre d'observation par traitement nécessaire par un test de taille d'échantillon fixe avec la même erreur de 1<sup>re</sup> espèce et la même puissance, donnée par la forme (3.4) et la taille de groupe nécessaire par bras de traitement convenablement arrondi est

$$l = \frac{R_B(K, \alpha, \beta)n_f(\alpha, \beta, \delta, \sigma^2)}{K}$$

**Exemple 3.5** Dans l'exemple précédant (3.1) de test  $H_0 : m_A = m_B$  avec erreur de 1<sup>re</sup> espèce  $\alpha = 0,05$  et puissance  $1 - \beta = 0,9$  à  $m_A - m_B = \pm 1$  quand les observations ont une variance  $\sigma^2 = 4$ . Et d'après la formule (3.4), la taille d'échantillon fixe requise observations par bras de traitement est  $n_f(0.05, 0.1, 1, 4) = 84,1$ . Et d'après le tableau (3.6) (voir l'annexe), la valeur de  $R_B(5, 0.05, 0.1) = 1.026$ .

La taille maximale de l'échantillon par bras de traitement dans test O'Brien et Fleming avec 5 groupes de les observations ont est

$$R_B(5, 0.05, 0.1)n_f(0.05, 0.1, 1, 4) = 1.026 \times 84,1 = 86,3$$

Et la taille de groupe nécessaire par traitement est  $l = \frac{86,3}{5} = 17,3$ , qui approximé à 18. Si

$$\left| \left( \sum_{i=1}^{18k} X_{Ai} - \sum_{i=1}^{18k} X_{Bi} \right) \right| \geq 2.040 \sqrt{\left(\frac{5}{k}\right)} \sqrt{(18k \times 2 \times 4)} = 54.74$$

Alors l'hypothèse nul  $H_0$  est rejet à l'analyse  $k$ ,  $k = 1, \dots, 5$ , sinon elle est acceptée à ce point.

La limite du test d'O'Brien et Fleming pour cet exemple est illustrée dans Figure 3.2.

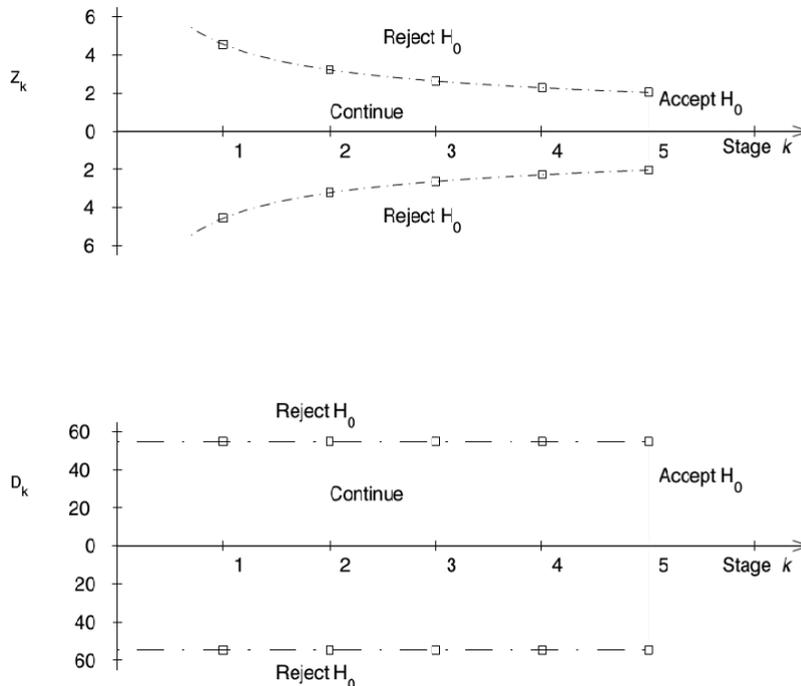


FIGURE 3.2 – Test O'Brien et Fleming pour 5 groupes d'observations

On remarque que

1. Dans le premier diagramme, le test est décrit en termes de critique valeurs pour les statistiques standardisées  $Z_k$ , les valeurs limites diminuent nettement aux analyses successives.
2. Le deuxième diagramme, montre le même test exprimé en termes des différences non normalisées des sommes d'échantillon, et sur cette échelle le test O'Brien et Fleming a limites.

$$D_k = \sum_{i=1}^{mk} X_{Ai} - \sum_{i=1}^{mk} X_{Bi}, \quad k = 1, \dots, 5$$

3. Le nombre total maximum de sujets sur les deux bras de traitement  $A$  et  $B$  pouvant être requis par le test de O'Brien et Fleming est de 180 a plus grand par rapport aux 170 pour un test d'échantillon fixe.

**Remarque 3.3** (Chr. Jennison. 2000)

Le test d'O'Brien et Fleming demande une taille de groupe plus petite que test de Pocock pour remplir la même exigence de puissance. Puisque la conséquence de  $m_A - m_B$  sur  $E(Z)$  croissance avec  $k$ .

## **3.4 Comparaison des tests séquentiels**

### **3.4.1 Comparaison de test de Pocock et de test d'O'Brien et Fleming**

A partir de la comparaison des tests séquentiels de groupe peuvent être considérées un certain nombre de propriétés.

1. Pour la comparaison qualitative entre le teste de Pocock et le teste O'Brien et Fleming sont représenter sur les figures (2.1) et (2.2) :  
le test de Pocock a premièrement des limites plus étroites, ce qui possibilité d'un arrêt très pressée. Tandis que le test O'Brien et Fleming a limites plus étroites à la fin des analyses et une plus petite taille maximale de l'échantillon.
2. Pour la comparaison quantitative, un certain nombre des essais sont disponibles. Les probabilités d'erreur sont une propriété déterminante de tout test, et il est important de comparer correctement des tests "apparié". Cependant, les probabilités d'erreur 1<sup>re</sup> espèce et de 2<sup>e</sup> espèce sont précisées, et la puissance du test est énormément déterminée sur la plage des valeurs des paramètres possibles.
3. La taille de l'échantillon fixe de 170 sujets présente un point de référence par rapport la comparaison des tests séquentiels, la tailles d'échantillon à  $m_A - m_B = 0$  ont plus attendu que cette valeur de référence, puisque pour accepter  $H_0$ , les tests séquentiels du groupe doivent continuer à leur la taille maximale d'échantillon c'est pour le test de Pocock est 210, et le test de O'Brien et Fleming est 180. Et la différence de traitement est grande il est important pour l'éthique d'arrêter et puisque la décision précoce évite la randomisation.

### **3.4.2 La taille maximale de l'échantillon**

On représente un tableau (3.7) (voir l'annexe) contient la distribution complète de la taille de l'échantillon à la fin l'étude des tests du Pocock et O'Brien et Fleming pour 5 groupes d'observations, pour avoir l'erreur 1<sup>re</sup> espèce  $\alpha = 0,05$  et puissance  $1 - \beta = 0,9$  à  $|m_A - m_B| = 1$ . Avec observations sur chaque traitement sont normaux avec variance  $\sigma^2 = 4$ . Les deux tests donnent lieu à une variation aléatoire considérable de la taille de l'échantillon et une probabilité substantielle de continuer à observer tous les 5 groupes

La taille de l'échantillon du test de Pocock Dans quelque applications, admet la plus grande variance désagréable et sa taille d'échantillon maximale plus élevée désavantage.

### 3.4.3 Le nombre des groupes

Le nombre des groupes dans lesquels les observations sont décidées par le testeur, la majorité d'un plus grand nombre des groupes impliquent des plus petites tailles d'échantillon aux valeurs de  $m_A - m_B$  est intéressantes.

Les tableaux (3.8) et (3.9) présentent la taille des échantillons maximum et attendus considérés varie en fonction du nombre des groupes dans les tests de Pocock et O'Brien et Fleming. Les résultats sont pour des tests avec  $\alpha = 0,05$  et puissance  $1 - \beta = 0,9$  à  $m_A - m_B = \pm\delta$ . Les entrées dans le tableau sont  $100\times$  les ratios des tailles d'échantillon maximum et attendues à  $n_f(\alpha, \beta, \delta, \sigma^2)$ .

Pour le test Pocock, la taille maximale de l'échantillon et la taille attendue de l'échantillon dans  $m_A - m_B$  est proche de zéro sont augmenter avec l'augmentation de  $K$  le nombre des groupes. Et d'autre part les tailles d'échantillon attendues de test d'O'Brien et Fleming aux plus grandes valeurs de  $|m_A - m_B|$  continuent de diminuer lorsque que  $K$  augmente. Il est possible de choisir 10 groupes dans tests d'O'Brien et Fleming. Et ne prend pas plus de 5 groupes dans un test de Pocock.

### 3.4.4 Discussion

Les deux types de tests séquentiels de groupe possèdent des avantages qui les rendent meilleurs que d'autres.

1. Pour  $|m_A - m_B|$  est grande, dont le test de Pocock les tailles d'échantillon attendues plus faibles que les tests d'O'Brien et Fleming et l'impératif éthique pour un arrêt précoce est important.
2. Pour  $|m_A - m_B|$  est petite, le test de Pocock a des tailles d'échantillons maximales plutôt élevées et leurs tailles d'échantillons attendues sont supérieures à celles de l'échantillon fixe.
3. Le test O'Brien et Fleming appliquent des niveaux de signification nominale extrêmement bas lors des premières analyses, puisque ils donnent peu de puissance pour une taille d'échantillon maximale donnée, et car le niveau de signification nominal appliqué à l'analyse finale est proche de la probabilité d'erreur de 1<sup>re</sup> espèce dans situation de acceptée l'hypothèse nul lorsqu'une analyse d'échantillon fixe de l'ensemble final des données sont rejeté  $H_0$  est peu probable.

# Annexe : Les tableaux statistiques

## Théorème de la moyenne

**théorème 3.1** (*E. Hairer. 2000*)

Si  $f : [a, b] \rightarrow \mathbb{R}$  est une fonction continue, alors il existe un  $\xi \in [a, b]$  tel que

$$\int_a^b f(x)dx = f(\xi)(b - a)$$

TABLE 3.1: Les valeurs de  $\mathbf{P}(Z < z)$  en fonction de  $z$

$z$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5039	0.5079	0.5119	0.5159	0.5199	0.5239	0.5279	0.5318	0.5358
0.10	0.5398	0.5437	0.5477	0.5517	0.5556	0.5596	0.5635	0.5674	0.5714	0.5753
0.20	0.5792	0.5831	0.5870	0.5909	0.5948	0.5987	0.6025	0.6064	0.6102	0.6140
0.30	0.6179	0.6217	0.6255	0.6293	0.6330	0.6368	0.6405	0.6443	0.6480	0.6517
0.40	0.6554	0.6590	0.6627	0.6664	0.6700	0.6736	0.6772	0.6808	0.6843	0.6879
0,50	0.6914	0.6949	0.6984	0.7019	0.7054	0.7088	0.7122	0.7156	0.7190	0.7224
0.60	0.7257	0.7290	0.7323	0.7356	0.7389	0.7421	0.7453	0.7485	0.7517	0.7549
0.70	0.7580	0.7611	0.7642	0.7673	0.7703	0.7733	0.7763	0.7793	0.7823	0.7852
0.80	0.7881	0.7910	0.7938	0.7967	0.7995	0.8023	0.8051	0.8078	0.8105	0.8132
0.90	0.8159	0.8185	0.8212	0.8238	0.8263	0.8289	0.8314	0.8339	0.8364	0.8389
1.00	0.8413	0.8437	0.8461	0.8484	0.8508	0.8531	0.8554	0.8576	0.8599	0.8621
1.10	0.8643	0.8665	0.8686	0.8707	0.8728	0.8749	0.8769	0.8789	0.8810	0.8829
1.20	0.8849	0.8868	0.8887	0.8906	0.8925	0.8943	0.8961	0.8979	0.8997	0.9014
1.30	0.9031	0.9049	0.9065	0.9082	0.9098	0.9114	0.9130	0.9146	0.9162	0.9177
1.40	0.9192	0.9207	0.9221	0.9236	0.9250	0.9264	0.9278	0.9292	0.9305	0.9318
1.50	0.9331	0.9344	0.9357	0.9369	0.9382	0.9394	0.9406	0.9417	0.9429	0.9440
1.60	0.9452	0.9463	0.9473	0.9484	0.9494	0.9505	0.9515	0.9525	0.9535	0.9544
1.70	0.9554	0.9563	0.9572	0.9581	0.9590	0.9599	0.9607	0.9616	0.9624	0.9632
1.80	0.9640	0.9648	0.9656	0.9663	0.9671	0.9678	0.9685	0.9692	0.9699	0.9706
1.90	0.9712	0.9719	0.9725	0.9731	0.9738	0.9744	0.9750	0.9755	0.9761	0.9767
2.00	0.9772	0.9777	0.9783	0.9788	0.9793	0.9798	0.9803	0.9807	0.9812	0.9816
2.10	0.9821	0.9825	0.9829	0.9834	0.9838	0.9842	0.9846	0.9849	0.9853	0.9857
2.20	0.9860	0.9864	0.9867	0.9871	0.9874	0.9877	0.9880	0.9883	0.9886	0.9889
2.30	0.9892	0.9895	0.9898	0.9900	0.9903	0.9906	0.9908	0.9911	0.9913	0.9915
2.40	0.9918	0.9920	0.9922	0.9924	0.9926	0.9928	0.9930	0.9932	0.9934	0.9936
2.50	0.9937	0.9939	0.9941	0.9942	0.9944	0.9946	0.9947	0.9949	0.9950	0.9952
2.60	0.9953	0.9954	0.9956	0.9957	0.9958	0.9959	0.9960	0.9962	0.9963	0.9964
2.70	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9971	0.9972	0.9973
2.80	0.9974	0.9975	0.9975	0.9976	0.9977	0.9978	0.9978	0.9979	0.9980	0.9980

## ANNEXE : LES TABLEAUX STATISTIQUES

---

TABLE 3.1: Les valeurs de  $\mathbf{P}(Z < z)$  en fonction de  $z$  (suite)

$z$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.90	0.9981	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0,9986
3.00	0.9986	0.9986	0.9987	0.9987	0.9988	0.9988	0.9988	0.9989	0.9989	0.9989
3.10	0.9990	0.9990	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992
3.20	0.9993	0.9993	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9994
3.30	0.9995	0.9995	0.9995	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996
3.40	0.9996	0.9996	0.9996	0.9996	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997
3.50	0.9997	0.9997	0.9997	0.9997	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.60	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.70	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.80	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.90	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

**ANNEXE : LES TABLEAUX STATISTIQUES**

TABLE 3.2: Les fractiles de la loi normale

$P(Z < z)$	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	0.010	
0.00	$\infty$	3.0902	2.8782	2.7478	2.6521	2.5758	2.5121	2.4573	2.4089	2.3656	2.3263	0.99
0.01	2.3263	2.2904	2.2571	2.2262	2.1973	2.1701	2.1444	2.1201	2.0969	2.0748	2.0537	0.98
0.02	2.0537	2.0335	2.0141	1.9954	1.9774	1.9600	1.9431	1.9268	1.9110	1.8957	1.8808	0.97
0.03	1.8808	1.8663	1.8522	1.8384	1.8250	1.8119	1.7991	1.7866	1.7744	1.7624	1.7507	0.96
0.04	1.7507	1.7392	1.7279	1.7169	1.7060	1.6954	1.6849	1.6747	1.6646	1.6546	1.6449	0.95
0.05	1.6449	1.6352	1.6258	1.6164	1.6072	1.5982	1.5893	1.5805	1.5718	1.5632	1.5548	0.94
0.06	1.5548	1.5464	1.5382	1.5301	1.5220	1.5141	1.5063	1.4985	1.4909	1.4833	1.4758	0.93
0.07	1.4758	1.4684	1.4611	1.4538	1.4466	1.4395	1.4325	1.4255	1.4187	1.4118	1.4051	0.92
0.08	1.4051	1.3984	1.3917	1.3852	1.3787	1.3722	1.3658	1.3595	1.3532	1.3469	1.3408	0.91
0.09	1.3408	1.3346	1.3285	1.3225	1.3165	1.3106	1.3047	1.2988	1.2930	1.2873	1.2816	0.90
0.10	1.2816	1.2759	1.2702	1.2646	1.2591	1.2536	1.2481	1.2426	1.2372	1.2319	1.2265	0.89
0.11	1.2265	1.2212	1.2160	1.2107	1.2055	1.2004	1.1952	1.1901	1.1850	1.1800	1.1750	0.88
0.12	1.1750	1.1700	1.1650	1.1601	1.1552	1.1503	1.1455	1.1407	1.1359	1.1311	1.1264	0.87
0.13	1.1264	1.1217	1.1170	1.1123	1.1077	1.1031	1.0985	1.0939	1.0893	1.0848	1.0803	0.86
0.14	1.0803	1.0758	1.0714	1.0669	1.0625	1.0581	1.0537	1.0494	1.0451	1.0407	1.0364	0.85
0.15	1.0364	1.0322	1.0279	1.0237	1.0194	1.0152	1.0110	1.0069	1.0027	0.9986	0.9945	0.84
0.16	0.9945	0.9904	0.9863	0.9822	0.9782	0.9741	0.9701	0.9661	0.9621	0.9581	0.9542	0.83
0.17	0.9542	0.9502	0.9463	0.9424	0.9385	0.9346	0.9307	0.9269	0.9230	0.9192	0.9154	0.82
0.18	0.9154	0.9116	0.9078	0.9040	0.9002	0.8965	0.8927	0.8890	0.8853	0.8816	0.8779	0.81
0.19	0.8779	0.8742	0.8706	0.8669	0.8632	0.8596	0.8560	0.8524	0.8488	0.8452	0.8416	0.80
0.20	0.8416	0.8381	0.8345	0.8310	0.8274	0.8239	0.8204	0.8169	0.8134	0.8099	0.8064	0.79
0.21	0.8064	0.8030	0.7995	0.7961	0.7926	0.7892	0.7858	0.7824	0.7790	0.7756	0.7722	0.78
0.22	0.7722	0.7688	0.7655	0.7621	0.758	0.7554	0.7521	0.7488	0.7454	0.7421	0.7388	0.77
0.23	0.7388	0.7356	0.7323	0.7290	0.7257	0.7225	0.7192	0.7160	0.7128	0.7095	0.7063	0.76
0.24	0.7063	0.7031	0.6999	0.6967	0.6935	0.6903	0.6871	0.6840	0.6808	0.6776	0.6745	0.75
0.25	0.6745	0.6713	0.6682	0.6651	0.6620	0.6588	0.6557	0.6526	0.6495	0.6464	0.6433	0.74
0.26	0.6433	0.6403	0.6372	0.6341	0.6311	0.6280	0.6250	0.6219	0.6189	0.6158	0.6128	0.73
0.27	0.6128	0.6098	0.6068	0.6038	0.6008	0.5978	0.5948	0.5918	0.5888	0.5858	0.5828	0.72
0.28	0.5828	0.5799	0.5769	0.5740	0.5710	0.5681	0.5651	0.5622	0.5592	0.5563	0.5534	0.71
0.29	0.5534	0.5505	0.5476	0.5446	0.5417	0.5388	0.5359	0.5330	0.5302	0.5273	0.5244	0.70
0.30	0.5244	0.5215	0.5187	0.5158	0.5129	0.5101	0.5072	0.5044	0.5015	0.4987	0.4958	0.69
0.31	0.4958	0.4930	0.4902	0.4874	0.4845	0.4817	0.4789	0.4761	0.4733	0.4705	0.4677	0.68
0.32	0.4677	0.4649	0.4621	0.4593	0.4565	0.4538	0.4510	0.4482	0.4454	0.4427	0.4399	0.67
0.33	0.4399	0.4372	0.4344	0.4316	0.4289	0.4261	0.4234	0.4207	0.4179	0.4152	0.4125	0.66
0.34	0.4125	0.4097	0.4070	0.4043	0.4016	0.3989	0.3961	0.3934	0.3907	0.3880	0.3853	0.65
0.35	0.3853	0.3826	0.3799	0.3772	0.3745	0.3719	0.3692	0.3665	0.3638	0.3611	0.3585	0.64

**ANNEXE : LES TABLEAUX STATISTIQUES**

TABLE 3.2: Les fractiles de la loi normale (suite)

$P(Z < z)$	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	0.010	
0.36	0.3585	0.3558	0.3531	0.3505	0.3478	0.3451	0.3425	0.3398	0.3372	0.3345	0.3319	0.63
0.37	0.3319	0.3292	0.3266	0.3239	0.3213	0.3186	0.3160	0.3134	0.3107	0.3081	0.3055	0.62
0.38	0.3055	0.3029	0.3002	0.2976	0.2950	0.2924	0.2898	0.2871	0.2845	0.2819	0.2793	0.61
0.39	0.2793	0.2767	0.2741	0.2715	0.2689	0.2663	0.2637	0.2611	0.2585	0.2559	0.2533	0.60
0.40	0.2533	0.2508	0.2482	0.2456	0.2430	0.2404	0.2378	0.2353	0.2327	0.2301	0.2275	0.59
0.41	0.2275	0.2250	0.2224	0.2198	0.2173	0.2147	0.2121	0.2096	0.2070	0.2045	0.2019	0.58
0.42	0.2019	0.1993	0.1968	0.1942	0.1917	0.1891	0.1866	0.1840	0.1815	0.1789	0.1764	0.57
0.43	0.1764	0.1738	0.1713	0.1687	0.1662	0.1637	0.1611	0.1586	0.1560	0.1535	0.1510	0.56
0.44	0.1510	0.1484	0.1459	0.1434	0.1408	0.1383	0.1358	0.1332	0.1307	0.1282	0.1257	0.55
0.45	0.1257	0.1231	0.1206	0.1181	0.1156	0.1130	0.1105	0.1080	0.1055	0.1030	0.1004	0.54
0.46	0.1004	0.0979	0.0954	0.0929	0.0904	0.0878	0.0853	0.0828	0.0803	0.0778	0.0753	0.53
0.47	0.0753	0.0728	0.0702	0.0677	0.0652	0.0627	0.0602	0.0577	0.0552	0.0527	0.0502	0.52
0.48	0.0502	0.0476	0.0451	0.0426	0.0401	0.0376	0.0351	0.0326	0.0301	0.0276	0.0251	0.51
0.49	0.0251	0.0226	0.0201	0.0175	0.0150	0.0125	0.0100	0.0075	0.0050	0.0025	0.0000	0.50
	0.010	0.009	0.008	0.007	0.006	0.005	0.004	0.003	0.002	0.001	0.000	$P(U < u)$

$C_P(K, \alpha)$			
$K$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
1	2.576	1.960	1.645
2	2.772	2.178	1.875
3	2.873	2.289	1.992
4	2.939	2.361	2.067
5	2.986	2.413	2.122
6	3.023	2.453	2.164
7	3.053	2.485	2.197
8	3.078	2.512	2.225
9	3.099	2.535	2.249
10	3.117	2.555	2.270
11	3.133	2.572	2.288
12	3.147	2.588	2.304
15	3.182	2.626	2.344
20	3.225	2.672	2.392

TABLE 3.3 – **Test de Pocock** : Les constantes  $C_P(K, \alpha)$

**ANNEXE : LES TABLEAUX STATISTIQUES**

$R_P(K, \alpha, \beta)$						
$K$	$1 - \beta = 0.8$			$1 - \beta = 0.9$		
	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
1	1.000	1.000	1.000	1.000	1.000	1.00
2	1.092	1.110	1.121	1.084	1.100	1.110
3	1.137	1.166	1.184	1.125	1.151	1.166
4	1.166	1.202	1.224	1.152	1.183	1.202
5	1.187	1.229	1.254	1.170	1.207	1.228
6	1.203	1.249	1.277	1.185	1.225	1.249
7	1.216	1.265	1.296	1.197	1.239	1.266
8	1.226	1.279	1.311	1.206	1.252	1.280
9	1.236	1.291	1.325	1.215	1.262	1.292
10	1.243	1.301	1.337	1.222	1.271	1.302
11	1.250	1.310	1.348	1.228	1.279	1.312
12	1.257	1.318	1.357	1.234	1.287	1.320
15	1.272	1.338	1.381	1.248	1.305	1.341
20	1.291	1.363	1.411	1.264	1.327	1.367

TABLE 3.4 – **Test de Pocock** : les constantes  $R_P(K, \alpha, \beta)$ .

$C_B(K, \alpha)$			
$K$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
1	2.576	1.960	1.645
2	2.580	1.977	1.678
3	2.595	2.004	1.710
4	2.609	2.024	2.733
5	2.621	2.040	2.751
6	2.631	2.053	2.765
7	2.640	2.063	2.776
8	2.648	2.072	2.786
9	2.654	2.080	2.794
10	2.660	2.087	2.801
11	2.665	2.092	2.807
12	2.670	2.098	2.813
15	2.681	2.110	2.826
20	2.695	2.126	2.842

TABLE 3.5 – **Test d'O'Brien et Fleming** : Les constantes  $C_B(K, \alpha)$ .

**ANNEXE : LES TABLEAUX STATISTIQUES**

		$R_B(K, \alpha, \beta)$					
$K$	$\alpha$	$1 - \beta = 0.8$			$1 - \beta = 0.9$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.10$
1	1.000	1.000	1.000	1.000	1.000	1.00	
2	1.001	1.008	1.016	1.001	1.007	1.014	
3	1.007	1.017	1.027	1.006	1.016	1.025	
4	1.011	1.024	1.035	1.010	1.022	1.032	
5	1.015	1.028	1.040	1.014	1.026	1.037	
6	1.017	1.032	1.044	1.016	1.030	1.041	
7	1.019	1.035	1.047	1.018	1.032	1.044	
8	1.021	1.037	1.049	1.020	1.034	1.046	
9	1.022	1.038	1.051	1.021	1.036	1.048	
10	1.024	1.040	1.53	1.022	1.037	1.049	
11	1.025	1.041	1.054	1.023	1.039	1.051	
12	1.026	1.042	1.055	1.024	1.040	1.052	
15	1.028	1.045	1.058	1.026	1.042	1.054	
20	1.030	1.047	1.061	1.029	1.045	1.057	

TABLE 3.6 – **Test d’O’Brien et Fleming** : Les constantes  $R_B(K, \alpha, \beta)$

<i>teste de Pocock</i>		Probabilité de taille d’échantillon =					
		42	84	126	168	210	
$ \mu_A - \mu_B  =$	0.0	0.016	0.012	0.009	0.007	0.956	
	0.5	0.055	0.071	0.075	0.076	0.723	
	1.0	0.214	0.268	0.210	0.138	0.171	
	1.5	0.507	0.352	0.110	0.025	0.006	
<i>teste d’O’Brien et Fleming</i>		Probabilité de taille d’échantillon =					
		36	72	108	144	180	
$ \mu_A - \mu_B  =$	0.0	0.000	0.001	0.008	0.017	0.974	
	0.5	0.000	0.015	0.078	0.134	0.773	
	1.0	0.001	0.134	0.354	0.282	0.229	
	1.5	0.010	0.472	0.415	0.089	0.013	

TABLE 3.7 – La distribution de la taille de l’échantillon de Pocock et O’Brien et Fleming pour 5 groupes d’observations

**ANNEXE : LES TABLEAUX STATISTIQUES**

$\alpha = 0.05$						
$K$	Taille de l'échantillon en pourcentage de la taille de l'échantillon fixe					
	Taille maximale de l'échantillon	Taille d'échantillon attendue à $m_A - m_B =$				
0		$\pm 0.5\delta$	$\pm \delta$	$\pm 1.5\delta$		
$1 - \beta = 0.9$						
1	100.0	100.0	100.0	100.0	100.0	100.0
2	110.0	108.4	100.9	77.6	59.2	
3	115.1	112.8	102.6	72.1	48.2	
4	118.3	115.6	104.1	69.7	43.7	
5	120.7	117.7	105.2	68.5	41.2	
10	127.1	123.4	109.0	66.6	36.7	
15	130.5	126.4	111.2	66.4	35.4	
20	132.7	128.4	112.8	66.5	34.8	

TABLE 3.8 – Test Pocock : taille d'échantillon maximale et attendue de test  $K$  groupes d'observations

$\alpha = 0.05$						
$K$	Taille de l'échantillon en pourcentage de la taille de l'échantillon fixe					
	Taille maximale de l'échantillon	Taille d'échantillon attendue à $m_A - m_B =$				
0		$\pm 0.5\delta$	$\pm \delta$	$\pm 1.5\delta$		
$1 - \beta = 0.9$						
1	100.0	100.0	100.0	100.0	100.0	100.0
2	100.7	100.5	98.2	85.1	63.3	
3	101.6	101.1	96.9	79.9	61.0	
4	102.2	101.6	96.4	76.7	57.3	
5	102.6	101.9	96.1	75.0	54.8	
10	103.7	102.8	95.6	71.8	50.8	
15	104.2	103.2	95.5	70.8	49.5	
20	104.5	103.4	95.5	70.3	48.9	

TABLE 3.9 – Tests O'Brien et Fleming : taille d'échantillon maximale et attendue de test avec  $K$  groupes d'observations

# Conclusion Générale

La mathématique est la base des sciences on la trouve dans tous les domaines, puisque grâce à elle on trouve tous les solutions des problèmes expérimentales.

Les essais cliniques se sont les preuves dans la médecine par des méthodes séquentielles.

Dans ce mémoire on simplifions le but d'analyse séquentielle est déterminer le nombre d'échantillon nécessaire pour chaque étape de l'expérience jusqu'à l'attente de l'objectif que est l'erreur de premier espèce  $\alpha$  fixée entre les traitements ou les produits médicaments médicales pour déterminer le traitement efficace contre les malades rares et très grave afin d'administré toute la population générale des patients en futur. L'utilisation de test séquentiel est diminuer la taille de l'échantillon. Puisque la distribution de la taille de l'échantillon dépend de la valeur de  $m_A - m_B$ , et il faut considérer la taille de l'échantillon sur une plage des valeurs de  $m_A - m_B$  lors de réaliser des comparaisons.

# Liste des tableaux

2.1	Décisions et probabilités . . . . .	20
3.1	Les valeurs de $\mathbf{P}(Z < z)$ en fonction de $z$ . . . . .	50
3.1	Les valeurs de $\mathbf{P}(Z < z)$ en fonction de $z$ (suite) . . . . .	51
3.2	Les fractiles de la loi normale . . . . .	52
3.2	Les fractiles de la loi normale (suite) . . . . .	53
3.3	<b>Test de Pocock</b> : Les constantes $C_P(K, \alpha)$ . . . . .	53
3.4	<b>Test de Pocock</b> : les constantes $R_P(K, \alpha, \beta)$ . . . . .	54
3.5	<b>Test d'O'Brien et Fleming</b> : Les constantes $C_B(K, \alpha)$ . . . . .	54
3.6	<b>Test d'O'Brien et Fleming</b> : Les constantes $R_B(K, \alpha, \beta)$ . . . . .	55
3.7	La distribution de la taille de l'échantillon de Pocock et O'Brien et Fleming pour 5 groupes d'observations . . . . .	55
3.8	Test Pocock : taille d'échantillon maximale et attendue de test $K$ groupes d'observations . . . . .	56
3.9	Tests O'Brien et Fleming : taille d'échantillon maximale et attendue de test avec $K$ groupes d'observations . . . . .	56

# Table des figures

2.1	Test unilatéral à droite . . . . .	18
2.2	Test unilatéral à gauche . . . . .	19
2.3	Test bilatéral . . . . .	19
2.4	Les zones de espace des paramètres . . . . .	26
2.5	Plan d'échantillonnage simple . . . . .	28
2.6	Plan tronqué . . . . .	29
2.7	Courbe d'efficacité . . . . .	29
2.8	Courbe d'efficacité idéale pour $[\theta_0, \theta_1]$ . . . . .	30
2.9	La fonction de l'effectif de moyen de l'échantillon dans le cas de plan tronqué	30
2.10	Frontières fermées . . . . .	31
2.11	Frontières ouvertes . . . . .	32
3.1	Test de Pocock avec 5 groupe d'observation . . . . .	43
3.2	Test O'Brien et Fleming pour 5 groupes d'observations . . . . .	45

# Bibliographie

- [1] *A. Wald.* (2013). Abraham Wald-A Statistician As A Key Figure For Modern Econometrics. Université de Hamburg. pp 201-241.
- [2] *B. D. BOUHANICK.* (2011-2012). Placebo Effet Placebo. pp 1-9
- [3] *C. Millier* (1967). Un Méthode Statistique : L'analyse Progressive. Annales des sciences forestières. INRA/EDP Sciences, 1967. 24(4). <10.1051/forest/19670403>. <hal-00881979>. pp.327-343.
- [4] *Cha. H. Evans, Fr, and S. T. Ildstad.* (2001). Small Clinical Trials Issues And Challenges. The National Academy of Sciences. ISBN : 0-309-07333-2. pp 12. 20. 137.
- [5] *Chr. Jennison, B. W. Turnbull.* (2000). Groupe Sequential Methods With Applications To Clinical Trails. CHAPMAN and HALL/CRC. ISBN : 0-03493-0316-8. pp 3-15. 20-35.
- [6] *D. Foata, A. Fuchs.* (1998). Calcul Des Probabilités.DUNOD. ISBN : 210 007547 0. pp 46.
- [7] *E. Hairer, G. Wanner.* (2000). L'analyse Au Fil De L'histoire. Université de Genève section de mathématique. ISBN : 3-540-67463-2. Springer. Verlag Berlin Heidelberg New York. pp 229.
- [8] *G. Saporta.* (2006). Probabilités Analyse Des Données Statistique. TECHNIP. ISBN : 2-7108-0814-5. pp 330-331.
- [9] *J. M. JOLION.* (2006). Probabilités et Statistique. Département Génie Industriel. pp I-2. II-11.
- [10] *Je. J. Aulas.* (2002). L'effet Placebo Et Ses Paradoxes. Psychiatre Et Psychopharmacologue clinicien. Science et pseudo-sciences. N 252. pp 10.
- [11] *P. A. Cornillon, E. M. Løober.* (2007). Régression Théorie Et Applications. ISBN-10 : 2-287-39692-6. ISBN-13 : 978-2-287-39692-2. Springer. Paris. Berlin. Heidelberg. New York. pp 287.
- [12] *P. Gabe.* (1969). Methodes Séquentielles En Statistique. DUNOD. ISBN : 103 141263 5. pp 4-19.

## ***CONCLUSION GENERALE***

---

- [13] *R. Veysseyre.* (2001-2006). Aide-Mémoire Statistique et Probabilités Pour L'ingénieur. DUNOD. ISBN : 2 10 0499947. pp 100-103. 255-261. 443. 453-454.
- [14] *Société Canadienne Du Cancer.*(2007). Les Essais Cliniques. Guide à l'intention des personnes atteintes du cancer. ISBN : 1 888 939 3333. pp 4-6. 12-13.

## Résumé

Dans ce travail comporte l'analyse séquentielle appliquée dans les essais cliniques. Nous étudions les tests classiques et les tests séquentielles et en particulier le test séquentiel de probabilité dans le domaine statistique.

Nous intéressons à le teste des Bocoock et le test de d'O'Brien et Fleming et pour l'erreur de 1<sup>re</sup> espèce fixée  $\alpha = 0.05$  à fin de diminuer le nombre de la taille d'échantillon dans le but de trouver l'efficacité de traitement.

## Les mots clés

Les essais cliniques, l'efficacité de traitement, les tests statistiques, les tests séquentielle, la taille d'échantillon.

## Abstract

In this work involves sequential analysis in clinical trials. We study classical tests and sequential tests and in particular the sequential probability test in statistics.

We are interested in the Bocoock test and the O'Brien and Fleming test and for the 1<sup>re</sup> species error fixed  $\alpha = 0.05$  in order to decrease the number of the sample size in the goal of finding treatment efficiency.

## Key words

Clinical trials, treatment efficiency, statistical tests, sequential testing, sample size.

## ملخص

هذا العمل يشمل التحليل المتسلسل المطبق في التجارب السريرية. نحن ندرس الاختبارات الكلاسيكية والاختبارات المتسلسلة، وبشكل خاص اختبار الاحتمال المتسلسل في مجال الإحصاء. نحن مهتمون باختبار بكووك واختبار أوبراين وفليمينغ وبالنسبة لخطأ النوع الاول الثابت  $\alpha = 0.05$  من أجل تقليل عدد حجم العينة للغرض ايجاد كفاءة العلاج.

## الكلمات المفتاحية

التجارب السريرية، كفاءة المعالجة، الاختبارات الإحصائية، الاختبارات المتسلسلة، حجم العينة.