

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE
LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ FRÈRES MENTOURI CONSTANTINE
FACULTÉ DES SCIENCES EXACTES
DÉPARTEMENT DE MATHÉMATIQUES



N° d'ordre : 53/D3C/2020

N° de série : 02/Math/2020

THÈSE

en vue de l'obtention du diplôme de

DOCTORAT 3^{ème} cycle

Option : Statistique appliquée

Présentée par : *BOUREZAZ Houda*

Titre du sujet de thèse

*LES PLANS BAYÉSIENS
DANS LES ESSAIS EXPÉRIMENTAUX*

Dirigée par : MERABET Hayet

Membres du jury :

Président	Mr BESSILA Khaled	MCA	U. F. Mentouri Cne
Rapporteur	Mme MERABET Hayet	Prof	U. F. Mentouri Cne
Examineur	Mr HAMRI Nasr-Eddine	Prof	C. U. de Mila
Examineur	Mr BERKANE Abdelhak	MCA	U. F. Mentouri Cne
Examineur	Mr ABDELOUAHAB Med Salah	MCA	C. U. de Mila

Soutenue le 03/12/2020

Intitulé de la thèse :

*LES PLANS BAYÉSIENS
DANS LES ESSAIS EXPÉRIMENTAUX*

Réalisée par :

BOUREZAZ Houda

Année : *2020*

Remerciements

Tout d'abord, nous tenons à remercier Allah, le tout puissant et le miséricordieux, de nous avoir donnés la santé, la volonté et la patience pour mener à terme ce travail.

Ce travail ne serait pas aussi riche et n'aurait pas pu avoir le jour sans l'aide et l'encadrement de Madame « MERABET Hayet » Professeur au département de mathématiques à l'université des frères Mentouri Constantine, je vous remercie madame pour votre encadrement, pour votre patience, votre judicieux conseils, vos encouragements et votre disponibilité durant ma préparation de cette thèse.

Je remercie infiniment Monsieur « DRUILHET Pierre » Professeur des universités « Blaise Pascal de Clermont Ferrand » mon responsable des stages de recherche d'avoir accepté de m'accueillir au sein du Laboratoire de Mathématiques, je vous remercie monsieur pour votre qualité d'encadrement exceptionnelle, pour votre temps que vous m'avez accordée, votre disponibilité et vos explications dans le domaine de l'analyse Bayésienne.

Mes remerciements vont aussi à Messieurs les membres du jury pour avoir accepté de se déplacer de si loin et venir examiner et scruter de si près le contenu de ma thèse de Doctorat.

Je tiens à remercier Mr « B'ESSILA Khaled » Maître des conférences A-UFM à l'université Frères Mentouri Constantine d'avoir accepté de présider le jury de mon travail, qu'il trouve ici toutes mes expressions respectueuses.

Mes remerciements vont aussi à Mr « BERKANE Abdelhak » Maître des conférences A-UFM à l'université Frères Mentouri Constantine pour le temps consacré à examiner ce modeste travail.

Je remercie également Messieurs « HAMRI Nasr-Eddine » professeur et « ABDELOUAHAB Mohamed Salah » Maître des conférences A au centre universitaire de Mila de m'avoir fait l'honneur d'être membres du jury et d'examiner ce travail. tous mes remerciements les plus sincères pour vous.

Je remercie ici tous ceux qui directement ou indirectement ont contribué à la réalisation de cette étude.



Dédicaces

C'est grâce à ALLAH que j'ai eu le courage et la force d'accomplir ce modeste travail.

*A mes chers parents,
Aucune dédicace ne saurait exprimer mon respect, mon amour éternel et ma considération pour les sacrifices que vous avez consentis pour mon instruction et mon bien être. Je vous remercie pour tout le soutien et l'amour que vous me portez depuis mon enfance et j'espère que votre bénédiction m'accompagnera toujours.*

*A Mes sœurs Meroua et Oumeïma,
Qui n'ont pas cessé de me soutenir, de m'encourager et de croire en moi, que Dieu, le tout puissant, vous protège et vous garde.*

A mon encadrant Mme MÉRABET Hayet qui m'a accompagnée tout au long du chemin, merci énormément.

A mon encadrant de stage Mr DRUILHET Pierre qui je veux remercier bien de m'avoir aidée et orientée.

A mes amies, mes collègues et toute ma grande famille qui n'ont pas cessé de m'aider et de me conseiller pour accomplir mon travail.

Houda,



Résumé

Les travaux présentés dans cette thèse portent sur les problèmes méthodologiques qui se posent dans le contexte des essais expérimentaux, nous avons exposé les méthodes de l'analyse séquentielle mise en place pour l'utiliser dans l'évaluation des plans d'expérience dont elle permet d'arriver à des plans d'expérimentation efficaces.

Nous considérons l'utilisation de la méthode d'assurance dans la planification des essais cliniques : Dans la méthode de l'assurance, qui est une alternative au calcul de la puissance, nous calculons la probabilité qu'un essai clinique aboutisse à un résultat positif, en obtenant une distribution de probabilité a priori sur l'effet du traitement pertinent.

Dans ce travail, nous nous intéressons à la modélisation de la prédiction Bayésienne pour évaluer un indice de satisfaction après une première étape d'expérimentation afin de décider d'arrêter à la deuxième étape ou de continuer. Nous appliquons cette méthode aux résultats distribués de Poisson et Gamma dans de nombreux domaines tels que la fiabilité ou l'analyse de survie pour une interruption précoce en raison de la futilité ou de l'efficacité. Nous examinons deux types de prise de décision : une approche hybride Bayésienne-fréquentiste ou une approche Bayésienne complète.

Ce travail a permis aussi de déterminer la taille de l'échantillon efficace d'un paramètre a priori dans un modèle Bayésien.

Mots clés : Distribution prédictive Bayésienne, indices de satisfaction, assurance Bayésienne, approche hybride Bayésienne-fréquentiste, approche Bayésienne, analyse séquentielle en deux étapes, essais expérimentaux, résultats Poisson et Gamma, taille de l'échantillon efficace.

ملخص

يتعلق العمل المقدم في هذه الرسالة بالمشكلات المنهجية التي تنشأ في سياق الاختبارات التجريبية ، وقد كشفنا عن طرق التحليل المتسلسل المعدة لاستخدامها في تقييم التصاميم التي تسمح بها للوصول إلى خطط تجريبية فعالة. نقوم بدراسة استخدام طريقة التأمين في التخطيط للتجارب السريرية. في طريقة التأكيد، التي تعد بديلاً لحساب القوة ، نقوم بحساب احتمال أن تؤدي التجربة السريرية إلى نتيجة إيجابية، والحصول على توزيع احتمالي مسبق على تأثير العلاج ذي الصلة. في هذا العمل ، نحن مهتمون بنمذجة التنبؤ البايزية لتقييم مؤشر الرضا بعد المرحلة الأولى من التجريب من أجل اتخاذ قرار بالتوقف عند المرحلة الثانية أو المتابعة ، نطبق هذه الطريقة على نتائج بواسون و جاما الموزعة في العديد من المجالات مثل الوثوقية أو تحليل البقاء على قيد الحياة للانقطاع المبكر بسبب عدم الجدوى أو الكفاءة. ندرس نوعين من اتخاذ القرار نهج هجين بايزي متكرر أو نهج بايزي كامل. هذا العمل جعل من الممكن أيضًا تحديد حجم العينة الفعال لمعامل مسبق في نموذج بايزي.

كلمات مفتاحية:

توزيع بايزي التنبؤي، مؤشرات الرضا ، ضمان بايزي، نهج هجين بايزي متكرر، نهج بايزي، تحليل متسلسل من خطوتين، الاختبارات التجريبية نتائج بواسون و جاما، حجم العينة الفعال.

Abstract

The work presented in this thesis relates to the methodological problems which arise in the context of experimental tests, we have exposed the methods of the sequential analysis setting up to use it in the evaluation of experimental designs which allows to arrive at effective experimental plans.

We are considering the use of the assurance method in the planning of clinical trials : In the assurance method, which is an alternative to the power calculation, we calculate the probability that a clinical trial will lead to a positive result, via eliciting a priori probability distribution about the pertinent treatment effect.

In this work, we are interested to Bayesian prediction modeling to evaluate a satisfaction index after a first stage of experiment in order to decide to stop at the second stage or continue. We apply this method to Poisson and Gamma distributed outcomes in many fields such as reliability or survival analysis for early termination due to either futility or efficacy. We look at two kinds of decisions making : an hybrid Bayesian-frequentist or a full Bayesian approach.

This work also made it possible to determine the effective sample size of an a priori parameter in a Bayesian model.

Keywords : Bayesian predictive distribution, satisfaction index, Bayesian assurance, Bayesian-frequentist hybrid approach, Bayesian approach, two-steps sequential analysis, experimental trials, Poisson and Gamma results, effective sample size.

Table des matières

Résumé	5
Introduction générale	12
1 Notions sur l'approche Bayésienne	16
1.1 Essais expérimentaux	16
1.2 Essais cliniques	17
1.3 Conception adaptative Bayésienne en deux étapes	19
1.4 Approche Bayésienne	19
1.4.1 Intérêts de l'approche Bayésienne	19
1.4.2 Difficultés de l'approche Bayésienne	21
1.4.3 Distributions a priori et méthodes Bayésiennes	22
1.5 Expression des résultats et règles de décision	27
2 Analyse séquentielle	29
2.1 Introduction	29
2.2 L'a priori séquentiel	29
2.2.1 L'importance de l'exhaustivité	31
2.2.2 Détermination d'une statistique exhaustive	32
2.2.3 Notion de séquentialité	32
2.2.4 Point de départ : Choix de l'a priori	36
2.3 Les essais séquentiels par groupe	38
2.3.1 Les essais de phase I	38
2.3.2 Les essais de phase II	39
2.3.3 Les essais de phase II-III	40
2.3.4 Les essais de phase III	42
2.4 Prédiction séquentielle	42

2.4.1	Le modèle Poisson-Gamma	43
2.5	Tests séquentiels	45
2.5.1	Inférence non-décisionnelle	46
2.6	Test séquentiels dans les essais cliniques à un seul bras	48
2.6.1	Cadre général	49
2.6.2	Design à une seule étape	49
2.6.3	Design à deux étapes	50
3	Assurance Bayésienne	52
3.1	Introduction	52
3.2	Taille d'échantillon des essais non séquentiels	54
3.2.1	Approches alternatives à l'évaluation de la taille de l'échantillon	55
3.3	Puissance statistique conventionnelle à l'assurance	59
3.3.1	La puissance statistique conventionnelle et ses limites	59
3.4	Assurance dans les essais cliniques	61
3.4.1	Incertitude des paramètres	62
3.4.2	« Puissance classique » : méthode hybride Bayésienne - fréquentiste supposant la normalité	64
3.4.3	Puissance Bayésienne	68
3.4.4	Formules d'ajustement pour différentes hypothèses	71
3.4.5	Distribution prédictive de la puissance et taille d'échantillon nécessaire	77
3.5	Calculs d'assurance pour les modèles paramétriques de survie	79
3.5.1	Schémas de Randomisation	80
3.5.2	Distribution exponentielle	82
3.5.3	Construire les a priori	83
3.5.4	Logiciel supporté	85
3.5.5	Exemple numérique	86
3.6	Conclusion	88

4	Modélisation de la prédiction Bayésienne en deux étapes	89
4.1	Introduction	89
4.2	La fiabilité dans les essais expérimentaux	91
4.3	Procédure de test dans un cadre hybride	93
4.3.1	Test de Neymann-Pearson	94
4.3.2	Indice de satisfaction	95
4.3.3	Prédiction de satisfaction	96
4.4	Conception en deux étapes dans le cadre séquentiel	98
4.4.1	Inférence statistique pour un plan expérimental en deux étapes	98
4.4.2	Prévision de l'indice de satisfaction et des règles d'arrêt	101
4.5	Conception en deux étapes pour les résultats de Poisson et de Gamma	102
4.5.1	Une stratégie Bayésienne en deux étapes pour les résultats de Poisson	102
4.6	Une stratégie Bayésienne en deux étapes pour les résultats de Gamma	107
4.7	Résultats de simulation	110
4.7.1	Résultats pour les données de Poisson	110
4.7.2	Résultats pour les données de Gamma	113
4.7.3	Analyse des données réelles	114
4.8	Conclusion	117
5	Détermination de la taille effective de l'échantillon d'un paramètre a priori	118
5.1	Introduction	118
5.2	La taille effective de l'échantillon	120

5.3	Déterminer la taille effective de l'échantillon pour quelques modèles	122
5.3.1	Le modèle Bêta- Binomiale	123
5.3.2	Le modèle Poisson-Gamma	125
5.3.3	Le modèle Gamma-Gamma	126
5.4	Conclusion	130
	Conclusion générale	131
	Perspectives	132
	Annexes	133
	Bibliographie	134

Introduction générale

Plusieurs auteurs défendent l'intérêt de l'approche Bayésienne dans le contexte des essais expérimentaux. Sa capacité à intégrer des connaissances a priori sur l'effet du traitement dans l'analyse de l'essai et à exprimer les résultats sous la forme de distributions de probabilités, rendent cette approche particulièrement intéressante dans ce contexte.

Récemment, des méthodes ont été proposées afin d'incorporer des opinions d'experts a priori dans l'analyse de l'essai, les méthodes Bayésiennes sont déjà bien implémentées dans la pratique des essais cliniques de phase I. Cependant, cela ne semble pas être le cas pour les essais cliniques des phases II et III, où l'on cherche à évaluer l'efficacité et l'applicabilité d'un traitement [Saville et al., 2014, Hamada et al., 2008, brard, 2018, Djeridi 2019].

Le principe de l'analyse Bayésienne permet d'actualiser l'information disponible sur l'étendue plausible des valeurs de l'effet traitement avant l'essai (exprimée par une probabilité dite a priori, ou prior) par les données recueillies au cours de l'essai (caractérisées par la fonction de vraisemblance) pour produire une estimation révisée de l'étendue plausible des valeurs de l'effet traitement (exprimée par une probabilité dite a posteriori, ou posterior). La largeur (ou variance) de la distribution a priori représente le niveau de certitude de l'effet traitement. Augmenter la variance de cette distribution correspond à étendre les valeurs possibles de l'effet traitement. L'aire sous la courbe à gauche d'une valeur donnée (distribution cumulée) représente la probabilité que l'effet traitement soit inférieur à cette valeur. Le calcul de la distribution a posteriori peut se faire de manière analytique lorsque les lois a priori et a posteriori sont de la même famille, on parle alors de loi a priori conjuguée, cependant, l'analyse Bayésienne nécessite le calcul d'intégrales. Ce calcul ne peut pas être réalisé de façon analytique lorsque la loi a priori n'est pas conjuguée. Cela nécessite alors de réaliser soit une approximation par intégration numérique, soit une estimation par simulations de Monte Carlo [brard, 2018].

En particulier, nous nous intéressons à l'utilisation de cette approche dans la prévision au contexte des essais expérimentaux en raison du rôle essentiel que joue la probabilité prédictive dans la conception et la surveillance des essais (fiabilité des systèmes, médecine, biologie, écologie,...). Les procédures prédictives Bayésiennes sont des méthodes très attrayantes pour répondre à des questions essentielles dont la méthode fréquentiste traditionnelle ne peut pas répondre, elles 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 [Djeridi, 2019].

Dans les cas de la vérification d'efficacité d'un traitement dans un essai clinique, les analyses sont effectuées progressivement. En utilisant l'indice de satisfaction défini dans un cadre hybride Bayésien-fréquentiste on peut évaluer à quel point nous sommes satisfaits de l'efficacité de ce traitement. Labdaoui (2015) a démontré que l'utilisation de la prédiction de satisfaction donne de bons résultats dans le design séquentiel que dans la planification expérimentale à cause d'accumulation progressive d'informations. Dans cet égard, l'objectif principal de ce travail est de fournir une procédure hybride Bayésienne-fréquentiste et une procédure Bayésienne complète pour les conceptions à deux étapes dans le cadre séquentiel afin de tester l'efficacité d'un nouveau traitement. Ces procédures sont basées sur le concept d'indice de satisfaction [Merabet, 2013, Merabet et al., 2017, Djeridi, 2019]; où nous envisageons, compte tenu des données disponibles, de calculer une satisfaction prédite de cet indice par l'approche Bayésienne, en tenant compte des observations précédentes. De nombreux auteurs ont recommandé l'utilisation des indices dans de telles situations comme [Lecoutre et al., 1995, Merabet et Labdaoui, 2015, Djeridi, 2019] en raison de la simplicité et de la souplesse dont ils disposent pour mesurer le degré de satisfaction dans le cas d'obtenir un résultat significatif. Cet indice est utilisé comme une règle d'arrêt pour la conception d'essais cliniques de phase II.

La valeur prédite généralise la puissance du test dans la logique de l'indice de satisfaction proposé, c'est pour cela; nous avons consacré tout

un chapitre pour bien illustrer le concept de la puissance de test et de l'assurance Bayésienne pour avoir un résultat souhaité et déterminer la taille des échantillons.

Cette thèse est constituée de cinq chapitres :

Le premier chapitre du travail repose sur une revue systématique de la littérature afin de faire un état des lieux de l'utilisation des méthodes Bayésiennes dans l'analyse des essais expérimentaux et cliniques ayant pour but d'évaluer l'effet et l'applicabilité de nouvelles modalités de traitement.

Nous donnons le principe des méthodes séquentielles au chapitre 2 notamment le choix de l'a priori initial ainsi le choix de la règle d'arrêt pour les tests séquentiels fréquentistes et Bayésiens.

Au chapitre 3, nous donnons des outils théoriques et la différence entre les calculs de puissance traditionnels et les calculs d'assurance pour déterminer la taille des échantillons. Nous décrivons la transition de la puissance statistique conventionnelle à l'assurance nouvellement développée, nous examinons comment l'assurance est calculée pour déterminer la probabilité inconditionnelle afin d'avoir un résultat souhaité et discutons les calculs d'assurance par des exemples donnés.

Au chapitre 4, nous considérons la modélisation de prédiction Bayésienne pour évaluer un indice de satisfaction après une première étape d'expérimentation afin de décider d'arrêter à la deuxième étape ou de continuer. Nous appliquons cette méthode aux résultats distribués de Poisson et Gamma dans de nombreux domaines tels que l'analyse de la fiabilité ou de la survie pour un arrêt précoce dû soit à la futilité, soit à l'efficacité. Nous examinons deux types de prise de décision : Une approche hybride Bayésienne-fréquentiste ou une approche Bayésienne complète adaptée aux conceptions séquentielles.

Finalement au chapitre 5, nous présentons une définition de la taille effective de l'échantillon d'une distribution a priori paramétrique dans un modèle Bayésien et proposons des méthodes pour calculer cette taille dans une variété de paramètres. Nous calculons ensuite une distance entre chaque a posteriori et le paramètre a priori, définie en termes de cour-

bure du logarithme de chaque distribution, et la distance a posteriori minimisant la distance définissant la taille effective de l'échantillon de l'a priori.

Chapitre 1

Notions sur l'approche Bayésienne

1.1 Essais expérimentaux

Une phase expérimentale est souvent nécessaire lorsqu'un phénomène physique doit être modélisé. L'expérimentation permet en effet d'avoir une vision réaliste du phénomène étudié, ce qui est primordial pour une modélisation correcte. Si nous pouvons bien mettre en œuvre une conception d'expérience, l'expérience est probablement la conception la plus solide, alors les données générées sont valides et une analyse appropriée des données fournit les inférences statistiques valides de manière efficace et économique. Si l'expérience n'est pas bien conçue, la validité des inférences statistiques est discutable et peut être invalide. Il est important de comprendre d'abord les terminologies de base utilisées dans la conception expérimentale.

Dans certaines expériences, la prédiction peut être plus importante que l'estimation. Dans le contrôle de la qualité et dans les essais cliniques, la prédiction des observations futures peut être particulièrement intéressante. Dans ces cas, l'approche Bayésienne utilise une analyse prédictive qui peut également être utile dans la conception de l'expérience.

Lors de la conception d'une expérience, des décisions doivent être prises avant la collecte des données qui est bornée par des ressources limitées. Vu que l'information est généralement disponible avant l'expé-

rimentation et, en fait, motive souvent la réalisation de l'expérience, les méthodes Bayésiennes sont parfaitement adaptées pour contribuer à la conception expérimentale. La théorie de la décision Bayésienne motive également la spécification précise de la raison pour laquelle l'expérience est menée. Comme la plupart des domaines des statistiques Bayésiennes, la conception expérimentale Bayésienne a gagné en popularité au cours des deux dernières décennies. Mais comme de nombreux domaines de la statistique Bayésienne, les applications aux expériences réelles sont encore en retard par rapport à la théorie.

1.2 Essais cliniques

Les essais cliniques sont utilisés pour créer des options de prévention, de diagnostic ou de traitement pour les personnes atteintes d'une maladie ou d'une affection donnée. L'aspect le plus important d'un essai clinique est que l'essai vise à utiliser les résultats d'un échantillon limité de participants pour voir si un traitement est sûr et efficace ou comparable à un témoin. Les chercheurs adaptent souvent de nouvelles conceptions ou modifient les conceptions actuelles afin d'économiser du temps et de l'argent. Cela peut être effectué par plusieurs manières. Parfois, cela se fait par la détermination de la taille de l'échantillon pour trouver la plus petite taille d'échantillon possible sans perte de puissance ou d'erreur de type I. D'autres fois, les cliniciens veulent adapter une conception afin de pouvoir terminer un essai plus tôt en raison de l'efficacité ou de la futilité. Ou peut-être que les chercheurs aimeraient utiliser la conception qui leur donne les meilleures chances de succès. Quelle que soit la raison, les chercheurs veulent toujours trouver la « meilleure » conception pour leur problème d'intérêt tout en économisant le plus de temps et d'argent.

Un essai clinique se compose généralement de trois phases, chacune aussi importante que l'autre. La phase I comprend un petit nombre d'individus et des tests de sécurité générale et de profils pharmacologiques. La phase II compte également un petit nombre de patients par contre

les chercheurs sont intéressés par une évaluation pilote de l'efficacité. La dernière phase avant l'approbation, la phase III, est une étude beaucoup plus vaste où les chercheurs évaluent de manière approfondie l'innocuité et l'efficacité. Avec autant d'étapes, les cliniciens souhaitent minimiser le nombre de participants inclus dans leurs études ou minimiser le temps nécessaire à l'étude. Encore une fois, l'un ou l'autre cas peut être effectué en adaptant de nouveaux modèles ou en modifiant les modèles actuels. Récemment, les méthodes Bayésiennes sont devenues d'un grand intérêt pour les chercheurs. Les méthodes Bayésiennes permettent aux chercheurs d'incorporer plus facilement des informations a priori ou des données historiques que les méthodes fréquentistes traditionnelles. Cet avantage par rapport aux méthodes fréquentistes séduit de nombreux chercheurs qui souhaitent intégrer leurs connaissances du sujet traité. Cependant, les chercheurs ne sont pas tous intéressés par les méthodes Bayésiennes car l'influence potentielle de l'incorporation d'informations a priori peut avoir sur les conclusions. Les distributions a priori sont un moyen pour les chercheurs d'incorporer leurs croyances ou connaissances a priori sur un paramètre d'intérêt. L'incorporation de l'a priori doit être justifiée et, par conséquent, l'influence de l'a priori sur la distribution a posteriori et les conclusions doivent être appropriées et justifiées. Nous pouvons non seulement inclure les connaissances ou les croyances a priori d'un paramètre d'intérêt dans notre distribution a priori, mais nous pouvons également incorporer l'incertitude ou la certitude que nous avons de nos connaissances ou de nos croyances. En règle générale, plus nous sommes confiants dans nos informations, plus nous utilisons un a priori informative, et de même, moins nous sommes confiants dans nos informations a priori, moins nous utilisons un a priori informative, souvent désigné comme un « non informatif » [Hand 2011].

1.3 Conception adaptative Bayésienne en deux étapes

Les conceptions des essais cliniques adaptatives peuvent rendre les études plus efficaces en adaptant les éléments de conception au cours de l'essai en fonction d'informations inconnues au départ. Une adaptation largement utilisée consiste à ajuster la taille totale de l'échantillon de l'étude en fonction des taux d'événements agrégés et observés lors des analyses intermédiaires prévues. Nous présentons une nouvelle conception adaptative en deux étapes pour les essais cliniques de phase II avec des résultats distribués de Poisson qui permet des ajustements de la taille de l'échantillon et une interruption précoce en raison de la futilité ou de l'efficacité. De nombreux modèles adaptatifs fréquentistes et Bayésiens ont été proposés dans la littérature. Certains modèles ne permettent pas d'ajuster la taille de l'échantillon après la première étape. Cette restriction limite la flexibilité dans le plan d'étude. La conception proposée, cependant, permet une telle flexibilité en basant la taille de l'échantillon de la deuxième étape sur des résultats de la première étape. Nous mettons en œuvre notre conception adaptative prédictive Bayésienne en deux étapes à l'aide d'un hypothétique essai de phase II pour plusieurs valeurs de l'hyper-paramètre [Hand 2011].

1.4 Approche Bayésienne

1.4.1 Intérêts de l'approche Bayésienne

Nous nous placerons dans le contexte d'un essai clinique avec une perspective opérationnelle : nous souhaitons prendre une décision quant à l'efficacité d'un nouveau traitement en considérant toute l'information disponible. Dans cette perspective, l'un des aspects les plus intéressants de l'approche Bayésienne est la possibilité de prendre en compte des informations extérieures pour l'analyse du critère de jugement principal

de l'essai, à travers de la distribution a priori [Berry DA., 2006, Lee JJ. et Chu CT., 2012]. Ces informations peuvent être de différents types :

- Opinions d'experts.
- Données historiques.

L'utilisation d'une distribution a priori non-informative est également possible. De plus, l'approche Bayésienne permet d'exprimer les résultats sous la forme de probabilités a posteriori et d'intervalle de crédibilité, plus facilement interprétables par le clinicien et le patient qu'une p-valeur ou un intervalle de confiance comme mesures d'incertitude de l'effet du traitement considéré.

Ainsi, le calcul et l'interprétation de la p-valeur correspondent à un raisonnement par l'absurde [Caroline Brard. 2018]. Quand on conclut à une différence significative entre deux groupes de traitement au seuil 5%, cela signifie que les données observées sont très « improbables » sous l'hypothèse nulle H_0 et que l'on peut donc rejeter H_0 et conclure à un effet du traitement, au risque de se tromper 5 fois sur 100. Ce concept est loin d'être compris dans la communauté médicale.

L'intervalle de confiance à 95% qui permet de donner une mesure de la précision de l'estimation, est également souvent interprété à tort comme l'intervalle de valeurs ayant 95 chances sur 100 de contenir la vraie valeur, ce qui correspond stricto sensu à l'intervalle de crédibilité. Expliquer la signification de l'intervalle de confiance est complexe : Il correspond à un intervalle tel que, si le même essai était répété 100 fois, 95 des intervalles de confiance obtenus contiendraient la vraie valeur de l'estimation. La simplicité d'interprétation de l'intervalle de crédibilité contraste avec la complexité de l'intervalle de confiance. Par ailleurs, quand les experts proposent de prendre en compte la magnitude du bénéfice clinique dans la décision, il nous semble que les règles proposées se définissent plus facilement dans le cadre Bayésien.

L'approche Bayésienne peut également être particulièrement intéressante pour « monitorer » les essais. En effet, elle permet d'arrêter l'essai lorsqu'il y a une forte probabilité d'un effet du traitement cliniquement si-

gnificatif [Gsponer et al. 2014, Saville et al. 2014], ou au contraire lorsque la probabilité que l'effet du traitement soit cliniquement significatif est faible. Ces dernières années, l'utilisation des méthodes Bayésiennes a été facilitée par l'amélioration et l'implémentation des techniques de calculs et la disponibilité des ordinateurs performants et des programmes informatiques permettant de réaliser des analyses Bayésiennes [Brard, 2018].

1.4.2 Difficultés de l'approche Bayésienne

Des difficultés techniques existent dans la construction de la distribution a priori, en particulier quand le critère étudié est une donnée censurée. Dans l'analyse habituelle fréquentiste des essais cliniques sur données censurées, un modèle de Cox semi-paramétrique est souvent utilisé pour évaluer l'effet relatif d'un traitement par le ratio d'hasard (HR). Il ne nécessite pas de modéliser la fonction de survie sous-jacente (risque de base).

Dans le cadre Bayésien, il est possible de réaliser une analyse Bayésienne simple que nous avons appelée « hybride ». En effet, en considérant que le $\text{Log}(\text{HR})$ suit une loi normale [Brard, 2018], l'utilisation d'une loi a priori conjuguée permet d'intégrer des données historiques sur l'effet relatif du traitement. Cette approche hybride consiste à mettre à jour la distribution a priori avec l'estimation du $\text{Log}(\text{HR})$ obtenue à partir d'une analyse fréquentiste. La distribution a posteriori du paramètre peut alors être déterminée de manière analytique [Spiegelhalter et al., 2004, Fayers et al., 1997].

En revanche, si l'on souhaite intégrer des informations sur une fonction de survie, par exemple parce qu'on dispose de données historiques individuelles pouvant enrichir l'information sur le groupe contrôle ; il devient nécessaire de préciser un processus stochastique particulier, donc de modéliser la fonction de survie ; la spécification du modèle le plus adéquat pour le risque de base peut être difficile en particulier dans le contexte des pathologies rares.

Dans le cadre des essais cliniques visant à évaluer de nouveaux traite-

ments, outre les difficultés techniques pour les statisticiens, la principale difficulté pour les destinataires des analyses (communauté médicale et scientifique, décideurs...) concerne la subjectivité dans l'interprétation de l'essai. Celle-ci est liée à l'intégration d'une distribution a priori informative pour des paramètres inconnus comme l'effet du traitement, les paramètres du groupe contrôle ou les effets des covariables. Bien qu'elle reconnaisse l'intérêt de l'approche Bayésienne pour les essais cliniques dans les maladies rares, l'European Medical Agency (EMA) exprime certaines réserves quant à l'incorporation de croyances a priori, et au fort impact qu'elles peuvent avoir sur les résultats et sur la décision finale corrélée [Brard, 2018]. Même si les distributions a priori sont basées sur des données réelles, l'investigateur doit décider quels jeux de données vont être incorporés et comment les pondérer en fonction de leur commensurabilité anticipée ou observée avec les futures données de l'essai [Lunn et al., 2012], on devrait quantifier l'incertitude sur les différences possibles entre la population source et la population cible. Lorsqu'il n'existe aucun jeu de données pertinent, les distributions a priori peuvent être basées sur des opinions d'experts. Cependant, il peut être difficile de savoir qui considérer comme un « expert » et de ce fait là, la subjectivité de ce type de distribution a priori est fortement critiquée [Brard, 2018].

1.4.3 Distributions a priori et méthodes Bayésiennes

La construction de la distribution a priori est une étape importante de l'inférence Bayésienne. Il faut en effet caractériser ce qui est connu, ou ce qui est pensé sur l'effet du traitement. Cette distribution peut être ensuite convertie en « pseudo-observations » afin de mieux visualiser ce que représente la quantité d'information contenue dans cette distribution a priori. La distribution a priori peut être informative ou non-informative. Il existe différents types de distribution a priori informatives et différentes méthodes pour les incorporer.

Distribution a priori informative

La distribution a priori informative peut être constituée d'opinions particulières, sceptiques ou enthousiastes, d'opinions d'un panel d'experts ou construit à partir de différentes sources telles que des données historiques issues de précédents essais ou études. Ainsi, des distributions informatives de type archétypale (from the shelf en anglais) peuvent être utilisées, choisies pour représenter des opinions sceptiques ou enthousiastes par rapport à l'effet du traitement [Parmar et al., 1994, Brard, 2018]. Concernant la distribution a priori du $\text{Log}(\text{HR})$ dans un essai contrôlé randomisé d'efficacité, la pratique traditionnelle est de définir une distribution a priori sceptique avec une moyenne centrée sur l'absence d'effet du traitement et une distribution a priori enthousiaste centrée sur un HR caractérisant une hypothèse alternative donnée, correspondant le plus souvent à l'hypothèse alternative du calcul du nombre de sujets nécessaire de l'approche fréquentiste. Ainsi, dans le cadre du monitoring d'un essai, on abandonnera un traitement dont l'efficacité s'avère insuffisante, même aux yeux des cliniciens enthousiastes, tandis que l'on conclura à son intérêt si on arrive à convaincre les cliniciens sceptiques [Parmar et al., 1994]

Une distribution a priori peut être également basée sur des opinions d'un panel d'experts interrogés dans le contexte de la problématique considérée, opinions qui sont ensuite transformées en distribution mathématique.

Cette approche nécessite à la fois une bonne pédagogie de la part du statisticien qui recueille les opinions et une certaine compréhension des statistiques de la part des experts interrogés. Différents processus pour l'élicitation d'opinions d'experts ont été développés [Thall et al., 2017, Hampson et al., 2015]. Une seule distribution a priori consensuelle peut être dérivée au travers d'une agrégation mathématique ou comportementale (où les experts sont amenés à un consensus ; par exemple la méthode Delphi). Alternativement, un panel d'opinions peut être enregistré pour capturer la diversité des opinions soutenues par la communauté clinique,

conduisant à l'élaboration de plusieurs distributions a priori différentes.

Le troisième type de distribution a priori informative repose sur des données dites « historiques », extérieures à l'essai, généralement connues au moment de la planification de l'essai. En effet aujourd'hui, de plus en plus de données sont accessibles aux chercheurs. Elles peuvent être sous forme de données individuelles ou d'informations agrégées extraites de publications. Elles peuvent provenir d'essais cliniques antérieurs ou être générées par des études observationnelles (données de registres, bases de données médico-administratives...). Elles peuvent porter sur l'effet relatif du traitement expérimental en comparaison au traitement contrôle ou sur les paramètres d'un des groupes de traitement seulement, en particulier le groupe contrôle. Les informations extérieures sur l'effet possible du nouveau traitement sont généralement utilisées pour justifier l'essai et le calcul du nombre de sujets nécessaire : Données d'essai de phase II a priori, données d'un essai réalisé sur une population proche, données d'un essai évaluant un traitement présentant de grandes similitudes avec le traitement évalué par exemple.... Il est également fréquent de disposer de l'informations sur le traitement servant de traitement de référence pour le groupe contrôle.

Dès 1976, Pocock a suggéré l'intégration de données historiques sur le groupe contrôle dans l'analyse d'un essai randomisé. Il a proposé six critères pour s'assurer de la pertinence des données contrôles historiques par rapport au nouvel essai [Pocock, 1976] sans pour autant garantir une parfaite compatibilité des données.

1. Le groupe contrôle historique doit avoir reçu le même traitement standard que le groupe contrôle du nouvel essai.
2. Le groupe contrôle historique doit faire partie d'un essai clinique récent contenant les mêmes critères d'éligibilité pour les patients.
3. La méthode d'évaluation du traitement doit être la même.
4. Les distributions des caractéristiques des patients dans le groupe contrôle historique doivent être comparables avec celles du nouvel essai.

5. L'essai historique doit être réalisé dans la même organisation avec globalement les mêmes investigateurs.
6. Il ne doit pas y avoir d'autres indications laissant penser que les résultats de l'essai historique et du nouvel essai puissent être différents (par exemple un recrutement plus rapide dans la nouvelle étude).

Dans le même esprit, Tan et al.(2003) proposent une stratégie pour sélectionner les études à inclure dans la construction de la distribution a priori en tenant compte de la pertinence, de la validité et de la précision. Ces études sont alors pondérées en fonction de ces trois critères.

Qu'il s'agisse des données historiques sur l'effet relatif du traitement évalué ou des données sur l'un des groupes de traitement, la commensurabilité des données historiques avec les données de l'essai actuel est un point de discussion important au moment de la planification et de la réalisation de l'analyse Bayésienne. À noter que le terme de commensurabilité est largement utilisé dans les articles rapportant des méthodes Bayésiennes sans qu'existe une définition formelle. On entend ici par commensurabilité entre deux jeux de données, le fait que l'on puisse considérer que ces jeux de données historique et actuel sont générés par des processus statistiques reposant sur des paramètres similaires. Le degré de commensurabilité peut être défini a priori par les experts. Il peut également être évalué à travers des tests.

Dans le contexte des pathologies rares, pour ce qui est de données historiques pouvant enrichir l'information du groupe contrôle, le traitement standard reste souvent stable au cours du temps puisque les avancées vers de nouvelles options sont plus lentes ; nous pouvons donc atteindre une certaine commensurabilité entre la performance du traitement contrôle dans les données historiques et celle dans le nouvel essai si les critères de sélection des patients et d'évaluation restent stables. En 2006, la Food and Drug Administration (FDA) publie ses recommandations pour l'utilisation des statistiques Bayésiennes dans le cadre des essais cliniques portant sur les dispositifs médicaux [Food and drug administration, 2010].

Ces recommandations soulignent l'avantage apporté par l'intégration de données historiques dans la distribution a priori du paramètre d'intérêt tout en insistant sur l'importance de pondérer cette information.

Concernant les méthodes Bayésiennes, les auteurs distinguent les méthodes dynamiques et les méthodes non dynamiques. Dans les méthodes dynamiques, les données extérieures sont pondérées selon leur commensurabilité avec les données actuelles en utilisant un paramètre aléatoire. Dans les méthodes non-dynamiques, le poids accordé aux données extérieures à l'essai est soit préspecifié dans le protocole, sur la base d'opinions d'experts ou d'une étude de calibration par simulations, soit fixé au moment de l'analyse au vu de la commensurabilité observée [Ibrahim et al., 2012].

Parmi les méthodes utilisables pour des critères de survie avec le Log (HR) comme paramètre d'intérêt, une possibilité est d'utiliser un power prior, développé par Ibrahim et al. 2000 notamment pour les modèles de régression, qui élève à la puissance α_0 la vraisemblance de l'essai historique, afin de pondérer ces données. La valeur de α_0 peut être fixe où une distribution a priori peut être définie pour ce paramètre [Brard, 2018].

Distribution a priori non-informative

De nombreux cliniciens et statisticiens, ainsi que les autorités de santé, trouvent difficile d'accepter l'inclusion d'une distribution a priori basée sur des éléments subjectifs [Gelman et al., 2014, Moyé LA., 2008]. L'utilisation d'une distribution a priori non-informative, où toutes les tailles d'effet du traitement sont équiprobables, peut être proposée pour l'analyse principale ou secondaire d'un essai clinique. La distribution a posteriori coïncide alors avec la vraisemblance, ne reflétant que ce que les données de l'essai apportent à la connaissance. Elle peut également compléter la distribution a priori informative dans le cadre d'une analyse de sensibilité [Brard, 2018].

1.5 Expression des résultats et règles de décision

L'approche Bayésienne permet d'enrichir l'approche fréquentiste classique en fournissant une description complète de la distribution de l'effet d'un traitement, permettant de traduire les résultats dans une forme plus facilement interprétable par les cliniciens et les patients (probabilité d'une certaine quantité d'effet du traitement, ou intervalle de crédibilité), et de choisir d'autres règles de décision que le test habituel de rejet de l'hypothèse nulle (absence de l'effet du traitement) [Brard, 2018]. À la fin d'un essai, l'incertitude sur les bénéfices et les risques d'un traitement peut donc être quantifiée et ainsi contribuer aux futures prises de décision [Hughes, 1993].

Récemment, de nombreux experts ont insisté sur la nécessité d'évaluer l'effet d'un traitement non seulement à l'aune de la significativité du test mais également de la magnitude de l'effet clinique. Le seuil pour considérer la pertinence clinique peut prendre en compte le risque de base de la population à l'étude, la complexité de l'intervention évaluée, le risque d'effets indésirables et les autres critères cliniques [Brard, 2018].

Un groupe d'experts de l'ESMO a formalisé ce concept de pertinence clinique [Cherny et al., 2017]. Des recommandations ont été établies dans un cadre fréquentiste en comparant, à une certaine valeur seuil, la borne inférieure de l'intervalle de confiance à 95% pour prendre en compte ce qu'ils appellent la « variabilité » du HR. Ainsi, pour un essai utilisant la survie globale comme critère de jugement avec une médiane de survie dans le groupe contrôle de plus de 12 mois, les experts concluent à un bénéfice cliniquement pertinent si la borne inférieure de l'intervalle de confiance à 95% est inférieure à 0.7. L'interprétation de cette règle ne nous semble pas immédiate contrairement à une règle Bayésienne basée sur l'expression probabiliste de l'effet du traitement par rapport à un seuil. On pourrait ainsi proposer de considérer comme bénéfice cliniquement pertinent un traitement pour lequel la probabilité d'un $HR < 0.7$

serait supérieure à 2.5% (si l'on reste sur une traduction grossière de la règle fréquentiste évoquée ci-dessus, tout autre seuil peut être pris en considération. D'autres règles plus intuitives pourraient également être proposées [Brard, 2018].

Dans le cadre du monitoring des essais, des règles d'arrêt Bayésiennes peuvent être proposées pour les analyses intermédiaires basées sur les probabilités a posteriori de l'effet du traitement [Berry D, 1985, Johnson VE et Cook JD, 2009].

Chapitre 2

Analyse séquentielle

2.1 Introduction

L'analyse séquentielle est une approche consistant à appliquer une méthode statistique de façon répétée jusqu'à ce qu'une condition, en principe fixée au départ, soit réalisée. Cette condition peut être la précision d'un estimateur ou bien le résultat d'un test d'hypothèses données à l'avance pour déterminer s'il faut continuer ou arrêter l'essai, dans laquelle la taille et la composition des échantillons ne sont pas fixées à l'avance mais sont fonctions des observations elles-mêmes et qui nécessite une évaluation précise des risques. Plusieurs applications dans les essais cliniques de phase II permettent de dégager l'intérêt, les limites et le champ d'application d'une telle technique. Malgré ses imperfections, elle constitue un outil non négligeable et une façon différente de raisonner statistiquement, elle est utile dans l'évaluation des plans d'expérience dont elle permet d'arriver à des plans d'expérimentation efficaces et dans le contrôle de qualité des matériaux fabriqués.

2.2 L'a priori séquentiel

Considérons un modèle dominé dans lequel la loi de densité sous-jacente $f(x|\theta)$ est paramétrée par θ et un échantillon x de la variable

aléatoire X ayant cette loi de distribution. Nous acceptons, par la suite, l'hypothèse que'on connaît avec certitude la forme analytique de la densité $f(x|\theta)$ mais que nous ne connaissons pas avec précision son paramètre θ . Nous disposons d'une série de résultats d'expériences, identiquement distribués suivant une loi paramétrée par θ . Il existe plusieurs méthodes pour déterminer la distribution a priori du paramètre θ selon le type d'information disponible [djeridi, 2019].

Nous allons étudier et développer quelques unes de ces méthodes qu'on appelle séquentielle qui peut être utilisée avec profit lorsque nous nous limitons au cas où l'information résulte d'observations passées du processus aléatoire. L'évaluation de la distribution a priori est réalisée à l'aide des idées classiques et les analyses associées sont plutôt à la frontière entre statistiques Bayésiennes et classiques ; ainsi dans ce domaine, l'étude est peu développée. La méthode séquentielle peut être présentée avec une distribution, au départ, soit une mesure uniforme sur l'ensemble Θ , soit une distribution non informative de la variable aléatoire θ , elle permet de retrouver la même forme de distribution et le même paramètre que celle de la conjuguée naturelle avec la fonction de vraisemblance [djeridi, 2019]. Malgré des restrictions sur la famille de distributions conjuguées, la propriété de conjugaison permettra d'avoir une stabilité de la forme des distributions ; ainsi le passage de la distribution a priori à l'a posteriori se réduit à un changement de paramètres, afin de réaliser facilement les calculs.

Une fois la distribution a priori soit déterminée, nous pouvons examiner successivement :

- La distribution a posteriori de θ ;
- La distribution prédictive de la variable aléatoire X ;
- La distribution prédictive d'une nouvelle variable Y en donnant la valeur de la variable aléatoire X et puis nous passerons éventuellement à l'étape de l'analyse des tests d'hypothèses.

2.2.1 L'importance de l'exhaustivité

Dans un problème statistique où figure un paramètre θ inconnu, un échantillon nous apporte une certaine information sur ce paramètre. Lorsqu'on résume cet échantillon x par une statistique, il ne faut pas perdre cette information sur ce paramètre θ ; une statistique qui la conserve sera qualifiée d'exhaustive.

En terme mathématique, une statistique T est dite exhaustive par rapport à θ , si la distribution conditionnelle $P_T = P_\theta[. | T(X) = t]$ d'un n échantillon aléatoire X , sachant que $T(X) = t$, est indépendante du paramètre θ . Cette distribution conditionnelle est portée par l'ensemble $\Omega_t = \{x \in \Omega; t = T(X)\}$, où $\Omega = X_n$ est l'ensemble des observations. Si le modèle est dominé et que la statistique T est exhaustive, la fonction de vraisemblance de l'échantillon x se factorise sous certaines conditions de régularité de la manière suivante :

$$f(x|\theta) = h(x)g_\theta(T(x)) = h(x)g(T(x)|\theta) \quad (2.1)$$

où $h(x)$ est une fonction indépendante du paramètre θ . La notation $g(T(x)|\theta)$ indique que g dépend de θ , mais sa dépendance en x est seulement fonction de $t = T(x)$. Le plus souvent $h(x)$ peut être choisi comme une densité de P sur Ω_t et g comme une densité de la loi de $T(x)$. Dans ce cas, nous disons que la distribution conditionnelle $f(x|\theta)$ admet une statistique exhaustive T .

Ainsi, toute l'information sur θ contenue dans $f(x|\theta)$ se retrouve dans $g(T(x)|\theta)$. Nous pouvons donc dire qu'une fois $t = T(x)$ connue, aucune valeur de l'échantillon ou d'autres statistiques ne nous apporteront des renseignements supplémentaires sur θ . Ce principe d'exhaustivité est très répandu dans l'analyse statistique.

Le principe de factorisation nous donne donc un moyen de reconnaître si une statistique est exhaustive mais permet difficilement de la construire ou même de savoir s'il en existe. L'utilisation de la notion d'exhaustivité nous aidera à trouver une loi de distribution du paramètre θ dans l'approche Bayésienne et à simplifier les procédures de recueil des données

[djeridi, 2019].

2.2.2 Détermination d'une statistique exhaustive

Exemple 2.1. La loi de Poisson

La loi de Poisson est une loi de probabilité discrète qui décrit le comportement du nombre d'évènements se produisant dans un laps de temps fixé, si ces évènements se produisent avec une fréquence moyenne connue et indépendamment du temps écoulé depuis l'évènement précédent. Si le nombre moyen d'occurrences dans cet intervalle est θ (le paramètre de la loi), alors la probabilité qu'il existe exactement x occurrences ($x \in \mathbb{N}$) est

$$p_X(x; \theta) = P(X = x) = \frac{\theta^x}{x!} \exp[-\theta], \quad x = 0, 1, 2, \dots \quad (2.2)$$

La loi jointe pour l'échantillon i.i.d. (x_1, \dots, x_n) est donnée par :

$$\begin{aligned} p_X(x_1, \dots, x_n; \theta) &= \prod_{i=1}^n P(x_i; \theta) \\ p(x_1, \dots, x_n; \theta) &= P(X = x) = \prod_{i=1}^n \frac{\theta^{x_i}}{x_i!} \exp[-\theta] \\ p(x_1, \dots, x_n; \theta) &= P(X = x) = \theta^{\sum_{i=1}^n x_i} \exp[-n\theta] \frac{1}{\prod_{i=1}^n x_i!} \end{aligned} \quad (2.3)$$

où $h(x) = \frac{1}{\prod_{i=1}^n x_i!}$ et $g(T(x)|\theta) = \theta^{\sum_{i=1}^n x_i} \exp[-n\theta]$

A partir de ce résultat et d'après le théorème de factorisation (critère de Fisher-Neyman), on a : $\sum_{i=1}^n x_i$ est une statistique exhaustive pour θ .

2.2.3 Notion de séquentialité

Soit (x_1, x_2, \dots, x_n) une série d'échantillons successifs, de tailles k_i , relatifs à des expériences aléatoires identiques. Soit $p(\theta)$ une distribution a priori. L'hypothèse séquentielle consiste à considérer la distribution a posteriori $p(\theta|x_i)$ comme une nouvelle distribution a priori pour l'échantillon suivant x_{i+1} . La distribution de θ est alors réajustée lors de la réalisation de x_{i+1} par la formule :

$$p(\theta|x_{i+1}) \propto f(x_{i+1}|\theta)p(\theta|x_i)$$

Partant du premier résultat x_1 , nous calculons les distributions de θ l'une après l'autre et nous avons finalement :

$$p(\theta|x_1, x_2, \dots, x_n) = \frac{\prod_j f(x_j|\theta)p(\theta)}{\int_{\Theta} \prod_j f(x_j|\theta)p(\theta)d\theta}$$

Remarquons que, s'il n'existe pas de stabilité de la forme de la suite de distributions $p_i(\theta) = p(\theta|x_1, x_2, \dots, x_n), i = 1, 2, \dots, n$, la formule n'est guère exploitable de manière analytique sauf s'il existe une statistique exhaustive.

Soit y un nouvel échantillon, la distribution a posteriori $p(\theta|y)$ peut être considérée comme la $(n + 1) - i\grave{e}me$ distribution de la séquence :

$$p(\theta|y) = p(\theta|x_1, x_2, \dots, x_n, y)$$

Nous observons que cette méthode incorpore facilement les données consécutives et que quelle que soit la forme de la distribution a priori, si le nombre n des échantillons est grand, la modification de la distribution a posteriori est très faible. Nous sommes donc amenés, pour tenir compte des connaissances récentes, à limiter le nombre n , c'est à dire à ne pas trop cumuler d'informations disponibles [Zhu, 1991].

Exemple 2.2. Modèle de Poisson pour les données de comptage des pannes d'un superordinateur.

Envisager de modéliser le nombre mensuel de pannes des composants de superordinateur du Laboratoire National Blue Mountain du Los Alamos (processeurs à mémoire partagée, en anglais shared memory processors ou SMPs) par une distribution de Poisson. Le superordinateur se compose de 47 SMPs « identiques » [Hamada et al., 2008]. Soit x_1, x_2, \dots, x_{47} indiquent le nombre mensuel de défaillances (pannes) pour le premier mois de fonctionnement enregistrées pour les SMPs. Avec $t_i = 1$ mois, nous modélisons les données de comptage des défaillances par la distribution de Poisson :

$$X_i \sim Poisson(\theta t) = Poisson(\theta), \quad i = 1, \dots, 47,$$

où θ est le nombre mensuel moyen de défaillances.

Les ingénieurs de superordinateur s'attendent à ce qu'il n'y ait pas plus de 10 pannes pour chaque composant au cours du premier mois de fonctionnement. Une façon de représenter cette information a priori consiste à supposer une distribution a priori Gamma pour θ avec une moyenne de cinq. Nous pouvons exprimer ces informations a priori par :

$$\theta \sim \text{Gamma}(\alpha = 5, \beta = 1)$$

Notez que pour la distribution a priori $\text{Gamma}(5, 1)$, la probabilité que θ dépasse 10 est de 0,03. La distribution a posteriori de θ étant donnée les données des défaillances $x = (x_1, x_2, \dots, x_{47})$ est :

$$\begin{aligned} \theta|x &\sim \text{Gamma}(\alpha + \sum_{i=1}^n x_i, \beta + \sum_{i=1}^n t_i) \\ &= \text{Gamma}(5 + 132, 1 + 47) = \text{Gamma}(137, 48) \end{aligned}$$

La figure 2.1 représente les distributions a priori et a posteriori du nombre mensuel moyen des défaillances θ . Notez la distribution a priori relativement diffuse (ligne pointillée) et la distribution a posteriori très pointue (ligne continue), ce qui indique que les données de comptage des pannes fournissent des preuves substantielles d'un nombre mensuel moyen de pannes inférieur à celui attendu par les ingénieurs. Le nombre mensuel moyen des défaillances a posteriori est :

$$E(x|\theta) = \frac{\alpha^*}{\beta^*} = 137/48 = 2.85$$

L'écart type a postérieur est :

$$\sqrt{\text{Var}(x|\theta)} = \sqrt{\frac{\alpha^*}{\beta^{*2}}} = 0.24$$

et un intervalle crédible à 95% correspond à (2.40, 3.35) échecs mensuels.

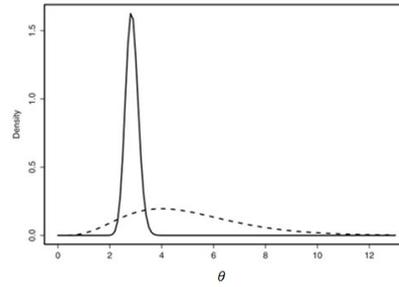


FIGURE 2.1 – Distribution a priori conjuguée Gamma $(\alpha; \beta)$ (ligne pointillée) et a posteriori Gamma $(\alpha + 132; \beta + 47)$ (ligne continue) du nombre mensuel moyen des pannes θ pour l'exemple du superordinateur.

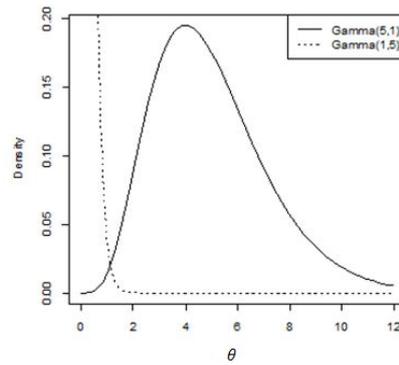


FIGURE 2.2 – Densités a priori de Gamma assez différentes.

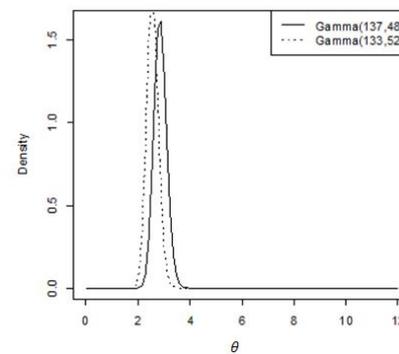


FIGURE 2.3 – Distributions a posteriori de Gamma pour les a priori de la figure 2.2.

Avec un échantillon aussi important, deux personnes ayant des opinions a priori plutôt divergentes s'accorderaient plutôt bien. Par exemple, supposons que la première personne choisira $(\alpha; \beta) = (5; 1)$ alors que l'autre choisit $(1; 5)$ (comme le montre la Figure 2.2). La première a posteriori est Gamma $(137; 48)$ tandis que la deuxième est Gamma $(133; 52)$, les deux étant représentées sur la Figure 2.3.

Cela démontre une autre caractéristique des méthodes Bayésiennes : Même des chercheurs ayant des croyances a priori extrêmement différentes peuvent finalement arriver à se mettre d'accord une fois que suffisamment de données se sont accumulées.

2.2.4 Point de départ : Choix de l'a priori

Pour utiliser la formule séquentielle, il nous faut une distribution a priori comme point de départ. Supposons maintenant que nous ne possédions pas tout renseignement sur les échantillons a priori. Avant le premier échantillon, l'information a priori se traduira par le fait que nous ne voudrions privilégier aucune valeur du paramètre inconnu θ . Dans ce cas, il est naturel que nous prenions l'a priori P comme une mesure uniforme sur l'espace mesurable (Θ, A) .

Malheureusement, lorsque Θ est un ensemble de mesure non finie, la mesure uniforme est une distribution impropre, puisque la probabilité totale (ou la masse) est infinie : $\int_{\Theta} P(\theta) d\theta = +\infty$

En cas unidimensionnel, et sans perdre de généralité, nous prendrons : $\Theta =]a; b[$ avec a et $b \in R \cup \{-\infty, +\infty\}$; la formule séquentielle devient :

$$p(\theta|x_1, x_2, \dots, x_n) = \lim_{\substack{x \rightarrow a \\ y \rightarrow b}} \frac{\frac{1}{y-x} \prod_j f(x_j|\theta)}{\int_x^y \frac{1}{y-x} \prod_j f(x_j|\theta) d\theta}$$

$$p(\theta|x_1, x_2, \dots, x_n) = \lim_{\substack{x \rightarrow a \\ y \rightarrow b}} \frac{\prod_j f(x_j|\theta)}{\int_x^y \prod_j f(x_j|\theta) d\theta}$$

$$\propto \prod_j f(x_j|\theta)$$

Bien entendu, nous devons supposer l'existence des limites mentionnées dans les expressions. Cette condition est vérifiée pour la famille des

distributions exponentielles, ce qui est le cas en général. Cette formule peut se généraliser dans le cas multidimensionnel [Zhu, 1991].

Ces distributions $\{p_i(\theta) = p(\theta|x_1, x_2, \dots, x_i); 1 \leq i \leq n\}$ appartiennent à une famille ξ de distributions conjuguées par rapport à la famille de fonctions de vraisemblance $\xi = \{f(x|\theta); x \in \mathbb{R}^k, k = 1, 2, \dots \text{ et } \theta \in \Theta\}$. En effet, un échantillon composé des n échantillons $x = x_1 \cup x_2 \cup \dots \cup x_n$. Compte tenu du fait que :

$$f(x_i|\theta) = \prod_j f(x_i^j|\theta); \quad j = 1, 2, 3, \dots, k_i$$

$$p_i(\theta) \propto f(x_i|\theta) \propto \prod_j f(x_i^j|\theta)$$

$$p_n(\theta) \propto \prod_i f(x_i|\theta) \propto \prod_i \prod_j f(x_i^j|\theta) = f(x|\theta)$$

d'où $\xi = \{f(x|\theta); x \in \mathbb{R}^k, k = 1, 2, \dots \text{ et } \theta \in \Theta\}$

Si la distribution sous-jacente $f(x|\theta)$ admet une statistique exhaustive (sous-entendu dans un modèle dominé) : $f(x|\theta) = g(T(x)|\theta)$ pour un échantillon quelconque x , alors il existe une statistique exhaustive pour l'échantillon composé (ou total) des n échantillons a priori : $x = x_1 \cup x_2 \cup \dots \cup x_n$, qui nous permet d'écrire à la fois les deux formules suivantes :

$$p_n(\theta) \propto \prod_i f(x_i|\theta) \propto \prod_i g(T(x_i)|\theta)$$

$$p_n(\theta) \propto f(x|\theta) \propto g(T(x)|\theta) = g(T(x_1 \cup x_2 \cup \dots \cup x_n)|\theta)$$

Pour la famille exponentielle, on obtient l'égalité suivante :

$$g(\tilde{T}(x_1 \cup x_2 \cup \dots \cup x_n)|\theta) = \prod_i g(\tilde{T}(x_i)|\theta) = g\left(\sum_i \tilde{T}(x_i)|\theta\right)$$

d'où l'opération de conjugaison linéaire :

$$\tilde{T}(x_1 \cup x_2 \cup \dots \cup x_n) = \sum_i \tilde{T}(x_i)$$

Nous obtenons donc la même forme de distribution a priori que par la méthode de la conjuguée naturelle (Section 2.2.1), ce qui en renforce l'intérêt.

2.3 Les essais séquentiels par groupe

Les analyses intermédiaires peuvent conduire à arrêter l'étude avant l'inclusion du nombre de sujets initialement prévus. Des plans expérimentaux ont été développés à partir de l'approche fréquentiste, autorisant des arrêts précoces pour efficacité, futilité ou toxicité. Dans l'approche Bayésienne, les résultats des analyses successives contribuent à la mise à jour des probabilités des hypothèses H_0 ou H_1 et les conséquences des observations à venir sur les résultats de fin d'étude sont pondérées par leurs probabilités Bayésiennes prédictives. Cette approche fournit une réponse pragmatique aux différentes questions posées. Quelle estimation de la probabilité a posteriori de l'hypothèse d'intérêt emporterait la conviction ? Quels résultats obtenus durant l'étape à venir seraient compatibles avec cette estimation ? Quelle est la probabilité de ces résultats compte tenu de l'information disponible ? L'approche Bayésienne s'appuie ici sur le calcul de la somme des probabilités prédictives des résultats de la seconde étape qui conduiraient à une estimation de la probabilité de l'hypothèse d'intérêt supérieure à une valeur seuil prédéfinie ; une fois l'étude terminée [Kramar et Mthoulin-Pélissier, 2011].

2.3.1 Les essais de phase I

Les essais de phase I sont généralement réalisés en monochimiothérapie, mais peuvent aussi servir à évaluer des combinaisons de médicaments déjà connus. Il s'agit de protocoles dont le but est :

- de déterminer les profils de toxicité et de sécurité d'emploi d'un nouvel agent ou d'une nouvelle combinaison d'agents connus.
- de décrire le profil pharmacocinétique du médicament après son administration.
- d'établir une dose optimale à administrer pour les études de phase II en découlant.

Dans le cas d'un agent cytotoxique, pour lequel le mécanisme d'action n'est pas spécifique, la toxicité est un élément prépondérant à étudier afin

d'établir des courbes dose-toxicité. Il faudra donc dans un premier temps définir la toxicité dose-limitante (TDL). Celle-ci pourra être constituée par un seul événement toxique ou par une combinaison d'événements ; par exemple, « tout événement toxique irréversible de grade supérieur ou égal à 2 » ou « thrombocytopenie de grade 4 pendant plus d'une semaine ou neutropénie avec nombre de leucocytes inférieur à 1 000 pendant plus de 2 semaines ou neutropénie fébrile durant plus d'une semaine ». Une fois la TDL déterminée, le principal objectif sera d'estimer la dose maximale tolérée (DMT), définie comme la dose induisant un risque prédéfini de toxicité sévère c-à-d une probabilité prédéfinie de présenter une toxicité dose-limitante. Il s'agit donc d'un protocole d'escalade de doses (en général monochimiothérapie) pour lequel il faut fixer la dose de départ et le mode d'escalade de doses [Tranchand et al., 2008].

En dehors des valeurs de DMT basses, le schéma d'escalade de dose non paramétrique classique 3+3 apparaît moins performant que la méthode de réévaluation séquentielle (continual reassessment method, CRM) pour estimer la DMT [O'Quigley et al., 1990, Kramar et Mthoulin-Pélissier, 2011]. Cette méthode repose sur une modélisation paramétrique de la relation dose-toxicité avec une attribution de dose et une estimation séquentielle du paramètre du modèle ; jusqu'à l'inclusion d'un nombre de sujets fixe et préétabli ou jusqu'à ce qu'une règle d'arrêt soit vérifiée [Kramar et Mthoulin-Pélissier, 2011]. La CRM a fait et fait encore aujourd'hui l'objet de nombreux travaux de recherche, en particulier pour l'estimation de la fonction dose-risque sous-jacente [Daimon et al., 2011].

2.3.2 Les essais de phase II

Les essais de phase II jouent un rôle important dans le processus de développement d'un nouveau traitement. L'efficacité thérapeutique est évaluée en phase II chez des patients recevant la DMT ; considérée comme la dose la plus efficace. L'utilisation fréquente de polychimiothérapies a justifié le développement de nouveaux schémas d'étude. En effet, seules les molécules ayant montré une efficacité anti tumorale suffisante à

l'issue de cette phase justifieront la poursuite de leur développement dans des essais comparatifs de phase III, essais à beaucoup plus grande échelle nécessitant de gros moyens. Les plans expérimentaux de phase II doivent donc être suffisamment performants pour éliminer rapidement les traitements inefficaces en utilisant au mieux les ressources disponibles, tout en minimisant le risque d'écarter une molécule ayant un effet anti-tumoral potentiellement intéressant. [Huang et al., 2007] ont ainsi proposé d'évaluer simultanément la sécurité et l'efficacité d'associations thérapeutiques dans un essai combiné de phase I-II. Ce schéma adaptatif inclut une phase initiale d'escalade de dose, puis une randomisation adaptative des combinaisons de doses admissibles. La randomisation adaptative privilégie les niveaux de dose les plus efficaces à partir des valeurs des probabilités a posteriori. L'essai est arrêté si la probabilité a posteriori de toxicité, d'efficacité ou de futilité dépasse un seuil préalablement défini.

Comparé à la mise en œuvre de deux essais distincts, le schéma combiné est plus puissant permet d'inclure moins de patients et alloue davantage de patients à des doses plus efficaces. Le risque de négliger d'éventuelles toxicités tardives a cependant été soulevé récemment et l'utilisation de ces méthodes doit être envisagée en fonction des situations spécifiques.

Les essais de phase II ont très souvent été élaborés à l'aide de plans expérimentaux à un seul bras (bras expérimental) où le taux de succès observés est comparé à un taux théorique ; sur un nombre limité de patients et en utilisant un critère d'évaluation binaire observable à court terme (typiquement, la réponse tumorale). Récemment, de nouvelles méthodes ont été utilisées afin de rendre cette phase importante du développement encore plus efficiente [Tranchand et al., 2008, Kramar et Mathoulin-Pélissier, 2011].

2.3.3 Les essais de phase II-III

Il se passe quelquefois un certain laps de temps entre la fin d'une étude de phase II ayant montré une efficacité anti-tumorale par exemple

et le début d'une étude de phase III. Le principe des plans de phase II-III est de réduire ce temps entre les essais en utilisant les données des patients inclus dans la phase II pour l'essai de phase III. Cette approche est possible en utilisant les critères d'arrêt précoce issus des essais de phase II dans des essais de phase III, ce qui pourrait être assimilable au principe d'une analyse intermédiaire. Un des intérêts de ce type de plan est de réaliser un essai de la phase II avec n bras expérimentaux [Tranchand et al., 2008], à chaque évaluation, la décision d'interrompre l'essai pour efficacité, de l'interrompre pour futilité ou de passer en phase III dépend des estimations des probabilités prédictives de sélectionner le plus efficace en tant que bras expérimental dans un essai comparatif de phase III. Le mélange des modèles paramétriques utilisé autorise des différences d'effets du traitement sur le taux de mortalité en fonction de l'obtention d'un contrôle local de la tumeur [Kramar et Mathoulin-Pélissier, 2011].

Très souvent, en effet, la réponse tumorale utilisée comme critère d'efficacité en phase II est abandonnée au profit de la survie globale ou sans progression en phase III [Tranchand et al., 2008].

Des simulations comparant le schéma proposé à une approche séquentielle groupée usuelle en phase III indiquent le gain de l'approche combinée en terme de durée d'étude et de nombre de sujets inclus. À côté de ce schéma complètement Bayésien, d'autres schémas « hybrides » pour des essais de phase II-III ont été proposés avec, par exemple, une analyse intermédiaire Bayésienne mais une analyse finale « fréquentiste » afin de prouver le contrôle du risque d'erreur de première espèce qui est une condition nécessaire imposée par les autorités de régulation pour les essais pivot [Schmidli H et al., 2007].

L'utilisation des méthodes Bayésiennes a aussi été proposée dans un schéma de phase II-III pour des thérapies ciblées en cancérologie ; avec sélection d'une sous-population à l'analyse intermédiaire [Brannath et al., 2009]. Lors de l'analyse finale, les tests sont ajustés de façon à contrôler le risque d'erreur de première espèce de l'ensemble de l'essai [Kramar et Mathoulin-Pélissier, 2011].

2.3.4 Les essais de phase III

L'essai de phase III est l'étude comparative d'efficacité proprement dite. Son but est de tester l'efficacité d'un nouveau traitement par rapport au meilleur traitement connu qui est le traitement de référence. C'est dans cette phase de développement que le rapport bénéfice/tolérance est estimé en situation réelle. Les études de phase III sont de grande envergure. Elles impliquent une centaine à plusieurs milliers de patients. Par rapport à la phase II, les durées d'exposition au médicament sont prolongées et une évaluation des interactions médicamenteuses peut être effectuée. C'est à l'issue de cette phase qu'une autorisation de mise sur le marché (AMM) pourra être donnée par les autorités de santé. Si les objectifs d'un essai de phase III sont bien définis, la méthodologie reste basée sur la minimisation des biais [Cucherat et al., 2004]. Cela conduit à réaliser des essais prospectifs, comparatifs, randomisés, si possible en double aveugle, avec un critère principal unique défini clairement, ainsi qu'à une analyse statistique réalisée en intention de traiter [Tranchand et al. 2008].

2.4 Prédiction séquentielle

Jusqu'à présent, nous avons utilisé les informations disponibles pour procéder à une évaluation de la plausibilité du paramètre inconnu mais ce n'est pas notre objectif final. 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, ce qui est le fond de l'approche Bayésienne [Djeridi, 2019].

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 E :

Définition 2.1. (La densité prédictive). La fonction de densité prédictive $f(y|x)$ est définie par :

$$f(y|x) = \int f(y|\theta, x)p(\theta|x)d\theta$$

$$f(y|x) = \frac{\int_{\theta} f(y|\theta)p(\theta)f(x|\theta)d\theta}{\int p(\theta)f(x|\theta)d\theta}$$

Un avantage de l'approche bayésienne est la capacité de considérer les probabilités des résultats non encore observés. A titre d'illustration, nous prenons le lecteur à travers les étapes formelles de la dérivation de la fonction de densité prédictive pour le modèle Poisson-Gamma :

2.4.1 Le modèle Poisson-Gamma

Prenons l'exemple d'une variable Poisson(θ). Soit un ensemble de N objets dont chacun est caractérisé par l'absence ou la présence d'une propriété A (par exemple, une malfaçon dans la fabrication). Désignons par θ la proportion d'objets présentant la propriété A. Le problème est d'estimer la vraie valeur de θ .

Considérons une variable aléatoire X , qui représente le nombre d'objets ayant la propriété A parmi k objets tirés au sort indépendamment. Alors la variable X admet une distribution Poisson de paramètres θ , avec $\theta \in \mathbb{N}$.

Un résultat $x \in \{0, 1, \dots, n\}$ a pour probabilité $x \sim \text{Poisson}(\theta)$:

$$f(x|\theta) = \exp(-\theta) \frac{\theta^x}{x!}$$

L'inconnu est le taux θ avec $\theta > 0$, partir des n expériences a priori la question qui se pose est : Quelle inférence peut-on faire sur le taux θ avant d'envisager une nouvelle expérience ? Et après avoir obtenu un nouveau résultat x quelle inférence peut-on faire sur le taux ? Dans le contexte prédictif, on suppose qu'il est possible de prendre des observations supplémentaires y . Comment s'en servir ?

Nous savons que chaque expérience x_i a pour probabilité :

$$f(x_i|\theta) = \exp(-\theta) \frac{\theta^{x_i}}{x_i!}$$

Il s'agit, en effet, de choisir dans un premier temps, la distribution a priori du taux θ à partir des n expériences et de déterminer, en second lieu, la distribution du taux θ conditionnelle à x , comme la distribution Gamma est la distribution a priori conjuguée pour θ ; nous avons l'observation x_1 de $\Gamma(\alpha, \beta)$. Une simple application du Théorème de Bayes donne : Nous savons que chaque expérience x_1 a pour probabilité :

$$f(\theta|x_1) = (\theta^{\alpha x_1}) \frac{\beta^\alpha \exp(-\beta\theta)}{\Gamma(\alpha)} \quad (2.4)$$

Après la première expérience x_1 , nous obtenons la distribution a posteriori $p(\theta)$ de forme suivante :

$$f(\theta|x_1) \propto \theta^{\alpha+x_1-1} \exp(-\theta(\beta+1)) \quad (2.5)$$

Si on effectue des observations indépendantes de même loi, le premier résultat est une suite de t_1 observations de loi de Poisson de paramètre θ et on choisit comme loi a priori pour θ la loi Gamma $Ga(\alpha, \beta)$, où $\alpha > 0$ et $\beta > 0$. Pour des raisons d'exhaustivité on pose $x = \sum_{i=1}^{t_1} x_i$. Ce qui donne une distribution a posteriori séquentielle comme suit :

$$f(\theta|x_t) \sim Ga(\alpha + \sum_{i=1}^{t_1} x_i, \beta + t_1) \quad (2.6)$$

Après avoir acquis une nouvelle expérience x , la distribution a posteriori de θ est la $(t_1 + 1)$ -ième distribution de la séquence

$$\{f_i(\theta)\} : f(\theta|X = x_{t_1+1}) = f_{t_1+1}(\theta).$$

Soit y le résultat de l'expérience future y , pour y_1 la densité prédictive est donnée par :

$$f(y|x) = \int_0^\infty f(y|\theta) f(\theta|x) d\theta$$

$$f(y_1|x_1) = \frac{(\beta+1)^{\alpha+x_1} \Gamma(\alpha+x_1+y_1)}{(\beta+2)^{\alpha+x_1+y_1} \Gamma(\alpha+x_1) y_1!} \quad (2.7)$$

D'une autre manière; on pose $y = \sum_{i=1}^{t_2} y_{t_2}$, la densité prédictive est donnée par (pour plus de détail voir chapitre 4) :

$$f(y|x) = \frac{(t_2)^y (\beta + t_1)^{\alpha+x} \Gamma(\alpha + x + y)}{(\beta + t_1 + t_2)^{\alpha+x+y} \Gamma(\alpha + x) y!} \quad (2.8)$$

2.5 Tests séquentiels

Étant donné une population P sur laquelle on mesure un caractère qui possède une fonction de répartition F , étant donné des hypothèses qu'on veut vérifier sur la loi F , les tests statistiques classiques procurent un moyen de vérifier ces hypothèses en tirant de P suivant des modalités à observer ou plan d'échantillonnage, un échantillon E de taille donnée en se donnant un risque de se tromper dans le refus ou l'acceptation d'une hypothèse. L'analyse séquentielle (en anglais « sequential analysis ») part d'un autre point de vue :

Soit à choisir entre deux hypothèses H_0 et H_1 , peut-on construire un plan d'échantillonnage consistant à tirer, un à un, des individus de P et à s'arrêter quand on a l'information nécessaire pour choisir entre les deux hypothèses H_0 et H_1 ?

La réponse est oui. La méthode que nous allons étudier est, en général, moins coûteuse en tirages que le plan d'échantillonnage classique; elle possède, néanmoins, des particularités qui restreignent sensiblement l'usage. Découverte pendant la guerre, cette technique fut, paraît-il, à cause de ce gain de temps donc d'argent (qu'on avait peut-être surestimé au départ) décrété secret militaire. Elle a fait l'objet d'une somme de travaux théoriques très importante depuis 1945 et elle est appliquée très souvent dans le contrôle des fabrications industrielles aux États-Unis [Millier, 1967].

Depuis quelques années, elle a fait son apparition dans les sciences biologiques, tout naturellement en médecine, il paraît en effet intéressant d'arrêter une expérimentation de médicament (ou autre) dès qu'on « voit » statistiquement que l'on s'est trompé.

2.5.1 Inférence non-décisionnelle

Il paraît nécessaire de définir la nature d'un test séquentiel par opposition aux tests classiques et pour cela, nous donnons les généralités sur les tests séquentiels. Supposons que la loi de probabilité de la variable aléatoire X dépende d'un paramètre inconnu θ . 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 (2007), une règle d'arrêt τ peut être définie comme suit :

Si les expériences ξ_i produisent des observations $x_i \in \chi_i$, avec $x_i \sim f(x_i|\theta)$, considérons la suite correspondante $A_i \subset \chi_1 \times \dots \times \chi_i$ telle que le critère τ prend la valeur n si $(x_1, x_2, \dots, x_n) \in A_n$, i.e., l'expérience s'arrête après la $n^{\text{ème}}$ observation seulement si les n premières observations sont dans A_n . La vraisemblance de (x_1, x_2, \dots, x_n) est alors :

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

et donc dépend seulement de τ via l'échantillon x_1, x_2, \dots, x_n . Ceci implique le principe suivant :

Principe de la règle d'arrêt

Si une suite d'expériences, $\xi_1, \xi_2, \xi_3 \dots$ admet une règle d'arrêt, τ , qui implique quand doivent s'arrêter les expériences, l'inférence sur θ ne doit dépendre de τ qu'à travers l'échantillon résultant.

Caractéristiques d'un test séquentiel

Définition 2.2. La fonction OC (de l'anglais « operating characteristic »). Soit H_0 une hypothèse prédéterminée, et θ la valeur du (ou des) paramètre (s) de la loi de probabilité telle la probabilité d'accepter H_0 est évidemment une fonction de θ , l'OC est elle que :

$$\begin{aligned} Pr\{\text{accepter } H_0 \text{ si } \theta\} &= L(\theta); \\ Pr\{\text{refuser } H_0 \text{ si } \theta\} &= 1 - L(\theta). \end{aligned}$$

Définition 2.3. La fonction ASN (de l'anglais « average sample number »)

Dans le test séquentiel, le nombre de tirages n'est pas fixe, c'est une variable aléatoire qui admet des moments et en particulier une moyenne. Cette moyenne dépend de la valeur θ du (ou des) paramètre (s); c'est l'ASN = $E(N|\theta)$ (nombre moyen d'observations à prendre en compte).

Influence du test sur l'OC et E(N)

Soit à tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1$ on a :

$$L(\theta_1) = Pr\{\text{accepter } H_0 \text{ si } \theta = \theta_1\} = \beta$$

de même

$$L(\theta_0) = Pr\{\text{accepter } H_0 \text{ si } \theta = \theta_0\} = 1 - \alpha.$$

$L(\theta)$ est donc asymptotique à $L(\theta) = 1$ et $L(\theta) = 0$ et passe par A si l'échantillonnage est classique, par A et par B si l'échantillonnage est séquentiel.

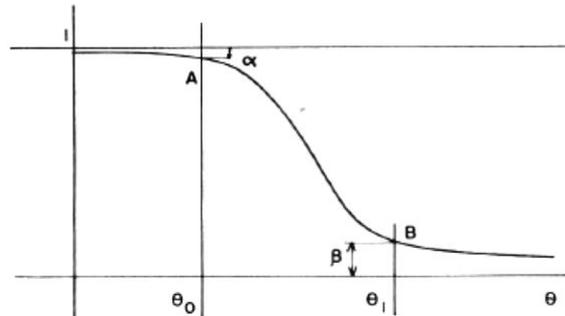


FIGURE 2.4 – Courbe OC

- Si l'échantillonnage est classique, le choix du test et le choix du nombre de tirages conditionnent la forme de la courbe et le point B .
- Si l'échantillonnage est séquentiel, le choix du test déterminera la forme de la courbe et le nombre moyen de tirages.

Quand nous comparons 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 (comparaison des courbes OC). Dans la suite, nous allons exposer les propriétés des tests séquentiels dans les essais cliniques de phase II à un seul bras.

2.6 Test séquentiels dans les essais cliniques à un seul bras

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 pour justifier la poursuite des études. Dans cette section, notre but est d'étudier les règles d'arrêt dans cette phase par démonstration de l'efficacité et la robustesse de l'inférence Bayésienne en utilisant la probabilité prédictive pour améliorer les conceptions (designs) de cette phase [Djeridi, 2019].

2.6.1 Cadre général

Une grande partie de la méthodologie statistique utilisée dans les essais avec un seul bras est dérivé de l'oncologie. De tels essais sont généralement des essais non contrôlés, à savoir un seul bras ; un seul traitement expérimental est évalué pour décider s'il a une efficacité suffisante pour justifier une étude plus approfondie. Dans cette section, nous nous concentrons sur les essais à un seul bras. Dans ce contexte, nous avons les hypothèses suivantes :

$$H_0 : \theta = \theta_0$$

- $H_0 : \theta \leq \theta_0$ où la vraie probabilité de réponse θ est inférieure à un certain niveau θ_0 inintéressant.
- $H_1 : \theta > \theta_1$ où la véritable probabilité de réponse est au moins une partie souhaitable θ_1 est un niveau cible.

Où

- θ_0 : est le taux de la réponse indésirable.
 - θ_1 : est le taux de la réponse désirable.
- tel que $\theta_1 > \theta_0$.

2.6.2 Design à une seule étape

Dans une conception en une seule étape 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]. Afin d'étudier l'efficacité d'un nouveau

médicament, Djeridi (2019) a donné une formule pour la taille de cette conception en une seule étape. Cette formule dépend du test choisi ; il existe deux méthodes différentes :

- La première est basée sur une approximation Normale.
- la seconde est basée sur les probabilités Binomiales exactes.

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

2.6.3 Design à deux étapes

Dans cette section, l'hypothèse nulle à tester est que la véritable probabilité de réponse (θ) est inférieure à un certain niveau θ_0 inintéressant : $H_0 : \theta \leq \theta_0$. L'hypothèse alternative est que la véritable probabilité de réponse est au moins une proportion souhaitable θ_1 (niveau cible) : $H_1 : \theta > \theta_1$. Lorsque H_0 ne peut être rejetée (on dit simplement « acceptée »), le médicament sera rejeté et ne nécessite pas une étude plus approfondie.

Afin de réduire le nombre de patients, en cas de signes précoces de (in) efficacité, une analyse intermédiaire est effectuée. Deux modèles différents sont examinés : Un modèle fréquentiste et une conception Bayésienne. Les caractéristiques communes de ces designs sont donnés ci-dessous.

Tout au long, on considère que n_1 , n_2 le nombre de patients dans la première et la seconde étape, respectivement, et x , y les nombres de réponses.

Après la première étape, lorsque les n_1 patients sont traités nous allons :

- H_0 si $x \leq \alpha_1$
- Rejeter H_0 si $x \geq r_1$
- continuer si $\alpha_1 < x < r_1$

En dernière analyse, lorsque $N_{max} = n_1 + n_2$ patients sont traités, nous allons :

- Accepter H_0 si $x + y \leq \alpha_2$
- Rejeter H_0 si $x + y \geq r_2$ avec $r_2 = \alpha_2 + 1$

Notons que dans l'analyse finale, une décision doit être prise, à savoir qu'il n'y a pas de nombre de réponses qui peuvent être mis à la fin de l'essai.

Chapitre 3

Assurance Bayésienne

3.1 Introduction

Un essai clinique bien conçu nécessite une taille d'échantillon appropriée avec une puissance statistique adéquate pour atteindre les objectifs de l'essai. La puissance statistique est traditionnellement définie comme la probabilité de rejeter l'hypothèse nulle avec un véritable effet de traitement clinique prédéfini c'est-à-dire la probabilité de rejeter la valeur nulle si le véritable effet du traitement de l'essai clinique est égal à une valeur a priori. Par conséquent, la puissance statistique est une probabilité conditionnelle à cette valeur de condition a priori inconnue comme discuté en détail dans [Chow et al., 2003], [Chen et Peace, 2011].

Cependant, la puissance d'un essai ne donne pas nécessairement la probabilité dont l'essai démontre un effet du traitement, car le véritable effet du traitement peut être différent de celui supposé dans le calcul de la puissance. Par conséquent, la puissance statistique traditionnelle utilisée pour « alimenter » un essai clinique ne peut pas réellement garantir sa réussite. Pour assurer la réussite d'un essai clinique plusieurs auteurs ont proposé une approche hybride Bayésienne-fréquentiste pour évaluer la probabilité d'un essai réussi, étant donné la taille de l'échantillon qui peut ensuite être utilisée pour éclairer les décisions.

La méthode hybride a été d'abord envisagée par [Spiegelhalter et

Freedman, 1986]. Ils ont construit une probabilité inconditionnelle pour avoir un résultat souhaité et ont appelé cette probabilité inconditionnelle la puissance moyenne. [O'Hagan et Stevens 2001] ont utilisé cette méthode pour choisir la taille des échantillons pour les essais cliniques de rentabilité. Ils ont qualifié la probabilité inconditionnelle de réussite de l'essai par l'assurance de l'essai, cette assurance peut alors être obtenue en tant que puissance attendue par rapport à la distribution de probabilité a priori de la valeur a priori (l'effet du traitement inconnu), ce qui conduit au paradigme Bayésien, elle est dite aussi La puissance attendue Bayésienne « en anglais The Bayesian expected power (BEP) ». [O'Hagan et al., 2005] ont étendu les méthodes d'assurance aux tests bilatéraux et aux essais d'équivalence, couvrant l'utilisation de distributions a priori non conjuguées pour des paramètres incertains. [Chuang-Stein 2006] a discuté la différence entre les calculs de puissance traditionnels et les calculs d'assurance pour déterminer la taille des échantillons, donnant un exemple de planification du prochain essai basé sur les résultats d'un essai précoce. [Chuang-Stein et Yang 2010] ont examiné le concept d'assurance et ont illustré son utilisation lors de la planification des essais de phase III. Ils ont également appliqué l'assurance d'étudier les plans lors de la ré-estimation de la taille d'un échantillon sur la base d'une analyse intermédiaire.

Un calcul d'assurance nécessite une distribution a priori de l'effet du traitement mais n'implique pas nécessairement une analyse Bayésienne des données d'essai. La méthode d'analyse, et en particulier les critères pour lesquels l'essai est considéré comme un « succès », sont déterminés en externe, par exemple, par un régulateur. Une fois les critères spécifiés, une distribution a priori est utilisée afin d'évaluer la probabilité pour que ces critères soient remplis. En règle générale, la distribution a priori ne sera utilisée qu'à l'étape de la conception et non de l'analyse. À l'étape de la conception, le risque d'échec de l'essai est principalement celui du sponsor de l'essai, il ne devrait donc pas être controversé pour un sponsor d'essai d'utiliser toutes ses connaissances a priori pour évaluer un tel risque [Ren et Oakley, 2013].

Nous considérons les essais cliniques dans lesquels le point final d'intérêt est un temps de survie : Pour les mesures des résultats dans le temps jusqu'à l'événement, les calculs de puissance et de taille d'échantillon ont été bien étudiés sous diverses hypothèses de modèle. Par exemple, [Schoenfeld et Richter, 1982] ont développé une fonction de puissance avec une période de recrutement limitée et une période de suivi pré-spécifiée en supposant que les temps de survie dans chaque groupe de traitement suivent des distributions exponentielles et que les patients entrent dans l'essai de manière uniforme. [Gross et Clark, 1975] ont fourni une méthode de calcul de la taille de l'échantillon en supposant que le temps de survie moyen de l'échantillon est approximativement normalement distribué sous les modèles de Weibull pour les temps de survie. [Spiegelhalter et al. 2004] ont établi une formule d'assurance en cas d'égalité de répartition et de suivi. La seule variable incertaine considérée était le rapport de risque logarithmique et un a priori normal a été supposé.

L'assurance, en tant qu'alternative à l'important concept de puissance statistique, est encore nouvelle pour de nombreux biostatisticiens, cliniciens et régulateurs gouvernementaux. D'autres illustrations de ce concept ainsi que la mise en œuvre de logiciels à usage public restent un besoin pratique non satisfaisant. Dans ce chapitre, nous décrivons la transition de la puissance statistique conventionnelle à l'assurance nouvellement développée, nous examinons comment l'assurance est calculée pour déterminer la probabilité inconditionnelle d'avoir un résultat souhaité et enfin nous discutons les calculs d'assurance par des exemples donnés.

3.2 Taille d'échantillon des essais non séquentiels

Dans cette section, nous considérons la contribution Bayésienne à la sélection de la taille de l'échantillon d'un essai clinique qui ne fera

pas l'objet d'un suivi intermédiaire : L'accent est mis en particulier sur les méthodes « hybrides » dans lesquelles des informations a priori sont formellement utilisées mais l'analyse finale est effectuée en un cadre classique. Dans certains contextes, cela peut être tout à fait approprié car il peut y avoir des informations a priori substantielles qui ne peuvent pas être incluses dans le rapport final à des fins réglementaires, par exemple. Cette section contient un certain nombre d'expressions assez complexes pour les quantités d'intérêt, les formules que nous présentons suivent la formulation traditionnelle dans laquelle l'intérêt se concentre sur un paramètre θ et $\theta > 0$ indique le bénéfice du traitement expérimental. Nous reconnaissons que dans beaucoup des exemples, $\theta < 0$ a représenté un tel avantage, et en outre, dans d'autres cas, nous pourrions utiliser des seuils autres que 0. Il faut donc être prudent lors de l'utilisation des formules, il peut être préférable de transformer d'abord le problème particulier étant analysé dans la formulation standard [Spiegelhalter et al., 2004 voir section 2.4].

On pourrait faire valoir que la recherche des croyances et des demandes a priori d'une large communauté de parties prenantes est nécessaire non seulement pour entreprendre une approche spécifiquement Bayésienne de la conception et de l'analyse mais aussi plus généralement dans le cadre de bonnes pratiques de recherche. Une conséquence potentielle de l'ignorance de cette source de jugement est que les essais peuvent être conçus sur la base des croyances et des demandes trop enthousiastes et ne parviennent donc pas à convaincre les autres et à modifier la politique ou la pratique des soins de santé.

3.2.1 Approches alternatives à l'évaluation de la taille de l'échantillon

Nous nous concentrons sur la façon dont les quatre points de vue principaux (en ignorant les approches Bayésienne de test d'hypothèse et les approches classiques de la théorie de décision) traitent la sélection de

la taille de l'échantillon d'une expérience de taille fixe, la conception et le suivi des études séquentielles. Une philosophie hybride est également incluse.

Fisherian : En principe, il n'est pas nécessaire de prévoir des tailles d'échantillon pré-planifiées mais un choix peut être fait en sélectionnant une précision de mesure particulière et en l'échangeant de manière informelle contre le coût de l'expérimentation.

Neyman - Pearson : La première étape consiste à définir une hypothèse nulle puis à spécifier une autre hypothèse $H_A : \theta = \theta_A$ que l'essai est conçu pour détecter. Diverses opinions ont été exprimées sur l'interprétation de θ_A (Spiegelhalter et al., 2004), notamment une différence « minimale cliniquement significative », une différence « valable » et une différence « pensée susceptible de se produire ». Ces idées ont tendance à confondre les exigences du nouveau traitement et les attentes de son bénéfice. Ce rôle combiné de l'alternative se reflète dans sa définition commune comme une différence « à la fois réaliste et importante » (dans un cadre Bayésien, ces propriétés sont clairement séparées).

La taille de l'échantillon est ensuite sélectionnée pour avoir une puissance raisonnable afin de détecter cette hypothèse alternative. La puissance est généralement réglée à 80% ou 90% :

La formule $n = (z_{1-\beta} - z_\epsilon)^2 \frac{\sigma^2}{\theta_A^2}$ peut être utilisée pour dériver la taille d'échantillon nécessaire dans des circonstances simples. Dans la pratique, le choix de l'alternative peut être influencé par les ressources disponibles.

Hybride Bayésienne-classique : Un intérêt considérable a été accordé à une approche hybride dans laquelle il est supposé qu'une analyse traditionnelle aura lieu à la fin de l'essai et que la distribution a priori est utilisée uniquement pour la conception.

Il peut être utile de considérer la distribution de probabilité conjointe des hypothèses et des résultats présentés dans le tableau 3.1. Dans un cadre traditionnel, ce sont des hypothèses ponctuelles et l'étude est conçue autour de l'erreur de type $\alpha = p(D_1|H_0)$ et de la puissance $1 - \beta = p(D_1|H_1)$.

TABLE 3.1 – Distribution de probabilité conjointe des hypothèses et des résultats d'un test d'hypothèse.

Résultat		Vérité		$p(D_0)$
		H_0	H_1	
D_0 : ne pas rejeter	H_0	$p(D_0, H_0) =$ $p(\text{exact négatif})$	$p(D_0, H_1) =$ $p(\text{faux négatif})$	
D_1 : rejeter	H_0	$p(D_1, H_0) =$ $p(\text{faux positif})$	$p(D_1, H_1) =$ $p(\text{exact positif})$	$p(D_1)$
		$p(H_0)$	$p(H_1)$	1

Cependant, si nous sommes prêts à reconnaître les probabilités a priori pour les hypothèses, il semblerait alors raisonnable de se concentrer également sur la probabilité de rejeter H_0 et que ce soit la bonne décision, c'est-à-dire la probabilité conjointe $p(D_1, H_1)$.

Puisque $p(D_1, H_1) = p(D_1|H_1)p(H_1) = (1 - \beta)p(H_1)$, cela signifie simplement ajuster la puissance par la probabilité initiale de H_1 : Le problème lié à l'utilisation de la seule puissance conditionnelle $p(D_1|H_1)$ est qu'il n'est pas tenu compte de la plausibilité de l'alternative et qu'il y a donc une tentation à de se tromper dans la conception d'essais pour détecter des hypothèses invraisemblables.

La probabilité inconditionnelle d'obtenir une conclusion « positive » peut être exprimée comme suit :

$$p(D_1) = p(D_1, H_0) + p(D_1, H_1)$$

et le premier terme, qui est la probabilité $p(D_1, H_0) = p(D_1|H_0)p(H_0)$ d'un résultat faussement positif, sera généralement très petit à condition que $\alpha = p(D_1|H_0)$ est faible et l'opinion a priori est en faveur de H_1 (comme ce sera souvent le cas avant un essai). Donc :

$$p(D_1) \approx p(D_1|H_1)p(H_1)$$

et donc la « puissance a priori ajustée » $(1 - \beta)p(H_1)$ sera souvent proche de la probabilité inconditionnelle dont l'essai a obtenu un résultat « significatif ». Les choses se compliquent un peu dans le cas le plus général lorsque les hypothèses sont composites, par exemple $H_0 : \theta < 0$ et $H_A : \theta > 0$. Ici la puissance classique est donnée par une courbe $p(D_1|\theta)$, et nous souhaitons utiliser une distribution a priori continue $p(\theta)$ [Spiegelhalter et al., 2004].

Un certain nombre de moyens d'incorporation de l'a priori sont possibles.

1. On peut tracer la courbe de puissance conditionnelle et superposer la distribution a priori comme un guide informel de la plausibilité relative des hypothèses alternatives. Cela pourrait empêcher qu'une étude soit conçue autour d'une alternative très optimiste.
2. La moyenne a priori μ pourrait simplement être considérée comme une hypothèse alternative ponctuelle θ_A , représentant une différence « plausible et valable » bien que cela ne reconnaisse pas l'incertitude actuelle sur θ exprimée par l'a priori.
3. Toute la courbe de la puissance classique $p(D_1|\theta)$ peut être moyennée par rapport à la distribution a priori pour obtenir une puissance classique « attendue » ou « moyenne » $p(D_1) = \int p(D_1|\theta)p(\theta)d\theta$. Cela donnera la probabilité inconditionnelle de rejeter H_0 . D'après la discussion ci-dessus, nous pouvons nous attendre à ce qu'il s'agisse d'une approximation raisonnable de la puissance ajustée a priori $p(D_1, H_1)$ si $p(\theta)$ ne donne pas une probabilité substantielle à des valeurs $\theta < 0$.
4. La courbe de la puissance classique peut être moyennée par rapport à la distribution a priori $p(\theta|H_1) = p(\theta|\theta > 0)$, c'est-à-dire conditionnelle à H_1 étant vraie (puisque $p(\theta|\theta > 0) = p(\theta, \theta > 0)/p(\theta > 0)$), ceci peut être obtenu en restreignant l'a priori à $\theta > 0$ et en le re-normalisant pour avoir la probabilité totale 1). Brown et al. (1987) recommandent cette technique comme prédisant la chance de détecter correctement une amélioration positive, plutôt que la chance

globale $p(D_1)$ d'obtenir un résultat positif quelle que soit la vérité. Mais cette méthode souffre de la même difficulté que le calcul de puissance classique d'origine, en ce sens qu'il n'est pas tenu compte de la plausibilité de H_1 .

5. La distribution prédictive sur les puissances possibles pourrait être affichée pour aider à décider des tailles d'échantillon appropriées.

Nous illustrerons ces options dans les sections suivantes en utilisant des probabilités normales et des a priori.

3.3 Puissance statistique conventionnelle à l'assurance

3.3.1 La puissance statistique conventionnelle et ses limites

Généralement, l'objectif général d'un essai clinique est de comparer si un nouveau médicament est meilleur qu'un placebo. Afin de démontrer l'efficacité du nouveau médicament, il faut déterminer le nombre de patients devraient être inscrits à chaque traitement. En termes statistiques, l'hypothèse nulle H_0 est définie comme les deux traitements n'étant pas différents par rapport à l'hypothèse alternative H_A qui est le nouveau médicament étant meilleur que le placebo. Le test d'hypothèse consiste ensuite à tester s'il existe un effet de traitement statistiquement significatif entre ces deux traitements. Les concepts associés pour ce test d'hypothèse sont l'erreur de type I et l'erreur de type II. L'erreur de type I (α , généralement contrôlée à 5%) est définie comme la probabilité de rejeter l'hypothèse nulle lorsqu'elle est vraie et l'erreur de type II (β) est définie comme la probabilité de ne pas rejeter l'hypothèse nulle lorsqu'elle est fausse.

La puissance statistique (π) est ensuite définie comme la probabilité

de rejeter l'hypothèse nulle lorsqu'elle est fautive (c'est-à-dire $\pi = 1 - \beta = p(D_1|H_1)$), qui est généralement définie entre 0,8 et 0,9. La taille d'échantillon associée peut ensuite être déterminée en fonction de cette puissance et de l'erreur de type I).

Supposons qu'un essai contrôlé randomisé soit mené pour comparer un traitement expérimental et un traitement standard pour une maladie particulière. Un test d'hypothèse doit être effectué pour tester l'hypothèse nulle de l'effet du traitement $\theta = 0$ contre l'alternative que $\theta \neq 0$. Suivant les notations d'O'Hagan et al. (2005), nous notons R comme l'événement de rejet de l'hypothèse nulle. La définition de la puissance statistique conventionnelle est alors :

$$\begin{aligned}\pi(\theta) &= P(R|\theta) \\ &= P(\text{rejeter } H_0 | \theta = \theta_A)\end{aligned}\tag{3.1}$$

La puissance du test $\pi(\cdot)$ fournit la probabilité de rejeter avec succès l'hypothèse nulle si la vraie valeur de θ est la spécifiée θ_A . Comme la vraie valeur de θ peut être très différente de θ_A , la probabilité réelle de rejeter avec succès l'hypothèse nulle peut être très différente de la puissance. On peut voir que la puissance statistique définie dans l'équation 3.1 est une probabilité conditionnelle de R conditionnée au vecteur de paramètre inconnu θ . La valeur de cette puissance ainsi que le calcul de la taille de l'échantillon associé dépendent alors du vecteur de paramètre inconnu θ [Chen et Ho, 2016].

Généralement, ce vecteur de paramètre ne peut pas être fourni précisément dans les essais cliniques pratiques comme indiqué dans [O'Hagan et al., 2005] et autres. Par conséquent, la puissance statistique, en tant que l'un des concepts les plus importants des essais cliniques, a traditionnellement été citée comme une probabilité fixe basée sur une valeur de paramètre pré-requis de l'espace de paramètres d'hypothèse alternative inconnue. Il est rare que les données observées coïncident avec la valeur du paramètre pré-requis, ce qui conduit souvent à un problème de sur-

alimentation « en anglais : over-powering » ou de sous-alimentation « en anglais : under-powering » d'un essai clinique [Chen et Ho, 2016].

3.4 Assurance dans les essais cliniques

Pour éliminer ces limitations de la puissance statistique conventionnelle, O'Hagan et Stevens (2001) ont recommandé « l'assurance » notée par γ comme une alternative à cette puissance statistique, qui est définie comme une probabilité inconditionnelle de rejeter l'hypothèse nulle, c'est-à-dire la probabilité inconditionnelle que l'essai se termine avec le résultat désiré ($\gamma = P(R)$) où R est le rejet de l'hypothèse nulle. L'assurance peut alors être considérée comme la puissance attendue de l'espace vectoriel des paramètres de θ , que nous dérivons via

$$P(\textit{essai réussi}) = \int P(\textit{essai réussi}|\theta)P(\theta)d\theta$$

$$\gamma = P(R) = \int P(R, \theta)d\theta = \int P(R|\theta)P(\theta)d\theta = E_{\theta}(P(R|\theta)) \quad (3.2)$$

où l'espérance est la distribution de probabilité (a priori) de l'espace vectoriel des paramètres de θ et $P(\theta)$ est la distribution a priori du véritable effet du traitement θ . Avec cette définition, l'assurance fournit un pont entre l'approche fréquentiste de la puissance statistique et le paradigme Bayésien en moyenne ou intégration de la puissance statistique conditionnelle avec toutes les valeurs (a priori) possibles de l'espace vectoriel des paramètres de θ . Cette assurance peut alors fournir une probabilité inconditionnelle ou des preuves pour évaluer le succès d'un essai clinique et elle est donc plus réaliste et robuste que celle de la puissance statistique conventionnelle. Si un essai réussi correspond simplement au rejet d'une hypothèse nulle d'absence d'effet du traitement, alors l'assurance en 3.2 peut être considérée comme une puissance attendue (interprétant θ_A en 3.1 comme la vraie valeur de l'effet du traitement, plutôt qu'une différence minimale cliniquement pertinente).

Si notre résultat souhaité est de rejeter l'hypothèse nulle avec des données favorisant le traitement expérimental, alors l'assurance est donnée par :

$$\begin{aligned}\gamma &= P(\text{'essai réussi'}) \\ &= P(\text{rejeter } H_0, \hat{\theta} > 0) \\ &= \int P(\text{rejeter } H_0, \hat{\theta} > 0 | \theta) f(\theta) d\theta\end{aligned}$$

avec $\hat{\theta} > 0$ indiquant que les données favorisent le traitement expérimental. Nous soulignons à nouveau que la spécification de ce qui constitue un « essai réussi » ne fait pas partie de la méthode d'assurance ; les critères déterminant la réussite d'un essai sont définis en externe et l'idée d'assurance est d'utiliser des informations a priori pour déterminer la probabilité que ces critères soient remplis. La puissance d'un essai clinique peut, en théorie, être rendue aussi grande que souhaité en augmentant la taille de l'échantillon. Il n'en va pas de même pour l'assurance. Pour un échantillon suffisamment grand, nous « observerons » le véritable effet du traitement, de sorte que l'assurance converge vers la probabilité a priori que le nouveau traitement soit suffisamment efficace et la variance a priori approche de zéro, l'assurance définie dans l'équation 3.2 se rapprocherait de la puissance statistique conventionnelle définie dans l'équation 3.1.

D'un autre côté, si cette probabilité a priori est faible, la variance a priori serait grande et l'assurance définie dans l'équation 3.1, qui fait la moyenne de toutes les valeurs potentielles de cette vague distribution a priori ; serait plus appropriée que la puissance statistique conventionnelle pour évaluer la probabilité d'un essai réussi.

3.4.1 Incertitude des paramètres

L'utilisation de l'assurance à l'étape de la conception implique une perspective Bayésienne car elle nécessite la spécification d'une distribu-

tion a priori pour les paramètres inconnus d'une population. La distribution a priori exprime l'incertitude a priori quant à la valeur vraie fixée mais inconnue de θ correspondant à l'effet du traitement spécifié (et tout autre paramètre requis) et décrit la plausibilité relative des différentes valeurs de paramètre. La probabilité a priori que, disons, un effet du traitement dépasse une certaine valeur δ_0 représente un degré de confiance que cette valeur critique sera dépassée.

L'adoption d'une perspective Bayésienne à l'étape de la conception est tout à fait raisonnable car la conception d'un essai est la décision du promoteur « en anglais : sponsor's decision » et son résultat est le risque du promoteur, la conception doit donc prendre en compte toutes les informations dont dispose le promoteur pour augmenter les chances que l'essai aboutisse à un résultat satisfaisant. En général, nous supposerons qu'à l'étape de l'analyse, l'analyse sera menée sur la base de tests de significations classiques ou d'intervalles de confiance sans incorporer les informations a priori aux données de l'essai, y'on a aussi des cas où nous indiquons comment l'assurance peut également être utilisée conjointement avec une analyse Bayésienne des données de l'essai. La distribution a priori est utilisée simplement pour déterminer la chance d'obtenir des résultats d'essai pertinents [O'Hagan et al., 2005].

Si les informations a priori pour l'effet du traitement inconnu sont suffisamment solides, alors la taille d'échantillon requise peut être nulle parce que les informations a priori suffisent à elles seules à produire un résultat positif sans avoir besoin de données d'échantillonnage supplémentaires, il est clairement indésirable et irréaliste d'exprimer des informations a priori aussi solides, en particulier lorsque l'intention est de générer des données à soumettre à une autorité de réglementation pharmaceutique. D'un autre côté, si les informations a priori sont faibles, même dans des tailles d'échantillon limitées, il peut ne pas suffire pour obtenir l'assurance souhaitable d'un résultat positif. En fait, la distribution a priori fournit une limite supérieure sur l'assurance réalisable d'un résultat positif, contrairement à la puissance où il est possible de générer n'importe quelle valeur souhaitable avec une spécification appropriée des

hypothèses.

La spécification de la distribution a priori peut être un peu plus qu'un jugement approximatif, auquel les assurances dérivées seront également approximatives mais peuvent néanmoins avoir une valeur réelle. Lorsque les assurances sont importantes, peut-être parce que l'essai clinique est suffisamment grand pour qu'il y ait un risque financier s'il a une faible probabilité d'aboutir à un résultat positif, ou parce qu'il contribue aux décisions stratégiques concernant la poursuite du développement, il conviendra de consacrer des efforts substantiels à la spécification de la distribution a priori. Un facilitateur expérimenté pourrait travailler avec un certain nombre d'experts pour obtenir les meilleurs jugements cliniques et scientifiques disponibles [O'Hagan et al., 2005].

3.4.2 « Puissance classique » : méthode hybride Bayésienne - fréquentiste supposant la normalité

Nous supposons que nous avons une distribution a priori à utiliser dans la conception de notre étude mais que les conclusions de l'étude seront entièrement classiques et ne feront pas usage de l'a priori, peut-être en raison d'une soumission à une autorité réglementaire. Supposons que nous ayons un a priori normal $\theta \sim N[\mu, \sigma^2/n_0]$ et que nos données futures Y_n aient une distribution $Y_n \sim N[\mu, \sigma^2/n]$ et que nous souhaitons calculer la probabilité prédictive d'obtenir un résultat classiquement « significatif » lors du test de l'hypothèse nulle $\theta < 0$. Dans une analyse classique, H_0 sera rejeté lorsque l'estimation du paramètre Y_n obéit

$$Y_n > \frac{-1}{\sqrt{n}} \sigma z_\epsilon \quad (3.3)$$

On note la propriété utile $\Phi(z) = 1 - \Phi(-z)$

$$P(Y \leq y) = P\left(\frac{Y - \theta}{\sigma} \leq \frac{y - \theta}{\sigma}\right) = P\left(Z \leq \frac{y - \theta}{\sigma}\right) = \Phi\left[\frac{y - \theta}{\sigma}\right]$$

où $z = (y - \theta)/\sigma$

Supposons que nous concevons un essai clinique avec la taille proposée n pour détecter une hypothèse alternative $H_1 : \theta > 0$, L'hypothèse nulle sera rejetée en faveur de H_1 à condition que (3.1), qui à partir des formules précédentes se produira avec probabilité

$$1 - \Phi\left[\frac{\frac{-\sigma z_\epsilon}{\sqrt{n}} - \theta}{\sigma\sqrt{n}}\right] = 1 - \Phi\left[-z_\epsilon - \frac{\theta\sqrt{n}}{\sigma}\right] = \Phi\left[\frac{\theta\sqrt{n}}{\sigma} + z_\epsilon\right]$$

cet événement, noté S_ϵ^c , se produira avec probabilité

$$P(S_\epsilon^c|\theta) = \Phi\left[\frac{\theta\sqrt{n}}{\sigma} + z_\epsilon\right] \quad (3.4)$$

La probabilité qu'un essai de n observations conduise à une conclusion statistiquement significative au niveau 2ϵ , étant donné que l'hypothèse alternative est vraie, est connue comme la puissance de l'étude, classiquement notée $1 - \beta$ et donc :

$$1 - \beta = \Phi\left[\frac{\theta_A\sqrt{n}}{\sigma} + z_\epsilon\right] \quad (3.5)$$

Nous pouvons tracer (3.4) superposé à l'a priori $p(\theta)$, ce qui peut révéler la plausibilité relative des hypothèses alternatives potentielles et suggérer si l'essai est basé sur des hypothèses trop optimistes (voir l'exemple 3.1). Si nous voulons calculer la probabilité inconditionnelle globale d'un résultat « significatif » S_ϵ^c on peut intégrer (3.4) par rapport à l'a priori. Cependant, il est analytiquement plus simple d'utiliser la distribution prédictive

$$Y_n \sim N \left[\mu, \sigma^2 \left(\frac{1}{n} + \frac{1}{n_0} \right) \right]$$

d'évaluer directement la probabilité que l'événement critique (3.3) se produise, qui peut être démontré

$$P(S_\epsilon^c) = \Phi \left[\sqrt{\frac{n_0}{(n_0 + n)}} \left(\frac{\mu\sqrt{n}}{\sigma} + z_\epsilon \right) \right] \quad (3.6)$$

La relation avec la courbe de puissance (3.4) est claire. Comme $n_0 \rightarrow \infty$, l'a priori tend vers une masse sur μ et $P(S_\epsilon^c)$ tend vers la puissance classique évaluée à la moyenne a priori. Cependant, n_0 fini signifiera que la puissance attendue est inférieure à la puissance classique évaluée à la moyenne a priori, à condition que la puissance classique soit supérieure à 50%. Il peut s'agir d'une évaluation plus réaliste des chances que l'essai aboutisse à une conclusion positive [Spiegelhalter et al., 2004]. Nous notons que le tableau 3.1 peut être étendu pour permettre des décisions « équivoques » et que les probabilités nécessaires peuvent être calculées en utilisant les aires de queue de la distribution normale bivariée [Spiegelhalter et Freedman, 1986].

Exemple 3.1. Puissance Bayésienne : Choix de la taille de l'échantillon pour un essai

Supposons que nous concevons un essai pour un nouveau traitement contre le cancer qui devrait augmenter la survie à 5 ans de 20% à 40%. D'après l'analyse de la probabilité normale pour les données de survie sous deux interventions différentes avec des taux de risque $h_1(t)$ et $h_2(t)$, p_1 et p_2 sont les chances de survivre jusqu'à un certain temps fixe sous les deux interventions comparées, alors sous l'hypothèse des risques proportionnels, $HR = \log(p_1)/\log(p_2)$ avec son logarithme naturel $\log(HR) = \theta = \log(\log(p_1)/\log(p_2))$, cela équivaut à un ratio d'hasard de $\log(0.40)/\log(0.20) = 0.57$ lorsqu'on suppose des risques proportionnels ou un log (rapport de risque) de $\theta_A = -0.56$.

Nous notons que la discussion ci-dessus sur la puissance a supposé une hypothèse alternative $\theta_A > 0$, alors que notre θ_A est négatif. Cepen-

dant, nous pouvons simplement inverser le rôle des hypothèses nulles et alternatives et prendre $\theta_A = 0.56$: Cela équivaut à redéfinir le rapport de risque comme risque de contrôle divisé par un nouveau risque d'intervention au lieu de son inverse. En prenant $\sigma = 2$, la puissance d'une étude dans laquelle n événements se produisent est donnée par (3.5) : En supposant que $\epsilon = 0.025$ génère la courbe de puissance illustrée à la figure 3.1, 80% de la puissance est atteinte à :

$$\begin{aligned} n &= (z_{1-\beta} - z_\epsilon)^2 \frac{\sigma^2}{\theta_A^2} & (3.7) \\ &= \frac{(7.85 \times 2)^2}{(0.56^2)} = 100 \end{aligned}$$

La puissance monte lentement au-dessus de cette taille d'essai. Dans l'hypothèse alternative, nous nous attendons à une mortalité globale de 30% sur 5 ans dans l'essai et donc pour observer 100 décès, nous pourrions recruter environ 330 patients, 165 dans chaque bras et les suivre pendant environ 5 ans.

Nous avons pris l'hypothèse alternative comme $\theta > 0$, conduisant à une courbe de puissance qui augmente pour des valeurs croissantes de θ . Nous interprétons cet exemple, dans lequel un essai pour un nouveau traitement contre le cancer est conçu pour avoir une puissance de 80% pour détecter un log (rapport de risque en anglais : « hazard ratio ») $\theta_A = 0.56$, nécessitant 100 événements en supposant un bilatéral de 0,05. Considérons un a priori enthousiaste archétypal centré sur l'hypothèse alternative et avec 5% de probabilité a priori que $\theta < 0$ [Spiegelhalter et al., 2004]. D'où $Y_n \sim N[\mu, \sigma^2/n_0]$, $\mu = 0.56$, $\sigma = 2$ et $\mu - 1.645\sigma/n_0 = 0$, de sorte que $n_0 = 1.645^2\sigma^2/\mu^2 = 34.5$.

La courbe de puissance classique et l'a priori sont montrés sur la figure 3.2 : La puissance à la moyenne a priori est de 80% comme prévu, la puissance attendue (3.6) en moyenne sur toute la distribution a priori est de 0,66, montrant la baisse par rapport à la valeur conditionnelle de 0,80. Si nous avons adopté l'approche recommandée par Brown et al.

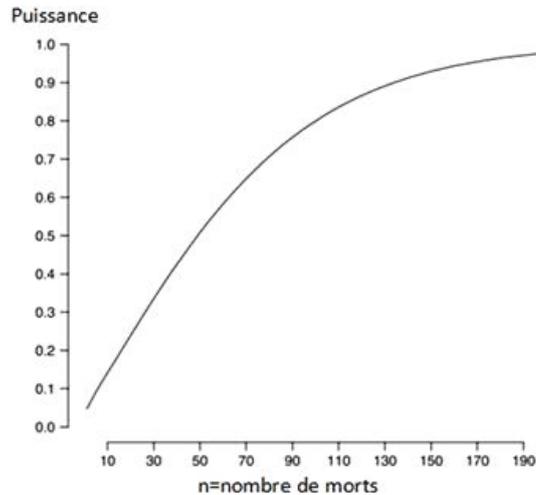


FIGURE 3.1 – Puissance d’un essai clinique dans lequel n événements doivent être observés et l’hypothèse alternative est une augmentation de 20% de survie à 40% de survie, équivalant à un rapport de risque (contrôle/nouveau) de $1/0,57$ ($\log(\text{ratio de hasard}) = \theta_A = 0,56$) : Puissance = $\Phi \left[\frac{\theta\sqrt{n}}{\sigma} + z_\epsilon \right]$. 80% de puissance est atteinte à $n = 100$.

(1987), nous ferions la moyenne de la courbe de puissance par rapport à l’a priori conditionnel $p(\theta|H_1) = p(\theta|\theta > 0)$; ce n’est pas simple à calculer et il est peut-être plus facile à l’évaluer en utilisant des méthodes de Monte Carlo, à partir desquelles nous trouvons, en utilisant la notation du tableau 3.1, que $p(H_1|D_1) = 0.70$. Une telle valeur aurait pu être prédite, car nous savons que $p(H_1) = 0.95$, $p(D_1) = 0.66$ et à partir de (3.1) que $p(D_1) \approx p(D_1|H_1)p(H_1)$

3.4.3 Puissance Bayésienne

Supposons que nous ayons les mêmes a priori et vraisemblances normaux qu’à la section 3.4.2 mais souhaitons maintenant effectuer une analyse entièrement Bayésienne dans laquelle l’a priori sera incorporé.

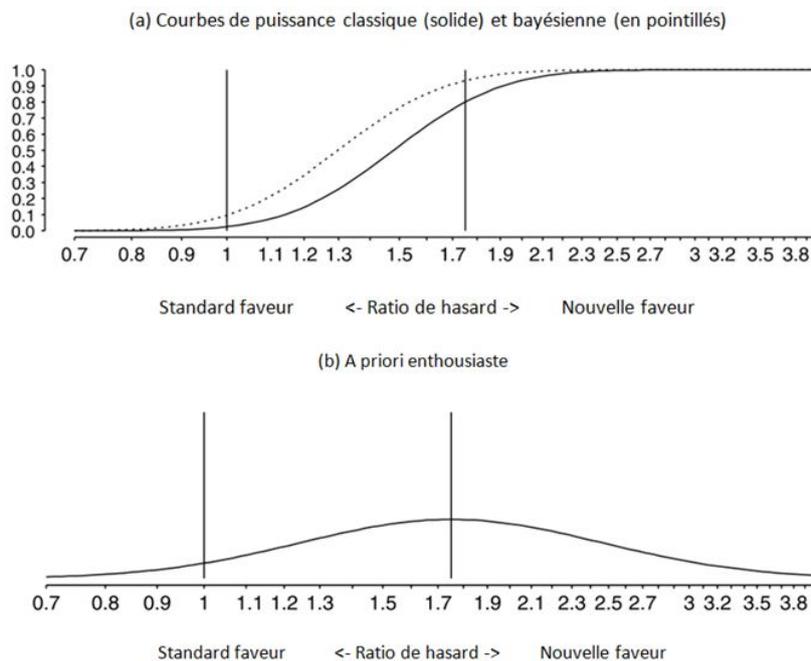


FIGURE 3.2 – Courbes de puissance (a) pour tester $H_1 : \theta > 0$, conçues pour avoir une puissance classique de 80% à $\theta_A = 0.56$ ($HR = 1.75$). La courbe de puissance Bayésienne en (a) suppose que l'a priori enthousiaste montré en (b) doit être inclus dans l'analyse.

Nous souhaitons calculer la probabilité prédictive d'obtenir un résultat Bayésien « significatif » lors du test de l'hypothèse nulle $\theta < 0$ par rapport à une alternative $\theta > 0$ et nous désignerons une telle « signification Bayésienne » comme $S_\epsilon^B = P(\theta < 0 | data) < \epsilon$.

En supposant une future estimation des paramètres Y_n , nous obtiendrons la distribution a posteriori

$$\theta | Y_n \sim N \left[\frac{n_0 \mu + n Y_n}{n_0 + n}, \frac{\sigma^2}{n_0 + n} \right],$$

et donc S_ϵ^B se produira lorsque l'estimation du paramètre Y_n obéit à

$$Y_n > \frac{-\sqrt{n_0 + n} z_\epsilon \sigma - n_0 \mu}{n} \quad (3.8)$$

Pour une vraie valeur particulière de θ , $Y_n \sim N[\mu, \sigma^2/n]$, il est donc

facile de montrer que cet événement se produira avec probabilité

$$P(S_\epsilon^B|\theta) = \Phi \left[\frac{\theta\sqrt{n}}{\sigma} + \frac{n_0\mu}{\sigma\sqrt{n}} + \frac{n_0+n}{n}z_\epsilon \right] \quad (3.9)$$

Avec un vague avis a priori, $n_0 \rightarrow 0$ et nous nous retrouvons avec la courbe de puissance classique standard donnée en (3.5). Tout comme dans la section 3.4.2, nous pouvons tracer (3.9) superposée a l'a priori $p(\theta)$: Pour calculer la probabilité inconditionnelle globale d'un résultat S_ϵ^B « significatif », il est à nouveau analytiquement plus simple d'utiliser la distribution prédictive de Y_n pour évaluer la probabilité que l'événement critique (3.8) se produise :

$$\begin{aligned} P(S_\epsilon^B) &= P\left(Y_n > \frac{-\sqrt{n_0+n} z_\epsilon \sigma - n_0\mu}{n}\right) \\ &= \Phi \left[\frac{\mu\sqrt{n_0+n}\sqrt{n_0}}{\sigma\sqrt{n}} + \frac{\sqrt{n_0}}{\sqrt{n}}z_\epsilon \right] \end{aligned} \quad (3.10)$$

Exemple 3.2. Puissance Bayésienne (suite) : Choix de la taille de l'échantillon pour un essai

Si nous sommes disposés à inclure la distribution a priori dans l'analyse, nous obtenons alors la courbe de puissance Bayésienne (3.9) représentée par une ligne en pointillés sur la figure 3.2 (a), qui est sensiblement plus élevée que la courbe de puissance classique en raison de la donnée a priori donnant un « bon départ ». La puissance de l'hypothèse alternative $\theta_A = 0.56$ est de 0,93, tandis que la chance d'un faux rejet de $\theta = 0$ est passé de 0,025 à 0,10, cette chance gonflée d'une erreur de type I illustre le risque d'obtenir le « mauvais » a priori. La puissance Bayésienne attendue (3.10), moyennée par rapport à la distribution a priori de la figure 3.2 (b), est de 0,78.

3.4.4 Formules d'ajustement pour différentes hypothèses

Toutes les formules fournies jusqu'à présent ont supposé que $\theta > 0$ indique une performance supérieure du traitement innovant et donc est l'hypothèse alternative d'intérêt, cela a simplifié l'exposition mais ne tiendra clairement pas dans toutes les situations. Une option consiste à redéfinir les mesures et les paramètres des résultats afin qu'ils aient les propriétés requises. Alternativement, on peut transformer les formules fournies et nous considérons maintenant les transformations nécessaires lorsque différentes hypothèses sont envisagées :

Seuil non nul. Supposons que l'hypothèse nulle soit $H_0 : \theta < 0$ et l'alternative $H_1 : \theta > 0$. Chacune des formules précédentes peut être transformées en remplaçant θ_0 de la moyenne a priori, de la statistique observée y_m et, dans les calculs de puissance conditionnelle, le paramètre θ . Par exemple, supposons que dans l'exemple 3.1, le seuil d'intérêt a été changé à $\theta = 0.2$, c'est-à-dire que l'intervalle a posteriori devrait se situer entièrement au-dessus d'un log (ratio de hasard) de 0.2 ($HR = 1.22$) avant que H_0 ne soit rejeté. La puissance conditionnelle à l'hypothèse alternative $\theta_A = 0.56$ n'est maintenant seulement 0.56, obtenue à partir de la transformation (3.9), tandis que la puissance attendue est trouvée à partir de (3.10) devient 0,53.

Renversement d'hypothèses. Comme nous l'avons vu dans les exemples précédents, il est courant d'exprimer le bénéfice de la nouvelle intervention comme une réduction du risque, et donc sur une échelle logarithmique pour fixer $H_1 : \theta < 0$. Ainsi, un résultat « significatif » sera obtenu si un intervalle final est entièrement inférieur à 0. Si, par exemple, nous adoptons une approche entièrement Bayésienne, cela doit être équivalent à l'événement $P(\theta > 0|data) < \epsilon$, que nous désignerons $S_\epsilon^{B^-}$ comme suit :

$$S_\epsilon^{B^-} \equiv [P(\theta > 0|data) < \epsilon] \equiv [P(\theta < 0|data) < 1 - \epsilon]$$

et donc, par exemple :

$$P(S_\epsilon^{B^-}) = 1 - P(S_{1-\epsilon}^B)$$

Par conséquent, les formules fournies peuvent être transformées en remplaçant $1 - \epsilon$ à ϵ et en remplaçant le résultat à partir de 1. Par exemple, supposons dans l'exemple 3.1 que le seuil d'intérêt a été changé en $\theta = 0.69$, $HR = 2$, et en outre, nous nous sommes intéressés à la puissance attendue pour rejeter l'hypothèse nulle $H_0 : \theta > 0$, c'est-à-dire que nous sommes intéressé par des valeurs avec un ratio d'hasard inférieur à 2. L'utilisation des deux transformations sur (3.9) conduit à :

$$P(S_\epsilon^{B^-}) = 1 - \Phi \left[\frac{(\mu - \theta_0)\sqrt{n_0 + n}\sqrt{n_0}}{\sigma\sqrt{n}} + \frac{\sqrt{n_0}}{\sqrt{n}} z_{1-\epsilon} \right] \quad (3.11)$$

$$\text{où } P(S_{1-\epsilon}^B) = \Phi \left[\frac{(\mu - \theta_0)\sqrt{n_0 + n}\sqrt{n_0}}{\sigma\sqrt{n}} + \frac{\sqrt{n_0}}{\sqrt{n}} z_{1-\epsilon} \right]$$

Ensuite, à partir de (3.11), nous trouvons que la puissance attendue est de 0,24 : Une valeur aussi faible pourrait être anticipée à partir du support a priori substantiel pour H_0 .

Exemple 3.3. Taille de l'échantillon pour un essai de chirurgie du cancer gastrique

Référence : Fayers et al. (2000), Spiegelhalter et al. (2004).

Intervention : Radicale (D2) par rapport à la chirurgie conventionnelle (D1) pour le cancer gastrique.

Objectif de l'étude : Des données provenant du Japon suggèrent qu'une chirurgie plus radicale était une explication possible des meilleurs taux de survie des patients atteints d'un cancer gastrique et le UK Medical Research Council a lancé un essai randomisé pour comparer la survie après une chirurgie radicale et conventionnelle.

Plan d'étude : RCT (radical cancer treatment) parallèle à deux groupes.

Mesure des résultats : Ratio d'hasard de décès ($HR > 1$ favorise un traitement radical).

Taille d'échantillon prévue : L'essai a été conçu en supposant que la différence minimale cliniquement significative était une amélioration de 13,5% de la survie à 5 ans de 20% à 33,5% chez les patients subissant une chirurgie conventionnelle; cette valeur pour l'hypothèse alternative était basée sur l'opinion de l'équipe d'essai. Cela équivaut à un rapport de risque de $\log(0.20) = \log(0.335) = 1.47$, ou $\log(HR) = 0.39$. Pour que l'essai puisse détecter une différence de 13,5% au niveau de signification de 5% avec une puissance de 90%, le nombre nécessaire d'événements (c-à-d décès) est $n = \sigma^2(1.96 + 1.28)^2/0.39^2 = 276$, en prenant $\sigma = 2$ qui est également un choix approprié dans l'analyse de survie). L'essai a été conçu pour avoir 200 patients par bras, ce qui devrait produire ce nombre d'événements.

Modèle statistique : Pour le but de la planification, l'approximation normale a été adoptée, tandis que pour l'analyse, une régression de Cox complète a été utilisée pour obtenir une probabilité de $\log(HR)$.

Distribution a priori : En complément aux trois membres chirurgicaux du comité de pilotage de l'essai, 23 chirurgiens supplémentaires avaient leurs convictions concernant les avantages / inconvénients probables de la chirurgie radicale par rapport à la chirurgie conventionnelle, tous les deux au début de l'essai et plus tard lorsque l'essai s'est arrêté mais n'a pas encore été publié. La distribution moyenne avait une moyenne a priori de 9,4% d'amélioration par rapport à leur contrôle moyen estimé à 5 ans de 21%, l'asymétrie dans les distributions donne lieu à une médiane d'environ 4%. En supposant une survie de base de 21%, la distribution d'une amélioration p peut être transformée en une échelle $\log(HR)$ par $\log(HR)/\log(\log(0.21)/\log(0.21 + p))$, l'ajustement d'une distribution normale à l'histogramme transformé donne un a priori avec une moyenne $\mu = 0.12$ et un écart type $\sigma/\sqrt{n_0} = 0.19$ et donc $n_0 = 4/0.19^2 = 111$. Cela correspond à un rapport de risque de 1,13 (95% d'intervalle de 0,78 à 1,64). Cette répartition est illustrée à la figure 2.3(a), révélant que la probabilité de dépasser l'hypothèse alternative de

$HR = 1.47$ est de 8%. Par conséquent, les croyances générales a priori des chirurgiens révèlent que l'essai a été conçu autour d'un objectif plutôt optimiste.

La figure 3.2 (b) montre la courbe de puissance (3.4) pour l'essai sur la base d'un $n=27$ événements attendus, avec une puissance de 90% à l'hypothèse alternative de 1,47. La figure 3.2 (a) montre que la croyance des chirurgiens est concentrée dans une zone de puissance assez faible. En effet, (3.6) montre que la puissance attendue n'est que de 30%, ce qui monte légèrement à 31% si une analyse Bayésienne finale est effectuée (3.10). Même si les chirurgiens étaient considérablement plus optimistes et que leur moyenne a priori était fixée à l'hypothèse alternative de $HR = 1.47$, la puissance attendue n'augmenterait que de 45%.

Fonction de perte ou demandes : Non, mais aussi bien que obtenir les croyances des chirurgiens, les auteurs ont obtenu leurs demandes de chirurgie radicale : Une amélioration d'environ 10% a été jugée nécessaire avant de vouloir mettre en œuvre systématiquement la chirurgie la plus radicale, qui est plus vaste et présente un risque supplémentaire des complications et d'utilisation des ressources.

Calcul / logiciel : Conjuguer le modèle normal.

Résultat de l'étude : L'essai a recruté les 200 patients sur chaque bras et finalement 281 événements ont été observés (137 sous $D1$, 144 sous $D2$), avec un résultat légèrement en faveur de la chirurgie conventionnelle. Le rapport de risque observé, basé sur une régression de Cox, était de 0,91 (IC à 95% de 0,72 à 1,15), équivalent à un $\log(HR)$ de -0.09 (erreur standard 0.11, équivalent à un nombre effectif d'événements de $m = \sigma^2/0.11^2 = 278$, presque exactement le même que le nombre réel d'événements observés). Le taux de survie à 5 ans chez les patients subissant une chirurgie conventionnelle était de 30%, considérablement plus élevé que les 20% attendus avant le début de l'essai. Cette probabilité est illustrée à la figure 3.3 (c).

Interprétation Bayésienne : La figure 2.3 (d) montre la distribution prédictive du rapport de risque observé, dérivée à l'aide des méthodes décrites à la section 2.4. La probabilité d'observer un résultat aussi ex-

trême que celui observé est de 0,32, deux fois la zone ombrée illustrée à la figure 2.3 (d). Il s'agit de la mesure de Box du conflit [Spiegelhalter et al., 2004, section 5.8] entre l'a priori et la vraisemblance et n'est pas particulièrement extrême même si l'attente a priori d'un bénéfice de D2 était en conflit avec le risque d'hasard observé.

Commentaires : Fayers et al. (2000) ont effectué un deuxième exercice de sollicitation une fois l'essai terminé avant l'annonce des résultats et ont constaté qu'il y avait encore un optimisme considérable parmi les collaborateurs cliniques. Ils concluent que bien que les opinions changent au fil du temps, les personnes impliquées dans un essai clinique ont tendance à être optimistes et si leurs attentes a priori sont utilisées comme base naïve pour les calculs de taille d'échantillon, l'essai pourrait entraîner une taille d'échantillon trop petite. Néanmoins, dans cet exemple, l'hypothèse alternative a été jugée optimiste même par les participants. Une évaluation plus réaliste des chances de succès de l'essai pourrait être réalisée en tenant compte de leur entière incertitude. Il est également important de surveiller un tel procès afin qu'il ne se poursuive pas inutilement. Dans cet exemple, l'essai aurait pu être arrêté et rejeté une « différence importante » quelque temps avant la conclusion finale. Cependant, il peut être plus approprié de surveiller l'utilisation de l'a priori clinique afin de s'assurer que le résultat négatif est convaincant même pour les enthousiastes [Spiegelhalter et al., 2004].

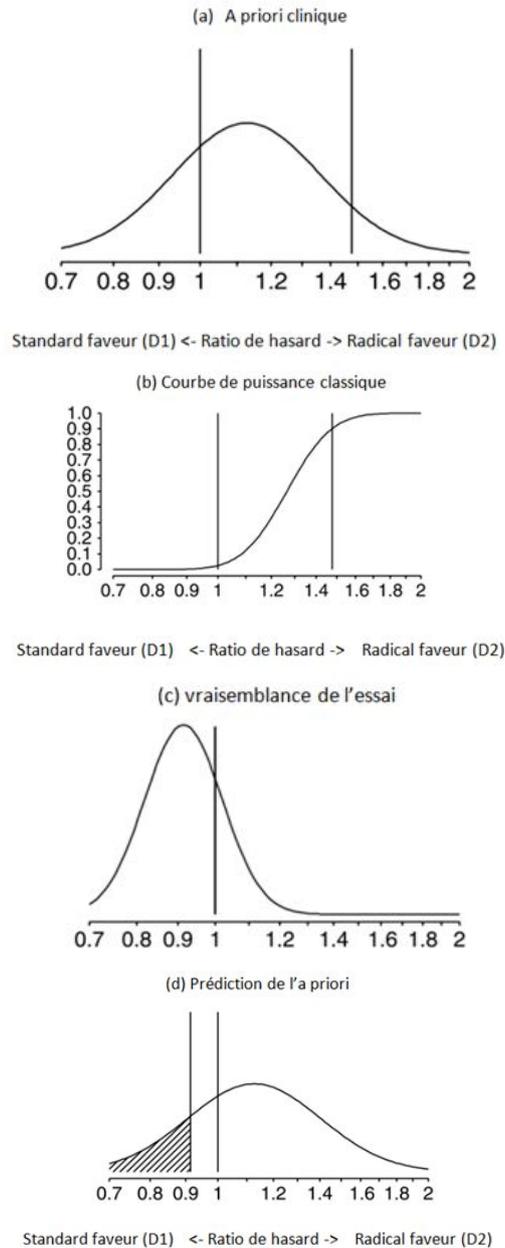


FIGURE 3.3 – L'évaluation a priori (a) de l'essai D2 en chirurgie du cancer gastrique montre une certaine attente de bénéfique mais l'hypothèse alternative de 1,47 autour de laquelle l'essai a été conçu est clairement très optimiste (b). Le résultat final de l'essai (c) n'a montré aucune preuve claire de l'avantage. La distribution prédictive dérivée de l'a priori (d) montre que le résultat observé (HR=0.91) n'était pas particulièrement surprenant, étant donné l'avis a priori exprimé par (a).

3.4.5 Distribution prédictive de la puissance et taille d'échantillon nécessaire

Considérons la formule de puissance classique donnée en (3.5). Si nous exprimons l'incertitude sur les paramètres comme une distribution a priori, alors la puissance peut être considérée comme une quantité inconnue avec une distribution induite par cet a priori. Cette distribution prédictive sur la puissance peut être obtenue au mieux par une méthodologie de simulation : Essentiellement les paramètres inconnus sont simulés à partir de leur distribution a priori, branchés dans la formule de la puissance et le résultat enregistré. Après de nombreuses itérations de cette procédure, une répartition sur les puissances possibles est obtenue. Il s'agit essentiellement d'une procédure de Monte Carlo qui est illustrée dans l'exemple 3.4.

Méthodes de Monte Carlo :

Les méthodes de Monte Carlo sont une boîte à outils de techniques qui ont toutes pour but d'évaluer des intégrales ou des sommes par simulation plutôt que par une analyse algébrique exacte ou approximative. Les méthodes de Monte Carlo seront utiles pour l'analyse Bayésienne à condition que la distribution préoccupante appartienne à une famille connue, cette distribution peut être a priori (si aucune donnée n'est disponible) ou a posteriori actuelle. Dans l'analyse Bayésienne conjuguée, il sera possible de dériver algébriquement une telle distribution a posteriori et donc d'utiliser les méthodes de Monte Carlo pour trouver les zones de queue (bien que ces zones de queue puissent également être directement obtenues dans le logiciel) ou plus utilement de trouver la distribution de fonctions complexes d'une ou plusieurs quantités inconnues comme dans l'analyse de sensibilité probabiliste. L'idée de base est de remplacer l'algèbre par la simulation, une application de ces idées dans les calculs de puissance est donnée dans l'exemple simple suivant :

Exemple 3.4. Incertitude : Distribution prédictive de la puissance

Supposons qu'un essai randomisé est prévu avec n patients dans chacun des deux bras, en utilisant une réponse avec un écart type $\sigma = 1$; par conséquent, la variance d'un contraste entre deux patients est $2\sigma^2$. L'essai vise à avoir une erreur de type I (bilatéral α) de 5% et une puissance de 80% pour détecter une vraie différence de $\theta = 0.5$ dans la réponse moyenne entre les groupes. À partir de (3.7), la taille d'échantillon nécessaire par groupe est :

$$n = (z_{0.8} - z_{0.025})^2 \frac{2\sigma^2}{\theta^2}$$

où $z_{0.8} = 0.842$, $z_{0.025} = -1.96$; notez que cela diffère légèrement de (3.7) car ici σ est l'écart type d'une réponse unique.

La taille d'échantillon nécessaire est $n = 63$. Supposons cependant que nous souhaitions exprimer une incertitude concernant à la fois θ et σ . Pour θ , nous évaluons une moyenne a priori de 0,5 et un écart-type a priori de 0,1, tandis que pour σ , nous supposons une moyenne a priori de 1 et un écart-type de 0,3. θ et σ sont supposés être indépendants et normalement distribués (sous réserve que la contrainte de σ soit positive).

En utilisant les méthodes de Monte Carlo, nous simulons les valeurs de θ et σ à partir de leurs distributions a priori, les remplacer dans la formule de taille d'échantillon ci-dessus et obtenons ainsi une distribution prédictive sur n . Cette distribution a les propriétés indiquées dans le tableau 3.2 et est représentée sur la figure 3.4, il est clair qu'il existe une énorme incertitude quant à la taille appropriée de l'échantillon.

TABLE 3.2 – Propriétés des distributions prédictives de la taille d'échantillon n nécessaire pour une puissance fixe de 80% et de la puissance pour une taille d'échantillon fixe $n=63$.

	Médiane	Intervalle 95%
n	62.5	9.3 à 247.2
Puissance %	80	29 à 100

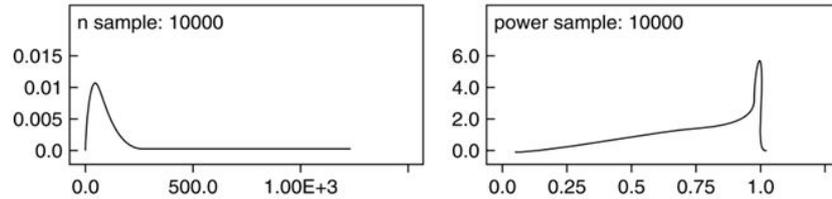


FIGURE 3.4 – Distributions prédictives de WinBUGS pour la taille d'échantillon n nécessaire pour atteindre une puissance de 80% et une puissance pour $n=63$ patients par groupe.

Pour n fixe, la puissance est :

$$Puissance = \Phi \left[\sqrt{\frac{n\theta^2}{2\sigma^2}} + z_{0.025} \right]$$

Si nous décidons d'utiliser 63 patients par groupe, nous pouvons simuler des valeurs potentielles pour la puissance en utilisant la même méthodologie. Les résultats sont à nouveau présentés dans le tableau 3.2 et représentés sur la figure 3.4 et montrent que bien que la puissance médiane soit de 80%, un essai de 63 patients par groupe pourrait être sérieusement sous-alimenté. Nous pouvons calculer d'autres quantités qui pourraient donner un aperçu de la taille d'échantillon prévue : Par exemple, il y a 37% de chances que la puissance soit inférieure à 70%.

3.5 Calculs d'assurance pour les modèles paramétriques de survie

Nous supposons maintenant que, dans chacun des deux groupes de traitement, la variable du résultat pour chaque patient est le temps de survie à un événement et considérons le modèle exponentiel pour les temps de survie. Pour le modèle, il faut d'abord choisir la méthode d'analyse et donc les critères de réussite de l'essai. Nous considérons ensuite les calculs d'assurance et les méthodes d'élicitation pour la distribution a priori requise de [Rena and Oakley, 2013] .

3.5.1 Schémas de Randomisation

La randomisation joue un rôle important dans la conduite des essais cliniques. La randomisation génère non seulement des groupes comparables de patients qui constituent des échantillons représentatifs de la population de patients visée mais permet également des tests statistiques valides pour l'évaluation clinique du médicament à l'étude. La randomisation dans les essais cliniques implique le recrutement aléatoire des patients dans la population cible de patients et l'affectation aléatoire des patients aux traitements.

Dans le cadre de la randomisation, une inférence statistique peut être établie sous une certaine hypothèse de distribution de probabilité de la population visée des patients. L'hypothèse de distribution de probabilité dépend de la méthode de randomisation sous un modèle de randomisation. Une étude sans randomisation entraîne la violation de l'hypothèse de distribution de probabilité et, par conséquent, aucune inférence statistique précise et fiable sur l'évaluation de l'innocuité et de l'efficacité du médicament à l'étude ne peut être tirée [Chow et al., 2003].

Considérons une conception en deux groupes parallèles pour comparer un médicament test et un témoin (placebo) où n_1 patients sont assignés au hasard au groupe du traitement et n_2 patients sont assignés au hasard au groupe témoin. Lorsque la randomisation est correctement appliquée, le modèle de population tient et les réponses des patients sont normalement distribuées. Considérons d'abord le cas le plus simple où deux populations de patients (traitement et contrôle) ont la même variance σ_1 et σ_2 est connu. Soit μ_1 et μ_2 les moyennes de population pour le traitement et le contrôle, respectivement. L'hypothèse nulle selon laquelle $\mu_1 = \mu_2$ (i-e., Il n'y a pas d'effet du traitement) est rejetée au niveau de signification α si :

$$T = \frac{|\bar{x}_1 - \bar{x}_2|}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > z_{\alpha/2}$$

où \bar{x}_1 est la moyenne de l'échantillon des réponses des patients du groupe de traitement, \bar{x}_2 est la moyenne de l'échantillon des réponses

des patients du groupe de contrôle et $z_{\alpha/2}$ est le centile supérieur $\alpha/2$ de la distribution normale standard. Intuitivement, mélanger les codes du traitement n'affecte pas le niveau de signification du test. La puissance du test, c'est-à-dire la probabilité de détecter correctement une différence de traitement lorsque $\mu_1 \neq \mu_2$, est :

$$P(\theta) = P\left(\frac{|\bar{x}_1 - \bar{x}_2|}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > z_{\alpha/2}\right) = \Phi[\theta - z_{\alpha/2}] + \Phi[-\theta - z_{\alpha/2}] \quad (3.12)$$

où Φ est la fonction de distribution normale standard et

$$\theta = \frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Cela découle du fait que sous le modèle de randomisation, $\bar{x}_1 - \bar{x}_2$ a la distribution normale avec une moyenne $\mu_1 - \mu_2$ et une variance $\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$.

Considérons la conception à deux groupes parallèles avec un inconnu σ^2 . La statistique de test est nécessairement modifiée en remplaçant $z_{\alpha/2}$ et σ^2 par $t_{(\alpha/2; n_1+n_2-2)}$ et

$$\hat{\sigma}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

où s_1^2 est la variance d'échantillon basée sur les réponses des patients du groupe de traitement, s_2^2 est la variance d'échantillon basée sur les réponses des patients du groupe témoin et $t_{(\alpha/2; n_1+n_2-2)}$ est la partie supérieure ($\alpha/2$) percentile de la t -distribution avec $n_1 + n_2 - 2$ degrés de liberté. Ensuite, le test t à deux échantillons de α niveau de signification et la puissance est donnée par :

$$1 - \Gamma_{(n_1+n_2-2)}(t_{(\alpha/2; n_1+n_2-2)}|\theta) + \Gamma_{(n_1+n_2-2)}(-t_{(\alpha/2; n_1+n_2-2)}|\theta)$$

où θ est défini par :

$$\theta = \frac{\mu_1 - \mu_2}{\hat{\sigma} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

et $\Gamma_{(n_1+n_2-2)}(\cdot|\theta)$ est la fonction de t -distribution non centrale avec $n_1 + n_2 - 2$ degrés de liberté et le paramètre de non-centralité θ .

3.5.2 Distribution exponentielle

Rena and Oakley, (2013) supposent d'abord que les temps de survie dans chaque groupe de traitement suivent une distribution exponentielle, avec des taux de risque λ_1 et λ_2 ($i = 1$ pour le groupe témoin et $i = 2$ pour le groupe expérimental) et permettent une période de recrutement limitée dans le temps 0 à R avec entrée uniforme du patient et T comme durée totale de l'essai. L'origine temporelle du temps de survie est le moment où un patient entre dans l'essai et non le début de l'essai, ici le travail est basé sur la considération de la méthode d'analyse basée sur [Schoenfeld et Richter, 1982].

L'hypothèse nulle est $m_1/m_2 = 1$, où m_i est le temps de survie médian dans le groupe i , contre l'alternative $m_1/m_2 = \phi$, où ϕ est la différence minimale cliniquement importante. À noter qu'en supposant un modèle exponentiel, les hypothèses énoncées précédemment sont équivalentes à $H_0 : \theta = 0$ contre $H_1 : \theta \neq 0$, où :

$$\theta = \log(\lambda_1/\lambda_2) \tag{3.13}$$

La statistique de test est :

$$T = \frac{\hat{\theta} - 0}{\sqrt{\frac{1}{d_1} + \frac{1}{d_2}}} \sim N(0, 1) \text{ sous } H_0 : \theta = 0$$

où d_i est le nombre d'événements dans le groupe i et $\hat{\theta}$ est l'estimation du maximum de vraisemblance de θ . Sur la base des propriétés asymptotiques de la statistique d'essai, la puissance π^E d'un essai bilatéral à

100 α % est

$$\begin{aligned} \pi^E &= P(\text{rejeter } H_0 | \lambda_1, \lambda_2) \\ &= \Phi \left[-\frac{\theta}{\sqrt{\frac{1}{N_1 P_{1e}} + \frac{1}{N_2 P_{2e}}}} - z_{\alpha/2} \right] + \Phi \left[\frac{\theta}{\sqrt{\frac{1}{N_1 P_{1e}} + \frac{1}{N_2 P_{2e}}}} - z_{\alpha/2} \right] \end{aligned} \quad (3.14)$$

où N_i est le nombre de patients du groupe i et P_{ie} est la probabilité qu'un patient individuel du groupe i subisse l'événement de résultat au cours de l'essai. Schoenfeld et Richter (1982) ont dérivé une formule exacte pour P_{ie} :

$$P_{ie} = 1 - \frac{\exp(-\lambda_i(T - R))}{(\lambda_i R)} \quad \text{pour } i = 1, 2 \quad (3.15)$$

L'assurance de rejeter l'hypothèse nulle avec des données favorisant le traitement expérimental est

$$\begin{aligned} \gamma^E &= P(\text{rejeter } H_0, \hat{\theta} < 0) \\ &= \int \Phi \left[-\frac{\theta}{\sqrt{\frac{1}{N_1 P_{1e}} + \frac{1}{N_2 P_{2e}}}} - z_{\alpha/2} \right] f(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2 \end{aligned} \quad (3.16)$$

où θ et P_{ie} sont des fonctions de λ_1 et λ_2 .

3.5.3 Construire les a priori

D'après (3.16), un a priori conjoint est requis pour λ_1 et λ_2 . [Kadane et Wolfson, 1998] ont fait valoir qu'il vaut mieux demander un avis sur les quantités observables plutôt que sur les paramètres dans les modèles statistiques et nous suivons leurs conseils ici. Nous obtenons $f(\lambda_1, \lambda_2)$ via des jugements sur les taux de survie à un moment précis. Pour construire un a priori pour chaque paramètre, notez d'abord que

$$\lambda_i = \frac{-\log(S_i(t_0))}{t_0} \quad \text{pour } i = 1, 2 \quad (3.17)$$

avec $S_i(t_0)$ le taux de survie du groupe i au temps t_0 . Par conséquent,

pour obtenir la distribution a priori conjointe, nous obtenons d’abord des jugements sur $S_1(t_0)$ au lieu de susciter directement des croyances sur λ_1 . Un expert peut juger $S_1(t_0)$ comme informatif pour $S_2(t_0)$, de sorte que λ_1 et λ_2 ne sont pas indépendants. Pour susciter cette dépendance, des jugements ont été proposés par [Rena and Oakley, 2013] sur la différence, $\rho = S_2(t_0) - S_1(t_0)$, et que ρ est supposé indépendant de $S_1(t_0)$.

Les méthodes pour obtenir des distributions univariées sont données dans [O’Hagan et al., 2006 les sections 6.3 et 6.4]. Voir également [Johnson et al., 2010] pour une revue systématique des méthodes d’éllicitation et aussi [Oakley et O’Hagan, 2010].

Une option consiste à obtenir une distribution bêta pour $S_1(t_0)$ et une distribution normale pour ρ , tronquant la priorité normale si nécessaire pour garantir $S_2(t_0) \in (0, 1)$. (Une alternative serait d’utiliser une distribution bêta décalée et mise à l’échelle).

À titre d’illustration, une méthode de « roulette d’essai » proposée par Gore (1987) est décrite pour obtenir un a priori normal pour ρ . La méthode est basée sur l’approche à intervalle fixe, dans laquelle l’expert est invité à fournir une probabilité que la quantité d’intérêt inconnue tombe dans un intervalle préfixe. En utilisant le package SHELF, le facilitateur, qui conduit l’éllicitation, obtient d’abord de l’expert les limites inférieures et supérieures de la gamme des valeurs plausibles pour ρ . Ensuite, l’animateur divise la gamme de la limite inférieure à la limite supérieure en 10 « bacs » de largeur égale. L’expert est invité à préciser sa probabilité de ρ mentir dans un bac particulier en plaçant des « jetons » dans ce bac, la proportion des jetons alloués représente la probabilité. Le nombre de jetons remis à l’expert est précisé par le facilitateur. Par exemple, si au total 20 jetons sont utilisés, chaque jeton représente une probabilité de 0,05. La méthode de la roulette d’essai est simple à utiliser et fournit à l’expert un affichage immédiat de ses jugements obtenus. Une distribution paramétrique peut être ajustée aux jugements obtenus en utilisant une procédure des moindres carrés : Les paramètres sont choisis pour rendre les probabilités ajustées aussi proches que possible des probabilités obtenues. Une rétroaction doit être fournie à l’expert pour vérifier

l'adéquation de la distribution obtenue [Rena et Oakley, 2013].

Une illustration de l'utilisation de l'élicitation dans les essais cliniques est donnée par Parmar et al. (1994), qui ont utilisé un questionnaire pour obtenir le log ratio d'hasard donné en (3.13) en obtenant une estimation ponctuelle pour $S_1(t_0)$ et une distribution a priori pour l'utilisation de la méthode de la roulette. Tan et al. (2003) et Hiance et al.(2009) ont adapté ce questionnaire et l'ont utilisé pour un essai de phase III. Le processus d'élicitation de [Rena et Oakley, 2013] est plus compliqué car ils considèrent l'incertitude dans $S_1(t_0)$ et ρ .

Étant donné les distributions obtenues $f(S_1(t_0))$ et $f(\rho)$. γ^E est estimée en utilisant la simulation de Monte Carlo :

$$= \frac{1}{M} \sum_{i=1}^M \Phi \left[-\frac{\theta^{(j)}}{\sqrt{\frac{1}{N_1 P_{1e}^{(j)}} + \frac{1}{N_2 P_{2e}^{(j)}}}} - z_{\alpha/2} \right] \quad (3.18)$$

où $\theta^{(j)}$ et $P_{1e}^{(j)}$ pour $i = 1, 2$ sont obtenus par les étapes suivantes :

1. Simulez $S_1^{(j)}(t_0)$ à partir de la distribution a priori obtenue $f(S_1(t_0))$.
2. Simulez $\rho^{(j)}$ à partir de la distribution a priori obtenue $f(\rho)$ et calculez $S_2^{(j)}(t_0) = S_1^{(j)}(t_0) + \rho^{(j)}$.
3. Calculez $\lambda_1^{(j)}$ et $\lambda_2^{(j)}$ en utilisant (3.17).
4. Calculez $\theta^{(j)}$ en utilisant (3.13).
5. Calculez $P_{1e}^{(j)}$ et $P_{2e}^{(j)}$ en utilisant (3.15).

Le processus de [Rena et Oakley, 2013] est rapide sur le plan des calculs, nous pouvons donc faire M très grande pour assurer la convergence.

3.5.4 Logiciel supporté

Shijie Rena et Jeremy E. Oakley (2013) ont mis à disposition un logiciel pour implémenter cette méthode. Le logiciel peut être téléchargé depuis www.jeremy-oakley.staff.shef.ac.uk/assurance.zip. Pour le cas exponentiel, Rena et Oakley (2013) ont écrit un outil d'élicitation interactif pour l'assurance informatique. Le code est écrit en R Development Core

Team (2011) et utilise le `rpanel` package de Bowman et Crawford (2008) et le `tkrplot` package de Tierney (2011) pour fournir des graphiques interactifs. L'outil aide à obtenir les distributions a priori du taux de survie de base $S_1(t_0)$ à un temps spécifié t_0 et la différence de survie entre le groupe expérimental et le groupe témoin au temps t_0 en utilisant la méthode de roulette d'essai. Une fois l'a priori spécifié, l'outil trace les courbes de puissance et d'assurance pour les distributions obtenues correspondantes.

3.5.5 Exemple numérique

Dans cette section, un exemple pour illustrer comment l'assurance est calculée est fourni pour éclairer le choix de la taille de l'échantillon sous différentes hypothèses du modèle. Pour présenter également comment l'a priori obtenu a un effet sur l'assurance. Dans cet exemple, l'essai contrôlé randomisé est supposé qu'il va être mené pour comparer deux effets du traitement avec un nombre égal de patients répartis dans chaque groupe du traitement. Les tailles d'échantillon déterminées à l'aide de calculs de puissance sont basées sur un test d'hypothèse bilatéral à 5%.

Modèle exponentiel

D'abord la taille des échantillons est considérée en fonction de la puissance. En supposant que l'essai a une période de recrutement de 3 ans avec une période de suivi de 2 ans et que 60% des patients recevant le traitement standard devraient être en vie après 5 ans, pour le calcul de la puissance, une augmentation absolue de 20% de la survie des patients est considérée pour le groupe expérimental. En utilisant l'équation (3.17), nous avons les paramètres du modèle $\lambda_1 = 0.102$ et $\lambda_2 = 0.0446$. Pour atteindre une puissance spécifiée π^* , la taille d'échantillon requise N est déterminée en résolvant

$$\pi^* = \pi^E(N)$$

où la fonction de puissance π^E est donnée dans l'équation (3.14).

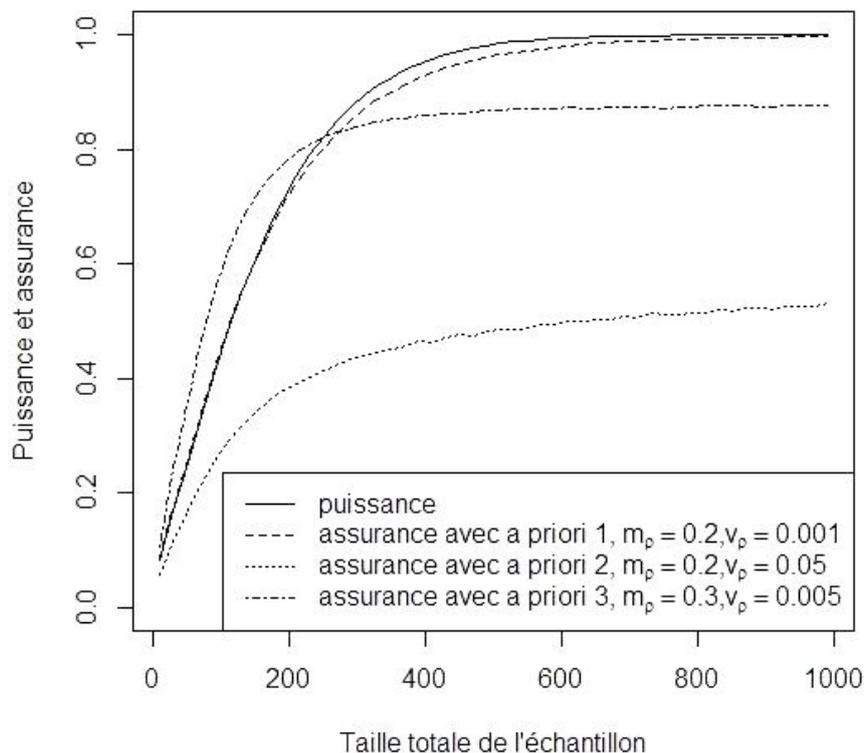


FIGURE 3.5 – La comparaison entre la puissance et l'assurance γ^E en considérant $\rho \sim N(m_\rho, v_\rho)$ et $S_1(5) \sim Be(60, 40)$.

Pour calculer l'assurance γ^E , les jugements d'un expert sur le taux de survie à 5 ans dans le groupe témoin $S_1(5)$ et la différence ρ de survie à 5 ans sont évalués à l'aide des méthodes univariées d'élicitation. Supposer que cela donne $\rho \sim N(m_\rho, v_\rho)$ et $S_1(5) \sim Be(a_s, b_s)$, dans ce qui suit, nous examinons trois scénarios pour les a priori :

Scénario 1 : $\rho \sim N(0.2, 0.001)$ et $S_1(5) \sim Be(60, 40)$.

Scénario 2 : $\rho \sim N(0.2, 0.05)$ et $S_1(5) \sim Be(60, 40)$.

Scénario 3 : $\rho \sim N(0.3, 0.005)$ et $S_1(5) \sim Be(60, 40)$.

Dans le scénario 1, l'a priori $\rho \sim N(0.2, 0.001)$ indique une forte croyance a priori que la différence de survie à 5 ans est d'environ 0,2. Dans le scénario 2, $v_\rho = 0.05$ implique que $P(S_2(5) - S_1(5) > 0) = 0.769$. Dans le scénario 3, l'a priori $\rho \sim N(0.3, 0.005)$ exprime la conviction

que $P(S_2(5) - S_1(5) < 0) = 0.00003$, c'est-à-dire que l'expérimentateur estime que le traitement expérimental a une très forte probabilité d'être supérieur.

La figure 3.5 montre en quoi les assurances diffèrent compte tenu des différentes distributions conjointes a priori. Lorsqu'un expert croit fermement (scénario 1) que l'effet du traitement sera proche de celui spécifié dans le calcul de la puissance, la taille d'échantillon requise informée par l'assurance est similaire à celle déterminée par le calcul de la puissance. Dans le scénario 2, l'assurance ne peut pas dépasser 80%, car la probabilité a priori que le nouveau traitement soit supérieur est de 76.9%. Dans le scénario 3, une taille d'échantillon plus petite peut être nécessaire pour atteindre une probabilité de 80% d'avoir un essai réussi, étant donné la très forte probabilité a priori que le traitement expérimental soit supérieur.

3.6 Conclusion

La simulation Bayésienne d'essai expérimental ou clinique est un outil générique qui peut calculer des assurances pour tout résultat d'essai, qu'il soit basé sur une analyse Bayésienne des données, des tests de signification fréquentiste ou une analyse de décision formelle telle qu'une décision d'un fournisseur de soins de la santé de rembourser l'utilisation d'un médicament pour des raisons de rentabilité. La fiabilité d'une probabilité d'assurance dépendra de la fiabilité de l'a priori obtenu et il sera donc important de vérifier la robustesse des assurances au choix de l'a priori. Il est clairement utile de connaître la probabilité qu'un essai produise un résultat positif, dans le contexte de la planification des essais cliniques ; l'effort supplémentaire requis pour utiliser la méthode d'assurance est relativement faible.

Chapitre 4

Modélisation de la prédiction Bayésienne en deux étapes

4.1 Introduction

La conception Bayésienne des expériences peut être considérée comme une combinaison naturelle de prédiction et de prise de décision, dans la mesure où les chercheurs recherchent la meilleure conception pour atteindre des objectifs ciblés, sur la base des connaissances a priori ou actualisées. Ils sont actuellement intéressés en raison de leur potentiel à économiser du temps et des ressources ainsi qu'à réduire le nombre d'événements indésirables [Spiegelhalter et al., 2004, Hand et al., 2016].

La procédure prédictive Bayésienne joue un rôle crucial dans différents domaines de la statistique appliquée, de l'épidémiologie, de la fiabilité et de l'analyse de survie dans le but de développer une conception adaptative. Plus généralement, ces procédures prédictives Bayésiennes sur les observations futures donnent au chercheur une méthode précise pour évaluer les chances que l'expérience aboutisse à un résultat concluant ou au contraire un résultat non concluant [Djeridi et Merabet, 2019]. Ils sont particulièrement adaptés aux conceptions séquentielles où l'analyse intermédiaire peut fournir des motifs pour mettre fin à l'étude de manière efficace en réduisant les coûts. De plus, elle peut bénéficier à la population

générale des patients en permettant une diffusion précoce de ses résultats. Les méthodes séquentielles du groupe fournissent un moyen d'équilibrer les avantages éthiques et financiers de l'arrêt précoce d'une étude par rapport au risque d'une conclusion incorrecte [Berry et al., 2011, Chow et al., 2007]. Les approches théoriques de décision pour la conception des essais tiennent compte du nombre fini des futurs patients pour lesquels les résultats de l'essai seront définitifs [Shih et Lavori, 2013].

Dans ce chapitre, nous exploitons la distribution prédictive Bayésienne dans des essais cliniques expérimentaux en deux étapes pour les données distribuées de Poisson ou de Gamma. Les distributions Gamma donnent des représentations utiles de nombreuses situations physiques. Par exemple, elles ont été utilisées pour apporter des ajustements réalistes aux distributions exponentielles dans la représentation des durées de vie [Blot et Meeter, 2016]. De plus, les résultats de Gamma ont une implémentation étendue dans divers domaines : En hydrologie, en finance, en sciences médicales, en psychologie et en fiabilité. On peut trouver d'autres exemples avec des données de Poisson et Gamma dans [Bakhshi, 2011 ; Hamada et a.l, 2008 ; Lakshmi et Vaidyanathan, 2015 ; Anisimov, 2008 ; Zaslavsky, 2010].

Nous nous intéressons aux données de dénombrement « en anglais : count data » comme résultat. La distribution de Poisson est un modèle naturel couramment utilisé. Par exemple, Hand et al. (2016) utilisent les données de dénombrement dans des études à un seul bras comme critère d'évaluation principal pour démontrer l'efficacité des préparations commerciales d'immunoglobulines intraveineuses (IGIV) afin de prévenir les infections bactériennes aiguës graves chez les patients présentant une immunodéficience primaire.

L'objectif principal de ce travail est de proposer une procédure basée sur le concept d'indice de satisfaction lié au test d'hypothèse qui est une fonction de la p-valeur ou de la probabilité d'observer un résultat comme extrême ou plus extrême que le résultat perçu sous l'hypothèse nulle. Compte tenu des données disponibles, nous calculons un indice de satisfaction prévu et conditionné aux observations précédentes pour

obtenir des résultats positifs à la fin de l'essai [Merabet, 2013]. Nous considérons le cas d'une procédure, en deux étapes avec des données distribuées de Poisson ou de Gamma, qui permet une interruption précoce en raison de la futilité ou de l'efficacité extrême.

4.2 La fiabilité dans les essais expérimentaux

La fiabilité peut se définir comme « l'aptitude d'une entité ou d'un système à accomplir une fonction requise dans des conditions données et pour une période de temps donnée ». Par exemple, un fabricant d'aspirateur souhaite que les composants électroniques qu'il contient soient garantis pendant 5 ans dans ses conditions d'utilisation [Gilles Rousseau, 2016].

La fiabilité a sans doute pris son développement depuis la dernière guerre mondiale. Elle est vite devenue une science à part entière dans les applications appartenant à de nombreux domaines. Elle a pour fondements mathématiques la statistique et le calcul des probabilités qui sont nécessaires à la compréhension et l'analyse des données de fiabilité [Bellaouar et Beleulmi, 2013]. De nombreuses lois sont appliquées à l'analyse de la fiabilité dans plusieurs domaines dont la loi de Poisson, loi exponentielle, loi de Gamma...ect

Plus les caractéristiques d'un produit ou d'un système sont appréhendées tôt dans son cycle de vie, moins les risques financiers ou reliés à la sûreté des installations dûs à la non réalisation des performances attendues sont élevés. Dans un contexte d'exigence de systèmes de plus en plus fiables, sûrs et de durées de garanties croissantes, il est impératif de vérifier le plus tôt possible que les performances des systèmes sont conformes au cahier des charges.

L'idéal, pour identifier la fiabilité du produit ou du système avant même sa fabrication en série, est de procéder de façon classique à des

séries d'essais sur des prototypes quand ils existent. Le problème est l'investissement demandé en temps et en quantité de matériel important car les matériels sont de plus en plus fiables, l'observation des défaillances est de moins en moins probable. Les industriels ne peuvent plus se permettre de tels coûts financiers. A l'extrême, certains systèmes se fabriquent à l'unité, ce qui rend les politiques d'essai difficiles.

Ainsi, cette problématique a été la source, pour la communauté scientifique, de nombreuses voies de recherche. Celles-ci sont basées principalement sur la modélisation stochastique des apparitions des défaillances au cours du temps et sur l'estimation statistique des paramètres des modèles à partir des résultats d'essai. Les premiers travaux ont consisté à étudier les essais réalisés en conditions nominales (on reproduit les conditions normales d'utilisation du produit) provoquant des temps d'essai importants avec des tailles d'échantillon élevées. Par la suite des travaux ont été menés pour réduire les tailles des échantillons et le temps d'essai [Guérin Fabrice, 2004].

Dans l'analyse de fiabilité Bayésienne, le modèle statistique se compose de deux parties : La fonction de vraisemblance et la distribution a priori. La fonction de vraisemblance est généralement construite à partir de la distribution d'échantillonnage des données, définie par la fonction de densité de probabilité supposée pour les données. Par exemple, il peut s'agir d'une distribution d'échantillonnage pour des données de la durée de vie. La distribution d'échantillonnage contient généralement des paramètres inconnus. Une fois que nous avons effectué l'expérience et observé son résultat, nous considérons la distribution d'échantillonnage comme une fonction des paramètres inconnus. Cette fonction (ou toute fonction qui lui est proportionnelle) est appelée fonction de vraisemblance [Hamada et al., 2008].

L'inférence Bayésienne est le seul cadre d'inférence statistique qui obéit systématiquement au principe de vraisemblance. En termes simples, le principe de vraisemblance stipule que toutes les informations contenues dans les données expérimentales sont contenues dans la densité d'échantillonnage des données observées.

Les méthodes Bayésiennes nous permettent d'intégrer et d'utiliser des informations autres que celles contenues dans les données expérimentales. Qu'un analyste de fiabilité dispose ou non de telles données de test, il disposera souvent d'autres informations pertinentes sur la valeur des paramètres de fiabilité inconnus. Ces informations pertinentes sont une composante extrêmement utile et puissante de l'approche Bayésienne, en outre les estimations des paramètres Bayésiens réfléchies reflètent cette connaissance. Ces informations pertinentes sont souvent dérivées de combinaisons de sources telles que la théorie physique / chimique, les résultats des tests d'ingénierie et de qualification, les données de fiabilité génériques à l'échelle de l'industrie, l'analyse informatique, l'expérience passée avec des dispositifs similaires, les résultats de tests précédents obtenus à partir d'un programme de développement de processus et jugement du personnel expérimenté.

Une fois les données de test obtenues, la distribution a posteriori décrit pleinement l'incertitude associée au paramètre. Nous calculons la distribution a posteriori via le théorème de Bayes en utilisant la fonction de vraisemblance et la distribution a priori. La séquence logique de la fonction de vraisemblance, de la distribution a priori, du théorème de Bayes et de la distribution a posteriori rend les méthodes de fiabilité Bayésiennes faciles à décrire et les estimations dérivées faciles à interpréter et à utiliser [Hamada et al., 2008].

4.3 Procédure de test dans un cadre hybride

Sachant que le contexte expérimental est constitué de deux expérimentations successives, les résultats $x \in \Omega'$ et $y \in \Omega''$ sont en général effectués indépendamment. Leurs distributions, construites dans le cadre d'un modèle bien établi, dépendent d'un paramètre $\theta \in \Theta$, seul y est utilisé pour trouver la conclusion officielle de l'étude et pour détermi-

ner la satisfaction de l'utilisateur notée $\varphi(y)$ (sur le choix de la fonction décroissante L sur laquelle nous reviendrons dans la Section 4.2.2).

Notons que :

P_Θ : La distribution a priori sur Θ .

P_Θ^x : La distribution a posteriori sur Θ , basée sur le résultat de la première étape.

$P_{\Omega''}^\theta$: La vraisemblance de la deuxième étape.

$P_{\Omega''}^x$: Distribution de probabilité sur Ω'' , conditionnée par le résultat de la première étape.

4.3.1 Test de Neymann-Pearson

Dans ce contexte expérimental, nous allons introduire le concept de l'indice de satisfaction relatif à un test d'hypothèse où l'hypothèse nulle est de type $\theta \leq \theta_0$ dans un cadre où un tel test peut être construit en utilisant une fonction de test raisonnable. Soit un modèle $(P_\theta)_{(\theta \in \Theta)}$ et testons une hypothèse nulle Θ_0 contre une alternative Θ_1 , définie pour une application $\psi(\Theta \rightarrow R)$: Supposons qu'il existe un point t_0 tel que :

$$\theta \in \Theta_0 \iff \psi(\theta) \leq t_0$$

Sinon, nous supposons qu'on a une application réelle $\xi(\Omega'' \rightarrow R)$ telle que :

$$\psi(\theta_1) < \psi(\theta_2) \implies \forall t; P_{\theta_1}[\xi \leq t] \geq P_{\theta_2}[\xi \leq t].$$

Alors un test de niveau α de Θ_0 contre $\Theta_1 = \{\theta; \psi(\theta) > t_0\}$ est défini par le rejet de l'hypothèse nulle si le résultat expérimental, y , vérifie que $\xi(y) > g(\alpha)$, où $g(\alpha)$ est $1 - \alpha$ quantile de la distribution de ξ quand $\psi(\theta) = t_0$. De plus, la région critique, R , de ce test est l'ensemble des observations, y , tel que $\xi(y) > g(\alpha)$. Puisque les distributions de ξ sont stochastiquement croissantes, pour tout θ_0 tel que $\psi(\theta_0) = t_0$:

$$\forall \theta \in \Theta_0, \quad P_\theta(R) \leq P_{\theta_0}(R) \leq \alpha$$

En outre, si $\theta_1 \in \Theta_1$ et $\theta_2 \in \Theta_1$ avec $\psi(\theta_1) < \psi(\theta_2)$, alors $P_{\theta_1}(R) \leq P_{\theta_2}(R)$ qui signifie que la fonction de puissance augmente avec $\psi(\theta)$.

4.3.2 Indice de satisfaction

Cette notion trouve son origine dans des situations où le statisticien, qui effectue un test, « souhaite » détecter un effet significatif, i.e. pour rejeter l'hypothèse nulle H_0 . De manière correspondante, ce statisticien est surtout plus satisfait si, en fonction des résultats expérimentaux, cet effet semble être plus significatif. En fait, on considère la satisfaction apportée par la seconde étape de l'expérimentation prédit à l'aide de la première, on fait les calculs avec deux étapes indépendantes mais c'est généralisable (en particulier la deuxième étape peut intégrer la première). On définit, tout d'abord, l'indice fruste de Grouin puis on donnera l'indice amélioré de Merabet.

Indice fruste

Étant fixé α , soit un test de niveau α défini par la région critique $\Omega_1''^{(\alpha)}$. Un premier indice de satisfaction, celui étudié par Grouin (1994), est défini par :

$$\varphi(y) = 1_{\Omega_1''^{(\alpha)}}(y)$$

à x fixé, la prédiction est, alors :

$$\pi(x) = P_{\Omega''}^x \left(\Omega_1''^{(\alpha)} \right) = \int_{\Omega_1''^{(\alpha)}} P_{\Omega''}^{\theta} \left(\Omega_1''^{(\alpha)} \right) P_{\Theta}^x(d\theta)$$

où $P_{\Omega''}^{\theta} \left(\Omega_1''^{(\alpha)} \right)$ est la valeur en θ de la puissance du test. Le défaut de l'indice fruste est qu'il exprime une satisfaction en « tout ou rien » (significatif ou non).

Indice amélioré

Il est plus intéressant de prendre en compte jusqu'à quel niveau le résultat apparaît toujours significatif. C'est ce que les utilisateurs soulignent en donnant, à la fin de la procédure de test, non seulement la conclusion en termes de « tout-rien » mais aussi la valeur du seuil inférieure pour laquelle le résultat obtenu sera considéré comme significatif, c'est-à-dire du point de vue de la théorie du test, la p-valeur qui est dans notre cas :

$$p = P_{\theta_0}(\xi > \xi(y))$$

Merabet (2004) a défini un indice amélioré (IS), pour ce test de niveau α , comme une application de l'ensemble de résultats dans R_+ ; telle que :

- Prend la valeur 0 si on ne rejette pas l'hypothèse nulle, i.e. si $\xi(y) \leq g(\alpha)$.
- Et si $\xi(y) \leq g(\alpha)$ est une fonction décroissante $P_{\theta_0}(\xi > \xi(y))$, qu'on notera $L(P_{\theta_0}(\xi > \xi(y))) = L(1 - F_{\theta_0}(\xi(y)))$, où F_{θ_0} est la distribution de ξ au frontière telle que $\psi(\theta_0) = t_0$.

4.3.3 Prédiction de satisfaction

Étant fixé α , soit un test de niveau α défini par la région critique $\Omega_1''^{(\alpha)}$. Il est plus intéressant de prendre en compte jusqu'à quel niveau le résultat apparaît toujours significatif. Nous utilisons l'indice de satisfaction qui est défini dans les tests Bayésiens, basé sur la même a priori P_{Θ} , comme suit :

$$\varphi(y) = \begin{cases} 0 & \text{si } y \in \Omega_0''^{(\alpha)} \\ 1 - \inf\{\beta, y \in \Omega_0''^{(\beta)}\} & \text{si } y \in \Omega_1''^{(\alpha)} \end{cases} \quad (4.1)$$

Une situation standard est qu'il existe une application $\psi(\Theta \rightarrow R)$ telle que $\Theta_0 = \{\theta; \psi(\theta) \leq t_0\}$ et où il existe, aussi, $\xi(\Omega'' \rightarrow R)$ et $g([0, 1[\rightarrow R)$

telles que :

$$\Omega_1''^{(\alpha)} = \{y; \xi(y) \leq g(\alpha)\}$$

où $g(\alpha)$ est le $(1-\alpha)$ -quantile de la distribution de ξ quand $\psi(\theta_0) = t_0$. De plus, on suppose que la distribution de ξ sous $P_{\Omega''}^\theta$ dépend seulement de $\psi(\theta)$ (qu'on notera $Q_{\psi(\theta)}$ qui est stochastiquement croissante, de tel sorte que ξ tend de plus en plus à prendre de grands valeurs quand $\psi(\theta)$ devient de plus en plus grand.

Un test de niveau α de Θ_0 est, alors, défini par le rejet l'hypothèse si le résultat expérimental, qui est y , vérifie $\xi(y) > g(\alpha)$. Il semble, donc, naturel d'utiliser la p-valeur pour définir les indices de satisfaction nuls si aucun effet significatif n'est détecté et, dans le cas contraire, une fonction croissante de l'indicateur classique de signification. Dans ce cas :

$$p(y) = P_{\theta_0}(\xi > \xi(y))$$

Un indice de satisfaction est, alors, défini naturellement par :

$$\varphi(y) = \begin{cases} 0 & \text{si } \xi(y) \leq g(\alpha) \\ L(p(y)) & \text{si } \xi(y) > g(\alpha) \end{cases}$$

où L est une fonction décroissante. Soit F_{θ_0} une distribution de ξ sur la frontière, i.e. pour tout θ_0 tel que $\psi(\theta_0) = t_0$, l'indice de satisfaction est défini par :

$$\varphi(y) = \begin{cases} 0 & \text{si } p(y) \geq 1 - \alpha \\ L(1 - F_{\theta_0}(y)) & \text{sinon} \end{cases}$$

La prédiction Bayésienne est, alors, donnée par :

$$\begin{aligned} \pi(x) &= \int_{\Omega_1''^{(\alpha)}} \varphi(y) P_{\Omega''}^x(dy) \\ &= \int_{\Theta} \left(\int_{\Omega_1''^{(\alpha)}} \varphi(y) P_{\Omega''}^\theta(y) \right) P_{\Theta''}^x(d\theta) \end{aligned}$$

$$= \int_{\{y; \xi(y) > g(\alpha)\}} L(1 - F_{\theta_0}(y)) P_{\Omega''}^x(dy)$$

Notons que $\left(\int_{\Omega_1^{(\alpha)}} \varphi(y) P_{\Omega''}^{\theta}(y) \right)$ généralise la puissance de test dans la logique d'indice de satisfaction proposé. On peut généraliser cette procédure à une famille d'indices limités définie par $L(p) = (1 - p)^l$ où $l \geq 0$. Il est préférable de choisir les indices limités car ils sont faciles à interpréter.

Remarque 1. 1. Si $l = 1$, $1 - \varphi(y)$ est la p -valeur.

2. Dans le cas où $l = 0$, on trouve la fonction indicatrice de la région critique qui est l'indice fruste de Grouin (Lecoutre et al. (1995)).

4.4 Conception en deux étapes dans le cadre séquentiel

4.4.1 Inférence statistique pour un plan expérimental en deux étapes

Ces dernières années, le concept d'un design en deux étapes a conduit au développement des conceptions séquentielles de groupe, où il existe un ensemble d'expériences disponibles qui peuvent être menées.

Après chaque étape, la décision est prise de poursuivre ou non l'expérience, et si c'est le cas, de construire un nouveau design pour l'étape suivante. Étant donné que certaines expériences peuvent être plus « informatives » que d'autres, il est possible d'économiser en utilisant des règles de décision qui choisissent des expériences plutôt qu'en utilisant des règles de décision qui prennent généralement une expérience de chaque type [Blot et Meeter, 2016].

Dans la conception expérimentale séquentielle, nous choisissons séquentiellement une expérience à effectuer et de nouveaux résultats sont

observés. Nous utilisons ici deux expériences de résultats successives et notons $x \in X$ et $y \in Y$; les résultats de chaque expérience qui sont supposés indépendants. Les distributions de x et y dépendent d'un paramètre commun $\theta \in \Theta$.

Nous proposons un indice de satisfaction basé à la fois sur la première et la deuxième phase de l'expérience [Merabet et al, 2017]. Cette expérience est utilisée pour établir la conclusion finale de l'étude et déterminer la satisfaction de l'utilisateur, notée $\varphi(y)$. Dans notre étude, la prédiction est réalisée dans un contexte Bayésien ; c'est-à-dire basée sur le choix d'une probabilité a priori sur Θ . Étant donné que les indices de satisfaction sont essentiels à l'étude de test fréquentiste et Bayésien, nous considérons à la fois une approche hybride fréquentiste-Bayésienne et une approche entièrement Bayésienne. Ces indices sont utilisés comme règle d'arrêt pour la conception des essais cliniques de phase II.

Indice Fréquentiste

Nous définissons l'indice de satisfaction fréquentiste dans le test classique comme une mesure du degré de satisfaction avec un résultat donné. Nous voulons tester l'hypothèse nulle $H_0 : \theta \in \Theta_0$ contre $H_1 : \theta \in \Theta \setminus \Theta_0$ sur le paramètre θ . La région de rejet R^α de test fréquentiste habituel au niveau α est définie par :

$$R^\alpha = \{z ; p(z) \leq \alpha\}$$

où $p(z)$ est la p-valeur du test.

Un indice de satisfaction de base est défini par :

$$\varphi_0(z) = 1_{R^\alpha}(z) \tag{4.2}$$

où $1_A(z) = 1$ si $z \in A$ et 0 sinon.

L'indice $\varphi_0(z)$ a une forme qui exprime une satisfaction « tout ou rien ». Cependant, il est intéressant de prendre en considération le niveau auquel le résultat apparaît significatif. Il apparaît donc naturel de

considérer comme des indices de satisfaction nuls si aucun effet significatif n'est détecté, et dans le cas contraire comme une fonction croissante de l'indicateur classique de signification qui est la p-valeur [Saville et al, 2014]. À cette fin, la p-valeur est considérée comme une mesure de crédibilité à attacher à l'hypothèse nulle que les praticiens utilisent souvent pour répondre à plusieurs critiques et inconvénients de l'approche de Neyman-Pearson. Par conséquent, nous proposons un indice de satisfaction considéré comme amélioré pour son intérêt dans le concept de prédiction de la satisfaction et qui est défini comme une fonction décroissante de la mesure concluante p après le traitement des données de la manière suivante :

$$\varphi(z) = \begin{cases} 0 & \text{si } p(z) \geq \alpha \\ 1 - p(z) & \text{sinon} \end{cases} \quad (4.3)$$

Indice Bayesian

Dans un contexte Bayésien, soit $\Pi(\Theta|z)$ la probabilité a posteriori étant donnée l'observation z . Il est classique dans les statistiques Bayésiennes de traiter le test de situation de Θ_0 contre Θ_1 en fournissant $\Pi(\Theta_1|z)$. On note \tilde{R}^α la région de rejet de test Bayésien, au niveau α , est définie par :

$$\tilde{R}^\alpha = \{z; \Pi(\Theta_1|z) \geq 1 - \alpha\}$$

De même à l'approche fréquentiste, nous considérons les deux indices de satisfaction :

$$\tilde{\varphi}_0(z) = 1_{\tilde{R}^\alpha}(z) \quad (4.4)$$

et

$$\tilde{\varphi}(z) = \begin{cases} 0 & \text{si } z \notin \tilde{R}^\alpha \\ \Pi(\Theta_1|z) & \text{si } z \in \tilde{R}^\alpha \end{cases} \quad (4.5)$$

4.4.2 Préviation de l'indice de satisfaction et des règles d'arrêt

La demande des règles d'efficacité où le bénéfice du traitement peut être basé soit sur des distributions Bayésiennes a posteriori (entièrement Bayésiennes), soit sur des critères fréquentistes tels que les p-valeurs (approche hybride Bayésienne-fréquentiste) [Saville et al., 2014]. Les principales questions sur lesquelles se concentrer sont les raisons éligibles d'arrêter ou de poursuivre une étude. Les raisons de l'arrêt peuvent inclure [Todd et al., 2011] :

- Le traitement expérimental est évidemment pire que le contrôle.
- Le traitement expérimental est déjà évidemment mieux.
- Les chances de montrer que le traitement expérimental est meilleur sont faibles.

La poursuite peut inclure les raisons suivantes :

- Un avantage modéré du traitement expérimental est probable et il est souhaitable d'évaluer la magnitude attentivement.
- Le taux d'événements est faible et plus de patients sont nécessaires pour atteindre une puissance donnée.

De nombreux auteurs ont été forts et cohérents en recommandant l'utilisation des probabilités prédictives pour prendre des décisions basées sur l'accumulation des données d'essais expérimentaux. Dans la logique de l'introduction de l'indice de satisfaction, il est naturel de caractériser la valeur de la procédure de test au lieu de la fonction puissance, un indice de prédiction, c'est-à-dire l'anticipation mathématique par rapport à la probabilité prédictive sur les données complètes conditionnées par le résultat de la première étape [Merabet, 2013], où une expérience en deux étapes doit être menée :

- Un premier résultat x , détermine si nous continuons ou non l'expérience.

- Si l'expérimentateur est très satisfait et que nous continuons effectivement l'expérience, alors le test final est basé sur le résultat des deux étapes plutôt que sur le résultat de la seconde étape unique.

Soit $p(z|x)$ la probabilité prédictive de $z = (x, y)$ après la première étape, c'est-à-dire conditionnellement sur x . La valeur prédite de l'indice φ est définie par :

$$\eta(x) = E(\varphi(z)|x) = \int \varphi(z)p(z|x)dx \quad (4.6)$$

L'indice prédit $\eta(x)$ généralise la puissance du test dans la dialectique de l'indice de satisfaction. Le praticien décide d'un indice prédit au-dessus duquel l'expérience se poursuit. D'un point de vue purement Bayésien, l'indice prévu est :

$$\tilde{\eta}(x) = E(\tilde{\varphi}(z)|x) = \int \tilde{\varphi}(z)p(z|x)dx \quad (4.7)$$

4.5 Conception en deux étapes pour les résultats de Poisson et de Gamma

4.5.1 Une stratégie Bayésienne en deux étapes pour les résultats de Poisson

Dans les essais cliniques expérimentaux utilisant des données de dénombrement d'événements rares, la distribution de Poisson est un modèle naturel et couramment utilisé. L'un des objectifs essentiels de la plupart des études cliniques de phase II est de décider de poursuivre un essai de phase III randomisé à grande échelle ou de reconsidérer ou d'abandonner le traitement en raison d'une absence d'efficacité ou d'évaluer la toxicité [Stallard, 1998]. Si un traitement peut donner des résultats inférieurs à ceux attendus, certaines études de phase II utilisent un plan en deux étapes pour permettre une interruption précoce. Nous considérons le cas

d'un essai de Phase II dont le critère principal est le nombre d'événements observés sur une période de temps fixe, où ce dénombrement a une distribution de Poisson [Hand et al., 2016]. Nous supposons également que les événements considérés ici montrent un résultat négatif pour les patients et, par conséquent, un grand nombre d'événements indiquent une absence d'efficacité. Par conséquent, le plan en deux étapes pour les données de Poisson est le suivant :

Soit n_1 et n_2 le nombre de sujets inclus dans la première et la deuxième étape et $n = n_1 + n_2$. Soit x_j , resp. y_j , le nombre d'occurrence des événements d'intérêt pour le $j^{\text{ème}}$ sujet à la première étape, resp. la deuxième étape, pendant une période de temps t_{1j} , resp. t_{2j} . Nous supposons que x_j et y_j sont indépendants et distribués par Poisson avec les paramètres $\theta \times t_{1j}$ et $\theta \times t_{2j}$. Soit $x = \sum_{j=1}^{n_1} x_j$ et $y = \sum_{j=1}^{n_2} y_j$ le nombre total d'événements observés à la première et à la deuxième étape et $t_i = \sum_{i=1}^{n_i} t_{ij}$, où $i = 1, 2$. On définit ici $z = x + y$ comme le nombre d'événements sur les deux étapes et $t = t_1 + t_2$. Soit s et s_1 des seuils pré-spécifiés tels que :

1. Si $x \geq s_1$, alors l'essai est arrêté par manque d'efficacité.
2. sinon, l'essai de phase II se poursuit jusqu'à la deuxième étape.

Si l'essai continue jusqu'à la deuxième étape, n_2 patients sont inscrits :

1. Si $z \geq s$, l'essai est arrêté par manque d'efficacité.
2. Autrement, l'essai de phase II se poursuit par un essai de phase III.

Dans un autre sens, sur la base des données accumulées à l'étape 1, on peut arrêter l'essai par manque d'efficacité convaincante. Cela se produit lorsque l'indice de satisfaction prévu est inférieur à un seuil spécifié. Sinon, l'expérience continue. Plus rarement, un essai peut être interrompu pour de solides preuves d'efficacité, c'est-à-dire lorsque l'indice de satisfaction prévu est supérieur à un seuil donné. Un arrêt précoce signifie que les ressources ne sont pas gaspillées et, en cas de preuve d'efficacité, permet un développement plus rapide.

Dans la modélisation Bayésienne, le choix d'une distribution a priori est crucial car elle a potentiellement une grande influence sur la densité

a posteriori, surtout lorsque les comptes (en anglais counts) d'intérêt observés et collectés sont faibles. Les chercheurs tentent de trouver une distribution a priori qui résume les informations disponibles et reflète avec précision l'incertitude. Dans l'analyse a posteriori, nous souhaitons généralement que la probabilité domine et, par conséquent, encourageons l'utilisation d'un a priori relativement non informatif. Ici, nous utilisons une famille Gamma conjuguée d'a priori pour θ avec les paramètres (a,b) [Hand et al, 2016].

Soit $x \sim Poisson(t_1\theta)$ avec fonction de probabilité :

$$f(x|t_1, \theta) = \frac{(t_1\theta)^x}{x!} \exp[-t_1 \theta], \quad x = 0, 1, 2, \dots \quad (4.8)$$

La distribution a posteriori $\Pi(\theta|x, t_1)$ est une distribution Gamma $(x + a, t_1 + b)$:

$$\Pi(\Theta_1|z) = \int_0^\infty \frac{(t_1\theta)^x \exp[-t_1 \theta]}{x!} \left[\frac{b^a \theta^{(a-1)}}{\Gamma(a)} \exp[-\theta b] \right] d\theta$$

$$\Pi(\Theta_1|z) = \frac{t_1^x b^a}{x! \Gamma(a)} \int_0^\infty \theta^x \exp[-t_1 \theta] \theta^{(a-1)} \exp[-\theta b] d\theta$$

$$= \frac{t_1^x b^a}{x! \Gamma(a)} \int_0^\infty \theta^{(a+x-1)} \exp[-\theta(b + t_1)] d\theta$$

$$= \frac{t_1^x b^a}{x! \Gamma(a)} \frac{\Gamma(a + x)}{(b + t_1)^{a+x}}$$

$$\Pi(\Theta_1|z) = \frac{(b_1)^{a_1}}{\Gamma(a_1)} \theta^{(a_1-1)} \exp[-\theta(b_1)]$$

Donc on a l'a posteriori $\Pi(\theta|x, t_1) \sim Ga(a_1, b_1)$ où $a_1 = a + x$ et $b_1 = b + t_1$

Si $x < s_1$, nous passons à la deuxième étape où les comptes y sont observés pendant le temps de personne t_2 .

Ainsi $y \sim \text{Poisson}(t_2\theta)$ avec fonction de probabilité :

$$f(y|t_2, \theta) = \frac{(t_2\theta)^y}{y!} \exp[-t_2\theta], \quad y = 0, 1, 2, \dots \quad (4.9)$$

Considérons le test unilatéral $H_0 : \theta \geq \theta_0$ contre $H_1 : \theta < \theta_0$ où θ_0 est le seuil d'efficacité souhaité. Ensuite, nous déterminons t et s , étant donné que la première étape est terminée et que x ont été observés avec $x < s_1$.

Le test habituel sur les résultats z de la première et la deuxième étape est défini par :

$$R^\alpha = \{z; z < q_\alpha\}$$

avec $q_\alpha = \sup\{c; Pr(z < c|\theta_0) \leq \alpha\}$

Dans l'approche Bayésienne :

$$\tilde{R}^\alpha = \{z; z < \tilde{q}_\alpha\}$$

avec $\tilde{q}_\alpha = \sup\{z; \Pi(\theta < \theta_0|z) \geq 1 - \alpha\}$

Nous avons ces formules pour les conceptions séquentielles :

$$\varphi_0(z) = 1_{R^\alpha(z)},$$

$$\tilde{\varphi}_0(z) = 1_{\tilde{R}^\alpha(z)},$$

$$\varphi(z) = (1 - \Phi_{\theta_0}^P(z)) 1_{z < q_\alpha},$$

$$\tilde{\varphi}(z) = \tilde{\Phi}_{a+z; b+t}(\theta_0) 1_{z < \tilde{q}_\alpha},$$

où $\Phi_{\theta_0}^P(z)$ est la fonction de distribution cumulative continue à gauche d'une distribution de Poisson et $\tilde{\Phi}_{(a_1, b_1)}$ est la fonction de distribution cumulative d'une distribution $\text{Gamma}(a_1, b_1)$. La distribution prédictive de y étant donnée x est une distribution Gamma-Poisson donnée par :

$$f(y|x) = \int_0^\infty f(y|\theta)f(\theta|x)d\theta$$

$$\begin{aligned}
 f(y|x) &= \int_0^\infty \frac{(t_2\theta)^y \exp[-t_2\theta]}{y!} \frac{(b_1)^{a_1}}{\Gamma(a_1)} \theta^{(a_1-1)} \exp[-\theta(b_1)] d\theta \\
 &= \frac{t_2^y b_1^{a_1}}{y! \Gamma(a_1)} \int_0^\infty \theta^y \exp[-t_2\theta] \theta^{(a_1-1)} \exp[-\theta b_1] d\theta \\
 &= \frac{t_2^y b_1^{a_1}}{y! \Gamma(a_1)} \int_0^\infty \theta^{a_1+y-1} \exp[-\theta(b_1 + t_2)] d\theta \\
 f(y|x) &= \frac{t_2^{(y)} (b + t_1)^{(a+x)} \Gamma[a + x + y]}{(b + t_1 + t_2)^{(a+x+y)} \Gamma[a + x] y!} \quad (4.10)
 \end{aligned}$$

Ensuite, nos indices de satisfaction prédictifs pour l'approche fréquentiste sont :

$$\eta_0(x) = E(\varphi_{0(\alpha)}(z)|x) = \sum_{y=0}^{q_\alpha-x-1} \frac{t_2^{(y)} (b + t_1)^{(a+x)} \Gamma[a + x + y]}{(b + t_1 + t_2)^{(a+x+y)} \Gamma[a + x] y!} \quad (4.11)$$

$$\eta(x) = E(\varphi_{(\alpha)}(z)|x) = \sum_{y=0}^{q_\alpha-x-1} (1 - \Phi_{\theta_0}^P(x+y)) \frac{t_2^{(y)} (b + t_1)^{(a+x)} \Gamma[a + x + y]}{(b + t_1 + t_2)^{(a+x+y)} \Gamma[a + x] y!} \quad (4.12)$$

Pour l'approche Bayésienne, nous avons :

$$\tilde{\eta}_0(x) = E(\tilde{\varphi}_{0(\alpha)}(z)|x) = \sum_{y=0}^{\tilde{q}_\alpha-x-1} \frac{t_2^{(y)} (b + t_1)^{(a+x)} \Gamma[a + x + y]}{(b + t_1 + t_2)^{(a+x+y)} \Gamma[a + x] y!} \quad (4.13)$$

$$\tilde{\eta}(x) = E(\tilde{\varphi}_{(\alpha)}(z)|x) = \sum_{y=0}^{\tilde{q}_\alpha-x-1} \tilde{\Phi}_{a+x+y; b+t_1+t_2}(\theta_0) \frac{t_2^{(y)} (b + t_1)^{(a+x)} \Gamma[a + x + y]}{(b + t_1 + t_2)^{(a+x+y)} \Gamma[a + x] y!} \quad (4.14)$$

4.6 Une stratégie Bayésienne en deux étapes pour les résultats de Gamma

Nous considérons ici le cas des familles conjuguées Gamma-Gamma. La distribution Gamma est une famille flexible de distribution pour les variables aléatoires continues non négatives. Les distributions Gamma sont utilisées dans de nombreux domaines tels que la finance, la science médicale, la modélisation des temps d'attente, la fiabilité et la modélisation des temps de service.

À la première étape, les sujets n_1 et n_2 sont inclus dont les réponses individuelles sont x_i et y_i . Soit $x = \sum_{i=1}^{n_1} x_i$, $y = \sum_{i=1}^{n_2} y_i$, $N_1 = n_1 p$ et $N_2 = n_2 p$, avec p connu. La distribution de x est une distribution Gamma avec des paramètres $(n_1 p, \theta)$ et une fonction de probabilité :

$$f(x, n_1 p, \theta) = \frac{\theta^{n_1 p}}{\Gamma(n_1 p)} x^{n_1 p - 1} \exp[-\theta x], \quad x = 0, 1, 2, \dots \quad (4.15)$$

Nous supposons la distribution a priori $\theta \sim \text{Gamma}(a, b)$. La distribution a posteriori sur θ après la première étape est une distribution Gamma $(a + n_1 p, b + x)$.

$$\begin{aligned} f(\theta|x) &= \int_0^\infty \frac{\theta^{n_1 p}}{\Gamma(n_1 p)} x^{n_1 p - 1} \exp[-\theta x] \left(\frac{b^a}{\Gamma(a)} \theta^{a-1} \exp[-b\theta] \right) d\theta \\ &= \frac{b^a x^{n_1 p - 1}}{\Gamma(a)\Gamma(n_1 p)} \int_0^\infty \theta^{n_1 p} \exp[-\theta x] \theta^{a-1} \exp[-b\theta] d\theta \\ &= \frac{b^a x^{n_1 p - 1}}{\Gamma(a)\Gamma(n_1 p)} \int_0^\infty \theta^{a+n_1 p - 1} \exp[-\theta(x+b)] d\theta \\ f(\theta|x) &= \frac{(b+x)^{a+n_1 p}}{\Gamma(a+n_1 p)} \theta^{a+n_1 p - 1} \exp[-\theta(x+b)] \end{aligned}$$

Bien qu'il existe une vaste littérature disponible sur l'estimation des paramètres d'une loi Gamma en utilisant l'approche fréquentiste, peu de

travaux ont été effectués sur l'inférence Bayésienne des paramètres d'une loi Gamma.

L'échantillonnage de la deuxième étape est également une distribution gamma $y \sim \text{gamma}(n_2p, \theta)$:

$$f(y, n_2p, \theta) = \frac{\theta^{n_2p}}{\Gamma(n_2p)} y^{n_2p-1} \exp[-\theta y], \quad y = 0, 1, 2, \dots \quad (4.16)$$

Définissez $z = x + y$ qui est une statistique suffisante Gamma distribuée avec des paramètres (np, θ) et $n = n_1 + n_2$ est le nombre total de sujets à traiter ou d'événements jusqu'à l'échec en fonction du domaine d'application :

$$f(z, np, \theta) = \frac{\theta^{np}}{\Gamma(np)} z^{(np)-1} \exp[-(\theta z)], \quad x = 0, 1, 2, \dots \quad (4.17)$$

Nous considérons le test unilatéral $H_0 : \theta \leq \theta_0$ contre $H_1 : \theta > \theta_0$ où θ_0 est le seuil d'efficacité souhaité.

Le test habituel sur les résultats z de la première et la deuxième étape est défini par :

$$R^\alpha = \{z; z < q_\alpha\}$$

avec $q_\alpha = \sup\{c; Pr(z > c | \theta_0) \geq 1 - \alpha\}$

Dans l'approche Bayésienne

$$\tilde{R}^\alpha = \{z; z < \tilde{q}_\alpha\}$$

avec $\tilde{q}_\alpha = \sup\{z; \Pi(\Theta_1 | z) \geq 1 - \alpha\}$

Nous avons ces formules pour les conceptions séquentielles :

$$\varphi_0(z) = 1_{R^\alpha(z)}$$

$$\tilde{\varphi}_0(z) = 1_{\tilde{R}^\alpha(z)}$$

$$\varphi(z) = 1 - \Phi_{\theta_0}^G(z)_{1_{z < q_\alpha}}$$

$$\tilde{\varphi}(z) = 1 - \tilde{\Phi}_{a+n_1p; b+z}(\theta_0)_{1_{z < \tilde{q}_\alpha}}$$

où $\Phi_{\theta_0}^G(z)$ est la fonction de distribution cumulative d'une distribution Gamma et $\tilde{\Phi}_{(a_2, b_2)}$ est la fonction de distribution cumulative d'une distribution Gamma (a_2, b_2) .

La distribution prédictive de y étant donnée x est une distribution bêta inverse donnée par :

$$\begin{aligned} f(y|x) &= \frac{(b+x)^{(\alpha+n_1p)} y^{n_2p-1}}{\Gamma(n_2p)\Gamma(a+n_1p)} \int_0^\infty \theta^{n_2p} \exp[-\theta y] \theta^{a+n_1p-1} \exp[-\theta(b+x)] d\theta \\ &= \frac{(b+x)^{(\alpha+n_1p)} y^{n_2p-1}}{\Gamma(n_2p)\Gamma(a+n_1p)} \int_0^\infty \theta^{a+n_1p+n_2p-1} \exp[-\theta(b+x+y)] \\ f(y|x) &= \frac{y^{(n_2p-1)} (b+x)^{(a+n_1p)}}{\beta(n_2p, a+n_1p)(b+x+y)^{(a+n_1p+n_2p)}} \sim InBeta(y, a+n_1p, b+x) \end{aligned} \quad (4.18)$$

Ensuite, nos indices de satisfaction prédictifs dans l'approche fréquentiste sont :

$$\eta_0(x) = E(\varphi_{0^{(\alpha)}}(z)|x) = \int_0^{q_{\alpha-x}} \frac{y^{(n_2p-1)} (b+x)^{(a+n_1p)}}{\beta(n_2p, a+n_1p)(b+x+y)^{(a+n_1p+n_2p)}} dy \quad (4.19)$$

$$\eta(x) = E(\varphi^{(\alpha)}(z)|x)$$

$$\eta(x) = \int_0^{q_{\alpha-x}} 1 - \Phi_{\theta_0}^G(z) \frac{y^{(n_2p-1)} (b+x)^{(a+n_1p)}}{\beta(n_2p, a+n_1p)(b+x+y)^{(a+n_1p+n_2p)}} dy \quad (4.20)$$

Pour l'approche Bayésienne, nous avons :

$$\tilde{\eta}_0(x) = E(\tilde{\varphi}_{0^{(\alpha)}}(z)|x) = \int_0^{\tilde{q}_{\alpha-x}} \frac{y^{(n_2p-1)} (b+x)^{(a+n_1p)}}{\beta(n_2p, a+n_1p)(b+x+y)^{(a+n_1p+n_2p)}} dy \quad (4.21)$$

$$\tilde{\eta}(x) = E(\tilde{\varphi}^{(\alpha)}(z)|x)$$

$$\tilde{\eta}(x) = \int_0^{\tilde{q}_{\alpha-x}} 1 - \tilde{\Phi}_{(a+n_1p; b+z)}(\theta_0) \frac{y^{(n_2p-1)}(b+x)^{(a+n_1p)}}{\beta(n_2p, a+n_1p)(b+x+y)^{(a+n_1p+n_2p)}} dy \quad (4.22)$$

4.7 Résultats de simulation

Nous montrons dans cette section les résultats de notre simulation, qui présentent une étude de comparaison de l'indice de satisfaction prévu associé au test fréquentiste et test Bayésien pour plusieurs valeurs d'hyperparamètre. Nous affichons également la taille effective de l'échantillon (ESS) associée à l'a priori [voir chapitre 5]

4.7.1 Résultats pour les données de Poisson

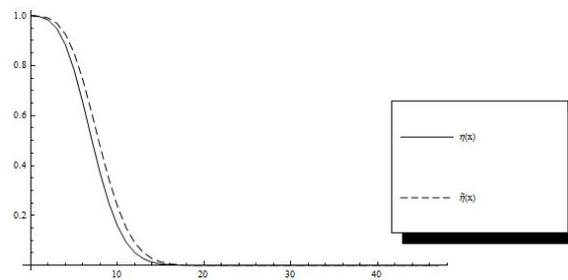


FIGURE 4.1 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \geq \theta_0 = 0.5$ avec $\alpha = 0.05, a = 1, b = 2, t_1 = 15, t_2 = 5, q_\alpha = 15, \tilde{q}_\alpha = 17, (ESS = 2)$.

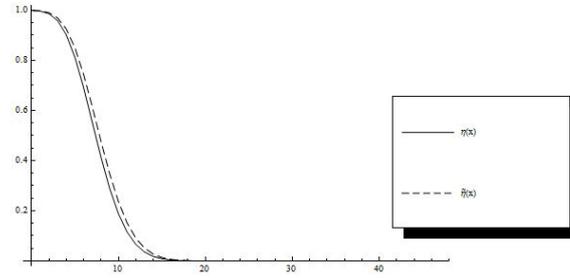


FIGURE 4.2 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \geq \theta_0 = 0.5$ avec $\alpha = 0.05, a = 1, b = 2, t_1 = 15, t_2 = 20, q_\alpha = 25, \tilde{q}_\alpha = 26, (ESS = 2)$.

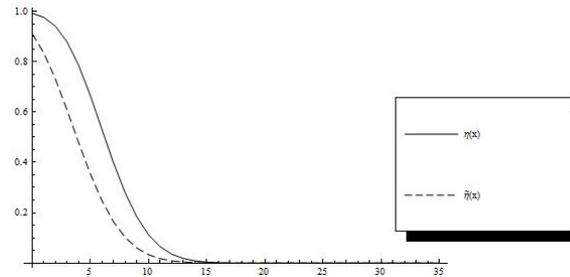


FIGURE 4.3 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \geq \theta_0 = 0.5$ avec $\alpha = 0.05, a = 6, b = 4, t_1 = 15, t_2 = 5, q_\alpha = 15, \tilde{q}_\alpha = 13, (ESS = 4)$.

Dans la figure 4.1 avec une valeur temps de personne t_2 petite et un a priori centré sur le seuil θ_0 , nous pouvons observer que les graphiques d'indice de satisfaction prédits correspondant au test fréquentiste et test Bayésien peuvent être différents. Dans la figure 4.2, nous observons que lorsque la valeur t_2 du temps de personne est grande, les indices prédits au test fréquentiste et au test Bayésien sont très proches. La figure 4.3 montre, comme prévu, que lorsque l'a priori favorise l'efficacité du traitement, l'indice de satisfaction prévu, basé sur le test Bayésien, est plus élevé que celui basé sur le test fréquentiste. La figure 4.4 et la figure 4.5 montrent l'effet de paramètre a priori sur la distribution prédictive.

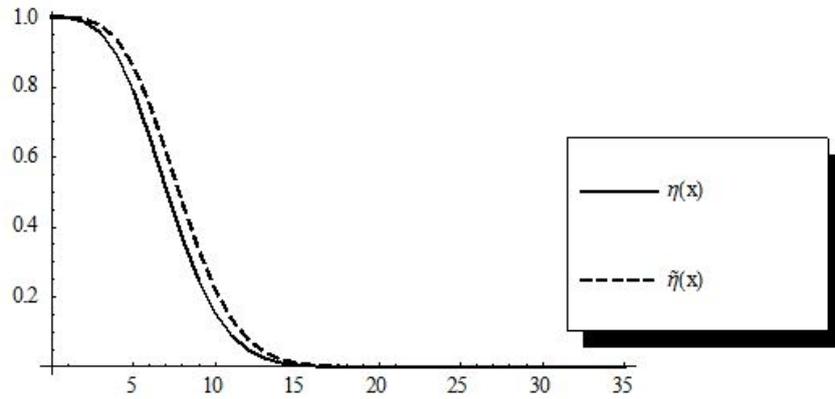


FIGURE 4.4 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \geq \theta_0 = 0.5$ avec $\alpha = 0.05, a = 0.2, b = 0.3, t_1 = 15, t_2 = 5, q_\alpha = 15, \tilde{q}_\alpha = 16, (ESS = 0.3)$.

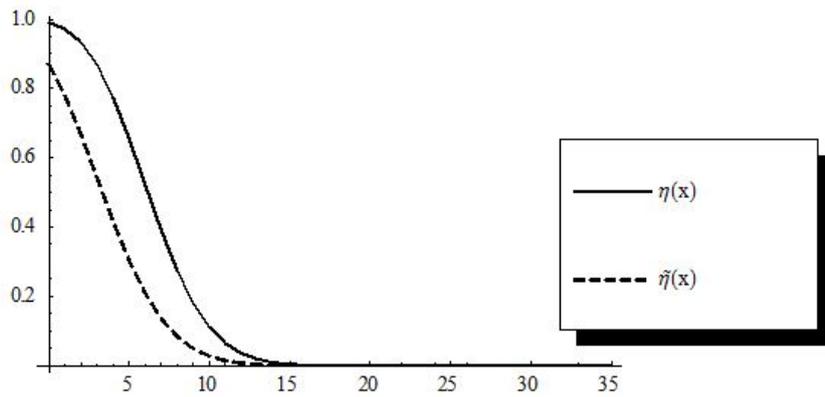


FIGURE 4.5 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \geq \theta_0 = 0.5$ avec $\alpha = 0.05, a = 7, b = 5, t_1 = 15, t_2 = 5, q_\alpha = 15, \tilde{q}_\alpha = 13, (ESS = 5)$.

4.7.2 Résultats pour les données de Gamma

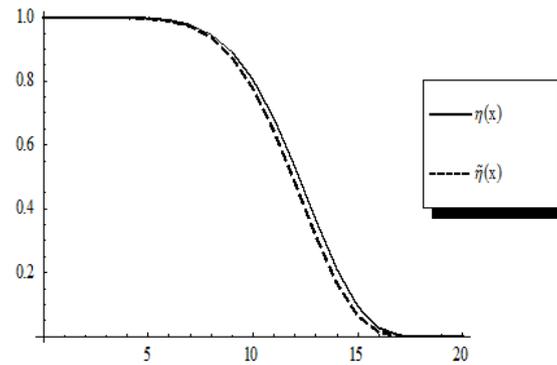


FIGURE 4.6 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \leq \theta_0 = 0.5$ avec $\alpha = 0.05, a = 1, b = 2, n_1 = 10, n_2 = 5, p = 1, q_\alpha = 18.4927, \tilde{q}_\alpha = 18.0719, (ESS = 1)$.

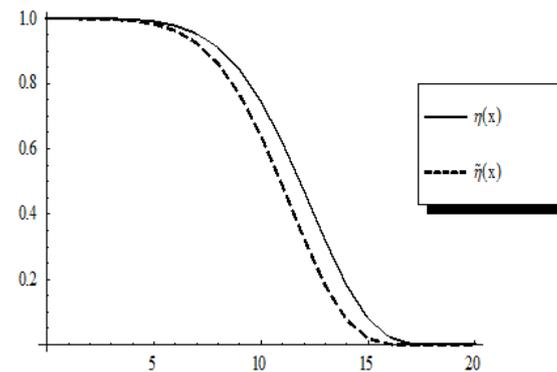


FIGURE 4.7 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \leq \theta_0 = 0.5$ avec $\alpha = 0.05, a = 3, b = 6, n_1 = 10, n_2 = 5, p = 1, q_\alpha = 18.4927, \tilde{q}_\alpha = 17.2686, (ESS = 3)$.

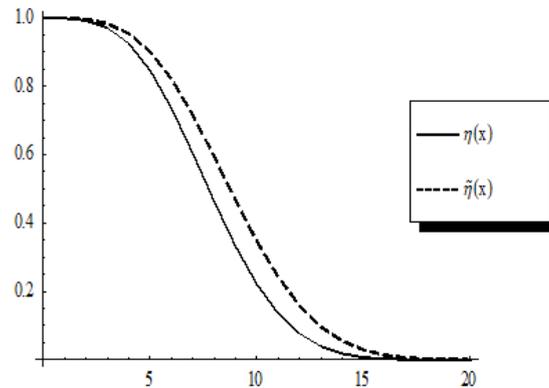


FIGURE 4.8 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \leq \theta_0 = 0.5$ avec $\alpha = 0.05$, $a = 3$, $b = 2$, $n_1 = 5$, $n_2 = 15$, $q_\alpha = 26.5093$, $\tilde{q}_\alpha = 29.4390$, ($ESS = 3$).

La figure 4.6 correspond à un a priori informatif faible avec $ESS = 1$ centré sur la valeur seuil θ_0 , tandis que la figure 4.7 correspond à un a priori plus informatif centré sur la même valeur. La figure 4.8 correspond à un a priori qui favorise l'efficacité du traitement.

4.7.3 Analyse des données réelles

Le modèle Poisson-Gamma a une application à la fois dans les essais cliniques et la fiabilité. Par exemple, tester les performances d'une nouvelle valve cardiaque ; en modélisant les données du temps d'échec « temps à mort » utilisées dans l'analyse de survie, et « temps d'interruption » utilisées dans la fiabilité. Nous testons la pertinence du modèle de poisson précédemment observé sur les données de dénombrement réel.

Nous considérons les données présentées dans [Hamada et al., 2008] qui consiste à modéliser le nombre mensuel de pannes des composants de superordinateur du Laboratoire National Los Alamos Blue Mountain (processeurs à mémoire partagée ou SMP) par une distribution de Poisson avec un nombre moyen inconnu de pannes θ_0 . Le superordinateur se compose de 47 SMP identiques et le tableau suivant présente leur nombre mensuel de pannes pour le premier mois de fonctionnement.

TABLE 4.1 – Nombre mensuel de pannes pour 47 composants de super-ordinateur

1	5	1	4	2	3	1	3	6	4	4	4
2	3	2	2	4	5	5	2	5	3	2	2
3	1	1	2	5	1	4	1	1	1	2	1
3	2	5	3	5	2	5	1	1	5	2	

Pour modéliser ces données, le nombre mensuel de défaillances est supposé suivre une distribution de Poisson. Pour représenter cela, l'information a priori pour le paramètre est supposée être une distribution a priori Gamma conjuguée avec une moyenne de 5, c'est-à-dire $\theta \sim \text{Gamma}(5, 1)$.

On choisit $\theta_0=1.2$ et d'après le calcul, on obtient $q_\alpha=69$, $\tilde{q}_\alpha=66$.

Les indices prévus dans la conception séquentielle pour le test fréquentiste et le test Bayésien sont donnés dans la figure et le tableau suivants :

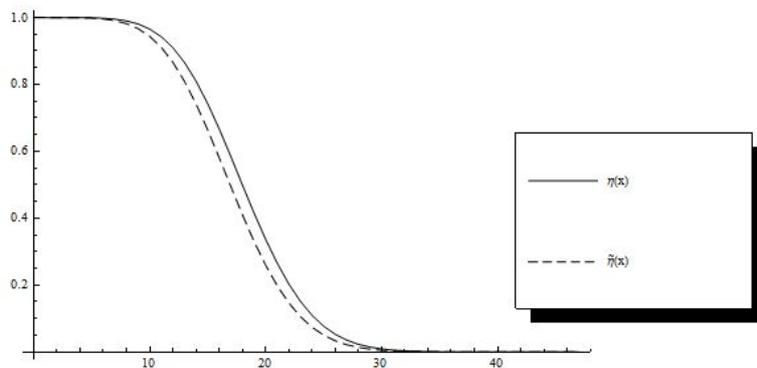


FIGURE 4.9 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien $H_0 : \theta \geq \theta_0 = 1.2$ avec $\alpha = 0.05$, $a = 5$, $b = 1$, $t_1 = 17$, $t_2 = 30$, $q_\alpha = 69$, $\tilde{q}_\alpha = 66$, ($ESS = 1$).

TABLE 4.2 – Indice de prédiction en $\eta(x)$ et $\tilde{\eta}(x)$ dans la conception séquentielle pour le test fréquentiste et le test Bayésien
 $H_0 : \theta \geq \theta_0 = 1.2$ avec $\alpha = 0.05, a = 5, b = 1, t_1 = 17,$
 $t_2 = 30, q_\alpha = 69, \tilde{q}_\alpha = 66, (ESS = 1)$

x	$\eta(x)$	$\tilde{\eta}(x)$
0	1	1
1	0.999992	0.999979
2	0.999969	0.999920
3	0.999894	0.999739
4	0.999675	0.999250
5	0.999114	0.998076
6	0.997817	0.995541
7	0.99510	0.990563
8	0.989891	0.981602
9	0.980688	0.966713
10	0.965612	0.943742
11	0.94260	0.910663
12	0.909716	0.866007
13	0.865554	0.865554
14	0.809618	0.741181
15	0.742586	0.663776
16	0.666373	0.580180
17	0.583949	0.494194
18	0.498964	0.409772
19	0.415244	0.330489
20	0.336296	0.259134
21	0.264905	0.197484
22	0.202896	0.146266
23	0.151084	0.105292
24	0.109380	0.073684
25	0.077001	0.050141
26	0.052723	0.033190
27	0.035121	0.021379
28	0.022771	0.013407
29	0.014374	0.008189
30	0.008838	0.004874
31	0.005296	0.002828
32	0.003094	0.00160
33	0.001763	0.000883
34	0.000980	0.000476
35	0.000532	0.000250

Cette analyse des données réelles conduit à la conclusion que les résultats obtenus à partir des indices prédits dans la conception séquentielle pour le test Bayésien et le test fréquentiste sont proches les uns des autres, mais l'approche purement Bayésienne est meilleure que l'approche fréquentiste. De plus, on observe, en regardant le tableau dans une approche Bayésienne complète, une bonne satisfaction par rapport à l'approche fréquentiste parce qu'elle conduit à fournir des inférences précises pour un paramètre d'intérêt et rend le praticien très satisfait.

4.8 Conclusion

L'approche prédictive Bayésienne est un outil utile dans la conception séquentielle de groupe pour évaluer la force de l'efficacité du traitement, qui est basée sur les observations disponibles et futures. Elle peut être utilisée pour l'évaluation Bayésienne ou fréquentiste de l'efficacité d'un traitement et permet d'arrêter précocement une expérience soit par manque, soit au contraire par prévision satisfaisante. Un des principaux avantages de la méthode prédictive Bayésienne est qu'elle prend automatiquement en compte le niveau d'information disponible et l'information prédite apportée par les observations futures. Un autre avantage est qu'elle peut être utilisée dans l'approche hybride Bayésienne-fréquentiste ou dans l'approche Bayésienne complète. Nous avons montré dans ce travail comment utiliser ces procédures dans les modèles Gamma-Poisson et Gamma-Gamma.

Chapitre 5

Détermination de la taille effective de l'échantillon d'un paramètre a priori

5.1 Introduction

Une question fondamentale dans toute analyse Bayésienne est la quantité d'informations contenues dans l'a priori. Pour de nombreux modèles couramment utilisés, la réponse semble simple.

Par exemple, on peut argumenter une loi Bêta(a, b) qui a une taille effective de l'échantillon (ESS) $a + b$. Ceci est basé sur le fait qu'une variable binomiale F d'un échantillon de taille n avec une probabilité de succès θ suivant une loi a priori Bêta(a, b), implique une loi a posteriori Bêta($a + x, b + n - x$). En d'autres termes, étant donné un échantillon de taille n , la somme $a + b$ devient la somme a posteriori $a + b + n$. Ainsi, dire qu'une loi Bêta (a, b) donnée a ESS $m = a + b$ nécessite le raisonnement implicite que la loi Bêta (a, b) peut être identifiée avec une loi a posteriori Bêta ($c + x, d + m - x$) résultante d'une loi a priori Bêta (c, d) ayant une très petite quantité d'informations. Une façon simple de formaliser ceci est de mettre $c - f - d = \epsilon$ pour une valeur arbitrairement petite $\epsilon > 0$ et de résoudre pour $m = a + b(c + d) = a + b - \epsilon$ [Morita et al., 2008].

Plus généralement, on peut correspondre a un a priori donné $p(\theta)$ avec le $q_m(\theta|x)$ a posteriori provenant d'un a priori précédent $q_0(\theta)$ qui

est choisi pour être vague dans un sens approprié et qui a été mis à jour par un échantillon de taille m en considérant m comme ESS de $p(\theta)$. Dans cette formulation générale, $p(\theta)$, $q_0(\theta)$ et $q_m(\theta|x)$ jouent des rôles analogues à ceux de $\text{Bêta}(a, b)$, $\text{Bêta}(c, d)$ et $\text{Bêta}(a + x, b + n - x)$ des distributions données ci-dessus. Dans certains cas, on peut trouver les hyper-paramètres de $q_m(\theta|x)$ en fonction de m , comparer $q_m(\theta|x)$ avec $p(\theta)$ et résoudre pour m analytiquement. Cependant, pour de nombreux modèles Bayésiens paramétriques, cette approche analytique ne fonctionne pas, et il n'est pas évident de déterminer l'ESS de l'a priori.

Dans ces cas, il n'y a pas de réponse évidente à la question de savoir ce que peut être ESS d'a priori. De plus, pour de nombreux choix couramment utilisés de $q_0(\theta)$, l'a priori joint est $p(\theta)$ dont il ne peut pas être associé à $q_m(\theta|x)$ analytiquement. Comprendre l'ESS a priori est important lors de l'application des méthodes Bayésiennes dans des contextes avec une taille d'échantillon faible à modérée. Par exemple, en ajustant un modèle Bayésien à un ensemble de données de 10 observations, une ESS a priori de 1 est raisonnable, alors qu'une ESS a priori de 20 implique que l'a priori, plutôt que les données, domine les inférences a posteriori. Si l'a priori est obtenu auprès d'un expert du domaine, un a priori informatif est souhaitable [Morita et al., 2008].

En revanche, si l'a priori n'est qu'un choix techniquement pratique, alors la compréhension d'ESS peut inciter l'investigateur à reconsidérer le choix a priori. Ainsi, il est important d'avoir une bonne idée de l'ESS de l'a priori pour interpréter ses inférences. Ceci est particulièrement important du point de vue de la défense des méthodes Bayésiennes contre la crainte que l'a priori puisse introduire de manière inappropriée des informations artificielles.

Dans ce chapitre, nous présentons une définition de l'ESS d'un a priori $p(\theta)$ dans un modèle paramétrique Bayésien, et nous fournissons une méthode pour calculer l'ESS dans une grande variété de contextes. Notre approche repose sur l'idée de construire un « ϵ -information » a priori $q_0(\theta)$, en considérant un échantillon X de taille m et l'a posteriori $q_m(\theta|X)$; et en calculant une distance entre $q_m(\theta|X)$ et $p(\theta)$ en termes de courbure

(dérivées secondes) de $\log\{p(\theta)\}$ et de $\log\{q_m(\theta|X)\}$. La valeur de m minimisant la distance est l'ESS a priori. Nous présentons une application motivante qui définit les a priori de l'information et l'ESS.

5.2 La taille effective de l'échantillon

La motivation intuitive de la construction suivante est la justification mathématique, donnée précédemment, de la raison pour laquelle l'ESS d'une loi Bêta(a, b) est égale à $a + b$. En tant que cadre Bayésien général, soit $f(x|\theta)$ qui désigne la fonction de distribution de probabilité d'un vecteur aléatoire s -dimensionnel x , et soit l'a priori $p(\theta|\tilde{\theta})$ sur le vecteur de paramètres $\theta = \theta_1, \theta_2, \dots, \theta_d$, où θ désigne le vecteur des hyperparamètres. La probabilité d'un échantillon indépendant et identiquement distribué (i.i.d.) $X_m = x_1, x_2, \dots, x_m$ est alors donnée par $f_m(X_m|\theta) = \sum_{i=1}^m f(x_i|\theta)$. Nous définissons une ϵ -information a priori $q_0(\theta|\tilde{\theta}_0)$ en exigeant d'avoir la même moyenne, $E_{q_0}(\theta) = E_p(\theta)$ et corrélations $Cor_{q_0}(\theta_j, \theta_{j'}) = Cor_p(\theta_j, \theta_{j'})$, $j \neq j'$.

Comme $p(\theta|\tilde{\theta})$, tout en augmentant les variances des éléments de θ pour que $Var_{(q_0)}(\theta_j) \gg Var_p(\theta_j)$, de sorte que $q_0(\theta|\tilde{\theta}_0)$ a une petite information mais $Var_{q_0}(\theta_j)$ doit exister pour $j = 1, \dots, d$. Le tableau 5.1 montre comment spécifier $q_0(\theta|\tilde{\theta}_0)$ pour plusieurs a priori paramétriques standards. Étant donné la vraisemblance $f_m(X_m|\theta)$ et l' ϵ -information a priori $q_0(\theta|\tilde{\theta}_0)$, nous désignons l'a posteriori par :

$$q_m(\theta|\tilde{\theta}_0, X_m) \propto q_0(\theta|\tilde{\theta}_0)f_m(X_m|\theta) \quad (5.1)$$

et la distribution marginale sous $p(\theta|\tilde{\theta})$ par :

$$f_m(X_m|\tilde{\theta}) = \int f_m(X_m|\theta)p(\theta|\tilde{\theta})d\theta \quad (5.2)$$

Lorsque $\tilde{\theta}_0$ est fixé, nous écrivons $f_m(X_m)$ par souci de brièveté. Pour définir la distance entre $p(\theta|\tilde{\theta})$ et $q_m(\theta|\tilde{\theta}_0, X_m)$, l'idée de base est de trouver la taille de l'échantillon, m , qui serait impliquée par les approxima-

tions normales de l'a priori $p(\theta)$ et l'a posteriori $q_m(\theta|\tilde{\theta}_0, X_m)$. Ceci nous a conduits à utiliser les dérivées secondes des densités logarithmiques pour définir la distance. Cependant, la vraie validation et la justification de notre définition proviennent de la comparaison des valeurs ESS résultantes avec l'ESS couramment rapportée dans les paramètres standards. Soit $\bar{\theta} = E_p(\theta)$ la moyenne a priori sous $p(\theta|\tilde{\theta})$. Nous définissons :

$$D_{p,j}(\theta) = -\frac{\partial^2(\log(p(\theta|\tilde{\theta})))}{\partial\theta_j^2} \quad (5.3)$$

et

$$D_{q,j}(m, \theta, X_m) = -\frac{\partial^2(\log(q_m(\theta|\tilde{\theta}_0, X_m)))}{\partial\theta_j^2}, \quad j = 1, 2, \dots, d \quad (5.4)$$

Notons :

$$D_{p,+}(\theta) = \sum_{j=1}^d D_{p,j}(\theta)$$

et

$$D_{q,+}(m, \theta) = \sum_{j=1}^d \int D_{q,j}(m, \theta, X_m) f_m(X_m) dX_m$$

Où $f_m(X_m)$ est la distribution marginale de X_m pour l'a priori $q_0(\theta)$.

Nous définissons une distance entre $p(\theta|\tilde{\theta})$ et $q_m(\theta|\tilde{\theta}_0, X_m)$ pour la taille de l'échantillon m comme une différence de la trace des deux matrices d'information :

$$\delta(m, \bar{\theta}, p, q_0) = \left| D_{p,+}(\bar{\theta}) - D_{q,+}(m, \bar{\theta}) \right| \quad (5.5)$$

C'est-à-dire que nous définissons la distance en termes de trace de la matrice d'information (dérivée seconde de la densité logarithmique) de l'a priori $p(\theta|\tilde{\theta})$ et la matrice d'information attendue de l'a posteriori $q_m(\theta|\tilde{\theta}_0, X_m)$ où l'attente (en anglais expectation) est par rapport au marginal $f_m(X_m)$. Quand $d = 1$, parce que l'indice + est superflu, on écrit $D_p(\theta)$ et $D_q(m, \bar{\theta})$.

Définition 5.1. L'ESS de $p(\theta|\tilde{\theta})$ par rapport à la vraisemblance $f_m(X_m|\theta)$ est l'entier m qui minimise la distance $\delta(m, \bar{\theta}, p, q_0)$.

La définition de la distance implique des choix arbitraires. Nous avons choisi cette définition après une étude empirique approfondie des formulations alternatives. Au lieu d'évaluer la courbure à la moyenne précédente, on peut utiliser le mode a priori.

De même, on peut marginaliser θ par rapport a l'a priori, moyenné sur X_m par rapport à $f_m(X_m|\theta)$ plutôt que la marginale $f_m(X_m)$, ou utiliser le déterminant plutôt que la trace de la matrice d'information. On peut aussi définir $\delta(\cdot)$ en termes de divergence de Kullback-Liebler ou de variance, la distance proposée $\delta(\cdot)$ correspondait le mieux aux résultats qui sont couramment utilisés comme valeurs ESS [Morita et al., 2008].

5.3 Déterminer la taille effective de l'échantillon pour quelques modèles

Nous déterminons ici la taille effective de l'échantillon d'une distribution a priori pour les modèles Bêta-Bernoulli, Gamma-Poisson et Gamma-Gamma comme proposé par [Morita et al., 2008].

L'idée est de faire correspondre un a priori $p(\theta)$ donné avec l'a posteriori $q_m\{\theta|x\}$ issu d'un a priori précédent $q_0(\theta)$ qui est choisi d'être vague dans un sens approprié et qui est mis à jour par un échantillon x de taille m . La valeur de m qui minimise la distance entre $p(\theta)$ et $q_m\{\theta|x\}$ sera l'échantillonnage effectif (ESS) associé à l'a priori $p(\theta)$.

La distance entre $q_m\{\theta|X_m\}$ et $p(\theta)$ est définie en termes de courbure (dérivées secondes) de $\log\{p(\theta)\}$ et $\log\{q_m\{\theta|X_m\}\}$. Étant donné la vraisemblance $f_m(X_m|\theta)$ et l'apriori $p(\theta|\tilde{\theta})$, nous désignons l'a posteriori par :

$$q_m(\theta|\tilde{\theta}, X_m) \propto q_0(\theta|\tilde{\theta})f_m(X_m|\theta). \quad (5.6)$$

Où $\bar{\theta} = E_p(\theta)$ désigne la moyenne a priori sous $p(\theta|\tilde{\theta})$.

On note :

$$D_p(\theta) = -\frac{\partial^2(\log(p(\theta|\tilde{\theta})))}{\partial\theta^2}$$

et

$$D_q(m, \theta) = -\frac{\partial^2(\log(q_m(\theta|\tilde{\theta}, X_m)))}{\partial\theta^2}$$

Nous définissons la distance entre $p(\theta|\tilde{\theta})$ et $q_m\{\theta|\tilde{\theta}, X_m\}$, pour la taille de l'échantillon m , par :

$$\delta(m, \bar{\theta}, p, q) = |D_p(\bar{\theta}) - D_q(m, \bar{\theta})| \quad (5.7)$$

L'ESS est obtenue en calculant les tailles d'échantillons implicites dans des modèles standards (Tableau 5.1) pour lesquels il existe des tailles d'échantillons a priori équivalentes communément rapportées.

5.3.1 Le modèle Bêta- Binomiale

Nous travaillons dans le cadre du modèle d'échantillonnage où la probabilité d'un échantillon indépendants et identiquement distribués (i.i.d.) X_m suit une loi Binomiale de paramètre (m, θ) . Si nous choisissons comme distribution a priori pour θ une loi Bêta est alors connue que la distribution a posteriori est toujours une loi Bêta.

L'a priori $p(\theta|\tilde{\theta}) \sim \text{Bêta}(\tilde{a}, \tilde{b})$ est :

$$p(\theta|\tilde{\theta}) \propto \theta^{(\tilde{a}-1)}(1-\theta)^{(\tilde{b}-1)}$$

$$\log p(\theta|\tilde{\theta}) = \log(\theta^{(\tilde{a}-1)}(1-\theta)^{(\tilde{b}-1)})$$

$$\log p(\theta|\tilde{\theta}) = (\tilde{a}-1)\log\theta + (\tilde{b}-1)\log(1-\theta)$$

$$\frac{\partial}{\partial\theta} = (\tilde{a}-1)\frac{1}{\theta} + (\tilde{b}-1)\frac{1}{1-\theta}$$

$$\frac{\partial}{\partial\theta^2} = -(\tilde{a}-1)\frac{1}{\theta^2} - (\tilde{b}-1)\frac{1}{(1-\theta)^2}$$

$$D_p(\theta) = -\frac{\partial}{\partial \theta^2} = (\tilde{a} - 1)\frac{1}{\theta^2} + (\tilde{b} - 1)\frac{1}{(1 - \theta)^2}$$

L'a posteriori $q_m(\theta|\tilde{\theta}_0, X_m) \sim \text{Bêta}(\frac{\tilde{a}}{c} + X_m, \frac{\tilde{b}}{c} + m - X_m)$,
où $X_m = \sum_{i=1}^m x_i$

$$q_m(\theta|\tilde{\theta}_0, X_m) \propto (\theta)^{\frac{\tilde{a}}{c} + X_m - 1} (1 - \theta)^{\frac{\tilde{b}}{c} + m - X_m - 1}$$

$$\log q_m(\theta|\tilde{\theta}_0, X_m) = \log \left((\theta)^{\frac{\tilde{a}}{c} + X_m - 1} (1 - \theta)^{\frac{\tilde{b}}{c} + m - X_m - 1} \right)$$

$$= \left(\frac{\tilde{a}}{c} + X_m - 1 \right) \log(\theta) + \left(\frac{\tilde{b}}{c} + m - X_m - 1 \right) \log(1 - \theta)$$

$$\frac{\partial}{\partial \theta} = \left(\frac{\tilde{a}}{c} + X_m - 1 \right) \frac{1}{\theta} + \left(\frac{\tilde{b}}{c} + m - X_m - 1 \right) \frac{1}{1 - \theta}$$

$$\frac{\partial}{\partial \theta^2} = - \left(\frac{\tilde{a}}{c} + X_m - 1 \right) \frac{1}{\theta^2} - \left(\frac{\tilde{b}}{c} + m - X_m - 1 \right) \frac{1}{(1 - \theta)^2}$$

$$D_q(m, \theta) = -\frac{\partial}{\partial \theta^2} = \left(\frac{\tilde{a}}{c} + X_m - 1 \right) \frac{1}{\theta^2} + \left(\frac{\tilde{b}}{c} + m - X_m - 1 \right) \frac{1}{(1 - \theta)^2}$$

On met $\bar{\theta} = E_p(\theta) = \frac{\tilde{a}}{\tilde{a} + \tilde{b}}$ donc la distance est :

$$\delta(m, \bar{\theta}, p, q) = |D_p(\bar{\theta}) - D_q(m, \bar{\theta})|$$

$$= \left| \frac{\tilde{a} - 1}{\bar{\theta}^2} + \frac{\tilde{b} - 1}{(1 - \bar{\theta})^2} - \left(\frac{\frac{\tilde{a}}{c} + \sum_{X=0}^m X f_m(X_m) - 1}{\bar{\theta}^2} + \frac{\frac{\tilde{b}}{c} + m - \sum_{X=0}^m X f_m(X_m) - 1}{(1 - \bar{\theta})^2} \right) \right|$$

$$= \left| \frac{\tilde{a} - 1}{\bar{\theta}^2} + \frac{\tilde{b} - 1}{(1 - \bar{\theta})^2} - \left(\frac{\frac{\tilde{a}}{c} + m \frac{\tilde{a}}{\tilde{a} + \tilde{b}} - 1}{\bar{\theta}^2} + \frac{\frac{\tilde{b}}{c} + m - m \frac{\tilde{a}}{\tilde{a} + \tilde{b}} - 1}{(1 - \bar{\theta})^2} \right) \right|$$

$$= \left| \frac{\tilde{a} - 1 - \frac{\tilde{a}}{c} - m \frac{\tilde{a}}{\tilde{a} + \tilde{b}} + 1}{\bar{\theta}^2} + \frac{\tilde{b} - 1 - \frac{\tilde{b}}{c} - m + m \frac{\tilde{a}}{\tilde{a} + \tilde{b}} + 1}{(1 - \bar{\theta})^2} \right|$$

$$\delta(m, \bar{\theta}, p, q) = \left| \frac{\tilde{a} - \frac{\tilde{a}}{c} - m \frac{\tilde{a}}{\tilde{a} + \tilde{b}}}{\bar{\theta}^2} + \frac{\tilde{b} - \frac{\tilde{b}}{c} - m + m \frac{\tilde{a}}{\tilde{a} + \tilde{b}}}{(1 - \bar{\theta})^2} \right|$$

$$ESS = m = \tilde{a} + \tilde{b}$$

5.3.2 Le modèle Poisson-Gamma

Nous travaillons dans le cadre du modèle d'échantillonnage où la probabilité d'un échantillon indépendants et identiquement distribués (i.i.d.) X_m suit une loi de Poisson de paramètre θ . Si nous choisissons comme distribution a priori pour θ une loi de Gamma est alors connue que la distribution a posteriori est toujours une loi de Gamma.

L'a priori $p(\theta|\tilde{\theta}) \sim \text{Gamma}(\tilde{a}, \tilde{b})$ est :

$$p(\theta|\tilde{\theta}) \propto \theta^{(\tilde{a}-1)} e^{(-\tilde{b}\theta)}$$

$$\log p(\theta|\tilde{\theta}) = \log(\theta^{(\tilde{a}-1)} e^{(-\tilde{b}\theta)})$$

$$\log p(\theta|\tilde{\theta}) = (\tilde{a} - 1) \log \theta - (\tilde{b})\theta$$

$$\frac{\partial}{\partial \theta} = (\tilde{a} - 1) \frac{1}{\theta} - (\tilde{b})$$

$$\frac{\partial}{\partial \theta^2} = -(\tilde{a} - 1) \frac{1}{\theta^2}$$

$$D_p(\theta) = -\frac{\partial}{\partial \theta^2} = (\tilde{a} - 1) \frac{1}{\theta^2}$$

L'a posteriori $q_m(\theta|\tilde{\theta}_0, X_m) \sim \text{Gamma}(\frac{\tilde{a}}{c} + X_m, \frac{\tilde{b}}{c} + m)$,
où $X_m = \sum_{i=1}^m x_i$

$$q_m(\theta|\tilde{\theta}_0, X_m) \propto (\theta)^{\frac{\tilde{a}}{c} + X_m - 1} e^{-\theta(\frac{\tilde{b}}{c} + m)}$$

$$\log q_m(\theta|\tilde{\theta}_0, X_m) = \log \left((\theta)^{\frac{\tilde{a}}{c} + X_m - 1} e^{-\theta(\frac{\tilde{b}}{c} + m)} \right)$$

$$= \left(\frac{\tilde{a}}{c} + X_m - 1 \right) \log(\theta) + \theta \left(\frac{\tilde{b}}{c} + m \right)$$

$$\frac{\partial}{\partial \theta} = \left(\frac{\tilde{a}}{c} + X_m - 1 \right) \frac{1}{\theta} + \left(\frac{\tilde{b}}{c} + m \right)$$

$$\frac{\partial}{\partial \theta^2} = - \left(\frac{\tilde{a}}{c} + X_m - 1 \right) \frac{1}{\theta^2}$$

$$D_q(m, \theta) = - \frac{\partial}{\partial \theta^2} = \left(\frac{\tilde{a}}{c} + X_m - 1 \right) \frac{1}{\theta^2}$$

On met $\bar{\theta} = E_p(\theta) = \frac{\tilde{a}}{\tilde{b}}$ donc la distance est

$$\delta(m, \bar{\theta}, p, q) = \left| D_p(\bar{\theta}) - D_q(m, \bar{\theta}) \right|$$

$$= \left| \frac{\tilde{a} - 1}{\bar{\theta}^2} - \frac{\frac{\tilde{a}}{c} + \sum_{X=0}^m X f_m(X_m) - 1}{\bar{\theta}^2} \right|$$

$$= \left| \frac{\tilde{a} - 1}{\bar{\theta}^2} - \frac{\frac{\tilde{a}}{c} + m \frac{\tilde{a}}{\tilde{b}} - 1}{\bar{\theta}^2} \right|$$

$$= \left| \frac{\tilde{a} - 1 - \frac{\tilde{a}}{c} - m \frac{\tilde{a}}{\tilde{b}} + 1}{\bar{\theta}^2} \right|$$

$$\delta(m, \bar{\theta}, p, q) = \left| \frac{\tilde{a} - \frac{\tilde{a}}{c} - m \frac{\tilde{a}}{\tilde{b}}}{\bar{\theta}^2} \right|$$

$$ESS = m = \tilde{b}$$

5.3.3 Le modèle Gamma-Gamma

Nous travaillons dans le cadre du modèle d'échantillonnage où la probabilité d'un échantillon indépendants et identiquement distribués (i.i.d.) X_m suit une loi de Gamma de paramètre (mp, θ) . Si nous choisissons comme distribution a priori pour θ une loi de Gamma est alors connue que la distribution a posteriori est toujours une loi de Gamma.

L'a priori $p(\theta|\tilde{\theta}) \sim \text{Gamma}(\tilde{a}, \tilde{b})$ est :

$$p(\theta|\tilde{\theta}) \propto \theta^{(\tilde{a}-1)} e^{(-\tilde{b}\theta)}$$

$$\log p(\theta|\tilde{\theta}) = \log(\theta^{(\tilde{a}-1)} e^{(-\tilde{b}\theta)})$$

$$\log p(\theta|\tilde{\theta}) = (\tilde{a} - 1) \log \theta - (\tilde{b})\theta$$

$$\frac{\partial}{\partial \theta} = (\tilde{a} - 1) \frac{1}{\theta} - (\tilde{b})$$

$$\frac{\partial}{\partial \theta^2} = -(\tilde{a} - 1) \frac{1}{\theta^2}$$

$$D_p(\theta) = -\frac{\partial}{\partial \theta^2} = (\tilde{a} - 1) \frac{1}{\theta^2}$$

L'a posteriori $q_m(\theta|\tilde{\theta}_0, X_m) \sim \text{Gamma}(\frac{\tilde{a}}{c} + mp, \frac{\tilde{b}}{c} + X_m)$,

où $X_m = \sum_{i=1}^m x_i$

$$q_m(\theta|\tilde{\theta}_0, X_m) \propto (\theta)^{\frac{\tilde{a}}{c} + mp - 1} e^{-\theta(\frac{\tilde{b}}{c} + X_m)}$$

$$\log q_m(\theta|\tilde{\theta}_0, X_m) = \log \left((\theta)^{\frac{\tilde{a}}{c} + mp - 1} e^{-\theta(\frac{\tilde{b}}{c} + X_m)} \right)$$

$$= \left(\frac{\tilde{a}}{c} + mp - 1 \right) \log(\theta) + \theta \left(\frac{\tilde{b}}{c} + X_m \right)$$

$$\frac{\partial}{\partial \theta} = \left(\frac{\tilde{a}}{c} + mp - 1 \right) \frac{1}{\theta} + \left(\frac{\tilde{b}}{c} + X_m \right)$$

$$\frac{\partial}{\partial \theta^2} = - \left(\frac{\tilde{a}}{c} + mp - 1 \right) \frac{1}{\theta^2}$$

$$D_q(m, \theta) = -\frac{\partial}{\partial \theta^2} = \left(\frac{\tilde{a}}{c} + mp - 1 \right) \frac{1}{\theta^2}$$

On met $\bar{\theta} = E_p(\theta) = \frac{\tilde{a}}{b}$ donc la distance est :

$$\delta(m, \bar{\theta}, p, q) = |D_p(\bar{\theta}) - D_q(m, \bar{\theta})|$$

$$= \left| \frac{\tilde{a} - 1}{\bar{\theta}^2} - \frac{\frac{\tilde{a}}{c} + mp - 1}{\bar{\theta}^2} \right|$$

$$= \left| \frac{\tilde{a} - 1 - \frac{\tilde{a}}{c} - mp + 1}{\bar{\theta}^2} \right|$$

$$\delta(m, \bar{\theta}, p, q) = \left| \frac{\tilde{a} - \frac{\tilde{a}}{c} - mp}{\bar{\theta}^2} \right|$$

$$ESS = m = \frac{\tilde{a}}{p}$$

TABLE 5.1 – L'a priori, la vraisemblance et q_m a posteriori correspondant par rapport à l'information a priori, et la taille d'échantillon effectif a priori traditionnellement rapportée, ESS pour certains modèles, où l'hyper-paramètre c est de très grandes constantes choisies pour faire augmenter les variances des éléments de θ sous le q_0 .

$p(\theta \tilde{\theta})$	$q_0(\theta \tilde{\theta}_0)$	$f(X_m \theta)$	$q_m(\theta \tilde{\theta}, X_m)$	ESS
$Be(\tilde{a}, \tilde{b})$	$Be\left(\frac{\tilde{a}}{c}, \frac{\tilde{b}}{c}\right)$	$Bin(m, \theta)$	$Be\left(\frac{\tilde{a}}{c} + x, \frac{\tilde{b}}{c} + m - x\right)$	$\tilde{a} + \tilde{b}$
$Ga(\tilde{a}, \tilde{b})$	$Ga\left(\frac{\tilde{a}}{c}, \frac{\tilde{b}}{c}\right)$	$Pois(\theta)$	$Ga\left(\frac{\tilde{a}}{c} + \sum_i x_i, \frac{\tilde{b}}{c} + m\right)$	\tilde{b}
$Ga(\tilde{a}, \tilde{b})$	$Ga\left(\frac{\tilde{a}}{c}, \frac{\tilde{b}}{c}\right)$	$Ga(mp, \theta)$	$Ga\left(\frac{\tilde{a}}{c} + mp, \frac{\tilde{b}}{c} + \sum_i x_i\right)$	$\frac{\tilde{a}}{p}$

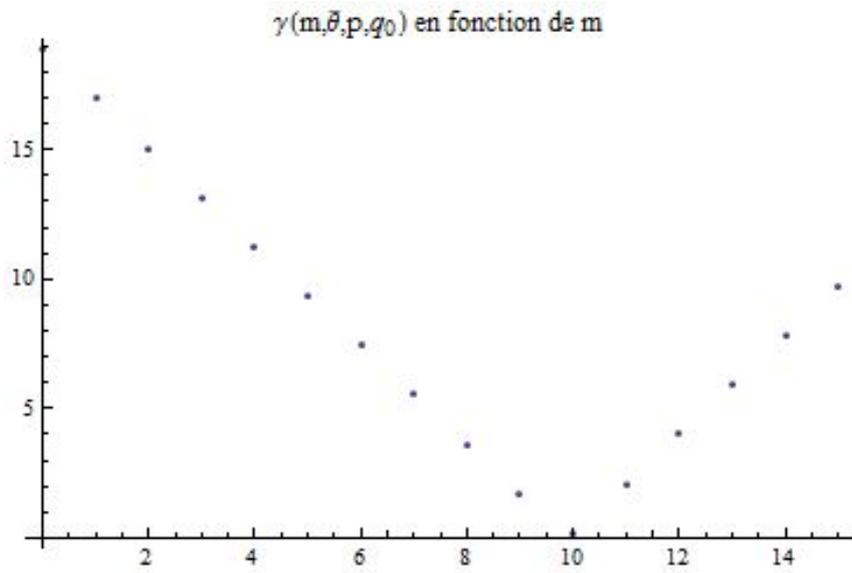


FIGURE 5.1 – La taille de l'échantillon effectif a priori pour le modèle Bêta-Binomiale avec $\tilde{\theta} = (\tilde{a}, \tilde{b}) = (3, 7)$, ESS=10.

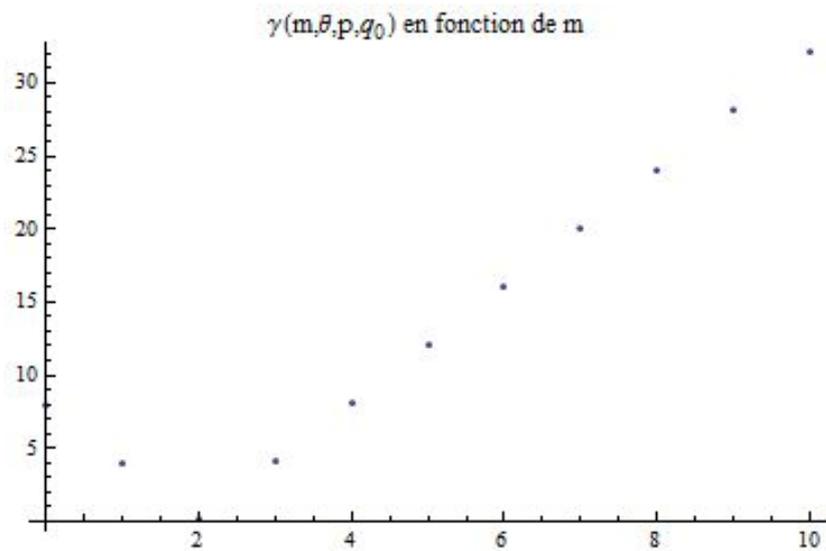


FIGURE 5.2 – La taille de l'échantillon effectif a priori pour le modèle Poisson-Gamma avec $\tilde{\theta} = (\tilde{a}, \tilde{b}) = (1, 2)$, ESS=2.

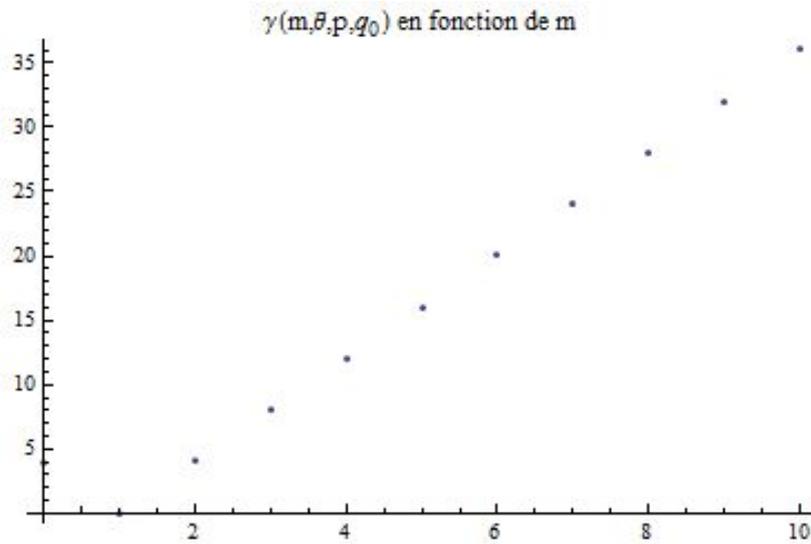


FIGURE 5.3 – La taille de l'échantillon effectif a priori pour le modèle Gamma-Gamma avec $\tilde{\theta} = (\tilde{a}, \tilde{b}) = (1, 2)$, $p=1$, $ESS=1$.

5.4 Conclusion

La méthode proposée est utile dans l'analyse Bayésienne, en particulier dans les contextes avec des a priori induits ; ou lorsque les données consistent en un nombre relativement restreint d'observations. En calculant des ESS, on peut éviter l'utilisation d'un a priori trop informatif dans le sens où l'inférence est dominée par l'a priori plutôt que par les données. Comme indiqué dans les applications, d'autres utilisations des valeurs ESS comprennent l'interprétation ou la révision des analyses Bayésiennes ou des conceptions des autres, l'utilisation des valeurs ESS elles-mêmes pour effectuer des analyses de sensibilité dans le caractère informatif de l'a priori et le calibrage des paramètres des conceptions Bayésiennes adaptatives aux résultats. Il s'agit d'un domaine potentiellement important pour les futures recherches, car il serait utile de calculer les valeurs ESS dans de tels contextes.

Conclusion générale

L'acceptation et l'applicabilité des méthodes Bayésiennes ont augmenté ces dernières années par rapport aux techniques classiques. Aujourd'hui, avec les progrès du calcul et de la méthodologie, nous utilisons des méthodes Bayésiennes pour résoudre une variété croissante de problèmes complexes.

Dans ce travail nous avons souligné l'importance des procédures prédictives Bayésiennes qui apportent une contribution précieuse aux méthodes inférentielles et au design séquentiel des essais expérimentaux. Dans cet aspect, les probabilités prédictives Bayésiennes peuvent être utilisées pour la planification d'essais expérimentaux afin que l'expérience aboutisse à un résultat statistiquement significatif, si l'essai devait se poursuivre jusqu'à le nombre de sujets prédéfini de l'échantillon.

De plus, notre recherche avait pour but de donner un aperçu général de la différence entre les calculs de puissance traditionnels et les calculs d'assurance pour déterminer la taille des échantillons, en donnant des exemples de planification que l'essai se termine avec le résultat souhaité.

La principale contribution de ce travail est de proposer et développer une conception séquentielle de groupe pour évaluer la force de l'efficacité du traitement ; qui est basée sur les observations disponibles et futures. La théorie présentée peut être facilement utilisée pour l'évaluation Bayésienne ou fréquentiste de l'efficacité d'un traitement et permet d'arrêter précocement une expérience soit par manque, soit au contraire par prévision satisfaisante. Nous avons montré dans ce travail comment utiliser cette procédure dans les modèles Gamma-Poisson et Gamma-Gamma.

D'autre part, nous avons proposé des méthodes pour calculer la taille effective de l'échantillon dans une variété de paramètres dans un modèle Bayésien.

En conclusion, ce travail nous a permis de mieux découvrir les avantages de l'approche prédictive Bayésienne que ce soit dans la construction

d'un nouvel essai dans une approche hybride Bayésienne-fréquentiste ou dans une approche Bayésienne complète.

Perspectives

Nous considérons l'approche prédictive Bayésienne dans la conception séquentielle de groupe pour évaluer la force de l'efficacité du traitement en l'appliquant dans une approche hybrides Bayésienne-fréquentiste ou dans une approche Bayésienne complète.

Nous avons montré dans ce travail comment utiliser cette procédure dans les modèles Gamma-Poisson et Gamma-Gamma pour trouver les frontières d'arrêt dans le cas des petits échantillons.

Une autre perspective possible à ce travail serait de comment passer de la puissance statistique à l'assurance statistique et aussi une autre perspective est de trouver la taille effective de l'échantillon d'une distribution a priori paramétrique dans un modèle Bayésien.

Annexes

Programme 01

Exemple3.1 : Puissance bayésienne: choix de la taille de l'échantillon pour un essai

```
theta=0.56 ; sigma=2; epsilon=0.025; beta=0.2;
```

```
zbeta= qnorm(1-beta) ; zepsilon= qnorm(epsilon);
```

```
n=( zbeta- zepsilon)^2 *{(sigma)^2/(theta)^2}; n
```

```
[1] 100.1133
```

```
round(n)
```

```
[1] 100
```

```
gamma=theta*(sqrt(n)/sigma)+zepsilon; gamma; pnorm(gamma)
```

```
T=200;
```

```
ndeath= matrix(0, ncol = T, nrow = length(1))
```

```
  for(i in 0:T){
```

```
ndeath[,i]=i+1
```

```
  }
```

```
ndeath = apply(ndeath, 2, prod); ndeath
```

```
gamma=theta*(sqrt(ndeath)/sigma)+zepsilon; gamma;
```

```
power<-pnorm(gamma)
```

```
power;
```

```
plot(ndeath,power)
```

Programme 02

Exemple 3.4 Incertitude: distribution prédictive de la puissance

```
theta=0.5 ; sigma=1; epsilon=0.025; beta=0.2;
```

```
zbeta= qnorm(1-beta) ; zepsilon= qnorm(epsilon);
```

```
n=(zbeta- zepsilon)^2 *(2*(sigma)^2/(theta)^2); n
```

```
[1] 62.79104
```

```
round(n)
```

```
[1] 63
```

```
T=100
```

```
npatients= matrix(0, ncol = T, nrow = length(1))
```

```
  for(i in 0:T){
```

```
npatients[,i]=i+1
```

```
  }
```

```
npatients = apply(npatients, 2, prod); npatients
```

```
gamma=sqrt((npatients*(theta)^2)/(2*(sigma)^2))+zepsilon; gamma;
```

```
power<-pnorm(gamma); power;
```

```
plot(npatients,power,type="l",xlab="n=patients",ylab="Puissance",lty=1,lwd=2)
```

Bibliographie

- [1] Anisimov, V. V. (2008). Using Mixed Poisson Models in Patient Recruitment in Multi centre Clinical Trials. *Proceedings of the World Congress on Engineering*, 2 :1046-1049.
- [2] Bakhshi, A. (2011). *Modelling and predicting patient recruitment in multi-centre clinical trials*. University of Glasgow : Master of Science, diss.
- [3] Bellaouar, A. et Beleulmi, S. (2013) *fiabilité, maintenabilité et disponibilité* . Faculté des Sciences de la Technologie Département Génie des Transports Université Constantine 1.
- [4] Berry, DA. (1985). Interim analyses in clinical trials : Classical vs. Bayesian approaches. . *Stat Med* , 4 :521-6.
- [5] Berry, DA. (2006). Bayesian clinical trials. *Nat Rev Drug Discov* , 5(1) : 27-36.
- [6] Berry, S.M., Carlin, B.P., Lee, J.J and Muller, P. (2011). *Bayesian Adaptive Methods for Clinical Trials*. New York : Biostatistics Chapman and Hall/CRC.
- [7] Blot, W.J. and Meeter, D.A. (1973). Sequential Experimental Design Procedures. *the American Statistical Association*, 68 :586-593.
- [8] Bowman AW, Crawford E. (2008). R package rpanel : simple control panels (version 1.0-5) *University of Glasgow, UK, 2008*. <http://www.stats.gla.ac.uk/~adrian/rpanel>.
- [9] Brannath, W., Zuber, E., Branson, M., Branson, M., Bretz, M., Gallo, P., Posch, M., Racine-Poon, A. (2009). Confirmatory Adap-

- tive Designs With Bayesian Decision Tools for a Targeted Therapy in Oncology. *Stat Med* , 28(10) :1445-1463.
- [10] Brard C. (2018). Approche Bayésienne de la survie dans les essais cliniques pour les cancers rares *Thèse de doctorat de l'Université Paris-Saclay*
- [11] Brown, B. W., Herson, J., Atkinson, E. N. et Rozell, M. E. (1987). Projection from previous studies - a Bayesian and frequentist compromise. *Controlled Clinical Trials*, 8 : 29-44.
- [12] Chen Ding Geng (Din) et Peace Karl E. (2003) Clinical Trial Data Analysis Using R *International statistical review* , 81 (1) : 151-173.
- [13] Chen Ding Geng (Din) et Peace Karl E. (2011) *Clinical trial methodology*. U.S : Chapman and Hall/CRC Biostatistics series.
- [14] Chen Ding-Geng (Din) et Ho Shuyen (2016). From statistical power to statistical assurance : It's time for a paradigm change in clinical trial design *Communications in Statistics - Simulation and Computation* , 46(10) :1532-4141.
- [15] Cherny NI, Dafni U, Bogaerts J et al. (2017). ESMO-Magnitude of Clinical Benefit Scale version 1.1. *Ann Oncol* , 28 :2340-2366.
- [16] Chow Shein-Chung, Chang Mark (2007). *Adaptive Design Methods in Clinical Trials*. New York : Biostatistics Chapman and Hall/CRC.
- [17] Chow Shein-Chung , Shao Jun and Wang Hansheng (2003). *Sample size calculations in clinical trial research*. New York : Biostatistics : A Series of References and Textbooks.
- [18] Chuang-Stein C. (2006). Sample size and the probability of a successful trial *Pharmaceutical Statistics* , 5 :305-309.
- [19] Chuang-Stein C, Yang R. (2010). A revisit of sample size decisions in confirmatory trials. *Statistics in Biopharmaceutical Research* , 2 :239-248.
- [20] Core Team. R, (2011). R Development Core Team. R : a language and environment for statistical computing, R Foundation for Statis-

tical Computing *Vienna, Austria, 2011*. <http://www.R-project.org/>, ISBN 3-900051-07-0.

- [21] Cucherat, M., Lièvre, M., Leizorovicz, A., Boissel,JP. (2004). *Lecture Critique et Interprétation des Résultats des Essais Cliniques pour La Pratique Médicale* . Medecine sciences publications
- [22] Daimon T, Zohar S, O’Quigley J. (2011). Posterior Maximization and Averaging for Bayesian Working Model Choice in The Continual Reassessment Method. *Stat Med* , 30 (13) : 1563-73.
- [23] Djeridi, Z. (2019) *Choix de l’a priori en statistique Bayésienne dans les essais expérimentaux*. . Thèse de Doctorat Ès-Sciences Université frères Mentouri Constantine 1.
- [24] Djeridi, Z., Merabet, H. (2019). A hybrid Bayesian-frequentist predictive design for monitoring multi-stage clinical trials. *Sequential Analysis*, 38(3) :301-317.
- [25] Djeridi, Z., Merabet, H. (2016). Choice of the prior in the bayesian design for the clinical trials. *Sciences et technologies A*, 44 :27-31.
- [26] Fayers, P. M., Cuschieri, A., Fielding, J., Craven, J., Uscinska, B. and Freedman, L. S. (2000). Sample size calculation for clinical trials : the impact of clinician beliefs. *British Journal of Cancer*, 82 :213-9.
- [27] FOOD AND DRUG ADMINISTRATION. Guidance for industry and FDA staff : Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials. (2010) URL : <http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071121.pdf>.
- [28] Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubinl, D.B. (2014). *Bayesian Data Analysis*. 3rd ed. U.S : Taylor and Francis Group.
- [29] Gilles Rousseau. (2016) *Vieillessement du TRIAC soumis à des essais de fiabilité du type HTRB* . Conservatoire national des arts et métiers centre regional associé de Orléans.

- [30] Gore SM.(1987). Biostatistics and the medical research council. *Medical Research Council News* , 35 : 19-20.
- [31] Gross AJ, Clark VA. (1975). *Survival Distributions : Reliability Applications in the Biomedical Sciences*. New York : Wesley .
- [32] Gsponer T, Gerber F, Bornkamp B et al. (2014) A practical guide to Bayesian group sequential designs. *Pharm Stat*, 13 :71-80
- [33] Hamada, M.S., Wilson, A.G., Reese, C.S and Martz, H.F. (2008). *Bayesian Reliability*. New York : Springer Series in Statistics.
- [34] Hampson LV, Whitehead J, Eleftheriou D et al. (2015) Elicitation of Expert Prior Opinion : Application to the MYPAN Trial in Childhood Polyarteritis Nodosa *Plos One* 2015 ;10 :e0120981.
- [35] Hand, A.L., (2011). Bayesian Sample-size Determination and Adaptive Design for Clinical Trials with Poisson Outcomes. *Faculty of Baylor University in Partial Fulfillment, Department of Statistical Science*.
- [36] Hand, A.L., Scott, J.A., Young, P.D., Stamey, J.D. and Young, D.M. (2016). Bayesian adaptive two-stage design for determining person-time in Phase II clinical trials with Poisson data. *Journal of Applied Statistics*, 43 :1625-1635.
- [37] Herbert, R. (1952). Aspects of The Sequential Design of Experiments. *Bull.Amer.Math.Soc* , 58(5) : 527-535.
- [38] Hiance A, Chevret S, Lévy V. (2009). A practical approach for eliciting expert prior beliefs about cancer survival in phase III randomized trial. *Journal of Clinical Epidemiology*, 62 (4) : 431-437.
- [39] Huang, X., Biswas, S., Oki, Y., Issa, JP., Berry, DA. (2007). A Parallel Phase I/II Clinical Trial Design for Combination Therapies *Biometrics* , 63(2) : 429-36.
- [40] Hughes MD. (1993). Reporting Bayesian analyses of clinical trials. *Stat Med* , 12 :1651-1663.
- [41] Ibrahim JG, Chen MH. (2000). Power prior distributions for regression models. *Statist Sci* , 15 : 46-60.

- [42] Ibrahim JG, Chen MH et Sinha D. (2001) *Bayesian Survival Analysis*. . New-York : Springer.
- [43] Ibrahim JG, Chen MH et Chu H. (2012). Bayesian methods in clinical trials : a Bayesian analysis of ECOG trials E1684 and E1690. *BMC Med Res Methodol* , 12 :183.
- [44] Jennison, C., Turnbull, B.(1999). *Group sequential methods with applications to clinical trials*. U.S.A : Chapman and Hall/CRC Interdisciplinary Statistics.
- [45] Johnson, N.L., Kemp, A.W., Kotz, S. (2005). *Univariate Discrete Distributions*. New Jersey : John Wiley and Sons.
- [46] Johnson, VE. et Cook, JD. (2009). Bayesian Design of Single-arm Phase II Clinical Trials with Continuous Monitoring *Clin Trials*, 6 : 217-26.
- [47] Johnson SR, Tomlinson GA, Hawker GA, Granton JT, Grosbein HA, Feldman BM. (2010). Methods to elicit beliefs for Bayesian priors : a systematic review. *Journal of Clinical Epidemiology*, 63 (4) : 355-369.
- [48] Kachiashvili, K.J. (2014). The Methods of Sequential Analysis of Bayesian Type for the Multiple Testing Problem. *Sequential Analysis : Design Methods and Applications*, 33 :23-38.
- [49] Kadane JB, Wolfson LJ. (1998). Experiences in elicitation. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 47 (1) : 3-19.
- [50] Koop, G., Poirier, D.J. and Tobias, J.L. (2007). *Bayesian Econometric Methods*. New York : Cambridge University Press.
- [51] Kramar, A. et Mathoulin-Pélissier, S. (2011) *Méthodes Biostatistiques Appliquées à la Recherche Clinique en Cancérologie* . John Libbey Eurotext.
- [52] Labdaoui, A. (2015) *Analyse Bayésienne dans les essais cliniques* . Thèse de Doctorat LMD Université frères Mentouri Constantine 1.

- [53] Lai, T.L. (2001). Sequential analysis : some classical problems and new challenges. *Statistica Sinica*, 11 :303-351.
- [54] Lakshmi, R.V., Vaidyanathan, V.S. (2015). Parameter Estimation in Gamma Mixture Model using Normal-based Approximation. *Statistical Theory and Applications*, 15 :25-35.
- [55] Lasonos, A., Wilton, AS., Riedel, ER., Seshan, VE., Spriggs, DR. (2008). A Comprehensive Comparison of The Continual Reassessment Method to The Standard 3+3 Dose Escalation Scheme in Phase I Dose-Finding Studies. *Clin Trials* , 5 (5) : 465-77.
- [56] Lecoutre, B., Derzko, G., Grouin, J.M. (1995). Bayesian predictive approach for inference about proportion. *Statistics in Medicine*, 14 :1057-1063.
- [57] Lee JJ. et Chu CT. (2012) Bayesian clinical trials in action *Stat Med*, 31 (25) : 2955-72.
- [58] Lellouch, J. (1969). L'application de L'analyse Séquentielle aux Essais Thérapeutiques. *Revue de statistique appliquée*, 17(1) :41-68.
- [59] Lunn D, Jackson C, Best N et al. (2012). The BUGS Book : A Practical Introduction to Bayesian Analysis. Chapman and Hall/CRC Press
- [60] Merabet, H. (2004). Index and prevision of satisfaction in exponential models for clinical trials. *Statistica*, LXIV :441-453.
- [61] Merabet, H. (2013). Bayesian sequential analysis of clinical trials feedback. *Statistica*, LXXIII :363-377.
- [62] Merabet, H., Labdaoui. A. and Druilhet, P. (2017). Bayesian prediction for sequential analysis in clinical trials designs. *Communications in Statistics-Theory and methods*, 46 :9807-9816.
- [63] Millier. C. (1967). Une Méthode Statistique : L'analyse Progressive. *Annales des sciences forestières, INRA / EDP Sciences*, 24(4) :327-343.
- [64] Monfort, (2020) Analyse Séquentielle *Dicostat2005, 2005-2020*.

- [65] Morita, S., Thall, P. F. and Muller, P. (2008). Determining the effective sample size of parametric prior. *Biometrics*, 64 :595-602.
- [66] Moyé LA. (2008). Bayesians in clinical trials : Asleep at the switch. *Stat Med*, 27 :469-82.
- [67] Nadarajah, S. (1973). Reliability for Lifetime Distributions. *Mathematical and Computer Modelling*, 37 : 683-688.
- [68] Nitis, M. and Basil, M. S. (2009). *Sequential methods and their applications*. U.S : Chapman and Hall/CRC, Taylor and Francis Group.
- [69] Oakley JE, O'Hagan A. (2010). Shelf : The SHEffield ELicitation Framework (version 2.0) School of Mathematics and Statistics, University of Sheffield, 2010. [http ://tonyohagan.co.uk/shelf](http://tonyohagan.co.uk/shelf).
- [70] O'Hagan A, Stevens JW. (2001). Bayesian assessment of sample size for clinical trials of cost-effectiveness. Medical Decision. *Medical Decision Making* , 21 (3) : 219-230.
- [71] O'Hagan A, Stevens JW, Campbell MJ. (2005). Assurance in clinical trial design. *Pharmaceutical Statistics* , 4 (3) : 187-201.
- [72] O'Hagan A, Buck CE, Daneshkhah A, Eiser JE, Garthwaite PH, Jenkinson DJ, Oakley JE, Rakow T. (2006). Uncertain Judgements : Eliciting Expert Probabilities. John Wiley and Sons Ltd : England
- [73] O'Quigley, J., Pepe, M., Fisher, L. (1990). Continual Reassessment Method : A Practical Design for Phase 1 Clinical Trials in Cancer. *Biometrics* , 46 (1) : 33-48.
- [74] Pauwels, E., Lajaunie, C. and Vert, J.P. (2014). A Bayesian active learning strategy for sequential experimental design in systems biology. *BMC Systems Biology*, 8 :1-11.
- [75] Parmar MKB, Spiegelhalter DJ, Freedman LS.(1994). The CHART trials : Bayesian design and monitoring in practice. *Statistics in Medicine* , 13 (13-14) : 1297-1312.
- [76] Pocock, SJ.(1976) The combination of randomized and historical controls in clinical trials. *J Chronic Dis*, 29 :175-88.

- [77] Ren Shijie et Oakley Jeremy E. (2013). Assurance calculations for planning clinical trials with time-to-event outcomes *Statistics in Medicine* , 33 : 31-45
- [78] Robert, C. P. (2007). *Bayesian Choice : From Decision-Theoretic Foundations to Computational Implementation*. New York : Springer-Science+Business Media..
- [79] Sackrowitz, H., Cahn, E.S. (1999). P Values as Random Variables-Expected P Values. *The American Statistician*, 53 :326-331.
- [80] 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*, 11 :485-493.
- [81] Schmidli H, Bretz F, Racine Poon A. (2007). Bayesian Predictive Power for Interim Adaptation in Seamless Phase II/III Trials Where the Endpoint is Survival up to Some Specified Time Point. *Stat Med* , 26 (27) : 4925-38.
- [82] Schoenfeld DA, Richter JR. (1982). Nomograms for calculating the number of patients needed for a clinical trial with survival as an endpoint *Biometrics* , 38 (1) : 163-170.
- [83] Shih, M.C. and Lavori, P.W. (2013). The utility of Sequential methods for comparative effectiveness experiments : point of care clinical trials. *Statistica Sinica*, 23 :1775-1791.
- [84] Spiegelhalter DJ, Freedman LS. (1986). A predictive approach to selecting the size of a clinical trial based on subjective clinical opinion *Statistics in Medicine* , 5 : 1-13
- [85] Spiegelhalter, D.J., Abrams, K.R. and Myles, J.P. (2004). *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. England : John Wiley and Sons.
- [86] Stallard, N. (1998). Sample size determination for Phase II clinical trials based on Bayesian decision theory. *Biometrics*, 54 :279-294.
- [87] Tan SB, Chung YFA, Tai BC, Cheung YB, Machin D.(2003). Elicitation of prior distributions for a phase III randomized controlled

- trial of adjuvant therapy with surgery for hepatocellular carcinoma. *Controlled Clinical Trials*, 24 (2) : 110-121.
- [88] Tanwir Akhtar and Athar Ali Khan. (2017). Bayesian analysis of poisson reliability model with r and jags. *International Journal of Recent Scientific Research* , 8(11) : 21837-21841.
- [89] Thall PF, Wathen JK. (2005). Covariate-Adjusted Adaptive Randomization in a Sarcoma Trial With Multi-Stage Treatments. *Stat Med* , 24 (13) : 1947-64.
- [90] Thall PF, Wathen JK. (2007). Practical Bayesian Adaptive Randomisation in Clinical Trials. *Eur J Cance*, 43 (5) : 859-66.
- [91] Thall PF, Ursino M, Baudouin V et al. (2017). Bayesian Treatment Comparison Using Parametric Mixture Priors Computed from Elicited Histograms. *Stat Methods Med (in press)*.,
- [92] Tierney L. (2011). tkrplot : Tk rplot <http://CRAN.R-project.org/package=tkrplot>, R package version 0.0-23.
- [93] Todd, S., Whitehead, A., Stallard, N. and Whitehead, J. (2011). Interim analyses and sequential designs in phase III studies. *Br J Clin Pharmacol*, 51 :394-399.
- [94] Tranchand, B., Chabaud, S., Colomban, O., Dussart, S., Pérol, D. (2008). Méthodologie des Essais Cliniques en Cancérologie. *La Lettre du Pharmacologue* , 22(3) : 98-110.
- [95] Traonmilin, Y. et Richou, A. (2018). Introduction aux Statistiques Bayésiennes. *Basées sur les notes de cours de Charles Dossal et de Jérémie Bigot*.
- [96] Vanlier, J., Tiemann, C.A., Hilbers, P.A.J. and van Riel, N.A.W. (2012). A Bayesian approach to targeted experiment design. *bioinformatics*, 28 :1136-1142.
- [97] Zaslavsky, B.G. (2010). The utility of Empirical Bayes models of Poisson clinical trials and sample size determination. *Pharm. Stat* , 9 :133-141.

- [98] Zhu, Q.Y., (1991) Modèles Bayésiens et application à l'estimation des caractéristiques de produits finis et aux contrôle de la qualité, Thèse de DOCTORAT de l'École Nationale des Ponts et Chaussées.

LES PLANS BAYÉSIENS DANS LES ESSAIS EXPÉRIMENTAUX

Thèse en vue de l'obtention du diplôme de DOCTORAT 3^{ème} cycle

Option : Statistique appliquée

Présentée par : **BOUREZAZ Houda**

Résumé :

Les travaux présentés dans cette thèse portent sur les problèmes méthodologiques qui se posent dans le contexte des essais expérimentaux, nous avons exposé les méthodes de l'analyse séquentielle mise en place pour l'utiliser dans l'évaluation des plans d'expérience dont elle permet d'arriver à des plans d'expérimentation efficaces.

Nous considérons l'utilisation de la méthode d'assurance dans la planification des essais cliniques : Dans la méthode de l'assurance, qui est une alternative au calcul de la puissance, nous calculons la probabilité qu'un essai clinique aboutisse à un résultat positif, en obtenant une distribution de probabilité a priori sur l'effet du traitement pertinent.

Dans ce travail, nous nous intéressons à la modélisation de la prédiction Bayésienne pour évaluer un indice de satisfaction après une première étape d'expérimentation afin de décider d'arrêter à la deuxième étape ou de continuer. Nous appliquons cette méthode aux résultats distribués de Poisson et Gamma dans de nombreux domaines tels que la fiabilité ou l'analyse de survie pour une interruption précoce en raison de la futilité ou de l'efficacité. Nous examinons deux types de prise de décision : une approche hybride Bayésienne-fréquentiste ou une approche Bayésienne complète.

Ce travail a permis aussi de déterminer la taille de l'échantillon efficace d'un paramètre a priori dans un modèle Bayésien.

Mots clés : Distribution prédictive Bayésienne, indices de satisfaction, assurance Bayésienne, approche hybride Bayésienne-fréquentiste, approche Bayésienne, analyse séquentielle en deux étapes, essais expérimentaux, résultats Poisson et Gamma, taille de l'échantillon efficace.

Membres du jury :

Président	Mr BESSILA Khaled	MCA	U. F. Mentouri Cne
Rapporteur	Mme MERABET Hayet	Prof	U. F. Mentouri Cne
Examineur	Mr HAMRI Nasr-Eddine	Prof	C. U. de Mila
Examineur	Mr BERKANE Abdelhak	MCA	U. F. Mentouri Cne
Examineur	Mr ABDELOUAHAB Med Salah	MCA	C. U. de Mila

Soutenue le 03/12/2020