

République Algérienne Démocratique et Populaire  
Ministère de l'enseignement Supérieur  
et de la Recherche Scientifique  
Université Mohammed Seddik Ben Yahia - Jijel



Faculté des Sciences Exactes et Informatique  
Département de Mathématiques

## Mémoire de fin d'études

Présenté pour l'obtention du diplôme de

### Master

Spécialité : Mathématiques Fondamentales

Option : Probabilités et Statistique

### Thème

**Plans prédictifs Bayésiens pour les tests  
de supériorité dans les essais  
expérimentaux.**

**Présenté par :**

Bouanani Nedjwa

**Devant le jury :**

Président : Dr Boudjerda Khawla (M.C.B) Université de Jijel

Encadreur : Dr Djeridi Zahra (M.C.B) Université de Jijel

Examinatrice : *M<sup>me</sup>* Ghouil Djoweyda (M.A.A) Université de Jijel

Promotion 2019/2020



## Remerciements

*Tout d'abord, je remercie Allah le tout puissant et le miséricordieux, pour me donner la santé, la volonté et la patience pour mener à terme mon formation de master.*

*C'est avec un grand honneur que je remercie mon enseignante et promoteur Madame Djeridi Zahra pour me conduire à la réalisation de ce travail, pour ces précieux conseils et ces encouragements.*

*Je tiens aussi à remercier les membres de jury qui acceptent à examiner mon mémoire de fin d'étude;*

*Je tiens à remercier Dr Boujarda Khawla d'avoir accepté de présider le jury de mon travail, qu'elle trouve ici toutes mes expressions respectueuses.*

*Je remercie également M<sup>me</sup> Gouil Djoweyda de m'avoir fait l'honneur de faire partie des membres du jury et d'examiner ce travail.*

*Je tiens à vous remercier.*

*Pour finir, je souhaite exprimer mon remerciements à: ma famille surtout Mes parents pour leurs intérêts envers mon travail et pour leurs aide; mes amies, pour leurs soutien moral et leurs présence dans les bons et les mauvais moments.*

❖ Nedjwa

# Table des matières

<b>Notations</b>	<b>iii</b>
<b>Introduction Générale</b>	<b>1</b>
<b>1 Statistique prédictive dans les essais clinique.</b>	<b>3</b>
1.1 Les essais clinique . . . . .	3
1.1.1 Différentes phases d'un essai clinique . . . . .	4
1.1.2 Essais de phase II . . . . .	6
1.2 Test d'hypothèses dans les essais cliniques de phase II . . . . .	8
1.2.1 Types de tests . . . . .	8
1.2.2 Puissance d'un essai et le nombre de sujets nécessaire . . . . .	11
1.2.3 Tests séquentielles . . . . .	14
1.3 Approche bayésienne . . . . .	16
1.3.1 Loi a priori et loi a posteriori . . . . .	16
1.3.2 Loi prédictive . . . . .	17
1.3.3 Exemple . . . . .	18
<b>2 Conceptions bayésiennes des essais de phase IIA.</b>	<b>22</b>
2.1 Probabilité Prédictive design . . . . .	22

2.1.1	Règle de décision . . . . .	23
2.1.2	Calcul de la probabilité prédictive . . . . .	24
2.1.3	Exemple . . . . .	24
2.1.4	Exemple . . . . .	26
2.2	Prédiction de satisfaction design . . . . .	29
2.2.1	Définition de l'indice de satisfaction . . . . .	29
2.2.2	Prédiction de satisfaction . . . . .	30
2.2.3	Procédure du PS-design . . . . .	31
2.3	Comparaison des différentes métriques bayésiennes . . . . .	35
<b>3</b>	<b>Conceptions bayésiennes des essais de phase IIB.</b>	<b>37</b>
3.1	Randomisation . . . . .	37
3.1.1	Différents types de randomisation . . . . .	38
3.1.2	Randomisation adaptative bayésienne . . . . .	38
3.2	Probabilité prédictive design pour les essais de supériorité . . . . .	39
3.3	Prédiction de satisfaction design pour les essais de supériorités . . . . .	43
	<b>Bibliographie</b>	<b>58</b>

# Notations

$C_n^x$	Coefficient binomial.
$x \sim p(x/\theta)$	$x$ est distribuée suivant la loi de densité.
$\propto$	Proportionnelle à.
$PP$	Probabilité prédictive.
$PS$	Prédiction de satisfaction.
$\pi(\theta)$	La loi a priori.
$\pi(\theta/x)$	La loi a posteriori.
$p(y/x)$	La loi prédictive.
$p - valeur$	La probabilité observer.
$1_{\{\cdot\}}$	Fonction indicatrice.
$E(N/p)$	Taille de l'échantillon espérée (prévue).
$PET(p)$	probabilité d'arrêt précoce.

# Introduction Générale

L'approche bayésienne apporte une grande souplesse à la méthodologie statistique des essais expérimentaux. Un des atouts majeurs du paradigme bayésien est la facilité avec laquelle on peut faire des prédictions sur les observations futures. En particulier, nous nous intéressons à l'utilisation de cette approche dans la prévision dans le contexte d'essais clinique en raison du rôle essentiel que joue la probabilité prédictive dans la conception et la surveillance des essais ( fiabilité des systèmes, médecine, biologie, écologie, ... ) [5].

Les procédures prédictives bayésiennes ont largement contribué à l'inférence et à l'analyse des données. Dans cette perspective, les probabilités prédictives bayésiennes constituent un outil particulièrement utile pour communiquer avec le chercheur. Ils leur donnent une méthode très attrayante pour répondre à des questions essentielles telles que : "Compte tenu des données actuelles, quelle est la probabilité que le résultat final soit en quelque sorte concluant, ou au contraire non concluant?" Cette question est inconditionnelle en ce qu'elle nécessite toutes les valeurs possibles des paramètres. les probabilités prédictives leur donnent des réponses directes et naturelles. En particulier, les procédures prédictives peuvent être utilisées pour illustrer les effets de la planification d'une expérience avec un très petit échantillon et pour aider à la décision d'abandonner précocement une expérience[5].

Les essais cliniques sont une étape obligatoire et systématique du développement d'un produit ou d'une intervention de santé. Ils permettent de préciser l'effet de cette intervention chez l'Homme, d'en déterminer l'efficacité ainsi que les éventuels effets indésirables. Le déroulement d'un essai clinique est formalisé sous forme de phases cliniques. Celles-ci ne peuvent être réalisées chez l'Homme qu'après de multiples étapes de développement pré-clinique.

Dans de nombreuses expériences ou enquêtes au cours desquelles les données sont accumulées de manière constante sur une période donnée, il est judicieux de surveiller les résultats au fur et à mesure de leurs apparitions afin de prendre des mesures telles que l'arrêt précoce ou une modification du plan d'étude. Des méthodes statistiques séquentielles ont été développées afin d'obtenir des avantages. Pour un essai avec résultat positif, un arrêt précoce signifie qu'un nouveau produit peut être exploité plus tôt. Dans le cas contraire, un arrêt précoce garantit que les ressources ne sont pas gaspillées ; Gould (1983) a qualifié cette situation "d'abandonner une cause perdue". Les méthodes séquentielles permettent généralement d'économiser la taille d'échantillon, le temps et le coût par rapport aux procédures standard d'échantillon fixe[5].

La mise en marché d'un médicament passe par plusieurs études différentes (ou essais clinique) avec chacune différents objectifs. Dans notre cas nous nous intéressons à une étude de phase II avec différents objectifs classés par ordre d'importance.

Notre but, dans ce mémoire, est de révéler cette notion dans la phase II où nous cherchons généralement à établir l'efficacité d'un nouveau traitement ou une nouvelle thérapie (phase II A), on compare deux traitements (standard et expérimentale) dans la phase IIB. On utilise l'approche prédictive car elle montre plus de flexibilité dans la conduite des essais et elles reposent sur le calcul de la probabilité prédictive ( $PP$ ), et la prédiction de satisfaction( $PS$ ). L'utilisation de cette approche est de permet au chercheur de prendre en compte l'incertitude des résultats expérimentaux qui n'ont pas été observés.

Notre recherche se fait dans le cadre Bayésien ; où nous utilisons les probabilité prédictives pour trouver les paramètres de la conception (design) de la prédiction de satisfaction dans les essais de phase IIB (chapitre 3).

Ce mémoire est divisé en trois chapitres ; le **premier chapitre** a pour objet d'introduire la méthodologie des essais clinique surtaux les essais de phase II, et nous nous sommes concentrés sur les types de tests des essais de phase II et la notion de puissance, nombre de sujet nécessaire pour les tests séquentiels et la prédiction bayésienne.

Le **deuxième** et le **troisième chapitre** sont consacré aux conceptions bayésiennes des essais de phase II A et phase IIB respectivement par le calcul de la probabilité prédictive ( $PP$ ) et la prédiction de satisfaction ( $PS$ ), en démontrons l'efficacité et la flexibilité de la prédiction dans la conception des essais de phase II.

# Statistique prédictive dans les essais clinique.

La recherche clinique est une activité médicale visant à améliorer la connaissance soit d'une maladie, soit d'une thérapeutique. La recherche clinique concerne l'être humain. En pharmacologie, Cette recherche est dominée par les études statistiques (notement la statistique inférentielle) du médicament administré à l'homme dans le cadre des essais cliniques (évaluations cliniques).

Dans ce chapitre, nous allons introduire la notion des essais clinique avec ses différentes phases surtout la mise en œuvre d'un essai clinique de phase II en utilisant les tests statistiques séquentiels. Nous explicitons aussi la prédiction bayésienne.

## 1.1 Les essais clinique

Un **essai clinique**, ou **étude clinique**, ou encore **essai thérapeutique**, est une étude scientifique réalisée en thérapeutique médicale humaine pour évaluer l'efficacité et la tolérance d'une méthode diagnostique ou d'un traitement [16]. Le comité international des rédacteurs de revue médicales [4] en donne la définition suivante : « Tout projet de recherche qui affecte de façon prospective des sujets humains à des groupes d'intervention et de comparaison afin d'étudier la relation de cause à effet entre un acte médical et l'évolution d'un état de santé ».

Les essais cliniques sont le plus souvent pratiqués en recherche médicale ou biomé-

dicale, mais d'autres disciplines cliniques, telle que la psychologie, mènent des travaux de recherche dans le but d'évaluer ou habituellement de comparer des interventions. Les recherches qui réalisent des essais cliniques poursuivent divers objectifs de différentes façons. De tels essais peuvent comprendre des questions indirectement reliées à une thérapie (rapport coût/efficacité, méthodisme des médicaments,  $\dots$ , etc ) [17].

### 1.1.1 Différentes phases d'un essai clinique

#### **Études pré-clinique :**

Les études pré-cliniques se déroulent in vitro sur des modèles cellulaires puis in vivo chez l'animal. Elles comportent des études de pharmacocinétique ( absorption, distribution, métabolisation, excrétion du principe actif et de ses métabolites ), de pharmacodynamie ( effets pharmacologiques, relation dose-effet), des études de toxicité aigüe, subaigüe puis chronique, de tératogénèse, de mutagénèse, et de cancérogénèse.

$\Rightarrow$  Sur la base de leurs objectifs, chaque phase permet de fournir des réponses spécifique sur le produit testé.

#### **Essai de phase I : Le produit est-il sûr ?**

Le produit expérimentale est testé chez un petit groupe de personnes ( 20 à 80 participants ) pour la première fois dans le but d'étudier comment il se comporte dans l'organisme humain.

L'essai de phase I a pour but de répondre aux questions suivantes :

- Quelle est la dose maximale tolérée ( études de tolérance et de toxicité ) ?
- Comment est-il assimilé dans l'organisme : administration , diffusion, métabolisme, excrétion ( étude de biodisponibilité ou de pharmacocinétiques clinique ) ?
- Quels en sont les effets secondaires dans ces conditions d'utilisation ?

Cette phase permet également d'étudier la pharmacocinétique et métabolisme chez l'homme de substance étudiée. Une étude de phase I est le préliminaire à l'étude d'efficacité d'un médicament.

#### **Essai de Phase II : Le produit a-t-il un effet bénéfique ?**

Les essais de phase II sont généralement les premières études d'un nouveau médicament

sur des patients. Ils sont habituellement réalisés sur un petit nombre de patients qui sont suivis de près.

Les études de phase II sont conçues pour découvrir si le médicament a un effet bénéfique sur la maladie en question : Elles peuvent comparer le nouveau médicament avec un traitement existant ou un placebo. Elles cherchent également à déterminer la meilleure marge posologique (un dosage sain) et à quelle fréquence le médicament devrait être administré. Elles étudient en outre la meilleure manière de gérer tout effet secondaire.

### **Essai de phase III : Le produit a-t-il un effet thérapeutique sur un plus grand effectif?**

L'essai de phase III se fait sur une grande population de volontaires atteints de la maladie ciblée.

L'objectif est de confirmer la tolérance (en surveillant l'apparition d'éventuels effets secondaires qui seraient passés inaperçus lors des phase précédentes) et l'efficacité du candide médicament (en comparant son efficacité à celle d'un placebo ou d'un traitement existant).

Le but est de déterminer dans quelles conditions l'utilisation du produit est le plus efficace.

### **Essai de phase IV : Phase de pharmacovigilance**

Ces essais sont réalisés après la mise sur le marché du médicament. Cette phase de pharmacovigilance se déroule dans les conditions habituelles d'emploi définies par l'autorisation de mise sur le marché du médicament.

Les objectifs poursuivis sont :

- Affiner la connaissance du médicament.
- Affiner la connaissance de la fréquence des effets indésirables.
- Évaluer la stratégie du traitement.
- Adapter la posologie pour des cas particuliers non pris en compte lors des essais précédents.

## 1.1.2 Essais de phase II

### ★ *Question : Pourquoi les essais de phase II ?*

L'étude de phase II des essais cliniques joue un rôle central dans le développement des médicaments. Le but de cette phase est de détecter les traitements montrant une efficacité suffisante (en vérifiant le rapport bénéfice/risque favorable) [9] pour justifier la poursuite des études [5].

La phase II ou **étude pilote** consiste à déterminer la dose optimale efficace du médicament et ses éventuels effets indésirables. Elle est subdivisée en deux phases : Les phases IIA et IIB

#### ● **Phase IIA**

Les études IIA sont généralement des phases d'expansion de dose suivant la phase I d'escalade de dose. Cette phase est non comparative et peut être organisée par cohortes. Elle consiste à tester la dose recommandée et surtout l'efficacité dans des indications précises.

- **Population éligible** : La phase IIA apprécie l'efficacité de la molécule sur un nombre limité (de 100 à 200) de patients.

#### ● **Phase IIB**

Les essais IIB s'intéressent généralement à une indication précise, afin de vérifier sur un nombre plus important de patients l'efficacité du médicament.

-**Population éligible** : La phase IIB détermine la dose thérapeutique de la molécule sur une plus grande échelle (de 100 à plus de 300 patients).

Il existe plusieurs schémas d'études pour les essais cliniques de phase II. Ces schémas peuvent être divisés en deux catégories : [10]

- **Les essais à un seul bras** : Ce sont les plus fréquemment utilisés en oncologie, contrairement aux autres aires thérapeutiques où leur utilisation est rare. Ils ont l'avantage d'être simples et rapides et de minimiser le nombre de patient à inclure. Les données obtenues sont comparées à des données historiques ce qui permet de comparer le taux d'efficacité à un taux de référence. Cependant ces données historiques peuvent parfois ne plus correspondre à la pratique clinique courante. De plus, ces essais sont très exposés aux biais de sélection, de suivi et d'évaluation. Afin d'éviter ces biais, les critères d'éligibilité

doivent être explicites, la taille de l'étude adaptée et les critères de jugement utilisés doivent être des critères de référence.

- **Les essais à plusieurs groupes et randomisés** : Deux types de schéma d'étude sont possibles : comparaison à un bras contrôle ou étude avec uniquement des bras expérimentaux :

- *Comparaison à un bras contrôle* : Cette méthode est la plus utilisée dans les autres aires thérapeutiques. L'utilisation de ce schéma est rare en oncologie car les études de phase II en oncologie portant sur des patients pour lesquels des alternatives thérapeutiques seraient encore disponibles. Dans ce cas là uniquement un bras contrôle sera estimé éthique. L'utilisation d'un placebo dans ce bras est très rare (uniquement en cas de nécessité de maintien d'aveugle) et les meilleurs soins de soutien (best supportive care) devront toujours être mis en place pour prendre en charge les patients inclus dans ce bras contrôle.

- *Étude avec plusieurs bras expérimentaux* : Ce design permet par exemple de comparer différents schémas d'administration, différentes doses ou encore différents sous-populations tumorales. Les données obtenues étant comparées à des données historiques, les mêmes problématiques que pour les essais à un seul bras seront rencontrées. Cependant, ce schéma d'étude permet tout de même d'obtenir des données plus robustes que les essais à un seul bras.

Les essais en plusieurs groupes et randomisés peuvent inclure le traitement évalué avec un groupe témoin ou plusieurs groupes de traitement évalués par rapport à des données historiques. Le schéma d'étude peut reposer sur un schéma avec une taille d'étude fixe dès le départ. Comme pour toute question de recherche, différents éléments guident la stratégie dans le choix du schéma d'étude, ainsi que le choix des critères de jugement précisés dans le protocole de l'essai précédemment [16].

Les études de phase II en oncologie peuvent être réalisées en une unique étape ou en plusieurs étapes :

- **Schéma en une étape** : (*One-stage design*) Ce schéma sera particulièrement utilisé si le critère d'évaluation est temps-dépendant. Tous les patients sont inclus, suivis et analysés. Il n'y a pas d'interruption de l'étude si l'efficacité semble insuffisante. Ces méthodes peuvent donc poser un problème éthique auquel les études de phase II en plusieurs étapes apportent une solution.

- **Schéma en plusieurs étapes** : ( $\geq$  *Two-stage design*) L'objectif est donc

de minimiser le nombre de patient exposés à un traitement inefficace. L'essai pourra être interrompu à la fin de chaque étape d'après des règles prédéfinies. Le nombre de patients inclus dans chaque étape dépendra du schéma choisi.

## 1.2 Test d'hypothèses dans les essais cliniques de phase II

Dans la majorité des schémas des essais de phase II, le critère de jugement principal est la réponse à un traitement définie comme un succès. L'efficacité est évaluée par un paramètre  $\pi$  qui correspond à la probabilité réelle du succès. Considérons les données suivantes fixées a priori :

- $K$  le nombre d'étapes du plan considéré ;
- $p_0$  la probabilité d'inefficacité maximale : par inefficacité maximale, on spécifie une probabilité de réponse  $p_0$  qui, si elle est vraie, implique clairement que le produit ne mérite plus d'être investigué. Cette probabilité correspond habituellement aux données historiques de l'efficacité du traitement de référence que l'on cherche à supplanter ;
- $p_1$  la probabilité d'efficacité minimale : par efficacité minimale, on spécifie une probabilité de réponse  $p_1$  telle que, si elle est vraie, implique que le produit a une efficacité suffisante et mérite d'être étudié.

### 1.2.1 Types de tests

À partir des définitions précédentes, on peut définir naturellement deux zones selon que le taux de succès observés  $\pi$  soit inférieur ou supérieur aux bornes  $p_0$  et  $p_1$  respectivement (FIGURE 1.1). En fait, ces zones d'inefficacité et d'efficacité correspondent aux zones de rejet des hypothèses nulle  $H_0$  et alternative  $H_1$  définies comme suit :  $H_0 : \pi \leq p_0$  versus  $H_1 : \pi \geq p_1$ .

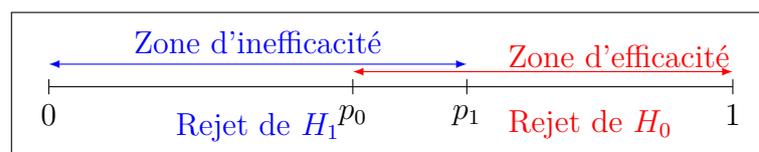


FIGURE 1.1 – Définition des zones d'efficacité et d'inefficacité.

## Définitions fondamentales pour les tests d'hypothèses

**Définition 1.2.1.** On dispose d'une variable aléatoire vectorielle  $X$  dont on sait que la loi de probabilité appartient à  $\mathcal{P} = \{\mathbb{P}_\theta, \theta \in \Omega\}$ ,  $\theta$  étant un paramètre de la loi.

**Définition 1.2.2.** On appelle fonction vraisemblance la fonction  $L : (x, \theta) \mapsto f_\theta(x)$  où  $f_\theta : x \mapsto f_\theta(x)$  est la fonction densité de la loi de probabilité  $\mathbb{P}_\theta$ .

**Remarque 1.2.1.** On emploie souvent le logarithme de la vraisemblance  $\ln \circ L : (x, \theta) \mapsto \ln(f_\theta(x))$ . En effet,  $X$  est en principe un  $n$ -uplet de variables aléatoires réelles indépendantes ; la densité de  $X$  est donc le produit de leurs densités.

L'hypothèse  $H_0$  est  $\theta \in \Omega_{H_0}$  et l'hypothèse  $H_1$  est sa négation.  $\Omega_{H_1}$  est l'ensemble des hypothèses alternatives. On note  $R_1$  et on appelle **région critique** du test l'ensemble des valeurs possibles de  $X$  pour lesquelles la décision de rejeter  $H_0$  serait prise, et  $R_0$  l'ensemble complémentaire.  $\Omega$  est partitionné en  $\Omega_{H_0}$  et  $\Omega_{H_1}$ .

**Définition 1.2.3.** On choisit un réel  $\alpha$  compris entre 0 et 1, appelé le niveau du test ou borne du risque de première espèce (seuil de signification), et on impose, si possible, que pour tout  $\theta \in \Omega_{H_0}$ ,  $\mathbb{P}_\theta\{X \in R_1\} = \mathbb{P}_\theta(R_1) \leq \alpha$ . Ou encore

$$\alpha = Pr\{\text{rejeter } H_0/H_0 \text{ vraie}\}.$$

**Définition 1.2.4.** Pour tout  $\theta \in \Omega_{H_1}$ ,  $\mathbb{P}_\theta\{X \in R_1\}$ , appelé puissance du test contre l'alternative  $\theta$ , soit maximisé du point de vue du choix de  $R_1$ . Ceci définit la fonction puissance du test  $\bar{\beta} : \mathbb{P}_\theta\{X \in R_1\} = \bar{\beta}(\theta)$  et le risque de seconde espèce en  $\theta$  C'est le risque de la démonstration de l'efficacité d'un produit réellement efficace,  $\beta(\theta) = 1 - \bar{\beta}(\theta)$ . Ou encore

$$\beta = Pr\{\text{ne pas rejeter } H_0/H_1 \text{ vraie}\}.$$

**Remarque 1.2.2.** La puissance d'un test  $1 - \beta$  est la probabilité de conclure à l'existence d'une différence entre les effets de deux traitements si cette différence existe réellement. Elle est l'aptitude d'un essai à démontrer l'efficacité ou la supériorité d'un produit donné.

Les différentes situation sont résumées dans le tableau suivant :

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

**Remarque 1.2.3.** La probabilité  $\alpha$ , (appelée aussi risque de client), est choisie a priori par l'utilisateur. Les valeurs les plus utilisées pour  $\alpha$  sont 0.05 et 0.01.

L'hypothèse  $H_0$  est une hypothèse de prudence, par exemple, pour tester l'efficacité d'un nouveau procédé ou d'un nouveau médicament, on part d'une hypothèse défavorable au nouveau produit.

**Définition 1.2.5.** L'hypothèse  $H_0$  est dite simple si  $\Omega_{H_0}$  n'a qu'un élément; sinon, on dit qu'elle est composite. De même pour  $H_1$ .

**Définition 1.2.6.** Si  $H_1$  est composite et s'il existe un test vérifiant les deux conditions ci-dessus, on dira que ce test est uniformément plus puissant (UPP, UMP en Anglais).

**Définition 1.2.7.** On dit qu'un test est non biaisé si

- pour tout,  $\theta \in \Omega_{H_0}$ ,  $\mathbb{P}_\theta\{X \in R_1\} = \mathbb{P}_\theta(R_1) \leq \alpha$
- pour tout,  $\theta \in \Omega_{H_1}$ ,  $\mathbb{P}_\theta\{X \in R_1\} = \mathbb{P}_\theta(R_1) \geq \alpha$ .

**Remarque 1.2.4.** Un test UPP est non biaisé (comparaison avec le test trivial). Un test biaisé peut, cependant, être puissant pour certaines valeurs de  $\theta$ . On appelle test UPPNB (UMPU) un test qui vérifie (2) ci-dessus seulement par rapport aux tests non biaisés.

## Des théorèmes

**lemme 1.2.1. ( Lemme fondamental de Neyman-Pearson)**

Lorsque  $H_0$  et  $H_1$  sont toutes les deux simples, notons  $\Omega_{H_0} = \{\theta_0\}$  et  $\Omega_{H_1} = \{\theta_1\}$ . Alors, quel que soit  $\alpha \in [0;1]$  il existe un UPP de niveau  $\alpha$  (UPP : de puissance maximale).

Ce test est défini par la région critique  $R_1 : R_1 = \{x, \frac{L(x, \theta_0)}{L(x, \theta_1)} \leq k\}$  où  $k$  est choisi de sorte que  $\theta \in \Omega_{H_0}$ ,  $\mathbb{P}_\theta\{X \in R_1\} = \mathbb{P}_\theta(R_1) = \alpha$ .

Les théorèmes de *Lehmann* prolongent le lemme fondamental, mais seulement lorsque  $\mathcal{P} = \{\mathbb{P}_\theta, \theta \in \Omega\}$  est une famille à vraisemblance monotone, c'est-à-dire lorsqu'il existe une statistique réelle  $T$  ( $T = g \circ X$ ) telle que  $\frac{L(x, \theta')}{L(x, \theta)}$  avec  $\theta' > \theta$ , soit une fonction croissant au sens large de  $T$ .

**Théorème 1.2.1. (Théorème de Lehmann)**

Pour  $H_0 : \theta_1 \leq \theta \leq \theta_2$  contre  $H_1 : \theta < \theta_1$  ou  $\theta > \theta_2$ ; région critique  $R^\alpha = \{x, T(x) < k_1$  ou  $T(x) > k_2\}$  où  $k_1$  et  $k_2$  sont définis par  $\mathbb{P}_\theta(T(x) < k_1) = \alpha$  et  $\mathbb{P}_\theta(T(x) > k_2) = \alpha$ ; ce test est seulement UPPSB; il n'existe pas de test UPP.

## La p-valeur

Les fréquentistes (et praticiens) ont tenté de compenser les inconvénients de l'approche de Neyman-Pearson en supprimant le niveau de signification  $\alpha$  et en proposant une réponse prenant ses valeurs dans  $[0, 1]$  et, de façon plus importante, dépendant des observations de manière plus adaptative qu'une acceptation ou un rejet établis en comparant  $T(x)$  à un seuil donné. La notion suivante a été introduite pour la première fois par Fisher (1956) [3].

**Définition 1.2.8.** *La p-value associée à un test est le niveau de signification  $\alpha$  le plus petit pour lequel l'hypothèse nulle est rejetée.*

*Une définition générale pour les hypothèses nulles ponctuelles (Thompson, 1989) est qu'une p-value est une statistique admettant une loi uniforme sous l'hypothèse nulle, se pose alors le difficile problème du choix de l'une de ces statistiques, comme d'ailleurs pour le test introduit dans la définition ci-dessus. En réalité, si un test de région critique  $R^\alpha$  est disponible pour tout niveau de signification  $\alpha$  et si ces régions sont imbriquées (c'est à dire si  $R^\alpha \subset R^\beta$  pour  $\beta > \alpha$ ), la procédure*

$$p(x) = \inf\{\alpha; x \in R^\alpha\}$$

**Exemple 1.2.1.** *Soit La région critique (qui est la région de rejet pour  $H_0$ ) est*

*$\{|x| > k\}$ , une p-value usuelle donnée par :*

$$\begin{aligned} p(x) &= \inf\{\alpha, |x| > k_\alpha\} \\ &= P^X(|X| > |x|), \quad X \sim N(0, 1) \\ &= 1 - \Phi(|x|) + \Phi(|x|) \\ &= 2[1 - \Phi(|x|)] \end{aligned}$$

*Par conséquent, si  $x = 1.68$ ,  $p(x) = 0.10$  et, si  $x = 1.96$ ,  $p(x) = 0.05$ .*

### 1.2.2 Puissance d'un essai et le nombre de sujets nécessaire

Le calcul du nombre de sujets d'un essai de phase II dépend également des risques de première espèce  $\alpha$  (le plus souvent en situation unilatérale) et de deuxième espèce  $\beta$  qui sont fixés a priori. Pour obtenir des résultats cohérents et être en mesure de convaincre d'autres investigateurs de leur validité, il faut alors définir des règles de décision basées

sur des résultats observés et attendus en réduisant les risques d'erreurs et les biais à un minimum.

Une grande partie de la méthodologie statistique utilisée dans les essais avec un seul traitement expérimental est évalué pour décider s'il a une efficacité suffisante pour justifier une étude plus approfondie. Ces essais utilisent, souvent, un résultat « réponse » binaire, à savoir le médicament "fonctionne" ou il ne le fait pas.

Dans cette exemple, nous nous concentrons sur un seul traitement expérimental.

Dans ce contexte, nous avons les hypothèses suivantes :

$H_0 : p \leq p_0$  où la vraie probabilité de réponse  $p$  est inférieur à un certain niveau  $p_0$  inintéressant.

$H_1 : p > p_1$  où la véritable probabilité de réponse est au moins une partie souhaitable  $p_1$  niveau cible.

où

$p_0$  : est le taux de la réponse indésirable.

$p_1$  : est le taux de la réponse désirable. tel que  $p_1 > p_0$ .

Dans un seul traitement expérimental un nombre prédéterminé de patients sont inscrits et l'hypothèse est testée une seule fois, à savoir, à la fin de l'essai (Schultz, 1973)[4].

Afin d'étudier l'efficacité du nouveau médicament, nous donnons une formule pour la taille d'un seul traitement expérimental. Cette formule dépend du test choisi ; il existe deux méthodes différentes :

- La première est basée sur une approximation *Normale*.
- Le second est basé sur les probabilités *Binomiales* exactes.

Les deux méthodes, utilisent le nombre de succès  $X$  comme point de départ pour trouver la taille d'échantillon.

### **L'approximation normale**

Cette procédure en une étape est basée sur la statistique  $\hat{p} = X/N$ , où  $N$  est la taille de l'échantillon et  $X$  suit une distribution *Binomiale* de paramètres  $(N; p)$ . De plus, nous supposons que pour  $N$  suffisamment grand,  $\hat{p}$  suit approximativement, une distribution *Normale* avec les paramètres  $(p; p(1 - p)/N)$  et

$$Z = \frac{(\hat{p} - p)}{se(p)} \sim Normal(0, 1); \text{ avec } se(p) = \sqrt{\frac{p(1-p)}{N}}$$

La procédure, en une seule étape, sera, alors comme suit : rejeter  $H_0$  lorsque  $Z > Z_\alpha$  ; ou de façon équivalente lorsque :

$$X > Np_0 + Z_\alpha \sqrt{Np_0(1-p_0)}$$

Il est facile de montrer que la taille de l'échantillon de cette conception avec le niveau de signification et la puissance  $(1 - \beta)$  sur  $H_1$  est approximativement égale à :

$$N_{max} = \frac{(Z_\alpha + Z_\beta)^2 \bar{p}(1-\bar{p})}{(p_1 - p_0)^2} \text{ avec } \bar{p} = \frac{(p_1 + p_0)}{2}$$

**Exemple 1.2.2.** *Supposons que nous voulons tester un traitement sur une maladie donnée et que 20% serait un taux de réponse intéressant. Alors, si la réponse est seulement de 10%, le traitement sera rejeter et on n'aura pas besoin à une étude plus approfondie.*

*Pour  $\alpha = 0.05$  (unilatéral),  $\beta = 0.20$ ,  $H_0 : p = 0.10$ , et  $H_1 : p = 0.20$ , nous constatons que le nombre de patients  $N_{max}$  est égal à 79, (Voir le programme 2 dans l'annexe I).*

### La distribution binomiale exacte

La taille de l'échantillon peut également être obtenue en utilisant la distribution Binomiale exacte de  $X$ .

Pour  $p_0$ ,  $p_1$ ,  $\alpha$  et  $\beta$  spécifiés nous déterminons la taille optimale de l'échantillon par énumération, qui peuvent être résumées dans les étapes suivantes :

- Prendre la valeur obtenue dans la section précédente  $N_{max}$ , et de varier la taille de l'échantillon  $N$  de  $(3/4) N_{max}$  à  $(5/4) N_{max}$ .
- Pour chaque  $N$ , nous déterminons le minimum de  $k$  dans  $(0, N)$  pour lesquels  $Pr(X \geq k) < \alpha$  sous l'hypothèse nulle ( $p_0$ ).
- Ensuite, prendre le minimum de  $N$  pour le quelle  $Pr(X \geq k) > 1 - \beta$  (sous  $p_1$ ).
- La taille de l'échantillon nécessaire sera, alors, le plus petit  $N$  et la valeur critique  $k$  correspondante qui satisfait les deux contraintes de probabilité :
  1.  $\alpha_k = Pr(\text{rejeter } H_0/p_0) < \alpha$
  2.  $1 - \beta_k = Pr(\text{rejeter } H_0/p_1) > 1 - \beta$

**Exemple 1.2.3.** *Si l'on calcule, maintenant, la taille de l'échantillon en utilisant les probabilités Binomiales exactes, nous voyons dans le (Table 1.1) que  $N_{max} = 79$  ( $k = 12$ )*

est le premier  $N$  dont la puissance est plus de 80%. Toutefois, pour  $N = 82$  ( $k = 13$ ) la puissance baisse et ne satisfait pas la condition requise, ce qui est un résultat de la discontinuité de la distribution Binomiale.

Tandis que  $N = 89$ , est le premier dont la puissance satisfait la condition et ne descend pas en dessous de 80%, (Voir le programme 3 dans l'annexe I).

$p_1$	$p_0$	$N_{max}$	$MinN$	$Mink$	$Alphaa$	$Puissance$
0.2	0.1	79	79	12	0.049	0.822
0.2	0.1	79	82	13	0.032	0.785
0.2	0.1	79	86	13	0.046	0.841
0.2	0.1	79	88	14	0.028	0.793
0.2	0.1	79	89	14	0.030	0.807
0.2	0.1	79	90	14	0.033	0.821

### 1.2.3 Tests séquentielles

Supposons que la loi de probabilité de la variable aléatoire  $X$  dépende d'un paramètre inconnu  $\theta$ . On peut faire sur la valeur de ce paramètre plusieurs hypothèses.

#### Principe des tests statistiques séquentiels

Supposons que deux hypothèses  $H_0$  et  $H_1$  soient à tester.

1. Le nombre d'observations n'est pas fixe,
2. On a préalablement déterminé dans l'espace des échantillons, trois sous-ensembles mutuellement exclusifs pour chaque valeur du nombre de tirages  $m$ ,

$R_m^0$  : accepter  $H_0$ ,

$R_m^1$  : accepter  $H_1$ ,

$R_m$  : continuer, c'est-à-dire procéder à un tirage supplémentaire.

On observe  $E = \{x_1, x_2, \dots, x_n\}$ ; suivant que  $E$  appartient à  $R_m^0$ ,  $R_m^1$  ou  $R_m$ , on applique les décisions qui en découlent.

#### Règle d'arrêt

D'après Robert, (2006), une règle d'arrêt  $\tau$  peut être définie comme suit : Si les expériences  $\xi_i$  produisent des observations  $x_i \in X_i$ , avec  $x_i \sim f(x_i/\theta)$ , considérons la suite

correspondante  $A_i \subset X_1 \times \cdots \times X_i$  telle que le critère  $\tau$  prend la valeur  $n$  si  $(x_1, \dots, x_n) \in A_n$ , c.à.d, l'expérience s'arrête après la  $n^{eme}$  observation seulement si les  $n$  premières observations sont dans  $A_n$ . La vraisemblance de  $(x_1, \dots, x_n)$  est alors :

$$f(x_1, \dots, x_n/\theta) = f(x_1/\theta)f(x_2/x_1, \theta) \cdots f(x_n/x_1, x_2, \dots, x_{n-1}, \theta)1_{A_n}(x_1, \dots, x_n)$$

et donc dépend seulement de  $\tau$  via l'échantillon  $(x_1, \dots, x_n)$ . Ceci implique le principe suivant.

### Principe de la règle d'arrêt

Si une suite d'expériences,  $\xi_1, \xi_2, \dots$ , admet une règle d'arrêt,  $\tau$ , qui implique quand doivent s'arrêter les expériences, l'inférence sur  $\theta$  ne doit dépendre de  $\tau$  qu'à travers l'échantillon résultant.

Quand nous comparions deux tests classiques, pour un nombre de tirages donné, nous préférons le test le plus puissant ( $1 - \beta$  maximum), celui qui minimisait le risque de deuxième espèce. Ici, nous choisirons le test qui demande un nombre moyen de tirages minimum.

### Caractéristiques d'un test séquentiel

#### Le nombre de sujet nécessaire

Le nombre de sujet nécessaires n'est plus un nombre fixe et déterminé puisque l'essai peut être arrêté après des tailles d'échantillon variables [4]. Deux notions sont couramment utilisées :

**Définition 1.2.9.** *Le nombre maximal de sujets nécessaires est le nombre de sujets qui sera inclus si toutes les analyses intermédiaires sont négatives et si l'essai n'est pas arrêté précocement.*

**Définition 1.2.10.** *Le nombre moyen (ou attendu) de sujets est le nombre de sujets qui sera inclus en moyenne avant de stopper l'essai. Ce nombre peut se calculer sous l'hypothèse nulle d'absence d'effet traitement, sous l'hypothèse alternative. c'est l'ASN =  $E(N/\theta)$  (nombre moyen d'observations à prendre en compte de l'anglais  $\ll$  average sample number  $\gg$ ).*

Pour comparer les tests séquentielle, on définit ( $PET$ ) la probabilité de terminer tôt l'essai après une analyse intermédiaires et  $E(N/\theta)$  la taille d'échantillon espérée ou la taille moyenne de l'échantillon.

Basé sur ces règles d'arrêt,  $PET$  et  $E(N)$  sous  $H_0$  peuvent être calculer par :

$PET(\theta_0) = Pr(EarlyTermination/H_0) = Pr(X_1 \leq x_1/H_0)$  où,  $X_1$  est le nombre de réponses dans  $m_1$  ( $m_1 < N_{max}$ ) patients à partir du première étape, et  $X_1 \sim Binomial(m_1; \theta_0)$

$$E(N/\theta_0) = m_1 + [1 - PET(\theta_0)] \times (N_{max} - m_1).$$

Parmi toutes les conceptions qui satisfont à la contrainte des taux d'erreur des types I et II, la conception optimale en deux étapes est obtenue lorsque  $E(N/\theta_0)$  est la plus petite. D'autre part, la conception minimax est définie lorsque la taille maximale de l'échantillon est la plus petite.

## 1.3 Approche bayésienne

L'approche bayésienne peut être présentée comme une généralisation de l'approche classique. Les paramètres ne sont plus des valeurs inconnues fixées, mais des variables aléatoires. Nous considérons que l'objet principal de la statistique est de tirer, au vu d'observations d'un phénomène aléatoire, une inférence au sujet de la loi générant ces observations; afin, soit d'analyser un phénomène passé, soit de prévoir un événement futur, en nous attachant tout particulièrement aux aspects non décisionnels de l'inférence bayésienne.

Dans cette partie, nous allons donner quelques notions de bases qui concernent la statistique bayésienne (la loi a priori, la loi a posteriori, la loi prédictive) [17].

### 1.3.1 Loi a priori et loi a posteriori

L'aspect de l'analyse bayésienne, qui est à la fois le plus délicat et le plus critiqué, est certainement celui du choix de la loi a priori des paramètres.

#### **Définition 1.3.1. (*La loi a priori*)**

*On entend par information a priori sur le paramètre  $\theta$  toute information disponible sur  $\theta$  en dehors de celle apportée par les observations.*

*L'information a priori sur  $\theta$  est entachée d'incertitude (si ce n'était pas le cas, le paramètre  $\theta$  serait connu avec certitude et on n'aurait pas à l'estimer). Il est naturel de modéliser cette information a priori a travers d'une loi de probabilité appelée loi a priori,*

notée  $\pi(\theta)$  sur  $\Theta$ . Il existe plusieurs types de loi a priori (voir [3]) :

- Loi a priori conjuguée,
- Loi a priori non-informative,
- Loi a priori impropres.

**Définition 1.3.2. ( La loi a posteriori )**

C'est la loi conditionnelle de  $\theta$  sachant  $x$ , sa densité notée  $\pi(\theta/x)$ , en vertu de la formule de Bayes, et on a aussi :

a) **La loi du couple  $(\theta, x)$**  : Sa densité est notée  $\varphi(\theta ; x)$  on a donc

$$\varphi(\theta, x) = p(x/\theta) \pi(\theta).$$

b) **La loi marginale de  $x$**  : Sa densité est notée  $m(x)$ , on a donc

$$m(x) = \int_{\Theta} p(x/\theta) \pi(\theta) d\theta.$$

La loi a posteriori est :

$$\pi(\theta/x) = \frac{p(x/\theta)\pi(\theta)}{\int_{\Theta} p(x/\theta)\pi(\theta)}$$

### 1.3.2 Loi prédictive

Dans la pratique, nous sommes intéressés par les données de l'expérience future  $F$  et on veut utiliser nos reconnaissances sur le paramètre inconnu pour prévoir les probabilités du future échantillon, et ce, est le fond de l'approche bayésienne.

Le concept des problèmes de prédiction est de trouver la distribution de probabilité d'observation  $y$  d'un échantillon future  $F$  en donnant l'échantillon  $x$  de l'expérience informative  $F$ .

**Définition 1.3.3.** Soit la fonction de densité prédictive pour une expérience futur  $F$  de fonction de densité

$$\{p(y/\theta) : \theta \in \Theta\} \text{ sur } \mathcal{Y}$$

Une expérience informative  $E$  de classe de fonction de densité

$$\{p(x/\theta) : \theta \in \Theta\} \text{ sur } \mathcal{X}$$

Et une fonction de densité  $\pi(\theta)$  sur  $\Theta$ , la fonction de densité prédictive  $p(y/x)$  pour  $F$  est définie par :

$$\begin{aligned}
p(y/x) &= \int p(y/\theta, x) \pi(\theta/x) d\theta \\
&= \frac{\int p(y/\theta) \pi(\theta) p(x/\theta) d\theta}{\int \pi(\theta) p(x/\theta) d\theta}
\end{aligned}
\tag{1.1}$$

Le schéma suivant résume le mécanisme bayésien conduisant à la densité prédictive.

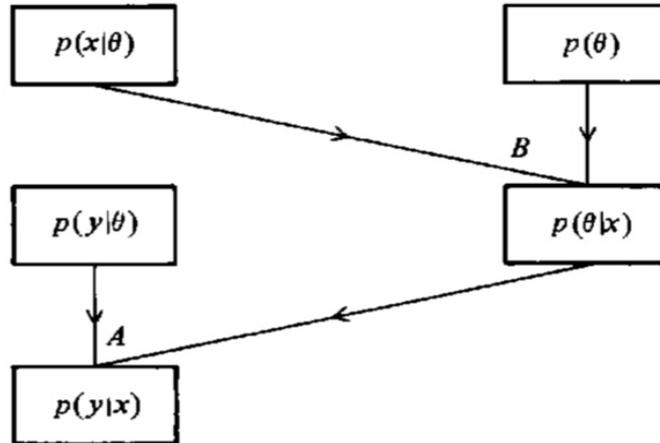


FIGURE 1.2 – Étapes de base conduisant à la densité prédictive.

### 1.3.3 Exemple

Considérons un  $n$ -échantillon de variable aléatoire  $X \sim \text{Binomial}(n; \theta)$  et pour tout  $\theta \in [0, 1]$  et tout  $x = (x_1, \dots, x_r) \in \{0, 1, \dots, n\}$

a) Sa fonction de vraisemblance est définie par :

$$p(x|\theta) = \prod_{i=0}^r p(x_i) = \left[ \prod_{i=0}^r C_n^{x_i} \right] \theta^{\sum_{i=1}^r x_i} (1 - \theta)^{rn - \sum_{i=1}^r x_i}$$

b) La loi a priori :

Dans le cas d'une loi binomial de paramètre  $(n, \theta)$  la loi a priori  $\pi$  est  $\text{Beta}(a, b)$  donnée par :

$$\pi(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1}$$

où :

$$0 < \theta < 1 \quad \text{et : } (a,b) > 0$$

c) La loi a posteriori :

$$\begin{aligned} \pi(\theta/x) &\propto p(x/\theta)\pi(\theta) \\ &\propto C_n^x \theta^x (1-\theta)^{n-x} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} \\ &\propto \theta^{x+a-1} (1-\theta)^{n-x+b-1} \\ &\propto \text{Beta}(x+a, n-x+b) \end{aligned}$$

d) Sa distribution prédictive suit une distribution *Beta-Binomial* telle que :

$$Y/(X=x) \sim \text{Beta-Binomial}(N-n, a+x, b+n-x)$$

On a pour un exemple numérique une étude publiée par **Freireich et Al** (1963) à été conçue pour évaluer les effectives de chimio thérapeutique **6-mercaptopurine** (6 - MP) pour le traitement de la leucémie aigue. Les patients ont été randomisés pour thérapie par paires. Soit  $\theta$  la proportion de la population de couples dans lesquels la 6 - MP séjour des patient en rémission plus longue que le patient du groupe placebo, (Pour distinguer une probabilité  $\theta$  à partir d'une distribution de probabilité sur  $\theta$ , nous appellerons cela une proportion de la population ou une propension). L'hypothèse nulle est  $H_0 : \theta = \frac{1}{2}$  c'est à dire pas d'efficacité de 6 - MP ; et  $H_1 : \theta \neq \frac{1}{2}$ . Il y avait 21 paires de patient dans l'étude, et 18 d'entre eux favorisaient 6 - MP. Supposons que la distribution a priori est *Uniforme* sur  $(0, 1)$ ,  $\pi(\theta) = U_{[0,1]} = \text{Beta}(1, 1)$ . La mise à jour de la distribution  $\text{Beta}(a, b)$  après les  $y$  succès et  $n - y$  échecs est facile, car la version *Beta* est une probabilité a priori conjuguée.

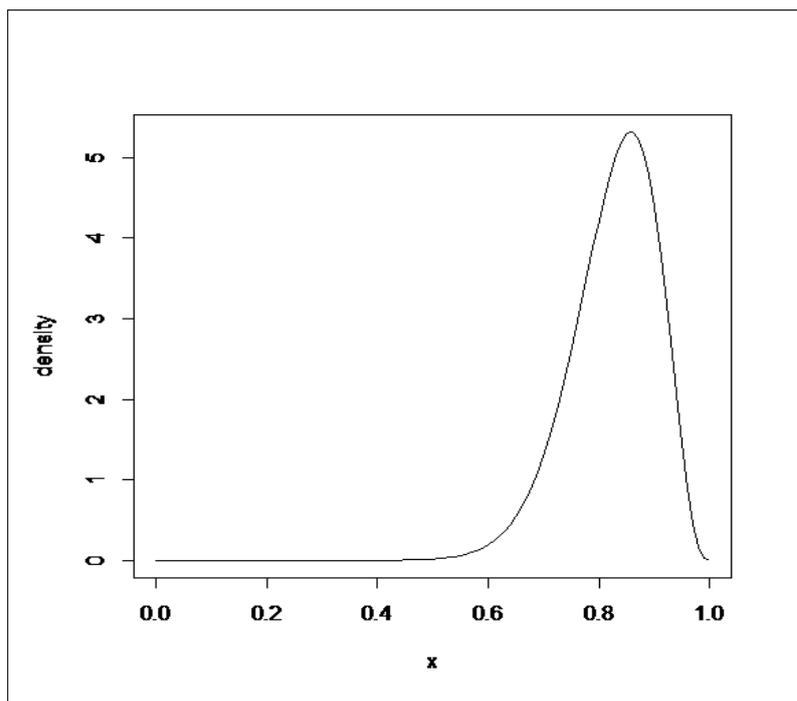


FIGURE 1.3 – La densité a posteriori de  $\theta$  est  $Beta(19, 4)$ .

Dans notre cas ( $y = 18, n - y = 3$ ), en utilisant  $b$ ) la distribution a posteriori est  $Beta(19, 4)$ ,

Remarquons que cette distribution est unimodale avec une moyenne de  $\frac{19}{(19+4)} = 0.8261$ .

Dans le contexte prédictif, on suppose qu'il est possible de prendre 5 observations supplémentaires. Comment s'en servir ?.

Une façon de répondre est d'évaluer la conséquence d'obtenir  $k$  succès au cours des 5 prochaines paires de patients (pour un total de 26 paires) pour  $k = 1, \dots, 5$ , puis de peser ces conséquences par les probabilités de prévisions des valeurs possibles de  $k$ .

Pour calculer les probabilités de futures observations, on trouve premièrement ces probabilités en supposant que les paramètres sont connus, et puis on trouve la distribution a posteriori des paramètres. Nous avons ensuite la moyenne des distributions conditionnelles à l'égard de la distribution a posteriori des paramètres. Cela donne la distribution inconditionnelle prédictive d'intérêt.

Pour illustrer ; premièrement, envisager une seule paire de patient, la paire de 22 patients, ayant un résultat binaire variable  $x_{22}$ . Supposons que cette paire est échangeable avec les 21 premières paires,  $x = (x_1, x_2, \dots, x_{21})^t$ , pour laquelle  $y = \sum_{i=1}^{21} x_i = 18$ . Un membre de la nouvelle paire est affecté à la 6 - MP, et l'autre est affecté au placebo.

La probabilité prédiction que le patient 6 - MP restera dans la remise est plus que la moyenne de la distribution a posteriori de  $\theta$ . Pour la distribution  $Beta(19, 4)$ , considérée séparément, la probabilité prédictive de succès est donnée par la règle du placebo de la succession :  $\frac{19}{(19+4)} = 0.8261$ . Par la formule (1,1) la distribution prédiction de  $x_{22}$  est :

$$g(x_{22}/x) = \int p(x_{22}/\theta)\pi(\theta/x)d\theta = \frac{\int \theta^{x_{22}}(1-\theta)^{1-x_{22}}}{\Gamma(19)\Gamma(4)}\theta^{18}(1-\theta)^3d\theta$$

Donc la chance que la 22<sup>eme</sup> observation soit un succès ( $x_{22} = 1$ ) est :

$$\begin{aligned} P(x_{22} = 1/x) &= \int p(x_{22}/\theta)\pi(\theta/x)d\theta \\ &= \frac{\Gamma(23)}{\Gamma(19)\Gamma(4)} \frac{\Gamma(20)\Gamma(4)}{\Gamma(20)} \int \frac{\Gamma(24)}{\Gamma(20)\Gamma(4)}\theta^{19}(1-\theta)^3d\theta \\ &= \frac{\Gamma(23)\Gamma(20)}{\Gamma(19)\Gamma(24)} \\ &= \frac{19}{23} = 0.8261 \end{aligned}$$

La 3<sup>eme</sup> égalité a résulté puisque l'intégrale fixée égale à 1, et la 4<sup>eme</sup> a résulté puisque  $\Gamma(z) = (z-1)\Gamma(z-1)$  pour tout  $z$ .

Maintenant supposons que deux paires additionnelles de patients (les deux sont échangeables par les 21 premières paires,  $x = (x_1, x_2, \dots, x_{21})^t$ ) sont traitées, un membre de chaque paire par 6 - MP n'est pas le carré de 0.8261, mais plutôt la probabilité que la 1<sup>ere</sup> paire soit un succès multipliée par la probabilité que la 2<sup>eme</sup> paire soit un succès sachant le résultat de la 1<sup>ere</sup>. En utilisant la formule (1.1) deux fois, nous avons :

$$P(x_{22} = 1 \wedge x_{23} = 1/x) = P(x_{22} = 1/x)P(x_{23} = 1/x_{22} = 1, x) = \frac{19}{23} \frac{20}{24} = 0.6884$$

Pour 5 paires additionnelles, la probabilité prédictive pour 4 succès pour 6 - MP est :

$$5\left(\frac{22 \times 21 \times 20 \times 19 \times 4}{27 \times 26 \times 25 \times 24 \times 23}\right) = 0.3624$$

Le prépondérant  $\ll 5 \gg$  est survenu puisque il y a 5 ordres possibles auxquels les 4 succès peuvent être arrivés(SSSSE, SSSSE,  $\dots$ , ESSSS), dont tous les événements ont la même probabilité globale.

# Conceptions bayésiennes des essais de phase IIA.

Les études de phase II peuvent être divisées en deux parties : phase IIA et Phase IIB. L'évaluation initiale de l'efficacité du médicament est le principal objectif des essais de phase IIA. En règle générale, les essais de phase IIA sont menés sous forme d'études à un seul bras évaluant l'efficacité de nouveaux médicaments, dans le but d'éliminer ceux qui ne sont pas efficaces[2].

Dans ce chapitre, nous allons présenter deux conceptions bayésienne qui utilisent la loi prédictive. Le premier est le *PP – design* due à **JJ. Lee and DD. Liu** (2008) [11], et le 2<sup>eme</sup> sera le *PS – design* due à **Z.Djeridi and H.Merabet** (2019) [6].

## 2.1 Probabilité Prédictive design

Dans cette section, nous présentons la probabilité prédictives, et nous décrivons leur utilisation dans la conception des tests cliniques de la phase IIA comme règle d'arrêt. Un avantage de cette approche est qu'elle imite le processus décisionnel clinique basé sur les données a priori. La probabilité prédictive est obtenue en calculant la probabilité d'une conclusion positive (le rejet de l'hypothèse nulle) si l'essai se déroule à la taille maximale de l'échantillon prévue. Donc, on peut prendre une décision de continuer ou d'arrêter l'essai en fonction de la puissance de cette probabilité prédictive.

Notre objectif est d'évaluer la proportion  $p$  pour un nouveau médicament en testant

les hypothèses

$$H_0 : p \leq p_0 \quad \text{contre} \quad H_1 : p \geq p_1$$

dans la phase *IIA* d'un essai.

D'après **J Jack Lee et Diane D Liu** (2008) [11], en donnant  $p_0, p_1$ , la distribution a priori du taux de réponse  $\pi(p)$  et la taille de cohorte de contrôle intérimaire  $m_i$ , nous cherchons les paramètres de design,  $y$  compris la taille de l'échantillon maximal  $N_{max}, \theta_L, \theta_T$  et  $\theta_U$  pour obtenir un design satisfaisant les contraintes de l'erreur du 1<sup>ère</sup> et du 2<sup>ème</sup> espèces simultanément.

### 2.1.1 Règle de décision

La règle de décision peut être construite comme suit :

**Algorithm 2.1.1.** (*Phase IIA basée sur le PP-design* )

**Étape1** : Si  $PP < \theta_L$ , arrêter l'essai et rejeter l'hypothèse alternative.

**Étape2** : Si  $PP > \theta_U$ , arrêter l'essai et rejeter l'hypothèse nulle.

**Étape3** : Sinon, passez à l'étape suivante jusqu'à ce qu'il atteigne  $N_{max}$  patients.

Typiquement,  $\theta_L$  et  $\theta_U$  sont deux nombres positives choisis compris entre 0 et 1, avec  $\theta_L$  soit très proche de 0 et  $\theta_U$  soit très proche de 1.

$PP < \theta_L$  signifie qu'il est improbable que le taux de succès serait supérieur à  $p_0$ , d'ici à la fin de l'essai, compte tenu des données actuelles. Dans ce cas, nous pouvons également arrêter l'essai et rejeter l'hypothèse alternative à ce moment là.

$PP > \theta_U$  , les données actuelles suggèrent que, si le même mouvement s'étend, nous aurons une probabilité élevée pour conclure que le traitement est efficace à la fin de l'étude.

Ceci résulte, et conduit à arrêter prématurément l'essai en raison d'efficacité.

En choisissant  $\theta_L > 0$  et  $\theta_U < 1$ , l'essai peut se terminer prématurément en raison soit de futilité soit d'efficacité. Pour les essai de la phase *IIA*, nous préférons souvent de choisir  $\theta_L > 0$  et  $\theta_U = 1$ , pour privilégier un arrêt prématuré en raison de futilité, mais non plus d'efficacité pour avoir un bon estimateur du taux de succès réel.

## 2.1.2 Calcul de la probabilité prédictive

Supposons que la distribution a priori de  $p$  est  $\pi(p)$ , suit à une distribution  $Beta(a, b)$ , avec la quantité  $\frac{a}{(a+b)}$  qui est la moyenne a priori, depuis les quantités  $a$  et  $b$  peuvent être considérées comme le nombre de réponses a priori efficaces et non-réponses, respectivement,  $a + b$  peut être considéré comme une mesure de précision a priori.

Supposons que nous avons fixé un nombre maximum de patients  $N_{max}$ , et supposons que  $X$  est le nombre de réponse chez les patients parmi les  $n$  patients en cours ( $n \leq N_{max}$ ) qui suit une distribution  $Binomial(n, p)$ , alors la distribution a posteriori de  $p$  serai :

$$p/x \sim Beta(a + x, b + n - x)$$

L'approche de probabilité prédictive regarde vers l'avenir sur la base des données actuelles observées pour projeter si une conclusion positive à la fin de l'étude est susceptible ou non, puis prendre une décision raisonnable à l'heure actuelle par conséquence.

Soit  $Y$  le nombre de réponses dans le potentiel de futurs patients  $m = N_{max} - n$ . Supposons que notre conception est de déclarer l'efficacité, si  $P(p > p_0/x, Y) > \theta_T$ , on sait que  $Y/X = x \sim Beta - Binomial(m, a + x, b + n - x)$ .

Quand  $Y = i$ , la distribution a posteriori est  $p/(X = x, Y = i) \sim Beta(a + x + i, b + N_{max} - x - i)$ . La probabilité prédictive ( $PP$ ) du succès de l'essai peut être alors calculée comme suit :

Soit  $B_i = P(p > p_0/x, Y = i)$  et  $1_i = 1_{\{B_i > \theta_T\}}$ , on a :

$$\begin{aligned} PP &= E[1_{\{P(p > p_0/x, Y) > \theta_T\}}/x] \\ &= \int 1_{\{P(p > p_0/x, Y) > \theta_T\}} dP(Y/x) \\ &= \sum_{i=0}^m P(Y = i/x) \times 1_{\{P(p > p_0/x, Y=i) > \theta_T\}} \\ &= \sum_{i=0}^m P(Y = i/x) \times 1_{\{B_i > \theta_T\}} \\ &= \sum_{i=0}^m P(Y = i/x) \times 1_i \end{aligned}$$

## 2.1.3 Exemple

Supposons qu'un chercheur envisage de recruter un maximum de  $N_{max} = 40$  patients dans l'étude de phase IIA. À un moment donné,  $x = 16$  réponses sont observées dans

$n = 23$  patient. Que sera la  $P(p > 60\%)$ ? Choississant une a priori vague pour  $p$ ;

$p \sim \text{Beta}(0.6, 0.4)$  et soit  $Y$  le nombre de réponse dans  $m = 17$  futures observations, donc

$$Y/X = 16 \sim \text{Beta} - \text{Beta}(17, 16.6, 7.4)$$

Pour chaque valeur possible de  $Y = i$ , l'a posteriori conditionnelle de  $p$  suit une distribution de  $\text{Beta}$ ,  $p/x$ ,

$$Y = i \sim \text{Beta}(16.6 + i, 24.4 - i)$$

Dans cet exemple, on prend  $\theta_T = 0.90$ . Comme on peut le voir dans le Tableau (2.1), lorsque  $Y = \overline{0, 11}$ , on a

$$0.0059 \leq P(p > 0.60) \leq 0.8314$$

Par conséquent nous concluons  $H_0$  pour  $Y \leq 11$ . D'autre part, lorsque  $Y = \overline{12, 17}$ ,

$$0.9022 \leq P(p > 0.60) \leq 0.9989$$

Dans ce cas, on favorises  $H_1$  pour  $y \geq 12$ . Apartir de la formule (2.1) Le calcul, donne  $PP = 0.5656$ . Si  $\theta_L = 0.10$ , l'essai ne serait pas arrêté à cause de la futilité car  $PP$  est supérieur à  $\theta_L$ , (voir le programme 1 dans l'annexe I) [2].

$Y = i$	$Pr(Y = i/x)$	$B_i = P(p > 0.6/x, Y = i)$	$1_{(B_i > 0.90)}$
0	0.0000	0.0059	0
1	0.0000	0.0120	0
2	0.0001	0.0263	0
3	0.0006	0.0526	0
4	0.0021	0.0963	0
5	0.0058	0.1624	0
6	0.0135	0.2529	0
7	0.0276	0.3651	0
8	0.0497	0.4910	0
9	0.0794	0.6186	0
10	0.1129	0.7354	0
11	0.1426	0.8314	0
12	0.1587	0.9022	1
13	0.1532	0.9488	1
14	0.1246	0.9760	1
15	0.0811	0.9901	1
16	0.0381	0.9964	1
17	0.0099	0.9989	1

TABLE 2.1 – Calcul de la  $PP$

### 2.1.4 Exemple

(Un essai sur le cancer du poumon)

L'objectif principal de cette étude est d'évaluer l'efficacité d'une thérapie de combinaison en tant que le traitement de première ligne chez les patients atteints d'un cancer du poumon avancé. L'étude implique la combinaison d'un anticorps de facteur de croissance endothéliale vasculaire, plus un facteur de croissance épidermique de la tyrosine kinase de récepteur inhibiteur. Le critère principal est le taux de réponse clinique (c-à-d., le taux de réponse complète et de réponse partielle combinée) au nouveau régime.

Le traitement standard courant rapporte un taux de réponse approximativement de 20% ( $p_0$ ). Le taux de réponse cible du nouveau régime est 40% ( $p_1$ ).

Prenant le *PP – design*, nous supposons que le taux de réponse  $p \sim \text{Beta}(0.2, 0.8)$ . L'essai est surveillé continûment après avoir évalué les réponses de 10 premiers patients. Pour chaque  $N_{max}$  entre 25 et 50, nous cherchons l'espace  $\theta_L$  et  $\theta_T$  pour générer des modèles qui satisfont à la fois, le  $\alpha$  et  $\beta$  moins de 0.10. Présentés dans le tableau (2.2) les résultats sont dans l'ordre de la taille maximale de l'échantillon,  $N_{max}$ . Parmi tous ces designs, le design avec  $N_{max} = 36$  (en gras) est le design avec le plus petit  $N_{max}$  qui satisfait à la fois  $\alpha$  et  $\beta$  contrôlés sous 0.1.

Sur la base de ce paramètre,  $\theta_L$  et  $\theta_T$  sont déterminés à 0.001 et  $[0.852, 0.922]$ , respectivement. Les régions de rejet correspondantes (en nombre de réponse /  $n$ ) sont 0/10, 1/17, 2/21, 3/24, 4/27, 5/29, 6/31, 7/33, 8/34, 9/35, et 10/36. L'essai sera arrêté et le traitement est considéré comme inefficace lorsque le nombre de réponses sont dans la région de rejet. Basé sur ces frontières, si le taux de réponse réel est de 20%, la probabilité d'accepter le traitement est 0.088. Par contre, si le taux de réponse réel est 40%, la probabilité d'accepter le traitement est 0.906. La probabilité d'arrêter l'essai précocement est 0.86 et la taille de l'échantillon espérée est 27.67 lorsque le vrai taux de réponse est 20%.

**Table 2.1.** Les caractéristique de fonctionnement du *PP* – design avec  $\alpha$  et  $\beta \leq 0.10$ ,  $p \sim \text{Beta}(0.2, 0.8)$ ,  $p_0 = 0.2$ ,  $et p_1 = 0.4$ . les intervalles dans la deuxième colonne indique que  $\theta_T$  peut prendre n'importe quel valeur dans l'intervalle fermé donné.

$\theta_L$	$\theta_T$	$r/N_{max}$	$PET(p_0)$	$E(N/p_0)$	$\alpha$	$\beta$
		<i>NA/35</i>			0.126	0.093
		<i>NA/35</i>			0.074	0.116
<b>0.001</b>	<b>[0.852, 0.922]</b>	<b>10/36</b>	<b>0.86</b>	<b>27.67</b>	<b>0.088</b>	<b>0.094</b>
0.011	[0.830, 0.908]	10/37	0.85	25.13	0.099	0.084
0.001	[0.876, 0.935]	11/39	0.88	29.24	0.073	0.092
0.001	[0.857, 0.923]	11/40	0.86	30.23	0.086	0.075
0.003	[0.837, 0.910]	11/41	0.85	30.27	0.100	0.062
0.043	[0.816, 0.895]	11/42	0.86	23.56	0.099	0.083
0.001	[0.880, 0.935]	12/43	0.88	32.13	0.072	0.074
0.001	[0.862, 0.924]	12/44	0.87	33.71	0.085	0.059
0.001	[0.844, 0.912]	12/45	0.85	34.69	0.098	0.048
0.032	[0.824, 0.898]	12/46	0.86	26.22	0.098	0.068
0.001	[0.884, 0.936]	13/47	0.89	35.25	0.071	0.058
0.001	[0.868, 0.925]	13/48	0.87	36.43	0.083	0.047
0.001	[0.850, 0.914]	13/49	0.86	37.86	0.095	0.038
0.020	[0.832, 0.901]	13/50	0.86	30.60	0.100	0.046

La conception *PP* permet une surveillance continue de l'essai. Sous la *PP* approche, l'essai peut être arrêté lorsqu'il n'y a pas de réponses dans le premier 10 – 16 patients, 1 réponse chez les 17 – 20 premiers patients, 2 réponse dans le 21 – 23 premiers patients, et ainsi de suite. Ainsi, la conception *PP* permet une plus grande flexibilité et une surveillance fréquente.

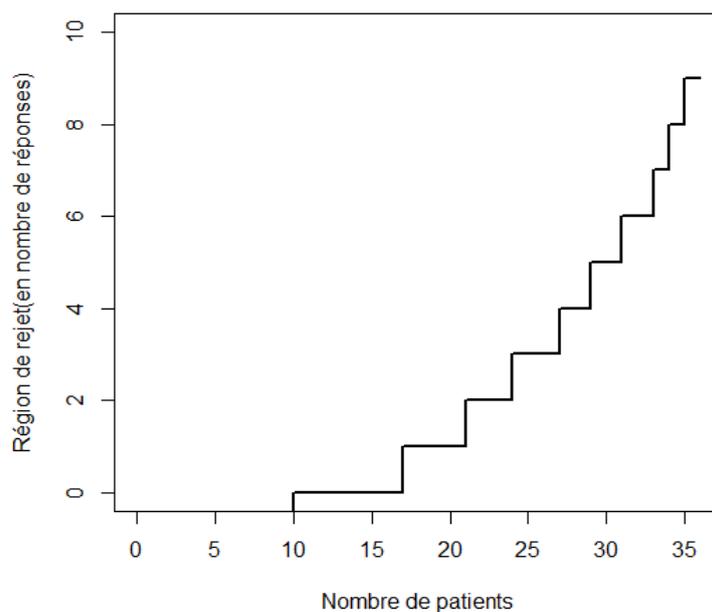


FIGURE 2.1 – Régions d’arrêt pour la probabilité prédictive design.

De plus, par rapport aux conception en deux étapes, il est plus difficile d’arrêter un essai sous la conception  $PP$  chez 17 ou 19 patients. Par exemple, avec 0, 1, 2, ou 3 réponses chez 17 patients, l’essai d’être arrêté seulement avec 0 ou 1 réponses à ce stade. La conception  $PP$  permet à l’essai de être arrêté à tout moment si l’accumulation de preuves ne permet pas de l’efficacité du nouveau traitement par rapport à la thérapie standard.

## 2.2 Prédiction de satisfaction design

L’indice de satisfaction trouve son origine dans les situations où le statisticien qui effectue un test « souhaite » détecter un effet significatif ; c’est-à-dire, rejeter l’hypothèse nulle  $H_0$ .

### 2.2.1 Définition de l’indice de satisfaction

Sachant que le contexte expérimental consiste en deux expérience successives avec des résultats  $x \in X$  et  $y \in Y$  (indépendants). On suppose que les distributions des deux résultats dépendent d’un seul paramètre  $\theta \in \Theta$ . Considérons  $z = (x, y)$ , qui est utilisé pour

établir la conclusion finale de l'étude et déterminer la satisfaction de l'utilisateur, qui est indiquée par  $\Phi(z)$  [12]. Cependant, sur la base du résultat  $x$  de l'essai clinique de première étape, il est utile d'anticiper ce que la satisfaction sera bien après la deuxième étape. Dans notre étude, cette prédiction est réalisée dans un contexte bayésien ; c'est-à-dire, basé sur le choix d'une probabilité à priori sur  $\Theta$ .

Avec  $\alpha$  fixe, soit un test d'erreur du 1<sup>ère</sup> espèce  $\alpha$  être défini par une région critique  $R^{(\alpha)}$ . Dans ce cas, l'indice de satisfaction a été défini par **Merabet et al. (2017)** comme suit :

$$\phi(z) = \begin{cases} 0 & \text{si : } z \notin R^{(\alpha)} \\ 1 - p(z) & \text{si : } z \in R^{(\alpha)} \end{cases} \quad (2.1)$$

où  $p(z) = \inf\{\beta, z \in R^{(\alpha)}\}$  est la p-valeur du test.

## 2.2.2 Prédiction de satisfaction

L'utilisation de l'approche prédictive permet au chercheur de prendre en compte l'incertitude du résultats expérimentaux qui n'ont pas été observés. Il est plus intéressant d'anticiper le degré de satisfaction par la prédiction de la satisfaction où l'on doit mener une expérience en deux étapes :

- Un premier résultat  $x$  détermine si nous continuons ou non l'expérience.
- Si l'expérimentateur est très satisfait et que nous continuons efficacement l'expérience, alors le résultat  $z$  des premier et deuxième étapes doit être la base du test.

Soit  $p(z/x)$  : la probabilité prédictive de  $z$  conditionnellement sur  $x$ , le  $PS$  (la prédiction de satisfaction) est alors donné par :

$$\pi(x) = E(\phi(z)/x) = \int \phi(z)p(z/x)dz \quad (2.2)$$

On constate que  $\pi(x)$  généralise la puissance du test dans la logique de l'indice de satisfaction proposée.

Cette procédure est généralisée pour les  $K$  étapes. On peut évaluer les chances que l'essai montre un résultat concluant à la fin de l'étude compte tenu des informations actuelle. Ensuite, la décision pour suivre ou l'arrêt l'essai peut se faire en fonction de la force du prédiction de satisfaction.

Dans ce cadre, un nombre maximum de groupe  $K$  et une taille de groupe de  $m_k$  sont choisis les données accumulées sont analysées après chaque groupe de réponses  $m_k$ . Pour un test avec un seuil  $\alpha$ , où l'hypothèse nulle est de type  $p \leq p_0$ , on calcule l'indice de satisfaction et sa prédiction. Pour les essais cliniques la distribution a priori du paramètre inconnu est une a priori non informative dans une séquence (cette procédure est vraie même dans le cas d'une a priori conjuguée ) [6].

### 2.2.3 Procédure du PS-design

Supposons que toutes les variables aléatoires soient indépendants  $X_{l_1}^1 (1 \leq l_1 \leq m_1), \dots, X_{l_k}^k (1 \leq l_k \leq m_k)$  d'une distribution de *Bernoulli*( $\theta$ ) où  $\theta$  est inconnu. Soit  $x_k = \sum_{l_k=1}^{m_k} X_{l_k}^k$  (pour :  $k = 1, \dots, K$ ) qui suit une distribution *Binomiale*( $m_k, \theta$ ) représentant le nombre de résultats positifs.

Au début, nous pouvons choisir une loi a priori *Beta*(1, 1) pour  $\theta$ , ce choix apparaît raisonnable en pratique, car souvent aucune information n'est disponible sur les paramètres, alors cette technique bayésienne spécifie l'état d'ignorance. L'essai est surveillé en permanence afin que la probabilité a posteriori bayésienne est mise à jour après avoir observer chaque nouveau résultat. Les décisions sont prises de manière adaptative tout au long de l'essai jusqu'à ce que la taille maximale de l'échantillon soit  $N_{max} = \sum_{k=1}^K m_k$ . L'approche adaptée ici a été formulée en terme de détermination de la prédiction de satisfaction lors d'une analyse intermédiaire ( $k^{ème}$  étape).

L'algorithme suivant suppose que le critère de rejet de l'hypothèse nulle à l'analyse intermédiaire est que la prédiction de satisfaction est un scalaire compris entre 0.5 et 1 qui a été choisie par les enquêteurs de l'étude [11].

1. Après la première étape :

La loi a priori est :  $\theta \sim \text{Beta}(a, b)$

La loi a posteriori :  $\theta/x_1 \sim \text{Beta}(x_1 + a, m_1 - x_1 + b)$

Nous voulons planifier un deuxième échantillon de taille  $N_{max} - m_1$  du résultat  $y$ . La distribution prédictive de la deuxième étape étant donné le premier résultat  $x_1$  est :

$$p(y/x_1) = \frac{C_{N_{max}-m_1}^y \beta(y + x_1 + a, N_{max} - (y + x_1) + b)}{\beta(x_1 + a, m_1 - x_1 + b)}$$

où :

$$\beta(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \text{ et } C_{N_{max}-m_1}^y = \frac{x!}{y!(x-y)!}$$

Après la deuxième étape, posons  $z = (x_1, y)$  la région de rejet est définie par :

$R^{(\alpha)} = \{z, z \geq q_2\}$  avec :  $q_2 = \inf\{u, Pr(Z \geq u/\theta_0) \leq \alpha\}$ , et l'indice de satisfaction sera :

$$\phi(z) = \Phi_{\theta_0}^B(z)I_{z > q_2}$$

où :  $\Phi_{\theta_0}^B$  est la fonction de distribution cumulée d'une *Binomiale*( $N_{max}, \theta_0$ ).

La prédiction de satisfaction sera :

$$\begin{aligned} \pi(x_1) &= \sum_{y=(q_2-x_1)^+}^{N_{max}-m_1} \phi(x_1 + y)p(y/x_1) \\ &= \sum_{y=(q_2-x_1)^+}^{N_{max}-m_1} \Phi_{\theta_0}^B(x_1 + y) \frac{C_{N_{max}-m_1}^y \beta(y + x_1 + a, N_{max} - (y + x_1) + b)}{\beta(x_1 + a, m_1 - x_1 + b)} \end{aligned}$$

où :  $u^+ = \max(0, u)$ .

Si la prédiction de satisfaction  $\pi(x_1)$  est inférieure au seuil spécifié  $\gamma$ , l'étude est arrêté et le médicament est abandonné, sinon l'essai se poursuit.

2. Supposons que nous décidions de pour suivre l'essai afin d'obtenir deux observation  $x_1$  et  $x_2$  dans la deuxième analyse intermédiaire, par conséquent nous devons planifier un troisième échantillon  $Y$  de taille  $N_{max} - m_1 - m_2$  ( en supposant que la prédiction de satisfaction est un nombre entre 0.5 et 1 pour inscrire d'autres patients dans l'étude ).

Par conséquent, la distribution a priori sera mise à jour à la lumière des analyses intermédiaires ; c'est-à-dire, que l'a priori sera :

$$\theta/x_1 \sim Beta(x_1 + a, m_1 - x_1 + b)$$

La posteriori est la distribution de  $\theta/x_1, x_2$  qui est :

$$\pi(\theta/x_1, x_2) \propto Beta(x_1 + x_2 + a, m_1 + m_2 - (x_1 + x_2) + b)$$

La distribution prédictive de la troisième étape étant donné le premier résultat  $x_1$  et le deuxième  $x_2$  est :

$$\begin{aligned} p(y/x_1, x_2) &= \int f(y/\theta) f(\theta/x_1, x_2) d\theta \\ &= \frac{C_{N_{max}-\sum_{i=1}^2 m_i}^y \beta\left(y + \sum_{i=1}^2 x_i + a, N_{max} - y - \left(\sum_{i=1}^2 x_i\right) + b\right)}{\beta\left(\sum_{i=1}^2 x_i + a, \sum_{i=1}^2 m_i - \left(\sum_{i=1}^2 x_i\right) + b\right)} \end{aligned}$$

L'indice de satisfaction après la troisième étape ( pour :  $z = (x_1, x_2, y)$  ) est :

$$\phi(z) = \Phi_{\theta_0}^B(z)I_{z > q_3}$$

où :

$$q_3 = \inf\{u, Pr(Z \geq u/\theta_0) \leq \alpha\}$$

Enfin, la prédiction de satisfaction pour la troisième étape sera :

$$\begin{aligned} \pi(x_1, x_2) &= \sum_{y=(q_3 - \sum_{i=1}^2 x_i)^+}^{N_{max} - \sum_{i=1}^2 m_i} \phi(x_2 + x_1 + y)p(y/x_1, x_2) \\ &= \sum_{y=(q_3 - \sum_{i=1}^2 x_i)^+}^{N_{max} - \sum_{i=1}^2 m_i} \Phi_{\theta_0}^B\left(\sum_{i=1}^2 x_i + y\right) \frac{C^y \beta\left(y + \sum_{i=1}^2 x_i + a, N_{max} - (y + \sum_{i=1}^2 x_i) + b\right)}{\beta\left(\sum_{i=1}^2 x_i + a, \sum_{i=1}^2 m_i - \left(\sum_{i=1}^2 x_i\right) + b\right)} \end{aligned}$$

3. Pour une analyse intermédiaire  $k - 1; k = 2, \dots, K$ , nous avons des données intermédiaires  $x_1, \dots, x_{k-1}$ , la même procédure peut être réalisée si la prédiction de satisfaction  $\pi(x_1, \dots, x_{k-2})$  est dans l'intervalle  $[0.5, 1]$ .

La loi a priori de cette étape est choisi comme la loi a posteriori de la dernière étape, qui est

$$\pi(\theta/x_1, \dots, x_{k-2}) \propto \text{Beta}\left(\sum_{i=1}^{k-2} x_i + a, \sum_{i=1}^{k-2} m_i - \sum_{i=1}^{k-2} x_i + b\right)$$

La distribution a posteriori de  $\theta/x_1, \dots, x_{k-1}$

$$\begin{aligned} \pi(\theta/x_1, \dots, x_{k-1}) &\propto p(x_{k-1}/\theta)p(x_{k-2}/\theta) \cdots p(x_2/\theta)p(x_1/\theta)\pi(\theta) \\ &\propto \text{Beta}\left(\sum_{i=1}^{k-1} x_i + a, \sum_{i=1}^{k-1} m_i - \sum_{i=1}^{k-1} x_i + b\right) \end{aligned}$$

La distribution prédictive de la  $k^{\text{ème}}$  étape étant donne les résultats  $x_1, \dots, x_{k-1}$  est :

$$\begin{aligned} p(y/x_1, \dots, x_{k-1}) &= \int p(y/\theta)\pi(\theta/x_1, \dots, x_k)d\theta \\ &= \frac{C^y \beta\left(y + \sum_{i=1}^{k-1} x_i + a, N_{max} - (y + \sum_{i=1}^{k-1} x_i) + b\right)}{\beta\left(\sum_{i=1}^{k-1} x_i + a, \sum_{i=1}^{k-1} m_i - \left(\sum_{i=1}^{k-1} x_i\right) + b\right)} \end{aligned}$$

Pour la  $k^{\text{ème}}$  étape avec  $z = (x_1, \dots, x_{k-1}, y)$  et la région critique  $R^\alpha = \{z, z \geq q_k\}$ ,

et avec  $q_k = \inf\{u, Pr(Z \geq u/\theta_0) \leq \alpha\}$ , l'indice de satisfaction sera :

$$\phi(z) = \Phi_{\theta_0}^B(z)I_{z > q_k}$$

Enfin, la prédiction de satisfaction pour la  $k^{\text{ème}}$  étape sera :

$$\begin{aligned} \pi(x_1, \dots, x_{k-1}) &= \sum_{y=(q_k - \sum_{i=1}^{k-1} x_i)^+}^{N_{max} - \sum_{i=1}^{k-1} m_i} \phi\left(\sum_{i=1}^{k-1} x_i + y\right) p(y/x_1, \dots, x_{k-1}) \\ &= \sum_{y=(q_k - \sum_{i=1}^{k-1} x_i)^+}^{N_{max} - \sum_{i=1}^{k-1} m_i} \Phi_{\theta_0}^B\left(\sum_{i=1}^{k-1} x_i + y\right) \frac{C^y \beta\left(y + \sum_{i=1}^{k-1} x_i + a, N_{max} - (y + \sum_{i=1}^{k-1} x_i) + b\right)}{\beta\left(\sum_{i=1}^{k-1} x_i + a, \sum_{i=1}^{k-1} m_i - \left(\sum_{i=1}^{k-1} x_i\right) + b\right)} \end{aligned}$$

Dans l'analyse finale, il faut prendre une décision, c'est-à-dire à la fin de l'essai, il n'y a pas de zone grise où le nombre de réponses pourrait se situer.

**Exemple 2.2.1.** Reprenons l'exemple 2.1.1 (*un essai sur le cancer du poumon*)

Pour les PS designs, deux scénarios ont été pris en compte : Dans le premier (SC1), nous supposons que le taux de réponse  $\theta$  suit a priori de Jeffreys Beta(1/2; 1/2). L'essai est surveillé continument après l'évaluation des réponses du 10 premiers patients. Selon les valeurs spécifiées de  $\theta_0$ ,  $\theta_1$ ,  $\alpha$  et  $\beta$ , les caractéristiques de la conception optimale sont les suivantes :

$m_1 = 24; x_1 = 8; N_{max} = 37; x_1 + x_2 = 11; PET(\theta_0) = 0.9638$  et  $E(N/\theta_0) = 24.47$ ; avec  $\alpha = 0.0892$  et  $\beta = 0.0960$ .

Avec une prédiction de satisfaction 0.5263, le design minimax correspondant donne :

$m_1 = 24; x_1 = 7; N_{max} = 35; x_1 + x_2 = 10; PET(\theta_0) = 0.9108$  et  $E(N/\theta_0) = 24.98$ ; avec  $\alpha = 0.0892$  et  $\beta = 0.0960$ .

Avec une prédiction de satisfaction, dans ce cas, 0.5998.

En appliquant la même méthode, nous pouvons trouvé les caractéristiques des PS designs pour le deuxième scénario (SC2) lorsque le taux de réponse  $\theta$  suit une a priori uniforme Beta(1; 1). Les résultats sont présentés dans la **Table 2.3** (voir l'annexe I le programme 4).

**Table 2.2.** *Caractéristiques de fonctionnement des conceptions en deux étapes de PS avec les taux d'erreur de type I et de type II  $\leq 0.10$ ,  $\theta_0 = 0.2$ ,  $\theta_1 = 0.4$ , a priori pour  $\theta \sim \text{Beta}(0.5 ; 0.5)$  (SC1), a priori pour  $\text{Beta}(1 ; 1)$  (SC2)*

	$x_1/m_1$	$(x_1 + x_2)/N_{max}$	$PET(\theta_0)$	$E(N/\theta_0)$	$\alpha$	$\beta$	$PS$
<i>PS-designs</i>							
<i>Optimal(SC1)</i>	8/24	11/37	0.9638	24.47	0.0892	0.0960	0.53
<i>Minimax (SC1)</i>	7/24	10/35	0.9108	24.98	0.0892	0.0960	0.60
<i>Optimal(SC2)</i>	7/24	10/34	0.9638	24.47	0.0892	0.0960	0.53
<i>Minimax (SC2)</i>	7/24	10/34	0.9108	24.89	0.0892	0.0960	0.60

## 2.3 Comparaison des différentes métriques bayésiennes

Dans cette section, nous comparons les différentes mesures bayésiennes utilisées pour le suivi intermédiaire, telles que : La prédiction de satisfaction, les probabilités a posteriori bayésiennes et la probabilité prédictive car le choix de la règle d'arrêt est très important dans l'analyse intermédiaire.

**Exemple 2.3.1.** *Supposons que l'essai soit conçu avec la planification de quatre analyses intermédiaires pour futilité ; elles sont effectuées lorsque les données sont disponibles pour 20, 50, 75 et 90 patients, respectivement. Supposons que lors de la première analyse intermédiaire, 12 ont été observées sur 20 patients (60%) ( $p$ -valeur unilatérale = 0.25), tandis que 48 réponses ou plus sont nécessaires chez les 80 patients restants afin de considérer la réussite de l'essai. Avec une a priori uniforme, la probabilité a posteriori d'intérêt est*

$$Pr(\theta > 0.50 | x_1 = 12, m_1 = 20) = 0.82$$

*La distribution prédictive a posteriori bayésienne des observations futures  $y_1 \sim \text{Beta} - \text{Binomial}(80, 13, 9)$ , et la probabilité prédictive de succès est égale à 0.54, ce qui est la probabilité d'observer 48 réponses ou plus chez les 80 patient restants les données observées. Supposons que les deuxième, troisième et quatrième analyses intermédiaires donnent 28 succès/ 50 (56%), 41/75 (55%) et 49/90 (54%), avec des probabilités posteriori de 0.81, 0.79 et 0.80, respectivement. Compte tenu de ces résumés presque identiques de preuves pour bénéfice du traitement, il n'est pas évident de savoir si l'essai doit être arrêté ou poursuivi après chacune des analyses intermédiaires. En revanche, les probabilités prédictives et les prédictions de satisfaction varie considérablement selon les évaluations intermédiaires, comme le montre **Table 2.4**.*

Puisque le *PS* est lié au taux d'erreur de type I, nous avons comparé le *PP* et le *PS* en deux cas :  $\alpha = 0.05$  et  $\alpha = 0.1$ .

**Table 2.3.** Comparaison entre différentes métriques bayésiennes :  $\sum_{l=1}^k m_l$  et  $\sum_{l=1}^k x_l$  : le nombre de patients et succès lors de l'analyse intermédiaire  $k$  ;  $N_{max} - \sum_{l=1}^k m_l$  : nombre de patients restants à l'analyse intermédiaire  $k$  ;  $y_i^*$  : nombre minimum de succès requis pour réussir.

$\sum_{l=1}^k m_l$	$\sum_{l=1}^k x_l$	$N_{max} - \sum_{l=1}^k m_l$	$y_i^*$	$P(\theta > 0.5 x)$	<i>PP</i>	<i>PS</i> ( $\alpha = 0.05$ )	<i>PS</i> ( $\alpha = 0.1$ )
20	12	80	48	0.81	0.62	0.58	0.64
50	28	50	31	0.80	0.45	0.37	0.51
75	41	25	18	0.79	0.26	0.16	0.36
90	49	10	10	0.80	0.10	0.03	0.25

À mesure que le nombre des analyses intermédiaires augmente, la prédiction de satisfaction lorsque  $\alpha = 0.1$  diminue légèrement par rapport à la prédiction de satisfaction lorsque  $\alpha = 0.05$  et à la *PP*.

Les critères d'arrêt spécifiques sont généralement propres à chaque essai et incluent des considérations commerciales, telles que les considérations risques-avantages des patients, disponibles ressources, coût d'opportunité et puissance statistique globale.

Si la futilité est définie comme un essai ayant peu de chances d'atteindre son objectif, alors il s'agit d'un problème de prédiction ; la meilleure solution consiste à prédire la satisfaction ou la probabilité prédictive [14].

# Conceptions bayésiennes des essais de phase IIB.

Les essais de phase IIB sont des études multi-bras pour comparer l'efficacité du nouveau médicament par rapport au traitement standard ou d'autres médicaments expérimentaux, afin que le plus prometteur puisse être sélectionné pour une évaluation à grande échelle dans les études de phase tardive. Le problème de toxicité du nouveaux agents peuvent également être évalués de plus dans des études de phase II [2].

Dans ce chapitre, nous allons donner deux conceptions en utilisant les probabilités prédictives des données non observés qui sont le PP-design et le PS-design. Mais il faut tout d'abord indiquer la notion de randomisation pour l'affectation des patients dans les différents bras d'étude.

## 3.1 Randomisation

Un essai clinique tente de discriminer les effets de différents traitements rapidement, de sorte qu'un plus grand nombre de patients en dehors de l'essai bénéficieraient. Non seulement la randomisation peut réduire efficacement les biais et les effets de confusion, mais il constitue également la base fondamentale de l'analyse statistique [8].

### 3.1.1 Différents types de randomisation

En général, la randomisation peut être classée comme suit :

- La randomisation simple est le moyen le plus simple d'attribution des patients, qui fixe les probabilités de randomisation mais peut conduire à des déséquilibrés groupes.
- La randomisation par blocs permutés garantit le nombre de sujets dans chaque groupe pour être le même pendant l'essais.
- La randomisation stratifiée stratifie d'abord les participants par certains facteurs pronostiques, puis les répartit dans chaque strate. Donc au moins, le nombre de patients pour ces facteurs de stratification peut être équilibré.
- La randomisation adaptative permet à la probabilité d'allocation pour chaque bras de changer à mesure que l'essai progresse. Il peut être adaptatif aux covariables pour équilibrer les facteurs pronostiques de base ou adaptatifs aux résultats pour affecter plus de patients aux meilleurs bras de traitement.

### 3.1.2 Randomisation adaptative bayésienne

Pour une comparaison plus objective des différents traitements, les essais de phase II avec plusieurs bras sont souvent randomisé. Les patients peuvent être répartis au hasard chaque bras de traitement avec une probabilité fixe tout au long de l'essai (par exemple, la randomisation égal attribue les patients à chaque bras avec la même probabilité), ou, la probabilité de randomisation peut changer de manière adaptative en fonction des données pendant l'essai. La randomisation adaptative (AR) basée sur la réponse ou les résultats tend à affecter plus de participants à l'essai de meilleurs traitements, car chaque nouveau patient a une probabilité plus élevée de recevoir le traitement le plus efficace basé sur les données recueillis dans l'essai en cours. Dans ces deux designs bayésiens considérons seulement le cas de la randomisation adaptative simple.

## 3.2 Probabilité prédictive design pour les essais de supériorité

Supposons que nous comparions deux traitements dans un essai randomisé de phase II à deux bras  $l = 1, 2$ .

Soit  $p_l$  le taux de réponse du traitement  $l$ , et effectuons à  $p_l$  une distribution a priori  $Beta(a_l, b_l)$  pour  $l = 1, 2$ . Si parmi  $n_l$  sujets traités dans le bras  $l$ , nous observons  $x_l$  réponses, alors :  $X_l \sim Binomial(n_l, p_l)$  et la distribution a posteriori du  $p_l$  est :

$$p_l / (X_l = x_l) \sim Beta(a_l + x_l, b_l + n_l - x_l).$$

Si la taille maximale de l'échantillon dans le bras  $l$  est  $N_l$  alors le nombre de réponses chez les futurs patients  $N_l - n_l$  suit une distribution  $Beta - Binomial$  :

$$Y_l / (X_l = x_l) \sim Beta - Binomial(N_l - n_l, a_l + x_l, b_l + n_l - x_l).$$

Lorsque  $Y_l = y_l$ , la distribution a posteriori du taux de réponse compte tenu du courant et du future est :

$$p_l / (X_l = x_l, Y_l = y_l) \sim Beta(a_l + x_l + y_l, b_l + n_l - x_l - y_l).$$

Nous spécifions une différence de traitement cliniquement significative  $\delta$  et une probabilité seuil  $\theta_T$  si :

$$P(|p_2 - p_1| > \delta / X_1 = x_1, X_2 = x_2, Y_1 = y_1, Y_2 = y_2) \geq \theta_T.$$

Nous affirmons la non-équivalence des deux traitements, c'est-à-dire qu'un traitement est supérieur à l'autre. Cependant,  $Y_1$  et  $Y_2$  sont les données futures, qui n'ont toujours pas été observées au stade actuel de la prise de décision. Nous pouvons faire la moyenne du caractère dans  $Y_1$  et  $Y_2$  en calculant le  $PP$  comme suit :

$$\begin{aligned} PP &= E [I \{P(|p_2 - p_1| > \delta / X_1 = x_1, X_2 = x_2, Y_1, Y_2 \geq \theta_T)\}] \\ &= \sum_{Y_1=0}^{N_1-n_1} \sum_{Y_2=0}^{N_2-n_2} P(Y_1 = y_1 / X_1 = x_1) P(Y_2 = y_2 / X_2 = x_2) \\ &\quad \times I_{\{P(|p_2 - p_1| > \delta / X_1 = x_1, X_2 = x_2, Y_1, Y_2 \geq \theta_T)\}} \end{aligned}$$

où :  $I_{\{\cdot\}}$  est la fonction d'indicateur.

$PP$  : désigne la probabilité prédictive pour revendiquer la supériorité à la fin

d'essais.

Selon les travaux de **JJ Lee and DD Liu** (2008) [11], nous devons spécifier les probabilité d'arrêt inférieure et supérieure pour la prise de décision adaptative dans la conduite de l'essai. Les règles de décision basées sur la  $PP$  sont les suivantes :

1. **Arrêt d'équivalence** : *Si  $PP < \theta_L$ , alors nous arrêtons l'essais et acceptons l'hypothèse nulle pour revendiquer l'équivalence de traitement.*
2. **Arrêt du supériorité** : *Si  $PP < \theta_U$ , alors nous arrêtons l'essai et rejetons l'hypothèse nulle pour revendiquer un bras de traitement supérieur.*

Nous pouvons maintenir les taux d'erreur fréquentistes de type  $I$  et de type  $II$ . Les paramètres de conception  $(N_{max}, \delta, \theta_T, \theta_L, \theta_U)$ , où  $N_{max}$  la taille maximale de l'échantillon de l'essai  $N_{max} = N_1 + N_2$ .

Une conception d'essai basée sur la  $PP$  permet une surveillance continue. Si les deux traitements ont des effets d'efficacité similaires ou si un traitement est largement meilleur que l'autre, l'essai peut être arrêté tôt lorsque des preuves suffisantes se sont accumulées. Il en résulterait une taille d'échantillon attendue plus petite, et donc un essai plus efficace. À la fin de l'essai, nous déclarons soit qu'un traitement est meilleur que l'autre, soit l'équivalence de deux traitements.

**Exemple 3.2.1.** *Supposons que nous comparions deux traitements en formulant une hypothèse du test fréquentiste comme suite :*

$$H_0 : p_1 = p_2 \text{ contre } H_0 : p_1 \neq p_2$$

*pour chaque paire de données futures  $(Y_1 = y_1, Y_2 = y_2)$ , nous pouvons tirer une conclusion de savoir si le test d'hypothèse donnerait une différence de traitement significative en la fin l'essais. En résumant tous les résultats futurs possibles  $(x_1, x_2)$ , la probabilité prédictive de rejeter  $H_0$  (c'est-à-dire, qu'il y a une différence de traitement significative) est donnée par :*

*Pr (une différence significative à la fin de l'essai /données)*

$$= \sum_{y_1=0}^{N_1-n_1} \sum_{y_2=0}^{N_2-n_2} P(y_1/x_1)P(y_2/x_2)I_{\{\text{rejetant } H_0\}}$$

*où la fonction d'indicateur  $I_{\{\cdot\}}$  caractérise si le test binomial habituel pour deux proportions est significatif. Sur la base des données actuelles et futures, l'estimation du taux de réponse pour le bras  $l$  est*

$$\hat{p} = \frac{y_l + x_l}{N_l}, \quad l = 1, 2$$

Sous l'hypothèse nulle  $H_0 : p_1 = p_2 = p$ , les échantillons regroupés sur les deux bras produisent une estimation de  $p$  comme suit :

$$\hat{p} = \frac{y_1 + x_1 + y_2 + x_2}{N_1 + N_2}$$

et

$$\text{Var}(\hat{p}_1 - \hat{p}_2) = \frac{p_1(1-p_1)}{N_1} + \frac{p_2(1-p_2)}{N_2} = p(1-p)\left(\frac{1}{N_1} + \frac{1}{N_2}\right).$$

En utilisant une approximation normale, la statistique fréquentiste à deux échantillons pour tester deux proportions est donnée par :

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/N_1 + 1/N_2)}},$$

qui suit asymptotiquement la distribution normale standard sous l'hypothèse nulle. Nous rejetons l'hypothèse nulle si  $|Z| \geq z_{\alpha/2}$ , où  $z_{\alpha/2}$  est le  $100(1 - \alpha/2)$  centile de la distribution normale standard. Compte tenu l'intérimaires de données, la probabilité prédictive de rejeter l'hypothèse nulle à la fin de l'épreuve est :

$$\sum_{y_1=0}^{N_1-n_1} \sum_{y_2=0}^{N_2-n_2} P(y_1/x_1)P(y_2/x_2)I_{\{|Z| \geq z_\alpha\}}.$$

Le programme de surveillance intérimaire susmentionné associe la prédiction bayésienne avec une procédure de test d'hypothèse fréquentiste. D'autre part, nous pouvons mettre en œuvre une procédure de surveillance intermédiaire entièrement bayésienne en utilisant la probabilité prédictive. Compte tenu des données actuelles  $(x_1, x_2)$  et des données futures  $(y_1, y_2)$ , nous calculons la probabilité prédictive,

$$\text{Pr}(p_1 > p_2/y_1, y_2, x_1, x_2) = \int_0^1 \int_{p_2}^1 f(p_2/y_2, x_2) f(p_1/y_1, x_1) dp_1 dp_2,$$

où  $f(p_l/y_l, x_l)$  est la fonction de densité de probabilité de  $p_l$  avec la distribution  $p_l \sim \text{Beta}(\alpha_l + y_l + x_l, \beta_l + N_l - y_l - x_l)$ , pour  $l = 1, 2$ . Traitement 1 est réputé supérieur au traitement 2 si

$$Pr(p_1 > p_2/y_1, y_2, x_1, x_2) \geq \theta_T \quad (3.1)$$

où la probabilité de coupure  $\theta_T$  prend généralement une valeur comprise entre 0.85 et 0.95, en fonction du degré certitude dont nous disposons pour revendiquer la supériorité du traitement 1 basé sur (3.1). Cependant, les données futures  $(y_1, y_2)$  n'ont pas encore été observées. Ainsi, compte tenu des données observées  $(x_1, x_2)$ , la probabilité prédictive de réclamer le traitement 1 supérieur au traitement 2 à la fin de l'essai est  $Pr(\text{revendiquant la supériorité du traitement 1/données})$

$$= \sum_{y_1=0}^{N_1-n_1} \sum_{y_2=0}^{N_2-n_2} P(y_1/x_1)P(y_2/x_2)I_{\{Pr(p_1 > p_2/y_1, y_2, x_1, x_2) \geq \theta_T\}}$$

Nous montrons dans cette exemple comment utiliser la probabilité prédictive pour le suivi des essais dans un essai randomisé à deux bras. La taille d'échantillon prévue pour chaque bras de traitement est le même; c'est-à-dire  $N_1 = N_2 = N_{max}/2$ . Pour le test d'hypothèse fréquentiste  $H_0 : p_1 = p_2$  vs  $p_1 \neq p_2$ , nous utilisons un test binomial bilatéral au niveau de signification de  $\alpha = 0.05$ . Pour la méthode bayésienne, nous prenons les distributions a priori pour  $p_1$  et  $p_2$  comme  $Beta(0.2, 0.8)$  et définissons la probabilité de coupure  $\theta_T = 0.95$ .

Dans le tableau suivant, nous présentons la taille totale de l'échantillon  $N_{max}$ ; le nombre de réponses sur le nombre actuel de sujets dans le bras  $l$ ,  $x_l/n_l$ ; et les probabilités prédictives en faveur du 1<sup>er</sup> bras ou du 2<sup>ème</sup> bras utilisant les approches fréquentistes [8] et bayésiennes, respectivement. A titre d'illustration, nous interprétons les résultats avec  $N_{max} = 40$  dans la première rangée. Nous avons prévu d'inscrire 20 sujets dans chaque bras pour comparer les deux traitements. Après 10 patients traités dans chaque bras, 5 patients ont répondu parmi les 10 patients de 1<sup>er</sup> bras et 2 ont répondu dans la 2<sup>ème</sup> bras. À ce stade, la probabilité prédictive de privilégier le traitement 1 à l'issue de l'essai est d'environ 51% en utilisant le test fréquentiste à deux échantillons, et 67% en utilisant l'approche entièrement bayésienne. De la ligne 1 à la ligne 4, il y a une augmentation de la probabilité prédictive de favoriser le bras 1 lorsque  $N_{max}$  augmente, ce qui atteint rapidement un plateau. En d'autres termes, s'il y a plus de données à collecter dans l'avenir, la différence de traitement actuellement observée pourrait être améliorée.

**Illustration de la surveillance intermédiaire avec la Probabilité prédictive dans un essai de phase IIB randomisé à deux bras.**

$N_{max}$	<i>bras1</i> <i>bras2</i>		$Pr(p_1 > p_2)$	
	$x_1/n_1$	$x_2/n_2$	Frequentist	Bayesian
40	5/10	2/10	0.5062	0.6702
60	5/10	2/10	0.6266	0.7225
80	5/10	2/10	0.6915	0.7567
100	5/10	2/10	0.7291	0.7815
100	10/20	4/20	0.8415	0.8999
100	15/30	6/30	0.9306	0.9735
100	20/40	8/40	0.9910	0.9993
100	10/20	8/20	0.2167	0.2821
100	10/20	9/20	0.1157	0.1573

### 3.3 Prédiction de satisfaction design pour les essais de supériorités

Dans cette section, on utilise une prédiction séquentielle bayésienne de l'indice de satisfaction. Pour une analyse intermédiaire qui traite des hypothèses de prédiction, telles que la surveillance de la futilité avec des résultats retardés, la prédiction de satisfaction tient correctement compte de la quantité de données restant à observé dans un essai clinique et a la flexibilité d'incorporer des informations via des variables auxiliaires [6].

Sachant que le contexte expérimental consiste en deux expérience successives (deux étapes), avec des résultats  $x = (x_1, x_2) \in X$  et  $y = (y_1, y_2)$  qui sont indépendants.

$$\text{telle que : } X_1 \sim \text{Binomiale}(n_1, p_1) \quad , 1 \leq n_1 \leq N_1.$$

$$X_2 \sim \text{Binomiale}(n_2, p_2) \quad , 1 \leq n_2 \leq N_2.$$

où

$X_1$  : Le nombre de réponse pour traitement 1.

$X_2$  : Le nombre de réponse pour traitement 2.

$n_l$  : Le nombre des sujets traités dans le bras  $l(l = 1, 2)$ .

Et après la deuxième étape, posons  $z = (z_1, z_2)$  telle que  $z_1 = (x_1, y_1)$  et  $z_2 = (x_2, y_2)$

Alors  $Z \sim \text{Poisson-Binomiale}(\underbrace{p_1, p_1, \dots, p_1}_{n_1 \text{ fois}}, \underbrace{p_1, p_1, \dots, p_1}_{N_1 - n_1 \text{ fois}}, \underbrace{p_2, p_2, \dots, p_2}_{n_2 \text{ fois}}, \underbrace{p_2, p_2, \dots, p_2}_{N_2 - n_2 \text{ fois}})$

(voir l'annex II)

On test les hypothèses suivante :

$$\begin{cases} H_0 : |p_1 - p_2| \leq \delta \\ H_1 : |p_1 - p_2| > \delta \end{cases}$$

telle que  $p_1$  est le taux de réponse pour le traitement 1, et  $p_2$  le taux de réponse pour le traitement 2.  $0 < \delta < 1$  est l'incrément minimale acceptable du taux de réponses pour le médicament expérimental par rapport au médicament standard.

On peut récrire les hypothèses comme suit :

$$\begin{aligned} \begin{cases} H_0 : |p_1 - p_2| \leq \delta \\ H_1 : |p_1 - p_2| > \delta \end{cases} &\Leftrightarrow \begin{cases} H_0 : -\delta \leq p_1 - p_2 \leq \delta \\ H_1 : p_1 - p_2 > \delta \text{ ou } p_1 - p_2 < -\delta \end{cases} \\ &\Leftrightarrow \begin{cases} H_0 : p_2 - \delta \leq p_1 \leq p_2 + \delta \\ H_1 : p_1 < p_2 - \delta \text{ ou } p_1 > p_2 + \delta \end{cases} \end{aligned}$$

D'après la **théorème 1.3.1 de Lehmann**(le chapitre 1) la région critique est donnée par :

$$R^{(\alpha)} = \{z, T(z) < q_1\} \cup \{T(z) > q_2\} \text{ telle que } T(z) = z_1 + z_2.$$

Et  $q_1$  et  $q_2$  vérifie :

$$P_{p_2 - \delta}(R^{(\alpha)}) = \frac{\alpha}{2} \quad \text{et} \quad P_{p_2 + \delta}(R^{(\alpha)}) = \frac{\alpha}{2}.$$

donc :

$$\begin{cases} q_1 = \inf\{u; Pr_{p_2 - \delta}(T(z) \leq u | (p_1, p_2)) \leq \frac{\alpha}{2}\} \\ q_2 = \inf\{u; Pr_{p_2 + \delta}(T(z) \geq u | (p_1, p_2)) \leq \frac{\alpha}{2}\} \end{cases}$$

**Remarque 3.3.1.** Dans le cas d'un test bilatérale la  $p$ -valeur peut être définie par :

$$p(z) = 2 \times \min\{P_{H_0}(T(z) < q_1), P_{H_0}(T(z) > q_2)\}$$

avec  $q_1$  et  $q_2$  vérifier :

$$\begin{cases} P_{H_0}(T(z) < q_1) = \frac{\alpha}{2} \\ P_{H_0}(T(z) > q_2) = \frac{\alpha}{2} \end{cases}$$

Alors l'indice de satisfaction est donné par :

$$\phi(z) = \begin{cases} 0 & \text{si } z \notin R^{(\alpha)} \\ 1 - p(z) & \text{si } z \in R^{(\alpha)} \end{cases}$$

**Après la 1<sup>ère</sup> étape :**

On planifie deux échantillons de taille  $N_1 - n_1$  du résultat  $y_1$  après  $x_1$ , et  $N_2 - n_2$  du résultat  $y_2$  après  $x_2$  respectivement, telle que la taille maximale de l'échantillon dans le 1<sup>er</sup> bras est  $N_1$ , et la taille maximale de l'échantillon dans le 2<sup>ème</sup> bras est  $N_2$ .

Dans ce cas :

La loi a priori est :  $p_l \sim \text{Beta}(a_l, b_l)$ ,  $l = 1, 2$ .

La distribution à posteriori du  $p_l$  ( $l = 1, 2$ ) est :

$$p_l | (X = x_l) \sim \text{Beta}(a_l + x_l, b_l + n_l - x_l).$$

Et on a  $y_1 \sim \text{Binomiale}(N_1 - n_1, p_1)$  et  $y_2 \sim \text{Binomiale}(N_2 - n_2, p_2)$ .

Soit  $y_l$  le nombre de réponse parmi les  $N_l - n_l$  sujets restants dans le  $l^{\text{ème}}$  bras, le nombre de réponses chez les futurs patients  $N_l - n_l$  ( $l = 1, 2$ ) suit une *Beta-Binomiale* pour le 1<sup>ère</sup> résultat. Donc la distribution prédictive :

$$p(y_1/x_1) = \frac{C_{N_1-n_1}^{y_1} \beta(y_1 + x_1 + a_1, N_1 - (y_1 + x_1) + b_1)}{\beta(x_1 + a_1, n_1 - x_1 + b_1)}.$$

$$p(y_2/x_2) = \frac{C_{N_2-n_2}^{y_2} \beta(y_2 + x_2 + a_2, N_2 - (y_2 + x_2) + b_2)}{\beta(x_2 + a_2, n_2 - x_2 + b_2)}.$$

Et l'indice de satisfaction sera :

$$\phi(z) = \Phi_{p_1, p_2}^{PB}(z) I_{R^{(\alpha)}}$$

où :

$\Phi_{p_1, p_2}^{PB}$  est la fonction de distribution cumulée d'une

Poisson-Binomiale( $\underbrace{p_1, p_1, \dots, p_1}_{n_1 \text{ fois}}, \underbrace{p_1, p_1, \dots, p_1}_{N_1 - n_1 \text{ fois}}, \underbrace{p_2, p_2, \dots, p_2}_{n_2 \text{ fois}}, \underbrace{p_2, p_2, \dots, p_2}_{N_2 - n_2 \text{ fois}}$ ).

La prédiction de satisfaction sera :

$$\begin{aligned} \pi(x_1, x_2) &= \sum_{y_1=(x_1-q_1)^+}^{N_1-n_1} \sum_{y_2=(q_2-x_2)^+}^{N_2-n_2} \phi(x_1 + y_1 + x_2 + y_2) p(y_1 \setminus x_1) p(y_2 \setminus x_2) \\ &= \sum_{y_1=(x_1-q_1)^+}^{N_1-n_1} \sum_{y_2=(q_2-x_2)^+}^{N_2-n_2} \Phi_{p_1, p_2}^{PB}(x_1 + y_1 + x_2 + y_2) \\ &\quad \times \frac{C_{N_1-n_1}^{y_1} \beta(y_1 + x_1 + a_1, N_1 - (y_1 + x_1) + b_1)}{\beta(x_1 + a_1, n_1 - x_1 + b_1)} \times \frac{C_{N_2-n_2}^{y_2} \beta(y_2 + x_2 + a_2, N_2 - (y_2 + x_2) + b_2)}{\beta(x_2 + a_2, n_2 - x_2 + b_2)}. \end{aligned}$$

où :  $u^+ = \max(0, u)$ .

Si la prédiction de satisfaction  $\pi(x_1, x_2)$  est inférieure au seuil spécifié  $\gamma$  (entre 0.5 et 1) , l'étude est arrêté et nous acceptons l'hypothèse nulle pour revendiquer un bras de traitement supérieur sinon on poursuit l'étude et collectons plus de données.

# Conclusion Générale

La probabilité a posteriori bayésienne, qui correspond à la probabilité que le paramètre soit contenu dans une région significative, est le meilleur outil pour répondre à la question traitée par la surveillance intermédiaire suivante : « Existe-t-il des preuves convaincantes en faveur de l'hypothèse nulle ou l'alternative? ». D'autre part, en utilisant des méthodes de stochastique curtalement telles que la probabilité prédictive et la prédiction de satisfaction, nous donnons des réponses à la question : « L'essai est-il susceptible de présenter des preuves convaincantes en faveur de l'hypothèse alternative si l'on recueille des données supplémentaires? ». Parce que nous traitons ici de la prévision des preuves qui seront disponibles aux étapes ultérieures de l'étude.

Afin d'utiliser au mieux les ressources et de sélectionner avec soin les traitements candidats prometteurs pour un essai de phase III, il existe un besoin croissant de plans d'essais randomisés de phase II telle que les essais de supériorité. L'utilisation de la probabilité prédictive et la prédiction de satisfaction pour guider la conception de l'essai de phase II est attrayante pour les chercheurs cliniques. Il est souhaitable de mettre fin à un essai si la preuve cumulative est suffisamment solide pour tirer une conclusion définitive au milieu du déroulement de l'essai. L'ajout de la randomisation adaptative améliore encore l'éthique individuelle de l'essai clinique en attribuant plus de patients à des traitements plus efficaces, et cela se traduit par une augmentation de la réponse globale de l'essai. Les conceptions qui évaluent les réponses à court terme, telles que les résultats binaires, sont idéales pour l'application de la prédiction bayésienne, qui peut être mise en œuvre presque en temps réel. Nous avons proposé deux approches différentes ; le *PP – design* et le *PS – design*, pour démontrer l'efficacité et la flexibilité de l'approche bayésienne dans la conception et l'analyse des essais cliniques, qui conduisent toutes deux à des caractéristiques de fonctionnement d'essais essentiellement intéressantes. Cependant, plusieurs

paramètres de conception peuvent être calibrés pour répondre aux objectifs de différentes conceptions et pour cela on peut distinguer les caractéristiques suivantes des deux designs :

- Le *PP – design* n’est pas lié à l’erreur du type *I* tandis que le *PS – design* l’est.
- Dans la procédure du *PP – design*, on intègre les patients un à un mais on fait l’analyse intermédiaire, mais dans le *PS – design* on intègre par cohorte (groupe).
- Dans la procédure de *PS – design*, à chaque analyse intermédiaire on considère qu’on refait l’expérience à deux étapes ce qui assure que l’erreur du deuxième espèce ne dépasse pas le seuil  $\alpha$ .
- Pour les deux designs, le choix de l’a priori initiale joue un rôle très important dans la conception de l’essai.
- Les deux designs respectent les caractéristiques fréquentistes (les erreurs du type *I* et du type *II*).

Pour une perspective de ce travail, on suggère la simulation de l’essai de supériorité avec la prédiction de satisfaction aussi l’extension de cette procédure avec la randomisation adaptative bayésienne.

# Résumé

Dans notre travail nous avons étudié l'importance de l'approche bayésienne dans le concept des essais cliniques en utilisant la prédiction notamment la probabilité prédictive ( $PP$ ) et la prédiction de satisfaction ( $PS$ ) pour prendre une meilleure décision de d'efficacité/futilité d'un traitement. Nous avons aussi appliquer la procédure d'indice de satisfaction pour les tests de supériorité dans les essais cliniques de phase IIB.

# Abstract

In this work we have studied the importance of the Bayesian approach in the concept of clinical trials by using prediction including predictive probability and the prediction of satisfaction make a better decision of efficacy/futility of a treatment. We also applied the satisfaction index procedure for superiority testing in phase IIB clinical trials.

# Annexe I

## Programme 1

```
> # Pour  $Pr(y = i/x)$ 
>  $p0 = 0.6; n = 23; Nmax = 40; m = (Nmax - n); x = 16; a = 0.6; b = 0.4$ 
>  $ab = c(x + a, n - x + b)$ 
>  $ys = 0 : m$ 
>  $pred = pbetap(ab, m, )$ 
>  $dist = round(cbind(ys, pred), 4)$ 
> # Pour  $Bi = Pr(p > p0/x, y = 0)$ 
>  $f = 1 - (pbeta(p0, a + x + 0, b + Nmax - x - 0))$ 
>  $fdis = round(cbind(1, f), 4)$ 
> # Pour  $Bi = Pr(p > p0/x, y = i)$ 
>  $i = 1 : m; p0 = 0.6; n = 23; Nmax = 40; m = (Nmax - n); x = 16; a = 0.6; b = 0.4$ 
>  $g = numeric(length(m + 1))$ 
>  $for(i in 1 : m)\{g[i] = 1 - (pbeta(p0, a + x + i, b + Nmax - x - i)); g[i]\}$ 
>  $dis = round(cbind(1 : m, g), 4)$ 
>  $Bi = c(fdis[, 2], dis[, 2])$ 
>  $pred = c(dist[, 2])$ 
>  $mat1 = cbind(ys, pred, Bi)$ 
>  $mat1$ 
> # Pour la probabilité prédictive
>  $theta.t = 0.90$ 
>  $k = NA$ 
>  $for(i in 1 : (m + 1))\{if(mat1[i, 3] > theta.t)\{k[i] = 1\} else \{k[i] = 0\}\}$ 
>  $mat2 = cbind(mat1, k)$ 
>  $colnames(mat2) = c("i", "Pr(Y = i/x)", "Bi", "1(Bi > 0.90)")$ 
>  $mat2$ 
>  $PP = 0$ 
>  $for(i in 1 : (m + 1))\{if(mat2[i, 4] == 1)\{PP = PP + mat2[i, 2]\}\}$ 
>  $PP$ 
[1]0.5655589
```

## Programme 2

```
>  $\alpha = 0.05, \beta = 0.20; p0 = 0.10; p1 = 0.20$ 
```

```

> zalpha = qnorm(1 - alpha); zalpha; zbeta = qnorm(1 - beta)
> zbeta; pbar = (p0 + p1)/2; pbar
[1] 1.644854
[1] 0.8416212
[1] 0.15
> Nmax = (zalpha + zbeta)^2 * (pbar * (1 - pbar))/(p1 - p0)^2
> Nmax
[1] 78.8276

```

### Programme 3 (beta = puissance)

```

> alpha = 0.05; beta = 0.80; p0 = 0.10; p1 = 0.20
> zalpha = qnorm(1-alpha); zalpha; zbeta = qnorm(beta); zbeta; pbar = (p0+p1)/2; pbar
[1] 1.644854
[1] 0.8416212
[1] 0.15
> Nmax = (zalpha + zbeta)^2 * (pbar * (1 - pbar)) = (p1 - p0)^2
> Nmax
[1] 78.8276
> N1 = round((3/4) * Nmax); N1; N2 = round((5/4) * Nmax); N2
[1] 59
[1] 99
> N = 79
> for(kin0 : N)
{alphak [k] = 1 - pbinom(k; N; p0); betak [k] = 1 - pbinom(k; N; p1)}
> for(s in 1 : N)
{
if(alphak [s] < alpha)
{alphaa [s] = alphak [s]} else {alphaa [s] = 0}
if(betak [s] > beta)
{betaa [s] = betak [s]}else{betaa [s] = 0}
}

```

### Programme 4

```

> N = 100; p0 = 0.5; p1 = 0.65; alpha = 0.05;
> q0 = qbinom(1 - alpha, N, p0)

```

```

> q0
> preject2p0=NULL
> preject2p1=NULL
> zz = 0
> for(m1in5 : (N - 5))
{
> p = numeric(m1)
> m2 = N - m1
> for(r1 in 1 : m1)
{
ab = c(r1 + 0.5, m1 - r1 + 0.5)
ys = 1 : m2
pred = pbetap(ab, m2, ys)
if((q0 - r1) > 0)x = ys[ys >= (q0 - r1)] else 0
pi = NULL
for(i in 1 : length(x))pi[i] = pbinom(x[i], m2, p0) * pred[x[i]]
p[r1] = sum(pi)
}
> print("m1 = "); print(m1);
> print("p = "); print(p);
> xr = min(which(p >= 0.5))
> print("xr = "); print(xr)
> print("p[xr]"); print(p[xr])
> # probabilité de rejeter H0 après la 2ème étape
> ipreject2p0 = 1;
> for(xrr in xr : m1)
{
> preject2p0 = 0
> for(kin(xrr + 1) : q0)
{
px = numeric(length(q0 - xrr - 1))
if(k > 0)px = (pbinom(k, m1, p0) - pbinom(k - 1, m1, p0)) * pbinom(q0 - k - 1, N - m1, p0)
else
px = pbinom(k, m1, p0) * pbinom(q0 - k - 1, N - m1, p0)
preject2p0 = preject2p0 + px
}
}
}

```

```

}
>preject2p0[ipreject2p0] = preject2p0;
>ipreject2p0 = ipreject2p0 + 1;
>preject2p0
}
> ipreject2p0 - 1
> preject2p0_
> preject2p0_
> ipreject2p1 = 1;
> for(xrr in xr : m1)
{
>preject2p1 = 0
> for(k in (xrr + 1) : q0)
{
>px = numeric(length(q0 - xrr - 1))
> if(k > 0)
>px = (pbinom(k, m1, p1) - pbinom(k - 1, m1, p1)) * pbinom(q0 - k - 1, N - m1, p1)
> else
>px = pbinom(k, m1, p1) * pbinom(q0 - k - 1, N - m1, p1)
>preject2p1 = preject2p1 + px
}
>preject2p1_[ipreject2p1] = preject2p1;
>ipreject2p1 = ipreject2p1 + 1;
>preject2p1
}
>ipreject2p1 - 1
> preject2p1_
> #alpha et la puissance après la 2ème étape
> alpha22 = NULL
> power22 = NULL
> ialpha22 = 1;
> ipower22 = 1;
> for(xrr in xr : m1)
{
> alpha21 = 1 - pbinom(xrr, m1, p0)

```

```

> aalpha22 = alpha21 - preject2p0
> alpha22[ialpha22] = aalpha22
> ialpha22 = ialpha22 + 1
> power21 = 1 - pbinom(xrr - 1, m1, p1)
> ppower22 = power21 - preject2p1
> power22
> power22[ipower22] = ppower22
> ipower22 = ipower22 + 1
}
> print("alpha22"); print(alpha22)
> print("power22"); print(power22)
> beta22 = 1 - power22
> tabf = round(cbind(xr : m1, alpha22, power22, p[xr : m1]), 4)
> tabf
> alphaa = numeric(m1 - xr)
> powera = numeric(m1 - xr)
> pii = numeric(m1 - xr)
> for(s in 1 : (m1 - xr))
{
> if(p[s] >= 0.5)
> pii[s] = p[s] else pii[s] = 0
> if(alpha22[s] <= 0.1)
> alphaa[s] = alpha22[s] else alphaa[s] = 0
> if(power22[s] >= 0.7)
> powera[s] = power22[s] else powera[s] = 0
> dist = cbind(pii, alphaa, powera)
}
> dist
> m1; xr; q0; N
}
> plot(1 : m1, p, col = "red", pch = 2)
> # points(xr : m1, tabf[, 3])

```

## Annexe II

### Loi poisson binomiale

Paramètres  $\mathbf{p} \in [0, 1]^n$ -probabilités de succès pour chacun des  $n$  essais

Support  $k \in \{0, \dots, n\}$

Fonction de masse  $\sum_{A \in F_k} \left( \prod_{i \in A} p_i \prod_{j \in A^c} (1 - p_j) \right)$

En théorie des probabilités et en statistique, la **loi poisson binomiale** est une loi de probabilité discrète de la somme d'épreuves de Bernoulli indépendantes.

En d'autres termes, c'est la loi de probabilité du nombre de succès (nombre de pile) d'une suite de  $n$  lancers de pile ou face dont les probabilités de succès (d'obtenir pile) sont  $p_1, p_2, \dots, p_n$ . La loi binomiale ordinaire est un cas spécial de la loi poisson binomiale lorsque toutes les probabilités sont les mêmes :  $p_1, p_2, \dots, p_n$ .

**package sinib** (somme de variables aléatoires *Binomiales* indépendantes non identiques)

Puisque la distribution de la somme des Binômes indépendants non identiques est définie par un vecteur d'essais et probabilités de paire (chaque paire pour un binôme constituant), il fallait redéfinissez ces arguments dans

les fonctions **sinib**. Par conséquent, les deux arguments suivants doivent être redéfinis :

- taille : vecteur entier du nombre d'essais.
- prob : vecteur numérique des probabilités de succès.

Tous les autres arguments restent les mêmes. Il est à noter que lorsque la taille et le prob les arguments sont donnés sous forme de vecteurs de longueur 1, les fonctions **sinib** se transforment en fonctions binom.

# Bibliographie

- [1] Aitchison, J; Dunsmore, I.R.(1975). *Statistical Prediction Analysis* .  
New York : Compridge University Press. ISBN : 978-0-521-20692-1.
- [2] Chow S. C. (2010). Bayesian adaptive methods for clinical trials. Chapman Hall/CRC Biostatistics Series, USA ; ISBN : 1-58488-776-1.
- [3] Christian P. Robert. (2006) Le choix bayésien principes et pratique, ISBN-10 : 2-287-25173-1 Springer Paris Berlin Heidelberg New York.
- [4] Comité international des rédacteurs de revue médicales.  
<http://www.drirc.ap-hop-paris.fr/publications/def-registre.pdf>.
- [5] Djeridi Z. (2019). Choix de l'a priori en statistique Bayésienne dans les essais expérimentaux, Thèse de Doctorat Ès-Sciences. Université de Constantine 1.
- [6] Djeridi, Z. and Merabet, H. (2019). A Hybrid Bayesian-Frequentist Predictive Design for Monitoring Multi-Stage Clinical Trials, *Sequential Analysis journal* , DOI :10.1080/07474946.2019.1648919, 2019.
- [7] Guosheng Yin, Nan Chen, and J. Jack Lee. Phase II trial design with Bayesian adaptive randomization and predictive probability.
- [8] Guosheng Yin, (1976) United States. Clinical trial design : Bayesian and frequentist adaptive methods. (Wiley series in probability and sta-

tistics), Includes bibliographical references and indexes. ISBN 978-0-470-58171-1 (cloth).

- [9] ICH. Harmonised tripartite guideline-General considerations for clinical trials - E8. 17 Juillet 1997. [Accès le 30 Mai 2014.].
- [10] Kramar Andrew, Simone Mathoulin-Pélissier, (2011), (Société française du cancer). *Méthodes biostatistiques appliquées à la recherche clinique en cancérologie*. Société Française Editions John Libbey Eurotext.
- [11] Lee J.J and Liu D.D. (2008). A predictive probability design for phase II cancer clinical trials, *Clinical Trials*, 5(2) : 93-106.
- [12] Merabet, H. and Labdaoui, A. (2015). Bayesian Prediction Using Two-Stages Design in Experimental Trials, *Nature and Technologie A-Sciences fondamentals et engineering* 12 : 11-23.
- [13] Rogatko A, Schoeneck D, Jonas W, Tighiouart M, Khuri FR, Porter A. Translation of innovative designs into phase I trials. *J Clin Oncol* (2007) ; 25 (31) : 4982-6.
- [14] Saville, B. R., Connor, J. T., Ayers, G. D., and Alvarez, J. (2014). The Utility of Bayesian Predictive Probabilities for Interim Monitoring of Clinical Trials, *Clinical Trials* DOI : 10.1177/1740774514531352.
- [15] Simon R (1989). Optimal two stage designs for phase II clinical trials. *Control Clin Trials* ; 10 : 1-10.
- [16] Simone Bateman (2009), *Les Essais cliniques : pourquoi sont-ils aujourd'hui critiqués, pour la science*.
- [17] Société canadienne du cancer. (2007). *Les essais cliniques ( guide à l'intention des personnes atteintes du cancer )*, canada ; 311-846.
- [18] Tourneau C, Jack Lee J, Siu LL (2009). Dose escalation methods in Phase I cancer clinical trials. *J Natl Cancer Inst* ; 101 : 708-20.

- [19] Thomas Quertermous and Boxiang Liu, December 6, 2017, Approximating the Sum of Independent Non-Identical Binomial Random Variables, Departments of Biology, Pathology, Genetics, Medicine - Cardiovascular Medicine, Stanford University fbliu2,tomq1g@stanford.edu, arXiv :1712.01410v1 [stat.CO].
- [20] Wenpin Tang, 02 September 2019, The poisson binomial distribution—Old & New, Columbia University.
- [21] Zerari A. (2013). Approche Bayésienne aux essais clinique séquentiels, Magister, université Constantine 1.