



THÈSE

Présentée en vue de l'obtention du diplôme de

Doctorat

ESTIMATION BAYESIENNE EN PRESENCE DE CENSURES PROGRESSIVES

Option : Probabilités et statistique

Par:

AIACHI Hiba

Sous la direction de

Directeur de these: CHADLI Assia

Prof. U.B.M. Annaba

Co-directeur de these : TALHI Hamida

MCA U.B.M. Annaba

Devant le jury

PRESIDENT : SEDDIK-AMEUR Nacira

Prof. U.B.M. Annaba

EXAMINATEUR : SADKI Ourida

Prof. USTHB. Alger

EXAMINATEUR : MERZOUGUI Mouna

Prof. U.B.M. Annaba

EXAMINATEUR : LASKRI Yamina

Prof. E.S.T.I. Annaba

Année : 2021/2022

Table des matières

Dédicace	5
Remerciement	6
Résumé	7
Abstract	9
Introduction	11
1 Estimation Bayésienne	14
1.1 Rappels sur la statistique classique	14
1.1.1 Notions de Base	14
1.1.2 Échantillonnage	15
1.1.3 Estimation	15
1.2 Analyse de survie	17
1.2.1 Les fonctions de base	18
1.2.2 Types de censures	19
1.2.3 Censure progressive de type II	21
1.3 Estimation Bayésienne	25
1.3.1 Introduction à la théorie de la décision Bayésienne	25
1.3.2 Spécification de la distribution a priori	27
1.3.3 La loi a priori informative	28
1.3.4 la loi a priori non informative	28

1.3.5	La loi a postérieure	30
1.3.6	Estimateur Bayésien	31
1.3.7	Prédiction Bayésienne pour les statistique d'ordre	35
2	Les méthodes de Monte Carlo par chaîne de Markov	36
2.1	Introduction	36
2.2	Chaînes de Markov	37
2.2.1	Classification d'état	40
2.2.2	Théorème ergodique	41
2.3	Méthodes de Monte Carlo	42
2.3.1	Application sur les intégrales	43
2.3.2	Génération de nombres aléatoires	44
2.4	Méthodes de Monte Carlo par chaînes de Markov	46
2.4.1	Une méthode MCMC : Algorithme de Metropolis Hastings	47
2.4.2	Une méthode MCMC : Algorithme de Gibbs	56
3	L'estimation Bayésienne en présence de censure : Modèle de Bertholon et Modèle Tronqué	58
3.1	L'estimation de la distribution de Bertholon avec des données censurées à droite	58
3.1.1	La distribution Bertholon \mathcal{B}	59
3.1.2	Estimation des paramètres par le maximum de vraisemblance	60
3.1.3	Estimation Bayésienne des paramètres sous différentes fonctions de perte	62
3.1.4	Simulation	65
3.2	L'estimation de la distribution de Zeghdoudi tronquée à droite	74
3.2.1	La distribution de Zghdoudi tronquée à droite	75
3.2.2	Estimation des paramètres par le maximum de vraisemblance	76
3.2.3	Estimation Bayésienne sous différentes fonctions de perte	77
3.2.4	Simulation	79
3.2.5	Application sur des données réelles	85
4	La distribution de Burr XII à trois paramètres	87
4.1	La distribution Burr XII (θ, β)	88

4.1.1	Propriétés statistiques et mathématiques	90
4.1.2	Estimation des paramètres	93
4.2	Distribution de Burr XII à trois paramètres	93
4.2.1	Propriétés statistiques et mathématiques	94
4.3	Estimation des paramètres par le maximum de vraisemblance en présence de données censurées progressives de type II	97
4.4	Estimation Bayésienne sous différentes fonctions de perte en présence de données censurées progressives de type II	98
4.4.1	Prédiction Bayésienne	102
4.5	Simulation	103
4.5.1	Résultats de l'estimation	104
4.5.2	Comparaison des estimations	109
4.6	Application sur des données réelles	112
	Conclusion et Perspectives	116
	Annexe	118

Dédicace

Je rends grâce à Dieu qui m'a donné la volonté, la patience et le courage d'accomplir ce modeste travail. Je dédie ce travail de recherche à la mémoire de ma mère que Dieu ait pitié de son âme.

À mon père, merci pour les efforts déployés jour et nuit pour mon éducation et mon bien-être. Pour mon frère "Yahya", les mots n'exprimeront pas l'amour et l'affection que je lui porte. Je dédie ce travail avec tous mes vœux de bonheur, de santé et de réussite, un avenir plein de joie et de sérénité à tous les membres de la famille "Aiachi" petits et grands. J'adresse mes sincères remerciements et mon amour à mes chers amis en particulier, "Aliouèche Meryem", "Talhi Kawther" et "Hedadji Chima", Ce sont des sœurs et des amies sur lesquelles je peux toujours compter, et n'oublions pas de remercier mes collègues, je ne trouve pas les mots justes et sincères pour leur exprimer mon affection et mes pensées.

Remerciement

Je suis très reconnaissante à ma directrice de thèse, Madame Chadli Assia et ma co-directrice Madame Hamida Talhi de l'université Badji-Mokhtar de Annaba et membres du laboratoire LaPS, pour m'avoir fait confiance en me proposant ce travail de thèse. Je les remercie sincèrement pour la disponibilité, la gentillesse et leurs conseils avisés.

Je tiens à exprimer toute ma gratitude à Madame Seddik-Ameur Professeur à l'université Badji-Mokhtar d'Annaba, d'avoir acceptée d'évaluer mon travail et de présider le jury de ma thèse. Je suis très honorée par la présence au sein du jury de ma thèse, en tant qu'examineurs, des professeurs :

- Sadki Ourida. Pr. USTHB. Alger.
- Merzougui Mouna. Pr. U.B.M. Annaba.
- Laskri Yamina. Pr. E.S.T.I. Annaba.

Qu'elles trouvent, ici, l'expression de ma profonde gratitude. Enfin, Mes remerciements sont également adressés à tous mes professeurs, collègues et amis de la Faculté de Mathématiques de l'UBM, les membres du laboratoire Laps, et de la Faculté des Sciences de l'université de Constantine.

Résumé

On considère dans cette thèse, l'inférence Bayésienne pour certains modèles récents de durées de survie en présence de censures. Les modèles auxquels on s'est intéressé sont le modèle de Bertholon à défaillances concurrentes, le modèle de Zeghdoudi assujetti à une troncature et enfin le modèle de Burr XII à trois paramètres. Pour ces trois modèles, on a procédé à l'estimation des paramètres et de certaines caractéristiques de fiabilité, à savoir la fonction de survie et le taux de hasard à l'aide d'une approche Bayésienne sous différentes fonctions de perte aussi bien symétriques (perte quadratique) qu'asymétriques (perte LINEX et entropie) et avec des données censurées. Les censures envisagées dans cette thèse sont des censures à droite de type II pour les modèles de Bertholon et Zeghdoudi, et des censures progressives pour le modèle de Burr XII.

Dans le cas de l'approche classique, les estimateurs sont des solutions d'un système non linéaire dont les solutions ne sont pas explicites analytiquement ; des méthodes numériques ont été adoptées. Dans l'approche Bayésienne, les estimateurs sont donnés sous forme d'un rapport d'intégrales, les méthodes de Monte-Carlo par Chaîne de Markov (MCMC) et en particulier l'algorithme Metropolis -Hastings ont été utilisées pour procéder à des simulations et à une analyse de données permettant d'illustrer les résultats obtenus.

Dans le cas du modèle de Burr XII, outre l'estimation, le problème de la prédiction de la k -ième statistique d'ordre d'un échantillon futur indépendant de l'échantillon des observations a été étudié.

Finalement, Nous avons utilisé le critère de Pitman et le critère "integrated mean square error" (IMSE) pour comparer les estimateurs Bayésiens et les estimateurs du maximum de vraisemblance.

Mots-clés et expressions : La distribution de Burr XII à trois paramètres, La distribution de Bertholon, , la distribution de Zeghdoudi, troncature, données censurées progressives de type II, estimation du maximum de vraisemblance, estimation Bayésienne, méthodes MCMC, densité prédictive.

Abstract

We consider in this thesis, the Bayesian inference for some recent survival times models in the presence of censorship. We have been interested in the Bertholon model with concurrent failures, the Zeghdoudi model subject to truncation, and finally the Burr XII model with three parameters. For these three models, parameters and specific reliability characteristics have been estimated, namely survival function and failure rate using a Bayesian approach under different loss functions, symmetrical (quadratic loss), and asymmetrical (LINEX loss and entropy) with censored data. The censorship studied in this thesis is type II right censorships for Bertholon and Zeghdoudi model, and progressive censorship for the Burr XII model.

In the classical approach, estimators are solutions of a non-linear whose solutions are not analytically explicit so numerical methods were adopted. In the Bayesian approach, estimators are given by the ratio of integrals, the Monte Carlo methods by Markov Chain (MCMC) and in particular, the Metropolis-Hastings algorithm has been used to carry out simulations and data analysis to illustrate the results obtained. In the case of Burr XII's model, in addition to estimation, the problem of prediction of the order statistical k -ième of a future sample independent of the sample of observations was studied. Finally, we used the Pitman criterion and the "integrated mean" criterion. square error" (IMSE) to compare Bayesian and maximum likelihood estimators.

Keywords and phrases : Three parameters Burr XII distribution, Bertholon distribution, Zeghdoudi distribution, truncation, progressive type II censored data, Maximum likelihood estimation, Bayesian estimations, MCMC methods, Posterior predictive density.

ملخص

نعتبر في هذه الأطروحة الاستدلال البايزي لبعض النماذج الحديثة أوقات البقاء على قيد الحياة في ظل وجود الرقابة. النماذج التي كنا مهتمين بها هي نموذج بيرثولون مع إخفاقات متزامنة ، نموذج زغدودي يخضع ل اقتطاع وأخيرا نموذج بور الثاني عشر مع ثلاثة معلمات. بالنسبة لهذه النماذج الثلاثة، تم تقدير المعلمات وبعض خصائص الموثوقية ، وهي وظيفة البقاء على قيد الحياة ومعدل الصدفة باستخدام نهج بايزي تحت وظائف فقدان مختلفة على حد سواء متماثلة (الخسارة التربيعية) و غير متماثلة (فقدان *LINEX* والانتروبيا) ومع البيانات الخاضعة للرقابة. الرقابة المتوخاة في هذه الأطروحة توجد رقابة يمينية من النوع الثاني على نماذج بيرثولون و زغدودي، والرقابة التقدمية على نموذج بور الثاني عشر.

في حالة النهج الكلاسيكي ، فإن المقدرين هم حلول لغير خطي حلوله ليست صريحة تحليليا ؛ الطرق العددية تم اعتمادها. في النهج البايزي ، يتم إعطاء المقدرين في من نسبة التكاملات ، طرق مونت كارلو بواسطة سلسلة ماركوف (*MCMC*) وعلى وجه الخصوص تم استخدام خوارزمية متروبوليس هاستينغز لتنفيذ المحاكاة وتحليل البيانات لتوضيح النتائج التي تم الحصول عليها. في حالة نموذج بور الثاني عشر ، بالإضافة إلى التقدير ، فإن مشكلة التنبؤ من ترتيب الكيم الإحصائي لعينة مستقبلية مستقلة عن عينة تمت دراسة الملاحظات. وأخيرا، استخدمنا معيار بيتمان ومعيار المتوسط المتكامل خطأ مربع (*IMSE*) لمقارنة المقدرين البايزيين ومقدي الاحتمالات القصوى.

الكلمات الرئيسية والعبارات: ثلاثة معلمات توزيع *BurrXII* ، توزيع بارثولون *Bertholon* ، توزيع زغدودي، ترونكاتور، بيانات خاضعة للرقابة من النوع الثاني التدريجي ، تقدير أقصى احتمال ، تقديرات بابز ، طرق *MCMC* ، الكثافة التنبؤية الخلفية.

Introduction

La modélisation dans l'analyse de survie est un aspect très important dans de nombreux domaines aussi variés que différents tels que l'industrie, la médecine, les finances.... L'analyse statistique, après modélisation, de ce que l'on appelle "durée de vie" ou "fiabilité" est devenue un sujet d'intérêt considérable pour les statisticiens. Les principaux problèmes abordés dans ce domaine sont la modélisation des phénomènes liés à la survie par des distributions aléatoires, ensuite il s'agit d'en estimer dans un cadre paramétrique les paramètres ou tout autre quantité qui en dépend. Cependant, l'approche non paramétrique ou semi paramétrique peut être envisagée par l'estimation directe de la distribution.

Les inférences sont souvent menées avec des données complètes ou censurées à droite, on s'intéresse dans cette thèse essentiellement au cas où les données sont progressivement censurées introduites par Balakrishnan [6].

Parmi les modèles, quelques distributions occupent une place centrale en raison de leur utilité démontrée dans différentes situations. On cite entre autres, la distribution de Weibull (Tse et al (2003)), et El Sagheer (2018); la distribution de Pareto (Raqab et al (2010) ainsi que les modèles exponentiels et leurs nombreuses généralisations.

Dans cette thèse, on s'est intéressé aux modèles de Bertholon (2006) qui est un modèle à défaillances concourantes, le modèle de Zeghdoudi (2021) et le modèle de Burr XII.

L'utilisation de modèles à risque concurrents est bien connue et une grande famille de distributions a notamment été présentée dans Park et Padgett (2004)[36]. Par conséquent, Bertholon, H. [14] a présenté la distribution de Bertholon comme le minimum entre la distribution exponentielle et la distribution Weibull. La distribution de Lindley avec un paramètre est un mélange de distributions exponentielles et gamma . elle est large-

ment utilisée dans la modélisation des données de durée de vie dont la série d'articles de Zeghdoudi et Nedjar,[30], La distribution de Zeghdoudi a été introduite par Messaadia et Zeghdoudi (2018)[31], ils ont suggéré une famille de distribution exponentielle à un paramètre qui est basée sur des mélanges de distribution gamma.

Tejeda et Barry [48] présentent le modèle de Burr (2008) avec ses propriétés statistiques et la relation avec les autres distributions qui est utilisé dans l'analyse de fiabilité. L'estimation ponctuelle et par intervalle des paramètres de la loi de Burr XII est obtenue par la méthode du maximum de vraisemblance sur les données censurées et complètes. [21],[50].

Pour les trois modèles mentionnés ci-dessous, nous utilisons deux approches, l'approche classique du maximum de vraisemblance et l'approche Bayésienne pour estimer les paramètres de la distribution, la fonction de fiabilité ainsi que la fonction taux de panne. Les estimateurs de Bayes et les risques à posteriori (PR) correspondants ont été obtenus sous différentes fonctions de perte aussi bien symétriques (perte quadratique) qu'asymétriques (Linex et entropie). Les estimateurs ne peuvent pas être obtenus explicitement, nous utilisons donc la méthode de Monte Carlo (MCMC) pour approcher les solutions dans le cas d'une approche Bayésienne. Dans le cas de l'approche du maximum de vraisemblance, les estimateurs sont solutions de systèmes non linéaires et ont été approchées par l'utilisation de l'algorithme EM et le logiciel \mathcal{R} .

Nous utilisons également l'erreur quadratique moyenne intégrée (IMSE) et le critère de proximité de Pitman pour comparer les résultats des deux méthodes. Enfin, un ensemble de données réelles a été analysé pour étudier l'applicabilité et l'utilité du modèle de Burr XII.

Organisation de la thèse

Le chapitre un est consacré aux principaux outils mathématiques utilisés dans cette thèse, on cite l'approche Bayésienne sous ses différentes fonctions de perte ; L'analyse de survie et ses principales caractéristiques et enfin les différents types de censures. Au chapitre deux, on développe les méthodes de Monte Carlo par Chaîne de Markov *MCMC* ; on y décrit l'aspect théorique et les algorithmes qui en découlent ; on cite principalement les algorithmes de Metropolis-Astings et l'algorithme de Gibbs.

Au chapitre trois, on s'est intéressé à l'estimation des paramètres et des caractéristiques de fiabilité pour les modèles de Bertholon et Zeghdoudi. Les données sont supposées être

censurées à droite de type II, les lois à priori sur les paramètres sont des lois gamma et cela sous différentes fonctions de perte.

Le chapitre quatre traite le cas d'une distribution de Burr XII à trois paramètres, qui joue un rôle très important dans la modélisation des problèmes financiers. On y estime les paramètres, la fonction de survie et la fonction taux de hasard à l'aide de deux approches ; l'approche classique du maximum de vraisemblance et l'approche Bayésienne sous un pannel de fonctions de perte et avec des données progressivement censurées. Des simulations ont été menées afin de comparer les résultats obtenus par les deux approches à l'aide de l'erreur quadratique moyenne intégrée (IMSE) et du critère de proximité de Pitman. Enfin, un ensemble de données réelles a été analysé pour étudier l'applicabilité et l'utilité du modèle de Burr XII à trois paramètres.

Chapitre No. 1

Estimation Bayésienne

Ce chapitre traite de la statistique paramétrique dans un cadre Bayésien. La statistique Bayésienne est une théorie rivale à la statistique classique au sens que chacune d'elles propose vis-à-vis, d'un même problème ; une approche et une résolution différentes. Nous commençons par rappeler quelques définitions et propriétés des éléments théoriques de l'analyse classique, puis nous présentons les notions et les outils sur lesquels se fonde l'analyse Bayésienne.

1.1 Rappels sur la statistique classique

1.1.1 Notions de Base

Modèle statistique

La démarche statistique comporte généralement deux étapes. La première est une phase de modélisation, qui consiste à mettre un phénomène réel sous forme mathématique. En pratique, ceci revient à supposer que l'observation X est un objet aléatoire dont la loi P_X (inconnue) appartient à une famille de lois $(P_\theta)_{\theta \in \Theta}$ que l'on spécifie. Cette étape préliminaire, cruciale, est en grande partie une affaire de praticien : pour chaque domaine d'application (physique, chimie, biologie, etc.), ce sont les spécialistes du domaine qui fourniront cette modélisation.

Ceci étant supposé acquis, la seconde étape est celle qui nous occupe dans cette thèse, à savoir l'inférence statistique. Il s'agit, à partir du modèle $(P_\theta)_{\theta \in \Theta}$ et de l'observation

X , d'obtenir l'information la plus pertinente possible sur les paramètres en jeu dans le modèle, c'est-à-dire dans la loi de X . On rappelle que si X est un objet aléatoire (variable, vecteur, processus) à valeurs dans un espace mesurable (E, Σ) , sa loi P_X est définie pour tout A de E par :

$$P_X(A) = P(X \in A) = P(\{\omega \in \Omega : X(\omega) \in A\}),$$

probabilité que l'objet aléatoire X tombe dans l'ensemble A . Résumons ce qui vient d'être dit.

Définition 1.1. (*Expérience statistique*) Une expérience statistique est la donnée d'un objet aléatoire X à valeurs dans un espace mesurable (E, Σ) et d'une famille de lois $(P_\theta)_{\theta \in \Theta}$ sur cet espace, supposée contenir la loi P_X , et appelée modèle statistique pour la loi de X .

Dans cette définition, l'hypothèse fondamentale est bien entendu qu'il existe une valeur $\theta \in \Theta$ telle que $P_X = P_\theta$. Le vrai paramètre θ est inconnu mais l'espace Θ qui le contient est supposé connu.

1.1.2 Échantillonnage

Considérons l'expérience qui consiste à tirer n valeurs consécutives x_1, x_2, \dots, x_n d'une variable aléatoire X . Ces réalisations qui constituent un échantillon de X peuvent être aussi considérées comme une réalisation d'un n -uplet de variables aléatoires (X_1, \dots, X_n) indépendantes et de même loi, celle de X . Ce n -uplet est appelé un n -échantillon aléatoire de la variable sous-jacente X , ou plus simplement un échantillon aléatoire de la variable X .

1.1.3 Estimation

L'estimation statistique consiste à donner une valeur approchée à un paramètre inconnu θ , c'est ce qu'on appelle estimation ponctuelle. Mais il est souvent plus intéressant de rechercher un ensemble de valeurs contenu dans un intervalle ayant une forte probabilité de contenir la vraie valeur de θ appelé l'intervalle de confiance.

1.1.3.1 Estimation ponctuelle

Définition 1.2. *Un estimateur d'un paramètre θ est une statistique T ; fonction mesurable $T = T(X_1, X_2, \dots, X_n)$ à valeurs dans l'ensemble Θ*

IL existe plusieurs méthodes de l'estimation ponctuelle comme la méthode des moindres carrés, la méthode des moments et la méthode du maximum de vraisemblance. On va développer cette dernière :

L'estimation du maximum de vraisemblance

Dans cette partie, nous allons expliquer la méthode du maximum de vraisemblance qui est la plus couramment utilisée pour trouver une estimation d'un paramètre inconnue θ à partir d'un n-échantillon. Elle possède de bonnes propriétés de convergence et d'efficacité quand la taille d'échantillons est grande.

Définition 1.3. *On appelle la fonction de vraisemblance du n-échantillon, la densité de probabilité conjointe des n variables aléatoires X_1, X_2, \dots, X_n ; $L(x_1, x_2, \dots, x_n; \theta)$, on la considère comme une fonction de θ uniquement, on la note par $L(x, \theta)$.*

Soit (X_1, X_2, \dots, X_n) un n-échantillon d'une variable aléatoire X de loi $f(x, \theta)$, alors

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i, \theta)$$

La méthode du maximum de vraisemblance consiste à trouver les valeurs $\hat{\theta}$ qui maximisent la vraisemblance $L(x, \theta)$ i.e. :

$$\hat{\theta} = \theta \in \Theta \underset{\text{arg max}}{L(x, \theta)}, \quad (1.1)$$

alors $\hat{\theta}$ est appelée estimateur du maximum de vraisemblance (MV) et on le note $\hat{\theta}_{MV}$.

La monotonie de la fonction logarithme nous permet de trouver le maximum ($\hat{\theta}_{MV}$) en maximisant le logarithme de la fonction de vraisemblance appelée fonction log-vraisemblance notée $l(\theta)$, par cela on a l'avantage de calculer la somme au lieu du produit i.e. :

$$l(\theta) = \log(L(x|\theta)) = \sum_{i=1}^n \log(f(x_i|\theta)), \quad (1.2)$$

Propriétés d'un estimateur

Définition 1.4. (Biais)

Un estimateur θ^* est dit sans biais si, pour tout $\theta \in \Theta$, $E(\theta^*) = \theta$. on appelle **biais** la fonction $b(\theta) = E(\theta^*) - \theta$. On dit qu'un estimateur est biaisé lorsque le biais est non nul.

Définition 1.5. (Information de Fisher)

Soit X une variable aléatoire de densité $f(x, \theta)$, on appelle l'information de Fisher apportée par la variable aléatoire X sur le paramètre θ , la quantité suivante :

$$I_X(\theta) = E \left[\left(\frac{\partial}{\partial \theta} \log f(x, \theta) \right)^2 \right].$$

Sous certaines conditions, la quantité de l'information de Fisher peut être écrit sous cette forme

$$I_X(\theta) = -E \left[\left(\frac{\partial^2}{\partial \theta^2} \log f(x, \theta) \right) \right],$$

Définition 1.6. (Borne de Cramer Rao)

Soit θ^* un estimateur sans biais de θ , la borne de Cramer Rao (BCR) est définie par :

$$BCR = \frac{1}{I_n(\theta)}$$

où $I_n(\theta) = nI_X(\theta)$ est l'information de Fisher apportée par l'échantillon.

Définition 1.7. on dit qu'un estimateur est efficace s'il est sans biais et que sa variance atteint la borne de Cramer Rao.

1.2 Analyse de survie

L'objectif de notre travail est l'estimation dans un modèle de censure, pour cela on commence dans ce chapitre par introduire les données de survie et la censure avec ses différents types. L'analyse de survie est la modélisation du temps de survie d'un certain événement précis, l'événement étudié est le passage entre deux états spécifiques. On appelle ce temps la durée de vie ou la durée de survie, ou simplement une durée qui est une variable positive ou nulle, souvent supposé bornée. Parfois pour certaines raisons il arrive que les événements d'intérêts ne sont pas produit durant la période d'observation,

on appelle ce phénomène la censure de la série des données, autrement dit pour certaines données on possède seulement une borne inférieure ou supérieure et non pas la valeur exacte. Bien que cette théorie ait été développée en recherche biomédicale, on peut rencontrer des observations censurées dans plusieurs domaines. Par exemple en *médecine*, on utilise l'analyse de survie pour évaluer l'efficacité d'un traitement, c'est-à-dire estimer le temps de survie probable d'un patient. Pour cela, on prend un échantillon de patients, on peut avoir pour chacun d'eux soit :

1. La valeur exacte de la durée de survie (données complètes non censurées).
2. Une borne inférieure de cette durée (donnée censurée à droite).

Ce dernier cas se produit lorsqu'un patient ne peut pas être suivi jusqu'à la fin de l'étude Pour une raison spécifique, par exemple il a quitté le pays ou quand il est mort pour une cause indépendante de la maladie. En *ingénierie* l'analyse de survie est utilisée pour estimer la fiabilité des machines, de composantes électroniques ...etc .

En *démographie* l'analyse de survie sert à construire des tables de mortalité. Celles-ci sont utilisées pour déterminer le montant des assurances-vie entre autres.

1.2.1 Les fonctions de base

Soit T une variable aléatoire positive (appelée temps de survie) de fonction de répartition F , définie par

$$F : \mathbb{R} \mapsto [0, 1] , F(t) = P(T \leq t).$$

Si F est dérivable, la variable aléatoire T à une densité f (par rapport à la mesure de Lebesgue), f est telle que :

$$dF(x) = f(x)dx.$$

La fonction de survie au point t , notée $S(t)$, est donnée par la probabilité que l'individu survive au delà de t , c'est à dire :

$$S : \mathbb{R} \mapsto [0, 1] , S(t) = P(T > t) = 1 - F(t).$$

Il est clair que cette fonction est décroissante, continue à droite avec

$$\lim_{t \rightarrow 0} S(t) = 1 \text{ et } \lim_{t \rightarrow +\infty} S(t) = 0.$$

On appelle taux de hasard du temps de survie T , la fonction

$$\lambda(t) = \begin{cases} \frac{f(t)}{S(t)} & \text{si } S(t) \neq 0, \\ 0 & \text{sinon.} \end{cases}$$

Si f est continue sur \mathbb{R}_+^* alors :

$$\lambda(t) = \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} P(t < T \leq t + \Delta | T > t),$$

ce qui justifie l'appellation de taux.

La fonction de hasard cumulée de T est défini par :

$$\Lambda(t) = \int_0^t \lambda(s) ds.$$

Il est facile de montrer les résultats de la proposition suivante.

Proposition 1.1. *Si f est continue sur \mathbb{R}_+^* et si $A = \{t > 0 / S(t) \neq 0\}$ alors les propositions suivantes sont équivalentes.*

1. $\forall t \in A, \lambda(t) = \frac{f(t)}{S(t)}$.
2. $\forall t \in A, \lambda(t) = (-\log S(t))'$.
3. $\forall t \in A, S(t) = \exp \{-\Lambda(t)\}$.
4. $\forall t \in A, f(t) = \lambda(t) \exp \{-\Lambda(t)\}$.

1.2.2 Types de censures

Une des caractéristiques de l'analyse de survie est l'existence d'observations incomplètes à cause de la censure, donc au lieu d'observer un n -échantillon d'une variable aléatoire T , une variable de censure donne une information partielle permettant par exemple de fixer une borne inférieure (censure à droite) ou une borne supérieure (censure à gauche).

Il existe différents types de censures : censure de type I, si le temps de censure est fixé par le chercheur comme étant la fin de l'étude. La censure de type II, se caractérise par le fait que l'étude cesse aussitôt que se produit un nombre d'événements prédéterminé par l'expérimentateur. Une autre possibilité de censure ; est la censure aléatoire qui soit du cadre de cette thèse.

Censure à droite

La durée de vie est dite censurée à droite si l'individu n'a pas subi l'événement à sa dernière observation. En présence de censure à droite, les durées de vie (T) ne sont pas toutes observées ; pour certaines d'entre elles, on sait seulement qu'elles sont supérieures à une certaine valeur connue. Soit R une variable aléatoire de censure, au lieu d'observer la variable Y qui nous intéresse, on observe le couple de variables (Z, δ) avec $Z = \min(T, R)$ et $\delta = 1_{\{T \leq R\}}$. δ est appelé indicateur de censure puisque ses valeurs nous informent sur le fait que l'observation est complète (si $\delta = 1$) ou censurée à droite (si $\delta = 0$).

Un exemple illustratif est lorsqu'on s'intéresse à la durée de vie d'un genre de machines précis mais que ces dernières tombent en panne s'il se produit une surtension d'électricité. Ici, la durée de vie de la machine est censurée à droite par l'instant auquel se produit la surtension.

C'est le type de censure le plus fréquent dans la pratique.

Remarque 1.1. *Dans beaucoup de cas, l'hypothèse d'indépendance entre T et R peut être admise dans la pratique. C'est le cas pour l'exemple précédent ou dans le cas de malades qui sortent de l'étude suite à un déménagement.*

La fonction de vraisemblance pour les données censurées à droite est

$$\begin{aligned} L(\eta_0, \eta_1, \beta | X) &= \prod_{i=1}^n f_B(x_i)^{\delta_i} S_B(x_i)^{1-\delta_i} \\ &= \prod_{i=1}^n h_B(x_i)^{\delta_i} S_B(x_i). \end{aligned}$$

Censure à gauche

La censure à gauche correspond au cas où l'individu a déjà subi l'événement avant qu'il ne soit observé. On sait uniquement que la valeur d'intérêt est inférieure à une certaine

valeur connue représentée par une variable aléatoire L . Pour chaque individu, on peut associer un couple de variables aléatoires (Z, δ) telles que $Z = \max(Y, L)$ et $\delta = \mathbf{1}_{\{Y \geq L\}}$. Un des premiers exemples de censure à gauche rencontré dans les recherches concerne le cas d'observateurs qui s'intéressent à l'horaire où les babouins descendent de leurs arbres pour aller manger (les babouins passent la nuit dans les arbres). Le temps d'événement (descente de l'arbre) est observé si le babouin descend de l'arbre après l'arrivée des observateurs. Par contre, la donnée est censurée si le babouin est descendu avant l'arrivée des observateurs : dans ce cas on sait uniquement que l'horaire de descente est inférieur à l'heure d'arrivée des observateurs. On observe donc le maximum entre l'heure de descente des babouins et l'heure d'arrivée des observateurs.

Remarque 1.2. *Très peu de travaux s'intéressent à la seule censure à gauche car beaucoup moins fréquente et constitue un phénomène symétrique à celui de la censure à droite. Certains auteurs ont proposé de renverser l'échelle de temps.*

1.2.3 Censure progressive de type II

Dans la fiabilité, les études médicales, les tests sur les produits industriels, où ce qui nous intéresse est le temps d'échec (mort, défaut), parfois le taux d'échec est faible, donc le temps nécessaire pour observer la n -ième échec sera très long, cependant l'expérience doit être finie dans un temps prédéfini. Une méthode pour éviter ce problème est l'utilisation de la censure progressive.

La censure progressive est une censure où les unités en cours d'essai sont retirés de l'essai de vie dans un instant préfixé ou aléatoire. Il y a plus de types de la censure progressive, mais l'idée principale peut être exprimée par la censure progressive de type II, qui est considéré comme la plus utilisée.

1.2.3.1 La censure progressive de type II

On va faire un test sur n unités, simultanément, alors on note par X le temps d'échec, seulement m unités sont complètement observées jusqu'à l'échec, ensuite, avec un plan de censure, $\mathcal{R} = (r_1, \dots, r_m)$ avec

$$\sum_{i=1}^m r_i = n - m$$

- Dans le premier temps d'échec $X_{1,n}$, r_1 unités parmi les $n - 1$ restants sont aléatoirement retirés.

- Dans le deuxième temps d'échec $X_{2,n}$, r_2 unités parmi les $n - 2 - r_1$ restants sont aléatoirement retirés.

- Finalement dans le m -ième temps d'échec, tous les $n - m - r_1 - \dots - r_m$ unités sont retirés.

Donc, $X = (X_{1,m,n}, X_{2,m,n}, \dots, X_{m,m,n})$ est un échantillon progressivement censuré d type II.

On remarque si $r_1 = r_2 = \dots = r_m = 0$, il n y a pas de censure (cas complet).

La densité de probabilité conjointe (Balakrishnan, N., Aggarwala, R) est donnée par

$$f_{X_{1,n,m}, \dots, X_{m,n,m}}(x_{1,n,m}, \dots, x_{m,n,m}) = A \prod_{i=1}^m f(x_{1,n,m}, \dots, x_{m,n,m}) [1 - F(x_{1,n,m}, \dots, x_{m,n,m})]^{r_i} \quad (1.3)$$

$$A = n(n - r_1 - 1)(n - r_1 - r_2 - 2) \dots \left(1 - \sum_{i=1}^{m-1} (r_i + 1)\right). \quad (1.4)$$

La génération de la censure progressive de type II

Dans la censure progressive de type II, les unités sont retirées dans chaque temps d'échec, pour cela le nombre des observation est préfixé, tandis que la durée de l'expérience est aléatoire, la génération de statistiques d'ordre censurées progressivement de type II est comme suit

Soit les variables aléatoires X_1, X_2, \dots, X_n sur l'espace probabilité (Ω, \mathcal{A}, P) et $\mathcal{R} = (r_1, \dots, r_m)$ un vecteur de \mathbb{N}^* ,

Pour $\omega \in \Omega$ l'échantillon progressivement censuré d type II est :

$$X_{1,m,n}^{\mathcal{R}}(\omega), X_{2,m,n}^{\mathcal{R}}(\omega), \dots, X_{m,m,n}^{\mathcal{R}}(\omega)$$

généralisé de $X_1(\omega), X_2(\omega), \dots, X_n(\omega)$ par les étapes suivantes

1. On calcule la statistique d'ordre : $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$.
2. $N_1 = \{1, \dots, n\}, i = 1$

3. $k_i = \min N_i$, et prend $X_{i,m,n}(\omega) = X_{(k_i)}(\omega)$.
4. Choisi par hasard et sans remplacement un échantillon $R_i \subseteq N_i \setminus \{k_i\}$ avec $|R_i| = r_i$.
5. Si $i < m$ alors, $N_{i+1} = N_i \setminus (\{k_i\} \cup r_i)$ et retourne à 3, sinon arrête.

Ainsi

$$X_{1,m,n}^{\mathcal{R}}(\omega), \dots, X_{m,m,n}^{\mathcal{R}}(\omega) = (X_{(1)}(\omega), \dots, X_{(m)}(\omega))$$

Théorème 1.1. On pose $X_{1,m,n}, \dots, X_{m,m,n}$ et $U_{1,m,n}, \dots, U_{m,m,n}$ sont des statistiques d'ordre progressivement censurées de type II, selon une loi de fonction de répartition F et une loi uniforme respectivement, alors,

$$(X_{i,m,n})_{1 \leq i \leq m} \stackrel{d}{=} (F^{-1}(U_{i,m,n}))_{1 \leq i \leq m}$$

Théorème 1.2. Soit les statistique d'ordre $X_{1,m,n}, \dots, X_{m,m,n}$ de fonction de répartition quelconque F et les variables $U_1, \dots, U_m \stackrel{iid}{\sim} U(0,1)$, alors,

$$X_{i,m,n} \stackrel{d}{=} F^{-1} \left(1 - \prod_{j=1}^i U_j^{1/\gamma_j} \right) \quad 1 \leq i \leq m$$

où γ_j représente le nombre d'objets survivants avant le j -ième temps d'échec, i.e

$$\gamma_j = \sum_{i=j}^m (r_i + 1)$$

L'étude de simulation

La génération des statistiques d'ordre progressivement censurées de type II a été introduite la première fois par Balakrishnan et Sandhu [8], voir aussi [7]. Ils avaient proposé la méthode plus utilisée, cependant l'approche évident pour généré les statistiques d'ordres avec une censure progressive de type II est obtenue par l'algorithme que l'on peut appliquer pour toutes les distributions, on a l'échantillon X_1, \dots, X_n , les étapes sont données par

Algorithme 1

1. On calcule les statistiques d'ordres $X_{(1)}, \dots, X_{(n)}$ de l'échantillon X_1, \dots, X_n .
2. $N_1 = \{1, \dots, n\}$
3. Pour $i = 1, \dots, m$ faire
4. $k_i = \min N_i$, et prend $X_{i,m,n} = X_{(k_i)}$.
5. Choisi par hasard et sans remplacement un échantillon $R_i \subseteq N_i \setminus \{k_i\}$ avec $|R_i| = r_i$.
6. Si $i < m$ alors, $N_{i+1} = N_i \setminus (\{k_i\} \cup r_i)$ et retourne à 4, sinon arrête.

Tout l'échantillon X_1, \dots, X_n est nécessaire, comme n est d'habitude plus large que m , cette méthode prend une longue durée de temps, pour cela la deuxième méthode (introduite par Balakrishnan, sandhu) est basée sur le théorème 1, qui fournit la génération des statistiques d'ordre d'une loi uniforme.

Donc pour un n -échantillon X_1, \dots, X_n , tout simplement on utilise le théorème 2 ,c'est à dire,

$$X_{i,m,n} \stackrel{d}{=} F^{-1} \left(1 - \prod_{j=1}^i U_j^{1/\gamma_j} \right)$$

où les U_1, \dots, U_m sont des variables indépendantes et identiquement distribuées d'une loi uniforme $U(0,1)$.

Algorithme 2

1. généré m variables aléatoires iid $U_1, \dots, U_m \sim U(0,1)$.
2. calculer $B_j = U_j^{1/\gamma_j}, j = 1, \dots, m$
3. prend $V_0 = 1$, et calcule $V_k = B_k V_{k-1}, k = 1, \dots, m$.
4. $U_{r;m,n} = (1 - V_r), r = 1, \dots, m$.
5. $X_{r;m,n} = F^{-1}(U_{r;m,n}) r = 1, \dots, m$.

donc, dans cet approche il suffit de calculer F^{-1} , il est claire que Cet algorithme est facile à utiliser dans le cas où la fonction de répartition a une représentation explicite.

1.3 Estimation Bayésienne

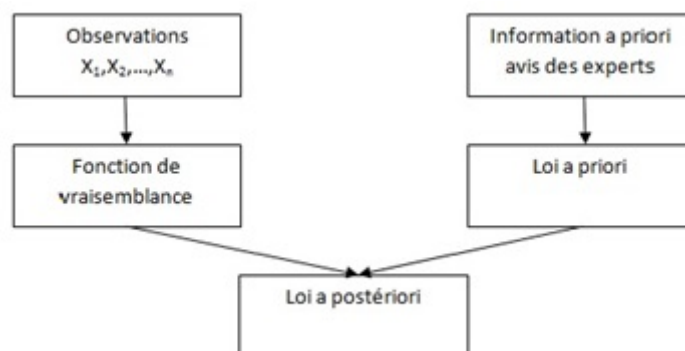
Dans un problème d'estimation d'un paramètre θ d'un certain modèle, la statistique classique considère θ comme un paramètre fixé appartenant à un espace Θ et on l'estime sur la base de l'échantillon observé, par contre, l'approche Bayésienne consiste à considérer θ comme une variable aléatoire qu'on estime sur la base des informations provenant de l'échantillon et en tenant compte de l'avis des experts, résumé par une loi sur θ appelée loi a priori.

Étant donné un modèle paramétrique d'observation $X \sim f(x/\theta)$ où f est la fonction de densité qui dépend d'un paramètre inconnu θ à estimer, $\theta \in \Theta$ espace de dimension fini qu'on appelle souvent ensemble des états.

Le point de vue Bayésien vise à exploiter le plus efficacement possible l'information apportée par X sur le paramètre θ , pour ensuite construire des procédures d'inférence. Bien que X ne soit qu'une réalisation aléatoire d'une loi gouvernée par θ , elle apporte une actualisation aux informations préalablement recueillies par l'expérimentateur.

L'analyse statistique Bayésienne est fondée essentiellement sur deux lois, la première est la loi a priori notée $\pi(\cdot)$ qui est le moteur de cette approche, elle résume l'information sur θ en dehors des connaissances apportées par les résultats expérimentaux. La deuxième est la loi a posteriori notée $\pi(\cdot|x)$ qui est un résumé complet de l'information disponible sur le paramètre θ .

Le schéma suivant résume le paradigme Bayésien :



1.3.1 Introduction à la théorie de la décision Bayésienne

Avant de parler de l'estimation, un passage sur la théorie de la décision est très utile, car comme nous allons le voir, déterminer un estimateur de Bayes revient à déterminer

une règle de décision.

En plus de l'espace des observations et l'espace des paramètres, un problème de décision est en général fondé sur les éléments suivants :

A : l'ensemble des règles de décisions (les estimateurs possibles).

D : l'espace des actions ou décisions, dont les éléments sont des images de l'observation par une règle de décision δ .

$L(\theta, \sigma)$: Fonction de perte évalue le coût d'une décision δ quand le paramètre vaut θ .

Elle permet donc, en quelque sorte, de quantifier la perte encourue par une mauvaise décision, une mauvaise évaluation de θ .

Définition 1.8. *Fonction de perte (de coût) on a δ un élément de D , on appelle la fonction de perte $L(\theta, \delta)$ la fonction mesurable de $(\Theta \times D)$ à valeurs dans \mathbb{R}_+ , telle que,*

1. $\forall (\theta, \delta) \in \Theta \times D, L(\theta, \delta) \geq 0$
2. $\forall \theta \in \Theta, \exists \delta^* \in D$ tels que : $L(\theta, \delta^*) = 0$

1.3.1.1 Le risque Bayésien moyen (fréquentiste)

Soit $L(\theta, \delta)$ une fonction de perte, on appelle risque Bayésien moyen associé à cette fonction, noté $R(\theta, \delta)$, la moyenne de la fonction de perte.

i.e :

$$\begin{aligned} R(\theta, \delta) &= E(L(\theta, \delta(x))) \\ &= \int_X L(\theta, \delta(x)) f(x|\theta) dx. \end{aligned}$$

1.3.1.2 Le risque de Bayes (intégré)

On peut considérer la moyenne du risque fréquentiste i.e. la moyenne du risque Bayésien moyen associé a la fonction de perte l par rapport à la loi a priori : $E^\pi[R(\theta, \delta)]$. Il s'agit du risque Bayésien ou risque de Bayes que l'on note $r(\pi, \delta)$.

On a :

$$\begin{aligned}
 r(\pi, \delta) &= E^\pi[R(\theta, \delta)] \\
 &= \int_{\Theta} R(\theta, \delta)\pi(\theta)d\theta \\
 &= \int_{\Theta} \int_X L(\theta, \delta(x))f(x|\theta)dx\pi(\theta)d\theta \\
 &= \int_{\Theta} \int_X L(\theta, \delta(x))\pi(\theta)f(x|\theta)dx d\theta.
 \end{aligned}$$

1.3.1.3 Le risque a postérieur

on définit le risque a postérieur noté $\rho(\pi, \delta)$ comme étant la moyenne de la fonction de perte par rapport à la loi a postérieur :

$$\rho(\pi, \delta|x) = E^{\pi(\cdot|x)}[L(\theta, \delta(x))] = \int_{\Theta} L(\theta, \delta(x))\pi(\theta|x)d\theta.$$

On a le résultat suivant en posant f_X la densité marginale de X :

$$\begin{aligned}
 r(\pi, \delta) &= \int_X \int_{\Theta} L(\theta, \delta(x))\pi(\theta|x)d\theta f_X(x)dx \\
 &= \int_X \rho(\pi, \delta|x)f_X(x)dx.
 \end{aligned} \tag{1.5}$$

i.e : le risque de Bayes $r(\pi, \delta)$ est la moyenne du risque a postérieur $\rho(\pi, \delta|x)$ suivant la loi marginale $f(x)$.

1.3.2 Spécification de la distribution a priori

le choix de la loi a priori est crucial pour l'analyse Bayésienne car elle touche d'une façon directe les distributions a posteriori. *Comment traduire alors l'information a priori en loi a priori ?*, Cette question est fondamentale dans la pratique car il est très rare que l'information a priori disponible soit suffisamment précise pour pouvoir déterminer exactement cette loi. Schématiquement, on peut mettre en évidence deux modes de pensée. Le premier est subjectif et considère que la distribution à priori traduit des connaissances

résultant d'expériences professionnelles et d'intuitions raisonnables avant l'observation des données. Ces informations sont exprimées par une loi dite "informatrice". Le second mode de pensée est plus objectif. Il est utilisé dans le cas où on a peu d'information. Il s'agit alors de pouvoir rester Bayésien en l'absence d'information a priori. On recherche donc des distributions a priori "non informatives" (ou vagues) exprimant l'ignorance a priori mais traitant les paramètres comme aléatoires.

1.3.3 La loi a priori informative

Elle repose sur l'information disponible sur le paramètre θ obtenu des opinions des chercheurs et experts. L'interprétation de l'information a priori est rarement assez précise pour conduire à la détermination d'une seule loi a priori. Il existe néanmoins des lois calibrées en fonction de la distribution des observations dites loi conjuguée.

Il existe plusieurs procédés pour obtenir des lois conjuguées, nous donnons ici l'un des plus intéressants qui est celui des familles conjuguées.

Définition 1.9. Une famille F de lois sur Θ est dite conjuguée pour la vraisemblance L (la loi de l'échantillon) si pour tout $\pi \in F$, $\pi(\cdot|x) \in F$.

Une famille conjuguée est un concept qui simplifie considérablement le calcul des distributions a posteriori.

Le tableau ci-dessous présente quelques familles de lois a priori conjuguée quand la loi de l'échantillon est choisie parmi quelques lois usuelles de la famille exponentielle.

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
Normale $N(\theta, \sigma^2)$	$N(\mu, r^2)$	$N(\Phi(\sigma^2\mu + r^2), \phi\sigma^2r^2)$ Avec $\Phi^{-1} = \sigma^2 + r^2$
Poisson $P(\theta)$	Gamma $\Gamma(\alpha, \beta)$	Gamma $\Gamma(\alpha + x, \beta + 1)$
Gamma $\Gamma(\mu, \theta)$	Gamma $\Gamma(\alpha, \beta)$	Gamma $\Gamma(\alpha + \mu, \beta + x)$

1.3.4 la loi a priori non informative

La théorie Bayésienne peut être appliquée même quand on ne dispose pas d'information a priori ou dans le cas où c'est difficile de traduire en terme de loi a priori l'information disponible sur le paramètre.

Dans de telles situations il est possible de bâtir une distribution a priori sur des considérations subjectives par des techniques qui intègrent notre ignorance sur les paramètres d'intérêt, de telles méthodes sont appelées non informatives.

le choix d'une loi a priori non informative conduit parfois a la spécification d'une mesure et non pas d'une probabilité, c'est-à-dire une loi impropre (une fonction mesurable définie sur l'ensemble \mathbb{R} de mesure π non intégrable i.e. : $\int_{\mathbb{R}} \pi(\theta)d\theta = +\infty$).

Nous donnons dans ce qui suit une des techniques les plus populaires dans la construction des lois a priori non informatives.

1.3.4.1 loi a priori de Jeffrey

Une méthode proposée par Jeffrey (1961) permet de fabriquer des lois a priori non informatives. Cette méthode utilise l'information de Fisher : $I(\theta)$, l'argument pourrait être le suivant : $I(\theta)$ représente une mesure de la quantité d'information sur θ contenue dans l'observation. Plus $I(\theta)$ est grande, plus l'observation apporte de l'information.

Il semble alors naturel de favoriser (au sens rendre plus probable suivant $\pi(\theta)$) les valeurs de θ pour lesquelles $I(\theta)$ est grande ; ce qui minimise l'influence de la loi a priori au profit de l'observation.

La règle de Jeffrey consiste donc à assigner à un modèle d'échantillonnage caractérisé par sa vraisemblance $L(x|\theta)$, la loi a priori de mesure :

$$\pi(\theta) = |\det I(\theta)|^{1/2},$$

où θ est un vecteur de paramètres inconnus et $I(\theta)$ est la matrice d'information de Fisher définie par :

$$I(\theta) = -\{E_{\theta}[\frac{\partial^2 \log L(\theta)}{\partial \theta_i \partial \theta_j}]\}_{1 \leq i, j \leq n}.$$

Exemple 1.3.1. (Berger et Yang 1995)

Soit $X \sim N(\mu, \sigma^2)$ où la moyenne μ et la variance σ^2 sont des paramètres inconnus.

La fonction de densité s'écrit :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\}.$$

Le paramètre d'intérêt est le vecteur $\theta = (\mu, \sigma)$.

La matrice d'information de Fisher est donnée comme suit :

$$I(\theta) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}.$$

La loi a priori de Jeffrey s'écrit :

$$\pi(\theta) = \frac{1}{\sqrt{2}\sigma^3}.$$

Il existe d'autres techniques pour construire des lois a priori non informatives comme celle de la loi a priori invariante et de référence, mais la loi de Jeffrey reste jusqu'à maintenant la plus utilisée.

1.3.5 La loi a postérieure

Une fois la loi a priori trouvée, nous pouvons maintenant construire la loi a postérieure, cela se fait à l'aide de la formule de Bayes, que nous rappelons.

La formule de Bayes :

Le théorème de Bayes est utilisé dans l'inférence statistique pour mettre à jour et actualiser les estimations d'une probabilité ou d'un paramètre quelconque, à partir des observations et des lois de probabilité de ces observations.

Le théorème de Bayes est d'abord donné, dans sa forme la plus simple par

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)},$$

où A et B sont des ensembles mesurables avec $P(B) \neq 0$.

En ce qui concerne les variables aléatoires. Il est connu que si (X, Y) est un vecteur de \mathbb{R}^n de densité conjointe f , alors la densité conditionnelle de Y sachant $X = x$ pour $f_X(x) \neq 0$ (f_X étant la densité marginale de X) est donnée par :

$$f_Y^{X=x}(y) = \frac{f(x, y)}{f_X(x)}.$$

On dispose d'un n -échantillon (X_1, X_2, \dots, X_n) i.i.d. de même loi d'une variable aléatoire dont la fonction de vraisemblance s'écrit

$$L(x|\theta) = \prod_{i=1}^n f(x_i|\theta) = f(x|\theta)^n.$$

la loi à postériori :

La distribution à postériori de θ notée $\pi(\theta|x)$ représente l'information dont on dispose sur θ après l'observation de x , par la formule de Bayes on trouve

$$\pi(\theta|x) = \frac{L(x|\theta)\pi(\theta)}{\int_{\Theta} L(x|\theta)\pi(\theta)d\theta} \propto L(x|\theta)\pi(\theta),$$

où \propto signifie proportionnel à.

$\pi(\cdot|x)$ véhicule une combinaison entre l'information a priori et l'information tirée de l'échantillon et sert à effectuer l'estimation visée.

1.3.6 Estimateur Bayésien

Comme on a déjà mentionné, toute estimation entraîne un coût qu'on mesure à l'aide de la fonction de perte. Intuitivement on cherche une décision qui minimise en moyenne cette dernière.

Définition 1.10. *Un estimateur Bayésien est la règle de décision σ^π qui minimise $r(\pi, \sigma)$, c'est-à-dire qui vérifie : $r(\pi, \sigma^\pi) = \text{Argmin}_{\sigma \in D} r(\pi, \sigma)$.*

D'après (1.5), minimiser $r(\pi, \sigma)$ pour toute valeur de x , est équivalent à minimiser la fonction de risque a posteriori.

Cette minimisation, peut se faire analytiquement comme on peut l'approcher numériquement selon la complexité de la fonction de perte l et de la loi a posteriori $\pi(\cdot|x)$.

Estimateur de Bayes sous des fonctions de perte classiques :

Pour des fonctions de perte classiques (quadratique, absolue,...) les estimateurs de Bayes correspondants sont des caractéristiques usuelles de la distribution a posteriori, médiane,...).

Fonction de perte quadratique

(La perte quadratique)

Une fonction de perte quadratique est une fonction $L : \Theta \times D \rightarrow \mathbb{R}_+$ donnée par :

$$L(\theta, \delta) = (\theta - \delta)^2.$$

Ainsi, soit :

$$\begin{aligned}\rho(\theta, \delta|x) &= E^{\pi(\cdot|x)}(L(\theta, \delta)) \\ &= \int_{\Theta} (\theta - \delta)^2 \pi(\theta|x) d\theta \\ &= \int_{\Theta} \theta^2 \pi(\theta|x) d\theta - 2\delta \int_{\Theta} \theta \pi(\theta|x) d\theta + \delta^2 \int_{\Theta} \pi(\theta|x) d\theta \\ &= E(\theta^2|x) - 2\delta E(\theta|x) + \delta^2.\end{aligned}$$

La décision δ qui minimise $\rho(\delta, x)$ est celle qui vérifie

$$\frac{d}{d\delta} \rho(\delta, x) = 0 \quad \text{et} \quad \frac{d^2}{d\delta^2} \rho(\delta, x) \geq 0.$$

Ce qui donne

$$-2E(\theta|x) + 2\delta = 0.$$

Et donc

$$\delta = E(\theta|x).$$

Donc pour la perte quadratique, l'estimateur de Bayes est la moyenne de la loi a posteriori.

La fonction de perte quadratique généralisée (GQ) est donnée par la formule

$$L(\theta, \delta) = \omega(\theta)(\theta - \delta)^2,$$

où $\omega(\theta)$ est une fonction non négative et l'estimateur de Bayes correspondant est

$$\hat{\delta}_{GQ} = \frac{E_{\pi}(\theta\omega(\theta))}{E_{\pi}(\omega(\theta))}.$$

Le risque a posteriori pour la fonction de perte GQ est $E_{\pi}(\omega(\theta)(\theta - \hat{\delta}_{GQ})^2)$,

où $\omega(\theta)$ est une fonction positive,

par exemple, $\omega(\theta) = \theta^{\gamma-1}$ et γ sont supposés être un entier.

Fonction de perte Linex (linéaire exponentielle)

La fonction de perte Linex (Linear Exponential) a été introduite par Varian (1975). Elle est asymétrique et très pratique. Cette fonction croît presque exponentiellement d'un côté de zéro et est approximativement linéaire de l'autre côté. Sous l'hypothèse que la perte minimale est obtenue pour $\tilde{\theta} = \theta$; la fonction de perte Linex pour θ s'exprime par

$$L(\Delta) \propto e^{a\Delta} - a\Delta - 1 \quad a \neq 0; \text{ ou } \Delta = \tilde{\theta} - \theta; \tilde{\theta} : \text{ l'estimateur de } \theta \quad (1.6)$$

Le signe et la norme de a représentent respectivement la direction et degré de symétrie.

$a > 0$: sur estimation plus grave que la sous estimation (et vice versa).

$a \approx 0$ Linex \rightarrow quadratique.

L'estimateur de Bayes associe à $L(\Delta)$ (l'estimateur qui minimise l'espérance a posteriori)

$$E_{\theta}(L(\tilde{\theta} - \theta)) \propto e^{-a\tilde{\theta}} E_{\theta}(e^{-a\theta}) - a(\tilde{\theta} - E_{\theta}(\theta)) - 1 \quad (1.7)$$

par la dérivation on trouve

$$\tilde{\theta}_L = -\frac{1}{a} \log E_{\theta}(e^{-a\theta}), \quad \text{ou } E_{\theta}(e^{-a\theta}) < \infty \quad (1.8)$$

Fonction de perte Entropie

La fonction de perte entropie découle de la fonction de perte Linex, et a été utilisée par Calabria et Pulcini (1994); elle est définie par :

$$L(\theta, \delta) = \left(\frac{\delta}{\theta}\right)^p - p \ln \left(\frac{\delta}{\theta}\right) - 1 \quad (1.9)$$

Le tableau suivant résume les différentes fonctions de perte utilisées

Exemple 1.3.2.

$$l(\theta, \delta) = \begin{cases} k_2(\theta - \delta(x)) & \text{si } \theta > \delta(x), \\ k_1(\delta(x) - \theta) & \text{sinon.} \end{cases}$$

Fonction de perte	Expression	Estimateur Bayésien	risque a posteriori
quadratique généralisée	$L(\theta, \delta) = \omega(\theta)(\theta - \delta)^2$	$\hat{\delta}_{GQ} = \frac{E_\pi(\omega(\theta)\theta)}{E_\pi(\omega(\theta))}$	$E_\pi(\omega(\theta)(\theta - \delta)^2)$
Entropie	$L(\theta, \delta) = \left(\frac{\delta}{\theta}\right)^p - p \log\left(\frac{\delta}{\theta}\right) - 1$	$\hat{\delta}_E = E_\pi(\theta^{-p})^{-\frac{1}{p}}$	$p[E_\pi(\log(\theta - \log(\hat{\delta}_E)))]$
Linex	$L(\theta, \delta) = \exp(r(\delta - \theta)) - r(\delta - \theta) - 1$	$\hat{\delta}_L = \frac{-1}{r} \log(E_\pi(\exp(-r\theta)))$	$r(\hat{\delta}_{GQ} - \hat{\delta}_L)$

TABLE 1.1 – Les estimateur et le risque des fonctions de perte

Calculons le risque a posteriori :

$$\begin{aligned}
E^{\pi(\cdot|x)}[l(\theta, \delta(x))] &= \int_{\Theta} l(\xi, \delta(x))\pi(\xi|x)d\xi \\
&= k_1 \int_{-\infty}^{\delta(x)} (\delta(x) - \xi)\pi(\xi|x)d\xi + k_2 \int_{\delta(x)}^{+\infty} (\xi - \delta(x))\pi(\xi|x)d\xi.
\end{aligned}$$

En remarquant que : $\pi(\xi|x)d\xi = dF(\xi|x) = -d(1 - F(\xi|x)) = dP^{\pi(\cdot|x)}(\theta \leq \xi)$ et en utilisant l'intégration par parties on écrit donc :

$$\begin{aligned}
\rho(\pi, \delta(x)) &= E^{\pi(\cdot|x)}[l(\theta, \delta(x))] \\
&= [k_1(\delta(x) - \xi)F(\xi|x)]_{-\infty}^{\delta(x)} + k_1 \int_{-\infty}^{\delta(x)} P^{\pi(\cdot|x)}(\theta \leq \xi)d\xi \\
&\quad - [k_2(\xi - \delta(x))(1 - F(\xi|x))]_{\delta(x)}^{+\infty} + k_2 \int_{\delta(x)}^{+\infty} P^{\pi(\cdot|x)}(\theta > \xi)d\xi \\
&= k_1 \int_{-\infty}^{\delta(x)} P^{\pi(\cdot|x)}(\theta \leq \xi)d\xi + k_2 \int_{\delta(x)}^{+\infty} P^{\pi(\cdot|x)}(\theta > \xi)d\xi.
\end{aligned}$$

Dérivant par rapport à $\delta(x)$, il vient :

$$\begin{aligned}
\frac{\partial \rho(\pi, \delta(x))}{\partial \delta(x)} &= k_1 P^{\pi(\cdot|x)}(\theta \leq \delta(x)) - k_2 P^{\pi(\cdot|x)}(\theta > \delta(x)) = 0 \\
&\iff k_1 P^{\pi(\cdot|x)}(\theta \leq \delta(x)) - k_2 (1 - P^{\pi(\cdot|x)}(\theta \leq \delta(x))) = 0 \\
&\iff (k_1 + k_2) P^{\pi(\cdot|x)}(\theta \leq \delta(x)) - k_2 = 0,
\end{aligned}$$

d'où

$$F(\delta|x) = P^{\pi(\cdot|x)}(\theta \leq \delta(x)) = \frac{k_2}{k_1 + k_2}.$$

1.3.7 Prédiction Bayésienne pour les statistique d'ordre

On suppose que la durée de vie $X_{1,m,n}, \dots, X_{m,m,n}$ est un échantillon censuré progressif de type II de taille m et de schéma de censure $R = (r_1, \dots, r_m)$, généré de de la distribution de Burr XII (α, θ, β) , cet échantillon représente l'échantillon informatif (passé), d'autre part on met Y_1, \dots, Y_N , un échantillon aléatoire de taille N d'observations futures (non observées) à partir de la même distribution, en supposant que les deux échantillons sont indépendants, notre objectif est de faire une prédiction Bayésienne pour le k -ème, $1 \leq k \leq N$ durée de vie ordonnée dans un futur échantillon de taille N .

La fonction de densité de la durée de vie ordonnée k -ième, Y_k dans l'échantillon futur (de taille N) est donnée par

$$g_{(k)}(y_k; \alpha, \theta, \beta) = k \binom{N}{k} [S(y_k; \alpha, \theta, \beta)]^{N-k} [F(y_k; \alpha, \theta, \beta)]^{k-1} f(y_k; \alpha, \theta, \beta).$$

Chapitre No. 2

Les méthodes de Monte Carlo par chaîne de Markov

Dans ce chapitre on rappelle les méthodes de Monte Carlo par chaînes de Markov (MCMC) et les différents algorithmes qui en , en fournissant suffisamment d'informations pour permettre de les appliquer de façon simple. Nous commençons par donner une introduction à ces méthodes, ensuite, nous décrivons les concepts importants pour l'application de ces méthodes.

2.1 Introduction

Les méthodes de Monte Carlo par chaînes de Markov ont été introduites dans les années 1950, elles consistent en une classe d'algorithmes permettant de simuler des variables aléatoires provenant de distributions complexes, autrement dit ces méthodes sont utilisées pour l'échantillonnage à partir de distributions de probabilité d'intérêt, des distributions qui souvent comportent plusieurs dimensions. Ces algorithmes se basent sur le parcours de chaînes de Markov ayant, pour lois stationnaires, les distributions à échantillonner. Le développement de cette méthodologie n'a pas seulement changé les solutions aux problèmes, mais a également modifié la façon d'aborder ces problèmes. Aujourd'hui, elles constituent un centre d'intérêt pour les chercheurs.

Les méthodes MCMC ont été créées peu après l'apparition des méthodes de Monte Carlo, mais leur impact dans le domaine de la statistique n'a pas été vraiment senti

jusqu'au début des années 1990. Aujourd'hui, malgré le nombre considérable d'algorithmes de simulation disponibles, les méthodes de Monte Carlo par chaîne de Markov constituent une des approches les plus utilisées dans la communauté statistique.

Il a fallu près de 40 ans pour que les méthodes MCMC pénètrent les statistiques générales. Leurs origines se trouvent dans la littérature de physique statistique, et ont été utilisées pendant une dizaine d'années dans les statistiques spatiales. Au cours des dernières années, les méthodes MCMC ont eu un effet profond sur la statistique Bayésienne, et ont également trouvé des applications dans les statistiques classiques. Des recherches récentes ont ajoutée considérablement à son ampleur la richesse de sa méthodologie, et ses fondements théoriques.

2.2 Chaînes de Markov

Les chaînes de Markov sont des processus aléatoires admettant un nombre d'états qui change au cours d'un temps discret, le nouvel état ne dépend que de l'état présent. Nous donnons les propriétés de ces chaînes, ensuite nous nous intéressons aux plus importantes définitions d'états (récurrent, ...), avec lesquelles nous déduirons les différents types des chaînes de Markov.

Définition 2.1. (Chaîne de Markov)

Soit E un ensemble fini ou dénombrable de N éléments, On appelle chaîne de Markov à valeur dans E toute suite de variables aléatoires $\{X_n\}_{n \geq 0}$, provenant d'un certain support S et respectant la propriété markovienne :

$$P(X_{n+1}|X_n, \dots, X_0) = P(X_{n+1}|X_n).$$

Autrement dit, la valeur d'une variable aléatoire de cette suite ne dépend que de celle qui la précède.

Définition 2.2. (Homogénéité)

Une chaîne de Markov est dite homogène pour toute suite d'état $(x_0, x_1, \dots, x_{n-1}, x_n, x_{n+1})$ de E si

$$P(X_{n+1} = x_{n+1}|X_n = x_n) = P(X_n = x_{n+1}|X_{n-1} = x_n)$$

Autrement dit que les probabilités $P(X_{n+1}|X_n)$ ne dépendent pas de n .

Définition 2.3. On appelle la probabilité de transition de X l'application p définie de $E \times E$ dans $[0, 1]$;

$$\forall x, y \in E, p(x, y) = P(X_{n+1} = y | X_n = x)$$

et on note $p(x, y) = p_{xy}$.

- La probabilité de transition p est dite symétrique si pour tout $x, y \in E$ $p(x, y) = p(y, x)$.

- La matrice de transition de X est la matrice $P = [p(x, y)]_{x_0 \leq x, y \leq x_{n+1}}$ de taille $N \times N$ où $N = n + 1$.

- $\mu(x) = P(X_0 = x)$ est la loi initiale de X , si μ et p sont connus, on peut calculer directement la loi de X_n

Proposition 2.1. Pour toute suite x_0, x_1, \dots, x_n dans E , on a

$$\begin{aligned} P(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) &= P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) \\ &\quad \times P(X_{n-1} = x_{n-1}, \dots, X_0 = x_0) \\ &= \mu(x_0)p(x_0, x_1)p(x_1, x_2)\dots p(x_{n-1}, x_n) \end{aligned}$$

Alors, on peut déduire une autre définition de la chaîne de Markov comme ce qui suit

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = \frac{\mu(x_0)p(x_0, x_1)p(x_1, x_2)\dots p(x_{n-1}, x_n)}{P(X_{n-1} = x_{n-1}, \dots, X_0 = x_0)} \quad (2.1)$$

Propriété 2.1. Toute matrice de transition P vérifie les propriétés suivantes :

1.

$$\forall x, y \in E; 0 \leq p_{xy} \leq 1.$$

2. Pour tout $x \in E$ on a

$$\sum_{y \in E} p_{xy} = 1$$

3. P admet la valeur propre 1, le vecteur $e = [1, \dots, 1]'$ étant un vecteur propre associé.

Démonstration. Soit $P = [p_{xy}]_{x_0 \leq x, y \leq x_{n+1}}$ une matrice de transition.

1. Pour tout couple (x, y) , la quantité p_{xy} est une probabilité, donc elle est bien comprise entre 0 et 1.

2. Pour tout $x \in E$:

$$\begin{aligned} \sum_{y \in E} p_{xy} &= \frac{1}{P(X_n = x)} \sum_{y \in E} P(X_{n+1} = y, X_n = x) \\ &= \frac{P(X_{n+1} \in E, X_n = x)}{P(X_n = x)} = \frac{P(X_n = x)}{P(X_n = x)} = 1 \end{aligned}$$

3. On a

$$P \times e = \begin{bmatrix} \sum_{i=x_0}^{x_{n+1}} p_{x_0 i} \\ \cdot \\ \cdot \\ \cdot \\ \sum_{i=x_0}^{x_{n+1}} p_{x_{n+1} i} \end{bmatrix} \stackrel{2}{=} \begin{bmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}$$

□

Définition 2.4. Soit π une probabilité sur E si ;

$$x, y \in E, \pi(x)p(x, y) = \pi(y)p(y, x)$$

alors π est dite *symétrique* (par rapport à p) ou *p-symétrique*.

Lemme 2.1. Si π est *p-symétrique*, alors elle est *p-invariante*.

Démonstration. Pour tout $y \in E$

$$\begin{aligned} (\pi p)(y) &= \sum_{x \in E} \pi(x)p(x, y) \\ &= \sum_{x \in E} \pi(y)p(y, x) \\ &= \pi(y) \end{aligned}$$

□

Exemple 2.2.1. Soit la chaîne de Markov ci-dessous qui représente la météo

$$P(\text{beau demain}/\text{beau aujourd'hui}) = 0.8.$$

$$P(\text{beau demain}/\text{pluie aujourd'hui}) = 0.6.$$

$$X_n = \begin{cases} 0 & \text{si beau le jour } n \\ 1 & \text{si pluie le jour } n \end{cases}$$

$$p_{00} = P(X_{n+1} = 0 | X_n = 0) = 0.8$$

$$p_{10} = P(X_{n+1} = 0 | X_n = 1) = 0.6$$

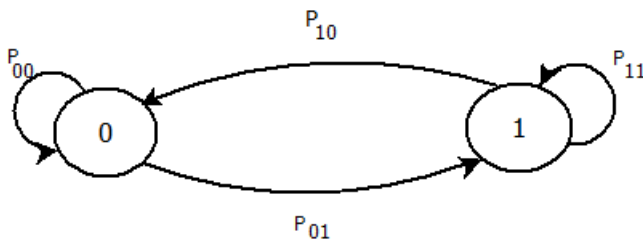
$$p_{00} + p_{01} = 1 \iff p_{01} = 1 - p_{00} = 1 - 0.8 = 0.2$$

$$p_{10} + p_{11} = 1 \iff p_{11} = 1 - p_{10} = 1 - 0.6 = 0.4$$

Donc la matrice de transition est :

$$\begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix}$$

le graphe de transition associe à la chaîne de Markov de cet exemple,



2.2.1 Classification d'état

Définition 2.5. (*État communiquent*) Soient $x, y \in E$, $f_{xy}^{(k)}$ est la probabilité d'aller de x à y en k étapes :

$$f_{xy}^{(k)} = P(X_n = x, X_{n+1} \neq y, \dots, X_{n+k-1} \neq y, X_{n+k} = y)$$

où $f_{xy} = \sum_{k \in \mathbb{N}} f_{xy}^{(k)}$.

Si $f_{xy} > 0$ on dit que l'état x mène à y ; on note $y \leftarrow x$. De plus si $x \leftarrow y$ et $y \leftarrow x$, on dit que x et y communiquent; on note $x \leftrightarrow y$.

La période de $x \in E$ est le PGCD de l'ensemble $M_x = \{n \geq 0 | P(X_n = x | X_0 = x) > 0\}$. Si $x \leftrightarrow y$; alors, x et y ont la même période, si la période est égal a 1, la classe est dite apériodique.

Définition 2.6. (*Chaîne de Markov irréductible*) La relation communiquer \leftrightarrow est une relation d'équivalence, quand on parle de classe, en parlant des états d'une chaîne de Markov, si tous les états communiquent, la chaîne de Markov est dite irréductible.

Définition 2.7. (*État récurrent et transient*) Soit T_x le temps de premier passage à x , on note $T_x^{(r)}$ par le temps de r -ième retour à x , définie par récurrence par

$$T_x^{(0)} = T_x, \quad T_x^{(r+1)} = \inf_n \{n > T_x^{(r)} : X_n = x\}$$

On dit que x est récurrent si la probabilité de retour égal a 1 ;

$$P(T_x^{(1)} < \infty) = 1$$

Si la propriété n'est pas vérifiée, on dit que l'état est transient.

Définition 2.8. (*Chaîne de Markov absorbante*) Soit l'état $x \in E$, on dit que x est absorbant si

$$p_{xx} = 1 ; \quad \forall j \neq x : p_{xy} = 0$$

Une chaîne de Markov est dit absorbante si pour tout état de E , il existe un état absorbant qui communique avec cet état.

Définition 2.9. (*Probabilité invariante*) une probabilité π sur E est dite invariante ou stationnaire pour une probabilité de transition p si $\forall x \in E; \pi(x) = \sum_{y \in E} \pi(y)p(x, y)$.

2.2.2 Théorème ergodique

Le théorème ergodique est une généralisation de la loi de grands nombres,

Théorème 2.1. Soit P la matrice de transition irréductible sur l'espace E , et soit π la probabilité invariante pour P , si on a

1. π est la seule probabilité invariante pour tout $x \in E$, de plus $\pi(x) > 0$.
2. $\forall x \in E$ x est récurrent.
3. $\forall x \in E$ et $\forall f : E \rightarrow \mathbb{R}$ telle que $\sum_{x \in E} |f(x)|\pi(x) < \infty$

alors,

$$\frac{1}{n} \sum_{k=1}^n f(X_k^x) \xrightarrow{n} \sum_{x \in E} f(x) \pi(x)$$

où $(X_n^x)_{n \in \mathbb{N}}$ est une chaîne de Markov d'état initiale x , et de matrice de transition P .

2.3 Méthodes de Monte Carlo

Les méthodes de Monte Carlo ont été créées en 1947 par N. Metropolis et publiées la première fois dans un article en 1949.

On réfère par le terme méthode de Monte-Carlo toute méthode visant à résoudre des problèmes centrés sur le calcul d'une valeur numérique en utilisant des procédures aléatoires, tout simplement, on utilise le hasard pour calculer une valeur numérique. La méthode de Monte-Carlo repose essentiellement sur les principes fondamentaux de la loi des grands nombres et du Théorème Central Limite.

Théorème 2.2. (*Loi forte de grands nombres*)

Si $(X_n)_{n>0}$ est une suite de variables aléatoires indépendantes identiquement distribuées telles que $E(|X_1|) < \infty$ alors

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{} E(X_1)$$

Théorème 2.3. (*Théorème Centrale Limite*)

Si $(X_n)_{n>0}$ est une suite de variables aléatoires indépendantes identiquement distribuées, supposons que l'espérance μ et l'écart-type σ existent et soient finis avec $\sigma \neq 0$, alors

$$\frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{i=1}^n X_i - \mu \right) \longrightarrow \mathcal{N}(0,1) \text{ (loi normale centrée réduite)}$$

Description de la méthode Monte Carlo

Si notre objectif est de trouver la quantité I , on pose que cette valeur est une espérance d'une variable aléatoire X , i.e. $I = E(X)$, si la simulation des variables X_1, X_2, \dots indépendantes identiquement distribuées (iid) selon la loi de X est possible, alors on approxime I par

$$\frac{X_1 + X_2 + \dots + X_N}{N} = \frac{1}{N} \sum_{i=1}^N X_i$$

Avec N grand sous réserve d'application de la loi des grands nombres. Ce type d'approximation est qu'on appelle "méthodes de Monte Carlo".

Remarque 2.1. *L'estimateur de Monte Carlo est un estimateur sans biais.*

Démonstration.

$$E\left(\frac{1}{N} \sum_{i=1}^N X_i\right) = \frac{1}{N} E\left(\sum_{i=1}^N X_i\right) = E(X)$$

□

2.3.1 Application sur les intégrales

Si l'on veut calculer l'intégrale

$$I = \int_{[0,1]^d} f(u_1, \dots, u_d) du_1, \dots, du_d$$

On pose que $X = f(U_1, \dots, U_d)$ où les variables U_1, \dots, U_d sont indépendants et suivent une loi uniforme sur $[0, 1]$, donc

$$I = E(f(U_1, \dots, U_d)) = E(X)$$

Si la simulation des variables (iid) de lois uniforme sur $[0, 1]$ $U_1^{(1)}, \dots, U_d^{(1)}, U_1^{(2)}, \dots, U_d^{(2)}, \dots$ est possible, alors, les variables $X_1 = f(U_1^{(1)}, \dots, U_d^{(1)})$, $X_2 = f(U_1^{(2)}, \dots, U_d^{(2)})$, ... sont indépendantes identiquement distribuées (iid),

donc on peut approximer I par

$$\frac{X_1 + X_2 + \dots + X_N}{N}.$$

Remarque 2.2. *Le type de la fonction à intégrer n'a rien à voir avec la qualité de l'approximation.*

Si l'intégrale I est de la forme

$$I = \int_{\mathbb{R}^d} g(x_1, \dots, x_d) f(x_1, \dots, x_d) dx_1, \dots, dx_d$$

avec f positive et $\int_{\mathbb{R}^d} f(x_1, \dots, x_d) dx_1, \dots, dx_d = 1$, on considère la fonction f la loi de la variable X à valeur dans \mathbb{R}^d , donc $I = E(g(X))$.

Alors si X_i est un échantillon tiré selon la loi f l'approximation de I est

$$\frac{1}{N} \sum_{i=1}^N g(X_i)$$

On remarque que toute simulation de Monte Carlo fait compte des nombres au hasard. A cet regard, On admet la disposition d'un générateur de nombres aléatoires qui est capable de simuler une suite de nombres au hasard.

2.3.2 Génération de nombres aléatoires

On appelle un générateur de nombres aléatoires la procédure qui permet de produire une séquence de nombres, de façon que cette séquence puisse être appelée « suite de nombres aléatoires », dont on ne peut pas facilement avoir des propriétés déterministes.

Il y a plusieurs méthodes pour simuler ou générer des variables aléatoires, on mentionne les méthodes classiques comme l'algorithme d'inversion de la fonction de répartition, l'algorithme de Box-Muller que l'on appelle générateurs pseudo-aléatoires.

2.3.2.1 Générateurs de nombres pseudo-aléatoires

on appelle un algorithme la suite des opérations que l'on applique à des paramètres pour obtenir un certain résultat. Un algorithme est déterministe de sens que l'on applique le même traitement aux mêmes paramètres, on obtient le même résultat. Pourtant quelques opérations sont raisonnablement imprévisibles pour offrir des résultats qui semblent aléatoires. Les nombres obtenus sont donc appelés pseudo-aléatoires.

On utilise tels nombres parce qu'ils sont faciles à produire. Il existe des domaines quand on peut utiliser ces nombres à la place de « vrais » nombres aléatoires. Ceci est possible à condition d'effectuer une étude numérique exacte pour le prouver.

- **Algorithme d'inversion de la fonction de répartition**

soit $f(x)$ la loi de probabilité avec la fonction de répartition $F(x) = \int_{-\infty}^x f(t)dt$, pour générer une distribution X selon cette loi, il suffit de calculer l'inverse de la fonction de répartition de la variable aléatoire U , qui suit une loi uniforme sur $[0, 1]$ i.e.

$$X = F^{-1}(U)$$

Alors F est de la fonction de répartition de la variable aléatoire X .

- **Algorithme de Box-Muller**

Si les variables aléatoires U et V indépendantes suivent la loi uniforme sur $[0, 1]$, alors

$$X = \sqrt{-2\log U} \cos(2\pi V)$$

$$Y = \sqrt{-2\log U} \sin(2\pi V)$$

sont des variables aléatoires indépendantes qui suivent la loi normale centrée réduite.

2.3.2.2 Les tests d'adéquation

On va essayer de vérifier si les séquences obtenues avec un générateur peuvent être considérées comme aléatoires, afin de dire si le générateur est lui-même aléatoire ou non, autrement dit on vérifie que ces séquences possèdent de différentes propriétés du hasard. Cependant, aucun générateur ne peut réussir tous les tests que l'on pense constitutif du hasard, car il n'existe aucune suite qui possède toutes ces propriétés.

Les tests d'adéquation servent à tester si une suite des nombres aléatoires est distribuée selon une loi de probabilité.

En présence d'une liste de résultats pour une variable aléatoire, on peut utiliser des tests non paramétriques pour comparer la répartition obtenue à partir des données avec une répartition connue ou une répartition théorique. Les deux tests d'adéquation les plus utilisés dans ce cas sont le test d'ajustement du Khi^2 et le test de Kolmogorov-Smirnov, il y a aussi d'autres tests qui font le même travail.

Ces tests calculent, à partir des écarts entre les valeurs réelles et les valeurs théoriques, une valeur que l'on compare à un seuil critique dans la table statistique correspondante.

Applications

L'utilisation des générateurs est fréquente dans plusieurs domaines et le nombre de leurs applications joue un rôle majeur en physique dans les domaines de la simulation. ils permettent aussi de calculer des intégrales, des valeurs approchées...etc.

1. Simulation

Les nombres aléatoires sont nécessaires pour simuler un phénomène physique, une expérience, la conduite, le pilotage ou n'importe quel jeu.

2. Analyse

A l'aide d'un échantillonnage bien choisi, aléatoire par exemple, on peut simplifier les analyses que l'on veut effectuer ; le calcul des valeurs numériques. Comme une intégrale en dimension supérieure à 1 ou une solution d'équation différentielle.

3. Sécurité informatique

Les applications sont là aussi nombreuses, notamment dans des procédures de tests, et autres méthodes afin de concéder de percer les failles d'un système informatique structuré. Réseau, logiciel...

Remarque 2.3. *La généralisation d'un variable aléatoire par les méthodes classiques qu'on a vu précédemment est facile dans la plupart des cas, mais dans d'autres cas la distribution de probabilité est plus compliquée et n'est pas facilement explicite. On va alors utiliser des méthodes plus sophistiquées pour générer des variables aléatoires selon cette distribution, (méthodes de Monte Carlo par chaînes de Markov).*

2.4 Méthodes de Monte Carlo par chaînes de Markov

Les méthodes de Monte Carlo par chaînes de Markov ou méthodes MCMC sont beaucoup utilisées dans la simulation ; dans le but de simuler des variables d'une distribution f ; une chaîne de Markov irréductible et apériodique dont la distribution stationnaire est f . Lorsque cette chaîne est produite, on obtient ainsi un échantillon de variables distribuées selon f , on peut ainsi utiliser l'estimateur de Monte Carlo pour approximer l'espérance mathématique.

2.4.1 Une méthode MCMC : Algorithme de Metropolis Hastings

La méthode de Metropolis-Hastings est historiquement la première des méthodes MCMC, elle fut d'abord publiée sous une première forme par Metropolis et al. (1953), puis généralisée par Hastings (1970).

Elle se fonde sur le choix d'une distribution de transition instrumentale conditionnelle, Si on peut trouver cette distribution pour laquelle la probabilité π est invariante, et facile à simuler. Cependant, les comportements pratiques et notamment la rapidité d'atteinte de l'état limite ergodique doivent être considérés avec attention.

L'algorithme de Metropolis-Hastings (MH) peut être présenté comme une généralisation de l'algorithme d'acceptation-rejet.

2.4.1.1 Description de la méthode

Soit π une distribution avec un grand nombre de dimensions, pour générer une suite de vecteurs suivant cette distribution, où dont on voudrait calculer un intégrale de la forme $\int_E f(x)\pi(x)dx$, la distribution π est appelée la loi cible.

On part d'une valeur x_0 choisit arbitrairement, à l'aide d'une loi instrumentale conditionnelle $p(y|x) = p(x, y)$ (la probabilité de transition) en effectue des sauts à partir du point initial x_0 .

Nous appelons p le noyau de proposition. On va construire une chaîne de Markov $(X_n)_{n \geq 0}$.

1. On prend $X_0 = x_0$ tel que $\pi(x_0) > 0$.
2. À l'étape $n + 1$ $X_n = x$ (on connaît X_n) On réalise un tirage aléatoire dans une loi uniforme et on note par $U_n \sim U([0, 1])$ et on génère le candidat Y_{n+1} qui s'appelle une proposition $Y_{n+1} \sim p(x, \cdot) = p(\cdot|x)$. On note que les variables U_{n+1}, Y_{n+1} sont indépendants (de plus indépendants des simulations passées).
3. décider :

$$X_{n+1} = \begin{cases} Y_{n+1} & \text{on conserve le candidat avec probabilité } \rho \\ X_n = x & \text{on reconduit la valeur précédente avec probabilité } 1 - \rho \end{cases}$$

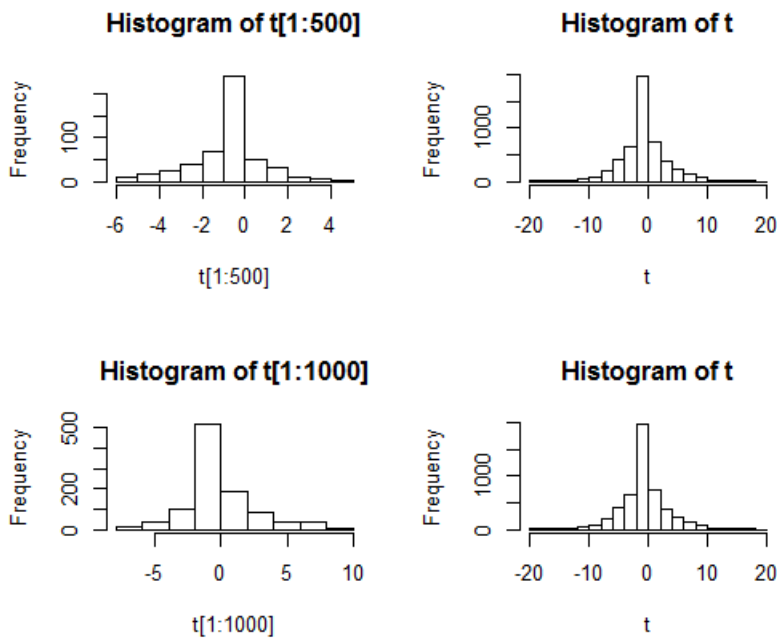
avec

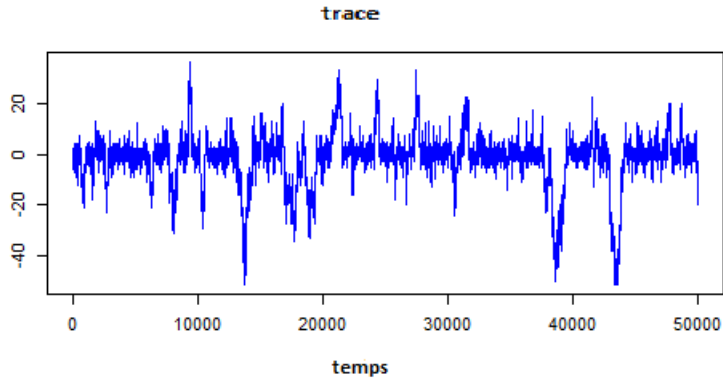
$$\rho = \rho(x, Y_{n+1}) = \min \left(\frac{\pi(Y_{n+1})p(x|Y_{n+1})}{\pi(x)p(Y_{n+1}|x)}, 1 \right)$$

4. Si $\rho(x, Y_{n+1}) \geq U_{n+1}$ on dit qu'on accepte la proposition, si non, on dit qu'on refuse la proposition.

Remarque 2.4. Dans le cas où p est symétrique, pour tout x, y ρ se simplifie en $\rho = \min \left(1, \frac{\pi(y)}{\pi(x)} \right)$.

Exemple 2.4.1. Les figures suivantes représentent un algorithme de Metropolis dans le cas où p est symétrique, et pour tout x sur E la loi cible est $\pi = C \exp(-\sqrt{|x|})$ (C est une constante telle que $\sum_{x \in E} \pi(x) = 1$).





Proposition 2.2. *La suite aléatoire $(X_n)_{n \geq 0}$ obtenue dans la partie précédente est une chaîne de Markov où sa matrice de transition M est donnée par :*

$$\begin{cases} M(x, y) = p(x, y)\rho(x, y) & \text{si } x \neq y \\ M(x, x) = 1 - \sum_{y \in E} M(x, y) & \text{sinon} \end{cases}$$

De plus la loi π est M -invariante.

Démonstration. On a pour tout $x \in E$, $P(X_{n+1}|X_0, X_1, \dots, X_n) = P(X_{n+1}|X_n)$, donc la suite $(X_n)_{n \geq 0}$ est bien une chaîne de Markov.

Soit $x \neq y$

$$\begin{aligned} P(X_{n+1} = y | X_n = x) &= P(Y_{n+1} = y, U_{n+1} \leq \rho(X_n, Y_{n+1}) | X_n = x) \\ &= P(Y_{n+1} = y, U_{n+1} \leq \rho(x, y) | X_n = x) \\ &= p(x, y)\rho(x, y) \end{aligned}$$

Il est évident pour $x = y$.

$$\begin{aligned}
\pi(x)M(x, y) &= \pi(x)p(x, y)\rho(x, y) \\
&= \pi(x)p(x, y) \min\left(1, \frac{\pi(y)p(y, x)}{\pi(x)p(x, y)}\right) \\
&= \min(\pi(x)p(x, y), \pi(y)p(y, x)) \\
&= \min(\pi(y)p(y, x), \pi(x)p(x, y)) \\
&= \pi(y)M(y, x)
\end{aligned}$$

Donc π est symétrique par rapport à M , D'après le lemme 1.1 ; on a $\pi M = \pi$. □

Lemme 2.2. *Lorsque $p(x, y)$ est irréductible tel que, for all $x, y \in E; p(x, y) \neq 0, p(y, x) \neq 0$, et $\pi(x) > 0$ pour tout x, y alors M est irréductible.*

Démonstration. Soient $x, y \in E, x \neq y$;

$$\exists n \in \mathbb{N} \text{ et } (x_0, x_1, \dots, x_n) \in E$$

$$\text{tel que } p(x, x_1)p(x_1, x_2)\dots p(x_n, y) > 0$$

Il est toujours possible de supposer $x \neq x_1, \dots, x_n \neq y$. On déduit que

$$M(x, x_1)M(x_1, x_2)\dots M(x_n, y) > 0$$

□

2.4.1.2 Algorithme Metropolis-Hastings de type marche aléatoire

L'algorithme de type marche aléatoire est un cas particulier de l'algorithme Metropolis-Hastings, pour lequel il faut prendre en compte la valeur présente de la chaîne de Markov à l'effet de générer un candidat pour le prochain intervalle de temps.

Il s'agit de simuler le candidat Y_{n+1}

$$Y_{n+1} = X_n + \epsilon_n$$

où ϵ_n indépendante de X_n est une perturbation aléatoire. Par exemple, $\epsilon_n \sim N(0, \sigma^2)$. La probabilité de transition $p(x, y) = p(y|x) = p(y - x)$.

La chaîne de Markov formée par les candidats $\{Y_n, n \geq 1\}$ est alors une marche aléatoire, par contre à cause de l'étape d'acceptation de l'algorithme Metropolis-Hastings

la chaîne de Markov de $\{X_n, n \geq 0\}$ n'est pas une marche aléatoire.

Si le noyau de proposition est symétrique, on obtient l'algorithme suivant :

Soit $X_n = x$,

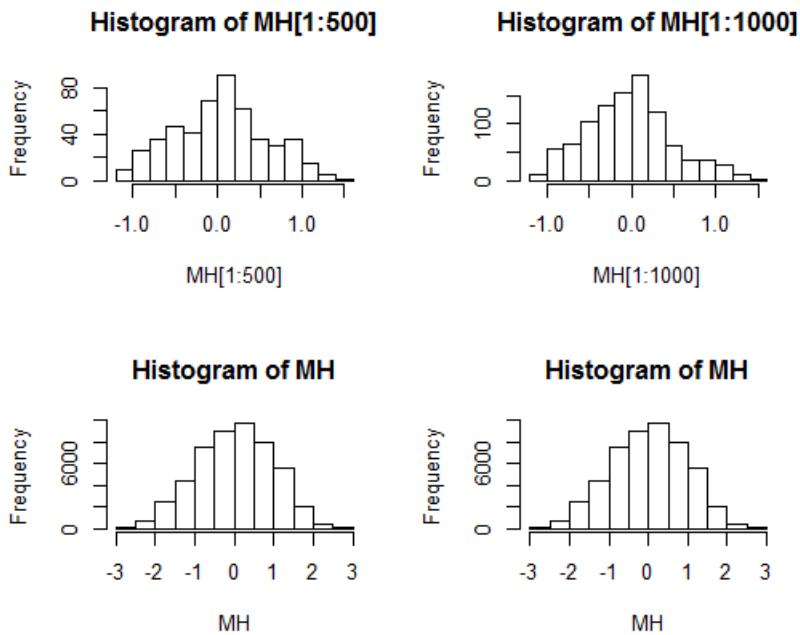
1. Générer une observation $Y_{n+1} = y$ qui suit la loi $p(x, \cdot)$
2. Calculer la probabilité d'acceptation :

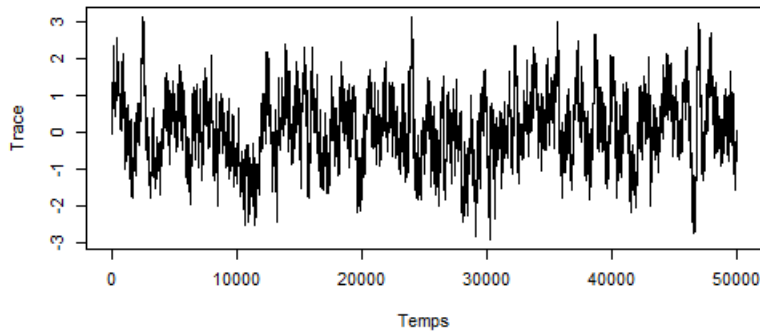
$$\rho(x, y) = \min \left(1, \frac{\pi(y)}{\pi(x)} \right).$$

3. Décider

$$X_{n+1} = \begin{cases} y & \text{avec probabilité } \rho(x, y) \\ x & \text{avec probabilité } 1 - \rho(x, y) \end{cases}$$

Les figures suivantes représentent une algorithme Metropolis-Hastings par marche aléatoire pour un noyau de proposition suit la loi normale $N(x, 0.1)$ et une loi cible, $N(0, 1)$.





2.4.1.3 Algorithme Metropolis-Hastings indépendant

Un cas particulier de l'algorithme Metropolis-Hastings est l'échantillon indépendant dans lequel la probabilité de transition ne dépend pas de l'état présent de la chaîne de Markov, i.e., $p(x, y) = p(y)$. On obtient alors l'algorithme suivant :

Au temps n ; $X_n = x$,

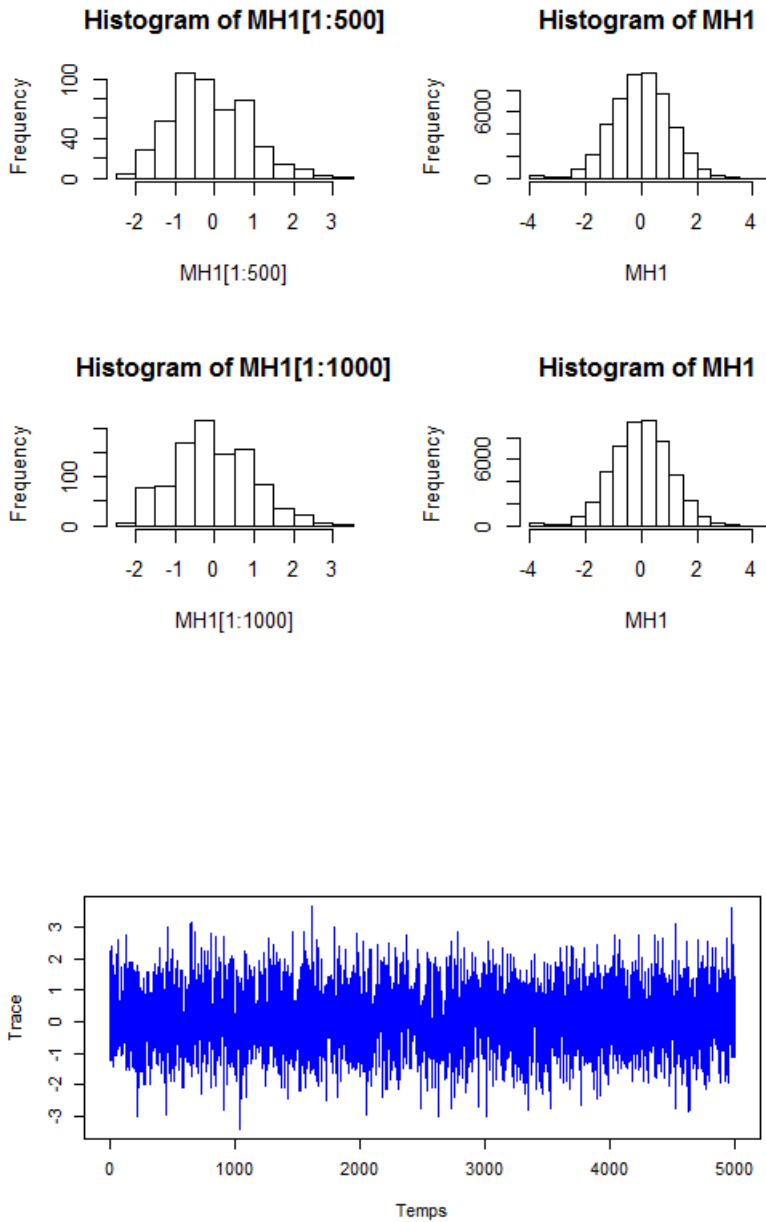
1. Générer une observation $Y_{n+1} = y$ qui suit la loi $p(x, \cdot)$
2. Calculer la probabilité d'acceptation :

$$\rho(x, y) = \min \left(1, \frac{\pi(y)p(x)}{\pi(x)p(y)} \right).$$

3. Décider

$$X_{n+1} = \begin{cases} y & \text{avec probabilité } \rho(x, y) \\ x & \text{avec probabilité } 1 - \rho(x, y) \end{cases}$$

Les figures suivantes représentent une algorithme Metropolis-Hastings par marche aléatoire pour un noyau de proposition suit la loi normale $N(1.2, 1)$ et une loi cible, $N(0, 1)$.



Exemple 2.4.2. *Simulation de données suivant la loi normale $N(0, 1)$.*

On applique Metropolis-Hastings Indépendant avec $Y_{n+1} \sim U[-3, +3]$

*De plus l'algorithme de Metropolis-Hastings Marche Aléatoire avec $\epsilon_n \sim U[-\delta, +\delta]$
(Hastings, 1970)*

La probabilité d'acceptation est

$$\rho = \min\left(1, \frac{\exp(X_n^2 - Y_{n+1}^2)}{2}\right)$$

FIGURE 2.1 – Les lois cibles

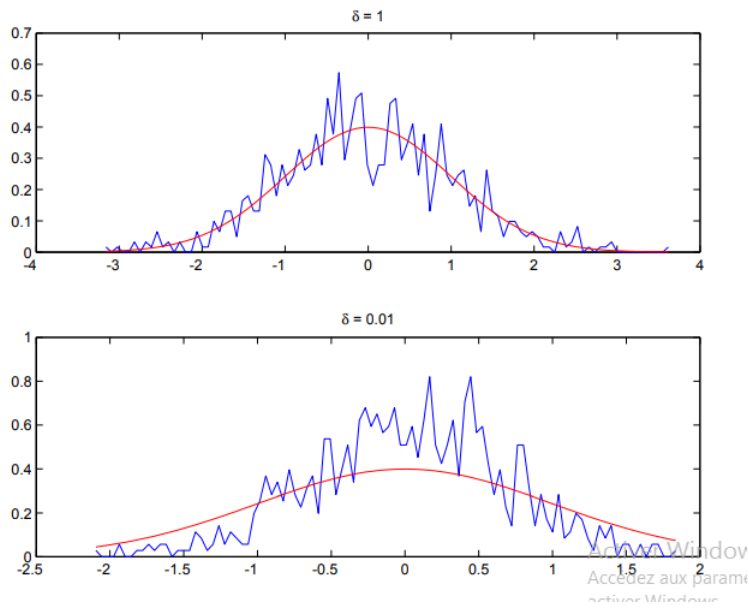


FIGURE 2.2 – Calcul des moyennes

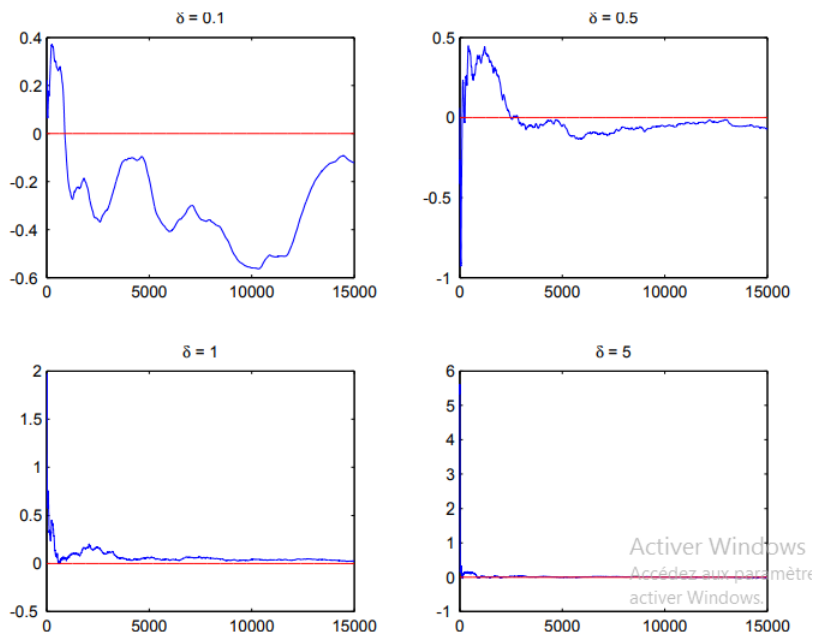
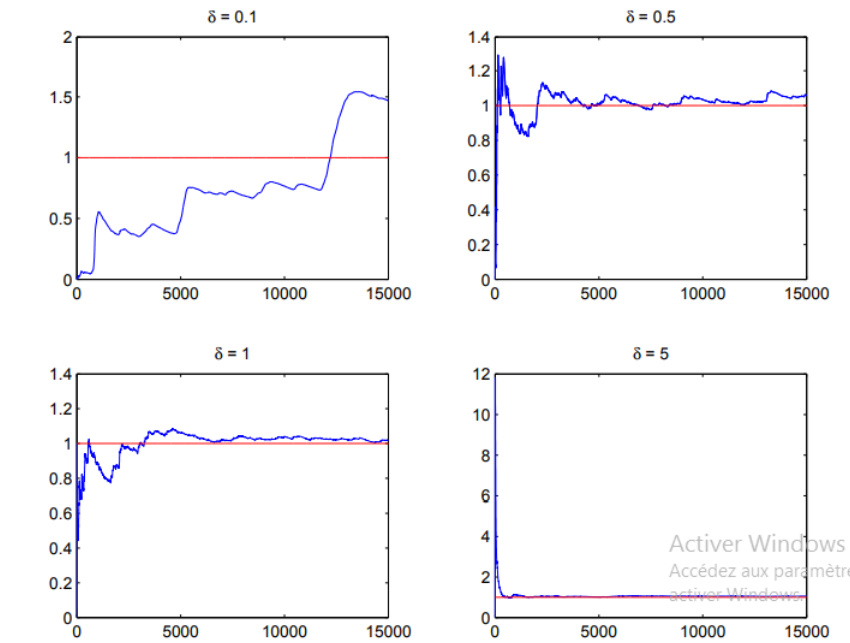


FIGURE 2.3 – Calcul des variances



2.4.1.4 Convergence

On possède maintenant un puissant outil d'inférence Bayésien. Un algorithme basé sur une propriété statistique (l'ergodicité d'une chaîne de Markov) est utilisé pour effectuer un calcul d'inférence Bayésienne (pour étudier une loi à posteriori). Malheureusement, il s'agit d'une méthode asymptotique; la convergence arrive à l'infini, mais combien de temps l'algorithme a-t-il fonctionné légitimement tout au long du temps, la solution de ce problème est de suivre graphiquement l'évolution de l'algorithme comme sur les figures précédentes.

On observe le graphe de série temporelle qui montre les réalisations de la chaîne de Markov à chaque itération par rapport aux nombres d'itérations, on appelle un tel graphe par la trace, cette méthode de diagnostic de convergence graphique est la plus courante, Elle est utilisée pour visualiser comment la chaîne de Markov se déplace dans l'espace d'état, c'est-à-dire comment elle se mélange.

Si la chaîne MCMC est coincée dans une partie de l'espace d'état, les traces montrent des bits plats indiquant une convergence lente. Un tel tracé est observé pour une chaîne MH si trop de propositions sont rejetées consécutivement.

Les tendances visibles ou les changements dans la propagation dans le tracé impliquent

que la stationnarité n'a pas encore été atteinte.

On dit souvent qu'un bon tracé devrait ressembler à une chenille poilue. Pour un algorithme MCMC efficace si la valeur initiale n'est pas dans la région à haute densité, on utilise un échantillon de brûlure (burn in sample). D'autre part, si le tracé montre un modèle similaire tout au long, alors il n'est pas utile de lancer un échantillon de brûlure.

2.4.2 Une méthode MCMC : Algorithme de Gibbs

Une autre méthode MCMC est l'échantillonnage de Gibbs, elle constitue une bonne manière d'échantillonner les distributions conjointes de plusieurs variables, autant que les distributions conditionnelles de la distribution cible soient disponibles, comme il est plus simple de générer des observations à travers les densités conditionnelles que les densités conjointes, l'échantillonnage de Gibbs est basé sur cette approche.

Soit la distribution d'un paramètre

$$\theta = (\theta_1, \dots, \theta_n) \in \Theta$$

, posons $\theta_{-i} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n)$.

On suppose que la simulation de la loi jointe $P(\theta|X)$ est difficile, et pour tout i la simulation de la loi $P(\theta_i|X, \theta_{-i})$ est disponible.

On part d'un point arbitraire $\theta^0 = (\theta_1^{(0)}, \dots, \theta_n^{(0)}) \in \Theta$, on a

$$\theta_1^{(t+1)} \sim P(\theta_1|X, \theta_2 = \theta_2^{(t)}, \dots, \theta_n = \theta_n^{(t)})$$

$$\theta_2^{(t+1)} \sim P(\theta_2|X, \theta_1 = \theta_1^{(t)}, \theta_3 = \theta_3^{(t)}, \dots, \theta_n = \theta_n^{(t)})$$

.

.

.

$$\theta_n^{(t+1)} \sim P(\theta_n|X, \theta_1 = \theta_1^{(t)}, \dots, \theta_{n-1} = \theta_{n-1}^{(t)})$$

2.4.2.1 Acceptation de l'échantillonnage Gibbs

Dans cette section, nous présentons comment l'échantillonnage de Gibbs est un cas particulier de l'algorithme de Metropolis-Hastings ; L'algorithme d'échantillonnage de Gibbs est la composition de p algorithmes de Metropolis Hastings de probabilité d'acceptation uniformément égale à 1.

La distribution de la proposition d'échantillonnage de Gibbs est donnée comme suit :

$$Q(\theta'_i, \theta_{-i} | \theta_i, \theta_{-i}) = P(\theta'_i | \theta_{-i})$$

L'application de cette proposition à l'algorithme de Metropolis-Hastings pour calculer les probabilités d'acceptation :

$$\begin{aligned} A(\theta'_i, \theta_{-i} | \theta_i, \theta_{-i}) &= \min \left(\frac{P((\theta'_i | \theta_{-i}) Q(\theta_i, \theta_{-i} | \theta'_i, \theta_{-i}))}{P((\theta_i | \theta_{-i}) Q(\theta'_i, \theta_{-i} | \theta_i, \theta_{-i}))}, 1 \right) \\ &= \min \left(\frac{P((\theta'_i | \theta_{-i}) P(\theta_i | \theta_{-i}))}{P((\theta_i | \theta_{-i}) P(\theta'_i | \theta_{-i}))}, 1 \right) \\ &= \min(1, 1) = 1 \end{aligned}$$

Chapitre No. 3

L'estimation Bayésienne en présence de censure : Modèle de Bertholon et Modèle Tronqué

Dans ce chapitre, on propose une analyse Bayésienne de distribution de Bertholon et la nouvelle distribution de Zeghdoudi tronquée à droite. Nous introduisons d'abord les distributions. Ensuite, nous dérivons les estimateurs du maximum de vraisemblance (MV) des paramètres inconnus, on utilise un échantillon de données censurées de type II. De plus, nous dérivons les estimateurs Bayésiens de ces paramètres sous la quadratique généralisée (GQ), l'entropie et les fonctions de perte de Linex. On effectue une expérience de simulation pour étudier le comportement des estimateurs proposés et les comparer avec l'estimateur du maximum de vraisemblance, en utilisant le critère de proximité de Pitman. Enfin, nous calculons l'erreur quadratique moyenne intégrée (IMSE).

3.1 L'estimation de la distribution de Bertholon avec des données censurées à droite

La distribution exponentielle $\varepsilon(\eta)$. avec la moyenne η est souvent utilisée pour modéliser les temps d'échec causés par des accidents ;car elle est l'unique distribution à taux de panne constant La fonction de survie de la distribution exponentielle est

$$S_E(t) = \exp\left(-\frac{t}{\eta}\right),$$

où le paramètre d'échelle η est l'inverse du taux de hasard constant λ .

La distribution de Weibull $\mathcal{W}(\eta, \beta)$, à la fonction de survie

$$S_W(t) = \exp\left[-\left(\frac{t}{\eta}\right)^\beta\right]$$

et le taux de hasard

$$h_W(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}$$

Cette distribution peut être utilisée pour modéliser les défauts de mortalité infantile lorsque le paramètre de forme $\beta < 1$ ou le vieillissement lorsque $\beta > 1$. Notez que lorsque $\beta = 1$, la distribution de Weibull se réduit à une distribution exponentielle.

Ainsi, une façon plus réaliste de modéliser les temps d'échecs est de considérer un modèle à risques concurrents qui tient compte du fait qu'une défaillance peut être causée par le vieillissement ou par accident. Ce modèle est défini comme suit ;

Un temps d'échec est la réalisation de la variable aléatoire $B = \min(E, W)$ où E suit la distribution exponentielle $E(\eta_0)$ et W suit la distribution Weibull $W(\eta_1, \beta)$ avec $\beta > 1$, tel que la variable aléatoire E et W soient indépendantes. Par conséquent, la distribution de B est caractérisé par les paramètres η_0, η_1 et β , on le note par $\mathcal{B}(\eta_0, \eta_1, \beta)$.

L'utilisation des modèles à risque concurrents est bien connue et une grande famille de distributions a notamment été présentée dans Park et Padgett (2004)[36].

3.1.1 La distribution Bertholon \mathcal{B}

On considère la variable aléatoire $B = \min(E, W)$; où la variable E a une distribution exponentielle avec la moyenne η_0 , la variable. W a une distribution Weibull avec le paramètre d'échelle η_1 et le paramètre de forme β , où E et W sont indépendants. Les principales caractéristiques de la distribution de probabilité \mathcal{B} sont les suivantes.

Le taux de hasard est

$$h_B(x) = \frac{1}{\eta_0} + \frac{\beta}{\eta_1} \left(\frac{x}{\eta_1}\right)^{\beta-1} \quad (3.1)$$

La fonction de survie est

$$S_B(x) = \exp\left[-\frac{x}{\eta_0} - \left(\frac{x}{\eta_1}\right)^\beta\right] \quad (3.2)$$

La densité de probabilité est donnée par

$$f_B(x) = \left(\frac{1}{\eta_0} + \frac{\beta}{\eta_1} \left(\frac{x}{\eta_1} \right)^{\beta-1} \right) \exp \left[-\frac{x}{\eta_0} - \left(\frac{x}{\eta_1} \right)^\beta \right] \quad (3.3)$$

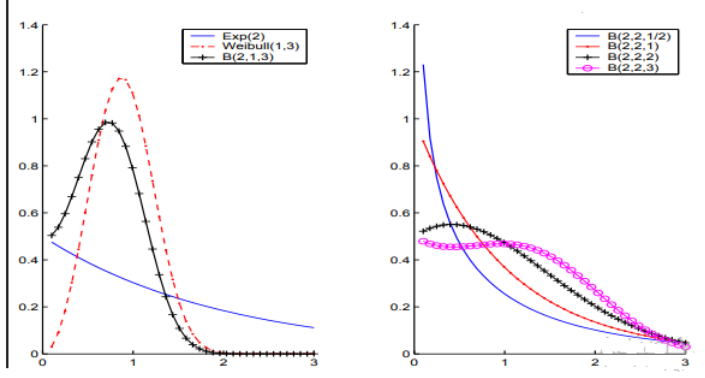


FIGURE 3.1 – Différents exemples de la densité de Bertholon

3.1.2 Estimation des paramètres par le maximum de vraisemblance

On considère un n échantillon (X_1, X_2, \dots, X_n) généré à partir de la distribution \mathcal{B} avec la densité (3.3), En supposant que les données sont censurées à droite,

En remplaçant par (3.3) et la fonction de survie (3.2), la vraisemblance est :

$$L(\eta_0, \eta_1, \beta | X) = \prod_{i=1}^n \left(\frac{1}{\eta_0} + \frac{\beta}{\eta_1} \left(\frac{x_i}{\eta_1} \right)^{\beta-1} \right)^{\delta_i} \exp \left[-\frac{\sum_{i=1}^n x_i}{\eta_0} - \sum_{i=1}^n \left(\frac{x_i}{\eta_1} \right)^\beta \right] \quad (3.4)$$

Étant donné que la variable aléatoire B est le résultat d'une concurrence entre E et W , alors, les observations sont des réalisations de l'une de ces deux variables aléatoires, il est souvent difficile de savoir à l'avance si une observation est une réalisation de E ou de W . Cela indique que la maximisation directe de la fonction de vraisemblance sera numériquement très instable. alors, on utilise l'algorithme EM pour maximiser la vraisemblance.

(cf. Dempster et al. (1977), Bousquet et al. (2006)[19, 14].

On procède comme suit. Définissant $z_i = (z_i^E, z_i^W)$ où $z_i^E = 1$ et $z_i^W = 0$ indiquent que l'observation associée provient d'un modèle exponentiel, et $z_i^W = 1$ et $z_i^E = 0$ provient de la distribution de Weibull. Par convention, les données complètes peuvent alors être écrites comme $y = (y_i = (x_i, z_i), i = 1, \dots, n) = (x, z)$. Ainsi, la densité résultante peut être écrite comme suit :

$$f(y_i) = h_E(x_i)^{z_i^E} h_W(x_i)^{z_i^W} S_E(x_i) S_W(x_i) \quad (3.5)$$

et le logarithme de la vraisemblance basée sur des données complètes $y = (y_1, \dots, y_n)$ est donnée par

$$\begin{aligned} l(\eta_0, \eta_1, \beta | y) &= \sum_{i=1}^n z_i^E \log(h_E(x_i)) + z_i^W \log(h_W(x_i)) \\ &\quad + \log(S_E(x_i)) + \log(S_W(x_i)). \end{aligned} \quad (3.6)$$

On pose $\Theta = (\eta_0, \eta_1, \beta)$ et soit $\tilde{\Theta}$ indique sa valeur actuelle. La valeur attendue de log-vraisemblance $Q(\Theta | \tilde{\Theta})$ est comme suit

$$\begin{aligned} Q(\Theta | \tilde{\Theta}) &= E(l(\eta_0, \eta_1, \beta | o) | x, \tilde{\Theta}) \\ &= \sum_{i=1}^n \tilde{p}_E(x_i) \log(h_E(x_i)) + \tilde{p}_W(x_i) \log(h_W(x_i)) \end{aligned} \quad (3.7)$$

$$+ \log(S_E(x_i)) + \log(S_W(x_i)). \quad (3.8)$$

où

$$\tilde{p}_E(x_i) = E(\tilde{p}_E(x_i | x, \tilde{\Theta}) = P(z_i^E = 1 | x, \tilde{\Theta}) = \frac{h_E(x_i)}{h_E(x_i) + h_W(x_i)}$$

$$\text{et } \tilde{p}_W(x_i) = 1 - \tilde{p}_E(x_i).$$

Ici, $\tilde{p}_E(x_i)$; ($\tilde{p}_W(x_i)$) désigne la probabilité que l'observation provien de la distribution exponentielle (Weibull). De plus, l'équation 3.7 a une structure additive qui résulte de la contribution des distributions exponentielles et de Weibull. Cette décomposition additive de 3.7 facilite la mise en œuvre de l'étape M, en ce sens qu'elle maximise séparément les termes correspondant aux distributions exponentielles et de Weibull. Le terme exponentiel

peut être maximisé par différenciation directe par rapport au paramètre η_0 , tandis que le terme de Weibull peut être maximisé en utilisant l'une des procédures itératives telles que la méthode de Newton Raphson (voir Mann et al.)[30], car il n'y a pas de forme des dérivés par rapport aux paramètres de Weibull η_1 et β , ces deux étapes peuvent être répétées jusqu'à ce que l'algorithme d'itération converge pour donner les estimations du maximum de vraisemblance souhaitées.

3.1.3 Estimation Bayésienne des paramètres sous différentes fonctions de perte

On suppose que les lois a priori sur η_0, η_1 sont des lois Gamma, $\eta_0 \sim G(a_1, b_1)$, $\eta_1 \sim G(a_2, b_2)$

$$\begin{aligned}\pi(\eta_0) &= \frac{a_1^{b_1}}{\Gamma(b_1)} \eta_0^{b_1-1} \exp[-a_1 \eta_0] \\ \pi(\eta_1) &= \frac{a_2^{b_2}}{\Gamma(b_2)} \eta_1^{b_2-1} \exp[-a_2 \eta_1],\end{aligned}$$

Alors que le paramètre β suit une distribution uniforme, $\beta \sim U(\beta_l, \beta_r)$,

$$\pi(\beta) = \frac{1}{\beta_r - \beta_l}, \quad \beta_r \leq \beta \leq \beta_l$$

De plus η_0, η_1 et β sont indépendants, ainsi, la distribution a priori conjointe de η_0, η_1 , et β est de suite

$$\pi(\eta_0, \eta_1, \beta) = \frac{a_1^{b_1} a_2^{b_2}}{\Gamma(b_1) \Gamma(b_2)} \eta_0^{b_1-1} \eta_1^{b_2-1} \exp[-a_1 \eta_0 - a_2 \eta_1] \frac{1}{\beta_r - \beta_l} \quad (3.9)$$

Les distributions gamma ont été choisies en raison de leur flexibilité, offrant un conjugué des distributions à priori. la distribution a posteriori conjointe est donc obtenue par

$$\pi(\eta_0, \eta_1, \beta | X) = \frac{L(\eta_0, \eta_1, \beta | X) \pi(\eta_0, \eta_1, \beta)}{\int \int \int_0^{+\infty} L(\eta_0, \eta_1, \beta | X) \pi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta}$$

alors, la distribution a posteriori conjointe de (η_0, η_1, β) est

$$\pi(\eta_0, \eta_1, \beta|x) = K \eta_0^{b_1-1} \eta_1^{b_2-1} \exp \left[-\frac{\sum_{i=1}^n x_i}{\eta_0} - \sum_{i=1}^n \left(\frac{x_i}{\eta_1} \right)^\beta - a_1 \eta_0 - a_2 \eta_1 \right] \prod_{i=1}^n \left(\frac{1}{\eta_0} + \frac{\beta}{\eta_1} \left(\frac{x_i}{\eta_1} \right)^{\beta-1} \right)^{\delta_i} \quad (3.10)$$

où K est un constante de normalisation.

Ensuite, on utilise les trois fonctions de perte, la fonction de perte quadratique généralisée (GQ), la la fonction de perte Linex et la fonction de perte entropie définies dans 1.1,

pour plus de commodité, on note

$$\phi(\eta_0, \eta_1, \beta) = \exp \left[-\frac{\sum_{i=1}^n x_i}{\eta_0} - \sum_{i=1}^n \left(\frac{x_i}{\eta_1} \right)^\beta - a_1 \eta_0 - a_2 \eta_1 \right] \prod_{i=1}^n \left(\frac{1}{\eta_0} + \frac{\beta}{\eta_1} \left(\frac{x_i}{\eta_1} \right)^{\beta-1} \right)^{\delta_i}$$

Sous la fonction de perte (GQ), les estimateurs Bayésiens de η_0, η_1 et β désignés respectivement par $\hat{\eta}_{0(GQ)}, \hat{\eta}_{1(GQ)}$ et $\hat{\beta}_{(GQ)}$ sont

$$\begin{aligned} \hat{\eta}_{0(GQ)} &= \frac{\int \int \int_0^{+\infty} \eta_0^{\gamma+b_1-1} \eta_1^{b_2-1} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta}{\int \int \int_0^{+\infty} \eta_0^{\gamma+b_1-2} \eta_1^{b_2-1} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta} \\ \hat{\eta}_{1(GQ)} &= \frac{\int \int \int_0^{+\infty} \eta_0^{b_1-1} \eta_1^{\gamma+b_2-1} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta}{\int \int \int_0^{+\infty} \eta_0^{b_1-1} \eta_1^{\gamma+b_2-2} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta} \\ \hat{\beta}_{(GQ)} &= \frac{\int \int \int_0^{+\infty} \beta^\gamma \eta_0^{b_1-1} \eta_1^{b_2-1} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta}{\int \int \int_0^{+\infty} \beta^{\gamma-1} \eta_0^{b_1-1} \eta_1^{b_2-1} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta} \end{aligned}$$

Les risques a posteriori correspondants sont alors,

$$\begin{aligned} PR(\hat{\eta}_{0(GQ)}) &= E_\pi(\eta_0^{\gamma+1}) - 2\hat{\eta}_{0(GQ)} E_\pi(\eta_0^\gamma) + \hat{\eta}_{0(GQ)}^2 E_\pi(\eta_0^{\gamma-1}) \\ PR(\hat{\eta}_{1(GQ)}) &= E_\pi(\eta_1^{\gamma+1}) - 2\hat{\eta}_{1(GQ)} E_\pi(\eta_1^\gamma) + \hat{\eta}_{1(GQ)}^2 E_\pi(\eta_1^{\gamma-1}) \\ PR(\hat{\beta}_{(GQ)}) &= E_\pi(\beta^{\gamma+1}) - 2\hat{\beta}_{(GQ)} E_\pi(\beta^\gamma) + \hat{\beta}_{(GQ)}^2 E_\pi(\beta^{\gamma-1}) \end{aligned}$$

De la même façon, sous la fonction de perte entropie, les estimateurs Bayésiens de η_0 , η_1 et β désignés respectivement par $\hat{\eta}_{0(E)}$, $\hat{\eta}_{1(E)}$ et $\hat{\beta}_{(E)}$ sont

$$\begin{aligned}\hat{\eta}_{0(E)} &= \left[K \int \int \int_0^{+\infty} \eta_0^{b_1-1-p} \eta_1^{b_2-1} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta \right]^{\frac{-1}{p}} \\ \hat{\eta}_{1(E)} &= \left[K \int \int \int_0^{+\infty} \eta_0^{b_1-1} \eta_1^{b_2-1-p} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta \right]^{\frac{-1}{p}} \\ \hat{\beta}_{(E)} &= \left[K \int \int \int_0^{+\infty} \beta^{-p} \eta_0^{b_1-1} \eta_1^{b_2-1} \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta \right]^{\frac{-1}{p}}\end{aligned}$$

Les risques a posteriori correspondants sont alors,

$$PR(\hat{\eta}_{0(E)}) = pE_{\pi}(\log(\eta_0) - \log(\hat{\eta}_{0(E)}))$$

$$PR(\hat{\eta}_{1(E)}) = pE_{\pi}(\log(\eta_1) - \log(\hat{\eta}_{1(E)}))$$

$$PR(\hat{\beta}_{(E)}) = pE_{\pi}(\log(\beta) - \log(\hat{\beta}_{(E)}))$$

Finalement, sous la fonction de perte linéaire, les estimateurs Bayésiens de η_0 , η_1 et β désignés respectivement par $\hat{\eta}_{0(L)}$, $\hat{\eta}_{1(L)}$ et $\hat{\beta}_{(L)}$ sont

$$\begin{aligned}\hat{\eta}_{0(L)} &= \frac{-K}{r} \log \left[\int \int \int_0^{+\infty} \eta_0^{b_1-1} \eta_1^{b_2-1} \exp[-r\eta_0] \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta \right] \\ \hat{\eta}_{1(L)} &= \frac{-K}{r} \log \left[\int \int \int_0^{+\infty} \eta_0^{b_1-1} \eta_1^{b_2-1} \exp[-r\eta_1] \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta \right] \\ \hat{\beta}_{(L)} &= \frac{-K}{r} \log \left[\int \int \int_0^{+\infty} \eta_0^{b_1-1} \eta_1^{b_2-1} \exp[-r\beta] \phi(\eta_0, \eta_1, \beta) d\eta_0 d\eta_1 d\beta \right]\end{aligned}$$

Les risques a posteriori sont

$$PR(\hat{\eta}_{0(L)}) = r(\hat{\eta}_{0(GQ)} - \hat{\eta}_{0(L)})$$

$$PR(\hat{\eta}_{1(L)}) = r(\hat{\eta}_{1(GQ)} - \hat{\eta}_{1(L)})$$

$$PR(\hat{\beta}_{(L)}) = r(\hat{\beta}_{(GQ)} - \hat{\beta}_{(L)})$$

Comme il est difficile d'obtenir des expressions de forme déterministe de tous ces estimateurs, dans la section suivante, on utilise les procédures MCMC pour les évaluer.

3.1.4 Simulation

On considère le modèle $\mathcal{B}(2, 1, 2)$. On procède à la simulation de $N = 10000$ échantillons du modèle censuré à droite avec différentes tailles $n = 10, n = 20$ et $n = 30$. On fixe le taux de censure respectivement à 10% et 20%, On obtient les résultats suivants,

3.1.4.1 Résultats de l'estimation

1. Dans les tableaux suivants, on affiche les valeurs des estimateurs à l'aide de l'algorithme EM pour le modèle \mathcal{B} lorsque 10% et 20% de données sont censurés, où un algorithme de Newton Raphson est appliqué à la distribution de Weibull et la maximisation de la vraisemblance (MV) directe est appliquée à la distribution exponentielle.

TABLE 3.1 – (Estimations du maximum de vraisemblance (MV) avec leurs erreurs Quadratiques (10%).

n	paramètre	MV
10	η_0	1.8904 (0.0003)
	η_1	0.6792 (0.0710)
	β	1.9512 (0.1843)
20	η_0	1.9158 (0.0085)
	η_1	0.7913 (0.1031)
	β	1.9923 (0.0001)
30	η_0	1.9985 (0.0002)
	η_1	0.8181 (0.0011)
	β	1.9491 (0.0005)

TABLE 3.2 – Estimations du maximum de vraisemblance (MV) avec leurs erreurs Quadratiques (20%).

n	paramètre	MV
10	η_0	1.9397 (0.0003)
	η_1	0.6641 (0.0152)
	β	2.0005 (0.0001)
20	η_0	1.9398 (0.0004)
	η_1	0.9013 (0.0020)
	β	1.9491 (0.0001)
30	η_0	1.8014 (0.0031)
	η_1	0.9485 (0.0011)
	β	1.9611 (0.0030)

Discussion : Pour les deux temps de censure, les valeurs estimées des paramètres sont proches des valeurs réelles. De plus, lorsque 10% des données sont censurées, la plus petite erreur quadratique correspond à l'échantillon le plus grand ($n = 30$).

2. Les estimateurs Bayésiens sont obtenus à l'aide des méthodes MCMC. Pour le choix des constants de la distribution à priori, on choisit les informations à priori suivantes. Pour le paramètre de forme β , on suppose que $[\beta_l, \beta_r] = [1, 5]$, pour le paramètre d'échelle η_0 du composant exponentiel, on suppose que $[a_1, b_1] = (1, 300)$, et pour le paramètre d'échelle η_1 du composant Weibull, $[a_2, b_2] = (1, 200)$.

TABLE 3.3 – Estimation Bayésienne sous la perte quadratique généralisée avec le risque a posteriori (entre parenthèses) .

n	.%	par	γ					
			-2	-1	-0.5	0.5	1	2
10	10%	η_0	2.0201 (0.0041)	2.0619 (0.0072)	2.0349 (0.0090)	2.1191 (0.0082)	2.1421 (0.0091)	2.1600 (0.0091)
		η_1	1.1014 (0.0061)	1.1014 (0.0711)	1.1216 (0.0991)	1.1425 (0.4005)	1.1323 (0.2905)	1.1338 (0.3136)
		β	1.9361 (0.0194)	1.9132 (0.0621)	1.8516 (0.8221)	1.7331 (0.0872)	1.7315 (0.0881)	1.8405 (0.1416)
	20%	η_0	2.0717 (0.0049)	2.0191 (0.0051)	2.0301 (0.0059)	2.0333 (0.0071)	2.0509 (0.0071)	2.0991 (0.0079)
		η_1	1.1609 (0.0612)	1.0861 (0.0914)	1.2615 (0.1009)	1.5441 (0.0991)	1.7822 (0.2923)	1.9001 (0.4105)
		β	1.8306 (0.0405)	1.8031 (0.0511)	1.8094 (0.0538)	1.8910 (0.0711)	1.9700 (0.0811)	2.0010 (0.1009)
20	10%	η_0	2.0001 (0.0001)	2.0002 (0.0015)	2.0001 (0.0018)	1.9901 (0.0019)	1.9491 (0.0049)	1.5800 (0.0080)
		η_1	0.6615 (0.0003)	0.6703 (0.0021)	0.5301 (0.0048)	0.5261 (0.0014)	0.6401 (0.0030)	0.7005 (0.0003)
		β	1.9013 (0.0001)	1.9005 (0.0029)	1.9305 (0.0005)	1.8001 (0.0007)	1.6609 (0.0039)	1.6712 (0.0059)
	20%	η_0	2.0201 (0.0041)	2.0482 (0.0070)	2.0561 (0.0073)	2.0531 (0.0061)	2.0677 (0.0069)	2.0823 (0.0072)
		η_1	0.9511 (0.0051)	0.8552 (0.0603)	0.8325 (0.0991)	0.8512 (0.1102)	0.8133 (0.1512)	0.9205 (0.2243)
		β	1.9431 (0.0254)	1.9005 (0.0491)	1.8522 (0.0605)	1.8914 (0.1231)	1.9233 (0.2215)	1.9705 (0.03105)
30	10%	η_0	2.1921 (0.0013)	2.1905 (0.0014)	2.1901 (0.0015)	2.1883 (0.0017)	2.1879 (0.0019)	2.1863 (0.0021)
		η_1	1.1205 (0.0004)	1.1201 (0.0004)	1.1182 (0.0005)	1.1173 (0.0007)	1.1145 (0.0029)	1.1129 (0.0031)
		β	1.9421 (0.0006)	1.9433 (0.0006)	1.9441 (0.0006)	1.9705 (0.0007)	1.9733 (0.0013)	1.9802 (0.0023)
	20%	η_0	2.1031 (0.0013)	2.0894 (0.0016)	2.0972 (0.0014)	2.0345 (0.0016)	2.0372 (0.0017)	2.0382 (0.0018)
		η_1	0.8305 (0.0002)	0.8313 (0.0003)	0.7805 (0.0006)	0.7134 (0.0007)	0.7235 (0.0018)	0.8302 (0.0034)
		β	1.9909 (0.0005)	1.9733 (0.0007)	1.9542 (0.0008)	1.9521 (0.0019)	1.9506 (0.0032)	1.9503 (0.0041)

TABLE 3.4 – Estimation Bayésienne sous la perte entropie avec le risque a posteriori (entre parenthèses)

n	.%	par	p					
			-2	-1	-0.5	0.5	1	2
10	10%	η_0	2.1093 (0.0020)	2.1098 (0.0007)	2.1046 (0.0081)	2.0985 (0.0007)	2.0941 (0.0019)	2.0920 (0.0143)
		η_1	0.8113 (0.1105)	0.7864 (0.0258)	0.7182 (0.0914)	0.6191 (0.0601)	0.5914 (0.3914)	0.5132 (0.0181)
		β	1.9633 (0.0532)	1.9103 (0.0051)	1.8901 (0.0301)	1.6515 (0.0068)	1.7314 (0.0313)	1.6105 (0.1104)
	20%	η_0	2.1001 (0.0051)	2.0931 (0.0009)	2.1032 (0.0066)	2.1013 (0.0006)	2.0909 (0.0040)	2.0891 (0.0021)
		η_1	0.8024 (0.0532)	0.8005 (0.0213)	0.7832 (0.1154)	0.7214 (0.1053)	0.6745 (0.0713)	0.6691 (0.0714)
		β	1.9501 (0.0609)	1.8295 (0.0071)	1.7917 (0.0614)	1.7431 (0.0615)	1.6913 (0.0061)	1.6565 (0.0324)
20	10%	η_0	2.1214 (0.0059)	2.1032 (0.0061)	2.1029 (0.0003)	2.0963 (0.0070)	2.0815 (0.0028)	2.0803 (0.0203)
		η_1	1.0995 (0.1414)	1.1015 (0.0001)	1.1005 (0.1405)	1.1001 (0.0739)	1.0993 (0.0729)	1.0957 (0.1223)
		β	2.1818 (0.0711)	2.1809 (0.0005)	2.1793 (0.0714)	2.1774 (0.0089)	2.1751 (0.2914)	2.1731 (0.1095)
	20%	η_0	2.1781 (0.0008)	2.1763 (0.0003)	2.1751 (0.0004)	2.1743 (0.0004)	2.1731 (0.0004)	2.1725 (0.0013)
		η_1	1.1535 (0.0006)	1.1529 (0.0006)	1.1719 (0.0010)	1.1521 (0.0007)	1.1509 (0.0005)	1.1502 (0.0032)
		β	2.1873 (0.0013)	2.1859 (0.0007)	2.1843 (0.0022)	2.1839 (0.0006)	2.1828 (0.0006)	2.1819 (0.0034)
30	10%	η_0	2.2011 (0.021)	2.2043 (0.0003)	2.2032 (0.0095)	2.2020 (0.0004)	2.2018 (0.0004)	2.2009 (0.0039)
		η_1	1.1738 (0.0007)	1.1729 (0.0003)	1.1715 (0.0008)	1.1709 (0.0006)	1.1707 (0.0006)	1.1702 (0.0020)
		β	2.1123 (0.0017)	2.0832 (0.0003)	2.0821 (0.0020)	2.0819 (0.0004)	2.0814 (0.0004)	2.0809 (0.0020)
	20%	η_0	2.0829 (0.0007)	2.0832 (0.0001)	2.0821 (0.0020)	2.0819 (0.0002)	2.0814 (0.0006)	2.0809 (0.0002)
		η_1	1.0734 (0.0008)	1.0629 (0.0005)	1.0583 (0.0008)	1.0453 (0.0006)	1.0423 (0.0006)	1.0417 (0.0001)
		β	1.9891 (0.0008)	1.9877 (0.0001)	1.9871 (0.0003)	1.9868 (0.0002)	1.9859 (0.0005)	1.9843 (0.0035)

TABLE 3.5 – Estimation Bayésienne sous la perte linex avec le risque a posteriori (entre parenthèses)

n	.%	par	r					
			-2	-1	-0.5	0.5	1	2
10	10%	η_0	2.1333 (0.0182)	2.1329 (0.0110)	2.1284 (0.0009)	2.1252 (0.0021)	2.1134 (0.0091)	2.1147 (0.0183)
		η_1	0.9784 (0.3412)	0.9523 (0.2656)	0.8613 (0.0411)	0.7924 (0.0420)	0.6813 (0.3214)	0.5211 (0.3929)
		β	2.0211 (0.1523)	2.0220 (1.1064)	1.9683 (0.0209)	1.8967 (0.0214)	1.8736 (0.1241)	1.7969 (0.1269)
	20%	η_0	2.0943 (0.0111)	2.0937 (0.0084)	2.0911 (0.0008)	2.0901 (0.0009)	2.0843 (0.2513)	2.0823 (0.3409)
		η_1	0.9651 (0.3518)	0.9117 (0.2614)	0.7643 (0.0021)	0.6853 (0.0527)	0.6067 (0.2729)	0.6095 (0.5114)
		β	1.9963 (0.1563)	1.9695 (0.0953)	1.9018 (0.0214)	1.9001 (0.0245)	1.8761 (0.0305)	1.8569 (0.1902)
20	10%	η_0	2.1426 (0.0113)	2.0805 (0.0071)	2.0793 (0.0008)	2.0754 (0.0008)	2.0721 (0.0021)	2.0711 (0.0043)
		η_1	0.9374 (0.4111)	0.8331 (0.3119)	0.7852 (0.0357)	0.6148 (0.0434)	0.4747 (0.3154)	0.4073 (0.6189)
		β	2.1921 (0.1616)	2.1920 (0.0024)	2.1893 (0.0003)	2.1884 (0.0003)	2.1864 (0.0020)	2.1859 (0.0063)
	20%	η_0	2.1073 (0.0028)	2.1051 (0.0023)	2.1039 (0.0006)	2.1021 (0.0007)	2.1014 (0.0009)	2.1009 (0.0018)
		η_1	1.1451 (0.0055)	1.1443 (0.0052)	1.1421 (0.0004)	1.1419 (0.0005)	1.1417 (0.0003)	1.1405 (0.0002)
		β	2.0213 (0.0028)	2.0205 (0.0032)	2.0202 (0.0005)	2.0183 (0.0006)	2.0167 (0.0006)	2.0166 (0.0020)
30	10%	η_0	2.1847 (0.0030)	2.1833 (0.0350)	2.1821 (0.0002)	2.1819 (0.0003)	2.1814 (0.0053)	2.1809 (0.0061)
		η_1	1.1753 (0.0052)	1.1733 (0.0029)	1.1725 (0.0005)	1.1.1801 (0.0059)	1.1793 (0.0029)	1.1722 (0.0020)
		β	2.1853 (0.0039)	2.1847 (0.0027)	2.1835 (0.0005)	2.1820 (0.0006)	2.1825 (0.0031)	2.1819 (0.0037)
	20%	η_0	2.1653 (0.0021)	2.1647 (0.0020)	2.1635 (0.0004)	2.1641 (0.0004)	2.1629 (0.0016)	2.1617 (0.0017)
		η_1	0.9872 (0.0057)	0.8822 (0.0023)	0.7581 (0.0040)	0.5979 (0.0004)	0.4421 (0.0202)	1.3771 (0.0075)
		β	2.2094 (0.0157)	2.2087 (0.0023)	2.2073 (0.0004)	2.2059 (0.0014)	2.2049 (0.0032)	2.2038 (0.0055)

Discussion : On remarque que la valeur $\gamma = -2$ donne le meilleur risque a posteriori associé a la fonction de perte quadratique généralisée. Dans les deux taux de censure, on obtient également le plus petit risque lorsque n est grande. Lorsque $p = -1$ et $n = 30$ on obtient le meilleur risque à posteriori sous la fonction de perte

entropie. Enfin, sous la fonction de perte Linex, le cas $r = -0,5$ fournit les meilleurs résultats.

Le tableau suivant illustre l'estimateur Bayésien sous les trois fonctions de perte,

TABLE 3.6 – l'estimation Bayésienne sous les trois fonctions de perte

n	.%	par	GQ($\gamma = -2$)	entropy ($p = -1$)	Linex ($r = -0.5$)
10	10%	η_0	2.0201(0.0041)	2.1098 (0.0007)	2.1284(0.0009)
		η_1	1.1014(0.0061)	0.7864(0.0258)	0.8613(0.0411)
		β	1.9361(0.0194)	1.9103(0.0051)	1.9683(0.0209)
	20%	η_0	2.0931(0.0009)	2.0954(0.0008)	2.0911(0.0008)
		η_1	1.1609(0.0612)	0.8005(0.0213)	0.7643(0.0021)
		β	1.8306(0.0405)	1.8295(0.0071)	1.9018(0.0314)
20	10%	η_0	2.0001(0.0001)	2.1032(0.0061)	2.0793(0.0008)
		η_1	0.6615(0.0003)	1.1015(0.0001)	0.7852(0.0375)
		β	1.9013(0.0001)	2.1809(0.0005)	2.1893(0.0003)
	20%	η_0	2.0201(0.0041)	2.1763(0.0003)	2.1039(0.0006)
		η_1	0.9511(0.0051)	1.1529(0.0006)	1.1421(0.0004)
		β	1.9381(0.0254)	2.1859(0.0007)	2.0202(0.0005)
30	10%	η_0	2.1921(0.0013)	2.2043(0.0003)	2.1821(0.0002)
		η_1	1.1205(0.0004)	1.1729(0.0032)	1.1725(0.0005)
		β	1.9421(0.0006)	2.1071(0.0003)	2.1835(0.0005)
	20%	η_0	2.1031(0.0013)	2.0832(0.0001)	2.1635(0.0004)
		η_1	0.8305(0.0002)	1.0629(0.0005)	0.7581(0.0040)
		β	1.9909(0.0005)	1.9877(0.0001)	2.2073(0.0004)

Discussion : On remarque que sous la fonction de perte entropie ; on obtient les meilleurs estimateurs des paramètres η_0 et β au sens où le risque a posteriori est plus petit. Alors que clairement les deux autres fonctions de perte (quadratique généralisée et Linex) donnaient le même niveau de performance.

3. Étant donné que les fonctions de survie et le taux de hasard dépendent du temps, on considère ici l'intervalle de temps $t = [1, 50]$, on trace alors le graphe de $h(t)$ et $S(t)$ avec leurs valeurs estimées.

on choisit la taille de l'échantillonnage $n = 30$, on utilise l'estimateur sous la fonction de perte entropie dans les deux cas de censure fois 10% et 20%. (le meilleur estimateur comme nous l'avons vu).

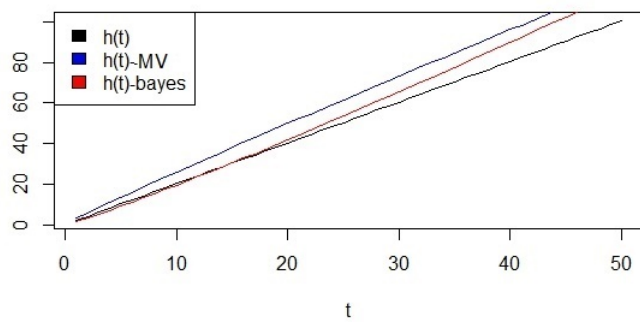


FIGURE 3.2 – Le taux de hasard avec 10%

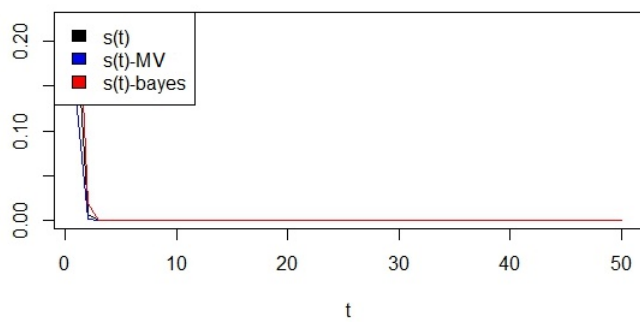


FIGURE 3.3 – La fonction de survie avec 10%

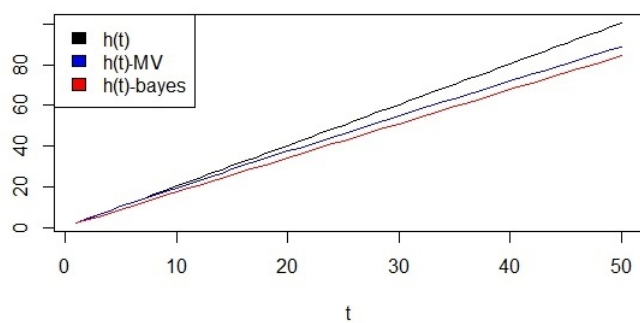


FIGURE 3.4 – Le taux de hasard avec 20%

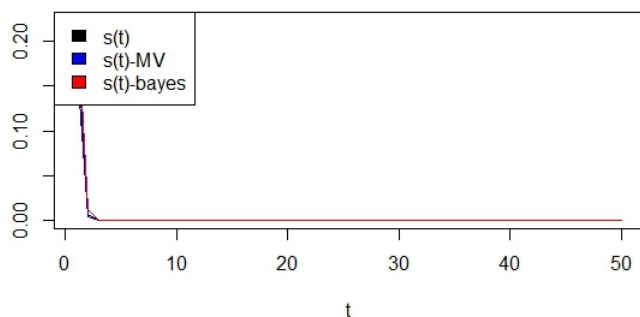


FIGURE 3.5 – La fonction de survie avec 20%

Discussion : On remarque que dans le cas de la fonction de survie, lorsque $t < 5$, les trois fonctions sont proches les unes des autres, mais elles coïncident lorsque $t > 5$, de plus la fonction de survie utilisant les estimations Bayésiennes est plus proche de la fonction de survie réelle, pour le taux de hasard, il est clair que l'estimateur Bayésien fonctionne mieux lorsque 10% des données sont censurées.

3.1.4.2 Comparaison des estimateurs

Il existe de nombreux critères pour comparer des estimateurs entre eux, souvent, la comparaison à partir de leurs variances est la plus utilisée, cependant le calcul des variances est généralement difficile, sinon impossible, alors, on utilise deux critères pour comparer les estimations Bayésiennes proposées avec l'estimation du maximum de vraisemblance (MV) ; le critère de proximité de Pitman (voir Pitman [37] et Jozani [26]) et l'erreur quadratique moyenne intégrée (IMSE) que l'on définit respectivement comme suit ;

1. Un estimateur θ_1 d'un paramètre θ préforme mieux qu'un autre estimateur θ_2 selon le critère de proximité de Pitman

$$P_{\theta}[|\theta_1 - \theta| < |\theta_2 - \theta|] > \frac{1}{2}.$$

2. Un estimateur $\theta_i (i = 1, l \dots, N)$ obtenu avec des échantillons de modèle N . L'erreur

carrée moyenne intégrée est

$$\text{IMSE} = \left(\sum_{i=1}^N (\theta_i - \theta)^2 \right) / N.$$

Dans le tableaux suivant, on présente les valeurs des probabilités de Pitman qui nous permettent de comparer les estimateurs Bayésiens avec l'estimateur du maximum de vraisemblance sous les trois fonctions de perte lorsque $\gamma = -2, p = -1$ et $r = -0.5$

TABLE 3.7 – Comparaison de Pitman

n	.%	par	GQ($\gamma = -2$)	entropy ($p = -1$)	Linex ($r = -0.5$)
10	10%	η_0	0.735	0.719	0.682
		η_1	0.352	0.205	0.226
		β	0.682	0.575	0.594
	20%	η_0	0.546	0.522	0.519
		η_1	0.366	0.312	0.299
		β	0.557	0.519	0.533
20	10%	η_0	0.134	0.148	0.175
		η_1	0.566	0.593	0.629
		β	0.335	0.394	0.341
	20%	η_0	0.288	0.275	2.242
		η_1	0.501	0.516	0.513
		β	0.205	0.201	0.197
30	10%	η_0	0.349	0.318	0.299
		η_1	0.561	0.601	0.621
		β	0.127	0.122	0.119
	20%	η_0	0.115	0.118	0.202
		η_1	0.231	0.209	0.285
		β	0.243	0.219	0.208

Discussion : Lorsque n est petit, les estimateurs Bayésiens de η_0 et β sont mieux que les estimateurs du maximum de vraisemblance.

De plus, on remarque que la fonction de perte quadratique fournit les meilleures valeurs, mais l'estimateur du maximum de vraisemblance de η_1 est meilleur que l'estimateur Bayésien. Lorsque n est grand les estimateurs du maximum de vraisemblance des trois paramètres sont meilleurs que les estimateurs Bayésiens.

Dans le tableau suivant, on présente les valeurs de l'erreur quadratique moyenne intégrée des estimateurs sous les trois fonction de pertes et l'estimateur du maximum de vraisemblance.

TABLE 3.8 – Comparaison IMSE

n	.%	par	MV	GQ($\gamma = -2$)	entropy ($p = -1$)	Linex ($r = -0.5$)
10	10%	η_0	0.1481	0.0061	0.0106	0.0081
		η_1	0.0053	0.2914	0.2143	0.2205
		β	0.1934	0.0714	0.0721	0.0034
	20%	η_0	0.0713	0.0007	0.0031	0.0033
		η_1	0.0529	0.2305	0.2115	0.2101
		β	0.0105	0.0715	0.0063	0.0042
20	10%	η_0	0.0613	0.0043	0.0073	0.0031
		η_1	0.0359	0.2093	0.2063	0.2215
		β	0.1048	0.0011	0.0053	0.0043
	20%	η_0	0.0801	0.2301	0.2297	0.2543
		η_1	0.0563	0.0984	0.1001	0.0975
		β	0.0421	0.0405	0.0463	0.0441
30	10%	η_0	0.1622	0.0063	0.0501	0.0479
		η_1	0.0393	0.0152	0.0322	0.0310
		β	0.1682	0.0215	0.0329	0.0308
	20%	η_0	0.0751	0.0511	0.0521	0.0522
		η_1	0.4242	0.3952	0.4102	0.4088
		β	0.2905	0.2143	0.2184	0.2123

Discussion : Lorsque n est petit, les estimateurs Bayésiens de η_0 et β fournissent le plus petit IMSE par rapport aux estimateurs du maximum de vraisemblance. Mais pour η_1 , l'estimateur du maximum de vraisemblance est meilleur que l'estimateur Bayésien.

Lorsque n est grand, tous les estimateurs Bayésiens sont meilleurs que l'estimateur du maximum de vraisemblance, et nous pouvons remarquer que la fonction de perte quadratique généralisée fournit les meilleures valeurs de l'IMSE.

3.2 L'estimation de la distribution de Zeghdoudi tronquée à droite

La troncature est le processus d'exclusion et d'omission de toutes les valeurs qui se trouvent en dehors des limites prédéterminées dans une expérience statistique. Les points de données restants à l'intérieur de ces limites sont appelés données tronquées.

Une variable aléatoire X est tronquée à droite (à gauche), à un seuil donné c , si seulement les résultats de X pour lesquels $X \leq c$ ($X > c$) sont pris en compte, c'est-à-dire que nous omettons toutes les valeurs de X pour lesquelles $X > c$ ($X \leq c$). Par exemple,

l'âge est plus tronqué, car il commence à zéro ($c = 0$).

Les distributions tronquées sont utilisées dans un large éventail d'application. Les travaux récents incluent Bantan et al. (2019)[10], Boudjerda et al. (2016)[13], Balakrishnan et Mitra (2012)[5].

La distribution de Lindley avec un paramètre θ est un mélange de distributions exponentielles (θ) et gamma ($2, \theta$). Elle est largement utilisée dans la modélisation des données de durée de vie dont la série d'articles de Zeghdoudi et Nedjar [30], où il est démontré qu'elle correspond bien à une grande classe d'ensembles de données réelles. Messaadia et Zeghdoudi (2018)[31] ont suggéré une famille de distributions exponentielles à un paramètre qui est basé sur des mélanges de distribution gamma($2, \theta$) et gamma($3, \theta$), connue sous le nom de distribution de Zeghdoudi et ont montré qu'elle s'adapte bien aux ensembles de données à vie où la distribution de Lindley donnait un mauvais ajustement.

3.2.1 La distribution de Zeghdoudi tronquée à droite

La densité de probabilité de la distribution de Zeghdoudi (voir Messaadia et Zeghdoudi) est donnée par

$$f_{ZD}(x, \theta) = \frac{\theta^3 x(1+x)e^{-\theta x}}{2+\theta}, \quad x, \theta > 0, \quad (3.11)$$

et sa fonction de répartition est

$$F_{ZD}(x) = \frac{1 - (x^2\theta^2 + \theta(\theta+2)x + \theta + 2)}{(\theta+2)e^{-\theta x}}, \quad x, \theta > 0. \quad (3.12)$$

La densité de probabilité de la distribution de Zeghdoudi tronquée à droite avec le paramètre $\beta > 0$, est donné par

$$f_{TZ}(x, \theta) = f_{ZD}(x, \theta | X \leq \beta) = \frac{f_{ZD}(x, \theta)}{F_{ZD}(\beta)},$$

qui, compte tenu de (3.11) et (3.12), est explicitement donné par la formule

$$f_{TZ}(x, \theta) = \left(\frac{\theta^3 x(1+x)}{(\theta+2)e^{\beta\theta} - (\beta^2\theta^2 + \theta(\theta+2)\beta + \theta + 2)} \right) e^{(-\theta x + \beta x)}, \quad \theta \leq x \leq \beta. \quad (3.13)$$

Ainsi, la fonction de répartition est

$$F_{TZ}(x) = \left(\frac{(\theta + 2)e^{\theta x} - x^2\theta^2 - \theta(\theta + 2)x - \theta - 2}{(\theta + 2)e^{\beta\theta} - \beta^2\theta^2 - \theta(\theta + 2)\beta - \theta - 2} \right) e^{(-x\theta + \beta\theta)}. \quad (3.14)$$

3.2.2 Estimation des paramètres par le maximum de vraisemblance

Laissez l'échantillon (x_1, x_2, \dots, x_n) être généré à partir du modèle Zeghdoudi tronqué à droite, en supposant que les données sont censurées de type II, c'est-à-dire que nous n'observons que (x_1, x_2, \dots, x_m) , pour un $m \in \{1, 2, \dots, n\}$. La fonction de vraisemblance est

$$L(\theta, \beta|X) = \frac{n!}{(n-m)!} \prod_{i=1}^m f_{TZ}(x_i, \theta, \beta) [(1 - F_{TZ}(x_m))]^{n-m}.$$

où

$$A = n!/(n-m)!,$$

$$B = [(\theta + 2)\exp(\beta\theta) - (\beta^2\theta^2 + \theta(\theta + 2)\beta + \theta + 2)]^{-n},$$

$$C = [\beta^2\theta^2 + \theta(\theta + 2)\beta + \theta + 2 + (x_m\theta^2 + (\theta + 2)x_m + \theta + 2)e^{(-x_m\theta + \beta\theta)}],$$

La fonction de vraisemblance se lit

$$L(\theta, \beta|X) = AB^{-n}C^{n-m}\theta^{3m} \exp(m\beta\theta) \prod_{i=1}^m x_i(1 + x_i) \exp(\theta x_i). \quad (3.15)$$

passant par le logarithme, on trouve

$$\begin{aligned} l(\theta, \beta|X) &= \ln(A) - n \ln(B) + (n-m) \ln(C) + 3m \ln(\theta) + m\beta\theta \\ &+ \sum_{i=1}^m \ln(x_i) + \sum_{i=1}^m \ln(1 + x_i) + \theta \sum_{i=1}^m x_i. \end{aligned} \quad (3.16)$$

La solution du système non linéaire suivant donne les estimateurs du maximum de vraisemblance θ_{MV} et β_{MV} des paramètres θ, β , respectivement.

$$\begin{cases} \frac{\partial}{\partial \theta} l(\theta, \beta | X) = -n \frac{B_1}{B} + (n - m) \frac{C_1}{C} + \frac{3m}{\theta} + m\beta + \sum_{i=1}^m x_i = 0, \\ \frac{\partial}{\partial \beta} l(\theta, \beta | X) = -n \frac{B_2}{B} + (n - m) \frac{C_2}{C} + m\theta = 0, \end{cases} \quad (3.17)$$

dont

$$\begin{aligned} B_1 &= \frac{\partial B}{\partial \theta} = e^{\beta\theta} [\beta(\theta + 2) + 1] - 2\beta^2\theta - 2\beta(\theta + 1) - 1, \\ C_1 &= \frac{\partial C}{\partial \theta} = - [2\beta^2\theta + 2\beta(\theta + 1) + 1 + (2x_m^2\theta + 2x_m(\theta + 1) + 1)e^{(-x_m\theta + \beta\theta)} \\ &\quad + (\beta - x_m)(x_m^2\theta^2 + \theta(\theta + 2)x_m + \theta + 2)e^{(-x_m\theta + \beta\theta)}], \\ B_2 &= \frac{\partial B}{\partial \beta} = \theta(\theta + 2) \exp(\beta\theta) - (2\beta\theta^2 + \theta(\theta + 2)), \\ C_2 &= \frac{\partial C}{\partial \beta} = - (2\beta\theta^2 + \theta(\theta + 2) + \theta(x_m^2 + \theta^2 + \theta(\theta + 2)x_m + \theta + 2)e^{(-x_m\theta + \beta\theta)}). \end{aligned}$$

La solution du système (3.17) semble analytiquement insoluble. Nous nous appuyerons sur des méthodes numériques pour obtenir des solutions approximatives. On utilisera le paquet *BB* du logiciel *R* pour obtenir la valeur approximative de les estimations du maximum de vraisemblance θ_{MV} et β_{MV} des paramètres θ, β . Le paquet *BB* est utilisé avec succès pour résoudre un système d'équations non linéaire, voir Varadhan et Gilbert (2010).

3.2.3 Estimation Bayésienne sous différentes fonctions de perte

On suppose que les paramètres θ, β sont indépendants et suivent la distribution Gamma comme une distribution a priori,

$$\pi(\theta) = \frac{a^b}{\Gamma(b)} \theta^{b-1} \exp(-a\theta), \quad \beta, \theta > 0, a, b > 0, \quad (3.18)$$

$$\pi(\beta) = \frac{c^d}{\Gamma(d)} \beta^{d-1} \exp(-c\beta), \quad \beta, \theta > 0, c, d > 0. \quad (3.19)$$

telles que les constantes a, b, c, d sont appelées hyper-paramètres.

Il n'y a pas de motivation objective pour choisir la famille gamma comme distributions a priori, sauf pour leur traçabilité et pour être des préalables conjugués naturels pour les

distributions exponentielles, d'autres distributions a priori pourraient bien être utilisées.

La distribution a posteriori conjointe de (θ, β) est alors

$$\pi(\theta, \beta|x) = KB^{-n}C^{n-m}\theta^{b-1+3m}e^{(-a\theta-c\beta+m\beta\theta)} \prod_{i=1}^m x_i(1+x_i)e^{\theta x_i}, \quad (3.20)$$

où K est une constante de normalisation.

On procède comme la section précédente, Nous obtenons les estimateurs Bayésiens sous les trois fonctions de perte choisies (quadratique généralisé (GQ), Linex et l'entropie). 1.1

Dans le cas de la fonction de perte quadratique généralisée, les estimateurs de Bayes sont donnés par les formules :

$$\hat{\theta}_{GQ} = \frac{\int \int_0^{+\infty} B^{-n}C^{n-m}\theta^{b+3m+\gamma-1}e^{(-a\theta-c\beta+m\beta\theta)} \prod_{i=1}^m x_i(1+x_i)e^{\theta x_i} d\theta d\beta}{\int \int_0^{+\infty} B^{-n}C^{n-m}\theta^{b+3m+\gamma-2}e^{(-a\theta-c\beta+m\beta\theta)} \prod_{i=1}^m x_i(1+x_i)e^{\theta x_i} d\theta d\beta},$$

$$\hat{\beta}_{GQ} = \frac{\int \int_0^{+\infty} B^{-n}C^{n-m}\theta^{b_1-1+3m+\beta\gamma}e^{(-a_1\theta-a_2\beta+m\beta\theta)} \prod_{i=1}^m x_i(1+x_i)e^{\theta x_i} d\theta d\beta}{\int \int_0^{+\infty} B^{-n}C^{n-m}\theta^{b_1-1+3m}\beta^{\gamma-1}e^{(-a_1\theta-a_2\beta+m\beta\theta)} \prod_{i=1}^m x_i(1+x_i)e^{\theta x_i} d\theta d\beta}.$$

Les risques a posteriori correspondants sont alors,

$$PR(\hat{\theta}_{GQ}) = E_{\pi}(\theta^{\gamma+1}) - 2\hat{\theta}_{GQ}E_{\pi}(\theta^{\gamma}) + \hat{\theta}_{GQ}^2E_{\pi}(\theta^{\gamma-1}),$$

et

$$PR(\hat{\beta}_{GQ}) = E_{\pi}(\beta^{\gamma+1}) - 2\hat{\beta}_{GQ}E_{\pi}(\beta^{\gamma}) + \hat{\beta}_{GQ}^2E_{\pi}(\beta^{\gamma-1}).$$

Sous la fonction de perte d'entropie, on obtient les estimateurs suivants :

$$\hat{\theta}_E = \left[K \int \int_0^{+\infty} B^{-n}C^{n-m}\theta^{b-1+3m-p}e^{(-a\theta-c\beta+m\beta\theta)} \prod_{i=1}^m x_i(1+x_i)e^{\theta x_i} d\theta d\beta \right]^{-\frac{1}{p}},$$

$$\hat{\beta}_E = \left[K \int \int_0^{+\infty} B^{-n}C^{n-m}\theta^{b-1+3m}\beta^{-p}e^{(-a\theta-c\beta+m\beta\theta)} \prod_{i=1}^m x_i(1+x_i)e^{\theta x_i} d\theta d\beta \right]^{-\frac{1}{p}}.$$

Les risques a posteriori correspondants sont alors,

$$PR(\hat{\theta}_E) = PE_{\pi}(\log(\theta) - \log(\hat{\theta}_E)),$$

$$PR(\hat{\beta}_E) = PE_\pi(\log(\beta) - \log(\hat{\beta}_E)).$$

Finalement, Sous la fonction de perte Linex, on obtient les estimateurs suivants :

$$\hat{\theta}_L = \frac{K}{r} \ln \left[\int \int_0^{+\infty} B^{-n} C^{n-m} \theta^{b-1+3m} e^{(-a\theta-c\beta+m\beta\theta-r\theta)} \prod_{i=1}^m x_i (1+x_i) e^{\theta x_i} d\theta d\beta \right],$$

$$\hat{\beta}_L = \frac{K}{r} \ln \left[\int \int_0^{+\infty} B^{-n} C^{n-m} \theta^{b-1+3m} e^{(-a\theta-c\beta+m\beta\theta-r\theta)} \prod_{i=1}^m x_i (1+x_i) e^{\theta x_i} d\theta d\beta \right],$$

Et les risques a posteriori correspondants sont ,

$$PR(\hat{\theta}_L) = r(\hat{\theta}_{GQ} - \hat{\theta}_L),$$

$$PR(\hat{\beta}_L) = r(\hat{\beta}_{GQ} - \hat{\beta}_L).$$

Dans la section suivante, on utilise la méthode MCMC pour évaluer ces estimateurs.

3.2.4 Simulation

On compare la performance des estimateurs du maximum de vraisemblance et de Bayes des paramètres inconnus pour la distribution tronquée de Zeghdoudi sous les données censurées de type II.

Pour les hyper-paramètres donnés $a = b = c = d = 1$, et $\theta = 1, \beta = 1, 5$, en utilisant des échantillons $N = 5000$ de notre distribution , on obtient les résultats suivants.

3.2.4.1 Résultats de l'estimation

1. On a utilisé le paquet *BB* du logiciel *R* pour dériver les valeurs numériques des estimateurs du maximum de vraisemblance,

À partir de 3.9, nous voyons que les valeurs estimées de θ sont proches de sa valeur réelle. De plus, l'erreur quadratique est faible. Toutefois, les valeurs estimées de β ne sont pas proches de sa valeur réelle.

2. Pour évaluer les estimateurs Bayésiens, on utilise l'algorithme de Metropolis-Hastings. Dans le cas de la fonction de perte quadratique généralisée, nous utilisons $\gamma \in \{2, -1.5, -1, -0.5, 0.5, 1, 1.5, 2\}$, pour la fonction de perte d'entropie, nous utilisons

TABLE 3.9 – Estimations du maximum de vraisemblance (MV) avec leurs erreurs Quadratiques.

n	m	paramètre	MV
10	8	θ	0.7742 (0.0735)
		β	1.9654 (0.1763)
30	24	θ	0.9135 (0.0078)
		β	1.6302 (0.0152)
50	40	θ	0.9573 (0.0045)
		β	1.6873 (0.0078)
100	80	θ	0.9843 (0.0075)
		β	1.5259 (0.0005)
200	160	θ	0.9397 (0.0054)
		β	1.6434 (0.0217)

$p \in \{2, -1.5, -1, -0.5, 0.5, 1, 1.5, 2\}$. Enfin, pour la fonction de perte Linex, nous utilisons $r \in \{2, -1.5, -1, -0.5, 0.5, 1, 1.5, 2\}$.

TABLE 3.10 – Estimation Bayésienne sous la perte quadratique généralisée avec le risque a posteriori (entre parenthèses) .

n	m	par	γ							
			-2	-1.5	-1	-0.5	0.5	1	1.5	2
10	8	θ	1,0881 (0,0065)	1,0925 (0,0079)	1,0990 (0,0087)	1,0994 (0,0089)	1,11109 (0,0095)	1,1175 (0,0091)	1,1099 (0,0098)	1,1241 (0,01193)
		β	1,1941 (0,0432)	1,2191 (0,0593)	1,2657 (0,7091)	1,2999 (0,0825)	1,3891 (0,0909)	1,4228 (0,1094)	1,4768 (0,1241)	1,5198 (0,1451)
30	24	θ	1,0739 (0,0043)	1,0829 (0,0058)	1,0799 (0,0066)	1,0884 (0,0061)	1,0957 (0,0061)	1,0965 (0,0069)	1,1021 (0,0077)	1,0971 (0,0089)
		β	1,1805 (0,0429)	1,2201 (0,0509)	1,2297 (0,0671)	1,3011 (0,0711)	1,3915 (0,0802)	1,4135 (0,0999)	1,4841 (0,1097)	1,4955 (0,1228)
50	40	θ	1,0739 (0,0051)	1,0792 (0,0058)	1,0825 (0,0061)	1,0888 (0,0070)	1,0922 (0,0077)	1,0977 (0,0078)	1,0990 (0,0081)	1,1035 (0,0085)
		β	1,1670 (0,0439)	1,1799 (0,0499)	1,2039 (0,0633)	1,2701 (0,711)	1,3591 (0,995)	1,3803 (0,1071)	1,4191 (0,1181)	0,4702 (0,1201)
100	80	θ	1,2111 (0,0016)	0,2132 (0,0018)	0,2135 (0,0019)	0,2147 (0,0021)	0,2151 (0,0026)	0,2159 (0,0028)	0,2701 (0,0030)	0,2209 (0,0030)
		β	1,7218 (0,0006)	1,7220 (0,0007)	1,7234 (0,0008)	1,7245 (0,0011)	1,7258 (0,0023)	1,7268 (0,0028)	1,7301 (0,0035)	1,7290 (0,0048)
200	160	θ	1,2108 (0,0013)	1,2118 (0,0015)	1,2127 (0,0016)	1,2138 (0,0018)	2,1839 (0,0020)	2,1841 (0,0031)	2,1841 (0,0042)	2,1851 (0,0066)
		β	1,7105 (0,0005)	1,7115 (0,0007)	1,7120 (0,0008)	1,7131 (0,0012)	1,7139 (0,0019)	1,7149 (0,0025)	1,7158 (0,0033)	1,7169 (0,0046)

On note que la valeur choisie $\gamma = -2$ fournit les meilleurs risques a posteriori, ce qui

présente le meilleur estimateur pour la fonction de perte quadratique généralisée. De plus, lorsque n est grand, nous obtenons les plus petits risques a posteriori.

Le tableau suivant présente les valeurs des estimateurs Bayésiens et leurs risques a posteriori entre parenthèses sous la fonction de perte d'entropie.

TABLE 3.11 – Estimation Bayésienne sous la perte d'entropie généralisée avec le risque a posteriori (entre parenthèses) .

n	m	par	p							
			-2	-1.5	-1	-0.5	0.5	1	1.5	2
10	8	θ	1,1097 (0,0099)	1,1067 (0,0091)	1,1053 (0,0035)	1,1041 (0,0009)	1,0995 (0,0008)	0,0981 (0,0038)	0,0942 (0,008)	0,0911 (0,0155)
		β	1,4579 (0,0997)	1,4407 (0,0611)	1,4239 (0,0199)	1,4177 (0,0072)	1,3639 (0,0071)	1,3491 (0,0308)	1,3188 (0,0699)	1,2909 (0,1249)
30	24	θ	1,0999 (0,0191)	1,0981 (0,0065)	1,0898 (0,0042)	1,0863 (0,0007)	1,0831 (0,0008)	0,0817 (0,003)	1,0799 (0,009)	1,0781 (0,0111)
		β	1,4325 (0,1011)	1,4151 (0,0622)	1,3966 (0,0281)	1,3709 (0,0074)	1,3061 (0,0077)	1,3185 (0,0319)	1,2908 (0,0690)	1,2901 (0,1246)
50	40	θ	1,7644 (0,1173)	1,7188 (0,1443)	1,6701 (0,0667)	1,6205 (0,0171)	1,5809 (0,0729)	1,4830 (0,0733)	1,3990 (0,1644)	1,3458 (0,2849)
		β	1,4259 (0,0944)	1,4077 (0,0661)	1,3881 (0,0303)	1,3633 (0,0073)	1,3255 (0,0065)	1,3055 (0,319)	1,2839 (0,009)	1,2588 (0,1265)
100	80	θ	1,2101 (0,0038)	1,2096 (0,0022)	1,2089 (0,0009)	1,2084 (0,0002)	1,2073 (0,0003)	1,2067 (0,0009)	1,2061 (0,0032)	1,2056 (0,0038)
		β	1,7078 (0,0018)	1,7071 (0,0010)	1,7060 (0,0005)	1,7062 (0,0001)	1,7054 (0,0001)	1,7049 (0,0006)	1,7046 (0,0010)	1,7042 (0,0021)
200	160	θ	1,2188 (0,0031)	1,2179 (0,0017)	1,2169 (0,0009)	1,2167 (0,0001)	1,2154 (0,0001)	1,2148 (0,0009)	1,2144 (0,0019)	1,2139 (0,0032)
		β	1,7071 (0,0018)	1,7060 (0,0010)	0,7059 (0,0006)	1,7051 (0,0002)	1,7044 (0,0001)	1,7037 (0,0005)	1,7034 (0,014)	1,7030 (0,0019)

À partir de ce tableau, la valeur $p = -0,5$ fournit les meilleurs risques a posteriori, ce qui présente les meilleures estimations pour la fonction de perte d'entropie. De plus, lorsque $n = 100$ et $n = 200$, nous obtenons les plus petits risques a posteriori.

Enfin, on présente les valeurs des estimateurs Bayésiens et leurs risques a posteriori entre parenthèses, sous la fonction de perte Linex.

TABLE 3.12 – Estimation Bayésienne sous la perte Linex généralisée avec le risque a posteriori (entre parenthèses) .

n	m	par	r							
			-2	-1.5	-1	-0.5	0.5	1	1.5	2
10	8	θ	1,1251 (0,0195)	1,1228 (0,0105)	1,1201 (0,0039)	1,1189 (0,0011)	1,1132 (0,0012)	1,1108 (0,0049)	1,1082 (0,0107)	1,1052 (0,0185)
		β	1,5309 (0,1666)	1,5117 (0,1033)	1,4806 (0,0411)	1,4492 (0,0177)	1,4132 (0,0167)	1,3581 (0,0433)	0,3495 (0,01213)	0,3345 (0,03131)
30	24	θ	1,1058 (0,0147)	1,1041 (0,0085)	1,1022 (0,0039)	1,1005 (0,0007)	1,0991 (0,0007)	1,0948 (0,0038)	1,0930 (0,3577)	1,0912 (0,6107)
		β	1,4951 (0,1678)	1,4759 (0,1001)	1,4547 (0,1041)	1,4315 (0,0481)	1,3815 (0,0183)	1,3538 (0,0131)	1,3272 (0,0119)	1,3014 (0,1205)
50	40	θ	1,0939 (0,0151)	1,0919 (0,0081)	1,0815 (0,0038)	1,0861 (0,0009)	1,0813 (0,0007)	1,0811 (0,0038)	1,0799 (0,0081)	1,0771 (0,0111)
		β	1,4821 (0,1884)	1,4639 (0,0581)	1,4455 (0,0519)	1,4193 (0,0131)	1,3609 (0,0199)	1,3338 (0,0655)	1,3160 (0,1155)	1,2915 (0,2109)
100	80	θ	2,2070 (0,0057)	2,2061 (0,0015)	2,2041 (0,0013)	2,2039 (0,0003)	2,0019 (0,0004)	2,2017 (0,0015)	2,0011 (0,0032)	2,0915 (0,0056)
		β	1,7094 (0,0057)	1,7091 (0,0032)	1,7080 (0,0014)	1,7077 (0,0004)	1,7059 (0,0013)	1,7055 (0,0049)	1,7044 (0,0031)	1,7038 (0,0054)
200	160	θ	1,2191 (0,0049)	1,2181 (0,0027)	1,2179 (0,0012)	1,2174 (0,0003)	1,2153 (0,0003)	1,2151 (0,0012)	1,2133 (0,0025)	1,2129 (0,0049)
		β	1,7031 (0,0051)	1,7059 (0,0031)	1,7054 (0,0013)	1,7045 (0,0004)	1,7011 (0,0004)	1,7041 (0,0012)	1,7039 (0,0027)	1,7015 (0,0041)

Ici, la valeur $r = -0.5$ fournit les meilleurs risques a posteriori, ce qui présente les meilleures estimations pour la fonction de perte Linex. De plus, lorsque n est élevé, nous obtenons le plus petit risques a posteriori.

Dans le tableau qui ce suit, on présente les trois fonctions de perte correspondantes aux meilleurs résultats, c'est-à-dire lorsque $\gamma = -2, p = -0,5$ et $r = -0,5$.

TABLE 3.13 – Estimation Bayésienne sous les fonctions de perte avec le risque a posteriori (entre parenthèses) .

n	m	par	GQ($\gamma = -2$)	Entropie ($p = -0,5$)	Linex($r = 0,5$)
10	8	θ	1,0881 (0,0065)	1,1041 (0,0009)	1,1189 (0,0011)
		β	1,1941 (0,0432)	1,4177 (0,0072)	1,4492 (0,0177)
30	24	θ	1,0739 (0,0043)	1,0863 (0,0007)	1,1005 (0,0007)
		β	1,1805 (0,0429)	1,3709 (0,0074)	1,4315 (0,0481)
50	40	θ	1,0739 (0,0051)	1,6205 (0,0171)	1,0861 (0,0009)
		β	1,1670 (0,0439)	1,3633 (0,0073)	1,4193 (0,0131)
100	80	θ	1,2111 (0,0016)	1,2084 (0,0002)	2,2039 (0,0003)
		β	1,7218 (0,0006)	1,7062 (0,0001)	1,7077 (0,0004)
200	160	θ	1,2108 (0,0013)	1,2167 (0,0001)	1,2174 (0,0003)
		β	1,7105 (0,0005)	1,7051 (0,0002)	1,7045 (0,0004)

En résumé, à partir des valeurs calculées, il s'ensuit que les meilleurs résultats sont obtenus par la fonction de perte d'entropie.

3.2.4.2 Comparaison des estimateurs

Dans le tableau suivant, nous présentons les valeurs des probabilités de Pitman qui nous permettent de comparer les estimateurs Bayésiens avec les estimateurs du maximum de vraisemblance sous les trois fonctions de perte lorsque $\gamma = -2, p = -0,5, r = -0,5$.

TABLE 3.14 – Comparaison sous le critère de Pitman

n	m	par	GQ($\gamma = -2$)	Entropie ($p = -0,5$)	Linex($r = -0,5$)
10	8	θ	0,6791	0,635	0,659
		β	0,657	0,519	0,635
30	24	θ	0,579	0,558	0,569
		β	0,534	0,499	0,516
50	40	θ	0,539	0,519	0,600
		β	0,441	0,385	0,381
100	80	θ	0,130	0,1327	0,1809
		β	0,491	0,335	0,394
200	160	θ	0,2214	0,071	0,091
		β	0,288	0,249	0,2191

Selon le critère de Pitman, les estimateurs Bayésiens de θ sont meilleurs que θ_{MV} lorsque n est petit. De plus, la fonction de perte quadratique généralisée a les meilleures valeurs par rapport aux deux autres fonctions de perte. Cependant, β_{MV} est plus proche de la valeur réelle que tous les estimateurs Bayésiens. Lorsque n est grand, θ_{MV} et β_{MV} fonctionnent mieux que les estimateurs Bayésiens correspondants .

Dans le tableau ci-dessous, on présente les valeurs de l'erreur quadratique moyenne intégrée des estimateurs Bayésiens et les estimateurs du maximum de vraisemblance.

TABLE 3.15 – L'erreur quadratique moyenne intégrée.

n	m	par	MV	GQ	Entropie	Linex
10	8	θ	0,2589	0,0092	0,0224	0,0185
		β	0,3188	0,1065	0,1082	0,7141
30	24	θ	0,1045	0,0098	0,0158	0,0149
		β	0,3689	0,1159	0,1213	0,1179
50	40	θ	0,0801	0,0214	0,0189	0,1494
		β	0,2249	0,1214	0,1192	0,1147
100	80	θ	0,0954	0,0259	0,0397	0,0344
		β	0,2941	0,0421	0,0448	0,0423
200	160	θ	0,1808	0,1625	0,1673	0,1674
		β	0,4748	0,3367	0,3165	0,3154

On remarque que lorsque n est petit, les estimateurs Bayésiens de θ et β fournissent un petit $IMSE$ pour les paramètres par rapport à θ_{MV} et β_{MV} . De plus, les valeurs fournies par la fonction de perte quadratique généralisée sont relativement proches de celles fournies par les fonctions d'entropie et de perte de Linex. Pour conclure, les estimateurs Bayésiens fonctionnent mieux que les estimateurs du maximum de vraisemblance, et la

fonction de perte quadratique généralisée donne le plus petit *IMSE*.

3.2.5 Application sur des données réelles

Dans cette section, on étudie l'application de la distribution tronquée de Zeghdoudi en effectuant les estimations précédentes à l'aide d'un ensemble de données réelles.

L'ensemble de données comprend le nombre de particules d'or observées sur chaque unité de dystrophine (la dystrophine est un produit génétique d'importance possible dans les dystrophies musculaires) discuté dans Mathews et Appleton (1993) et M Cullen et al. (1990). Il est confirmé que la distribution de Zeghdoudi correspond à ces données en utilisant le test kolmogorov-Smirnov (K-S).

La valeur du test K-S est 0,012901, ce qui est inférieur à leur valeur critique correspondante au niveau de signification 5%, soit 0,025449 (pour $n = 198$). Sa *P*-valeur est égale à 0,793548.

Les observations complètes sont affichées dans le tableau suivant.

x_i	1	2	3	4	5	n
n_i	122	50	18	4	4	198

Il est clair que le point de troncature est égal à 5 dans le cas de données complètes. Nous l'assumons qu'il est égal à 3 pour les données censurées, en choisissant $m = 180$.

Le tableau suivant présente l'estimation Bayésienne des paramètres θ et β sous les trois fonctions de perte avec le risque à posteriori correspondant et les estimateurs du maximum de vraisemblance.

TABLE 3.16 – L'estimation avec les données réelles.

n	m	par	MV	GQ	Entropie	Linex
198	198	θ	0.9807	1.2108 (0,0010)	1.2018 (0,0003)	1.2174 (0,0005)
		β	1.5987	1.6405 (0,0005)	1.6051 (0,0009)	1.7045 (0,0006)
198	180	θ	0.9513	1.3408 (0,0017)	1.2027 (0,0008)	1.2174 (0,0099)
		β	1.6492	1.6605 (0,0008)	1.6151 (0,0005)	1.7045 (0,0009)

On note que les estimateurs basés sur l'ensemble complet des données présentent un risque à posteriori plus faible que les données censurées, ce qui est attendu puisque nous perdons une partie de l'information en utilisant des données censurées. Nous remarquons également que la fonction de perte d'entropie a le plus petit risque a posteriori.

Dans le tableau suivant, on présente les valeurs du critère de Pitman en utilisant l'ensemble des données réelles des estimateurs Bayésiens et du maximum de vraisemblance

TABLE 3.17 – Comparaison sous le critère de Pitman avec les données réelles

n	m	par	GQ($\gamma = -2$)	Entropie ($p = -0,5$)	Linex($r = -0,5$)
198	198	θ	0,213	0,1525	0,1920
		β	0,488	0,365	0,387
198	180	θ	0,2334	0,141	0,099
		β	0,308	0,321	0,2219

En comparant les estimateurs selon le critère de Pitman, l'estimateur du maximum de vraisemblance est meilleur que l'estimateur Bayésien.

Le tableau suivant présente les valeurs de l'erreur quadratique moyenne intégrée des estimateurs Bayésiens et des estimateurs du maximum de vraisemblance.

TABLE 3.18 – L'erreur quadratique moyenne intégrée avec les données réelles

n	m	par	MV	GQ	Entropie	Linex
198	198	θ	0.1893	0.1591	0.1613	0.1603
		β	0.5857	0.5417	0.5793	0.5734
198	180	θ	0.1903	0.1093	0.1190	0.1133
		β	0.5057	0.4350	0.5033	0.5040

On note que les valeurs fournies par la fonction de perte quadratique généralisée sont les plus petites et que tous les estimateurs Bayésiens fonctionnent mieux que les estimateurs du maximum de vraisemblance.

Chapitre No. 4

La distribution de Burr XII à trois paramètres

Les modèles statistiques décrivent un phénomène sous forme d'équations mathématiques. Beaucoup de modèles de durée de survie en théorie de la fiabilité est maintenant disponible dans le sujet pour décrire le comportement de survie d'un composant ou d'un système. La plupart des vies sont de nature continue et donc biens des distributions continues de survie ont été étudiées dans la littérature.

La distribution de Burr a été introduite pour la première fois par Burr (1942)[16]. La distribution la plus importante dans le système Burr est la distribution Burr Type XII. (tableau4). Elle a fait l'objet d'une attention particulière au cours des deux dernières décennies en raison de sa flexibilité qui lui permet d'être un bon modèle pour de vastes applications dans différents domaines tels que la mécanique, la biologie, l'économie et la finance (voir par exemple [32], [25], [4]), notamment dans le domaine de la fiabilité, de la modélisation du temps de survie et du plan d'échantillonnage d'acceptation, etc. Les douze types de distributions continues de Burr sont dans le tableau suivant

Type	$F(x)$	Intervalle
I	x	$[0,1]$
II	$(1 + e^{-x})^{-\beta}$	$[-\infty, \infty]$
III	$(1 + x^\theta)^{-\beta}$	$[0, \infty]$
IV	$\{1 + [x^{-1}(\theta - x)]^{1/\theta}\}^{-\beta}$	$[0, \theta]$
V	$\{1 + \theta e^{\tan(x)}\}^{-\beta}$	$[-\pi/2, \pi/2]$
VI	$\{1 + \theta e^{r \sinh(x)}\}^{-\beta}$	$[-\infty, \infty]$
VII	$2^{-\beta}(1 + \tanh(x))^\beta$	$[-\infty, \infty]$
VIII	$[2\pi^{-1} \arctan(e^x)]^{-\beta}$	$[-\infty, \infty]$
IX	$1 - 2\{2 + \theta[(1 + e^x)^\beta - 1]\}^{-1}$	$[-\infty, \infty]$
X	$(1 - e^{-x^2})^\beta$	$[0, \infty]$
XI	$[x - (2\pi)^{-1} \sin(2\pi x)]^\beta$	$[0, 1]$
XII	$1 - (1 + x^\theta)^{-\beta}$	$[0, \infty]$

L'inférence de la fiabilité du Burr XII sous les données de censure progressive a été étudiée par Gunaskera [23]. Soliman et al [45] ont effectué une inférence Bayésiennes sous échantillonnage censuré progressif de première défaillance. Shao [43] a dérivé l'estimateur du maximum de vraisemblance. de la distribution de Burr XII à trois paramètres.

Les modèles de Burr étendus sous différents schémas d'échantillonnage et pour plus de paramètres, n'ont été étudiés que récemment. El-sagheer et al. [21] ont étudié la distribution de Burr XII à trois paramètres dans le cadre d'un schéma hybride censuré. Afify et Mead [3] ont discuté des propriétés et des applications de la distribution Burr-XII à cinq paramètres.

Dans ce chapitre, on s'intéresse à l'étude de l'estimation paramétrique de la distribution de Burr à trois paramètres dans la présence de censure progressive de type II.

4.1 La distribution Burr XII (θ, β)

La distribution de Burr de type-XII a attiré l'attention depuis son introduction, elle comporte deux paramètres de forme $\theta > 0, \beta > 0$. d'où la fonction de répartition est :

$$F(x; \theta, \beta) = 1 - (1 + x^\theta)^{-\beta}$$

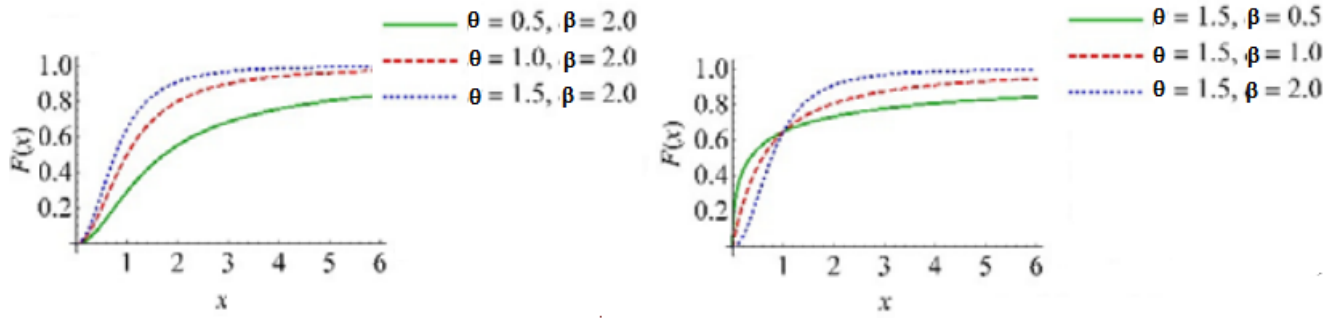


FIGURE 4.1 – La fonction de répartition du Burr XII (θ, β)

La densité de cette loi est donnée par

$$f(x; \theta, \beta) = \frac{dF(x; \theta, \beta)}{dx} = \theta\beta x^{\theta-1}(1+x^\theta)^{-(\beta+1)}, \quad x > 0. \quad (4.1)$$

On note que les figures 4.2 illustrent quelques-unes des formes possibles de la densité et la fonction de répartition de la distribution du Burr XII (θ, β) pour les valeurs sélectionnées des paramètres θ et β .

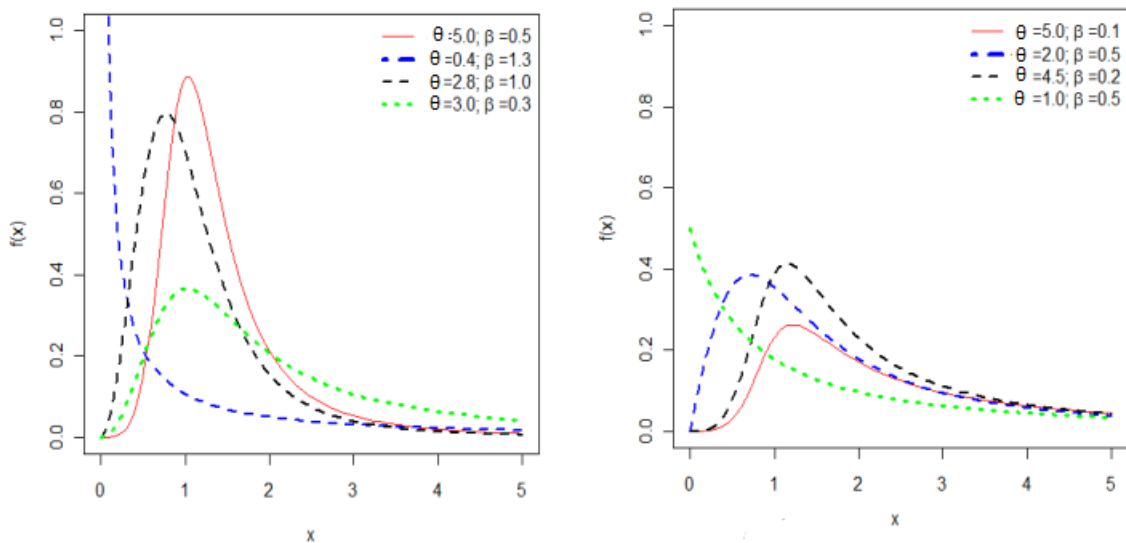


FIGURE 4.2 – La densité de Burr XII (θ, β)

4.1.1 Propriétés statistiques et mathématiques

Dans cette section, nous donnons quelques mesures statistiques et mathématiques importantes pour la distribution de Burr XII (θ, β) comme les quantiles, la médiane, le mode, les moments, le taux de hasard et la fonction de survie.

Les quantiles sont largement utilisées dans les statistiques générales, il s'ensuit que le p -ème quantile i.e. $Q(p) = F^{-1}(p)$ du Burr XII (θ, β) est donné par

$$Q(p) = \left[(1-p)^{\frac{1}{\beta}} - 1 \right]^{\frac{1}{\theta}}$$

En particulier, les premiers quantiles, Q_1, Q_3 , peuvent être obtenus en définissant $p = 0,25$ et $p = 0,75$ dans l'équation précédente.

La médiane de Burr XII (θ, β) est la valeur x à laquelle $F(x) = \frac{1}{2}$ autrement dit ; $Q(\frac{1}{2})$ d'où

$$Md = \left[\left(\frac{1}{2} \right)^{\frac{1}{\beta}} - 1 \right]^{\frac{1}{\theta}}$$

Le mode est la valeur de x pour laquelle la densité $f(x)$ est maximum, le mode est alors,

$$\begin{aligned} \frac{df(x; \theta, \beta)}{dx} &= [\theta \beta x^{\theta-1} (1+x^\theta)^{-(\beta+1)}] = 0 \\ \theta \beta x^{\theta-2} (1+x^\theta)^{-(\beta+2)} [(\theta-1) - x^\theta (1+\theta\beta)] &= 0 \\ \Rightarrow x^\theta &= \frac{(\theta-1)}{1+\theta\beta} \quad (x > 0, \alpha > 0) \text{ d'où} \end{aligned}$$

$$Mo = \left(\frac{\theta-1}{\beta\theta+1} \right)^{\frac{1}{\theta}} \quad \theta\beta > 1, \theta > 0, \beta > 0$$

Les moments sont nécessaires et plus importants dans toute analyse statistique, en particulier dans le travail appliqué. Certaines des caractéristiques les plus importantes d'une distribution peuvent être étudiées à travers des moments (par exemple, tendance, dispersion, asymétrie et aplatissement). Pour un entier positif fixe k , le moment d'ordre

k de la variable aléatoire X qui suit la loi de Burr XII(θ, β) est défini par ;

$$E(X^k) = \int_{-\infty}^{+\infty} x^k f(x; \theta, \beta) dx = \beta \int_0^{+\infty} x^k [(1+x^\theta)^{-(\beta+1)}] \theta x^{\theta-1} dx \quad (4.2)$$

Soit $u = (1+x^\theta)^{-1}$ d'où

$x = [(1-u)/u]^{1/\theta}$, alors,

$du = -\theta x^{\theta-1} (1+x^\theta)^{-2} dx = -\theta x^{\theta-1} u^2 dx$ donc,

$\theta x^{\theta-1} dx = -u^{-2} du$

Il est clair que $u \rightarrow 1$ comme $x \rightarrow 0$, et $u \rightarrow 0$ comme $x \rightarrow \infty$. Par conséquent, l'équation (4.30) devient

$$E(X^k) = \beta \int_1^0 \left[\left(\frac{1-u}{u} \right)^{1/\theta} \right]^k u^{\beta+1} (-u^{-2}) du = \beta \int_0^1 u^{(\beta-\frac{k}{\theta})-1} (1-u)^{(1+\frac{k}{\theta})-1} du \quad (4.3)$$

Lorsque $\beta - \frac{k}{\theta} > 0$ et $1 + \frac{k}{\theta} > 0$ (ou lorsque $k < \theta\beta$), la partie intégrale de l'équation (4.3) existe et est égale à $B(\beta - \frac{k}{\theta}, 1 + \frac{k}{\theta})$, où $B(.,.)$ est la fonction bêta. Parce que $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$, pour tout $a, b > 0$ où $\Gamma(.)$ est une fonction gamma, le moment d'ordre k de X peut être écrit comme

$$E(X^k) = \beta B(\beta - \frac{k}{\theta}, 1 + \frac{k}{\theta}) = \frac{\Gamma(\beta - \frac{k}{\theta})\Gamma(1 + \frac{k}{\theta})}{\Gamma(\beta)}, \quad k < \theta\beta \quad (4.4)$$

A partir de l'équation (4.4), pour $k = 1$, on obtient le premier moment qui est appelé moyenne et noté par μ . Par conséquent, la moyenne de X est

$$\mu = E(X) = \beta B(\beta - \frac{1}{\theta}, 1 + \frac{1}{\theta}) = \frac{\Gamma(\beta - \frac{1}{\theta})\Gamma(\frac{1}{\theta})}{\theta\Gamma(\beta)} \quad \theta\beta > 1$$

Pour $k = 2$, on obtient le second moment $E(X^2)$. D'autre part, le deuxième moment central ou la variance $\sigma^2 = E(X^2) - \mu^2$, par conséquent, nous obtenons

$$Var(X) = \sigma^2 = \frac{\Gamma(\beta - \frac{2}{\theta})\Gamma(1 + \frac{2}{\theta})\Gamma(\beta) - \Gamma^2(\beta - \frac{1}{\theta})\Gamma^2(1 + \frac{1}{\theta})}{\Gamma^2(\beta)} \quad \theta\beta > 2$$

Les mesures d'asymétrie (ou Skewness) et d'aplatissement (ou Kurtosis) peuvent main-

tenant être calculées à l'aide des relations suivantes ;

$$Skewness(X) = \frac{2E(X^3) - 3E(X)E(X^2) + E(X^3)}{Var(X)^{3/2}} \quad (4.5)$$

$$Kurtosis(X) = \frac{-3E(X)^4 + 6E(X)^2E(X^2) - 4E(X)E(X^3) + E(X^4)}{Var(X)^2} \quad (4.6)$$

La fonction de survie correspondante est donnée par :

$$S(t) = 1 - F(t; \theta, \beta) = (1 + t^\theta)^{-\beta} \quad (4.7)$$

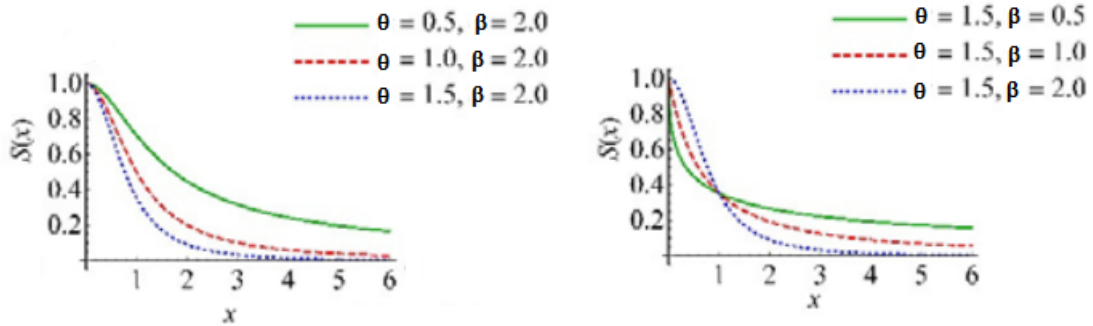


FIGURE 4.3 – La fonction de survie du Burr XII (θ, β)

Le taux de hasard à l'instant t est exprimé comme :

$$h(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = \frac{\theta\beta t^{\theta-1}}{1 + t^\theta} \quad (4.8)$$

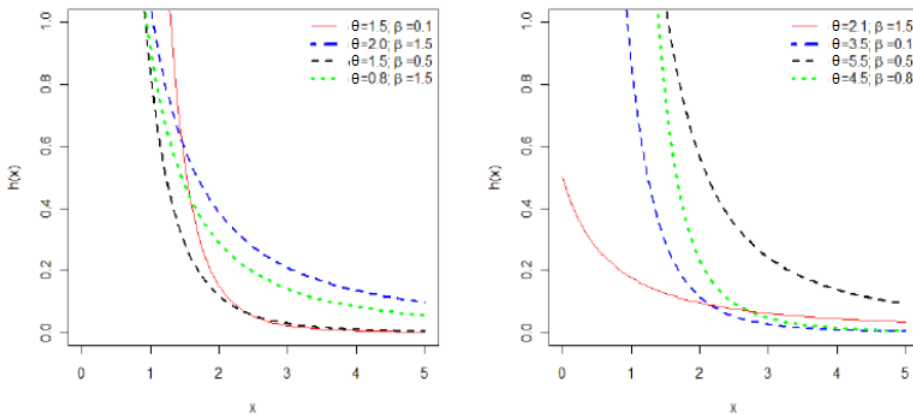


FIGURE 4.4 – Le taux de hasard du Burr XII (θ, β)

4.1.2 Estimation des paramètres

Beaucoup des travaux sont faits sur le modèle de Burr, présentant ses propriétés statistiques et la relation avec les autres distributions qui est utilisés dans l'analyse de fiabilité. L'estimation ponctuelle et par intervalle des paramètres de la loi de Burr XII est obtenue par la méthode du maximum de vraisemblance sur les données censurées et complètes. [21],[50].

Les estimateurs de Bayes, du maximum de vraisemblance sur les données censurées progressives de type II sont mentionnés dans [23],[7] utilisant les fonctions de perte symétriques et asymétriques.

4.2 Distribution de Burr XII à trois paramètres

Soit X une variable aléatoire ayant une distribution de Burr XII (α, θ, β) , Tejeda et Barry (2008)[48], la densité de cette loi est donnée par

$$f(x; \alpha, \theta, \beta) = \theta \beta \alpha^{-\theta} x^{\theta-1} \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-(\beta+1)}, \quad x > 0, \alpha, \theta, \beta > 0, \quad (4.9)$$

La fonction de répartition est comme suit :

$$F(x; \alpha, \theta, \beta) = \int_{-\infty}^{\infty} f(x; \alpha, \theta, \beta) = 1 - \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-\beta}, \quad x > 0, \alpha, \theta, \beta > 0. \quad (4.10)$$

Les figures 4.5 représentent des formes possibles de la densité et la fonction de répartition pour les valeurs sélectionnées des paramètres α pour $\theta = 5$ et $\beta = 2$.

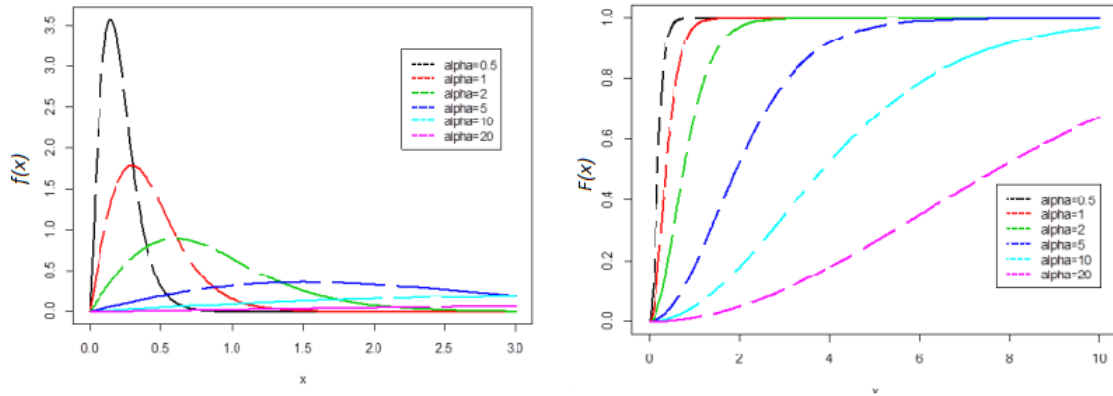


FIGURE 4.5 – La densité et la fonction de répartition de Burr XII $(\alpha, 5, 2)$

4.2.1 Propriétés statistiques et mathématiques

Les quantiles $Q(p) = F^{-1}(p)$; pour $(x > 0, \alpha, \theta, \beta > 0)$.

$$\begin{aligned}
 1 - \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-\beta} &= p \\
 1 + \left(\frac{x}{\alpha} \right)^\theta &= (1 - p)^{-\frac{1}{\beta}} \\
 x^\theta &= [(1 - p)^{-\frac{1}{\beta}} - 1] \alpha^\theta \\
 x &= \alpha [(1 - p)^{-\frac{1}{\beta}} - 1]^{\frac{1}{\theta}}
 \end{aligned}$$

donc

$$F^{-1}(p) = Q(p) = \alpha [(1 - p)^{-\frac{1}{\beta}} - 1]^{\frac{1}{\theta}} \quad (4.11)$$

Q_1, Q_3 , peuvent être obtenus par remplacer $p = 0,25$ et $p = 0,75$ dans l'équation précédente.

Comme la médiane est $Q(\frac{1}{2}) = F^{-1}(\frac{1}{2})$,

$$Md = \alpha [2^{\frac{1}{\beta}} - 1]^{\frac{1}{\theta}}$$

Le mode de la distribution Burr XII (α, θ, β) est

$$\begin{aligned}\frac{df(x; \alpha, \theta, \beta)}{dx} &= \left[\theta \beta \alpha^{-\theta} x^{\theta-1} \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-(\beta+1)} \right] = 0 \\ \theta \beta \alpha^{-2\theta} x^{\theta-2} \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-(\beta+2)} [\alpha^\theta(\theta-1) - x^\theta(1+\theta\beta)] &= 0\end{aligned}$$

comme $x > 0, \alpha > 0$ $x^\theta = \frac{\alpha^\theta(\theta-1)}{1+\theta\beta}$ d'où

$$Mo = \left(\frac{\alpha^\theta(\theta-1)}{1+\theta\beta} \right)^{1/\theta} \quad \theta\beta > 1, \alpha > 0, \theta > 0, \beta > 0$$

Le moment d'ordre k de la distribution Burr XII (α, θ, β)

$$E(X^k) = \int_{-\infty}^{+\infty} x^k f(x; \alpha, \theta, \beta) dx = \beta \int_0^{+\infty} x^k \theta \alpha^{-\theta} x^{\theta-1} \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-(\beta+1)} dx \quad (4.12)$$

Soit $u = \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-1}$ d'où

$x = \alpha [(1-u)/u]^{1/\theta}$, alors,

$du = -\theta \alpha^{-\theta} x^{\theta-1} \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-2} dx = -\theta \alpha^{-\theta} x^{\theta-1} u^2 dx$ donc,

$\theta \alpha^{-\theta} x^{\theta-1} dx = -u^{-2} du$,

alors, l'équation (4.12) devient (le même travail que (4.3))

$$E(X^k) = \beta \alpha^k \int_0^1 u^{(\beta - \frac{k}{\theta})-1} (1-u)^{(1 + \frac{k}{\theta})-1} du. \quad (4.13)$$

Le moment d'ordre k de X peut être écrit comme

$$E(X^k) = \beta \alpha^k B\left(\beta - \frac{k}{\theta}, 1 + \frac{k}{\theta}\right) = \alpha^k \frac{\Gamma(\beta - \frac{k}{\theta}) \Gamma(1 + \frac{k}{\theta})}{\Gamma(\beta)}, \quad k < \theta\beta. \quad (4.14)$$

La moyenne μ de la distribution de Burr XII (α, θ, β) est

$$\mu = E(X) = \beta \alpha B\left(\beta - \frac{1}{\theta}, 1 + \frac{1}{\theta}\right) = \alpha \frac{\Gamma(\beta - \frac{1}{\theta}) \Gamma(\frac{1}{\theta})}{\theta \Gamma(\beta)}, \quad 1 < \theta\beta \quad (4.15)$$

La variance σ^2 est

$$Var(X) = \sigma^2 = \alpha^2 \left(\frac{\Gamma(\beta - \frac{2}{\theta})\Gamma(1 + \frac{2}{\theta})\Gamma(\beta) - \Gamma^2(\beta - \frac{1}{\theta})\Gamma^2(1 + \frac{1}{\theta})}{\Gamma^2(\beta)} \right) \quad \theta\beta > 2$$

Les fonctions de survie et le taux de hasard sont données respectivement par

$$S(x; \alpha, \theta, \beta) = \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-\beta}, \quad x > 0, \alpha, \theta, \beta > 0, \quad (4.16)$$

$$h(x; \alpha, \theta, \beta) = \theta\beta\alpha^{-\theta}x^{\theta-1} \left[1 + \left(\frac{x}{\alpha} \right)^\theta \right]^{-1}, \quad x > 0, \alpha, \theta, \beta > 0. \quad (4.17)$$

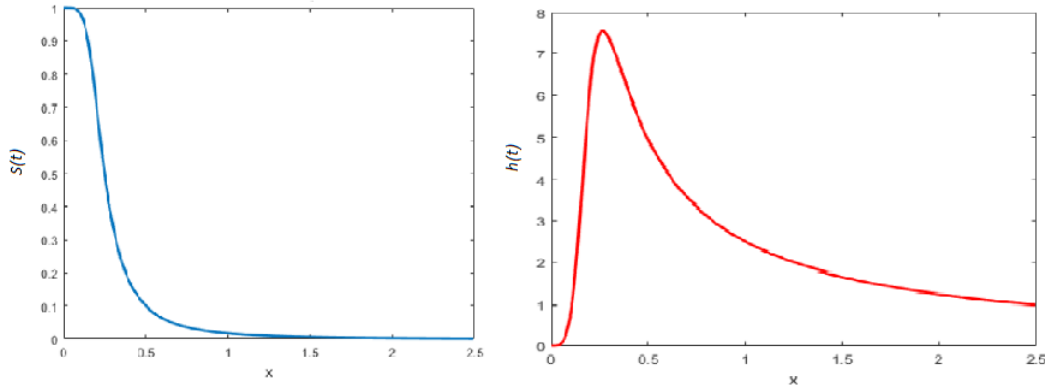


FIGURE 4.6 – La fonction de survie et le taux de hasard de Burr XII (0.2, 5, 0.5)

Dans ce qui suit, pour plus de commodité, on désigne $\Phi_{\alpha,\theta}(x) := 1 + \left(\frac{x}{\alpha}\right)^\theta$, ce qui nous permet de réécrire la densité, la fonction de répartition et la fonction de survie dans (4.9), (4.10) et (4.16) comme suit,

$$f(x; \alpha, \theta, \beta) = \theta\beta\alpha^{-\theta}x^{\theta-1}[\Phi_{\alpha,\theta}(x)]^{-(\beta+1)}, \quad (4.18)$$

$$F(x; \alpha, \theta, \beta) = 1 - [\Phi_{\alpha,\theta}(x)]^{-\beta}, \quad (4.19)$$

$$S(x; \alpha, \theta, \beta) = [\Phi_{\alpha,\theta}(x)]^{-\beta}. \quad (4.20)$$

4.3 Estimation des paramètres par le maximum de vraisemblance en présence de données censurées progressives de type II

On s'intéresse au temps de défaillance de n unités qui sont indépendantes, un n-échantillon (X_1, X_2, \dots, X_n) généré à partir de la loi (4.9). Soit $m \in \mathbb{N}$ fixé ($m < n$), les m variables aléatoires $X_{1,n,m}, \dots, X_{m,n,m}$ sont progressivement censurés de type II, où $R = (r_1, r_2, \dots, r_m)$ est le schéma de censure.

On obtient la fonction de vraisemblance de $(X_{i,m})_{i=1\dots m}$ en remplaçant (4.18) et (4.20) dans (1.3),

$$L(\underline{x}, \alpha, \theta, \beta) = A \prod_{i=1}^m \theta \beta \alpha^{-\theta} x_i^{\theta-1} [\Phi_{\alpha,\theta}(x_i)]^{-(\beta+1)} [\Phi_{\alpha,\theta}(x_i)]^{\beta r_i}$$

$$L(\underline{x}, \alpha, \theta, \beta) = A \theta^m \beta^m \alpha^{-\theta m} \prod_{i=1}^m x_i^{\theta-1} [\Phi_{\alpha,\theta}(x_i)]^{-(\beta+1+\beta r_i)}, \quad (4.21)$$

tel que $\underline{x} = (x_{(1)}, \dots, x_{(m)})$ est le vecteur des observations ordonnées.

Ainsi, le logarithme de la vraisemblance $l(x) = \ln L(x, \alpha, \theta, \beta)$ est proportionnelle à

$$l(x) \propto \ln A + m \ln \theta + m \ln \beta - \theta m \ln \alpha + (\theta - 1) \sum_{i=1}^m \ln x_i - \sum_{i=1}^m (\beta + 1 + \beta r_i) \ln [\Phi_{\alpha,\theta}(x_i)].$$

Les estimateurs du maximum de vraisemblance des paramètres inconnus α, θ et β sont les solutions du système non linéaire

$$\begin{cases} \frac{\partial l(x)}{\partial \alpha} = -\frac{\theta m}{\alpha} - \sum_{i=1}^m (\beta + 1 + \beta r_i) \frac{\frac{\partial \Phi_{\alpha,\theta}(x_i)}{\partial \alpha}}{\Phi_{\alpha,\theta}(x_i)} = 0, \\ \frac{\partial l(x)}{\partial \theta} = \frac{m}{\theta} + \sum_{i=1}^m \ln \frac{x_i}{\alpha^m} - \sum_{i=1}^m (\beta + 1 + \beta r_i) \frac{\frac{\partial \Phi_{\alpha,\theta}(x_i)}{\partial \theta}}{\Phi_{\alpha,\theta}(x_i)} = 0, \\ \frac{\partial l(x)}{\partial \beta} = \frac{m}{\beta} - \sum_{i=1}^m (1 + r_i) \ln [\Phi_{\alpha,\theta}(x_i)] = 0, \end{cases} \quad (4.22)$$

tel que

$$\frac{\partial \Phi_{\alpha,\theta}(x_i)}{\partial \alpha} = -\frac{\theta x_i}{\alpha^{\theta+1}} \quad \text{and} \quad \frac{\partial \Phi_{\alpha,\theta}(x_i)}{\partial \theta} = \left(\frac{x_i}{\alpha}\right)^{\theta} \ln \left(\frac{x_i}{\alpha}\right).$$

Le système (4.22) ne peut pas être résolu analytiquement. Ainsi, on calcule des ap-

proximations des solutions numériquement pour obtenir les estimateurs de vraisemblance $\tilde{\alpha}_{MV}$, $\tilde{\theta}_{MV}$ et $\tilde{\beta}_{MV}$ des paramètres inconnus α , θ et β respectivement.

4.4 Estimation Bayésienne sous différentes fonctions de perte en présence de données censurées progressives de type II

Dans cette section, on discutera de l'estimation des paramètres inconnues par l'approche Bayésienne au cas de données progressivement censurées.

On suppose que nous avons une information préalable (ou une distribution a priori) sur les paramètres à estimer. Ici, on suppose que les paramètres α , θ et β ont une distribution à priori gamma, c'est-à-dire

$$\pi_1(\alpha) = \frac{a_1^{b_1}}{\Gamma(b_1)} \alpha^{b_1-1} \exp(-b_1\alpha), \quad \alpha > 0, \quad (4.23)$$

$$\pi_2(\theta) = \frac{a_2^{b_2}}{\Gamma(b_2)} \theta^{b_2-1} \exp(-b_2\theta), \quad \theta > 0, \quad (4.24)$$

$$\pi_3(\beta) = \frac{a_3^{b_3}}{\Gamma(b_3)} \beta^{b_3-1} \exp(-b_3\beta), \quad \beta > 0, \quad (4.25)$$

où $a_1, a_2, a_3, b_1, b_2, b_3 > 0$ sont des constants connus.

En supposant que les paramètres α, θ et β sont indépendants, la distribution a priori conjointe est donnée par

$$\pi(\alpha, \theta, \beta) = \frac{a_1^{b_1} a_2^{b_2} a_3^{b_3}}{\Gamma(b_1)\Gamma(b_2)\Gamma(b_3)} \alpha^{b_1-1} \theta^{b_2-1} \beta^{b_3-1} \exp(-b_1\alpha - b_2\theta - b_3\beta). \quad (4.26)$$

Maintenant, on utilise la formule de Bayes combinée avec (1.3) pour dériver la distribution a posteriori conjointe pour α, θ et β

$$\pi(\alpha, \theta, \beta | \underline{x}) = \frac{L(\underline{x} | \alpha, \theta, \beta) \pi(\alpha, \theta, \beta)}{\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} L(\underline{x} | \alpha, \theta, \beta) \pi(\alpha, \theta, \beta) d\alpha d\theta d\beta}$$

$$= \frac{\theta^m \beta^m \alpha^{-\theta m} \prod_{i=1}^m x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \alpha^{a_1-1} \theta^{a_2-1} \beta^{a_3-1} \exp(-b_1 \alpha - b_2 \theta - b_3 \beta)}{\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \theta^m \beta^m \alpha^{-\theta m} \prod_{i=1}^m x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \alpha^{a_1-1} \theta^{a_2-1} \beta^{a_3-1} \exp(-b_1 \alpha - b_2 \theta - b_3 \beta)}$$

On note

$$\Psi_m(\alpha, \theta, \beta) := \alpha^{-\theta m + a_1 - 1} \theta^{m + a_2 - 1} \beta^{m + a_3 - 1} \exp(-b_1 \alpha - b_2 \theta - b_3 \beta)$$

et

$$J := \left(\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right)^{-1}.$$

$$\pi(\alpha, \theta, \beta | \underline{x}) \propto J \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right]. \quad (4.27)$$

Pour obtenir les estimateurs Bayésiens des paramètres inconnus α, θ et β on choisit trois différentes fonctions de perte (quadratique généralisée, linex, entropie).1.1

1. Sous fonction de perte quadratique généralisée GQ , les estimateurs Bayésiens de α, θ et β ; Notés respectivement par $\hat{\alpha}_{GQ}, \hat{\theta}_{GQ}$ et $\hat{\beta}_{GQ}$ sont,

$$\begin{cases} \hat{\alpha}_{GQ} = J_1 \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \alpha^\gamma \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} \left(1 + \left(\frac{x_i}{\alpha} \right)^\theta \right)^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \\ \hat{\theta}_{GQ} = J_2 \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \theta^\gamma \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} \left(1 + \left(\frac{x_i}{\alpha} \right)^\theta \right)^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \\ \hat{\beta}_{GQ} = J_3 \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \beta^\gamma \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} \left(1 + \left(\frac{x_i}{\alpha} \right)^\theta \right)^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \end{cases} \quad (4.28)$$

où

$$\begin{cases} J_1 = \left(\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \alpha^{\gamma-1} \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right)^{-1} \\ J_2 = \left(\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \theta^{\gamma-1} \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right)^{-1} \\ J_3 = \left(\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \beta^{\gamma-1} \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right)^{-1}. \end{cases}$$

Alors, les risques a posteriori correspondants sont

$$\begin{cases} PR(\hat{\alpha}_{GQ}) = E_{\pi}(\alpha^{\gamma+1}) - 2\hat{\alpha}_{GQ}E_{\pi}(\alpha^{\gamma}) + \hat{\alpha}_{GQ}^2E_{\pi}(\alpha^{\gamma-1}), \\ PR(\hat{\theta}_{GQ}) = E_{\pi}(\theta^{\gamma+1}) - 2\hat{\theta}_{GQ}E_{\pi}(\theta^{\gamma}) + \hat{\theta}_{GQ}^2E_{\pi}(\theta^{\gamma-1}), \\ PR(\hat{\beta}_{GQ}) = E_{\pi}(\beta^{\gamma+1}) - 2\hat{\beta}_{GQ}E_{\pi}(\beta^{\gamma}) + \hat{\beta}_{GQ}^2E_{\pi}(\beta^{\gamma-1}). \end{cases} \quad (4.29)$$

2. Sous fonction de perte entropie E , les estimateurs Bayésiens de α, θ et β ; Notés respectivement par $\hat{\alpha}_E, \hat{\theta}_E$ et $\hat{\beta}_E$ sont,

$$\begin{cases} \hat{\alpha}_E = \left[J \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \alpha^{-p} \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right]^{\frac{-1}{p}}, \\ \hat{\theta}_E = \left[J \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \theta^{-p} \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right]^{\frac{-1}{p}}, \\ \hat{\beta}_E = \left[J \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \beta^{-p} \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right]^{\frac{-1}{p}}. \end{cases} \quad (4.30)$$

Alors, les risques a posteriori correspondants sont

$$\begin{cases} PR(\hat{\alpha}_E) = pE_{\pi}(\ln \alpha - \ln \hat{\alpha}_E), \\ PR(\hat{\theta}_E) = pE_{\pi}(\ln \theta - \ln \hat{\theta}_E), \\ PR(\hat{\beta}_E) = pE_{\pi}(\ln \beta - \ln \hat{\beta}_E). \end{cases} \quad (4.31)$$

3. Sous fonction de perte linex L , les estimateurs Bayésiens de α, θ et β ; Notés respectivement par $\hat{\alpha}_L, \hat{\theta}_L$ et $\hat{\beta}_L$ sont,

$$\begin{cases} \hat{\alpha}_L = -\frac{J}{r} \ln \left[\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \exp(-r\alpha) \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right], \\ \hat{\theta}_L = -\frac{J}{r} \ln \left[\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \exp(-r\theta) \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right], \\ \hat{\beta}_L = -\frac{J}{r} \ln \left[\int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} \exp(-r\beta) \Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right] d\alpha d\theta d\beta \right]. \end{cases} \quad (4.32)$$

Alors, les risques a posteriori correspondants sont

$$\begin{cases} PR(\hat{\alpha}_L) = r(\hat{\alpha}_{GQ} - \hat{\alpha}_L), \\ PR(\hat{\theta}_L) = r(\hat{\theta}_{GQ} - \hat{\theta}_L), \\ PR(\hat{\beta}_L) = r(\hat{\beta}_{GQ} - \hat{\beta}_L). \end{cases} \quad (4.33)$$

Les estimateurs Bayésiens dans (4.28), (4.30) et (4.32) ne peuvent pas être calculés analytiquement, pour traiter cela nous proposons ensuite, une procédure MCMC pour les approximer.

La méthode MCMC a été utilisée dans le cadre Bayésien dans de nombreux travaux EL-Sagheer (13), Chadli, Talhi et Fellag (6), Achcar et Leonardo (2), maintenant, afin d'approximer les estimateurs Bayésiens, on doit échantillonner à partir de la distribution a posteriori ;

$$\pi(\alpha, \theta, \beta|x) \propto J\Psi_m(\alpha, \theta, \beta) \prod_{i=1}^m \left[x_i^{\theta-1} [\Phi_{\alpha, \theta}(x_i)]^{-(\beta+1+\beta r_i)} \right].$$

On choisit d'utiliser l'algorithme Metropolis Hastings pour générer l'échantillon a posteriori. Indique que $x := (\alpha, \theta, \beta)$ et partir d'une valeur choisie arbitrairement $x_0 = (\alpha_0, \theta_0, \beta_0)$. On sélectionne la distribution exponentielle comme la distribution instrumentale (ou proposition) $p(x, \cdot)$

Soit X^i la valeur retenue à l'étape i , on propose le candidat $\tilde{x} \rightarrow p(x, \cdot)$

1. Calculer

$$r = \min \left(1, \frac{\pi(\tilde{x})p(\tilde{x}, x)}{\pi(x)p(x, \tilde{x})} \right).$$

2. Générer $U_{[0,1]}$ notant $U^i = u$.

3. Si $r \geq u : X^{i+1} \leftarrow \tilde{x}$, sinon $X^{i+1} \leftarrow x$.

En répétant ces étapes pour $i = \overline{1, M}$, où M où M est le nombre des itérations.

La suite aléatoire $(X_i)_{i \geq 0}$ est une chaîne de Markov, puisque l'état de X_i ne dépend que de X_{i-1} suivant la distribution π .

La figure 4.7 montre que l'histogramme de l'échantillon généré des paramètres

$(\alpha_i, \theta_i, \beta_i), i = 1 \dots, M$ de $\pi(\alpha, \theta, \beta|X)$ en utilisant l'algorithme ci-dessus correspond assez bien à la courbe théorique de la densité a postérieure conjointe.

On exécute la chaîne de Markov pour 10000 itérations initialisées aux paramètres proposés (voir la section ci-dessous). La figure 4.8 illustre les traces et les diagrammes de densité des échantillons de la chaîne de Markov. On remarque que la trace se mélange bien pour les trois paramètres. De plus, à partir de l'intrigue de trace, on voit qu'il n'y

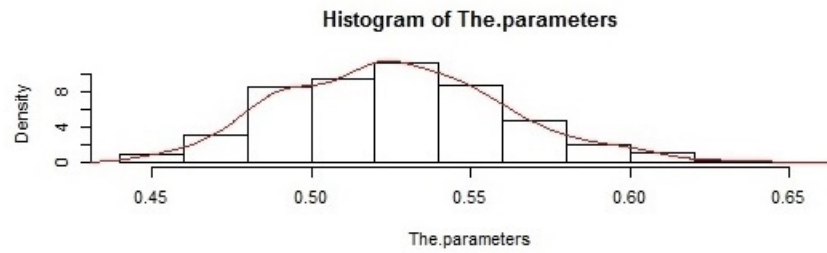


FIGURE 4.7 – L’histogramme des paramètres et la densité a postérieure conjointe a pas un besoin d’un échantillon de brûlure (burn in sample) et le choix des valeurs de départ semble être raisonnable.

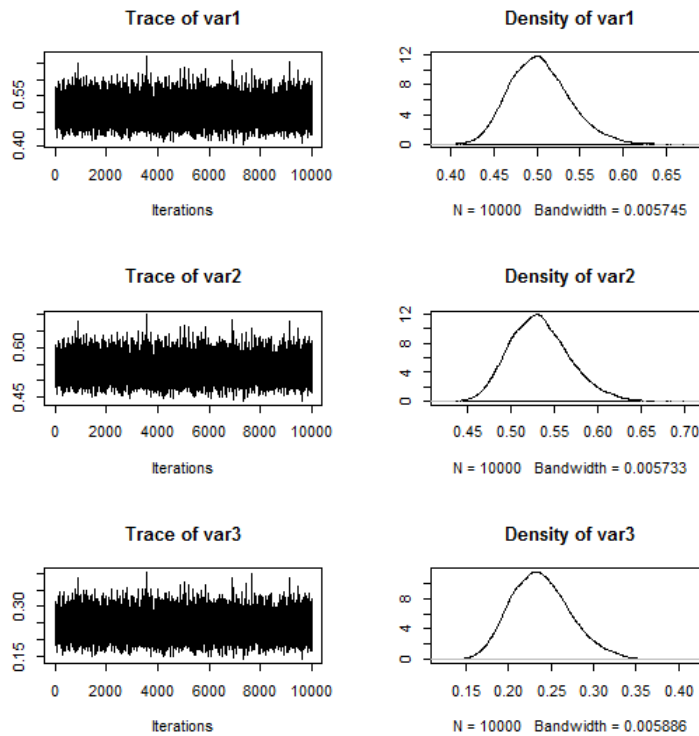


FIGURE 4.8 – La trace et la densité MCMC

4.4.1 Prédiction Bayésienne

on utilisant la densité, la fonction de répartition et la fonction de survie (4.18), (4.19) et (4.20), on obtient

$$g_{(k)}(y_k; \alpha, \theta, \beta) = k \binom{N}{k} \theta \beta \alpha^{-\theta} y_k^{\theta-1} [\Phi_{\alpha, \theta}(x)]^{\beta(N, k)} [1 - [\Phi_{\alpha, \theta}(x)]^{-\beta}]^{k-1}, \quad (4.34)$$

tel que $\beta(N, k) = -\beta(N - k - 1) - 1$, après l'utilisation de l'expansion binomiale, le résultat est

$$g_{(k)}(y_k; \alpha, \theta, \beta) = k \binom{N}{k} \theta \beta \alpha^{-\theta} y_k^{\theta-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} [\Phi_{\alpha, \theta}(x)]^{\beta_j(N, k)}, \quad (4.35)$$

où $\beta_j(N, k) = -\beta(N - k - 1 - j) - 1$.

Ensuite, la densité prédictive Bayésienne de Y_k est donnée par

$$g_{(k)}(y_k|X) = \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} g_{(k)}(y_k; \alpha, \theta, \beta) \pi(\alpha, \theta, \beta|X) d\alpha d\theta d\beta,$$

Lorsque $\pi(\alpha, \theta, \beta|X)$ est la densité a posteriori conjointe vue ci-dessus, il est clair que la densité prédictive Bayésienne ne peut pas être exprimée ou calculée analytiquement. Nous proposons l'utilisation d'un prédicteur naturel $\hat{Y}_k = E(Y_k)$ en utilisant la procédure d'échantillonnage MCMC décrite dans le chapitre précédent, afin qu'on puisse obtenir des estimations basées sur la simulation.

4.5 Simulation

Dans cette section, on étudie la performance des estimateurs mentionnés ci-dessus, en commençant par l'estimateur du maximum de vraisemblance puis l'estimateur Bayésien sous les trois fonctions de perte. Finalement, on compare les deux estimateurs.

Tout d'abord, nous générons un échantillon progressif de type II à partir de notre modèle, c'est-à-dire $\text{Burr}(\alpha, \theta, \beta)$ en appliquant l'algorithme suivant ;

1. Généré m les variables aléatoires indépendantes et identiquement distribuées (iid) $U_1, \dots, U_m \sim U(0, 1)$.
2. Pour une valeur donnée du schéma de censure $R = (r_1, r_2, \dots, r_m)$ calculer $B_i = U_i^{1/\sum_{j=i}^m (r_j+1)}$, $i = 1, \dots, m$.
3. Soit $V_0 = 1$, et calculer $V_k = B_k V_{k-1}$, $k = 1, \dots, m$.
4. $U_{i;m,n} = (1 - V_i)$, $i = 1, \dots, m$.
5. Ensuite, pour des valeurs données de α, θ et β . $X_{i;m,n} = \alpha \left((1 - U_{r;m,n})^{-\frac{1}{\beta}} - 1 \right)^{\frac{1}{\theta}}$, $i = 1, \dots, m$ est un échantillon censuré progressif de type II suit la lois de Burr XII(α, θ, β)

On effectue une étude de simulation en utilisant différentes tailles d'échantillons, $n = 30, 40$ en faisant correspondre différentes tailles d'échantillons effectifs, $m = 20, 30$, de plus différents schémas de censure, on prend $\alpha = 0.5, \theta = 0.5$ et $\beta = 0.25$. Les résultats sont obtenus après $M = 10000$ générations d'échantillons, pour $t = 10$ donc $S(t) = 0.6538$ et $h(t) = 0.0102$.

4.5.1 Résultats de l'estimation

1. Dans le tableau suivant, les valeurs résultantes sont affichées après l'utilisation du langage de programmation \mathcal{R} , plus précisément on a utilisé le paquet BB , abréviation de Barzilai-Brown qui est connu pour sa grande capacité à résoudre des systèmes non linéaires à grande échelle, pour plus de détails, on se réfère à Varadhan et Gilbert [49]

TABLE 4.1 – Estimations des paramètre à l'aide du maximum de vraisemblance (MV) avec leurs erreurs Quadratiques.

n	m	schémas	paramétré	MV	
30	20	$(5^1, 10^0, 5^1)$	α	0.6563	(0.2443)
			θ	0.5945	(0.0893)
			β	0.2135	(0.0133)
			$S(t)$	0.6823	(0.0081)
			$h(t)$	0.0160	$(1.2893.10^{-6})$
		$(10, 19^0)$	α	0.6961	(0.3848)
			θ	0.5711	(0.0506)
			β	0.2073	(0.0182)
			$S(t)$	0.7006	(0.2184)
			$h(t)$	0.0097	$(2.2849.10^{-6})$
40	20	$(5^2, 10^0, 5^2)$	α	0.6853	(0.3433)
			θ	0.5269	(0.0072)
			β	0.2665	(0.0027)
			$S(t)$	0.6213	(0.0106)
			$h(t)$	0.0090	$(1.4387.10^{-5})$
		$(20, 19^0)$	α	0.5302	(0.0091)
			θ	0.4656	(0.0118)
			β	0.3327	(0.0684)
			$S(t)$	0.5486	(0.1106)
			$h(t)$	0.0102	$(2.100.10^{-9})$
40	30	$(10^0, 10^1, 10^0)$	α	0.4881	(0.0014)
			θ	0.5136	(0.0018)
			β	0.3091	(0.0349)
			$S(t)$	0.6080	(0.0210)
			$h(t)$	0.0104	$(4.7001.10^{-7})$
		$(10, 29^0)$	α	0.4555	(0.0198)
			θ	0.4485	(0.0265)
			β	0.3176	(0.0457)
			$S(t)$	0.6243	(0.0087)
			$h(t)$	0.0090	$(0.4111.10^{-5})$

2. Les tableaux suivants affichent l'estimation Bayésienne sous les fonctions de perte quadratique généralisée, d'entropie et de Linex, en utilisant la procédure MCMC de la section 4, pour information préalable, on choisit $a_1 = a_2 = a_3 = 1$, et $b_1 = b_2 = 2, b_3 = 4$. Ces paramètres ont été choisis de manière à ce que la moyenne à priori devienne la valeur exceptée du paramètre.

TABLE 4.2 – Estimation Bayésienne sous la perte quadratique généralisée avec le risque à posteriori (entre parenthèses).

n	m	schémas	paramètre	Estimateur Bayésien (GQ)			
				$\gamma = -2$		$\gamma = 2$	
30	20	$(5^1, 10^0, 5^1)$	α	0.5272	(0.0014)	0.5117	(0.0222)
			θ	0.5563	(0.0025)	0.5418	(0.0398)
			β	0.2675	(0.0006)	0.5025	(0.1617)
			$S(t)$	0.6135	(0.0007)	0.4062	(0.0306)
			$h(t)$	0.0125	$(2.5210 \cdot 10^{-6})$	0.0226	$(7.7826 \cdot 10^{-5})$
		$(10, 19^0)$	α	0.5024	(0.0010)	0.4596	(0.0155)
			θ	0.5316	(0.0013)	0.5195	(0.0215)
			β	0.2424	(0.0006)	0.4828	(0.1568)
			$S(t)$	0.6503	$(6.2032 \cdot 10^{-6})$	0.4281	(0.0288)
			$h(t)$	0.0107	$(1.1777 \cdot 10^{-7})$	0.0208	$(5.5523 \cdot 10^{-5})$
40	20	$(5^2, 10^0, 5^2)$	α	0.5092	(0.0005)	0.5023	(0.0075)
			θ	0.5386	(0.0011)	0.5321	(0.0180)
			β	0.2460	(0.0002)	0.4967	(0.0608)
			$S(t)$	0.6443	$(4.5440 \cdot 10^{-5})$	0.4138	(0.0283)
			$h(t)$	0.0110	$(3.3240 \cdot 10^{-7})$	0.0220	$(6.8914 \cdot 10^{-5})$
		$(20, 19^0)$	α	0.5027	(0.0009)	0.4888	(0.0149)
			θ	0.5314	(0.0013)	0.5185	(0.0208)
			β	0.2481	(0.0006)	0.4766	(0.1605)
			$S(t)$	0.6438	$(5.0231 \cdot 10^{-5})$	0.4332	(0.0243)
			$h(t)$	0.0111	$(2.6986 \cdot 10^{-7})$	0.0204	$(5.2286 \cdot 10^{-5})$
40	30	$(10^0, 10^1, 10^0)$	α	0.5075	(0.0004)	0.5007	(0.0057)
			θ	0.5368	(0.0010)	0.5304	(0.0172)
			β	0.2445	(0.0002)	0.4964	(0.0623)
			$S(t)$	0.6465	$(2.6773 \cdot 10^{-5})$	0.4145	(0.0286)
			$h(t)$	0.0109	$(2.1789 \cdot 10^{-7})$	0.0219	$(6.7808 \cdot 10^{-5})$
		$(10, 29^0)$	α	0.5063	(0.0003)	0.5024	(0.0043)
			θ	0.5357	(0.0009)	0.5320	(0.0138)
			β	0.2426	(0.0001)	0.4998	(0.0384)
			$S(t)$	0.6490	$(1.1664 \cdot 10^{-5})$	0.4155	(0.0293)
			$h(t)$	0.0108	$(1.7563 \cdot 10^{-7})$	0.0221	$(7.0492 \cdot 10^{-5})$

TABLE 4.3 – Estimation Bayésienne sous la perte entropie avec le risque a posteriori (entre parenthèses).

n	m	schémas	paramètre	Estimateur Bayésien (Entropie)			
				$p = -2$		$p = 2$	
30	20	$(5^1, 10^0, 5^1)$	α	0.5404	(0.1554)	0.5375	(0.1448)
			θ	0.5699	(0.2617)	0.5672	(0.2521)
			β	0.2755	(0.1945)	0.2694	(0.1497)
			$S(t)$	0.6029	(0.0013)	0.6104	(0.0009)
			$h(t)$	0.0132	$(4.4529 \cdot 10^{-6})$	0.0128	$(3.4332 \cdot 10^{-6})$
		$(10, 19^0)$	α	0.5195	(0.0767)	0.5153	(0.0603)
			θ	0.5488	(0.1863)	0.5448	(0.0187)
			β	0.2566	(0.0522)	0.2477	(0.0187)
			$S(t)$	0.6296	(0.0003)	0.6408	$((8.5156 \cdot 10^{-5})$
			$h(t)$	0.1176	$(1.1969 \cdot 10^{-6})$	0.0113	$(5.4218 \cdot 10^{-7})$
40	20	$(5^2, 10^0, 5^2)$	α	0.5253	(0.0987)	0.5177	(0.0696)
			θ	0.5547	(0.0276)	0.5474	(0.1813)
			β	0.2618	(0.0926)	0.2495	(0.0038)
			$S(t)$	0.6221	(0.0005)	0.6382	(0.0001)
			$h(t)$	0.0121	$(1.8769 \cdot 10^{-6})$	0.0114	$(7.0722 \cdot 10^{-7})$
		$(20, 19^0)$	α	0.5192	(0.0754)	0.5182	(0.0714)
			θ	0.5489	(0.1865)	0.5479	(0.1830)
			β	0.2526	(0.0204)	0.2480	(0.0162)
			$S(t)$	0.6341	(0.0001)	0.6396	(0.0001)
			$h(t)$	0.0116	$(9.2919 \cdot 10^{-7})$	0.0113	$(6.3907 \cdot 10^{-7})$
40	30	$(10^0, 10^1, 10^0)$	α	0.5290	(0.1129)	0.5247	(0.0964)
			θ	0.5583	(0.0266)	0.5542	(0.0258)
			β	0.2660	(0.1244)	0.2578	(0.0615)
			$S(t)$	0.6165	(0.0007)	0.6268	(0.0003)
			$h(t)$	0.0124	$(2.4800 \cdot 10^{-6})$	0.0120	$(1.5096 \cdot 10^{-6})$
		$(10, 29^0)$	α	0.5387	(0.0490)	0.5318	(0.0233)
			θ	0.5681	(0.0255)	0.5616	(0.0232)
			β	0.2750	(0.0907)	0.2612	(0.0878)
			$S(t)$	0.6039	(0.0012)	0.6210	(0.0005)
			$h(t)$	0.0131	$(4.2347 \cdot 10^{-6})$	0.0123	$(2.759 \cdot 10^{-6})$

TABLE 4.4 – Estimation Bayésienne sous la perte Linex avec le risque a posteriori (entre parenthèses).

n	m	schémas	paramètre	Estimateur Bayésien (Linex)			
				$r = -2$		$r = 2$	
30	20	$(5^1, 10^0, 5^1)$	α	0.5252	(0.0040)	0.5207	(0.0179)
			θ	0.5546	(0.0035)	0.5501	(0.0167)
			β	0.2605	(0.0140)	0.2556	(0.1938)
			$S(t)$	0.6236	(0.0005)	0.6304	(0.0003)
			$h(t)$	0.0121	$(1.7570.10^{-6})$	0.0117	$(1.1771.10^{-6})$
		$(10, 19^0)$	α	0.5005	(0.0037)	0.4966	(0.0142)
			θ	0.5300	(0.0032)	0.5262	(0.0133)
			β	0.2354	(0.0140)	0.2310	(0.1035)
			$S(t)$	0.6588	$(1.2320.10^{-5})$	0.6648	$((6.0337.10^{-5})$
			$h(t)$	0.0103	$(9.8531.10^{-9})$	0.0101	$(9.2968.10^{-9})$
40	20	$(5^2, 10^0, 5^2)$	α	0.5083	(0.0017)	0.5067	(0.0085)
			θ	0.5378	(0.0015)	0.5361	(0.0080)
			β	0.2432	(0.0056)	0.2414	(0.1104)
			$S(t)$	0.6477	$(1.8845.10^{-5})$	0.6501	$(6.6034.10^{-6})$
			$h(t)$	0.0109	$(2.2602.10^{-7})$	0.0108	$(1.5187.10^{-7})$
		$(20, 19^0)$	α	0.5009	(0.0037)	0.4972	(0.0168)
			θ	0.5298	(0.0032)	0.5262	(0.0153)
			β	0.2406	(0.0150)	0.2359	(0.1816)
			$S(t)$	0.6528	$(5.1062.10^{-7})$	0.6592	$(1.4517.10^{-5})$
			$h(t)$	0.0158	$(6.7019.10^{-8})$	0.0103	$(2.4685.10^{-9})$
40	30	$(10^0, 10^1, 10^0)$	α	0.5066	(0.0018)	0.5048	(0.0051)
			θ	0.5360	(0.0016)	0.5342	(0.0076)
			β	0.2417	(0.0056)	0.2399	(0.1129)
			$S(t)$	0.6499	$(7.6542.10^{-6})$	0.6524	$(9.6471.10^{-7})$
			$h(t)$	0.0107	$(1.5765.10^{-7})$	0.0107	$(9.6471.10^{-8})$
		$(10, 29^0)$	α	0.5058	(0.0020)	0.5048	(0.0047)
			θ	0.5353	(0.0009)	0.5342	(0.0044)
			β	0.2410	(0.0032)	0.2399	(0.1186)
			$S(t)$	0.6510	$(4.1225.10^{-6})$	0.6523	$(1.0245.10^{-6})$
			$h(t)$	0.0107	$(1.3054.10^{-7})$	0.0107	$(2.7466.10^{-8})$

Discussion : À partir des tableaux ci-dessus, on remarque ce qui suit

- La plus petite erreur quadratique se produit lorsque n et m sont grands.
- La valeur $\gamma = -2$ associée à la fonction de perte quadratique généralisée donne le meilleur risque a posteriori.
- La valeur $p = 2$ associée à la fonction de perte d'entropie donne le meilleur risque a posteriori.

- La valeur $r = -2$ associée à la fonction de perte Linex donne le meilleur risque a posteriori.
- La modification du schéma de censure affecte les résultats.
- La fonction de perte linex donne le plus petit risque à posteriori quand il s'agit de $h(t)$.
- La fonction de perte quadratique généralisée donne le plus petit risque a posteriori parmi les trois fonctions de perte.
- On a le plus petit risque a posteriori lorsque n et m sont grandes.

4.5.2 Comparaison des estimations

Dans le tableau suivant, on compare les estimateurs Bayésiens sous les trois fonctions de perte lorsqu'il fonctionne le mieux, c'est-à-dire lorsque, $\gamma = -2$, $p = 2$ et $r = -2$ avec l'estimateur du maximum de vraisemblance.

TABLE 4.5 – Comparaison des estimateurs en utilisant le critère de Pitman.

n	m	schémas	paramètre	$GQ(\gamma = -2)$	Entropie($p = 2$)	Linex ($r = -2$)
30	20	$(5^1, 10^0, 5^1)$	α	0.999	0.989	0.909
			θ	0.989	0.990	0.997
			β	0.652	0.601	0.745
			$S(t)$	0.100	0.062	0.116
			$h(t)$	0.614	0.341	0.911
		$(10, 19^0)$	α	0.068	0.036	0.055
			θ	0.881	0.681	0.842
			β	0.546	0.531	0.651
			$S(t)$	0.640	0.472	0.701
			$h(t)$	0.512	0.290	0.935
40	20	$(5^2, 10^0, 5^2)$	α	0.933	0.793	0.905
			θ	0.990	0.990	0.991
			β	0.989	0.980	0.994
			$S(t)$	0.905	0.981	0.970
			$h(t)$	0.933	0.677	0.909
		$(20, 19^0)$	α	0.984	0.952	0.977
			θ	0.999	0.990	0.992
			β	0.979	0.981	0.959
			$S(t)$	0.997	0.978	0.994
			$h(t)$	0.931	0.664	0.900
40	30	$(10^0, 10^1, 10^0)$	α	0.939	0.780	0.916
			θ	0.604	0.283	0.508
			β	0.639	0.586	0.736
			$S(t)$	0.413	0.273	0.467
			$h(t)$	0.122	0.059	0.099
		$(10, 29^0)$	α	0.716	0.707	0.715
			θ	0.716	0.716	0.716
			β	0.458	0.422	0.537
			$S(t)$	0.338	0.238	0.301
			$h(t)$	0.200	0.093	0.334

Le tableau suivant affiche l'erreur quadratique moyenne intégrée pour tous les estimateurs, la même que ci-dessus lorsque $\gamma = -2, p = 2$ et $r = -2$ pour les estimateurs Bayésiens.

TABLE 4.6 – Comparaison IMSE .

n	m	schémas	paramètre	MV	$GQ(\gamma = -2)$	$En(p = 2)$	$Li(r = -2)$
30	20	$(5^1, 10^0, 5^1)$	α	0.02063	0.00186	0.00392	0.00221
			θ	0.05847	0.00341	0.00724	0.00430
			β	0.00189	0.00241	0.00380	0.00195
			$S(t)$	0.00208	0.00037	0.00800	0.00353
			$h(t)$	$5.8703.10^{-6}$	$9.0338.10^{-6}$	$2.9285.10^{-5}$	$9.0227.10^{-7}$
		$(10, 19^0)$	α	0.02393	0.00167	0.00349	0.00197
			θ	0.06285	0.00313	0.00668	0.00395
			β	0.00221	0.00232	0.00338	0.00180
			$S(t)$	0.00245	0.00352	0.00729	0.00324
			$h(t)$	$6.2572.10^{-6}$	$7.7780.10^{-6}$	$2.55816.10^{-5}$	$9.9834.10^{-7}$
40	20	$(5^2, 10^0, 5^2)$	α	0.09605	0.00176	0.00350	0.00204
			θ	0.44161	0.00315	0.00663	0.00395
			β	0.023111	0.00251	0.00337	0.00196
			$S(t)$	0.03636	0.00348	0.00726	0.00351
			$h(t)$	$1.8354.10^{-5}$	$8.0666.10^{-6}$	$2.5285.10^{-5}$	$8.5233.10^{-7}$
		$(20, 19^0)$	α	0.11257	0.00193	0.00380	0.00225
			θ	0.42645	0.00249	0.00709	0.00433
			β	0.03120	0.00344	0.00366	0.00201
			$S(t)$	0.04946	0.00388	0.00781	0.00372
			$h(t)$	$1.754503.10^{-5}$	$9.0657.10^{-6}$	$2.7850.10^{-5}$	$8.9946.10^{-7}$
40	30	$(10^0, 10^1, 10^0)$	α	0.01658	0.00121	0.00286	0.00145
			θ	0.04303	0.00268	0.00590	0.00338
			β	0.00154	0.00176	0.00281	0.00131
			$S(t)$	0.00160	0.00025	0.00620	0.00233
			$h(t)$	$4.1495.10^{-6}$	$5.5116.10^{-6}$	$2.1294.10^{-5}$	$1.6230.10^{-6}$
		$(10, 29^0)$	α	0.08360	0.07186	0.07305	0.07202
			θ	0.10528	0.07287	0.07523	0.07336
			β	0.01924	0.01903	0.01977	0.01870
			$S(t)$	0.12229	0.12327	0.12589	0.12307
			$h(t)$	$5.2454.10^{-5}$	$3.3614.10^{-5}$	$4.4837.10^{-5}$	$3.0816.10^{-5}$

Discussion :

- Lorsque n est petit, on peut voir que les estimateurs Bayésiens des paramètres α, θ et β sont meilleurs que les estimateurs du maximum de vraisemblance. Mais pas toujours pour $S(x)$ et $h(x)$. En outre, il est clair que la fonction de perte quadratique généralisée fournit les meilleures valeurs, mais lorsque n est grande, l'estimateur du maximum de vraisemblance de $S(t)$ et $h(t)$ est meilleur que les estimateurs Bayésiens.
- Lorsque n est petite, les estimateurs Bayésiens pour les trois fonctions de perte fournissent le plus petit IMSE par rapport au l'estimateur du maximum de vraisemblance. Mais lorsque n est grande, tous les estimateurs Bayésiens ne sont pas meilleurs que l'estimateur du maximum de vraisemblance, et nous pouvons remar-

quer que la fonction de perte quadratique généralisée fournit les meilleures valeurs de l'IMSE.

4.6 Application sur des données réelles

Dans cette section, on applique les méthodes proposées à un ensemble de données qui représente les montants de sinistres survenus entre 2015 et 2019 à partir d'un contrat d'assurance automobile, qui nous a été fourni par la société d'assurance automobile S.A.A. On utilise un échantillon complet pour les montants mensuels des sinistres. Les résultats de cette étude peuvent nous donner un nouvel aperçu de la prédiction des montants mensuels qui peuvent survenir dans un proche avenir.

L'ensemble de données se compose de 60 observations données dans le tableau suivant.

TABLE 4.7 – Montants mensuels des sinistres dans la société S.A.A.

Mois \ années	2015	2016	2017	2018	2019
Janvier	108359319.3	91251458.8	126252901.25	123854041.95	136724716.59
Février	108051012.2	108199232.29	111455462.91	106767459.16	155210984.05
Mars	132583400.9	112232638.39	118966778.27	140082764.19	161188834.33
Avril	131682245.2	101411375.93	131760329.15	133995516.96	145708124.7
Mai	138763777.5	112050919.56	121534252.2	134887487.43	174445210.24
Juin	127472451.4	92857094.48	106032084.34	156069160.99	155877782.25
Juillet	157461435.7	104336392.13	130928558.38	164424277.71	150030573.28
Aout	147181634.8	95439179.77	138498736.05	139323594.5	155365194.65
Septembre	133049464.3	87403403.19	114162673.88	121052740.22	131837711.49
Octobre	129306784.6	89656893.21	141433992.68	135579481.66	150634506.45
Novembre	146823384.7	102954049.91	118519128.69	139047972.12	140338111.28
Decembre	130188519.9	93239056.94	178298704.25	177643770.67	136827221.9

Pour confirmer si notre modèle correspond à ces données, on considère certains tests d'ajustement avec certaines distributions principalement connues pour ajuster les ensembles de données, qui sont, la distribution de Weibull à trois paramètres, les distributions de Pareto généralisées et Bêta Pareto, la distribution Log-normale et enfin la burr XII de la distribution de deux paramètres.

Dans cette partie, on choisit les tests Akaike Information Criterion (AIC), Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC) et Hannan-Quinn Information Criterion (HQIC). Indiquant par \bar{l} le logarithme de la vraisemblance correspondant à l'échantillon de taille n , et p le nombre de paramètres,

- AIC= $-2\bar{l} + 2p$
- CAIC= $-2\bar{l} + 2p/(n - p - 1)$
- BIC= $-2\bar{l} + p \log(n)$
- HQIC= $-2\bar{l} + 2p \log[\log(n)]$

De plus, on utilise le test de Cramér-von Mises pour voir dans quelle mesure la fonction de répartition correspond à la distribution empirique de l'ensemble de données.

$$W = \left(\frac{1}{2n} + 1 \right) \left[\sum_{i=1}^n \left(Z_i - \frac{2i-1}{2n} \right)^2 + \frac{1}{12n} \right],$$

où $z_i = F(y_i)$, notant que y_i sont les observations ordonnées. La distribution avec le meilleur ajustement correspond aux plus petits résultats. On utilise le langage de programmation \mathcal{R} pour obtenir les résultats suivants ;

TABLE 4.8 – Critères d'ajustement des données sur les montants des sinistres

Modèles	AIC	CAIC	BIC	HQIC	W
Burr-XII(α, θ, β)	2785.861	2786.29	2792.144	2788.319	15.23076
Pareto généralisée	6808.969	6809.397	6815.252	6811.426	20.1766
Bêta-Pareto	319028.1	319028.9	319036.5	319031.4	19.17514
Log-normale	81482.1	81482.53	81488.38	81484.55	20.041667
Weibull à trois paramètres	1927177	1927177	1927183	1927179	20.16667
Burr-XII(θ, β)	319024.1	319024.3	319028.3	319025.8	19.18277

On peut voir que la loi Burr XII(α, θ, β) convient le mieux à cet ensemble de données. Dans le but d'appliquer les méthodes mentionnées dans les sections ci-dessus, nous sélectionnons au hasard un échantillon censuré progressif de type II de taille $m = 40$ parmi les observations $n = 60$ présentées dans le tableau 4.7. Le schéma de censure et les observations sont présentés dans le tableau suivant.

TABLE 4.9 – Échantillon censuré progressif type II des données.

$x_{i,40,60}$	108051012.2	132583400.9	138763777.5	127472451.4	129306784.6	146823384.7	130188519.9
r_i	0	0	0	2	0	0	0
$x_{i40,60}$	108199232.29	101411375.93	112050919.56	104336392.13	87403403.19	89656893.21	102954049.91
r_i	2	0	0	0	0	2	0
$x_{i40,60}$	126252901.25	111455462.91	131760329.15	121534252.2	106032084.34	130928558.38	138498736.05
r_i	0	0	2	0	0	2	0
$x_{i40,60}$	141433992.68	118519128.69	123854041.95	140082764.19	133995516.96	134887487.43	164424277.71
r_i	0	0	0	2	0	0	0
$x_{i40,60}$	139323594.5	121052740.22	135579481.66	139047972.12	177643770.67	155210984.05	161188834.33
r_i	2	0	0	0	2	0	0
$x_{i40,60}$	145708124.7	174445210.24	150030573.28	131837711.49	150634506.45		
r_i	2	0	0	2	0		

Les résultats de l'estimation Bayésienne à l'aide de l'échantillon au-dessus sont donnés dans les tableaux suivants.

TABLE 4.10 – L'estimation Bayésienne et l'estimation du maximum de vraisemblance sur des données réelles

n	m	paramètre	MV	$GQ(\gamma = -2)$	$En(p = 2)$	$Li(r = -2)$
60	40	α	0.5563 (0.1593)	0.5235 (0.0031)	0.5384 (0.0148)	0.5245 (0.0052)
		θ	0.5756 (0.0503)	0.5145 (0.0013)	0.5641 (0.1372)	0.5233 (0.0048)
		β	0.2860 (0.0409)	0.2455 (0.0022)	0.2670 (0.0039)	0.2598 (0.0005)
		$S(t)$	0.6821 (0.0164)	0.6660 (0.0059)	0.6432 (0.0183)	0.6707 (0.0057)
		$h(t)$	0.0137 (0.0004)	0.0109 (9.1583.10 ⁻⁶)	0.0125 (0.0026)	0.0106 (9.8024.10 ⁻⁶)

TABLE 4.11 – Comparaison de Pitman

n	m	paramètre	$GQ(\gamma = -2)$	$En(p = 2)$	$Li(r = -2)$
60	40	α	0.909	0.908	0.871
		θ	0.990	0.802	0.969
		β	0.124	0.490	0.004
		$S(t)$	0.919	0.951	0.238
		$h(t)$	0.898	0.711	0.900

TABLE 4.12 – Comparaison IMSE

n	m	paramètre	MV	$GQ(\gamma = -2)$	$En(p = 2)$	$Li(r = -2)$
60	40	α	0.07999	0.00429	0.01016	0.00877
		θ	0.09501	0.00177	0.00501	0.00410
		β	0.00306	0.00470	0.01635	0.01461
		$S(t)$	0.02033	0.00769	0.02822	0.02497
		$h(t)$	$0.5.40392.10^{-5}$	$1.17735.10^{-5}$	$3.59688.10^{-5}$	$1/82144.10^{-5}$

De toute évidence, les résultats obtenus à partir des tableaux 4.10, 4.11 et 4.12 confirment ce que nous avons vu dans la section simulation et comparaison.

On a vu que on ne peut pas calculer la densité prédictive Bayésienne de manière analytique. Ainsi, on suppose que $(\alpha_i, \theta_i, \beta_i); i = 1, \dots, M$ sont des échantillons MCMC que nous avons obtenus à partir de l'estimation précédente en utilisant l'ensemble des données réelles au-dessus, on obtient le prédicteur $\hat{Y}_k = E(Y_k)$ lorsque $k = 1, N = 60$ c'est-à-dire les première et dernières statistiques d'ordre dans le future comme

$$\hat{Y}_1 = 119277838.3, \quad \hat{Y}_N = 133511790.2 .$$

Conclusion et Perspectives

Dans cet ouvrage, nous avons fait une inférence bayésienne fondée sur des données censurées ; il y a eu une censure à droite pour le modèle de Bertholon, une censure du type II pour la distribution de Zeghdoudi tronquée, une censure progressive pour la distribution de Burr XII à trois paramètres. Tout d'abord, les approximations des estimateurs de vraisemblance maximale ont été trouvées, et ensuite, les estimateurs bayésiens ont été dérivés sous les trois fonctions de perte.

Dans le modèle de Bertholon, l'algorithme EM a servi à obtenir l'estimation du maximum de vraisemblance. Les résultats ont été obtenus à partir de bases de données simulées de tailles 10, 20 et 30. et pour 10 %% et 20 %% des données sont censurées. Alors qu'en ce qui concerne la distribution Burr à trois paramètres et la distribution Zeghdoudi tronquée, les estimateurs du maximum de vraisemblance ont été obtenus par la commande *BB – solve* du logiciel R.

Les résultats de simulation pour l'estimation bayésienne dans les trois modèles ont été obtenus au moyen des méthodes MCMC ; on n'a remarqué que sous la fonction de perte quadratique généralisée ; Les estimateurs bayésiens ont généralement réalisé de meilleurs résultats que les estimateurs de vraisemblance maximale pour les distributions de Zeghdoudi et de Burr. D'un autre côté, pour le modèle Bertholon, la fonction de perte d'entropie donne les meilleurs résultats. Lorsque l'on compare les résultats, il est peu probable d'obtenir un résultat clair et parfait pour tous les paramètres combinés, compte tenu du risque a posteriori et de l'erreur quadratique. En outre, il a été noté que le choix des schémas et le taux de censure affectent les résultats d'estimation. La prédiction a été obtenue au moyen de l'ensemble de données réel pour la distribution de Burr.

Au cours des prochains travaux, nous examinerons le système de censure progressive de type II et choisirons une fonction de perte optimale pour une meilleure comparaison.

Annexe

```
# Algorithme de l'estimation du maximum de vraisemblance
rm(list=ls())
N<-10000
t<-10
p1<-numeric(N)
p2<-numeric(N)
p3<-numeric(N)
Sv<-numeric(N)
Hv<-numeric(N)
alpha<-0.5;theta<-0.5;beta<-0.25;
S<-(1+(t/alpha) ^ theta) ^ (-beta)
H<-theta*beta*alpha ^ (-theta)*t ^ (theta-1)*(1+(t/alpha) ^ theta) ^ (-1)
n<-40
m<-30
R<-c(rep(0,10),rep(1,10),rep(0,10))
Z<-numeric(m-1)
for(i in 1 :m-1){
  Z[i]<-R[i]+1}
C<-cumsum(Z)
D=numeric(m-1)
for(i in 1 :m-1){
  D[i]<-n-C[i]}
```

```

A<-n*prod(D)
for(h in 1 :N)
Xr<-numeric(m)
U<-runif(m,0,1)
B<-vector("numeric",m)
g<-rep(0,m)
for(k in 1 :m)g[k]<-(R[k]+1)
gamma<-rep(0,m)
for(i in 1 :m)
s<-0
for(j in i :m)
s<-s+g[j]
gamma[i]<-s
for(i in 1 :m){
B[i]<-U[i] ^ (1/gamma[i])
}
V<-rep(0,m)
V[1]<-B[1]
for(k in 2 :m)V[k]=B[k]*V[k-1]
Ur<-rep(0,m)
for(r in 1 :m)Ur[r]<-1-V[r]
for(r in 1 :m)Xr<-alpha*((1-Ur) ^ -1/beta)-1) ^ (1/theta);
Xr<-sort(Xr)
g<-function(q)log(A)+m*log(q[1])+m*log(q[3])-q[1]*m*log(q[2])+(q[1]-1)
*sum(log(Xr))-(q[3]+1)*sum(log(1+(Xr/q[2]) ^ q[1]))-q[3]*sum(R*log(1+(Xr/q[2]) ^ q[1]))
f<-rep(NA,length(q))
f[1]<- m/q[1]-m*log(q[2])+sum(log(Xr))-(q[3]+1)*sum(log(Xr/q[2])*(Xr/q[2]) ^ q[1]
/(1+(Xr/q[2]) ^ q[1]))-q[3]*sum(R*log(Xr/q[2])*(Xr/q[2]) ^ q[1]/(1+(Xr/q[2]) ^ q[1]))
f[2]<- -q[2]*(m/q[1])+(q[3]+1)*sum((-q[2]/q[1]*(Xr/q[1]) ^ q[2]
/(1+(Xr/q[1]) ^ q[2]))-q[3]*sum(R*-q[2]/q[1]*(Xr/q[1]) ^ q[2]/(1+(Xr/q[1]) ^ q[2]))
f[3]<- m/q[3]-sum(log(1+(Xr/q[1]) ^ q[2]))-sum(R*log(1+(Xr/q[1]) ^ q[2]))
f

```

```

library(BB)
v<-BBSolve(par=c(0.5,0.5,0.25), fn = g)$par
p1[h]<-v[1]
p2[h]<-v[2]
p3[h]<-v[3]
Sv[h]<-(1+(t/p2[h]) ^ p1[h]) ^ (-p3[h])
Hv[h]<-p1[h]*p3[h]*p2[h] ^ (-p1[h])*t ^ (p1[h]-1)*(1+(t/p2[h]) ^ p1[h]) ^ (-1)
alphamle<-mean(p2)
thetamle<-mean(p1)
betamle<-mean(p3)
Se<-mean(Sv)
He<-mean(Hv)
alphaER<-sum(p2-alpha) ^ 2/N
thetaER<-sum(p1-theta) ^ 2/N
betaER<-sum(p3-beta) ^ 2/N
SER<-sum(Sv-S) ^ 2/N
HER<-sum(Hv-H) ^ 2/N
alphamle
thetamle
betamle
alphaER
thetaER
betaER
Se
He
SER
HER
# Algorithme de l'estimation Bayésienne
rm(list=ls())
N<-10000 ;
to<-10
p1<-numeric(N)

```



```

p2<-numeric(N)
p3<-numeric(N)
p4<-numeric(N)
p5<-numeric(N)
alpha<-0.5;theta<-0.5;beta<-0.25;a1<-1;a2<-1;a3<-1;b1<-2;b2<-2;b3<-4;b<-3;
S<-(1+(to/alpha) ^ theta) ^ (-beta)
H<-theta*beta*alpha ^ (-theta)*to ^ (theta-1)*(1+(to/alpha) ^ theta) ^ (-1)
n<-40
m<-30
R<-c(10,rep(0,29))
for(h in 1 :N)
  Xr<-numeric(m)
  U<-runif(m,0,1)
  B<-numeric(m)
  g<-rep(0,m)
  for(k in 1 :m)g[k]<-(R[k]+1)
  gamma<-rep(0,m)
  for(i in 1 :m)
    s<-0
    for(j in i :m)
      s<-s+g[j]
    gamma[i]<-s
  for(i in 1 :m)
    B[i]<-U[i] ^ (1/gamma[i])
  V<-rep(0,m)
  V[1]<-B[1]
  for(k in 2 :m)V[k]=B[k]*V[k-1]
  Ur<-rep(0,m)
  for(r in 1 :m)Ur[r]<-1-V[r]
  vec<-alpha*((1-Ur) ^ (-1/beta)-1) ^ (1/theta);
  Xr<-sort(vec);
q<-function(alpha,theta,beta)b ^ 3*exp(-b*(alpha+theta+beta))

```

```

f<-function(alpha,theta,beta){alpha ^ (-theta*m+a1-1)*theta ^ (m+a2-1)*beta^(m+a3-
1)
*exp(-b1*alpha-b2*theta- b3*beta)*prod(Xr ^ (theta-1)
*(1+(Xr/alpha) ^ theta) ^ -(beta+1))
*(1+(Xr/alpha) ^ theta) ^ (-beta*R))}
M<-100; ind<-M*3;
X<-matrix(rep(0,ind),ncol=3,nrow=M)
Y<-numeric(3)
X[1,1]<-1;X[1,2]<-1;X[1,3]<-1
for(i in 2 :M)
Y<-rexp(1,1)+c(0.37,0.4,0.1)
val<-(f(Y[1],Y[2],Y[3])*q(X[i-1,1],X[i-1,2],X[i-1,3]))/
(q(Y[1],Y[2],Y[3])*f(X[i-1,1],X[i-1,2],X[i-1,3]))
val
t<-min(1,val)
t
u<-runif(1)
if (u<t)X[i,1]<-Y[1];X[i,2]<-Y[2];X[i,3]<-Y[3] else
X[i,1]<-X[i-1,1];X[i,2]<-X[i-1,2];X[i,3]<-
X[i-1,3]
AL<-(X[,1])
TH<-(X[,2])
BE<-(X[,3])
IMSEa=sum((p1-alpha) ^ 2)/N;
IMSEt=sum((p2-theta) ^ 2)/N;
IMSEb=sum((p3-beta) ^ 2)/N;
IMSEa
IMSEt
IMSEb
# Estimation par la fonction de perte quadratique généralisée
lambda=-c(-2,2)
p11q[h]<-mean(AL ^ (lambda[1]-1)*AL)/mean(AL ^ (lambda[1]-1))

```

```

p21q[h]<-mean(TH ^ (lambda[1]-1)*TH)/mean(TH ^ (lambda[1]-1))
p31q[h]<-mean(BE ^ (lambda[1]-1)*BE)/mean(BE ^ (lambda[1]-1))
r11q[h]<-mean(alpha ^ (lambda[1]-1)*(alpha-AL) ^ 2)
r21q[h]<-mean(theta ^ (lambda[1]-1)*(theta-TH) ^ 2)
r31q[h]<-mean(beta ^ (lambda[1]-1)*(beta-BE) ^ 2)
p41q<-(1+(to/p11q) ^ p21q) ^ (-p31q)
p51q<-p21q*p31q*p11q ^ (-p21q)*to ^ (p21q-1)*(1+(to/p11q) ^ p21q) ^ (-1)
r41q<-(p41q-S) ^ 2/2
r51q<-(p51q-H) ^ 2/2
p12q[h]<-mean(AL ^ (lambda[2]-1)*AL)/mean(AL ^ (lambda[2]-1))
p22q[h]<-mean(TH ^ (lambda[2]-1)*TH)/mean(TH ^ (lambda[2]-1))
p32q[h]<-mean(BE ^ (lambda[2]-1)*BE)/mean(BE ^ (lambda[2]-1))
r12q[h]<-mean(alpha ^ (lambda[2]-1)*(alpha-AL) ^ 2)
r22q[h]<-mean(theta ^ (lambda[2]-1)*(theta-TH) ^ 2)
r32q[h]<-mean(beta ^ (lambda[2]-1)*(beta-BE) ^ 2)
p42q<-(1+(to/p12q) ^ p22q) ^ (-p32q)
p52q<-p22q*p32q*p12q ^ (-p22q)*to ^ (p22q-1)*(1+(to/p12q) ^ p22q) ^ (-1)
r42q<-(p42q-S) ^ 2/2
r52q<-(p52q-H) ^ 2/2

```

Estimation par la fonction de perte Linex

```

a<-c(-2,2)
p11L[h]<-(-1/a[1])*(log(mean(exp(-a[1]*AL))))
p21L[h]<-(-1/a[1])*(log(mean(exp(-a[1]*TH))))
p31L[h]<-(-1/a[1])*(log(mean(exp(-a[1]*BE))))
r11L[h]<-a[1]*(p11q-P11L)
r21L[h]<-a[1]*(p21q-p21L)
r31L[h]<-a[1]*(p31q-p31L)
p41L<-(1+(to/p11L) ^ p21L) ^ (-p31L)
p51L<-p11L*p31L*p11L ^ (-p21L)*to ^ (p21L-1)*(1+(to/p11L) ^ p31L) ^ (-1)
r41L<-(p41L-S) ^ 2/2
r51L<-(P51L-H) ^ 2/2
p12L[h]<-(-1/a[2])*(log(mean(exp(-a[1]*AL))))

```

```

p22L[h]<-(-1/a[2])*(log(mean(exp(-a[1]*TH))))
p32L[h]<-(-1/a[2])*(log(mean(exp(-a[1]*BE))))
r12L[h]<-a[2]*(p12q-P12L)
r22L[h]<-a[2]*(p22q-p22L)
r32L[h]<-a[2]*(p32q-p32L)
p42L<-(1+(to/p11L) ^ p21L) ^ (-p31L)
p52L<-p22L1*p32L*p12L1 ^ (-p22L)*to ^ (p22L-1)*(1+(to/p12L) ^ p32L) ^ (-1)
r42L<-(p42L-S) ^ 2/2
r52L<-(P52L-H) ^ 2/2

```

Estimation par la fonction de perte Entropy

```

p<-c(-2,2)
p1[h]<-(-mean(AL ^ (-p[1]))) ^ (-1/p[1])
p2[h]<-(-mean(TH ^ (-p[1]))) ^ (-1/p[1])
p3[h]<-(-mean(BE ^ (-p[1]))) ^ (-1/p[1])
ral1<-p[1]*mean(log(alpha)-log(p1[h]))
rth1<-p[1]*mean(log(theta)-log(p2[h]))
rbe1<-p[1]*mean(log(beta)-log(p3[h]))
p4<-(-1+(to/p1[h]) ^ p2[h]) ^ (-p3[h])
p5<-p2[h]*p3[h]*p1[h] ^ (-p2[h])*to ^ (p2[h]-1)*(1+(to/p1[h]) ^ p2[h]) ^ (-1)
p10[h]<-(-mean(AL ^ (-p[2]))) ^ (-1/p[2])
p20[h]<-(-mean(TH ^ (-p[2]))) ^ (-1/p[2])
p30[h]<-(-mean(BE ^ (-p[2]))) ^ (-1/p[2])
ral2[h]<-p[2]*mean(log(alpha)-log(p1[h]))
rth2[h]<-p[2]*mean(log(theta)-log(p2[h]))
rbe2[h]<-p[2]*mean(log(beta)-log(p3[h]))
p40<-(-1+(to/p1[h]) ^ p2[h]) ^ (-p3[h])
p50<-p2[h]*p3[h]*p1[h] ^ (-p2[h])*to ^ (p2[h]-1)*(1+(to/p1[h]) ^ p2[h]) ^ (-1)
}

```

Les estimateurs Bayésiens GQ

```

alpha1q<-mean(p11q)
theta1q<-mean(p21q)
beta1q<-mean(p31q)

```

```

ralpha1q<-mean(r11q)
rtheta1q<-mean(r21q)
rbeta1q<-mean(r31q)
Se1q<-mean(p41q)
He1q<-mean(p51q)
alpha1q<-mean(p11q)
theta1q<-mean(p21q)
beta1q<-mean(p31q)
ralpha1q
rtheta1q
rbeta1q
Se1q
He1q
alpha2q<-mean(p12q)
theta2q<-mean(p22q)
beta2q<-mean(p32q)
ralpha1q<-mean(r12q)
rtheta1q<-mean(r22q)
rbeta1q<-mean(r32q)
Se1q<-mean(p42q)
He1q<-mean(p52q)
alpha1q<-mean(p12q)
theta1q<-mean(p22q)
beta1q<-mean(p32q)
ralpha2q
rtheta2q
rbeta2q
Se2q
He2q
# Les estimateurs Bayésiens Linex
alpha1L<-mean(p11L)
theta1L<-mean(p21L)

```

```

beta1L<-mean(p31L)
ralpha1L<-mean(r11L)
rtheta1L<-mean(r21L)
rbeta1L<-mean(r31L)
Se1L<-mean(p41L)
He1q<-mean(p51L)
alpha1L<-mean(p11L)
theta1L<-mean(p21L)
beta1L<-mean(p31L)
ralpha1L
rtheta1L
rbeta1L
Se1L
He1L
    alpha2L<-mean(p12L)
theta2q<-mean(p22L)
beta2L<-mean(p32L)

    ralpha1L<-mean(r12L)
rtheta1L<-mean(r22L)
rbeta1L<-mean(r32L)

    Se1L<-mean(p42L)
He1L<-mean(p52L)
alpha1L<-mean(p12L)
theta1L<-mean(p22L)
beta1L<-mean(p32L)

    ralpha2L
rtheta2L
rbeta2qL
Se2L

```

```

He2L
# Les estimateurs Bayésiens Entropy
alpha1<-mean(p1)
theta1<-mean(p2)
beta1<-mean(p3)
Se1<-mean(p4)
He1<-mean(p5)
Sr1<-mean(Se1-S)
Hr1<-mean(He1-H)
ralpha1<-mean(ral1)
rtheta1<-mean(rth1)
rbeta1<-mean(rbe1)
ve1<-c(alpha1,theta1,beta1,Se1,He1)
re1<-c(ralpha1,rtheta1,rbeta1,Sr1,He1)
alpha2<-mean(p10)
theta2<-mean(p20)
beta2<-mean(p30)
Se2<-mean(p40)
He2<-mean(p50)
Sr2<-mean(Se2-S)
Hr2<-mean(He2-H)
ralpha2<-mean(ral2)
rtheta2<-mean(rth2)
rbeta2<-mean(rbe2)
ve2<-c(alpha2,theta2,beta2,Se2,He2)
re2<-c(ralpha2,rtheta2,rbeta2,Sr2,He2)
ve1
re1
ve2
re2
# Le trace
rm(list=ls())

```

```

N<-10000;
to<-10
p1<-numeric(N)
p2<-numeric(N)
p3<-numeric(N)
Sv<-numeric(N)
Hv<-numeric(N)
alpha<-0.5;theta<-0.5;beta<-0.25;a1<-1;a2<-1;a3<-1;b1<-2;b2<-2;b3<-4;b<-4;
S<-(1+(to/alpha) ^ theta) ^ (-beta)
H<-theta*beta*alpha ^ (-theta)*to ^ (theta-1)*(1+(to/alpha) ^ theta) ^ (-1)
n<-30
m<-20
R<-c(10,rep(0,19))
for(h in 1 :N)
Xr<-numeric(m)
U<-runif(m,0,1)
B<-numeric(m)
g<-rep(0,m)
for(k in 1 :m)g[k]<-(R[k]+1)
gamma<-rep(0,m)
for(i in 1 :m)
s<-0
for(j in i :m)
s<-s+g[j]
gamma[i]<-s
for(i in 1 :m)
B[i]<-U[i] ^ (1/gamma[i])
V<-rep(0,m)
V[1]<-B[1]
for(k in 2 :m)V[k]=B[k]*V[k-1]
Ur<-rep(0,m)
for(r in 1 :m)Ur[r]<-1-V[r]

```



```

vec<-alpha*((1-Ur) ^ (-1/beta)-1) ^ (1/theta) ;
Xr<-sort(vec) ;
q<-function(alpha,theta,beta)b ^ 3*exp(-b*(alpha+theta+beta))
f<-function(alpha,theta,beta)alpha ^ (-theta*m+a1-1)*theta ^ m+a2-1)
*beta ^ (m+a3-1)*exp(-b1*alpha-b2*theta-
b3*beta)*prod(Xr ^ (theta-1)*(1+(Xr/alpha) ^ theta) ^ -(beta+1))
*(1+(Xr/alpha) ^ theta) ^ (-beta*R))
M<-100 ; ind<-M*3 ;
X<-matrix(rep(0,ind),ncol=3,nrow=M)
Y<-numeric(3)
X[1,1]<-1 ;X[1,2]<-1 ;X[1,3]<-1
for(i in 2 :M)
Y<-rexp(1,1)+c(0.37,0.4,0.1)
val<-(f(Y[1],Y[2],Y[3])*q(X[i-1,1],X[i-1,2],X[i-1,3]))/
(q(Y[1],Y[2],Y[3])*f(X[i-1,1],X[i-1,2],X[i-1,3]))
val
t<-min(1,val)
t
u<-runif(1)
if (u<t)X[i,1]<-Y[1] ;X[i,2]<-Y[2] ;X[i,3]<-Y[3] else
X[i,1]<-X[i-1,1] ;X[i,2]<-X[i-1,2] ;X[i,3]<-X[i-1,3]
p1[h]<-mean(X[,1])
p2[h]<-mean(X[,2])
p3[h]<-mean(X[,3])
m<-matrix(c(p1,p2,p3),ncol=3)
library('coda')
mcmc.trace <- mcmc(m)
summary(mcmc.trace)
effectiveSize(mcmc.trace)
plot(mcmc.trace)
# Algorithme de comparaison avec Pitman et IMSE
rm(list=ls())

```

```

N<-10000;
to<-10
alphaq<-numeric(N)
thetaq<-numeric(N)
betaq<-numeric(N)
Seq<-numeric(N)
Heq<-numeric(N)
alphal<-numeric(N)
thetal<-numeric(N)
betal<-numeric(N)
Sel<-numeric(N)
Hel<-numeric(N)
alphae<-numeric(N)
thetae<-numeric(N)
betae<-numeric(N)
See<-numeric(N)
Hee<-numeric(N)
alphaml<-numeric(N)
thetaml<-numeric(N)
betaml<-numeric(N)
Seml<-numeric(N)
Heml<-numeric(N)
n1<-0;n2<-0;n3<-0;n4<-0;n5<-0;n6<-0;n7<-0;n8<-0;n9<-0;n10<-0;
n11<-0;n12<-0;n13<-0;n14<-0;n15<-0;
alpha<-0.5;theta<-0.5;beta<-0.25;a1<-1;a2<-1;a3<-1;b1<-2;
b2<-2;b3<-4;b<-3;
S<-(1+(to/alpha)^theta)^(-beta)
H<-theta*beta*alpha^(theta)*to^(theta-1)*(1+(to/alpha)^theta)^-1)
n<-40
m<-30
R<-c(10,rep(0,29))
Z<-numeric(m-1)

```

```

for(i in 1 :m-1)
Z[i]<-R[i]+1
C<-cumsum(Z)
D=numeric(m-1)
for(i in 1 :m-1)
D[i]<-n-C[i]
A<-n*prod(D)
for(h in 1 :N)
Xr<-numeric(m)
U<-runif(m,0,1)
B<-numeric(m)
g<-rep(0,m)
for(k in 1 :m)g[k]<-(R[k]+1)
gamma<-rep(0,m)
for(i in 1 :m)
s<-0
for(j in i :m)
s<-s+g[j]
gamma[i]<-s
for(i in 1 :m)
B[i]<-U[i] ^ (1/gamma[i])
V<-rep(0,m)
V[1]<-B[1]
for(k in 2 :m)V[k]=B[k]*V[k-1]
Ur<-rep(0,m)
for(r in 1 :m)Ur[r]<-1-V[r]
vec<-alpha*((1-Ur) ^ (-1/beta)-1) ^ (1/theta) ;
Xr<-sort(vec) ;
q<-function(alpha,theta,beta)b ^ 3*exp(-b*(alpha+theta+beta))
f<-function(alpha,theta,beta)alpha ^ (-theta*m+a1-1)*theta ^ (m+a2-1)
*beta^(m+a3-1)*exp(-b1*alpha-b2*theta-b3*beta)
*prod(Xr ^ (theta-1)*(1+(Xr/alpha) ^ theta) ^ -(beta+1))

```

```

*(1+(Xr/alpha) ^ theta) ^ (-beta*R))
M<-100; ind<-M*3;
X<-matrix(rep(0,ind),ncol=3,nrow=M)
Y<-numeric(3)
X[1,1]<-1;X[1,2]<-1;X[1,3]<-1
for(i in 2 :M)
Y<-rexp(1,1)+c(0.37,0.4,0.1)
val<-(f(Y[1],Y[2],Y[3])*q(X[i-1,1],X[i-1,2],X[i-1,3]))/
(q(Y[1],Y[2],Y[3])*f(X[i-1,1],X[i-1,2],X[i-1,3]))
val
t<-min(1,val)
t
u<-runif(1)
if (u<t)X[i,1]<-Y[1];X[i,2]<-Y[2];X[i,3]<-Y[3]
else X[i,1]<-X[i-1,1];X[i,2]<-X[i-1,2];X[i,3]<-X[i-1,3]
lambda=-2
alphaq[h]<-mean(X[,1] ^ (lambda-1)*X[,1])/mean(X[,1] ^ (lambda-1))
thetaq[h]<-mean(X[,2] ^ (lambda-1)*X[,2])/mean(X[,2] ^ (lambda-1))
betaq[h]<-mean(X[,3] ^ (lambda-1)*X[,3])/mean(X[,3] ^ (lambda-1))
Seq[h]<-(1+(to/alphaq[h]) ^ thetaq[h]) ^ (-betaq[h])
Heq[h]<-thetaq[h]*betaq[h]*alphaq[h] ^ (-thetaq[h])*to ^ (thetaq[h]-1)
*(1+(to/alphaq[h]) ^ thetaq[h]) ^ (-1)
a<-2
alphal[h]<-(-1/a)*(log(mean(exp(-a*X[,1]))))
thetal[h]<-(-1/a) ^ (log(mean(exp(-a*X[,2]))))
betal[h]<-(-1/a) ^ (log(mean(exp(-a*X[,3]))))
Sel[h]<-(1+(to/alphal[h]) ^ thetal[h]) ^ (-betal[h])
Hel[h]<-thetal[h]*betal[h]*alphal[h] ^ (-thetal[h])
*to ^ (thetal[h]-1)
*(1+(to/alphal[h]) ^ thetal[h]) ^ (-1)
p<-2
alphae[h]<-(mean(X[,1] ^ (-p))) ^ (-1/p)
thetae[h]<-(mean(X[,2] ^ (-p))) ^ (-1/p)

```

```

betae[h]<-(mean(X[,3] ^ (-p[1]))) ^ (-1/p)
See[h]<-(1+(to/alphae[h]) ^ thetai[h]) ^ (-betae[h])
Hee[h]<-thetae[h]*betae*alphae[h] ^ (-thetae[h])
*to ^ thetai[h]-1*(1+(to/alphae[h]) ^ thetai[h]) ^ (-1)
g<-function(q)log(A)+m*log(q[1])+m*log(q[3])-q[1]*m*log(q[2])+(q[1]-1)
*sum(log(Xr))-(q[3]+1)*sum(log(1+(Xr/q[2]) ^ q[1]))-q[3]*
sum(R*log(1+(Xr/q[2]) ^ q[1]))
f<-rep(NA,length(q))
f[1]<- m/q[1]-m*log(q[2])+sum(log(Xr))-(q[3]+1)*sum(log(Xr/q[2])
*(Xr/q[2]) ^ q[1]/(1+(Xr/q[2]) ^ q[1]))-
q[3]*sum(R*log(Xr/q[2])*(Xr/q[2]) ^ q[1]/(1+(Xr/q[2]) ^ q[1]))
f[2]<- -q[2]*(m/q[1])+(q[3]+1)*sum((-q[2]/q[1]
*(Xr/q[1])^q[2]/(1+(Xr/q[1]) ^ q[2]))-q[3]*sum(R*-q[2]/
q[1]*(Xr/q[1]) ^ q[2]/(1+(Xr/q[1]) ^ q[2]))
f[3]<- m/q[3]-sum(log(1+(Xr/q[1]) ^ q[2]))-sum(R*log(1+(Xr/q[1]) ^ q[2]))
f
library(BB)
v<-BBsolve(par=c(0.5,0.5,0.25), fn = g)$par
alphaml[h]<-v[1]
thetaml[h]<-v[2]
betaml[h]<-v[3]
Seml[h]<-(1+(to/thetaml[h])^alphaml[h])^(-betaml[h])
Heml[h]<-alphaml[h]*betaml[h]*thetaml[h]^(-alphaml[h])
*to^(alphaml[h]-1)*(1+(to/thetaml[h])^alphaml[h])^(-1)
if(abs(alphaq[h]-alpha)<abs(alphaml-alpha))n1=n1+1
if(abs(thetaq[h]-theta)<abs(thetaml-theta))n2=n2+1
if(abs(betaq[h]-beta)<abs(betaml-beta))n3=n3+1
if(abs(Seq[h]-S)<abs(Seml-S))n4=n4+1
if(abs(Heq[h]-H)<abs(Heml-H))n5=n5+1
if(abs(alphal[h]-alpha)<abs(alphaml-alpha))n6=n6+1
if(abs(thetal[h]-theta)<abs(thetaml-theta))n7=n7+1
if(abs(betal[h]-beta)<abs(betaml-beta))n8=n8+1

```

```

if(abs(Sel[h]-S)<abs(Seml-S))n9=n9+1
if(abs(Hel[h]-H)<abs(Heml-H))n10=n10+1
if(abs(alphae[h]-alpha)<abs(alphaml-alpha))n11=n11+1
if(abs(thetae[h]-theta)<abs(thetaml-theta))n12=n12+1
if(abs(betae[h]-beta)<abs(betaml-beta))n13=n13+1
if(abs(See[h]-S)<abs(Seml-S))n14=n14+1
if(abs(Hee[h]-H)<abs(Heml-H))n15=n15+1
}
P1=n1/N
P2=n2/N
P3=n3/N
P4=n4/N
P5=n5/N
PQ=c(P1,P2,P3,P4,P5)
P6=n6/N
P7=n7/N
P8=n8/N
P9=n9/N
P10=n10/N
PL=c(P6,P7,P8,P9,P10)
P11=n11/N
P12=n12/N
P13=n13/N
P14=n14/N
P15=n15/N
PE=c(P11,P12,P13,P14,P15)
IMSE1=mean((alphaq-alpha)^2)
IMSE2=mean((thetaq-theta)^2)
IMSE3=mean((betaq-beta)^2)
IMSE4=mean((Seq-S)^2)
IMSE5=mean((Heq-H)^2)
IMSE6=mean((alphal-alpha)^2)

```

```

IMSE7=mean((thetal-theta) ^ 2)
IMSE8=mean((betal-beta) ^ 2)
IMSE9=mean((Sel-S) ^ 2)
IMSE10=mean((Hel-H) ^ 2)
IMSE11=mean((alphae-alpha) ^ 2)
IMSE12=mean((thetae-theta) ^ 2)
IMSE13=mean((betae-beta) ^ 2)
IMSE14=mean((See-S) ^ 2)
IMSE15=mean((Hee-H) ^ 2)
IMSE16=mean((alphaml-alpha) ^ 2)
IMSE17=mean((thetaml-theta) ^ 2)
IMSE18=mean((betaml-beta) ^ 2)
IMSE19=mean((Seml-S) ^ 2)
IMSE20=mean((Heml-H) ^ 2)
#pitman
PQ
PL
PE
#IMSE
IMSE1
IMSE2
IMSE3
IMSE4
IMSE5
IMSE6
IMSE7
IMSE8
IMSE9
IMSE10
IMSE11
IMSE12
IMSE13

```

IMSE14

IMSE15

IMSE16

IMSE17

IMSE18

IMSE19

IMSE20

La prédiction

rm(list=ls())

N<-10;

to<-10

L<-numeric(N)

U<-numeric(N)

delta=0.90

kk=2

alpha<-0.5 ;theta<-0.5 ;beta<-0.25 ;a1<-1 ;a2<-1 ;a3<-1 ;b1<-2 ;b2<-2 ;b3<-4 ;b<-3 ;

n=60

m=40

R=c(0,0,0,2,0,0,0,2,0,0,0,2,0,0,0,2,0,0,0,2,0,0,0,2,0,0,0,2,0,0,

0,0,2,0,0,0,2,0,0,0,2,0,0,2,0,0,2,0)

x=c(108359319.3,108051012.2,132583400.9,131682245.2,

138763777.5,127472451.4,157461435.7,

147181634.8,133049464.3,129306784.6,146823384.7,

130188519.9,91251458.84,108199232.29,

112232638.39,101411375.93,112050919.56,92857094.48,

104336392.13,95439179.77,

87403403.19,89656893.21,102954049.91,93239056.94,

126252901.25,111455462.91,

118966778.27,131760329.15,121534252.2,106032084.34,

130928558.38,138498736.05,114162673.88,141433992.68,

118519128.69,178298704.25,123854041.95,106767459.16,

140082764.19,133995516.96,134887487.43,156069160.99,


```

164424277.71,139323594.5,121052740.22,135579481.66,
139047972.12,177643770.67,136724716.59,155210984.05,
161188834.33,145708124.7,174445210.24,155877782.25,
150030573.28,155365194.65,131837711.49,150634506.45,
140338111.28,136827221.9)
Xr=c(108051012.2,132583400.9,138763777.5,127472451.4,129306784.6,
146823384.7,130188519.9,108199232.29,101411375.93,112050919.56,104336392.13,
87403403.19,89656893.21,102954049.91,126252901.25,111455462.91,131760329.15,
121534252.2,106032084.34,130928558.38,138498736.05,
141433992.68,118519128.69,123854041.95,140082764.19,
133995516.96,134887487.43,164424277.71,139323594.5,
121052740.22,135579481.66,139047972.12,177643770.67,
155210984.05,161188834.33,145708124.7,174445210.24,
150030573.28,131837711.49,150634506.45)
g<-rep(0,m)
for(k in 1 :m)g[k]<-(R[k]+1)
gamma<-rep(0,m)
for(i in 1 :m)
s<-0
for(j in i :m)
s<-s+g[j]
gamma[i]<-s
Ck=prod(gamma[1 :kk])
a=c(rep(1,kk-1))
aa=c(rep(1,kk))
for (i in 1 :kk)
for (l in 1 :kk-i)
for (j in i :kk)
if (j!=i)
aa[l]=(1/(gamma[j]-gamma[i]))
a[i]=prod(aa)
for(h in 1 :N)

```

```

q<-function(alpha,theta,beta)b ^ 3*exp(-b*(alpha+theta+beta))
f<-function(alpha,theta,beta)alpha ^ (-theta*m+a1-1)*theta ^ m+a2-1)
*beta^(m+a3-1)*exp(-b1*alpha-b2*theta-
b3*beta)*prod(Xr^(theta-1)*(1+(Xr/alpha) ^ theta) ^ -(beta+1))
*(1+(Xr/alpha) ^ theta) ^ (-beta*R))
M<-100; ind<-M*3;
X<-matrix(rep(0,ind),ncol=3,nrow=M)
Y<-numeric(3)
X[1,1]<-1;X[1,2]<-1;X[1,3]<-1
for(i in 2 :M)
Y<-rexp(1,1)+c(0.37,0.4,0.1)
val<-(f(Y[1],Y[2],Y[3])*q(X[i-1,1],X[i-1,2],X[i-1,3]))/
(q(Y[1],Y[2],Y[3])*f(X[i-1,1],X[i-1,2],X[i-1,3]))
val
t<-min(1,val)
t
u<-runif(1)
if (u<t)X[i,1]<-Y[1];X[i,2]<-Y[2];X[i,3]<-Y[3]
else X[i,1]<-X[i-1,1];X[i,2]<-X[i-1,2];X[i,3]<-X[i-1,3]
p1=X[,1]
p2=X[,2]
p3=X[,3]
g<-function(q)
f<-rep(NA,length(q))
f[1]<-sum(Ck*sum(a/gamma*(1+(q[1]/p1) ^ p2) ^ (-p3*gamma)))-1/2+delta/2
f[2]<-sum(Ck*sum(a/gamma*(1+(q[2]/p1) ^ p2) ^ (-p3*gamma)))-1/2-delta/2
f
library(BB)
v<-BBsolve(par=c(130000000.5,10000000.5), fn = g)$par
L[h]<-v[1]
U[h]<-v[2]
lower<-mean(L)

```

```

upper<-mean(U)
lower
upper
# Les testes d'ajustements
x=c(108359319.3,108051012.2,132583400.9,131682245.2,138763777.5
,127472451.4,157461435.7,147181634.8,
133049464.3,129306784.6,146823384.7,
130188519.9,91251458.84,108199232.29,
112232638.39,101411375.93,112050919.56,
92857094.48,104336392.13,95439179.77,87403403.19,
89656893.21,102954049.91,93239056.94,
126252901.25,111455462.91,118966778.27,
131760329.15,121534252.2,106032084.34,
130928558.38,138498736.05,114162673.88,
141433992.68,118519128.69,178298704.25,
123854041.95,106767459.16,140082764.19,
133995516.96,134887487.43,156069160.99,
164424277.71,139323594.5,121052740.22,
135579481.66,139047972.12,177643770.67,
136724716.59,155210984.05,161188834.33,145708124.7,
174445210.24,155877782.25,150030573.28,155365194.65,
131837711.49,150634506.45,140338111.28,136827221.9)
n=60
p=3
alpha=0.5
theta=0.5
beta=0.25
### l=n*log(alpha)-alpha*n*log(theta)+(alpha-1)*sum(log(x))-sum((x/theta)^alpha)
### l=-n*log(beta)-n/2*log(2*pi)-sum(log(x))-1/(2*beta^2)*sum((log(x)-theta)^2)
### l=-n*log(beta)-n/2*log(2*pi)-1/(2*beta^2)*sum((log(x)-theta)^2)
### l=n*log(alpha)-(1/theta+1)*sum(log(1+theta*((x-beta)/alpha)))
### l=n*log(k)-n*log(theta)-n*log(beta(apha,beta))+(alpha+1)

```

```

*sum(log(1+(x/theta) ^ k)-(k*beta-1)*sum(log(x/theta))
## l=-n*log(beta)+n*log(alpha)+(beta-1)*sum(log(x))-(alpha+1)*sum((log(1+x ^ beta))
l=n*log(theta)+n*log(beta)-theta*n*log(alpha)+(theta-1)*sum(log(x))-(beta+1)
*sum(log(1+(x/alpha) ^ theta))
Zi=numeric(n)
for (i in 1 :n)
Zi[i]=1-(1+(x[i]/alpha) ^ theta) ^ (-beta)
AIC=-2*l+2*p
CAIC=-2*l+2*p*n/(n-p-1)
BIC=-2*l+p*log(n)
HQIC=-2*l+2*p*log(log(n))
j=1 :n
k=n-j+1
Zk=Zi[n :1]
W=(1/(2*n)+1)*(sum((Zi-(2*j-1)/(2*n)) ^ 2)+1/(12*n))
A=(9/(4*n ^ 2)+3/(4*n)+1)*(n+1/n*sum((2*j-1)*log(Zi*(1-Zk))))
AIC
CAIC
HQIC
BIC
W
A

```

Bibliographie

- [1] Achcar J.A , and Leonardo R.A, (1998). Use of Markov Chain Monte Carlo methods in a Bayesian analysis of the Block and Basu bivariate exponential distribution. *Annals of the Institute of Statistical Mathematics*, 50, 403-416.
- [2] Afify A.Z, Aljohani H.M, Alghamdi A, and Gemeay A.M, (2021). A new two-parameter Burr-Hatke distribution : properties and Bayesian and non-Bayesian inference with applications. *Journal of Mathematics*, 2314-4629.
- [3] Afify A. Z, Mead M.E., (2017). On five-parameter burr XII distribution : properties and applications, *S Afr StatJ* 51 :67-80.
- [4] Al-Sabbah S.A.S. et al (2021). Comparison of methods for estimating the reliability of arriving for the Burr distribution. Comparison of methods for estimating the reliability of arriving for the Burr distribution. *J. Phys. : Conf. Ser.* 1897 012005.
- [5] Balakrishnan N, and Mitra D, (2012). Left truncated and right censored Weibull data and likelihood inference with an illustration, *computation statistics and data analysis.*56 (12), 4011-4025.
- [6] Balakrishnan N, and Aggarwala R, (2000). *Progressive Censoring : Theory, Methods and Applications*, Boston : Birkhauser.
- [7] Balakrishnan N, and Sandhu R.A, (1997). A simple simulation algorithm for generating progressive type II censored samples. *American Statistician* 49, 2, 229-230.

- [8] Balakrishnan N, and Sandhu R.A, (1996). Best linear unbiased and maximum likelihood estimation for exponential distributions under general progressive Type-II censored samples. *Sankhya B*, 58, 1–9.
- [9] Balakrishnan N, (2007). Progressive censoring mythology : an appraisal (with discussions). *TEST*, 16, 211-296.
- [10] Bantan R.A.R, Jamal F, Chesneau C, and Elgarhy M, (2019). Truncated Inverted Kumaraswamy Generated Family Of Distributions with Applications. *Entropy* 2019, 21(11), 1089.
- [11] Bergeron V.L, (2010). L'Algorithme Metropolis-Hastings Département de Mathématiques et Statistique Université de Montréal.
- [12] Berglund N, (2007). Chaînes de Markov Université d'Orléans.
- [13] Boudjerda K, Chadli A, Fellag H, (2016). Posterior Analysis of the Compound Truncated Weibull Under Different Loss Functions for Censored Data, *international journal of mathematics and computers in simulation*, Vol 10, 265-272.
- [14] Bousquet N, Bertholon H, and Celeux G, (2006). An alternative competing risk model to the Weibull distribution for modelling aging in lifetime data analysis, *Lifetime Data Analysis*, 12, p. 481-504.
- [15] Boreux J.J et Parent E. et Bernier J, (2010). *Pratique du calcul Bayésien*. Springer-Paris.
- [16] Burr I.W, (1942). Cumulative frequency functions. *Ann Math Stat* 13 :215-232.
- [17] Chadli A, Talhi H, Fellag H, (2013). Comparison of the maximum likelihood and bayes estimators for symmetric bivariate exponential distribution under different loss functions. *Afr. Stat.* 8 (1) 499 - 514.
- [18] Cohen A.C, (1963). Progressively censored samples in life Testing. *Technometrics*, 5, 3, 327-339.
- [19] Dempster A.P, Laird N.M, and Rubin D.B, (1977). Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Ser. B*, 39(1), p.1-38.

- [20] Droesbeke J.J, Fine J, et Saporta G, (2002). Méthodes Bayésiennes en statistique. Éditions TECHNIP.
- [21] EL-Sagheer R.M, Mahmoud M.A.W, and Hasaballah M.M., (2019). Bayesian estimation using MCMC approach under three parameter Burr XII distribution based on unified hybrid censored scheme. *Journal of statistical theory and practice*, 13(6),1-20.
- [22] EL-Sagheer R.M, (2018). Estimation of parameters of Weibull–Gamma distribution based on progressively censored data. *Stat Papers* 59, 725–757.
- [23] Gunasekera S, (2018). Inference for the Burr XII reliability under progressive censoring with random removals, *Math Comput Simulation*, 144, 182–195.
- [24] Groiez A, (2014). Recyclage des candidats dans l’algorithme Metropolis à essais multiples. Université de Montréal.
- [25] Hakim A.R, et al (2021). Properties of Burr distribution and its application to heavy-tailed survival time data. *J. Phys. : Conf. Ser.*, 1725 012016.
- [26] Jozani M.J, Davies K.F, Balakrishnan N, (2012). Pitman closeness results concerning ranked set sampling. *Statistics and Probability Letters*, 82, 12, 2260-2269.
- [27] Kast R, (2002) La théorie de la décision. Edition la découverte et Syros-Paris.
- [28] Kumar D, (2017). The Burr Type Xii Distribution with Some Statistical Properties. *Journal of data science : JDS*. 15. 10.6339 JDS.201707 15(3).0008.
- [29] Lebeau G, Introduction à l’analyse de l’algorithme de Metropolis Département de Mathématiques, Université de Nice Sophia-Antipolis Parc Valrose 06108 Nice Cedex 02, France. 2009.
- [30] Mann N.R, Schafer R.E, and Singpurwalla N.D, (1974). *Methods for Statistical Analysis of Reliability and Life Data*, John Wiley and Sons.
- [31] Messaadia H, And Zeghdoudi H,(2018). Zeghdoudi distribution and its applications, *Int. J. Computing Science and Mathematics*, Vol. 9, No. 1, pp. 58-65.
- [32] Moussa M.A.M.A, and Jaheen Z.F,(2002). Statistical inference for the Burr model based on progressively censored data. *Comput Math Appl*, 43,1441-1449.

- [33] Nofal Z.M, Afify A.Z, Yousof H.M, and G.M. Cordeiro,(2017). The generalized transmuted-G family of distributions. *Communications in Statistics - Theory and Methods*, 46, 8, 4119-4136.
- [34] Okasha M, and Matter M. (2015). On the Three-Parameter Burr Type XII Distribution and its Application to Heavy Tailed Lifetime Data.. *Advances in Mathematics*. 10. 3429-3442.
- [35] Parent E, et Bernier J, (2007). *Le raisonnement Bayésien-modélisation et inférence* Springer-Verlag France, Paris.
- [36] Park C, and Padgett W.J, (2004). Analysis of strength distributions of multimodal failures using the EM algorithm, Technical report No. 220, Department of Statistics, University of South Carolina.
- [37] Pitman E, (1937). The closest estimates of statistical parameters. *Mathematical proceeding of the Cambridge philosophical society*, 33(2).
- [38] Raqab M.Z, Asgharzadeh A, Valiollahi R, (2010) Prediction for Pareto distribution based on progressively Type-II censored samples.*Computational Statistics and Data Analysis*, 54, 7, 1732-1743.
- [39] Riou J, (2007), *Statistique Bayésienne et Algorithmes MCMC*. .
- [40] Robert C.P. et Marin J.M, (2007). *Les bases de la statistique Bayésienne*. Université Paris Dauphine.
- [41] Robert C. P, (2006) *Le choix Bayésien*. Springer.
- [42] Rossi V, (2006) *les méthodes de Monte Carlo*. Université Montpellier.
- [43] Shao Q, (2004). Notes n maximum likelihood estimation for the three parameter Burr XII distribution. *Comput Stat Data Anal* 45 :675-687.
- [44] Singh S. K, Singh U, Sharma V. K, (2014). The Truncated Lindley Distribution : Inference and Application, *Journal of Statistics Applications and Probability*, 3(2), 219-228.

- [45] Solman A.A, Abd Ellah A.H, Abou Elheggeag N.A, and Modhesh A.A, (2011). Bayesian inference and prediction of Burr type XII distribution for progressive first failure censored sampling. *Intelligent Information Management*, 3, 175-185.
- [46] Talhi H, and Aiachi H., (2021). On truncated Zeghdoudi distribution : Posterior Analysis under Different Loss Functions for Type II Censored Data. *Pakistan Journal of Statistics and Operation Research*, 17, 2, 497-508.
- [47] Tchuendom R.F, (2014). Estimation Bayésienne via une méthodes de Monte Carlo par chaînes de Markov. Inria-Grand Est Equipe Tosca poincaré.
- [48] Tejeda H. A, and Goodwin B. K,(2008). Modeling crop prices through a Burr distribution and analysis of correlation between crop prices and yields using a Copula method, annual meeting of the Agricultural and Applied Economics Association, Orlando, FL, USA, Vol. 83, No. 3, pp. 643–649.
- [49] Varadhan R, and Gilbert P,(2010). An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function. *Journal of statistical software*, 32, 4.
- [50] Wang F.K, Keats J.B, (1996). Maximum likelihood estimation of the Burr XII parameters with censored and uncensored data, *Microelectron. Reliab.* 36(3),359-362.