وزارة التعليم العالي والبحث العلمي

Université Badji Mokhtar Annaba

Badji Mokhtar University - Annaba

جامعة باجي مختار عنابة

Faculté des Sciences
Département de Mathématiques

Laboratoire de Probabilités et Statistique

THESE

Présenté en vue de l'obtention du diplôme de

Doctorat en Mathématiques

Option : Probabilités et Statistique

Etude de l'estimateur de Bayes sous différentes fonctions de perte

Par:
Boudjerda Khawla

Sous la direction de :
Prof. Chadli Assia
Et
Prof. Fellag Hocine

Devant le jury

PRESIDENT :	Seddik-Ameur Nacera	Prof	U.B.M.Annaba
Rapporteur :	Chadli Assia	Prof	U.B.M.Annaba
Co-Directeur:	Fellag Hocine	Prof	U. Tizi Ouzou
EXAMINATEUR:	Djellab Natalia	Prof	U. B.M.Annaba
EXAMINATEUR :	Atil Linda	MCA	U.Tizi-Ouzou

Année :2016-2017

Remerciements

Louange à Dieu qui nous a donné l'esprit, la volonté, le courage et le savoir.

Je tiens à exprimer mes respectueux remerciement et toute ma gratitude à ma directrice de thèse Madame Chadli Assia, professeur à l'université Badji Mokhtar Annaba, pour la confiance et la bienveillance qu'elle m'a témoignée et dont la disponibilité et l'indulgence m'ont permis de mener à bien cette étude. Je la remercie pour son aide précieux et les conseils qu'elle m'a prodiguée tout au long de mon parcours de doctorat.

J'adresse un remerciement chaleureux à monsieur Fellag Hocine, professeur à l'université Mouloud Mameri de Tizi Ouzou, mon co-encadreur, pour m'avoir soutenu d'une façon remarquable et pour avoir suivi de très près mon travail depuis le début. Pour ses nombreuses remarques pertinentes et pour toute l'aide qu'il m'a apporté. Merci infiniment Monsieur.

Je souhaite remercier Madame Seddik Ameur Nacira, professeur à l'université Badji Mokhtar Annaba, qui a honoré par la présidence du jury.

Je remercie vivement les membres examinatrices pour voir accepté d'évaluer ce travail ; particulièrement :

Madame Djellab Natalia, professeur à l'université Badji Mokhtar Annaba, qui a honoré ce travail en acceptant d'être membre de jury de cette thèse.

Madame Atil Linda, maître conférence classe A à l'université Mouloud Mameri Tizi-Ouzou, qui a honoré ce travail en acceptant d'être membre de jury de cette thèse, je la remercie également de me consacrer de son temps et d'avoir accepté de se déplacer.

J'aimerais aussi remercier Monsieur Rahmania Nadji pour son accueil chaleureux à l'université de Lille 1, pour ses encouragements, sa participation dans cette thèse, ainsi de ses nombreuses remarques qui ont contribué au bon déroulement de cette thèse.

Ces remerciement ne seraient pas complets sans mentionner :

Mes chers parents pour ces longues années de soutiens inconditionnel, pour leur confiance permanente et l'acceptation de mes choix. Ils m'ont offert la possibilité de réaliser mon rêve. Ils ont toujours témoigné de la plus grande des patiences et de la plus grande des compréhension. Merci aussi à mes frères et mes soeurs qui m'ont toujours encouragé.

Je remercie toute personne ayant participée de près ou de loin à l'élaboration et l'évaluation de ce travail.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

Résumé

Dans cette thèse, nous nous intéressons essentiellement à deux modèles : le modèle de Weibull tronqué et le modèle de Rayleigh. Les distributions tronquées surviennent quant une variable aléatoire X suit une distribution de probabilité connue, mais une portion de l'espace d'échantillon ne peut pas être observée. Si les valeurs de la variable aléatoire sont limitées à droite par une valeur T, on dit que la distribution est tronquée à droite. La distribution tronquée a laquelle on s'intéresse est la distribution de Weibull tronquée a droite ; cette dernière a la particularité d'avoir un taux de panne en baignoire. La distribution de Rayleigh trouve son application dans de nombreux domaines et en particulier en épidémiologie. Dans cette thèse, nous exposons l'étude des estimateurs de Bayes des paramètres, de la fonction de fiabilité et de la fonction taux de panne sous différentes fonctions de pertes et avec plusieurs type de données. Les lois a priori utilisées dans ce travail sont deux types : les lois a priori non informatives et les lois a priori conjuguées naturelles.

Dans une première partie, on utilise une approche classique du maximum de vraisemblance puis une approche Bayésienne dans le cas d'une distribution de Weibull tronquée à droite en considérant des données censurées de type II. Cette distribution dépend de trois paramètres : un paramètre lié a la troncature et les deux autres propres a la loi de Weibull et qui sont des paramètres de forme et d'échelle.

Les estimateurs Bayésiens et les risques a posteriori sont obtenus en utilisant des fonctions de pertes symétriques (la fonction de perte quadratique et la fonction de perte entropie) puis des fonctions de perte asymétriques (la fonction de perte Linex). Dans le cas de l'approche classique, les estimateurs sont solutions d'un système non linéaire dont les solutions ne sont pas explicites analytiquement ; des méthodes numériques ont été adoptées.

Dans l'approche Bayesienne, les estimateurs sont donnés sous forme d'un rapport d'intégrales, Nous appliquons les méthodes de Monte-Carlo et en particulier l'algorithme de *Metropolis-Hastings* pour procéder à des simulations et à une analyse des données. Finalement, on utilise le critère de Pitman et le critère integrated mean square error ($IMSE$) pour comparer les estimateurs bayésiens et les estimateurs du maximum de vraisemblance.

Dans la deuxième partie, on s'intéresse à l'estimation du paramètre σ et de quelques caractéristiques de fiabilité, comme la fonction de fiabilité $S(t)$ et la fonction de taux de hasard $h(t)$ de la loi de Rayleigh. On a utilisé l'approche Bayésienne sous différentes fonctions de pertes (la fonction de perte quadratique et la fonction de perte Linex), avec des données **censurées de type** II puis avec des données **progressivement censurées**. La loi a priori sur le paramètre est d'abord considérée comme non informative, puis nous considérons le cas d'une loi a priori conjuguée naturelle. Les estimateurs Bayésiens de σ, $S(t)$, et $h(t)$ ont été obtenus avec leurs expressions analytiques exactes. Les risques a posteriori sont calculés dans chaque cas. Une étude par simulation et une analyse de données réelles ont été réalisées afin de comparer les différents estimateurs à partir de leurs risques a posteriori.

Pour le modèle de Weibull tronqué, une publication intitulée : "Posterior Analysis of the compound truncated Weibull under different loss functions for censored data" est sortie dans la revue : International Journal of Mathematics and Compters in Simulation*, vol 10, 2016. pp 265-272.*

Pour le modèle de Rayleigh, un article intitulé : "*Bayesian estimation of the Rayleigh distribution under different loss function" a été soumis à la revue EJASA.*

Abstract

In this thesis, we are intrested in two models : the truncated Weibull and Rayleigh's distributions. The first occur when a random variable X follows a known probability distribution, but a portion of the sample space can not observed. If the values of the random variable are limited to the right by the value T, the distribution is said right truncated. The truncated distribution that we are intrested is the right truncated Weibull distribution ; the latter has the peculiarity of having a failure rate function in bathtub. The Rayleigh distribution finds its application in many fields and particulary in epidemiology.

We shall especially, study the Bayes estimators of the parameters, the reliability and the failure rate function under different loss functions and with several types of data. The prior laws used in this work are two types : non informative and natural conjugate.

In the first part, we use a classical approach of maximum likelihood and the Bayesian approach in the case of the right truncated Weibull distribution considering type II censored data. This distribution depends on three parameters : a parameter related to the truncation and two others specific to the Weibull distribution and have to do with form and scale.

The Baysian estimators and the posterior risks are obtained using symmetrical loss functions (the quadratic and the entropy loss functions) and then the asymmetrical loss function (the Linex loss function). In the case of the classical approach, the estimators are solutions of a non linear system whose solutions are not explicit analytically ; Numerical methods have been used.

In the Bayesian approach, the estimators are given in the form of a division of integrals. We have applied Monte-Carlo methods and in particular the Metropolis-Hastings algorithm to proceed with the simulations and the data analysis.

Finally, the Pitman and the integrated mean square error (IMSE) criteria are used to compare the Bayesian and the maximum likelihood estimators.

In the second part, we are intrested in estimating the parameter σ and some reliability features, such as the reliability function $S(t)$ and the failure rate function $h(t)$ of the Rayleigh distribution.

We have used the Bayesian approach under different loss functions (the quadratic and Linex loss functions), with type II then with progressive censored data. The prior distribution of the parameter is considered as a non informative law. Then we have also used a naturel conjucated law. The σ Bayesian estimators , $S(t)$ and $h(t)$, are obtained with their exact analytical expressions. The posterior risks are calculated in each case. A simulation study was carried out as well as real data analysis in order to compare between the different estimators through the consideration of their posterior risks. Our results leads us to conclude that the best estimator is obtained under Linex loss function.

ملخص

اعتمدنا في هذه الأطروحة على نموذجين : النموذج الأول هو نموذج weibull tronquée أما النموذج الثاني فهو نموذج Rayleigh.

التوزيعات المقتطعة تظهر لما يكون متغير عشوائي لا يتبع توزيع احتمالي, لكن جزء من فضاء العينة لا يمكن ملاحظته.

لما يكون لدينا متغير عشوائي محدود من اليمين بالقيمة T, نقول أن التوزيع مقتطع من اليمين, التوزيع الذي اهتممنا بدراسته هو توزيع Weibull tronquée à droite, الذي لديه خاصية انه يمتلك دالة نسبة العطب على شكل حوض.

توزيع Rayleigh له تطبيقات في مجالات عديدة خاصة في علم الأوبئة .

في هذه الأطروحة نهتم بدراسة مقدرات بايز للوسائط ,لدالة الموثوقية و لدالة نسبة العطب باستعمال عدة دوال أخطاء و عدة أنواع من المعطيات.

القانون البديهي المستعمل في هذه العمل هو نوعين:القانون البديهي الغامض و القانون البديهي المتزاوج الطبيعي.

في الجزء الأول استعملنا الطريقة الكلاسيكية (Méthode de maximum de vraisemblance) ثم الطريقة البايزية في حالة توزيع weibull tronquée à droite و باستعمال معطيات محذوفة من النوع II ,هذا التوزيع مرتبط بثلاث وسائط : الوسيط الأول عائد إلى الاقتطاع أما الآخرين فهما وسيطا نموذج Weibull العادي.

مقدرات بايز و الأخطار البعدية تحصلنا عليهم باعتمادنا على دوال أخطاء تناظرية (دالة الأخطاء التربيعية و دالة أخطاء الكون) ثم دوال أخطاء غير تناظرية (دالة أخطاء لينكس).في هذه الطريقة الكلاسيكية, المقدرات هم حلول لجملة معادلات غير خطية و ليس لديهم عبارات تحليلية واضحة و لهذا لإيجاد هذه المقدرات اعتمدنا على الطرق العددية.

في الطريقة البايزية , المقدرات عبارة عن نسب تكاملات , و قد استعملنا طرق Monte-Carlo بالخصوص خوارزم Metrpolis-Hastings للحصول على القيم العددية لهذه المقدرات ثم أجرينا محاكاة و تطبيق حقيقي لتوضيح النتائج و أخيرا استخدمنا مقاييس بيثمان و IMSE للمقارنة بين المقدرات الكلاسيكية و البايزية .

في الجزء الثاني اهتممنا بتقدير الوسيط σ , دالة الموثوقية R(t) و دالة نسبة العطب h(t) لنموذج Rayleigh. استخدمنا الطريقة البايزية تحت عدة دوال أخطاء (دالة الأخطاء التربيعية و دالة الأخطاء لينكس) مع معطيات محذوفة تدريجيا .

القانون البديهي الوسيط هو في الأول قانون بديهي غامض ثم قانون بديهي طبيعي مزدوج .

مقدرات بايز لh(t), R(t),σ متحصل عليهم في هذه الحالة بعبارات تحليلية واضحة .الأخطار البعدية تم حسابها في كل الحالات و تمت المقارنة بين المقدرات بالاعتماد على أخطارهم البعدية.

Introduction

La fiabilité a sans doute pris son développement depuis la dernière guerre mondiale. Elle est vite devenue une science a part entière dans les applications appartenant à de nombreux domaines. Elle a pour fondement mathématiques la statistique et le calcul de probabilité qui sont nécessaires à la compréhension et à l'analyse des données de fiabilité. Aujourd'hui, la fiabilité est devenue un paramètre clé de la qualité et d'aide à la décision, dans l'étude de la plupart des composants produit, et processus "grand public" : transport, énergie, composants électronique, composants mécanique...

Le terme fiabilité (ou la durée de survie) désigne le terme écoulé jusqu'à la survenue d'un évènement précis. l'évènement écoulé (communément appelé "décès") est le passage irréversible entre deux états (communément nommé "vivant" et "décès"). L'évènement terminal n'est pas forcément la mort : il peut s'agir de l'apparition d'une maladie (par exemple, le temps entre le diagnostic et la guérison), la panne d'une machine (durée de fonctionnement d'une machine, en fiabilité) ou la survenue d'un sinistre (temps entre deux sinistres, en actuariat). La fiabilité est l'une des composantes essentielles de la qualité d'un produit et elle est retenue entant que critère fondamental dés le stade de la conception.

La fiabilité est la caractéristique d'un dispositif exprimé par la probabilité que ce dispositif accomplisse une fonction requise dans des conditions d'utilisations et pour une période de temps déterminée. Cette thèse s'articule au tour de quatre chapitres, dans le premier chapitre, on donne quelques caractéristiques de la fiabilité puis on fait un rappel sur les méthodes numériques pour résoudre les systèmes d'équations non linéaire, et on termine ce chapitre en donnant quelques distributions de probabilités utiles en fiabilité, le deuxième chapitre, consiste à donner les principes de l'approche Bayésienne, et on explique comment faire un choix d'une distribution a priori finalement on donne quelques fonctions de perte symétriques et asymétriques. Le troisième chapitre est une étude des estimateurs de Bayes de la loi de Weibull tronquée sous différentes fonctions de pertes et en présence de données censurées de type II, on fait une étude de Monté-Carlo en utilisant l'algorithme de Metropolis-Hastings pour calculer ces estimateurs et on termine ce chapitre par une comparaison entre des deux approches (l'approche classique celle du maximum de vraisemblance et l'approche Bayésienne) à l'aide de critères de Pitman et IMSE, dans le dernier chapitre, on s'intéresse d'étudier l'estimateurs de Bayes de paramètre, de la fonction de fiabilité et de la fonction de taux de panne de la distribution de Rayleigh avec des données censurées de type II, enfin on fait une étude de simulation et une application par des données réelles pour terminer ce chapitre.

CHAPITRE 1

Outils mathématiques

Sommaire

1.1 Les 5 définitions équivalentes de la fiabilité

Cinq fonctions équivalentes définissent la loi de la durée : Supposons que la durée de survie X soit une variable positive ou nulle, et absolument continue. Alors sa loi de probabilité peut être définie par l'une des fonctions suivantes :

1- La fonction de survie S (La fonction de fiabilité)

Par définition :

$$S(t) = P(X \geq t); \ t \geq 0.$$

Pour t fixé, c'est la probabilité de survie jusqu'à l'instant t.

2- La fonction de répartition F

La fonction de répartition (f.r ou c.d.f en anglais pour "cumulative distribution function") est :

$$F(t) = P(X < t) = 1 - S(t).$$

Pour t fixé, c'est la probabilité de mourir avant l'instant t.

Il est arbitraire de décider que $S(t) = P(X \geq t)$ ou $S(t) = P(X > t)$ entraînant du même coup que $F(t) = 1 - S(t)$ vaut $F(t) = P(X < t)$ ou $F(t) = P(X \leq t)$. Lorsque la loi qui régit X continue, cela n'a aucune importance car ces deux quantités sont égales : $P(X \geq t) = P(X > t)$ et $P(X < t) = P(X \leq t)$. Cependant, dans les cas où S et donc F ont des sauts, ce qui arrive lorsque le temps est discret, compté en mois ou en semaine par exemple, on a quelque fois avantage à adopter la notion suivante qui évite toute ambiguïté :

$$S^-(t) = P(X \geq t) \qquad S^+(t) = P(X > t).$$

$$F^-(t) = P(X < t) \qquad F^+(t) = P(X \leq t).$$

3- La densité de probabilité f

C'est une fonction $f(t) \geq 0$, telle que pour tout $t \geq 0$,

$$F(t) = \int_0^t f(s)ds.$$

Si la fonction de répartition a une dérivée au point t alors :

$$f(t) = \lim_{dt \to 0} \frac{P(t \leq X < t + dt)}{dt} = F^{'}(t) = -S^{'}(t).$$

Pour t fixé, la densité de probabilité caractérise la probabilité de mourir dans un petit intervalle de temps après l'instant t.

4- Le risque instantané (le taux de hasard) h

Appelé "le taux de hasard" est défini comme :

$$h(t) = \lim_{dt \to 0} \frac{P(t \leq X < t + dt \mid X \geq t)}{dt} = \frac{f(t)}{S(t)}$$

Pour t fixé, $h(t)$ù caractérise la probabilité de mourir dans un petit intervalle de temps après l'instant t, conditionnellement au fait d'avoir survécu jusqu'à l'instant t. Aussi cela signifie-t-il le risque de mort instantané pour ceux qui ont survécus.

5- Le taux de hasard cumulé H

C'est l'intégrale du taux de hasard h :

$$H(t) = \int_0^t h(u)du = -ln(S(t)).$$

On peut déduire la fonction de taux de hasard cumulé grâce à la relation :

$$S(t) = exp(-H(t)) = exp(-\int_0^t h(u)du).$$

La définition de la distribution de probabilité de X repose sur l'une des cinq données suivantes, qui sont équivalentes : $S(t)$, $F(t)$, $f(t)$, $h(t)$ et $H(t)$.

Nous avons

$$S(t) = P(X > t)$$
$$= \mathbf{P}_{]0,t]}(1 - h(s)ds)$$
$$= exp(-\int_0^t h(s)ds)$$
$$= exp(-H(t)).$$

Par ailleurs les quantités suivantes sont parfois utilisées :

- la fonction de survie conditionnelle : $S(t \mid t_0) = P(X > t + t_0 \mid X > t_0)$.

- la durée de vie moyenne restante : $r(t) = E(X - t \mid X > t)$.

1.2 Quantités associées à la distribution de survie

1.2.1 Moyenne et variance de la durée de survie

Le temps moyen de survie $E(X)$ et la variance de la durée de survie $V(X)$ sont définis par les quantités suivantes :

$$E(X) = \int_0^\infty S(t)dt$$
$$V(X) = 2\int_0^\infty tS(t)dt - (E(X))^2$$

Ainsi on peut déduire l'espérance et la variance à partir de n'importe laquelle des fonctions : S, F, f, h, et H (mais pas l'inverse).

1.2.2 Quantiles de la durée de survie

- La médiane de la durée de survie est le temps t pour lequel la probabilité de survie $S(t)$ est égal à 0.5, c'est-à-dire, la valeur t_m qui satisfait $S(t_m) = 0.5$.

Dans le cas où l'estimateur est une fonction en escalier, il se peut qu'il y'ait un intervalle de temps vérifiant $S(t_m) = 0.5$. Il faut alors être prudent dans l'interprétation, si les deux évènements encadrant le temps médian sont éloignés. Il est possible d'obtenir un intervalle de confiance du temps médian. Soit $[B_i, B_s]$ un intervalle de confiance de niveau α du temps médian t_m est :

$$[S^{-1}(B_s), S^{-1}(B_i)]$$

-La fonction quantile de la durée de survie est définie par :

$$q(p) = inf(t : E(t) \geq p), \ \ 0 < p < 1.$$

$$= inf(t : S(t) \leq 1 - p)$$

Lorsque la fonction de répartition F est strictement croissante et continue alors :

$$q(p) = F^{-1}(p), \ \ 0 < p < 1.$$

$$= S^{-1}(1 - p).$$

Le quantile $q(p)$ est le temps où une proposition p de la population a disparue.

1.3 Censure et troncature

Une des caractéristiques des données de survie est l'existence d'observations incomplètes. En effet, les données sont souvent recueillies partiellement, notamment, à cause des processus de censure et troncature. Les données censurées ou tronquées proviennent du fait qu'on n'a pas accès à toute l'information : au lieu d'observer des réalisations indépendantes et identiquement distribuées (i.i.d) de durée X, on observe la réalisation de la variable X soumise à diverses perturbations indépendantes ou non du phénomène étudié.

1.3.1 Censure

La censure est le phénomène le plus couramment rencontré lors du recueil des données de survie :

Pour l'individu i, considérons :

- Son temps de survie X_i.

- Son temps de censure C_i.

- La durée réellement observée T_i.

Censure à droite

La durée de vie est dite censuré à droite si l'individu n'a pas subi l'évènement à sa dernière observation. En présence de censure à droite, les durées de vie ne sont pas toutes observées ; pour certaines d'entre elles, on sait seulement qu'elles sont supérieurs à une certaine valeur connue.

1- La censure de type I

Soit C une valeur fixée, au lieu d'observer les variables $X_1, ..., X_n$, on observe X_i uniquement lorsque $X_i \leq C$, sinon on sait uniquement que $X_i > C$. On utilise la notation suivante :

$$T_i = X_i \wedge C = min(X_i, C)$$

Ce mécanisme de censure est fréquemment rencontré dans les applications industrielles. Par exemple, on peut tester la durée de vie de n objets identiques (ampoules) sur un intervalle d'observations fixé $[0, u]$. En biologie, on peut tester l'efficacité d'une molécule sur un lot de souris (les souris au bout d'un temps u sont sacrifiées).

3- La censure de type II

Elle est présenté quand on décide d'observer les durées de survie de n patients jusqu'à ce que k d'entre eux soient décédés et d'arrêter l'étude à ce moment la. Soient $X_{(i)}$ et $T_{(i)}$ les statistiques d'ordre des variables X_i et T_i. La date de censure est donc $X_{(k)}$ et on observe les variables suivantes :

$$T_{(1)} = X_{(1)}$$

.

$$.$$

$$.$$

$$T_{(k)} = X_{(k)}$$

$$T_{(k+1)} = X_{(k)}$$

$$.$$

$$.$$

$$.$$

$$T_{(n)} = X_{(n)}$$

3- La censure de type III (ou censure aléatoire de type I)

Soient $C_1, ..., C_n$ des variables aléatoires i.i.d. On observe les variables

$$T_i = X_i \wedge C_i$$

L'information disponible peut être résumée par :

- La durée réellement observée T_i.

- Un individu $\delta_i = \mathbb{1}_{\{X_i \leq C_i\}}$.

- $\delta_i = 1$ si l'évènement est observé (d'où $T_i = X_i$). On observe les "vrais" durées ou les durées complètes.

- $\delta_i = 0$ si l'individu est censuré (d'où $T_i = C_i$). On observe des durées incomplètes (censurées).

La censure aléatoire est la plus courantes. Par exemple, lors d'un essai thérapeutique elle peut être engendrée par :

(a)- La perte de vue : le patient quitte l'étude en cours et on le revoit plus (à cause d'un déménagement, le patient décide de se faire soigner ailleurs). Ce sont des patient "perdue de vue".

(b)- L'arrêt ou le changement du traitement : les effet secondaires ou l'inefficacité du traitement peuvent entraîner un changement ou un arrêt du traitement. Ces patients sont exclus de l'étude.

(c)- La fin de l'étude : l'étude se termine alors que certains patients sont toujours vivants (ils n'ont pas subi l'évènement). Ce sont des patients "exclus-vivants". Les "perdus de vue" et les "exclus vivants" correspondent à des observations censurées mais les deux mécanismes sont de nature différentes (la censure peut être informative chez les "perdus de vue").

Censure à gauche

La censure à gauche correspond au cas où l'individu a déjà subi l'évènement avant que l'individu soit observé. On sait uniquement la date de l'évènement inférieure à une certaine date connue. Pour chaque individu, on peut associer un couple des variables aléatoire (T, δ) :

$$T = X \vee C = max(X, C).$$

$$\delta = \mathbb{1}_{X \geq C}.$$

Censure par intervalle

Une date est censurée par intervalle, si au lieu d'observer avec certitude le temps de l'évènement, la seule information disponible est qu'il a eu lieu entre deux dates connues.

1.3.2 Troncature

Les troncatures diffèrent des censures au sens où elles concernant l'échantillonnage lui même. Ainsi, une variable X est tronquée par un sous ensemble éventuellement aléatoire A de $\mathbb{R}^+$ si au lieu de X, on observe X uniquement si $X \in A$. Les points de l'échantillon "tronqué" appartiennent tous à A, et suivent donc la loi de T conditionnée par l'appartenance à A. Il ne faut pas confondre censure et troncature. S'il y a troncature, une partie des individus (donc des X_i) ne sont pas observables et on n'étudie qu'un sous échantillon (problème d'échantillonnage). Le biais de sélection est un cas particulier de troncature.

1- La troncature à gauche

Soit Z une variable aléatoire indépendante de X, on dit qu'il y'a troncature à gauche lorsque X n'est observable que si $X > Z$. On observe le couple (X, Z), avec $X > Z$.

2- La troncature à droite

De même, il y a troncature à droite lorsque X n'est observable que si $X < Z$.

3- La troncature par intervalle

Quand une durée est tronquée à droite et à gauche, on dit qu'elle est tronquée par intervalle.

1.3.3 Données progressivement censurées

Les données progressivement censurées peuvent être décrites par la méthode suivante : On suppose que n unités indépendantes sont mis dans un test et le censure $R = (r_1, r_2, ..., r_m)$ est déjà fixé. Lors du premier échec dit $X_{1:m:n}$, r_1 unités sont éliminées au hasard de $n-1$ unités restantes. Dans le deuxième échec, dit $X_{2:m:n}$, r_2 unités sont éliminés au hasard de $n - r_1 - 2$ unités restantes. On continue la procédure jusqu'au le m^{ime} échec, dit $X_{m:m:n}$, tous le reste $r_m = n - m - r_1 - ... - r_{m-1}$ sont éliminés. Puis, $X_{1:m:n} < X_{2:m:n} < ... < X_{m:m:n}$ sont appelés les statistiques d'ordre progressivement censurées.

Remarque

Pour générer les données progressivement censurées d'une distribution connue, on utilise l'algorithme de Balakrishnan et Sandhu avec les étapes suivantes :

1- générer m variables identiquement indépendantes distribuées $(u_1, u_2, ..., u_m)$ de la loi uniforme $U(0, 1)$.

2- Soit $z_i = -log(1 - u_i)$, z_i sont identiquement indépendantes distribuées de la distribution exponentielle standard.

3- En donnant les censures R $(R_1, R_2, ..., R_m)$, soit $y_1 = \frac{z_1}{m}$, et pour $i = 1, ..., m$

$$y_i = y_{i-1} + \frac{z_i}{n - \sum_{j=1}^{i-1} R_j - i - 1}$$

Donc, $(y_1, y_2, ..., y_m)$ sont des données progressivement censurées d'un échantillon $U(0, 1)$.

4- Soit $w_i = 1 - exp(-y_i)$.

5- Soit $x_i = F^{-1}(w_i)$ i.e $w_i = F(x_i)$, donc, $(x_1, x_2, ..., x_m)$ sont des données progressivement censurées de la distribution qu'on veut générer.

1.4 Méthodes d'estimation numériques

Dans cette section, on suppose que les données $x_1, ..., x_n$ sont n réalisations indépendantes d'une même variable aléatoire sous-jacente X. Il est équivalent de supposer $X_1, ..., X_n$ indépendantes et de même loi. Nous adopterons ici la seconde formule, qui est la plus pratique à manipuler.

Les techniques de statistique descriptive, comme l'histogramme ou le graphe de probabilité, permettent de faire des hypothèses sur la nature de la loi de probabilité X_i. Des techniques statistiques plus sophistiquées, les tests d'adéquation, permettent de valider ou pas ces hypothèses.

On supposera ici que ces techniques ont permis d'adopter une famille de loi de probabilité bien précise pour la loi des X_i.

On notera θ le paramètre inconnu, le problème traité est celui de l'estimation du paramètre θ. Comme on l'a déjà dit, il s'agit de donner, au vu des observations $x_1, ..., x_n$, une approximation ou une évaluation de θ que l'on espère la plus proche possible de la vraie valeur inconnue.

Il existe de nombreuse méthodes pour estimer un paramètre θ.

Par exemple, les estimations graphiques à partir des graphes de probabilité. Dans cette section, nous ne nous intéressons qu'aux deux méthodes d'estimation les plus usuelles, la méthodes des moments et la méthode de maximum de vraisemblance.

Mais il faut d'abord définir précisément ce que sont une estimation et surtout un estimateur.

Pour estimer θ on ne dispose que des données $x_1, ..., x_n$, donc une estimation de θ sera une fonction de ces observations. Une statistique t est une fonction des observations $x_1, ..., x_n$:

$$t : \mathbb{R}^n \to \mathbb{R}^m$$

$$(x_1, ..., x_n) \mapsto t(x_1, ..., x_n)$$

Pour simplifier les écritures, on note souvent $t_n = t(x_1, ..., x_n)$ et $T_n = t(X_1, ..., X_n)$.
Par abus, on donne le même nom de statistique aux deux quantités, mais dans une perspective d'estimation, on va nommer différemment t_n et T_n.

Un estimateur d'une grandeur θ est une statistique T_n à valeurs dans l'ensemble des valeurs possibles de θ. Une estimation de θ est une réalisation t_n de l'estimateur T_n

1.4.1 La méthode des moments

C'est la méthode la plus naturelle, que nous avons déjà utilisée sans la formaliser. L'idée de base est d'estimer une espérance mathématique par une moyenne empirique, une variance par une variance empirique, etc...
Si le paramètre à estimer est l'espérance de la loi de X_i, alors on peut l'estimer par la moyenne empirique de l'échantillon. Autrement dit, si $\theta = E(X)$, alors l'estimateur de θ par la méthode des moments (EMM) est $\widehat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$.
Plus généralement, pour $\theta \in \mathbb{R}$, si $E(\theta) = \varphi(\theta)$, où φ est une fonction inversible, alors l'estimateur de θ par la méthode des moments est $\widehat{\theta}_n = \varphi^{-1}(\overline{X}_n)$.
De la même manière, on estime la variance de la loi des X_i par la variance empirique de l'échantillon $S_n^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}_n^2$. Plus généralement, si la loi de deux paramètres θ_1 et θ_2 tels que $(E(X), var(X)) = \varphi(\theta_1, \theta_2)$, où φ est une fonction inversible alors les estimateurs de θ_1 et θ_2 par la méthode des moments sont $(\widehat{\theta}_{1n}, \widehat{\theta}_{2n}) = \varphi^{-1}(\overline{X_n}, S_n^2)$.
Ce principe peut naturellement se généraliser aux moment de tous ordres, centrés ou non centrés : $E((X - E(X))^k)$ et $E(X^k), k \geq 1$.

1.4.2 La méthode du maximum de vraisemblance

Soit $X_1, X_2, ..., X_n$, n observations indépendantes et de même loi (cas des données complètes), la fonction de maximum de vraisemblance est calculée comme suit :

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

Pour n quelconque, il est logique de dire que la valeur la plus vraisemblance est la valeur pour laquelle la probabilité d'observer $x_1, x_2, ..., x_n$ est la plus forte possible. Cela revient à faire comme si c'était l'éventualité la plus probable qui s'était produite au cours de l'expérience.

L'estimateur de maximum de vraisemblance de θ est la valeur $\widehat{\theta}_n$ de θ qui rend maximale la fonction de vraisemblance $L(\theta; x_1, ..., x_n)$

L'estimateur de maximum de vraisemblance (EMV) de θ est la variable aléatoire correspondante.

Donc $\widehat{\theta}_n$ sera en général calculé en maximisant la log-vraisemblance

$$\widehat{\theta}_n = argmax_{\theta} L(\theta; x_1, ..., x_n).$$

Quand $\theta = (\theta_1, ..., \theta_d) \in R^d$ et que toutes les dérivées partielles ci-dessus existent, $\widehat{\theta}_n$ est la solution du système d'équations appelées équations de vraisemblance :

$$\forall j \in \{1, ..., d\}, \frac{\partial}{\partial \theta_j} lnL(\theta; x_1, ..., x_n) = 0$$

A priori, une solution de ce système d'équations pourrait être un minimum de la vraisemblance. Mais on peut montrer que la nature d'une fonction de vraisemblance fait que c'est bien un maximum que l'on obtient. Il est fréquent que le système des équations de vraisemblance n'ait pas de solution explicite. Dans ce cas, on le résout par des méthodes numériques, comme la méthode de Newton-Raphson. En R, la maximisation numérique peut se faire à l'aide des commandes **optim** et **BBsolve**...

Lorsqu'on a des données censurées, la fonction de vraisemblance est calculée à l'aide de la formule suivante :

$$L(\theta; x_1, ..., x_n) = \frac{n!}{(n-m)!} \prod_{i=1}^{m} f(x_i)(1 - F(x_m))^{n-m}.$$

et lorsque les données sont progressivement censurées, la vraisemblance est :

$$L(\theta; x_1, ..., x_n) = A \prod_{i=1}^{m} f(x_i)(1 - F(x_i))^{R_i}$$

$$R = (R_1, R_2, ..., R_m)$$

$$A = n(n - 1 - R_1)(n - 2 - R_1 - R_2)...(n - \sum_{i=1}^{m}(R_i + 1)).$$

1.4.3 Estimateur sans biais et de variance minimale (ESBVM)

Un estimateur T_n de θ sera un bon estimateur s'il est suffisamment proche, en un certain sens de θ. Il faut donc définir une mesure de l'écart entre θ et T_n. On appelle cette mesure le risque de l'estimateur. On a intérêt à ce que le risque de l'estimateur soit le plus petit possible.

Par exemple, les risques $T_n - \theta$, $|T_n - \theta|$, $(T_n - \theta)^2$ expriment bien l'écart entre T_n et θ. Mais comme il est plus facile d'utiliser des quantités déterministes que les quantités aléatoires, on s'intéresse en priorité aux espérances des quantités précédentes. En particulier :
Le biais de T_n est $E(T_n) - \theta$
Le risque quadratique ou erreur quadratique moyenne est donné par :

$$EQM(T_n) = E[(T_n - \theta)^2]$$

Un estimateur T_n de θ est sans biais si et seulement si $E(T_n) = \theta$.
Il est biaisé si et seulement si $E(T_n) \neq \theta$. Le biais mesure une erreur systématique d'estimation de θ par T_n. Par exemple, si $E(T_n) - \theta < 0$, cela signifie que T_n aura tendance à sous-estimer θ.
L'erreur quadratique moyenne s'écrit :

$$EQM(T_n) = E[(T_n - \theta)^2] = E[(T_n - E(T_n) + E(T_n) - \theta)^2]$$

$$= E[(T_n - E(T_n))^2] + 2E[T_n - E(T_n)]E[E(T_n) - \theta] + E[(E(T_n) - \theta)^2].$$

$$= var(T_n) + E[E(T_n) - \theta]^2$$

$$= \text{variance de l'estimateur} + \text{carré de son biais.}$$

Si T_n est un estimateur sans biais, $EQM(T_n) = var(T_n)$. On a donc intérêt à ce qu'un estimateur soit sans biais et de faible variance. Par ailleurs, on en déduit immédiatement que de deux estimateurs, le meilleur est celui qui a la plus petite variance.

La variance d'un estimateur mesure sa variabilité. Si l'estimateur est sans biais. Cette variabilité est autour de θ. Si on veut estimer correctement θ, il ne faut pas que cette variabilité soit trop forte.

1.5 Méthodes pour résoudre les systèmes d'équation non linéaire

1.5.1 La méthode de Newton-Raphson

L'itération de Newton-Raphson est fréquemment utilisée dans la statistique dans l'estimation par la méthode de maximum de vraisemblance. Soit θ est le paramètre ; $l(\theta; x)$ est le log de la fonction de vraisemblance. On suppose que $l(\theta; x)$ est deux fois différentiable par rapport à θ.

On définit la fonction score :

$$U(\theta; x) = \frac{\partial l(\theta; x)}{\partial \theta}.$$

L'information observée est :

$$I(\theta; x) = -\frac{\partial^2 l(\theta, x)}{\partial \theta^2}$$

et l'information attendue est :

$$J(\theta) = E[I(\theta; x)].$$

L'estimateur de maximum de vraisemblance $\widehat{\theta}$ est la valeur de θ qui maximise $l(\theta; x)$. On suppose que la vraisemblance n'a pas un maximum global unique. Pour trouver l'estimateur de maximum de vraisemblance, on résout l'équation

$$U(\theta; x) = 0$$

par rapport à θ, on suppose que la solution est $\widehat{\theta}$. Si $I(\widehat{\theta}) > 0$ alors $\widehat{\theta}$ est le maximum. Si on sait que $l(\theta; x)$ a un seul maximum, donc $\widehat{\theta}$ doit être ce maximum, autrement $\widehat{\theta}$ peut être un maximum local.

L'itération de Newton-Raphson pour résoudre $U(\theta) = 0$ est :

$$\theta_{k+1} = \theta_k + \frac{U(\theta_k; x)}{I(\theta_k; x)}.$$

Où θ_0 est la valeur initiale de θ.

Il est souvent le cas que $J(\theta)$ est beaucoup plus facile à évaluer que $I(\theta; x)$. Pour cette raison, l'itération est modifiée pour l'utiliser $J(\theta)$ au lieu de $I(\theta, x)$ comme suit

$$\theta_{k+1} = \theta_k + \frac{U(\theta_k, x)}{J(\theta_k)}.$$

Cette méthode est appelée la méthode de Fisher-Scoring.

La méthode de Fisher-Scoring nécessite généralement plus d'itérations que la méthode de Newton-Raphson parce que $J()$ est plus simple à calculer que $I()$.

Dans certains cas $I(\theta; x)$ ne dépend pas de x donc $I(\theta; x) = J(\theta)$ et la méthode de Newton-Raphson est identique à la méthode de Fisher-Scoring.

Le cas multiparamétrique est une extension simple. Si θ est un vecteur $p \times 1$, $U(\theta; x)$ devient un vecteur de $p \times 1$, $I(\theta; x)$ et $J(\theta)$ sont des matrices $p \times p$. L'itération de Newton-Raphson est :

$$\theta_{k+1} = \theta_k + I(\theta_k; x)^{-1} U(\theta_k; x),$$

et Fisher-Scoring est :

$$\theta_{k+1} = \theta_k + J(\theta_k)^{-1} U(\theta_k; x),$$

et la matrice de variance de $\widehat{\theta}$ est estimé par $I(\widehat{\theta}; x)^{-1}$ ou $J(\widehat{\theta}; x)^{-1}$.

1.5.2 L'algorithme EM

L'algorithme EM (expectation maximization) est un outil très puissant et utile pour l'analyse des données incomplètes ; voir McLachlan et Krishnan (2008), pour une discussion détaillée sur cette méthode. L'algorithme consiste à deux étapes, l'étape "Expectation"

(E-step) et l'étape "Maximization" (M-step). Tout d'abord, la vraisemblance de données complètes pour le problème donné est formé. Puis, à l'étape E, l'espérance conditionnelle de log de vraisemblance avec des données complètes est obtenu, comptent des données incomplètes observée et la valeur courante du paramètre, scalaire ou vectorielle valeur. Ce log-vraisemblance attendue est essentiellement fonction du paramètre impliqué et la valeur actuelle du paramètre dans lesquelles l'espérance a été calculé. Ainsi, si la variable T est la durée de vie sous-jacente, et X est la variable observée, donc dans l'étape "E", notre but est d'obtenir :

$$Q(\theta, \theta^{(k)}) = E_{\theta}^{(k)}[logL_c(t;\theta)|x].$$

Où $L_c(t;\theta)$ est la vraisemblance avec des données complètes, θ est le vecteur de paramètre d'intérêt, et $\theta^{(k)}$ est la valeur du paramètre à la k^{ime} étape de l'algorithme (la valeur actuelle du paramètre).

Dans l'étape (M-step), la log-vraisemblance attendue pour des données complètes $Q(\theta, \theta^{(k)})$ est maximisé par-rapport à θ au-dessus l'espace de paramètre θ pour obtenir le meilleur estimateur du paramètre en tant que :

$$\theta^{(k+1)} = argmax_{\theta \in \Theta} Q(\theta, \theta^{(k)}).$$

1.6 Les distributions de probabilités utiles en fiabilité

Dans ce paragraphe, nous présenterons quelques distributions de vie qui interviennent le plus fréquemment dans l'analyse des données de vie et qui sont communes à plusieurs disciplines. Nous parlerons en particuliers des lois continues. Nous énoncerons les principales propriétés de ces lois (densité de probabilité, fonction de fiabilité et taux de défaillance) ainsi que leurs applications en fiabilité ([Afnor, 1988], [Ayyub and Mccuen, 1997],...).

1.6.1 La loi exponentielle

Cette loi a de nombreuses applications dans plusieurs domaines. C'est une loi simple, très utilisée en fiabilité dont le taux de défaillance est constant. Elle décrit la vie des matériels qui subissent des défaillances brutales.
La densité de probabilité d'une loi exponentielle de paramètre λ s'écrit :

$$f(t) = \lambda e^{-\lambda t}.$$

La fonction de fiabilité

$$S(t) = e^{-\lambda t}.$$

Le taux de défaillance est constant dans le temps

$$\lambda(t) = \lambda.$$

Propriétés sans mémoire de la loi exponentielle :
Une propriété principale de la loi exponentielle est d'être sans mémoire ou "Memory less property" en anglais, ((Bon, 1995), (Leemis, 1994)) :

$$P(X \geq t + \Delta t | X \geq t) = \frac{e^{-\lambda(t+\Delta t)}}{e^{-\lambda t}} = e^{-\lambda \Delta t} = P(X \geq \Delta t),\ t > 0,\ \Delta > 0.$$

Ce résultat montre que la loi conditionnelle de la durée de vie d'un dispositif qui a fonctionné sans tomber en panne jusqu'à l'instant t est identique à la loi de la durée de vie d'un nouveau dispositif est considéré comme neuf (ou "as good as new" en anglais), de durée de vie exponentielle de paramètre λ.

1.6.2 La loi Log-normale

Une variable aléatoire continue et positive X est distribuée selon une loi log-normale si son logarithme népérien est distribué suivant une loi normale. Cette distribution est largement utilisée pour modéliser des données de vie, en particulier les défaillances par fatigue en mécanique.

La densité de probabilité d'une loi log-normale de paramètres positifs μ et σ est :

$$f(t) = \frac{1}{\sigma t}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{log(t)-\mu}{\sigma})^2},\ t > 0.$$

La fonction de fiabilité

$$S(t) = 1 - \Phi(\frac{log(t)-\mu}{\sigma}),$$

$$S(t) = 1 - \frac{1}{\sigma\sqrt{2\pi}}\int_0^t \frac{1}{x}e^{-\frac{1}{2}(\frac{log(x)-\mu}{\sigma})^2}dx.$$

Φ : Fonction de répartition de la loi normale centrée réduite.

Le domaine de définition n'étant jamais négatif, il n'y a aucune limitation à l'emploi de la distribution log-normale en fiabilité. Le taux de défaillance est croissant dans le début de vie puis décroissant en tendant vers zéro et la distribution est très dissymétrique.

1.6.3 La loi de Weibull

C'est la plus populaire des lois utilisées dans plusieurs domaines (électronique, mécanique,...). Elle permet de modéliser en particulier de nombreuses situations d'usure de matériel. Elle caractérise le comportement du système dans les trois phases de vie, période de jeunesse, période de vie utile et période d'usure ou vieillissement dépend des trois paramètres suivants : β, η et γ. La densité de probabilité d'une loi de Weibull a pour expression

$$f(t) = \frac{\beta}{\eta}(\frac{t-\gamma}{\eta})^{\beta-1}e^{-(\frac{t-\gamma}{\eta})}, t \geq \gamma.$$

Où : β est le paramètre de forme ($\beta > 0$).

η est le paramètre d'échelle ($\beta > 0$).

γ est le paramètre de position ($\gamma \geq 0$).

La fonction de fiabilité s'écrit :

$$S(t) = e^{-(\frac{t-\gamma}{\eta})^\beta}$$

Le taux de défaillance est donné par :

$$\lambda(t) = \frac{\beta}{\eta}(\frac{t-\gamma}{\eta})^{\beta-1}$$

Suivant les valeurs de β, le taux de défaillance est soit décroissant ($\beta < 1$) soit constant ($\beta = 1$), soit croissant ($\beta > 1$).

La distribution de Weibull permet donc de représenter les trois périodes de la vie d'un dispositif décrites par la courbe en baignoire.
Le cas $\gamma > 0$ correspond à des dispositifs dont la probabilité de défaillance est nulle jusqu'à un certain âge γ.

1.6.4 La loi gamma

La loi gamma est la loi de l'instant d'occurrence du α^{me} évènement dans un processus de Poison.
Soit $T_{i_{\{i=1,\alpha\}}}$ le vecteur représentant les durées inter-évènements (les temps entre défaillances successives d'un système). Si ces durée sont dés variables aléatoires indépendantes et identiquement distribuées selon une loi exponentielle de paramètre β, alors le taux cumulé d'apparition de α défaillances suit une loi gamma de paramètres (α, β). Sa densité de probabilité s'écrit

$$f(t) = \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)},\ t \geq 0,\ \alpha \geq 1\ et\ \beta \geq 0.$$

Le taux de défaillance est donné par :

$$\lambda(t) = \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\int_t^\infty \Gamma(\alpha) f(u) du}.$$

La loi gamma est très utilisée dans l'approche Bayésienne, elle est la conjuguée naturelle de la loi exponentielle de paramètre λ. Un cas particulier intéressant consiste, pour un entier naturel n fixé, à choisir les paramètres : $\alpha = \frac{n}{2}$ et $\beta = \frac{1}{2}$ la loi obtenue est celle de Khi-deux à n degrés de liberté.

CHAPITRE 2

L'approche Bayésienne

Sommaire

Dans de nombreuses situations d'expériences aléatoires, il semble raisonnable d'imaginer que le praticien a une certaine idée du phénomène aléatoire qu'il est en train d'observer. Or, la démarche statistique classique repose essentiellement sur un principe de vraisemblance qui consiste à considérer que ce qui a été observé rend compte de manière exhaustive du phénomène. Mais l'observation ne fournit qu'une image et celle ci peut être mauvaise. Certes cet inconvénient est en générale gommé par les considérations asymptotiques et un certain nombre de théorème permettent d'évaluer la bonne qualité des estimateurs si le nombre d'observations est suffisant.

L'analyse Bayésienne des problèmes statistiques propose d'introduire dans la démarche d'inférence, l'information dont dispose a priori le praticien. Dans le cadre de la statistique paramétrique, ceci se traduira par le choix d'une loi sur le paramètre d'intérêt. Dans l'approche classique, le modèle paramétrique $(\mathcal{X}, \mathcal{A}, P_\theta, \theta \in \Theta)$. Ayant un a priori sur le paramètre, modélisé par une densité de probabilité que nous noterons $\pi(\theta)$, on "réactualise" cet a priori au vu de l'observation en calculant la densité a posteriori $\pi(\theta|x)$, et c'est à partir de cette loi que l'on mène l'inférence.

On peut alors, par exemple, de manière intuitive pour le moment retenir l'espérance mathématique ou encore le mode de cette densité a posteriori comme l'estimateur de θ.

Le paramètre θ devient donc en quelque sorte une variable aléatoire, à laquelle on associe une loi de probabilité dite loi a priori.

On sent bien d'emblée que les estimateurs bayésiens sont très dépendants du choix de la loi a priori.

Différentes méthodes existent pour déterminer ces lois a priori. On peut se référer à des techniques bayésiennes empiriques, où l'on construit la loi a priori sur la base d'une expérience passée, usant de méthodes fréquentistes, pour obtenir formes et valeurs des paramètres pour cette loi. Nous verrons que l'on peut aussi modéliser l'absence d'information sur le paramètre au moyen des lois dites non informative.

L'approche Bayésienne se différencie donc de l'approche classique dans le sens où le paramètre θ n'est plus considéré comme étant totalement inconnu, il est devenu une variable aléatoire dont le comportement est supposé connu. On fait intervenir dans l'analyse statistique une distribution associée à ce paramètre. On appelle modèle statistique bayésien ; la donnée d'un modèle statistique paramétrique $(\mathcal{X}, \mathcal{A}, P_\theta, \theta \in \Theta)$ avec $f(x|\theta)$ densité de P_θ et d'une loi $\pi(\theta)$ sur le paramètre.

La démarche de l'analyse Bayésienne conduit au calcul d'une loi a posteriori $\pi(\theta|x)$; actualisation de la loi a priori $\pi(\theta)$ au vu de l'observation.

Ce calcul repose sur la version continue du théorème de Bayes

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{f(x)},$$

$f(x|\theta)$ désignant la loi de l'observation ou vraisemblance et $f(x)$ la loi marginale ou prédictive

$$f(x) = \int_\Theta f(x|\theta)\pi(\theta)d\theta$$

2.1 Estimation ponctuelle

2.1.1 Coût et décision

Le problème très général auquel on s'intéresse ici est celui d'un individu plongé dans un environnement donné (nature) et qui, sur la base d'observations, est conduit à mener des actions et à prendre des décisions qui auront un coût.

Les espaces intervenant dans l'écriture d'un modèle de décision sont :

$\mathcal{X}$: l'espace des observations.

Θ : l'espace des états de la nature (l'espace des paramètres dans le cas d'un problème statistique)

$\mathcal{A}$: l'espace des actions ou décisions, dont les évènements sont des images de l'observation par une application δ appelée règle de décision (une statistique (i.e fonction des observations) dans le cas d'un problème statistique)

$\mathcal{D}$: l'ensemble des règles de décisions δ, applications de $\mathcal{X}$ dans $\mathcal{A}$ (les estimateurs possibles). On note a une action. On a $a = \delta(x)$.

L'inférence consiste à choisir une règle de décision $\delta \in \mathcal{D}$ concernant $\theta \in \Theta$ sur la base d'une observation $x \in \mathcal{X}$, x et θ étant liés par la loi $f(x|\theta)$.

En statistique, la règle de décision est un estimateur, l'action est une estimation (valeur de l'estimateur au point d'observation).

Pour choisir une décision, construit une relation de préférence en considérant une mesure du coût ou perte encourue lorsqu'on prend la décision $\delta(x)$ et que l'état de la nature est θ.

Pour faire on introduit la fonction L, appelée fonction de coût (ou de perte) définie de la manière suivante : On appelle fonction de coût, toute fonction L de $\Theta \times \mathcal{D}$ dans $\mathbb{R}$.

$L(\theta, a)$ évalue le coût d'une décision a 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. Il s'agit d'une fonction de θ. Un coût négatif correspond à un gain.

2.1.2 Risque fréquentiste

On dira qu'une décision est une bonne décision si elle conduit à un coût nul.

Autrement dit, une bonne décision est solution de l'équation

$$L(\theta, \delta(x)) = 0$$

θ étant inconnu, on ne peut évidemment pas résoudre cette équation. Classer les décisions par la seule considération du coût est donc impossible. Celui-ci ne prend pas compte l'information apportée par le modèle $f(x|\theta)$. Ces remarques conduisent à considérer la moyenne de la perte, c'est le risque fréquentiste.

Définition

On appelle risque fréquentiste le coût moyen (l'espérance mathématique) du coût d'une règle de décision

$$R(\theta, \delta) = E_\theta(L(\theta, \delta)) = \int L(\theta, \delta) dP_\theta(x)$$

On dira que δ_1 est préférable à δ_2 et on note $\delta_1 < \delta_2$ si :

$$R(\theta, \delta_1) \leq R(\theta, \delta_2),$$

cette définition permet d'établir un préodre sur l'ensemble $\mathcal{D}$ des décisions.

Cependant, ce préodre est partiel puisqu'il ne permet pas de comparer deux règles de décision telles que :

$$R(\theta_1, \delta_1) < R(\theta_1, \delta_2) \;\; et \;\; R(\theta_2, \delta_1) > R(\theta_2, \delta_2).$$

2.1.3 Risque de Bayes

Puisque l'approche Bayésienne met à la disposition du statisticien une loi a priori $\pi(\theta)$, on peut considérer la moyenne du risque fréquentiste i.e la moyenne du coût moyen suivant la loi a priori : $E^\pi(R(\theta, \delta(x)))$.

Il s'agit du risque bayésien ou risque de Bayes que l'on note $r(\pi, \delta)$. On a :

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

On définit alors le coût a posteriori $\rho(\pi, \delta(x))$ comme étant la moyenne du coût par rapport à la loi a posteriori

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

Il s'agit d'une fonction de x.

On a le résultat suivant : Le risque de Bayes $r(\pi, \delta)$ est la moyenne de coût a posteriori $\rho(\pi, \delta(x))$ suivant la loi marginale $f(x)$.

Preuve

$$r(\pi, \delta) = \int_{\Theta} \int_{\mathcal{X}} L(\theta, \delta(x)) f(x|\theta)\pi(\theta) dx d\theta$$

Où $f(x|\theta)\pi(\theta) = \pi(\theta|x)f(x)$, on a donc ;

$$\begin{aligned} r(\pi, \delta) &= \int_{\mathcal{X}} \int_{\Theta} L(\theta, \delta(x))\pi(\theta|x) d\theta f(x) dx \\ &= \int_{\mathcal{X}} \rho(\pi, \delta(x)) f(x) dx. \end{aligned}$$

2.2 Distribution a posteriori

La distribution a posteriori est la quantité la plus importante dans l'inférence bayésienne. Il contient toutes les informations disponibles sur le paramètre θ inconnu après avoir observé les données $X = x$.

(distribution a posteriori)

Soit $X = x$ désigne la réalisation observée d'une (éventuellement multi variée) variable aléatoire X avec la fonction de densité $\pi(x|\theta)$.

Soit $\pi(\theta)$ est la densité a priori qui nous permet de calculer la fonction de densité $\pi(\theta|x)$ de la distribution a posteriori en utilisant le théorème de Bayes :

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta) d\theta}$$

Le terme $f(x|\theta)$ est la fonction de vraisemblance. Puisque θ est aléatoire, nous conditionnons explicitement sur une valeur spécifique θ et on écrit $L(\theta) = f(x|\theta)$.

Le dénominateur peut être écrit comme :

$$\int f(x|\theta)\pi(\theta) d\theta = \int f(x, \theta) d\theta = f(x).$$

qui souligne que cela ne dépend pas de θ. La quantité $f(x)$ est connue comme la probabilité marginale et est important pour le choix du modèle Bayésien.

La densité de la distribution a posteriori est donc proportionnelle au produit de la vraisemblance est la densité a posteriori de la distribution avec une constante de proportionnalité $\frac{1}{f(x)}$. Ceci est habituellement noté

$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta).$$

Où $\propto$ représente "est proportionnel à" et implique que $\frac{1}{\int L(\theta)\pi(\theta)d\theta}$ est la constante de normalisation pour assurer que : $\int \pi(\theta|x)d\theta$=1. telle que $\pi(\theta|x)$ est une fonction de densité valide. L'inférence statistique à propos de θ est basée uniquement sur la distribution a posteriori. L'estimation ponctuelle appropriés sont les paramètres de la localisation, telles que la moyenne, médiane ou le mode de la distribution a posteriori.
La moyenne a posteriori $E(\theta|x)$ est l'espérance de la distribution a posteriori

$$E(\theta|x) = \int \pi(\theta|x)dx.$$

Le mode a posteriori $Mod(\theta|x)$ est le mode de la distribution a posteriori

$$Mod(\theta|x) = argmax_{\theta}\pi(\theta|x).$$

La médiane a posteriori $Med(\theta|x)$ est la médiane de la distribution a posteriori ce est le nombre α qui vérifié

$$\int_{-\infty}^{\alpha} \pi(\theta|x)d\theta = 0.5 \ \ et \ \ \int_{\alpha}^{\infty} \pi(\theta|x)d\theta = 0.5$$

Implicitement, on suppose souvent que la moyenne a posteriori est finie, auquel cas il est également unique. Cependant, le mode a posteriori et la médiane a posteriori ne sont pas nécessairement unique. En effet, une distribution a posteriori peut avoir plusieurs modes et donc est appelée multimodale. L'estimation Bayésienne par intervalle est également dérivée de la distribution a posteriori. Pour les distinguer des intervalles de confiance, qui ont une interprétation différente, ils sont appelés intervalles de crédibilité. Ici la définition est pour un paramètre scalaire θ.

(Intervalle de crédibilité)

Pour un $\gamma \in [0, 1]$ fixé, un $\gamma - 100$ intervalle de crédibilité est définie par deux nombres réels

t_l et t_u qui remplissent :

$$\int_{t_l}^{t_u} f(\theta|x)d\theta = \gamma.$$

La quantité γ est appelée le niveau crédible de l'intervalle de crédibilité $[t_l, t_u]$.

2.3 Choix de la distribution a priori

L'inférence Bayésienne permet la spécification probabiliste des croyances antérieures par le biais d'une distribution préalable.

Il est souvent utile et justifié de restreindre l'éventail des possibles distributions a priori a une famille spécifique avec un ou deux paramètres. Le choix de cette famille peut être basée sur le type de fonction de vraisemblance rencontré.

2.3.1 La distribution a priori conjuguée

Une approche pragmatique de choisir une distribution a priori est de sélectionner un membre d'une famille spécifique de distributions telles que la distribution a posteriori appartient à la même famille. Elle est appelée distribution a posteriori conjuguée.

(La distribution a priori conjuguée)

Soit $L(\theta) = f(x|\theta)$ est la fonction de maximum de vraisemblance basée sur l'observation $X = x$. La classe G des distributions est appelée conjuguée par rapport à $L(\theta)$, si la distribution a posteriori $f(\theta|x)$ est dans G pour tout x chaque fois que la distribution $\pi(\theta)$ est dans G.

La famille G ={toutes les distributions} est conjuguée trivialement par rapport à toutes fonction de vraisemblance. Dans la pratique, on essaie de trouver des petits ensembles G qui sont spécifiques à la probabilité $L_x(\theta)$.

Avant d'étudier les distributions a priori conjuguées, on note qu'il suffit d'étudier conjugaison pour un membre X_i d'un échantillon aléatoire $X_{1:n}$. En effet, si l'a priori est conjuguée, la partie a posteriori après avoir observé la première observation, est par définition, du

même type de sert de nouvelle distribution a posteriori, incorporant désormais la deuxième observation, n'est à niveau dans une classe conjuguée,... etc.
Seulement les paramètres de la distribution changeront dans un tel traitement séquentiel des données.
Le tableau suivant donne quelques exemples des distributions a priori avec la fonction de vraisemblance correspondante :

La vraisemblance	La distribution a priori conjuguée	La distribution a posteriori
$X\|\theta \sim Bin(n,\theta)$	$\theta \sim Be(\alpha,\beta)$	$\theta\|x \sim Be(\alpha+x,\beta+n-x)$
$X\|\theta \sim Geom(\theta)$	$\theta \sim Be(\alpha,\beta)$	$\theta\|x \sim Be(\alpha+1,\beta+x-1)$
$X\|\pi \sim Po(e\lambda)$	$\lambda \propto G(\alpha,\beta)$	$\lambda\|x \sim G(\alpha+x,\beta+e)$
$X\|\pi \sim exp(\lambda)$	$\lambda \sim G(\alpha,\beta)$	$\lambda\|x \sim G(\alpha+1,\beta+x)$
$X\|\mu \sim N(\mu,\sigma^2 connu)$	$\mu \sim N(\upsilon,\tau^2)$	$\mu\|X \sim N((\frac{1}{\sigma^2}+\frac{1}{\tau^2})^{-1}(\frac{x}{\sigma^2}+\frac{\upsilon}{\tau^2}),(\frac{1}{\sigma^2}+\frac{1}{\tau^2})^{-1})$
$X\|\sigma^2 \sim N(\mu connu,\sigma^2)$	$\sigma^2 \sim IG(\alpha,\beta)$	$\sigma^2 \sim IG(\alpha+\frac{1}{2},\beta+\frac{1}{2}(x-\mu)^2)$

TAB. 2.1 – **Quelques distributions a priori pour différentes fonctions de vraisemblance**

2.3.2 Distribution a priori impropre

La distribution a priori a un influence sur la distribution a posteriori. Si on veut minimiser l'influence de la distribution a priori, il est courant de spécifier une a priori "vague", par exemple avec une très grande variance.
Dans la limite cela peut conduire à une distribution préalable mauvaise avec une fonction de densité qui n'intègre pas. En raison de la constante de normalisation manquante, ces fonctions de densité sont généralement spécifiées en utilisant le signe de proportionnalité " $\propto$ ".
Si on utilise l'a priori impropre, alors il est nécessaire de vérifier que au moins la distribution a priori est propre. Si ça le cas, alors l'a priori impropre peuvent être utilisée dans l'analyse Bayésienne.

(Distribution a priori impropre)

Une distribution a priori avec une fonction de densité $\pi(\theta) \geq 0$ est appelée impropre si,

$$\int_{\Theta} \pi(\theta) d\theta = \infty$$

Ou

$$\sum_{\theta \in \Theta} \pi(\theta) = \infty$$

Pour un paramètre θ continue ou discret, respectivement.

2.3.3 Distribution a priori de Jeffrey

Dans certaines situations, il peut être utile de choisir une distribution a priori qui ne donne pas beaucoup d'informations sur le paramètre en raison de la connaissance préalable faible ou manquante.

Un premier choix naïf est un a priori uniforme $\pi_\theta(\theta) \propto 1$, dans ce cas, l'a posteriori est proportionnelle à la fonction de vraisemblance. Notez qu'un a priori uniforme localement sera incorrecte si l'espace de paramètre est pas borné.

Cependant, il existe des problèmes liés à cette approche. On suppose que $\phi = h(\theta)$ est une transformation différentiable de θ, qui a une loi a priori localement uniforme avec une densité $\pi_\theta(\theta) \propto 1$. On utilise un changement de variable, on obtient l'a priori correspondant à ϕ.

$$\pi_\phi(\phi) = \pi_\theta\{h^{-1}(\phi)\} \mid \frac{dh^{-1}(\phi)}{d\phi} \mid$$

$$\propto \mid \frac{dh^{-1}(\phi)}{d\phi} \mid .$$

On note que ce terme est nécessairement constant. En effet, $\pi_\phi(\phi)$ sera indépendante de ϕ seulement si h est linéaire. Si h est non linéaire, la densité a posteriori $\pi_\phi(\phi)$ dépend de ϕ et ne sera pas (localement) uniforme. Cependant, si nous avons choisi une para métrisation avec ϕ dès le départ, nous avons choisi une a priori uniforme (localement) $\pi_\phi(\phi) \propto 1$.

(La distribution a priori de Jeffrey)

Soit X une variable aléatoire avec une fonction de maximum de vraisemblance $L(x|\theta)$ où θ

est le paramètre inconnu.

L'a priori de Jeffrey est définie comme suit

$$\pi(\theta) \propto \sqrt{J(\theta)}$$

Où $J(\theta)$ est **l'information de Fisher**. Cette dernière équation est nommée : la règle de Jeffrey.

L'a priori de Jeffrey est proportionnelle à la racine carrée de l'information de Fisher, qui donne une distribution a priori impropre.

A première vue, il est surprenant que ce choix est censé être invariant par para métrisation, mais nous avons le voir dans le résultat suivant

Résultat (Invariance de la loi a priori de Jeffrey)

L'a priori de Jeffrey est invariant sous la para métrisation de θ si :

$$\pi_\theta(\theta) \propto \sqrt{J_\theta(\theta)}$$

Donc, la fonction de densité de $\phi = h(\theta)$ est :

$$\pi_\phi(\phi) \propto \sqrt{J_\phi(\phi)},$$

où $J_\phi(\phi)$ est l'information de Fisher de ϕ.

Preuve

On a $\pi_\theta \propto \sqrt{J_\theta(\theta)}$ et on fait un changement de variable, on a

$$\pi_\phi(\phi) = \pi_\theta(\theta) \mid \frac{dh^{-1}(\phi)}{d\phi} \mid$$

$$\propto \sqrt{J_\theta(\theta)} \mid \frac{dh^{-1}(\phi)}{d\phi} \mid = \sqrt{J_\theta(\theta)\{\frac{dh^{-1}(\phi)}{d\phi}\}^2}$$

$$= \sqrt{J_\theta(\theta)}$$

2.4 Fonctions de perte

2.4.1 Fonction de perte quadratique

La fonction de perte quadratique est la fonction définie par

$$L(\theta, d) = (\theta - d)^2$$

Une variante de cette fonction de perte est une fonction de perte quadratique pondérée (fonction de perte quadratique généralisée) de la forme

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

Sous l'hypothèse d'un coût quadratique, l'estimateur de Bayes $\delta^\pi(x)$ de θ associé à la loi a priori π est la moyenne a posteriori de θ

$$\delta^\pi(x) = E_{\pi(.|x)}(\theta) = \int_{\theta\in\Theta} L(\theta, \delta(x))\pi(\theta|x)d\theta$$

Preuve

Par définition, l'estimateur de Bayes minimise le coût a posteriori i.e $\rho(\pi, \delta) = E^{\pi(.|x)}(L(\theta, \delta(x)))$. Sous l'hypothèse d'un coût quadratique, on a :

$$\rho(\pi, \delta) = E^{\pi(.|x)}((\theta - \delta(x))^2)$$

$$= E^{\pi(.|x)}(\theta^2) - 2\delta(x)E^{\pi(.|x)}(\theta) + \delta^2(x).$$

Il s'agit d'un polynôme du second degré en $\delta(x)$. Il sera minimum en $E^{\pi(.|x)}(\theta)$.

2.4.2 Fonction de perte absolue

La fonction de perte absolue est la fonction définie par :

$$L(\theta, \delta(x)) = |\theta - \delta(x)|$$

Un estimateur de Bayes associé à π et la fonction de perte absolue, est un fractile d'ordre $\frac{k_2}{k_1+k_2}$ de $\pi(\theta|x)$.

Preuve

$$E^{\pi(.|x)}(L(\theta,\delta(x))) = \int_{\Theta} L(\zeta-\delta(x))\pi(\zeta|x)d\zeta$$

$$= \int_{\delta(x)}^{+\infty} k_2(\zeta-\delta(x))\pi(\zeta|x)d\zeta + \int_{-\infty}^{\delta(x)} k_1(\delta(x)-\zeta)\pi(\zeta|x)d\zeta$$

On remarque que :

$$\pi(\zeta|x)d\zeta = dF(\zeta|x) - d(1-F(\zeta|x)),$$

et on écrit, donc :

$$E^{\pi(.|x)}(L(\theta,\delta(x))) = [k_2(\zeta-\delta(x))(1-F(\zeta|x))]_{\delta(x)}^{+\infty} +$$

$$\int_{\delta(x)}^{+\infty} k_2 P^{\pi(.|x)}(\theta>\zeta)d\zeta + [k_1(\delta(x)-\zeta)F(\zeta|x)]_{-\infty}^{\delta(x)} + \int_{-\infty}^{\delta(x)} k_1 P^{\pi(.|\zeta)}d\zeta$$

$$= \int_{\delta(x)}^{+\infty} k_2 P^{(.|x)}(\theta>\zeta)d\zeta + \int_{-\infty}^{\delta(x)} k_1 P^{(.|x)}(\theta|\zeta)d\zeta$$

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

$$\frac{\partial}{\partial\delta(x)}\rho(\pi,\delta(x)) = -k_2 P^{\pi(.|x)}(\theta>\delta(x)) + k_1 P^{\pi(.|x)}(\theta<\zeta) = 0$$

$$\Leftrightarrow -k_2(1-P^{\pi(.|x)}(\theta>\delta(x))) + k_1 P^{\pi(.|x)}(\theta<\zeta) = 0$$

$$\Leftrightarrow (k_1+k_2)P^{\pi(.|x)}(\theta<\delta(x)) - k_2 = 0$$

d'où

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

et le coût est donc maximum pour $\widetilde{\theta} = \delta(x)$ tel que $P^{\pi(.|x)}(\theta<\delta(x)) = \frac{k_2}{k_1+k_2}$. On rappelle : si f est intégrable sur [a,b] alors $F(x) = \int_{-\infty}^{x} f(t)dt$, $a<x<b$, est différentiable.

2.4.3 Fonction de perte $0-1$

la fonction de perte $0-1$ est l'application L définie par

$$L(\theta,\delta(x)) = \begin{cases} 0, & si \;\; \theta \in \Theta_0 \\ 1, & si \;\; \theta \in \Theta_1 \end{cases}$$

On trouve en utilisant cette fonction de perte, les résultats de la théorie des tests d'hypothèses. Un problème de test est un problème de choix (de prise de décision) entre $H_0 : \theta \in \Theta_0$ et $H_1 : \theta \in \Theta_1$.

On définit donc de la manière suivante :

$\delta(x) = 1$: On accepte H_0

$\delta(x) = 0$: On rejette H_0 (ce ci ne dépend pas de θ)

On a un espace d'action de la forme $\mathcal{A} = \{0, 1\}$

Soit w la région de rejet i.e le sous-ensemble de $\mathcal{X}$ qui conduit à rejeter H_0. On peut construise une fonction de coût de la manière suivante : supposons $\theta \in \Theta_0$.

Si $X \in W$, on prend la décision de rejeter i.e $\delta(x) = 0$, mais la décision n'est pas bonne on va pénaliser et $L(\theta, \delta(x)) = 1$.

Si $X \ n'appartient \ pasdansW$, on ne rejette pas, on prend la décision $\delta(x) = 1$, la décision est bonne $L(\theta, \delta(x)) = 0$.

Le coût s'écrit donc :

$$L(\theta, \delta(x)) = \begin{cases} 1 - \delta(x), & si \ \theta \in \Theta_0 \\ \delta(x), & si \ \theta \in \Theta_1 \end{cases}$$

Ce qu'on peut écrire : $L(\theta, \delta(x)) = 1(x \in W)$ et on calcule la fonction de risque :

$$R(\theta, \delta) = E(L(\theta, \delta(x))) = \int_{\aleph} L(\theta, \delta(x)) dp_\theta(x) = P_\theta(x \in W), \theta \in \Theta_0.$$

On retrouve le risque de première espèce.

2.4.4 La fonction de perte Linex

Une fonction de perte asymétrique très pratique est la fonction de perte Linex (Linear exponential). Elle a été introduite par Varian (1975). Cette fonction de perte presque exponentiellement d'un coté de zéro sous l'hypothèse que la perte minimale est obtenu pour $\widehat{\delta}(x) = \theta$, la fonction de perte linex pour θ, soit a

$$L(\Delta) \propto e^{a\Delta} - a\Delta - 1, a \neq 0$$

Où : $\Delta = (\delta(x) - \theta)$ où $\delta(x)$ est un estimateur du θ. Le signe de a représentant respectivement la direction et le degré de symétrie ($a > 0$: la surestimation est plus grave que la sous-estimation et vice versa). Pour a proche de zéro, la perte Linex est approximativement la fonction de perte quadratique :

$$E_\theta(L(\delta(x)-\theta)) \propto e^{a\delta(x)}E_\theta(e^{-a\theta}) - a(\delta(x) - E_\theta(\theta) - 1)...(*)$$

Où $E_\theta(.)$ représente l'espérance a posteriori relative à la densité a posteriori de θ.

L'estimateur de Bayes $\delta_\pi(x)$ qui minimise $(*)$. Pour trouver l'estimateur, nous dérivons l'équation $(*)$ par rapport à $\delta(x)$, nous obtenons

$$\frac{d}{d\delta(x)}(E_\theta(L(\delta(x)-\theta))) = ae^{a\delta(x)}E_\theta(e^{-a\theta}) - a$$

En égalant cette expression à zéro, nous obtenons

$$e^{-a\delta(x)}E_\theta(e^{-a\theta}) = a.$$

Alors, l'estimateur de Bayes $\widehat{\delta}_L(x)$ sous la fonction de perte Linex est :

$$\delta(x) = -\frac{1}{a}log(E_\theta(e^{-a\theta}))$$

étant donné que $E_\theta(e^{-a\theta})$ existe et est finie.

2.4.5 Fonction de perte de DeGroot

DeGroot (1970) a introduit plusieurs types des fonctions de perte et est obtient les estimateurs de Bayes sous cette fonction de perte. Un exemple d'une fonction de perte symétrique est la fonction de perte de DeGroot définie par

$$L(\theta,\delta(x)) = (\frac{\theta-\delta(x)}{\delta(x)})^2.$$

Sous cette fonction de perte, l'estimateur de Bayes est

$$\delta_\pi(x) = \frac{E^\pi(\theta^2|x)}{E^\pi(\theta|x)}.$$

2.4.6 Fonction de perte d'entropie

Galabria et Pulcini (1994) ont proposé une fonction de perte qui découle de la fonction de perte Linex appelée la fonction de perte entropie est définie par

$$L_E(\theta,d) \propto (\frac{d}{\theta})^p - pln(\frac{d}{\theta}) - 1,$$

qui a minimum lorsque $d = \theta$.

L'estimateur de Bayes de paramètre θ sous cette fonction de perte est

$$\delta(x) = (E_\theta(\theta^{-p}))^{\frac{-1}{p}}$$

$a)-$ Lorsque $p = 1$, l'estimateur de Bayes coïncide avec l'estimateur de Bayes sous la fonction de perte quadratique pondéré $\frac{(d-\theta)^2}{\theta}$.

$b)-$ Lorsque $p = -1$, l'estimateur de Bayes coïncide avec l'estimateur de Bayes sous la fonction de perte quadratique.

2.5 Méthodes de simulation

2.5.1 Notions préliminaires

Chaîne de Markov

Nous résumons ici quelques propriétés des chaînes de Markov qui seront utiles pour la suite. On appelle chaîne de Markov à temps discret toute suite de variables aléatoires $\{X_n\}_{n\geq 0}$ provenant d'un certain support S et respectant la propriété markovienne :

$$P(X_n|X_{n-1}, ..., X_0) = P(X_n|X_{n-1})$$

Autrement dit, la valeur d'une variable aléatoire de cette suite ne dépend que de celle qui la précède. Nous nous intéresserons ici aux chaînes dites homogènes, où les probabilités $P(X_n|X_{n-1})$ sont indépendantes de la valeur de n. Ainsi, une chaîne de Markov homogène est complètement définie par la valeur ou la distribution de X_0 et les probabilités de transition en un pas $P_{i,j} = P(X_n = j|X_{n-a} = i), n \geq 1$, habituellement regroupées dans une matrice $P = \{P_{i,j}\}$, définie sur l'ensemble des valeurs $i, j \in S$ (l'espace d'états). Finalement on définit les probabilités de transition en k pas ainsi : $P^k_{i,j} = P(X_k = j|X_o = i)$.

2.5.2 Quelques critères de classification

Pour qu'une chaîne ait des propriétés asymétriques intéressantes, on souhaitera que tous ses états soient récurrents positif et apériodiques.
On dira alors que la chaîne est ergodique. Un état i est dit récurrent positifs si :

$$P_{ii}^{(n)} \nrightarrow 0, quand \ n \to \infty.$$

Un état est apériodique si sa période vaut 1, la période étant définie comme étant le plus grands commun diviseur de $n \in \mathbb{N} : P_{ii}^{n} > 0$. Pour vérifier ces propriétés, il sera plus simple de passer par notion de classe d'états, que nous définissons maintenant.
On dit qu'un état j est accessible à partir d'un état i, s'il est possible en partant de i d'atteindre l'état j en un nombre fini de transitions. Formellement :

$$i \to j \Leftrightarrow \exists n \geq 0, \ \ telque \ \ P_{i,j}^{(n)} > 0.$$

Si l'on a à la fois $i \to j$ et $j \to i$, on dira que i et j communiquent. On montre alors facilement que les états communiquant entre eux fortement une classe d'équivalence sur l'ensemble des états de la chaîne. Les états d'une même classe possèdent plusieurs propriétés si l'un d'entre eux l'est.

2.5.3 Théorème ergodique et distribution stationnaire

La distribution stationnaire lorsqu'elle existe est une distribution de probabilité prenant valeur sur S et respectant la condition :

$$\pi = \pi P \Leftrightarrow j : \pi_j = \sum_i \pi_i P_{ij}$$

Dans le cas d'une chaîne ergodique, la distribution stationnaire existera toujours, et on aura alors, pour tout les états : $i, j \in S$

$$P_{ij}^{(n)} \to \pi_j > 0.$$

Cette distribution et ses propriétés seront primordiales pour les développement qui ont suivre. Mentionnons seulement pour l'instant qu'une chaîne de Markov ergodique convergera en loi vers sa distribution stationnaire peu importe la valeur initiale X_0.

Pour un espace d'états continus, les distributions pourront être exprimées par les distributions continues conditionnelles de densité $P(x_n|x_{n-1}), n \geq 1$, ou encore sous forme de distributions conjointes de densité $P(x_{n-1}, x_n)$. Par souci de concision, nous ne reprendrons pas notions précédentes dans le cas d'un espace continu. Mentionnons seulement que toutes les propriétés et résultats énoncés ont leur équivalent, à quelques considérations techniques près dans cette nouvelle situation.

2.6 Méhodes de Monte-Carlo

Les méthodes de Monte-Carlo sont des techniques d'échantillonnage aléatoire numériques visant à calculer des intégrales. Le problème classique est le suivant : Soit $h(x), x \in \mathbb{R}^m$ une fonction quelconque et $f(x)$ une fonction de densité de support $\mathcal{X} \subset \mathbb{R}^m$.

On cherche à calculer :

$$I = E_f(h(x)) = \int_{A \in \mathcal{X}} h(x) f(x) dx$$

Pour ce faire, on génère un certain nombre n de variable $x_i, i = 1, ..., n$ i.i.d de densité f, avec les quelles on estime I par

$$\overline{h_n} = \frac{1}{n} \sum_{j=1}^{n} h(x_i)$$

Propriétés

1$-$ Par la loi des grands nombres, avec probabilité 1

$$\overline{h}_n \rightarrow E(h).$$

2$-$ A la condition que $E_g(h^2)$ soit finie

$$V(\overline{h}_n) = \frac{1}{n} V(h(x)) = \frac{1}{n} \int h^2(x) f(x) dx - E_f^2(h(x)).$$

3− Sous la même condition, par le théorème de la limite centrale

$$\frac{\overline{h}_n - 1}{\sqrt{V(h(x))}} \to \mathcal{N}(0,1).$$

Ces techniques seront surtout utiles dans les problèmes en grandes dimension, où les méthodes numériques traditionnelles perdent de leur efficacité. La difficulté sera de trouver une façon de générer efficacement un échantillon de variables i.i.d de densité f.

2.7 Méthodes MCMC

Les méthodes de Monte-Carlo par Chaîne de Markov permettent d'élargir grandement l'éventail des distributions pouvant être simulés numériquement. Elles sont relativement simples à implémenter et ne requièrent souvent qui le connaissance de la fonction de densité cible à une constante près, ce qui les rend intéressantes dans de nombreuses situations.
Cependant, une implémentation naïve peut mener à des temps de calcul très longs, puisque la convergence de cette méthodes est relativement lente lorsqu'elles ne sont pas bien calibrées à une situation donnée.
Nous verrons d'abord sommairement les justifications théoriques de ces méthodes, puis nous les illustrerons par la présentation de l'algorithme original de Metropolis-Hastings, pour ensuite voir les améliorations successives que l'on peut y apporter leurs particularités et leur validité théorique.
L'idée de base est de simuler une distribution de densité f en utilisant une chaîne de Markov ergodique $\{X_t\}$ dont la distribution stationnaire est f. Le théorème ergodique garantit alors la convergence en loi de $\{X_t\}$ vers une variable de densité f et par conséquent, pour presque toute valeur initiale X_0 :

$$\frac{1}{n}\sum_{t=1}^{n} h(X_t) \to E_f(h)$$

Nous appellerons MCMC toute méthode permettant de simuler une distribution en utilisant une chaîne de Markov ergodique ayant celle-ci comme distribution stationnaire. Pour construire un tel algorithme, il faut donc déterminer un ensemble de probabilités de transition P approprié, c'est-à-dire irréductible, ergodique et ayant la bonne distribution stationnaire.

Nous aurons besoin pour la suite du résultats suivant : Soit une chaîne de Markov ayant comme probabilité de transition P et une distribution de probabilité $\pi(.)$ définie sur le même espace d'états S. Si P possède la propriété de réversibilité par rapport à π :

$$\forall x, y \in S : \pi(x)P(x,y) = \pi(y)P(y,x),$$

alors la distribution stationnaire de la chaîne de Markov est π

Démonstration

On aura stationnarité si : $\forall y \in S$

$$\int_S \pi(dx)P(x,y) = \pi(y).$$

Or, sous l'hypothèse de réversibilité :

$$\forall y : \int_S \pi(dx)P(x,y) = \int_S \pi(y)P(y,dx) = \pi(y) \int_S P(y,dx) = \pi(y).$$

On utilisera cette propriété pour construire les probabilités de transition appropriées.

2.7.1 L'algorithme de Metropolis-Hastings

L'algorithme ne nécessite qu'une valeur de départ X_0 et le choix d'une distribution conditionnelle de densité $q(x|y) = q(y|x)$. A une étape donnée t, les manipulations suivants sont effectuées.

Algorithme MH

1− A partir de la valeur $X_t = x$, on génère $Y_{t+1} = y$ selon la distribution de densité $q(y|x)$.

2− On pose :

$X_{t+1} = Y_{t+1}$ avec probabilité $\alpha(x, y)$.

$X_{t+1} = X_t$ avec probabilité $1 - \alpha(x, y)$.

où les seuils α doivent avoir la forme générale :

$$\alpha(x, y) = \frac{s(x, y)}{1 + r(x, y)}.$$

$r(x, y)$ est le ratio $\frac{\pi(x)q(x,y)}{\pi(y)q(y,x)}$, et la fonction s est choisie de façon à ce que $s(x, y) = s(y, x)$ et $0 \leq \alpha(x, y) \leq 1$. Habituellement on utilise exclusivement

$$\alpha(x, y) = min\{1, \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)}\}.$$

qui correspond au choix $s(x, y) = min\{1 + r(x, y), 1 + r(y, x)\}$. Si en plus la densité q est symétrique ($q(y|x) = q(x|y)$), le rapport devient tout simplement :

$$\alpha(x, y) = min\{1, \frac{\pi(y)}{\pi(x)}\}.$$

Nous nommerons q la densité de proposition des candidats, ou densité instrumentale, et α les probabilités d'acceptation de ces derniers. Voyons maintenant les propriétés théoriques de cette procédure.

Propriétés

L'algorithme tel que défini génère une chaîne de Markov dont les probabilités de transition sont données par :

$$P(x, y) = q(x, y)\alpha(x, y), si \;\; x \neq y.$$

$$P(x, y) = 1 - \int P(x, y)dy, autrement.$$

Pour prouver que la distribution stationnaire de cette chaîne est π, il suffit de montrer qu'elle est réversible par rapport à π. Or

$$\pi(x)P(x, y) = \pi(x)q(x, y)\alpha(x, y).$$

$$= \frac{\pi(x)q(x, y)s(x, y)}{1 + \frac{\pi(x)q(x,y)}{\pi(y)q(y,x)}}.$$

$$= \frac{\pi(x)\pi(y)q(x, y)q(y, x)s(x, y)}{\pi(y)q(y, x) + \pi(x)q(x, y)}$$

$$= \frac{\pi(y)q(y, x)s(y, x)}{\frac{\pi(y)q(y,x)}{\pi(x)q(x,y)} + 1}$$

$$= \pi(y)q(y, x)\alpha(y, x) = P(y)P(y, x).$$

Maintenant, il faut s'assurer que la chaîne converge bien vers sa distribution stationnaire, c'est-à-dire qu'elle est ergodique. Or, ceci est facilement vérifié la plupart du temps.

Par exemple, si $q(x, y)$ est positive pour toute pair (x, y) appartenant au support de π,

$P(x, y)$ aussi sera toujours positive. Ainsi, à partir d'une valeur X_t donnée, toute valeur X_{t+1} sera atteignable en une seule étape avec une probabilité positive. La chaîne est donc irréductible. La chaîne sera aussi apériodique des moment qu'il existe au moins une paire (x, y) tel que $\alpha(x, y) < 1$, car on aura alors $P(x, x) > 0$.
Cela sera pratiquement toujours vrais et on conclut donc que la chaîne est effectivement ergodique.

2.7.2 L'échantillonnage de Gibbs

Introduit par Geman (1984) dans le cadre de la restauration d'images. Le principe repose encore sur une décomposition du problème générale (simuler suivant un certain loi) en une série de problèmes élémentaires (simuler suivant des lois conditionnelle). Considérons la densité $f(x, y_1, ..., y_p)$. On s'intéresse à la loi marginale

$$f(x) = \int ... \int f(x, y_1, ..., y_p) dy_1 ... dy_p.$$

En particulier, on souhaite obtenir son espérance mathématique et sa variance. On se place ici dans le cas où les intégrations impliquées dans le calcul de la marginale sont compliquées et difficiles à effectuer même numériquement. On supposera cependant que la densité conditionnelle est disponible.
L'échantillonneur de Gibbs va nous permettre de générer x suivant $f(x)$ sans utiliser directement son expression supposée difficile à manipuler, mais en utilisant les densités conditionnelles.
Ainsi fabriquant un échantillon $(x_1, ..., x_m)$ assez grand on pourra approximer la moyenne, la variance et autre caractéristique en utilisant une loi de grands nombre

$$\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} g(x_i) = E(g(x)).$$

Principe de l'algorithme

Considérons le cas élémentaire : $f(x, y)$. On suppose $f(x|y)$ et $f(y|x)$ disponibles. On peut alors générer ce qu'on appellera une séquence de Gibbs de la manière suivante : partant d'une valeur x_0, on génère y_0 avec $\pi(.|x_0)$, puis x_1 avec $\pi(.|y_0)$ puis y_1 avec $\pi(.|x_1)$ et ainsi de suite.

Après M itérations de ce schéma, il vient une séquence : $(x_0, y_0, x_1, y_1, ..., x_M, y_M)$. Pour M assez grand, x_M est une réalisation de X.

Dans le cadre Bayésien, l'algorithme de Gibbs va permettre d'obtenir une réalisation du paramètre $\theta = (\theta_1, ..., \theta_m)$ avant la loi a posteriori $\pi(\theta|x)$ dès que l'on est capable d'expression les lois conditionnelles : $\pi(\theta_i|\theta_j; x)$; $j \neq i$; l'échantillonnage de Gibbs consiste à :

Partant d'un vecteur initial $\theta^{(0)} = (\theta_1^{(0)}, ..., \theta_m^{(0)})$,

à la $(p+1)^{ime}$ étape, disposant du vecteur $\theta^{(p)} = (\theta_1^{(p)}, ..., \theta_m^{(p)})$,

simuler

$$\theta_1^{(p+1)} = \pi(\theta_1|\theta_2^{(p)}, \theta_3^{(p)}, ..., \theta_m^{(p)}; x)$$

$$\theta_2^{(p+1)} = \pi(\theta_2|\theta_1^{(p+1)}, \theta_3^{(p)}, ..., \theta_m^{(p)}; x)$$

$$.$$

$$.$$

$$.$$

$$\theta_m^{(p+1)} = \pi(\theta_m|\theta_1^{(p+1)}, \theta_2^{(p)}, ..., \theta_{m-1}^{(p)}; x)$$

Les itérations successives de cet algorithme génèrent successivement les états d'une chaîne de Markov $\theta^{(p)}, p > 0$ à valeur dans $\mathbb{N}^{\otimes m}$.

La probabilité de transition de $\theta^{'}$ vers θ a pour expression :

$$K(\theta^{'}, \theta) = \pi(\theta_1|\theta^{'}_2, ..., \theta^{'}_m) \times \pi(\theta_2|\theta_1, \theta^{'}_2, ..., \theta^{'}_m)$$

$$\times \pi(\theta_3|\theta_1, \theta_2, \theta^{'}_4, ..., \theta^{'}_m) \times ... \times \pi(\theta_m|\theta_1, ..., \theta_{m-1})$$

On montre que cette chaîne admet une mesure invariante qui est la loi a posteriori. Pour un nombre d'itération suffisamment grand, le vecteur θ obtenu peut donc être considéré comme étant une réalisation de la loi a posteriori.

Soit f une densité de probabilité. Etant donnés une densité g telle que

$$\int_Z g(x,z)dz = f(x).$$

et un entier $p > 1$ tel que les densités conditionnelles de $g(y) = g(y_1, ..., g(y_p))$, $g_1(y_1|y_2, ..., y_p)$, $g_2(y_2|y_1, y_3, ..., y_p)$,...,$g_p(y_p|y_1, ..., y_{p-1})$, soient simulables, l'algorithme de l'échantillonnage de Gibbs associé à cette décomposition est fourni par la transition $y^{(t)}$ à y^{t+1} suivante :

Simuler

$$y_1^{t+1} \sim g_1(y_1|y_2^{(t)}, ..., y_p^{(t)})$$

$$y_2^{t+1} \sim g_2(y_1|y_3^{(t)}, ..., y_p^{(t)})$$

$$.$$

$$.$$

$$.$$

$$y_p^{t+1} \sim g_p(y_1|y_1^{(t+1)}, ..., y_{p-1}^{(t+1)})$$

Lorsque la complétion de f en g n'est pas nécessaire, un choix relativement restreint consiste en la section du nombre p de composantes des sous vecteurs (y_1,y_2,...,y_p) (qui ne sont pas nécessairement scalaires). Le choix est cependant souvent limité pour des raisons pratiques de simulation, en particulier lorsque y est de petite dimension.

CHAPITRE 3

Modèle de Weibull tronqué à droite

Sommaire

3.1 Introduction

Les distributions tronquées surviennent quant une variable aléatoire X suit une distribution de probabilité connue, mais une portion de l'espace d'échantillon ne peut pas être observée. Si les valeurs de la variable aléatoire sont limitées à droite par une valeur T, on dit que la distribution est tronquée à droite.

La distribution tronquée à laquelle on s'intéresse est la distribution de Weibull tronquée à droite ; cette dernière a la particularité d'avoir un taux de panne en baignoire.

Plusieurs auteurs ont étudié l'estimation des paramètres et des caractéristiques de fiabilité dans le cas d'une loi de Weibull tronquée : Dallas R.Wingo (1988) a proposé l'estimation paramétrique de la distribution de Weibull tronquée et il a utilisé comme méthode d'estimation, la méthode de maximum de vraisemblance. Shalaby O.A et El-Youcef (1993) ont utilisé les données de Sinha et Kale (1980) pour trouver les estimateurs Bayesiens des paramètres et de la fonction de fiabilité sous la fonction de perte quadratique, Shalaby O.A (1992) a étudié le problème du risque a posteriori des paramètres de la loi de Weibull tronquée. Plus récemment ; Teiling Zhang et Min Xie (2012) ont étudié d'une façon très simplifié les caractéristiques de cette loi, et ils ont fait une comparaison entre la loi de Weibull tronquée, la loi de Weibull modifiée et la loi de Weibull étendue. Balackrishnan et Mitra (2012) ont étudié les problèmes de l'estimation de maximum de vraisemblance, la prédiction et les intervalles de confiances pour une loi de Weibull tronquée à gauche, Yeliz MertKantar et Ilham Usta (2015) ont traité la modélisation aléatoire dans un modèle de Weibull tronqué.

On s'intéresse dans ce chapitre à l'estimation des paramètres de la loi de Weibull tronquée à droite. En effet une telle loi dépend de trois paramètres ; un paramètre lié à la troncature et les deux autres propres à la loi de Weibull à deux paramètres. Les approches utilisées sont dans une première partie une approche classique celle du maximum de vraisemblance où on utilise des packages du R pour trouver les estimateurs de maximum de vraisemblance des paramètres. La deuxième approche est une approche bayesienne avec différentes fonctions de perte et une loi a priori sur les paramètres qui est une mixture entre loi a priori non-informative et loi a priori conjuguée naturelle.

Ce chapitre s'articule autour de quatre sections, dans la première section on définit les caractéristiques de la loi de Weibull tronquée à droite, la deuxième section porte sur l'estimation du maximum de vraisemblance des paramètres et la troisième section traite l'estimation Bayesienne des paramètres, sous différentes fonctions de pertes et une loi a priori. Dans la quatrième section, on fait une étude de Monte-Carlo et on utilise les méthodes MCMC (en particulier l'algorithme de Metropolis-Hastings) pour calculer les estimateurs bayesiens et les risques a posteriori des paramètres puis on a comparé entre les estimateurs sous les trois fonctions de pertes en considérant que le meilleur estimateur qui a le risque a posteriori minimum. Enfin, on fait une étude comparative entre les estimateurs de Bayes et de maximum de vraisemblance des paramètres à l'aide des critères de Pitman, Integrated mean square error (IMSE).

3.2 Le modèle

On considère une variable aléatoire de la loi de Weibull tronquée à droite avec les paramètres $(\alpha > 0, \lambda > 0)$, la fonction de densité est donnée par :

$$g(t; \alpha, \lambda, T) = \frac{f(t)}{F(T)}, \quad 0 \leq t \leq T,$$

où : f est la fonction de densité de la loi de Weibull de paramètres α et λ $(\alpha > 0, \lambda > 0)$ et F est sa fonction de répartition donc :

$$f(t; \alpha, \lambda, T) = (\frac{\lambda}{\alpha})(\frac{t}{\alpha})^{\lambda-1} e^{-(\frac{t}{\alpha})^\lambda}$$

$$F(t) = 1 - e^{-(\frac{t}{\alpha})^\lambda}$$

donc la densité de la loi de Weibull tronquée à droite est :

$$g(t; \alpha, \lambda, T) = \frac{(\frac{\lambda}{\alpha})(\frac{t}{\alpha})^{\lambda-1} e^{-(\frac{t}{\alpha})^\lambda}}{1 - e^{-(\frac{T}{\alpha})^\lambda}}, \quad 0 \leq t \leq T.$$

Et la fonction de répartition est :

$$G(t; \alpha, \lambda, T) = \frac{F(t)}{F(T)}$$

$$= \frac{1 - e^{-(\frac{t}{\alpha})^\lambda}}{1 - e^{-(\frac{T}{\alpha})^\lambda}}, 0 \leq t \leq T.$$

La fonction de fiabilité correspondante est définie comme suit

$$S(t; \alpha, \lambda, T) = 1 - G(t; \alpha, \lambda, T) = 1 - \frac{1 - e^{-(\frac{t}{\alpha})^\lambda}}{1 - e^{-(\frac{T}{\alpha})^\lambda}}$$

$$S(t; \alpha, \lambda, T) = 1 - G(t; \alpha, \lambda, T) = \frac{e^{-(\frac{t}{\alpha})^\lambda} - e^{-(\frac{T}{\alpha})^\lambda}}{1 - e^{-(\frac{T}{\alpha})^\lambda}}, \ \ 0 \leq t \leq T.$$

La fonction de taux de hasard est :

$$h(t; \alpha, \lambda, T) = \frac{g(t; \alpha, \lambda, T)}{S(t; \alpha, \lambda, T)}$$

$$h(t; \alpha, \lambda, T) = \frac{g(t; \alpha, \lambda, T)}{S(t; \alpha, \lambda, T)} = \frac{(\frac{\lambda}{\alpha})(\frac{t}{\alpha})^{\lambda-1} e^{-(\frac{t}{\alpha})^\lambda}}{e^{-(\frac{t}{\alpha})^\lambda} - e^{-(\frac{T}{\alpha})^\lambda}}, \qquad 0 \leq t \leq T.$$

Le moment d'ordre k de la distribution de Weibull tronquée est donné par la formule suivante :

$$E(t^k) = \int_0^\infty t^k g(t; \alpha, \lambda, T) dt = \frac{\alpha^k}{1 - e^{-\left(\frac{T}{\alpha}\right)^\lambda}} \Gamma\left(\frac{k}{\lambda} + 1\right) \qquad k = 1, 2, 3,$$

3.3 Estimation par la méthode de maximum de vraisemblance

On considère un n-échantillon $(t_1, t_2, ..., t_n)$ généré de la loi de Weibull tronquée à droite et soit m une constante fixé $m \in \{1, 2, ..., n\}$. Les données sont supposées censurées de type II c'est-à-dire seulement les observations $\{t_1, t_2, ..., t_m\}$ sont observés. La fonction de maximum de vraisemblance est donc :

$$L(t|\alpha, \lambda, T) = \frac{n!}{(n-m)!} \prod_{i=1}^m g(t_i, \alpha, \lambda, T)(1 - F(t_m; \alpha, \lambda, T))^{n-m}$$

$$= \frac{n!}{(n-m)!} \prod_{i=1}^m \frac{(\frac{\lambda}{\alpha})(\frac{t_i}{\alpha})^{\lambda-1} e^{-(\frac{t_i}{\alpha})^\lambda}}{1 - e^{-(\frac{T}{\alpha})^\lambda}} \left(\frac{e^{-(\frac{t_m}{\alpha})^\lambda} - e^{-(\frac{T}{\alpha})^\lambda}}{1 - e^{-(\frac{T}{\alpha})^\lambda}}\right)$$

$$L(\underline{t}|\alpha, \lambda, T) = \frac{n!}{(n-m)!} \lambda^m \alpha^{-m\lambda} \prod_{i=1}^m t_i^{\lambda-1} e^{-\sum_{i=1}^m (\frac{t_i}{\alpha})^\lambda} (e^{-(\frac{t_m}{\alpha})^\lambda} - e^{-(\frac{T}{\alpha})^\lambda})^{(n-m)} (1 - e^{-(\frac{T}{\alpha})^\lambda})^{-n}.$$

On pose les notations suivantes :

$$l(\underline{t}|\alpha, \lambda, T) = \ln L(\underline{t}|\alpha, \lambda, T), \qquad S = \sum_{i=1}^{m} \left(\frac{t_i}{\alpha}\right)^{\lambda}, \quad P = \prod_{i=1}^{m} t_i$$

et

$$\Phi = (e^{-(\frac{t_m}{\alpha})^{\lambda}} - e^{-(\frac{T}{\alpha})^{\lambda}}) \quad , \Psi = (1 - e^{-(\frac{T}{\alpha})^{\lambda}})$$

Donc, on obtient

$$L(\underline{t}|\alpha, \lambda, T) = \frac{n!}{(n-m)!} \lambda^m \alpha^{-m\lambda} P^{\lambda-1} e^{-S} \Phi^{n-m} \Psi^{-n}.$$

On prend le logarithme de la fonction de vraisemblance :

$$l(\underline{t}|\alpha, \lambda, T) = ln(n!) - ln((n-m)!) + m\ ln\lambda - m\ \lambda\ ln\alpha + (\lambda-1) \sum_{i=1}^{m}\ lnt_i - S + (n-m)\ ln\Phi - n\ ln\Psi.$$

Les estimateurs de maximum de vraisemblance des trois paramètres α, λ et T sont les solutions de l'équation

$$l(\underline{t}|\alpha, \lambda, T) = 0$$

Donc, l'équations de vraisemblance sont comme suit

$$\begin{aligned}
\frac{\partial l}{\partial \alpha} &= -\frac{m\lambda}{\alpha} - \frac{\partial S}{\partial \alpha} + (n-m)\frac{\Phi_1}{\Phi} - n\frac{\Psi_1}{\Psi} = 0 \\
\frac{\partial l}{\partial \lambda} &= \frac{m}{\lambda} + \sum_{i=1}^{m} \ln t_i - m \ln \alpha - \frac{\partial S}{\partial \lambda} + (n-m)\frac{\Phi_2}{\Phi} - n\frac{\Psi_2}{\Psi} = 0 \\
\frac{\partial l}{\partial T} &= (n-m)\frac{\Phi_3}{\Phi} - n\frac{\Psi_3}{\Psi} = 0
\end{aligned}$$

Où

$\Phi_1 = \dfrac{\partial \Phi}{\partial \alpha}$; $\Phi_2 = \dfrac{\partial \Phi}{\partial \lambda}$; $\Phi_3 = \dfrac{\partial \Phi}{\partial T}$; $\Psi_1 = \dfrac{\partial \Psi}{\partial \alpha}$; $\Psi_2 = \dfrac{\partial \Psi}{\partial \lambda}$ and $\Psi_3 = \dfrac{\partial \Psi}{\partial T}$.

Après quelques manipulations algébriques, on obtient le système d'équations non linéaire suivant :

$$\begin{cases}
\frac{-m\lambda}{\alpha} + \frac{\lambda}{\alpha}S + (n-m)\frac{\lambda}{\alpha}\frac{[(\frac{t_m}{\alpha})^{\lambda}e^{-(\frac{t_m}{\alpha})^{\lambda}} - (\frac{T}{\alpha})^{\lambda}e^{-(\frac{T}{\alpha})^{\lambda}}]}{\Phi} + n\frac{\lambda}{\alpha}\frac{(\frac{T}{\alpha})^{\lambda}e^{-(\frac{T}{\alpha})^{\lambda}}}{\Psi} = 0 \\
\frac{m}{\lambda} + \sum_{i=1}^{m} ln(\frac{t_i}{\alpha}) - \sum_{i=1}^{m} ln(\frac{t_i}{\alpha})(\frac{t_i}{\alpha})^{\lambda} + (n-m)\frac{[ln(\frac{T}{\alpha})(\frac{T}{\alpha})^{\lambda}e^{-(\frac{T}{\alpha})^{\lambda}} - ln(\frac{t_m}{\alpha})(\frac{t_m}{\alpha})e^{-(\frac{t_m}{\alpha})^{\lambda}}]}{\Phi} - n\frac{[ln(\frac{T}{\alpha})(\frac{T}{\alpha})^{\lambda}e^{-(\frac{T}{\alpha})^{\lambda}}]}{\Psi} = 0 \\
(n-m)\frac{\lambda}{T}\frac{(\frac{T}{\alpha})^{\lambda}e^{-(\frac{T}{\alpha})^{\lambda}}}{\Phi} - n\frac{\lambda}{T}\frac{(\frac{T}{\alpha})^{\lambda}e^{-(\frac{T}{\alpha})^{\lambda}}}{\Psi} = 0
\end{cases}$$

Il n'y a pas des solutions théoriques de ce système. Donc, on utilise les méthodes numériques pour obtenir les valeurs approximées aux estimateurs de maximum de vraisemblance $\widehat{\alpha}_{ML}$, $\widehat{\lambda}_{ML}$ et T_{ML} des paramètres α, λ et T respectivement. Dans ce chapitre, on utilise le package du R (le package BB) qui a une capacité pour résoudre les systèmes d'équations non linéaires.

3.4 Estimation Bayesienne sous différentes fonctions de perte

3.4.1 Les distributions a posteriori et a priori

Dans cette section, on considère la distribution a priori de (α, λ) définie comme suit

$$\pi(\alpha, \lambda) \propto \lambda^{-a}\alpha^{-b}e^{-\frac{c}{\alpha}}, \quad \alpha, \lambda > 0, \quad a > 1, \quad b, c > 0.$$

De plus, on choisit une distribution a priori impropre de T, qui ne dépend pas de (α, λ) et donnée par $\pi(T) = \frac{1}{T}$. Donc, l'a priori de (α, λ, T) est

$$\pi(\alpha, \lambda, T) = \pi(\alpha, \lambda)\pi(T) = \frac{1}{T}\lambda^{-a}\alpha^{-b}e^{-\frac{c}{\alpha}}$$

La densité a posteriori est calculée à l'aide de la formule suivante :

$$\pi(\alpha, \lambda, T|t) = \frac{L(t|\alpha, \lambda, T)\pi(\alpha, \lambda, T)}{\int\int\int(L(t|\alpha, \lambda, T)\pi(\alpha, \lambda, T))d\alpha d\lambda dT}$$

donc,

$$\begin{aligned}
&= \frac{\lambda^m\alpha^{-m\lambda}P^{\lambda-1}e^{-S}\Phi^{n-m}\Psi^{-n}\frac{1}{T}\lambda^{-a}\alpha^{-b}e^{-\frac{c}{\alpha}}}{\int\int\int\lambda^m\alpha^{-m\lambda}P^{\lambda-1}e^{-S}\Phi^{n-m}\Psi^{-n}\frac{1}{T}\lambda^{-a}\alpha^{-b}e^{-\frac{c}{\alpha}}d\alpha d\lambda dT} \\
&= \frac{\frac{1}{T}\lambda^{m-a}\alpha^{-m\lambda-b}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}}{\int_0^\infty\int_0^\infty\int_0^\infty\frac{1}{T}\lambda^{m-a}\alpha^{-m\lambda-b}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT} \\
&= K^{-1}\frac{1}{T}\lambda^{m-a}\alpha^{-m\lambda-b}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}
\end{aligned}$$

Où

$$K = \int_0^\infty\int_0^\infty\int_0^\infty\frac{1}{T}\lambda^{m-a}\alpha^{-m\lambda-b}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT.$$

3.4.2 Fonctions de perte

On considère la fonction de perte quadratique, Linex et entropie, le tableau suivant présente les trois fonctions de perte et l'expression de l'estimateur bayesien avec le risque a posteriori correspondant, Sous la fonction de perte quadratique généralisée, on suppose

Fonctions de perte	expréssion	Estimateurs de Bayes	Risque a posterio
quadratique généralisée	$L(\theta,\delta)=\tau(\theta)(\theta-\delta)^2$	$\widehat{\delta}_{GQ}=\frac{E_\pi(\tau(\theta)\theta)}{E_\pi(\tau(\theta))}$	$E_\pi(\tau(\theta)(\theta-\widehat{\delta}_{GQ})$
Entropie	$L(\theta,\delta)=(\frac{\delta}{\theta})^p-pln(\frac{\delta}{\theta})-1$	$\widehat{\delta}_E=[E_\pi(\theta)^{-p}]^{-1/p}$	$p[E_\pi(ln(\theta)-ln(\widehat{\delta}$
Linex	$L(\theta,\delta)=e^{r(\delta-\theta)}-r(\delta-\theta)-1$	$\widehat{\delta}_L=\frac{-1}{r}ln(E_\pi(e^{-r\theta}))$	$r(\widehat{\delta}_Q-\widehat{\delta}_L)$

TAB. 3.1 – **Les fonctions de perte et les estimateurs Bayésiens (avec les risques a posteriori) correspondantes**

que $\tau(\theta)=\theta^{\beta-1}$. Les estimateurs bayésiens des paramètres sont données par les formules suivantes

$$\widehat{\alpha}_{GQ}=\frac{\int\int\int\tau(\alpha)\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT}{\int\int\int\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT}$$

$$=\frac{\int_0^\infty\int_0^\infty\int_0^\infty\frac{1}{T}\alpha^{-m\lambda-b+\beta}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT}{\int_0^\infty\int_0^\infty\int_0^\infty\frac{1}{T}\alpha^{-m\lambda-b+\beta-1}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT}$$

$$\widehat{\lambda}_{GQ}=\frac{\int\int\int\tau(\lambda)\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT}{\int\int\int\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT}$$

$$=\frac{\int_0^\infty\int_0^\infty\int_0^\infty\frac{1}{T}\alpha^{-m\lambda-b}\lambda^{m-a+\beta}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT}{\int_0^\infty\int_0^\infty\int_0^\infty\frac{1}{T}\alpha^{-m\lambda-b}\lambda^{m-a+\beta-1}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT}$$

$$\widehat{T}_{GQ}=\frac{\int\int\int\tau(T)\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT}{\int\int\int\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT}$$

$$=\frac{\int_0^\infty\int_0^\infty\int_0^\infty T^{\beta-1}\alpha^{-m\lambda-b}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT}{\int_0^\infty\int_0^\infty\int_0^\infty T^{\beta-2}\alpha^{-m\lambda-b}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT}$$

Les risques a posteriori correspondants sont :

$$PR(\widehat{\alpha}_{QG})=E_\pi(\alpha^{\beta+1})-2\widehat{\alpha}_{QG}E_\pi(\alpha^\beta)+\widehat{\alpha}_{QG}^2E_\pi(\alpha^{\beta-1}).$$

$$PR(\widehat{\lambda}_{QG})=E_\pi(\lambda^{\lambda+1})-2\widehat{\lambda}_{QG}E_\pi(\lambda^\beta)+\widehat{\lambda}_{QG}^2E_\pi(\lambda^{\beta-1}).$$

$$PR(\widehat{T}_{QG})=E_\pi(T^{\beta+1})-2\widehat{T}_{QG}E_\pi(T^\beta)+\widehat{T}_{QG}^2E_\pi(T^{\beta-1}).$$

On note que lorsque $\beta = 1$, on obtient la fonction de perte quadratique usuelle.

Sous la fonction de perte entropie, les estimateurs sont

$$\widehat{\alpha}_E = (\int\int\int \alpha^{-p}\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT)$$

$$= K^{1/p}\left[\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{T}\alpha^{-m\lambda-b-p}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT\right]^{-1/p}$$

$$\widehat{\lambda}_E = (\int\int\int \lambda^{-p}\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT)$$

$$= K^{1/p}\left[\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{T}\alpha^{-m\lambda-b}\lambda^{m-a-p}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT\right]^{-1/p}$$

$$\widehat{T}_E = (\int\int\int T^{-p}\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT)$$

$$= K^{1/p}\left[\int_0^\infty\int_0^\infty\int_0^\infty T^{-p-1}\alpha^{-m\lambda-b}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT\right]^{-1/p}$$

Les risques a posteriori sont :

$$PR(\widehat{\alpha}_E) = pE_\pi(ln(\alpha) - ln(\widehat{\alpha}_E))$$

$$PR(\widehat{\lambda}_E) = pE_\pi(ln(\lambda) - ln(\widehat{\lambda}_E))$$

$$PR(\widehat{T}_E) = pE_\pi(ln(T) - ln(\widehat{T}_E))$$

Sous la fonction de perte Linex, les estimateurs bayésiens sont donnés par

$$\widehat{\alpha}_L = -\frac{1}{r}ln(\int\int\int e^{-r\alpha}\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT)$$

$$= -\frac{1}{r}\ ln\left[K^{-1}\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{T}\alpha^{-m\lambda-b}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}-r\alpha}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT\right]$$

$$\widehat{\lambda}_L = -\frac{1}{r}ln(\int\int\int e^{-r\lambda}\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT)$$

$$= -\frac{1}{r}\ ln\left[K^{-1}\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{T}\alpha^{-m\lambda-b}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}-r\lambda}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT\right]$$

$$\widehat{T}_L = -\frac{1}{r}ln(\int\int\int e^{-rT}\pi(\alpha,\lambda,T|\underline{t})d\alpha d\lambda dT)$$

$$= -\frac{1}{r}\ ln\left[K^{-1}\int_0^\infty\int_0^\infty\int_0^\infty \frac{1}{T}\alpha^{-m\lambda-b}\lambda^{m-a}P^{\lambda-1}e^{-S-\frac{c}{\alpha}-rT}\Phi^{n-m}\Psi^{-n}d\alpha d\lambda dT\right]$$

Les risques a posteriori correspondants sont :

$$PR(\widehat{\alpha}_L) = r(\widehat{\alpha}_Q - \widehat{\alpha}_L)$$

$$PR(\widehat{\lambda}_L) = r(\widehat{\lambda}_Q - \widehat{\lambda}_L)$$

$$PR(\widehat{T}_L) = r(\widehat{T}_Q - \widehat{T}_L)$$

On peut pas calculer l'expression analytique de ces estimateurs, on utilise les procédures MCMC comme l'algorithme de Metropolis-Hastings. Dans la section suivante.

3.4.3 Etudes de Monte-Carlo

Dans cette section, on propose une étude de Monte-Carlo, on suppose que $a = 2$, $b = c = 1$, $\alpha = 1$, $\lambda = 2$ et $T = 1.5$. On utilise $N = 1000$ échantillons de la loi de Weibull tronquée à droite, on obtient les résultats suivants

Estimation de maximum de vraisemblance

Puisque les formules analytiques ne sont pas disponibles, pour obtenir les estimateurs de maximum de vraisemblance, on utilise les procédures numériques. Dans cette section, on utilise le package BB du R qui présente des performances très élevées pour les systèmes d'équations non linéaire. Dans le particulier, on a besoin de la fonction BBsolve pour ce problème, les résultats sont dans le tableau suivant

n	m	paramètres	EMV
10	7	α	0.6951(0.0929)
		λ	1.9800(0.0003)
		T	1.9452(0.1660)
30	21	α	0.9036(0.0092)
		λ	2.0003(1.07 10^{-7})
		T	1.6203(0.0144)
50	35	α	0.9364(0.0040)
		λ	2.0367(0.0013)
		T	1.5945(0.0089)
100	70	α	0.9943(3.22 10^{-5})
		λ	2.0368(0.0013)
		T	1.5149(0.0002)
200	140	α	0.9268(0.0053)
		λ	2.0957(0.0091)
		T	1.6457(0.0212)

TAB. 3.2 – **Les estimateurs de maximum de vraisemblance des paramètres avec les erreurs quadratique**

On remarque que les estimateurs de α, λ sont proches des valeurs initiales des paramètres, par contre l'estimateur de T n'est pas proche de T.

L'estimation Bayésienne

Les estimateurs Bayésiens sont obtenus en utilisant les méthodes MCMC. Le tableau 3.3 présente les estimateurs Bayésiens et les risques a posteriori entre parenthèses, sous la fonction de perte quadratique généralisée. On remarque que la valeur $\beta = 2$ donne le meilleur risque, aussi on obtient les risques a posteriori petits.

Avec la fonction de perte entropie, on obtient le tableau 3.4 où la valeur $p = -0.5$ et les cas $n = 100$ et $n = 200$ donnent les meilleurs risques a posteriori.

Le tableau 3.5 donne les estimateurs de Bayes sous la fonctio de perte Linex et leurs risques a posteriori.

On peut noter que la valeur $r = -0.5$ donne le meilleur risque a posteriori. Si on compare les trois fonctions de pertes, on note que la fonction de perte entropie donne le meilleur estimateur Bayesien de α,λ et T. ça est illustré dans le tableau 3.6

3.4.4 Comparaison avec l'estimateur de maximum de vraisemblance

Dans cette section, on propose de comparer le meilleur estimateur Bayésien obtenue au-dessus avec l'estimateur de maximum de vraisemblance. Pour cela, on propose d'utiliser les critères suivant :

Le critère de Pitman (Pitman 1937, 1982 et Jozani 2012) et IMSE (intgrated mean square error) définis comme suit :

Définition Un estimateur $\widehat{\theta_1}$ de paramètre θ domine dans le sens de critère de Pitman un autre estimateur $\widehat{\theta_2}$, si, pour tout $\theta \in \Theta$:

$$P_\theta \left[|\hat{\theta}_1 - \theta| < |\hat{\theta}_2 - \theta| \right] > 0.5$$

On considère les estimateurs $\widehat{\theta_i}$ $(i = 1, ..., N)$ obtenue avec N échantillons du modèle.

Définition Integrated mean square error $(IMSE)$ est définie comme suit

$$IMSE = \frac{\sum_{i=1}^{N} (\hat{\theta}_i - \theta)^2}{N}$$

Dans la suite, nous présentons la valeur de probabilité de Pitman qui est utilisée pour comparer les estimateurs Bayésiens sous les trois fonctions de perte et le MLE où $\beta = -2$, $p = 0.5$ et $r = -0.5$. Le tableau 3.7 est lu comme suit :

− Lorsque n n'est pas grand, les estimateurs Bayésiens $\widehat{\alpha}_B$ et $\widehat{T}_B$ de α et T sont meilleurs que les MLE $\widehat{\alpha}MLE$ et $\widehat{T}_{MLE}$. La fonction de perte quadratique généralisée donne les meilleurs valeurs. Cependant $\widehat{\lambda}_{MLE}$ est plus proche à la valeur initiale du paramètre que l'estimateur Bayésien $\widehat{\lambda}_B$.

− Lorsque n est grand, les MLE de trois paramètres sont meilleurs que les estimateurs Bayésiens.

Le tableau 3.8 présente la valeur $(IMSE)$ des estimateurs Bayésiens des trois paramètres sous les trois fonctions de perte, et les estimateurs de maximum de vraisemblance.

Selon ce critère, lorsque n est petit, les estimateurs Bayésiens $\widehat{\alpha}_B$ et $\widehat{T}_B$ donnent le petit IMSE de α et T relativement à $\widehat{\alpha}_{MLE}$ et $\widehat{T}_{MLE}$. Aussi, les valeurs données avec la fonction de perte quadratique généralisée relativement équivalent à l'entropie et Linex. Mais, l'IMSE de $\widehat{\lambda}_{MLE}$ est plus petit que l'IMSE de l'estimateur Bayésien.
Si n est grand, tous les estimateurs bayésiens sont mieux que l'MLE et la fonction de perte quadratique généralisée donne la meilleur valeur de IMSE.

3.5 Conclusion

Dans ce chapitre, nous comparons les estimateurs Bayésiens de la loi de Weibull tronquée à droite sous différentes fonctions de perte. Dans l'estimation Bayésienne, pour chaque fonction de perte, on obtient le paramètre approprié, qui optimise l'estimation Bayésienne. Aussi, notre étude de Monte-Carlo montre que la fonction de perte entropie donne le petit risque a posteriori. Les estimateurs bayésien et les estimateurs de maximum de vraisemblance des paramètres sont comparés en utilisant le critère de Pitman et IMSE (Integrated mean square error). Après, en utilisant une procédure de Monte-Carlo, on obtient que lorsque n est petit, les estimateurs bayésiens sont meilleurs pour α et T et pas pour λ. Si n est grand, l'IMSE des estimateurs de maximum de vraisemblance est plus grand que les IMSE des estimateurs Bayésiens.

Interprétation des résultats Le tableau 3.7 présente les résultats de la comparaison entre les estimateurs Bayésiens et les estimaeurs de maximum de vraisemblance à l'aide de critère de Pitman. Ce tableau est lit comme suit : Lorsque la probabilité de Pitman est superieur à 0.5 on dit que l'estimateur bayésien est le meilleur, sinon, on dit que l'estimaeur de maximum de vraisemblance est le mailleur.
Pour cela, on remarque que lorsque n est petit, les estimateurs bayésien des paramètres α et T sont meilleurs que leurs estimateurs par la méthode de maximum de vraisemblance parce que la probabilité de Pitman est supérieur à 0.5 et la fonction de perte quadratique généralisée donne les mailleurs résultats parce que la probabilité de Pitman est plus grande. Alors que pour le paramètre σ les estimateurs par la méthode de maximum de vraisemblance

sont les meilleurs, parce que la probabilité de Pitman est inférieure à 0.5.

Lorsque n est grand, les estimateurs de maximum de vraisemblance des trois paramètres sont les meilleurs parce que la probabilité de Pitman est inférieure à 0.5.

Le tableau 3.8 présente les valeurs IMSE des stimateurs Bayésiens et de maximum de vraisemblance des paramètres. On dit qu'un estimateurs est meilleur losque il'a le plus petit IMSE. Pour cela, on remarque que lorsque n est petit, les estimateurs Bayésiens de paramètres α et T ont le plus petit IMSE par rapport à leurs estimateurs par la méthode de maximum de vraisemblance, alors que mle de λ a un IMSE plus petit que son estimateur bayésien.

Lorsque n est grand, les estimateurs Bayésiens des trois paramètres ont IMSE plus petit que leurs estimateurs par la méthode de maximum de vraisemblance.

			β							
n	m	paramèteres	-2	-1.5	-1	-0.5	0.5	1	1.5	2
10	7	α	1.0902	1.0943	1.0981	1.1026	1.1111	1.1155	1.1198	1.1242
			(0.0067)	(0.0071)	(0.0076)	(0.0081)	(0.0091)	(0.0096)	(0.0102)	(0.0108)
		λ	1.1873	1.2534	1.3350	1.4307	1.6448	1.7491	1.8430	1.9225
			(0.0720)	(0.0982)	(0.1309)	(0.1696)	(0.2583)	(0.3032)	(0.3453)	(0.3835)
		T	1.1928	1.2286	1.2675	1.3085	1.3921	1.4323	1.4699	1.5043
			(0.0436)	(0.0520)	(0.0614)	(0.0714)	(0.0924)	(0.1030)	(0.1134)	(0.1234)
30	21	α	1.0784	1.0816	1.0849	1.0882	1.0950	1.0984	1.1018	1.1052
			(0.0055)	(0.0058)	(0.0061)	(0.0064)	(0.0071)	0.0074	(0.0078)	(0.0082)
		λ	1.1633	1.2244	1.3014	1.3936	1.6072	1.7146	1.8128	1.8970
			(0.0708)	(0.0974)	(0.1310)	(0.1713)	(0.2651)	(0.3127)	(0.3567)	0.3951)
		T	1.1737	1.2075	1.2446	1.2842	1.3661	1.4065	1.4444	1.4793
			(0.0429)	(0.0514)	(0.0608)	(0.0708)	(0.0918)	(0.1023)	(0.1124)	(0.1218)
50	35	α	1.0750	1.0783	1.0817	1.0852	1.0922	1.0957	1.0993	1.1029
			(0.0057)	(0.0060)	(0.0063)	(0.0067)	(0.0074)	(0.0078)	(0.0082)	(0.0085)
		λ	1.1440	1.1997	1.2713	1.3590	1.5699	1.6802	1.7835	1.8740
			(0.0685)	(0.0951)	(0.1239)	(0.1713)	(0.2724)	(0.3256)	(0.3758)	(0.4201)
		T	1.1565	1.1886	1.2245	1.2634	1.3464	1.3880	1.4278	1.4650
			(0.0425)	(0.0514)	(0.0612)	(0.0720)	(0.0951)	(0.1069)	(0.1182)	(0.1290)
100	70	α	1.2021	1.2032	1.2044	1.2055	1.2078	1.2089	1.2100	1.2112
			(0.0015)	(0.0017)	(0.0018)	(0.0020)	(0.0025)	(0.0027)	(0.0030)	(0.0033)
		λ	2.2002	2.2008	2.2015	2.2021	2.2034	2.2041	2.2047	2.2053
			(0.0002)	(0.0003)	(0.0005)	(0.0008)	(0.0019)	(0.0028)	(0.0042)	(0.0062)
		T	1.7016	1.7024	1.7032	1.7040	1.7057	1.7065	1.7073	1.7081
			(0.0005)	(0.0007)	(0.0009)	(0.0012)	(0.0021)	(0.0027)	(0.0036)	(0.0047)
200	140	α	1.2109	1.2118	1.2128	1.2138	1.2157	1.2167	1.2176	1.2186
			(0.0013)	(0.0014)	(0.0015)	(0.0017)	(0.0021)	(0.0023)	(0.0025)	(0.0028)
		λ	2.1894	2.1901	2.1908	1.1915	2.1929	2.1936	2.1941	2.1949
			(0.0002)	(0.0004)	(0.0006)	(0.0009)	(0.0020)	(0.0030)	(0.0044)	(0.0065)
		T	1.7005	1.7013	1.7021	1.7028	1.7043	1.7051	1.7059	1.7066
			(0.0005)	(0.0006)	(0.0008)	(0.0011)	(0.0019)	(0.0025)	(0.0033)	(0.0044)

TAB. 3.3 – **Les estimateurs Bayésiens des paramètres et les risques a posteriori sous la fonction de perte quadratique généralisée.**

n	m	paramèteres	p							
			-2	-1.5	-1	-0.5	0.5	1	1.5	2
10	7	α	1.1198	1.1176	1.1155	1.1133	1.1090	1.1069	1.1047	1.1026
			(0.0155)	(0.0087)	(0.0038)	(0.0009)	(0.0009)	(0.0038)	(0.0086)	(0.0154)
		λ	1.8338	1.7939	1.7491	1.6998	1.5916	1.5362	1.4825	1.4321
			(0.2151)	(0.1284)	(0.0603)	(0.0158)	(0.0681)	(0.0694)	(0.1576)	(0.2793)
		T	1.4678	1.4507	1.4323	1.4129	1.3717	1.3504	1.3292	1.3083
			(0.1052)	(0.0612)	(0.0281)	(0.0072)	(0.0075)	(0.0307)	(0.0698)	(0.1248)
30	21	α	1.1018	1.1001	1.0984	1.0967	1.0933	1.0916	1.0899	1.0882
			(0.0123)	(0.0069)	(0.0031)	(0.0007)	(0.0007)	(0.0030)	(0.0069)	(0.0123)
		λ	1.8035	1.7616	1.7146	1.6633	1.5530	1.4977	1.4449	1.3961
			(0.2284)	(0.1359)	(0.0636)	(0.0166)	(0.0707)	(0.0716)	(0.1612)	(0.2837)
		T	1.4424	1.4250	1.4065	1.3870	1.3461	1.3252	1.3045	1.2843
			(0.1077)	(0.0625)	(0.0286)	(0.0073)	(0.0076)	(0.0308)	(0.0699)	(0.1245)
50	35	α	1.0993	1.0975	1.0957	1.0940	1.0904	1.0887	1.0869	1.0852
			(0.0129)	(0.0073)	(0.0032)	(0.0008)	(0.0008)	(0.0032)	(0.0072)	(0.0128)
		λ	1.7744	1.7297	1.6802	1.6268	1.5149	1.4603	1.4091	1.3625
			(0.2434)	(0.1442)	(0.0671)	(0.0174)	(0.0728)	(0.0731)	(0.1632)	(0.2848)
		T	1.4259	1.4075	1.3880	1.3676	1.3256	1.3045	1.2838	1.2639
			(0.1143)	(0.0661)	(0.0301)	(0.0077)	(0.0079)	(0.0318)	(0.0717)	(0.1269)
100	70	α	1.2100	1.2095	1.2089	1.2083	1.2072	1.2066	1.2061	1.2055
			(0.0037)	(0.0021)	(0.0009)	(0.0002)	(0.0002)	(0.0009	(0.0021)	(0.0037)
		λ	2.2047	2.2044	2.2041	2.2037	2.2031	2.2028	2.2024	2.2021
			(0.0011)	(0.0006)	(0.0002)	(0.00007)	(0.0002)	(0.0002)	(0.0006)	(0.0011)
		T	1.7073	1.7069	1.7065	1.7061	1.7053	1.7048	4.7044	1.7040
			(0.0019)	(0.0010)	(0.0004)	(0.0001)	(0.0001)	(0.0004)	(0.0010)	(0.0019)
200	140	α	1.2176	1.2171	1.2167	1.2162	1.2152	1.2147	1.2143	1.2138
			(0.0031)	(0.0017)	(0.0007)	(0.0001)	(0.0001)	(0.0007)	(0.0017)	(0.0031)
		λ	2.1942	2.1939	1.1936	1.1932	1.1925	1.1922	1.1918	1.1915
			(0.0012)	(0.0007)	(0.0003)	(0.00007)	(0.0003)	(0.0003)	(0.0007)	(0.0012)
		T	1.7059	1.7055	1.7051	1.7047	1.7040	1.7036	1.7032	1.7028
			(0.0017)	(0.0010)	(0.0004)	(0.0001)	(0.0001)	(0.0004)	(0.0010)	(0.0017)

TAB. 3.4 – **Les estimateurs Bayésiens des paramètres et le risque a posteriori sous la fonction de perte entropie**

n	m	paramèteres	r							
			-2	-1.5	-1	-0.5	0.5	1	1.5	2
10	7	α	1.1252	1.1228	1.1203	1.1179	1.1131	1.1107	1.1083	1.1059
			(0.0195)	(0.0109)	(0.0048)	(0.0012)	(0.0012)	(0.0048)	(0.0107)	(0.0191)
		λ	1.9693	1.9301	1.8809	1.8204	1.6703	1.5892	1.5120	1.4431
			(0.4403)	(0.2714)	(0.1317)	(0.0356)	(0.0394)	(0.1599)	(0.3556)	(0.4303)
		T	1.5205	1.5016	1.4805	1.4573	1.4059	1.3788	1.3515	1.3248
			(0.1764)	(0.1038)	(0.0481)	(0.0124)	(0.0131)	(0.0535)	(0.1212)	(0.2150)
30	21	α	1.1059	1.1040	1.1021	1.1002	1.0965	1.0946	1.0928	1.0909
			(0.0150)	(0.0084)	(0.0037)	(0.0009)	(0.0009)	(0.0037)	(0.3576)	(0.6106)
		λ	1.9446	1.9035	1.8518	1.7885	1.6339	1.5524	1.4762	1.4093
			(0.4599)	(0.2833)	(0.1372)	(0.0369)	(0.0403)	(0.1622)	(0.3576)	(0.6106)
		T	1.4949	1.4758	1.4546	1.4314	1.3814	1.3537	1.3271	1.3013
			(0.1767)	(0.1039)	(0.0480)	(0.0124)	(0.0130)	(0.00527)	(001190)	(0.2104)
50	35	α	1.1036	1.1016	1.0997	1.0977	1.0938	1.0918	1.0899	1.0880
			(0.0156)	(0.0088)	(0.0039)	(0.0009)	(0.0009)	(0.0038)	(0.0087)	(0.0154)
		λ	1.9271	1.8821	1.8252	1.7580	1.5972	1.5155	1.4411	1.3770
			(0.4939)	(0.3029)	(0.1457)	(0.0389)	(0.0414)	(0.1646)	(0.3586)	(0.6063)
		T	1.4822	1.4615	1.4388	1.4141	1.3609	1.3336	1.3068	1.2811
			(0.1884)	(0.1103)	(0.0508)	(0.0130)	(0.0135)	(0.0543)	(0.1217)	(0.2136)
100	70	α	1.2117	1.2110	1.2103	1.2096	1.2082	1.2075	1.2068	1.2062
			(0.0055)	(0.0031)	(0.0013)	(0.0003)	(0.0003)	(0.0014)	(0.0031)	(0.0056)
		λ	2.2069	2.2062	2.2055	2.2048	2.2034	2.2026	2.2019	2.2012
			(0.0056)	(0.0031)	(0.0014)	(0.0003)	(0.0003)	(0.0014)	(0.0031)	(0.0055)
		T	1.7093	1.7086	1.7079	1.7072	1.7058	1.7051	1.7044	1.7037
			(0.0056)	(0.0031)	(0.0013)	(0.0003)	(0.0003)	(0.0013)	(0.0031)	(0.0055)
200	140	α	1.2190	1.2184	1.2178	1.2173	1.2161	1.2155	1.2143	1.2143
			(0.0047)	(0.0026)	(0.0011)	(0.0002)	(0.0002)	(0.0011)	(0.0026)	(0.0046)
		λ	2.1965	2.1958	2.1951	2.1643	2.1928	2.1920	2.1913	2.1905
			(0.0059)	(0.0033)	(0.0014)	(0.0003)	(0.0003)	(0.0015)	(0.0034)	(0.0060)
		T	1.7077	1.7071	1.7064	1.7058	1.7015	1.7038	1.7031	1.7025
			(0.0051)	(0.0029)	(0.0012)	(0.0003)	(0.0003)	(0.0012)	(0.0029)	(0.0051)

TAB. 3.5 – **Les estimateurs bayésiens des paramètres et les risques a posteriori sous la fonction de perte Linex**

n	m	paramètresr	generalized quadratic($\beta = -2$)	Entropy(p=-0.5)	Linex(r=-0.5)
10	7	α	1.0902	1.1133	1.1179
			(0.0067)	(0.0009)	(0.0012)
		λ	1.1873	1.6998	1.8204
			(0.0720)	(0.0158)	(0.0356)
		T	1.1928	1.4129	1.4573
			(0.0436)	(0.0072)	(0.0124)
30	21	α	1.0784	1.0969	1.1002
			(0.0055)	(0.0007)	(0.0369)
		λ	1.1633	1.6633	1.7885
			(0.0429)	(0.0073)	(0.0124)
		T	1.1737	1.3870	1.4314
			(0.0429)	(0.0073)	(0.0124)
50	35	α	1.0750	1.0940	1.0977
			(0.0057)	(0.0008)	(0.0009)
		λ	1.1440	1.6368	1.7580
			(0.0685)	(0.0174)	(0.0389)
		T	1.1565	1.3676	1.4141
			(0.0425)	(0.0077)	(0.0130)
100	70	α	1.2021	1.2083	1.2048
			(0.0015)	(0.0002)	(0.0003)
		λ	2.2002	2.2037	2.2048
			(0.0002)	(0.00007)	(0.0003)
		T	1.7016	1.7061	1.7072
			(0.0001)	(0.0001)	(0.0003)
200	140	α	1.2109	1.2162	1.2173
			(0.0013)	(0.0001)	(0.0002)
		λ	2.1894	2.1932	2.1643
			(0.0002)	(0.00007)	(0.0003)
		T	1.7005	1.7047	1.7058
			(0.0005)	(0.0001)	(0.0003)

TAB. 3.6 – **Les estimateurs Bayésiens des paramètres et les risques a posteriori sous les trois fonctions de perte**

n	m	par	QG($\beta = -2$)	Entropie(p=-0.5)	Linex(r=-0.5)
10	7	α	0.699	0.638	0.654
		λ	0.249	0.130	0.159
		T	0.654	0.585	0.609
30	21	α	0.575	0.556	0.559
		λ	0.273	0.109	0.151
		T	0.532	0.489	0.501
50	35	α	0.535	0.514	0.518
		λ	0.231	0.087	0.117
		T	0.423	0.393	0.396
100	70	α	0.152	0.146	0.146
		λ	0.130	0.136	0.161
		T	0.413	0.376	0.383
200	140	α	0.097	0.095	0.095
		λ	0.221	0.077	0.086
		T	0.278	0.250	0.257

TAB. 3.7 – **Comparaison entre les estimateurs α, λ et T à l'aide du critère de Pitman**

n	m	paramètres	mle	QG	Entropie	Linex
10	7	α	0.2695	0.0080	0.0221	0.0190
		λ	0.1184	0.4333	0.3666	0.3696
		T	0.3278	0.1034	0.1070	0.7038
30	21	α	0.1031	0.0098	0.0168	0.0154
		λ	0.0686	0.3922	0.3938	0.3905
		T	0.2579	0.1036	0.1107	0.1087
50	35	α	0.0702	0.0114	0.0167	0.157
		λ	0.0551	0.4181	0.4275	0.4239
		T	0.2437	0.1125	0.1191	0.1175
100	70	α	0.0872	0.0380	0.0460	0.0445
		λ	0.0529	0.0191	0.0438	0.0362
		T	0.2840	0.0321	0.0449	0.0421
200	140	α	0.1908	0.1624	0.1672	0.1663
		λ	0.5467	0.5287	0.5493	0.5432
		T	0.4838	0.3167	0.3263	0.3242

TAB. 3.8 – **Le IMSE des estimateurs des paramètres α, λ et T**

CHAPITRE 4

La distribution de Rayleigh

Sommaire

Introduction

Lord Rayleigh (1880) a introduit la distribution de Rayleigh dans le cadre d'un problème dans le domaine d'acoustique. Depuis lors, d'importants travaux ont été lié à cette distribution dans les différentes domaines des sciences et technologies. Cette distribution a des relations avec certaines distributions bien connues comme les distributions de Weibull, Chi-deux,...

La distribution de Rayleigh est un cas particulier de la distribution de Weibull à deux paramètres est constitue un modèle approprié pour les études d'essai de survie. Polovoko (1968), Dyer et Whisnand (1973) ont démontré l'importance de cette distribution dans l'électro des dispositifs et l'ingénierie de communication.

La fonction de densité de probabilité, la fonction de fiabilité et la fonction de taux de hasard, sont données respectivement comme suit :

$$f(x,\sigma) = \frac{x}{\sigma^2} exp(-\frac{x^2}{2\sigma^2}), x > 0, \sigma > 0$$

$$R(x,\sigma) = exp(-\frac{x^2}{2\sigma^2})$$

et

$$h(x,\sigma) = \frac{x}{\sigma^2}$$

Où $\sigma > 0$ est le paramètre.

Dans ce chapitre, on s'intéresse dans un premier temps à l'estimation de la fonction de fiabilité et de la fonction de taux de hasard sous la fonction de perte quadratique et sous une fonction de perte asymétrique (Linex) en présence d'un échantillon censuré de type II avec une loi a priori non informative puis avec une loi a priori conjuguée naturelle. Dans une second partie, nous nous intéressons au cas où les données sont progressivement censurées ; toujours pour l'estimation des mêmes quantités

4.1 Inférences avec des données censurées de type II

4.1.1 Loi a priori vague

Soit $(X_1, X_2, ..., X_m, ..., X_n)$ un échantillon de taille n censuré à X_m.

La fonction de maximum de vraisemblance est

$$L(\sigma|x) = \frac{n!}{(n-m)!}\prod_{i=1}^{m} f(x_i,\sigma)(1-F(x_m,\sigma))^{n-m}$$

$$= \frac{n!}{(n-m)!}\prod_{i=1}^{m}\frac{x_i}{\sigma^2}exp(-\frac{x_i^2}{2\sigma^2})exp(-\frac{x_m^2}{2\sigma^2})$$

$$L(x \mid \sigma) \propto \frac{1}{\sigma^{2m}}exp(-\frac{T_m}{2\sigma^2})$$

Où

$$T_m = \sum_{i=1}^{m} x_i^2 + (n-m)x_m^2$$

Il est facile de démontrer que la distribution a priori définie par Jeffrey est :

$$\pi_1(\sigma) = |I_1(\sigma)|^{\frac{1}{2}} = |-E\frac{\partial^2 lnf}{\partial\sigma^2}| \propto \frac{1}{\sigma}.$$

La densité a posteriori est

$$\pi_1(\sigma|x) = \frac{L(x|\sigma)\pi_1(\sigma)}{\int_0^\infty L(x|\sigma)\pi_1(\sigma)d\sigma} = \frac{(T_m)^m}{\Gamma(m)}\frac{1}{2^{m-1}}\sigma^{-2m-1}exp(-\frac{T_m}{2\sigma^2}).$$

Estimation du paramètre σ

L'estimateur de maximum de vraisemblance noté $\widehat{\sigma}_{MV}$ est obtenue par la solution de l'équation $\frac{\partial lnL(x,\sigma)}{\partial\sigma} = 0$, donc

$$\widehat{\sigma}_{MV} = \sqrt{\frac{T_m}{2m}}.$$

Par rapport à la fonction de perte quadratique, et une loi a priori vague, l'estimateur Bayésien de σ noté $\widehat{\sigma}_{VQ}$ est obtenu en calculant la moyenne de la densité a posteriori

$$\widehat{\sigma}_{VQ} = \int_0^\infty \sigma\pi(\sigma|x)d\sigma = \frac{(T_m)^m}{\Gamma(m)}\frac{1}{2^{m-1}}\int_0^\infty \sigma^{-2m}exp(-\frac{T_m}{2\sigma^2})d\sigma = \frac{\Gamma(m-\frac{1}{2})}{\Gamma(m)}(\frac{T_m}{2})^{\frac{1}{2}}.$$

Le risque a posteriori de σ est donné par :

$$R(\widehat{\sigma}_{VQ}) = (\frac{T_m}{2})(\frac{\Gamma(m-1)}{\Gamma(m)} - \frac{\Gamma^2(m-\frac{1}{2})}{\Gamma^2(m)}).$$

Par rapport à une fonction de perte Linex, et une loi a priori vague, l'estimateur de σ noté $\widehat{\sigma}_{VL}$ est la solution de l'équation suivante :

$$E_p[\frac{\widehat{\sigma}_{VL}}{\sigma^2}exp(a(\frac{\widehat{\sigma}^2_{VL}}{\sigma^2}))] = \frac{(T^m_m)}{\Gamma(m)}\frac{\widehat{\sigma}_{VL}}{2^{m-1}}\int_0^\infty \frac{1}{\sigma^{2m+3}}exp(-\frac{1}{2\sigma^2}(T_m - 2a\widehat{\sigma}^2_{VL}))d\sigma$$

et

$$exp(a)E_p(\frac{\widehat{\sigma}_{VL}}{\sigma^2}) = \frac{(T_m)^m}{\Gamma(m)}\frac{\widehat{\sigma}_{VL}}{2^{m-1}}\int_0^\infty \frac{1}{\sigma^{2m+3}}exp(-\frac{1}{2\sigma^2(T_m)})d\sigma$$

Après quelques manipulations algébrique, on obtient

$$\widehat{\sigma}_{VL} = [\frac{T_m}{2a}(1 - exp(-\frac{a}{m+1}))]^{\frac{1}{2}}.$$

Le risque a posteriori du paramètre σ sous la fonction de perte Linex est donné par

$$R(\widehat{\sigma}_{VL}) = a(\widehat{\sigma}_{VQ} - \widehat{\sigma}_{VL}).$$

Estimation de la fonction de fiabilité

Pour obtenir l'estimateur $S_{MV}(t)$ de la fonction de fiabilité $S(t)$, on remplace σ par $\widehat{\sigma}_{MV}$ dans l'expression de $S(t)$, donc

$$S_{MV}(t) = exp(-\frac{mt^2}{T_m}).$$

L'estimateur Bayésien de $S(t)$ par rapport à la fonction de perte quadratique et une loi a priori vague est

$$S_{VQ}(t) = \int_0^\infty S(t)\pi_1(\sigma|x)d\sigma = (\frac{T_m}{T_m + t^2})^m.$$

Le risque a posteriori de la fonction de fiabilité sous la fonction de perte quadratique est donné par

$$R(S_{VQ}(t)) = (\frac{T_m}{T_m + 2t^2})^m - (\frac{T_m}{T_m + t^2})^{2m}.$$

Pour calculer l'estimateur Bayésien de $S(t)$ par rapport à la fonction de perte Linex noté $S_{LV}(t)$, on pose le changement de variable $S(t) = exp(-\frac{t^2}{2\sigma^2}) = \gamma \Rightarrow \sigma = (-\frac{t^2}{2ln\gamma})^{\frac{1}{2}}$ on écrit la densité a posteriori de γ, on obtient

$$\pi_1(\gamma|x) = \frac{(T_m)^m}{\Gamma(m)}\frac{1}{2^{m-1}}\frac{1}{2}(-\frac{t^2}{2ln\gamma})^{\frac{1}{2}}(\frac{t^2}{2\gamma(ln\gamma)^2})(-\frac{t^2}{2ln\gamma})^{-m-\frac{1}{2}}(\gamma)^{\frac{T_m}{t^2}}$$

$$= (\frac{T_m}{t^2})^m \frac{1}{\Gamma(m)}(\gamma)^{\frac{T_m}{t^2}-1}(-ln\gamma)^{m-1}; \;\; 0 \leq \gamma \leq 1$$

donc, l'estimateur Bayésien $\widehat{\gamma}_{LV}$ de γ est

$$\gamma_{LV} = -\frac{1}{a}lnE_p(exp(-a\gamma)) = -\frac{1}{a}ln[(\frac{T_m}{t^2})^m \frac{1}{\Gamma(m)}\int_0^\infty exp(-a\gamma)(\gamma)^{\frac{T_m}{t^2}-1}(-ln\gamma)^{m-1}d\gamma]$$

$$= -\frac{1}{a}ln[\sum_{j=0}^{k}\frac{(-a)^j}{j!}(1+j\frac{t^2}{T_m})^{-m}]$$

Le dernier résultat est obtenu en utilisant un développement de $(exp(-a\gamma))$ à l'ordre k au voisinage de zéro, et on fait le changement de variable $u = (-ln\gamma)$ pour calculer l'intégrale. Le risque a posteriori de la fonction de fiabilité sous la fonction de perte Linex est donné par

$$R(S_{VL}(t)) = a(S_{VQ}(t) - S_{VL}(t)).$$

Estimation de la fonction de taux de hasard

L'estimateur de maximum de vraisemblance de la fonction de taux de hasard $h(t)$ noté $h_{MV}(t)$ est obtenu lorsqu'on remplace σ par $\widehat{\sigma}_{MV}$ dans l'expression de $h(t)$ donc

$$h_{MV}(t) = 2m\frac{t}{T_m}.$$

Par rapport à une fonction de perte quadratique et une loi a priori vague de σ l'estimateur Bayésien de $h(t)$ noté $h_{VQ}(t)$ est

$$h_{VQ}(t) = \int_0^\infty h(t)\pi_1(\sigma|x)d\sigma = \frac{(T_m)^m}{\Gamma(m)}\frac{t}{2^{m-1}}\int_0^\infty \sigma^{-2m-3}exp(-\frac{T_m}{2\sigma^2})d\sigma$$

$$h_{VQ}(t) = 2m\frac{t}{T_m}.$$

Le risque a posteriori de la fonction de taux de hasard sous la fonction de perte quadratique

$$R(h_{VQ}(t)) = \frac{4t^2\Gamma(m+2)}{\Gamma(m)}(T_m)^{-2} - (2m)^2\frac{t^2}{(T_m)^2}.$$

L'estimateur de $h(t)$ obtenu par la méthode de maximum de vraisemblance et par l'approche Bayésienne avec une loi a priori vague et une fonction de perte quadratique sont identique.

4.1.2 Loi a priori conjuguée naturelle

La loi a priori conjuguée naturelle est définie par :

$$\pi_2(\sigma) \propto \frac{1}{\sigma^{\alpha+1}} exp(-\frac{\beta}{2\sigma^2});\;\; \alpha, \beta > 0$$

La densité a posteriori est donc

$$\pi_2(\sigma|x) = \frac{(T_m+\beta)^{m+\frac{\alpha}{2}}}{2^{m+\frac{\alpha}{2}-1}\Gamma(m+\frac{\alpha}{2})} \frac{1}{\sigma^{2m+\alpha+1}} exp(-\frac{1}{2\sigma^2}(T_m+\beta))$$

Estimation Bayésienne du paramètre σ

Toujours par rapport à une fonction de perte quadratique et une loi a priori conjuguée naturelle de σ, l'estimateur de σ noté $\widehat{\sigma}_{CQ}$ est obtenu en calculant sa moyenne par rapport à la densité a posteriori

$$\widehat{\sigma}_{CQ} = \int_0^{\infty} \sigma\pi_2(\sigma|x)dx = \sqrt{\frac{T_m+\beta}{2}} \frac{\Gamma(m+\frac{\alpha}{2}-\frac{1}{2})}{\Gamma(m+\frac{\alpha}{2})}.$$

Le risque a posteriori de paramètre σ est donné par la formule suivante :

$$R(\widehat{\sigma}_{CQ}) = \frac{T_m+\beta}{2}(\frac{\Gamma(m+\frac{\alpha}{2}-1)}{\Gamma(m+\frac{\alpha}{2})} - \frac{\Gamma^2(m-\frac{\alpha}{2}-\frac{1}{2})}{\Gamma^2(m-\frac{\alpha}{2})}).$$

L'estimateur Bayésien de σ par rapport à la fonction de perte Linex et une loi a priori conjuguée naturelle de σ, est la solution de l'équation

$$E_p[\frac{\widehat{\sigma}_{CL}}{\sigma^2} exp(a(\frac{\widehat{\sigma}^2_{CL}}{\sigma^2}))] = 2^{m+\frac{\alpha}{2}} \widehat{\sigma}_{CL}[\frac{\Gamma(m+\frac{\alpha}{2}+1)}{(T_m+\beta-2a\widehat{\sigma}^2_{CL})^{m+\frac{\alpha}{2}+1}}]$$

$$exp(a)E_p(\frac{\widehat{\sigma}_{CL}}{\sigma^2}) = \widehat{\sigma}_{CL} e^a \frac{\Gamma(m+\frac{\alpha}{2}+1)}{(T_m+\beta)^{m+\frac{\alpha}{2}+1}}$$

$$\widehat{\sigma}_{CL} = [\frac{T_m+\beta}{2a}(1-exp(-\frac{a}{m+\frac{\alpha}{2}+1}))]^{\frac{1}{2}}.$$

Le risque a posteriori de paramètre σ sous la fonction de perte Linex est donné par :

$$R(\widehat{\sigma}_{CL}) = a(\widehat{\sigma}_{CQ} - \widehat{\sigma}_{CL}).$$

Estimation de la fonction de fiabilité

L'estimateur Bayésien de $S(t)$ par rapport à la fonction de perte quadratique et une loi a priori conjuguée naturelle est $S_{CQ}(t)$

$$S_{CQ}(t) = \int_0^\infty S(t)\pi_2(\sigma|x)d\sigma = (\frac{T_m+\beta}{T_m+\beta+t^2})^{rm+\frac{\alpha}{2}}$$

$$S_{CQ}(t) = (\frac{T_m+\beta}{T_m+\beta+t^2})^{m+\frac{\alpha}{2}}.$$

Le risque a posteriori de la fonction de fiabilité sous la fonction de perte quadratique est donné par :

$$R(S_{CQ}(t)) = (\frac{T_m+\beta}{T_m+\beta+2t^2})^{m+\frac{\alpha}{2}} - (\frac{T_m+\beta}{T_m+\beta+t^2})^{2m+\alpha}.$$

Avec une loi a priori conjuguée naturelle, l'estimateur Bayésien sous la fonction de perte Linex est noté $S_{CL}(t)$, pour le calculer, on utilise le changement de variable suivant $S(t) = exp(-\frac{t^2}{2\sigma^2}) = \gamma \Rightarrow \sigma = (-\frac{t^2}{2ln\gamma})^{\frac{1}{2}}$, donc

$$\gamma_{CQ} = \frac{-1}{a} lnE_p(exp(-a\gamma))$$

$$= \frac{-1}{a} ln[\frac{(T_m+\beta)^{m+\frac{\alpha}{2}}}{\Gamma(m+\frac{\alpha}{2})} \frac{1}{(t^2)^{m+\frac{\alpha}{2}}} \int_0^\infty exp(-a\gamma)(\gamma)^{\frac{T_m+\beta}{t^2}-1}(-ln\gamma)^{m+\frac{\alpha}{2}-1}d\gamma]$$

$$= -\frac{1}{a} ln[\sum_0^k \frac{(-a)^j}{j!}(1+\frac{jt^2}{T_m+\beta})^{(-m-\frac{\alpha}{2})}].$$

Le risque a posteriori de la fonction de fiabilité sous la fonction de perte Linex est donné par

$$R(S_{CL}(t)) = a(S_{CQ}(t) - S_{CL}(t))$$

Estimation de la fonction de taux de hasard

Par rapport à la fonction de perte quadratique, l'estimateur Bayésien de $h(t)$ est

$$h_{CQ}(t) = \frac{2(m+\frac{\alpha}{2})}{T_m+\beta} t$$

Le risque a posteriori de la fonction de taux de hasard est donné par

$$R(h_{CQ}(t)) = (\frac{t}{2})^2 \frac{\Gamma(m+\frac{\alpha}{2}+2)}{\Gamma(m+\frac{\alpha}{2})}(T_m+\beta)^{-2} - h_{CQ}(t)^2.$$

Sous la fonction de perte Linex, l'estimateur Bayésien de la fonction de taux de hasard est donné par la formule suivante

$$h_{CL}(t) = a\frac{2t}{T_m + \beta}[1 - exp(-\frac{a}{m + \frac{\alpha}{2} + 1})]^{-1}.$$

Le risque a posteriori de la fonction de taux de hasard sous la fonction de perte Linex est

$$R(h_{CL}(t)) = a(h_{CQ}(t) - h_{CL}(t)).$$

4.2 Simulation

1$-$ On donne des valeurs à α et β, on génère la loi a priori conjuguée naturelle et on déduit la valeur de σ.

2$-$ On génère $N = 10000$ échantillons de taille n de la loi de Rayleigh de paramètre σ, on prend trois valeurs de n ($n = 20, 30, 50$) et deux valeurs de m,($m = 10, 15$).

3$-$ On calcule les différents estimateurs de σ, $S(t)$ et $h(t)$, avec la méthode de maximum de vraisemblance (notée MV), et avec l'approche Bayésienne sous une fonction de perte quadratique (noté MQ), et finalement sous une fonction de perte Linex (noté LI) pour trois valeurs de a, ($a = -0.5, -1, 1$).

4$-$ Pour chaque estimateur, on calcule l'erreur moyenne quadratique (MSE), ϕ peut être σ, $S(t)$ et $h(t)$, $\widehat{\phi}$ leurs estimateurs respectivement

$$MSE = \frac{1}{10000}\sum_{i=0}^{10000}(\phi - \widehat{\phi})^2.$$

(n,m)	paramètres	MV	quadratique	linex(a=-0.5)	linex(a=-1)	linex(a=1)
(20,10)	σ	1.7494(0.0005)	1.8187(0.0021)	1.6872(0.0072)	1.7084(0.0041)	1.6272(0.0210)
	$S(t)$	0.9058(0.0048)	0.9063(0.0049)	0.8841(0.0023)	0.8402(0.0017)	0.6825(0.0235)
	$h(t)$	0.2651(0.0006)	0.2651(0.0006)	0.2850(0.0021)	0.2791(0.0016)	0.3067(0.0046)
(30,10)	σ	1.7513(0.0044)	1.8106(0.0023)	1.6990(0.0069)	1.7081(0.0041)	1.6310(0.0199)
	$S(t)$	0.9059(0.0048)	0.9063(0.0049)	0.8841(0.0023)	0.8403(0.0002)	0.6827(0.0253)
	$h(t)$	0.2652(0.0007)	0.2652(0.0007)	0.2859(0.0021)	0.2788(0.0016)	0.3048(0.0043)
(50,10)	σ	1.7478(0.0006)	1.8170(0.0019)	1.6856(0.0075)	1.7090(0.0040)	1.6332(0.0193)
	$S(t)$	0.9059(0.0006)	0.9065(0.0049)	0.8845(0.0023)	0.8398(1.46e-05)	0.6824(0.0235)
	$h(t)$	0.2656(0.0007)	0.2656(0.0007)	0.2856(0.0021)	0.2780(0.0015)	0.3048(0.0043)
(20,15)	σ	1.7557(0.0002)	1.8012(0.0008)	1.7157(0.0032)	1.7285(0.0019)	1.6760(0.0092)
	$S(t)$	0.9086(0.0052)	0.9089(0.0053)	0.8867(0.0025)	0.8105(10-05)	0.6838(0.0231)
	$h(t)$	0.2565(0.0003)	0.2565(0.0003)	0.2684(0.0008)	0.2645(0.0006)	0.2812(0.0018)
(30,15)	σ	1.7625(9e-05)	1.8082(0.0012)	1.7200(0.0027)	1.7238(0.0018)	1.6710(0.0102)
	$S(t)$	0.9033(0.0053)	0.9096(0.0054)	0.8871(0.0026)	0.8426(4e-05)	0.6837(0.0231)
	$h(t)$	0.2543(0.0002)	0.2543(0.0002)	0.2671(0.0008)	0.2642(0.0006)	0.2830(0.0019)
(50,15)	σ	1.7551(0.0002)	1.8006(0.0007)	1.7119(0.0036)	1.7299(0.0018)	1.67421(0.0096)
	$S(t)$	0.9085(0.0003)	0.9088(0.0052)	0.8864(0.0025)	0.8427(4e-05)	0.6838(0.0237)
	$h(t)$	0.2566(0.0003)	0.2566(0.0003)	0.2693(0.0009)	0.2451(0.0006)	0.2818(0.0018)

TAB. 4.1: **Les estimateurs Bayésiens de la distribution de Rayleigh avec une loi a priori vague**

(n,m)	paramètres	quadratique	linex(a=-0.5)	linex(a=-1)	linex(a=1)
(20,10)	σ	1.7846(0.0788)	1.6615(0.0805)	1.6831(0.0767)	1.6123(0.0891)
	S(t)	0.9033(0.0054)	0.9011(0.0051)	0.8903(0.0037)	0.9640(0.0180)
	h(t)	0.2740(0.0102)	0.2936(0.0133)	0.2858(0.0118)	0.3119(0.0172)
(30,15)	σ	1.7867(0.0773)	1.6634(0.0787)	1.6824(0.0761)	1.6103(0.0900)
	S(t)	0.9037(0.0055)	0.9014(0.0051)	0.8903(0.0037)	0.9636(0.0179)
	h(t)	0.2729(0.0098)	0.2924(0.0128)	0.2859(0.0119)	0.3127(0.0174)
(50,10)	σ	1.7879(0.0781)	1.6646(0.0791)	1.6823(0.0793)	1.6120(0.0889)
	S(t)	0.9037(0.0055)	0.9015(0.0052)	0.8900(0.0037)	0.9640(0.0179)
	h(t)	0.2729(0.01000)	0.2924(0.0131)	0.2868(0.0124)	0.3117(0.0168)

(20,15)	σ	1.7856(0.0516)	1.7012(0.0518)	1.7074(0.0524)	1.6655(0.0576)
	S(t)	0.9077(0.0056)	0.9053(0.0053)	0.8932(0.0037)	0.9691(0.0186)
	H(t)	0.2599(0.0055)	0.2725(0.0067)	0.2707(0.0064)	0.2846(0.0028)
(30,15)	σ	1.7856(0.0516)	1.7012(0.0518)	1.7074(0.0524)	1.6655(0.0576)
	S(t)	0.9070(0.0055)	0.9046(0.0052)	0.8932(0.0037)	0.9689(0.0186)
	h(t)	0.2620(0.0055)	0.2747(0.0068)	0.2706(0.0065)	0.2852(0.0081)
(50,15)	σ	1.7782(0.0528)	1.6942(0.0540)	1.7113(0.0517)	1.6603(0.0574)
	S(t)	0.9069(0.0056)	0.9045(0.0052)	0.8936(0.0038)	0.9686(0.0185)
	h(t)	0.2423(0.0057)	0.2750(0.0069)	0.2695(0.0064)	0.2860(0.0081)

TAB. 4.2: **L'estimateurs Bayésiens de la loi de Rayleigh avec une loi conjuguée naturelle**

4.3 Application

On propose ici d'appliquer les méthodes des données réelles qui sont présentés dans Lawless. Les données ont été soulevées dans l'essai sur l'endurance des paliers à rainures profondes et sont discutées par Leiblein et Zalen. Il sont le nombre de tours à l'échec de chacun de $n = 23$ paliers dans le test de vie. Raqab et Madi ont indiqué que la distribution de Rayleigh à un seul paramètre est acceptable pour ces données. La taille d'échantillon censuré est $m = 13$, les observations sont :

0.1788,0.2892,0.3300,0.4212,0.4560,0.4848,0.5184,0.5196,0.6780,0.8412,0.9312,1.2792

L'estimateur de maximum de vraisemblance σ est égal à $\widehat{\sigma}_{MV} = 0.9175$

Pour des différentes valeurs de t, $t = 0.25, 0.5, 0.75, 1$, on obtient les estimateurs de maximum de vraisemblance de la fonction de fiabilité qui sont : $S_{MV}(t) = 0.9635, 0.8620, 0.7160, 0.5521$

L'estimateurs de maximum de vraisemblance de la fonction de taux de hasard pour différentes valeurs de t sont : $\widehat{h}_{MV}(t) = 0.2969, 0.5939, 0.8908, 1.1878$

Les estimateurs Bayésiens de paramètre σ, de la fonction de fiabilité et de la fonction de taux de panne avec une loi a priori vague puis avec une loi a priori conjuguée naturelle sont données dans les tableaux suivants : **Resultats** :

t	paramètres	quadratique(PR)	Linex[a=-1](PR)	Linex[a=-0.5](PR)	Linex[a=1](PR)
0.25	σ	0.9451(0.0187)	0.9001(0.0449)	0.8921(0.0264)	0.8686(0.0765)
	S(t)	0.9636($9.87*10^{-5}$)	0.9466(0.0169)	0.9605(0.0015)	0.9577(0.0818)
	h(t)	0.2969(0.0067)	0.3085(0.0115)	0.3141(0.0085)	0.3313(0.0343)
0.5	S(t)	0.8627(0.0012)	0.8514(0.0112)	0.8609(0.0008)	0.8514(0.0112)
	h(t)	0.5939(0.0271)	0.6170(0.0231)	0.6282(0.0171)	0.6170(0.0231)
0.75	S(t)	0.7190(0.0042)	0.7144(0.0045)	0.7189($2.14*10^{-5}$)	0.7378(0.0188)
	h(t)	0.8908(0.0610)	0.9255(0.0346)	0.9423(0.0257)	0.9940(0.1031)
1	S(t)	0.5594(0.0078)	0.5603(0.0008)	0.5609(0.0007)	0.5628(0.0033)
	h(t)	1.1878(0.1085)	1.2340(0.0462)	1.2564(0.0343)	1.3275(0.1375)

TAB. 4.3 – **L'estimateurs Bayésiens de la loi de Rayleigh avec une loi a priori vague**

t	paramètres	quadratique(RP)	Linex[a=-1](PR)	Linex[a=-0.5](PR)	Linex[a=1](PR)
0.25	σ	0.9473(0.0181)	0.6075(0.3397)	0.6051(0.1710)	0.5980(0.3493)
	S(t)	0.9638(0.0654)	0.9065(0.0581)	0.9176(0.0231)	0.9852(0.0214)
	h(t)	0.2949(0.0064)	0.3059(0.0110)	0.3113(0.0082)	0.3277(0.0328)
0.5	S(t)	0.8635(0.1843)	0.7124(0.1511)	0.7173(0.0731)	0.7124(0.1511)
	h(t)	0.5898(0.0257)	0.6117(0.0220)	0.6226(0.0164)	0.6118(0.0220)
0.75	S(t)	0.7205(0.2371)	0.4772(0.2432)	0.4776(0.1214)	0.4783(0.2421)
	h(t)	0.8847(0.0579)	0.9178(0.0331)	0.9339(0.0246)	0.9833(0.0986)
1	S(t)	0.5614(0.2015)	0.2731(0.2882)	0.2723(0.1445)	0.2697(0.2916)
	h(t)	1.1796(0.1030)	1.2237(0.0441)	1.2452(0.0328)	1.3111(0.1315)

TAB. 4.4 – **L'estimateurs Bayésiens de la loi de Rayleigh avec une loi a priori conjuguée naturelle.**

On conclut que l'stimateur d paramètre σ a un risque minimum lorsqu'on utilise la fonction de perte qudratique, et le meilleur estimateur de la fonction de fiabilité st obtenu avec la fonction de perte Linex ($a = -1$).

4.4 Inférences avec des données progressivement censurées

4.4.1 Estimation par la méthode de maximum de vraisemblance

Dans le cas des données progressivement censurées, la fonction de maximum de vraisemblance est calculée à l'aide de la formule suivante

$$L(x \mid \sigma) = A \prod_{i=0}^{m} f(x_i, \sigma)(1 - F(x_i, \sigma))^{R_i}$$

avec

$$A = n(n - R_1)(n - 2 - R_1 - R_2)...(n - \sum_{i=1}^{m-1}(R_i - 1))$$

$$L(x \mid \sigma) = A \prod_{i=1}^{m} \frac{x_i}{\sigma^2} exp(-\frac{x_i^2}{2\sigma^2}) exp(-\frac{x_i^2}{2\sigma^2})^{R_i}$$

$$= A \frac{1}{\sigma^{2m}} \prod_{i=1}^{m} exp(-\frac{1}{2\sigma^2} T_{m,R_i})$$

avec $T_{m,R_i} = \sum_{i=1}^{m} x_i^2 (1 - R_i)$

On prend le log de la vraisemblance, on obtient

$$l(x \mid \sigma) = ln(L(x \mid \sigma)) = lnA - 2mln\sigma + \sum_{i=1}^{m} lnx_i - \frac{T_{m,R_i}}{2\sigma^2}$$

On dérive par rapport à σ, on obtient

$$\frac{\partial l(x \mid \sigma)}{\partial \sigma} = \frac{-2m}{\sigma} + \frac{T_{m,R_i}}{\sigma^3}$$

En résolvant l'équation $\frac{\partial l(x|\sigma)}{\partial \sigma} = 0$, On obtient $\widehat{\sigma}_{MV,1} = \sqrt{\frac{T_{m,R_i}}{2m}}$ qui est l'estimateur de maximum de vraisemblance de paramètre σ avec des données progressivement censurées.

Pour trouver l'estimateur de maximum de vraisemblance de la fonction de fiabilité, on remplace la valeur de σ par la valeur de $\widehat{\sigma}_{MV,1}$ dans l'expression de $S(x, \sigma)$.

$S_{MV,1}(x) = exp(-\frac{mx^2}{T_{m,R_i}})$ est l'estimateur de maximum de vraisemblance de la fonction de fiabilité, avec des données progressivement censurées.

Aussi, pour trouver l'estimaeur de maximum de vraisemblance de la fonction de taux de panne, on remplace la valeur de σ par $\widehat{\sigma}_{MV,1}$ dans l'expression de $h(x,\sigma)$.

$h_{MV,1}(x) = \frac{2mx}{T_{m,R_i}}$ est l'estimateur de maximum de vraisemblance de la fonction de fiabilité, avec des données progressivement censurées.

4.4.2 Estimation Bayésienne

On prend une loi a priori conjuguée naturelle

$$\pi(\sigma) \propto \frac{1}{\sigma^{\alpha+1}} exp(-\frac{\beta}{2\sigma^2}), \alpha, \beta > 0.$$

Donc, la densité a posteriori est calculée à l'aide de la formule suivante

$$\pi_1(\sigma \mid x) = \frac{L(x \mid \sigma)\pi(\sigma)}{\int_0^\infty L(x \mid \sigma)\pi(\sigma)d\sigma}$$

La densité a posteriori est donnée par

$$\pi_1(\sigma \mid x) = \frac{(T_{m,R_i} + \beta)^{m+\frac{\alpha}{2}}}{2^{m+\frac{\alpha}{2}-1}\Gamma(m+\frac{\alpha}{2})} \sigma^{-2m-\alpha-1} exp(-\frac{1}{2\sigma^2}(T_{m,R_i} + \beta)).$$

Fonction de perte quadratique

L'estimateur du paramètre σ sous la fonction de perte quadratique est la moyenne a posteriori $\widehat{\sigma}_{Q,1} = E_{\pi_1}(\sigma \mid x)$, donc

$$\begin{aligned}\widehat{\sigma}_{Q,1} = E_{\pi_1}(\sigma \mid x) &= \int_0^\infty \sigma\pi_1(\sigma \mid x)d\sigma \\ &= \sqrt{\frac{T_{m,R_i} + \beta}{2}} \frac{\Gamma(m+\frac{\alpha}{2}-\frac{1}{2})}{\Gamma(m+\frac{\alpha}{2})}\end{aligned}$$

$\widehat{\sigma}_{Q,1}$ est l'estimateur Bayesien du paramètre σ sous la fonction de perte quadratique, avec des données progressivement censurée.

Le risque a posteriori du paramètre σ sous la fonction de perte quadratique est calculé comme suit :

$$R(\widehat{\sigma}_{Q,1}) = var_{\pi_1}(\sigma) = E_{\pi_1}(\sigma^2 \mid x) - E_{\pi_1}(\sigma \mid x)^2$$

Donc

$$R(\widehat{\sigma}_{Q,1}) = \frac{T_{m,R_i} + \beta}{2} (\frac{\Gamma(m + \frac{\alpha}{2} - 1)}{\Gamma(m + \frac{\alpha}{2})} - \frac{\Gamma^2(m + \frac{\alpha}{2} - \frac{1}{2})}{\Gamma^2(m + \frac{\alpha}{2})})$$

L'estimateur de la fonction de fiabilité est calculé comme suit

$$S_{Q,1}(x) = \int_0^\infty S(x)\pi(\sigma \mid x)d\sigma$$

$$S_{Q,1}(x) = (\frac{T_{m,R_i} + \beta}{T_{m,R_i} + \beta + x^2})^{m+\frac{\alpha}{2}}.$$

$S_{Q,1}(x)$, est l'estimateur Bayesien de la fonction de fiabilité sous la fonction de perte quadratique.

$$R(S_{Q,1}(x)) = var_\pi(S(x)) = E_{\pi_1}(S(x)^2 \mid x) - E_\pi(S(x) \mid x)^2$$

$$= \frac{(T_{m,R_i} + \beta)}{T_{m,R_i} + \beta + 2x^2}^{m+\frac{\alpha}{2}} - \frac{(T_{m,R_i} + \beta + x^2)^{2m+\alpha}}{T_{m,R_i} + \beta}$$

$R(S_{Q,1}(x))$ est le risque a posteriori de la fonction de fiabilité sous la fonction de perte quadratique avec des données progressivement censurées.

L'estimateur de la fonction de taux de panne est calculé comme suit

$$h_{Q,1}(x) = \int_0^\infty h(x)\pi_1(\sigma \mid x)d\sigma$$

$$= \frac{2x\Gamma(m + \frac{\alpha}{2} + 1)}{(T_{m,R_i} + \beta)\Gamma(m + \frac{\alpha}{2})}$$

$h_{Q,1}(x)$, est l'estimateur Bayesien de la fonction de taux de panne sous la fonction de perte quadratique avec des données progressivement censurées.

$$R(h_{Q,1}(x)) = var_{\pi_1}(\sigma \mid x) = E_{\pi_1}(h(x)^2 \mid x) - E_{\pi_1}(h(x) \mid x)^2$$

$$= \frac{4x^2\Gamma(m + \frac{\alpha}{2}) + 2}{\Gamma(m + \frac{\alpha}{2})}(T_{m,R_i} + \beta)^{-2} - \frac{4x^2\Gamma^2(m + \frac{\alpha}{2} + 1)}{\Gamma^2(m + \frac{\alpha}{2})}(T_{m,R_i} + \beta)^{-2}$$

$R(h_{Q,1}(x))$ est le risque a posteriori de la fonction de taux de panne sous la fonction de perte quadratique avec des données progressivement censurées.

Fonction de perte Linex

Considérons la fonction convexe donnée par

$$L(\Delta) \propto e^{a\Delta} - a\Delta - 1, a \neq 0$$

Le signe a et sa grandeur absolue représentent respectivement la direction et le degré d'asymétrie. Pour $a \to 0$, on retrouve la fonction de perte quadratique.

Considérons la fonction de perte $L(\Delta_1)$, où $\Delta_1 = (\frac{\widehat{\theta}}{\theta})^2 - 1$; cette fonction de perte a été utilisée par plusieurs auteurs dont Zellner (2006) et (2009). On minimise l'espérance a posteriori $E_\pi(L(\Delta_1))$

$$E_\pi(L(\Delta_1)) = E_\pi(exp((\frac{\widehat{\theta}}{\theta})^2 - 1) - a((\frac{\widehat{\theta}}{\theta})^2 - 1) - 1).$$

$$= exp(-a)E_\pi(exp(\frac{\widehat{\theta}}{\theta})^2) - a(\frac{\widehat{\theta}}{\theta})^2 - 1.$$

En dérivant $E_\pi(L(\Delta_1))$ par rapport à $\widehat{\theta}$ et en égalant à zéro, on obtient l'estimateur Bayesien $\widehat{\theta}_{L,1}$ de θ sous la fonction de perte $L(\Delta_1)$

$$\frac{\partial L_\pi(L(\Delta_1))}{\partial \theta} = 2aexp(-a)E_\pi(\frac{\widehat{\theta}}{\theta^2})) - 2aE_\pi(\frac{\widehat{\theta}}{\theta^2}) = 0$$

On obtient alors $\widehat{\theta}_{L,1}$ comme solution de l'équation

$$E_\pi(\frac{1}{\theta^2}exp(a(\frac{\widehat{\theta}_{L,1}}{\theta^2}))) = exp(a)E_\pi(\frac{1}{\theta^2}).$$

L'estimateur Bayesien de paramètre σ sous la fonction de perte Linex $\sigma_{L,1}$ est la solution de l'équation suivante :

$$E_{\pi_1}[\frac{\widehat{\sigma}_{L,1}}{\sigma^2}exp(a(\frac{\widehat{\sigma}^2_{L,1}}{\sigma^2}))] = 2^{m+\frac{\alpha}{2}}\widehat{\sigma}_{L,1}[\frac{\Gamma(m+\frac{\alpha}{2}+1)}{(T_{m,R_i}+\beta-2a\widehat{\sigma}_{L,1})^{m+\frac{\alpha}{2}+1}}]$$

$$exp(a)E_{\pi_1}(\frac{\widehat{\sigma}_{L,1}}{\sigma^2}) = \widehat{\sigma}_{L,1}e^a\frac{\Gamma(m+\frac{\alpha}{2}+1)}{(T_{m,R_i}+\beta)^{m+\frac{\alpha}{2}+1}}$$

$$\widehat{\sigma}_{L,1} = [\frac{T_{m,R_i}+\beta}{2a}(1-exp(-\frac{a}{m+\frac{\alpha}{2}+1}))]^{\frac{1}{2}}$$

Le risque a posteriori du paramètre σ sous la fonction de perte Linex est donné par

$$R(\widehat{\sigma}_{L,1}) = a(\widehat{\sigma}_{Q,1} - \widehat{\sigma}_{L,1})$$

Avec une loi a priori conjuguée naturelle, l'estimateur Bayesien par rapport à une fonction de perte Linex est noté $S_{L,1}(t)$, pour le calculer, on fait un changement de variable suivant : $S(t) = exp(-\frac{t^2}{2\sigma^2}) = \gamma \Rightarrow \sigma = (-\frac{t^2}{2ln\gamma})^{\frac{1}{2}}$; on réécrit la densité a posteriori donnée en fonction de γ

$$\gamma_{L,1} = \frac{-1}{a}lnE_\pi(exp(-a\gamma))$$

$$= \frac{-1}{a} ln[\frac{(T_{m,R_i}+\beta)^{m+\frac{\alpha}{2}}}{\Gamma(m+\frac{\alpha}{2})} \frac{1}{(t^2)^{m+\frac{\alpha}{2}}} \int_0^{\infty} exp(-a\gamma)(\gamma)^{\frac{T_{m+R_i+\beta}}{t^2}}(-ln\gamma)^{m+\frac{\alpha}{2}-1} d\gamma]$$

$$= -\frac{1}{a} ln[\sum_{j=0}^{k} \frac{(-a)^j}{j!}(1+\frac{jt^2}{T_{m,R_i}+\beta})^{(-m+\frac{\alpha}{2})}]$$

Ce dernier résultat est obtenu en utilisant un développement de $exp(-a\gamma)$ à l'ordre k dans un voisinage de zéro et en faisant le changement de variable $u = (-ln\gamma)$ pour le calcul de l'intégrale.

Le risque a posteriori de la fonction de fiabilité sous la fonction de perte Linex est donné par la formule suivante

$$R(S_{L,1}(t)) = a(S_{Q,1}(t) - S_{L,1}(t)).$$

Sous la fonction de perte Linex, l'estimateur Bayésien de la fonction de taux de panne est donné par la formule suivante

$$h_{L,1}(t) = a\frac{2t}{T_{m,R_i+\beta}}[1 - exp(-\frac{a}{m+\frac{\alpha}{2}} + 1)]^{-1}$$

Le risque a posteriori de la fonction de taux de panne sous la fonction de perte Linex est calculé à l'aide de la formule suivante

$$R(h_{L,1}(t)) = a(h_{Q,1}(t) - h_{L,1}(t)).$$

Conclusion et perspectives

L'évaluation de la fiabilité ou de la durée de survie suivant le contexte dans lequel on se place est indisponsable pour concevoir des systèmes plus performants. Les méthodes d'estimation des paramètres du modèle ajusté au problème étudié sont nombreuses et peuvent donner des résultats fortement biaisés.

Dans cette thèse, nous nous sommes intéréssés à l'estimation des paramètres, de la fonction de fiabilité et de la fonction taux de panne, à l'aide d'une approche Bayesienne. L'approche Bayesienne repose sur le choix d'une loi a priori qui résume l'information dont on dispose en aval et sur le choix d'une fonction de perte. La fonction de perte la plus utilisée est la fonction de perte quadratique, cependant, ces dernières années les fonctions de perte asymétriques de type LINEX ont été utilisées dans les problèmes de l'inférence statistique Bayesienne. De même, nous consid érons différents plans d'expérience. En effet, un plan d'expé rience complet est difficilement réalisable pour des raisons évidentes de coût ; c'est pourquoi, nous avons considéré des plans censurés de type II et des plans progressivement censurés.

Pour le modèle de Weibull tronqué et le modèle de Rayleigh, nous nous sommes intéréssés à l'estimation des caracté ristiques cités ci-dessus, avec une étude comparative à partir des erreurs quadratiques moyennes (pour l'approche du maximum de vraisemblance) et à partir du risque à postériori pour l'approche Bayesienne.

En conclusion, les estimateurs obtenus à partir d'une approche Bayesienne, sous des fonctions de perte asymétriques sont plus performants.

Une extension possible de ce travail est de considérer pour les modè les étudiés des plans d'expériences accélérés.

Annexe

Algorithme d'estimation des paramètres de la loi de Weibull tronquée avec la méthode de maximum de vraisemblance

$N = 1000$

$H1 = numeric(N)$

$H2 = numeric(N)$

$H3 = numeric(N)$

$nsims = 100; m = 70; alpha = 2; lambda = 1; T = 1.5$

$c1 = exp(-(T/alpha)$lambda)

t=numeric(N)

$for(kin1 : N)\{$

$u = runif(nsims, 0, 1)$

$vec = alpha\star$((-log(1-(1-c1)$\star$u)))$\star$(1\lambda)

vec=sort(vec)

t=vec

$\neq$ LV représente la fonction de log de vraisemblance

LV=function(par){log(factorial(nsims)-log(factorial(nsims-m))+(m$*$log(lambda))$*$(m$*$lambda$*$log(alp

1)$*$(sum(log(t))))-((alpha(-lambda))$*$sum(tlambda))+((nsims-m)$*$log(c1))-(nsims$*$log(c2))}

$\neq$ Q représente les paramètres de la loi

$\neq$ dd représente les fonctions de la dérivée de la log de vraisemblance g← function(Q){

b1=$((t[m] \setminus Q[1] \curlywedge Q[2]) \star (exp(-(t[m] \setminus Q[1]) \curlywedge Q[2])$

b2=$((Q[3] \setminus Q[1] \curlywedge Q[2]) \star (exp(-(Q[3] \setminus Q[1]) \curlywedge Q[2])$

b3=$((Q[3] \setminus Q[1] \curlywedge Q[2]) \star (exp(-(t[m] \setminus Q[1]) \curlywedge Q[2])$

b4=$1 - (exp(-(Q[3] \setminus Q[1]) \curlywedge Q[2])$

$dd \leftarrow rep(NA, lenght(Q))$

$dd[1] \leftarrow ((-m \star Q[2] \setminus Q[1]) + (((Q[1] \curlywedge (-Q[2])) \star (sum(t \curlywedge Q[2])))) \setminus Q[1]) + ((nsims - m) \star (((Q[2] \star b1) + (Q[1] + b3))) + ((nsims \star Q[2] \star b2) \setminus (Q[1] \star b4))))$

$dd[2] \leftarrow (m \setminus Q[2]) - (m \star log(Q[1])) + (sum(log(t))) - ((Q[1] \curlywedge (-Q[2])) \star sum(t \curlywedge Q[2])) + (Q[2] \star sum((t \curlywedge Q[2] \star log(t))) \star ((nsims - m) \star ((log(t[m] \setminus Q[1]) \star b1) - (log(Q[3] \setminus Q[1]) \star b2))) \setminus b3) - ((nsims \star Q[2] \star b2) \setminus (Q[3] \star b4))$

$dd[3] \leftarrow (((nsims - m) \star Q[2] \star b2) \setminus (Q[3] \star b3)) - ((nsims \star Q[2] \star b2) \setminus (Q[3] \star b3)) - ((nsims \star Q[2] \star b2) \setminus (Q[3] \star b4))$

}

$Q0 \leftarrow c(2, 1, 1.5)$

$library(BB)$

$g(Q0)$

$BBsolve(par = Q0, fn = g)$

Les estimateurs de Bayes des paramètres de la loi de Weibull tronqueé (L'algorithme de Metropolis-Hastings)

$N = 1000$

$H1 = numeric(N)$;$H2 = numeric(N)$;$H3 = numeric(N)$

$nsims = 100; m = 70; alpha = 1; lambda = 2; T = 1.5$

$c2 = exp(-(T \setminus alpha) \curlywedge lambda)$

$t = numeric(m)$

for (k in 1 :N){

u=runif(nsims,0,1)

$vec = alpha \star ((-log(1-(1-c2) \star u))) \curlywedge (1 \setminus lambda))$

vec=sort(vec) ;t=vec

$a = 2; b = 1; c = 1; aa = 4;$

$f \leftarrow function(x, y, z)(1 \setminus z) \star (y \curlywedge (m-a)) \star (x \curlywedge (-b-(m \star y))) \star (prod(t) \curlywedge (y-1)) \star exp((-c \setminus x$

$-(x \curlywedge (-y) \star sum(t \curlywedge y))) \star ((exp(-t[m] \setminus x) \curlywedge y) - exp(-(z \setminus x) \curlywedge y)) \curlywedge (nsims - m)) \star$

$((1 - exp(-(z \setminus x) \curlywedge y)) \curlywedge (-nsims))\}$

$q \leftarrow function(x, y, z)(x \curlywedge (aa \setminus 2 - 1) \star exp(-x \setminus 2)) \star exp(-x \setminus 2)) \star (y \curlywedge (aa \setminus 2 - 1) \star exp(-y \setminus 2))$

$M = 500; ind = N \star 3$

$X = matrix(rep(0, ind), ncol = 3, nrow = M);$

y

$X[1,1] = 1; X[1,2] = 1; X[1,3] = 1;$

for (i in 2 :M)

}

$y = rchisq(1,2) + c(1, 0.8, 1.5)$

$val = (f(y[1], y[2], y[3]) * q(X[i-1,1], X[i-1,2], X[i-1,3])) \setminus (q(y[1], y[2], y[3]) * f(X[i-$

$1,1], X[i-1,2], X[i-1,3]));$

$alpha0 = min(1, val); u = runif(1);$

$if(u < alpha0)$

$\{X[i,1] \leftarrow y[1]; X[i,2] \leftarrow y[2]; X[i,3] \leftarrow y[3]\}$

$else$

$\{X[i,1] \leftarrow X[i-1,1]; X[i,2] \leftarrow X[i-1,2]; X[i,3] \leftarrow X[i-1,3]\}$

}

$H1[k] = mean(X[,1])$

$H2[k] = mean(X[,2])$

$H3[k] = mean(X[,3])$

}

Comparaison à l'aide des critères de Pitman, IMSE

N=1000

alphagq=numeric(N) ; lambdagq=numeric(N) ; Tgq=numeric(N) ;

alphae=numeric(N) ; lambdae=numeric(N) ; Te=numeric(N) ;

alphal=numeric(N) ; lambdal=numeric(N) ; Tl=numeric(N) ;

num1=0 ; num2=0,num3=0 ; num4=0 ; num5=0 ; num6=0 ; num7=0 ; num8=0 ; num9=0 ;

nsims=100 ; m=70 ; alpha=1 ; lambda=2 ; T=1.5 ; beta=-2 ; p=-0.5 ; r=-0.5 ; a=1 ; b=1 ;

c=2 ; aa=4

$\neq$ Génération de la loi

c2=exp(-(T\\$alpha) \curlywedge lambda$)

$t = numeric(m)$

$for(kin1 : N)\{$

$u = runif(nsims, 0, 1)\{$

$vec = alpha * (((-log(1 - (1 - c2) * u))) \curlywedge (1 \setminus lambda)$

$vec = sort(vec)$

$t = vec$

$\neq$ Méthode de Metropolis-Hastings

f←function (x,y,z){ $(1\backslash z)*(y \curlywedge (m-a))*(x \curlywedge (-b-(m*y))*(prod(t) \curlywedge (y-1))*exp((-c\backslash x)*$

$(x \curlywedge (-y)*sum(t \curlywedge y)))*exp(-(z\backslash x) \curlywedge y)) \curlywedge (nsims-m))*((1-exp(-(z\backslash x) \curlywedge y)) \curlywedge (-nsims))\}$

$q \leftarrow$ function (x,y,z){ (x$\curlywedge((aa \setminus 2) - 1) * exp(-x \setminus 2)) * (y \curlywedge ((aa \setminus 2) - 1) * exp(-y \setminus 2)) *$

$(z \curlywedge ((aa \setminus 2) - 1) * exp(-z \setminus 2))\}$

$M = 500; ind = N * 3$

$X = matrix(rep(0, ind), ncol = 3, nrow = M);$

$y = numeric(3)$

y

$X[1, 1] = 1; X[1, 2] = 1; X[1, 3] = 1;$

$for(iin2 : M)$

$\{$

$y = rchisq(1,2) + c(1,2,1.5);$

$val = (f(y[1],y[2],y[3]) * q(X[i-1,1],X[i-1,2],X[i-1,3])) \setminus (q(y[1],y[2],y[3]) * f(X[i-1,1],X[i-1,2],X[i-1,3]))$

$alpha0 = min(1,val); u = runif(1)$

$if(u < alpha0)$

$\{X[i,1] \leftarrow y[1]; X[i,2] \leftarrow y[2]; X[i,3] \leftarrow Y[3]\}$

$else$

$\{X[i,1] \leftarrow y[1]; X[i,2] \leftarrow y[2]; X[i,3] \leftarrow y[3]\}$

$\neq$ Les estimateurs de Bayes des paramètres sous la fonction de parte quadratique généralisée

alphagq[k]=(mean(X[,1])$\curlywedge beta) \setminus (mean(X[,1] \curlywedge (beta-1)))$

$lambdagq[k] = (mean(X[,2]) \curlywedge beta) \setminus (mean(X[,2] \curlywedge (beta-1)))$

$Tgq[k] = (mean(X[,3]) \curlywedge beta) \setminus (mean(X[,3] \curlywedge (beta-1)))$

$\neq$ Les estimateurs de Bayes des paramètres sous la fonction de perte entropie

alphae[k]=(mean(X[,1]$\curlywedge(-p))) \curlywedge (-1 \setminus p)$

$lambdae[k] = (mean(X[,2] \curlywedge (-p))) \curlywedge (-1 \setminus p)$

$Te[k] = (mean(X[,3] \curlywedge (-p))) \curlywedge (-1 \setminus p)$

$\neq$ Les estimateurs de Bayes des paramètres sous la fonction de perte Linex

alphal[k]=(-1\$a) * (log(mean(exp(-a * X[,1]))))$

$lambdal[k] = (-1 \setminus a) * (log(mean(exp(-a * X[,2]))))$

$Tl[k] = (-1 \setminus a) * (log(mean(exp(-a * X[,3]))))$

$\neq$ Les estimateurs de maximum de vraisemblance

g$\leftarrow function(Q)\{$

$b1 = ((t[m] \setminus Q[1]) \curlywedge Q[2]) * (exp(-(t[m] \setminus Q[1]) \curlywedge Q[2]))$

$b2 = ((Q[3] \setminus Q[1]) \curlywedge Q[2]) * (exp(-(Q[3] \setminus Q[1]) \curlywedge Q[2]))$

$phi = (exp(-(t[m] \setminus Q[1]) \curlywedge Q[2])) - (exp(-(Q[3] \setminus Q[1]) \curlywedge Q[2]))$

$psi = 1 - exp(-(Q[3] \setminus Q[1]) \curlywedge Q[2]))$

$dd \leftarrow rep(NA, lengh(Q))$

$dd[1] \leftarrow ((-m * Q[2]) \setminus Q[1]) + (Q[2] * (-Q[1]-1)) * (sum(t \curlywedge Q[2]))) + (((nsims - m) *$

$(Q[2] \setminus Q[1]) * (b1 - b2)) \setminus phi) - (nsims * (Q[2] \setminus Q[1]) * b2) \setminus psi)$

$dd[2] \leftarrow (m \setminus Q[2]) - (m * log(Q[1])) + (sum(log(t))) + (Q[1] \curlywedge (-Q[2]) * (log(Q[1])) * (sum(t \curlywedge Q[2]))) - ((Q[1] \curlywedge (-Q[2])) * (sum(log(t) * (t \curlywedge Q[2])))) + ((nsims - m) * (((log(Q[3] \setminus Q[1])) * b2) - ((log(t[m] \setminus Q[1])) * b1))) \setminus phi) - (nsims * (log(Q[3] \setminus Q[1])) * b2) \setminus psi)$

$dd[3] \leftarrow ((nsims - m) * (Q[2] \setminus Q[3]) * b2 \setminus phi) - ((nsims * (Q[2] \setminus Q[3])) * b2) \setminus psi)$

$\}$

$libary(BB)$

$v = BBsolve(par = c(1, 2, 1.2), fn = g)\S par$

$alphamle[k] = v[1]$

$lambdamle[k] = v[2]$

$Tmle[k] = v[3]$

$if(abs(alphagq[k] - alpha) < abs(alphamle[k] - alpha))\{num1 = num1 + 1\}$

$if(abs(lambdagq[k] - lambda) < abs(lambdamle[k] - lambda))\{num2 = num2 + 1\}$

$if(abs(Tgq[k] - T) < abs(Tmle[k] - T))\{num3 = num3 + 1\}$

$if(abs(alphae[k] - alpha) < abs(alphamle[k] - alpha))\{num4 = num4 + 1\}$

$if(abs(lambdae[k] - lambda) < abs(lambdamle[k] - lambda))\{num5 = num5 + 1\}$

$if(abs(Te[k] - T) < abs(Tmle[k] - T))\{num6 = num6 + 1\}$

$if(abs(alphal[k] - alpha) < abs(alphamle[k] - alpha))\{num7 = num7 + 1\}$

$if(abs(lambdal[k] - lambda) < abs(lambdamle[k] - lambda))\{num8 = num8 + 1\}$

$if(abs(Tl[k] - T) < abs(Tmle[k] - T))\{num9 = num9 + 1\}$

$\}$

$\neq$ Comparaison entre les estimateurs mle et les estimateurs de Bayes sous la fonction de perte quadratique généralisée à l'aide de critère de Pitman

P1=num1$\setminus N$

$P2 = num2 \setminus N$

$P3 = num3 \setminus N$

$P1$

$P2$

$P3$

$\neq$ Comparaison entre les estimateurs mle et les estimateurs de Bayes sous la fonction de perte entropie à l'aide de critère de Pitman

P4=num4\N

$P5 = num5 \setminus N$

$P6 = num6 \setminus N$

$P4$

$P5$

$P6$

$\neq$ Comparaison entre les estimateurs mle et les estimateurs de Bayes sous la fonction de perte Linex à l'aide de critère de Pitman

P7=num7\N

$P8 = num8 \setminus N$

$P9 = num9 \setminus N$

$P7$

$P8$

$P9$

$\neq$ Integrated mean square error

$\neq$ Sous la fonction de perte quadratique généralisée

IMSE1=sum((alphagq-alpha)⅄2) $\setminus N$

$IMSE2 = sum((lambdagq - lambda) \curlywedge 2) \setminus N$

$IMSE3 = sum((Tgq - T) \curlywedge 2) \setminus N$

$IMSE1$

$IMSE2$

$IMSE3$

$\neq$ Sous la fonction de perte entropie

IMSE4=sum((alphae-alpha)⅄2) $\setminus N$

$IMSE5 = sum((lambdae - lambda) \curlywedge 2) \setminus N$

$IMSE6 = sum((Te - T) \curlywedge 2) \setminus N$

$IMSE4$

$IMSE5$

$IMSE6$

$\neq$ Sous la fonction de perte Linex

IMSE7=sum((alphal-alpha)$\curlywedge$2) $\setminus N$

$IMSE8 = sum((lambdal - lambda) \curlywedge 2) \setminus N$

$IMSE9 = sum((Tl - T) \curlywedge 2) \setminus N$

$IMSE7$

$IMSE8$

$IMSE9$

$\neq$ Les estimateurs de maximum de vraisemblance

IMSE10=sum((alphamle-alpha)$\curlywedge$2) $\setminus N$

$IMSE11 = sum((lambdamle - lambda) \curlywedge 2) \setminus N$

$IMSE12 = sum((Tmle - T) \curlywedge 2) \setminus N$

$IMSE10$

$IMSE11$

$IMSE12$

Bibliographie

Mitra,D. (2012). Left truncated and reight censored Weibull data and likelihood inference with an illustration, computation statistics and Data analysis 56, 4011-4025

Balakrishnan N,,Approximate MLE of the scale parameter of the Rayleigh distribution with censoring, *IEEE transaction on realiability* **38** (1989), 355–357

Bakker A, ROUX J, Mostert P,A generalization of the coumpound Rayleigh distribution using a Bayesian methods on cases survival times, *Communications in statistics, Theory and methods* **29** (2000), 1419–1433

Bosq, D et Lecoutre J.P,Théorie de l'estimation , *Economirica, Paris* (1987),

Bakker A, ROUX J, Mostert P,A generalization of the coumpound Rayleigh distribution using a Bayesian methods on cases survival times, *Communications in statistics, Theory and methods* **29** (2000), 1419–1433

Chansoo K, Keun L,Estimation of the scale parameter of the Rayleigh distribution under general progressive censoring, *Journal of the Korean statistical society* **28** (2009),239 –246

CHAO M.T AND LO S.H,Some representation of non parametric maximum likelihood estimators with truncated data, *Ann. Statist* **16**,661 –668
Dallas R, Wingo (1988). Parametric point estimator for a doubly truncated Weibull distribution, Microelectron. Reliab, 28, 613-617.

Dusit, C. and Cohen, A.C (1984). Estimation in the singly truncated weibull distribution with an unknown truncation point Communications in Statistics - Theory and Methods Volume 13, Issue 7

DYER DD,WHISENAND CW,Best linear unbiased of the parameter of the Rayleigh distribution-Part I : Small sample theory for censored order statistics, *IEEE transaction on reliability* **22** (1973),27 –34

FADDEN M.C ET RUUD P.A, Estimation by simulation , *review of economics and statistics* **76** (1994),591 –608

FERNANDEZ AJ,Bayesian inference from type II censored data , *statistics and probability letters* **22** (2006),393–399

Fuller, W.A (1982). Closeness of estimators. Encyclopedia of statistical sciences, vol 2, Wiley.

GAUSS CF,Least squares method for combinations of observations, *Mallet-Bacheleir-Paris* (1810),–

HILL C, AND ALL.,Analyse statistique des données de survie, *Collection statistique en biologie et médecine* (2000),–

HOWGOARD P, Analysis of muultivariate survival data, *Statistics for biology and health* **Springer** (2000),–

HOWLADER HA, HOSSAIN A,On Bayesian estimation and prediction from realibility based on type II censored data, *communications in statistics-Theory and methods* **24** (1995),2249–2259

JAMES RC,On a robustness property of the Rayleigh and buighan tests of uniformity, *statistics ans probability letters* (2016),55–59

JEAN DAVID FERMANIAN,Modèles des durées , France

Kantar, Y.M. and Usta, I. (2015). Analysis of the upper truncated Weibull distribution for wind speed, Energy conversion and Management 96 ,81-88

KUNDU D, RAQAB MZ,Generalized Rayleigh distribution : differents methods of estimation, *Computational statistics ans data analysis* **49** (2005),167–200

KUNDU D, RAQAB MZ,Discriminating between the log-normal and generalized distribution , *Statistics* **46** (2007),505–515

KLEIN P AND ALL, Survival analysis , *Statistics for biology and health*, (2005)

LAI T.L AND YING Z, Estimating a distribution function with truncated and censored data, *Ann Stat*, (1991), 417–442

LALITHA S, HISHRA A,Modufied maximum likelihood estimation for rayleigh distribution , *Communications in statistics-Theory and Methods***25** (1996),398–401

LAWLESS JF,Statistics and methods for lifetime data, *Weiley Newyork* (),–

Maltamo M, Eerikäinen K,Pitkänen J,Hyyppä J,Vehmas M.(2004). M.Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. Rem Sens Environ , 90 :319-30.

MOHAMED Z, RAQAB MZ, MADI MT,Inference for the generalized rayleigh distribution based on progressively censored data, *Journal of statistical planning abd inference***141** (2011),3313–3322

MOUSA MA, AL SAGHEER SA,Statistical inference for the Rayleigh model based on progressively type II censored data, *Statistics***40** (2006),149–157

Mital MM, Dahiya RC (1989). Estimating of truncated Weibull distribution, Commun Stat-Theory Meth, 18, 2027-2042.

Palahi M, Pukkala T, Blasco E (2007). Comparison of beta, Johnson's SB, Weibull and truncated Weibull functions for modeling the diameter distribution of forest stands in Catalonia (north-east of Spain). Eur J Forest Res , 126(4) : 563-71.

Philipe saint Pierre (2015). Introduction à l'analyse des données de survie. Université de Pièrre et Marie Curie.

Pitman, E. J. G. (1937). The closest estimates of statistical parameters, Proc. Cambridge Phil. Soc., 33, 212-222.

Povoko A.M, Fundamentals of reliability theory, *NewYork Academic Press* (1968),–

Mousa MA, Al Sagheer SA,Statistical inference for the Rayleigh model based on progressively type II censored data, *Statistics***40** (2006),149–157

Raqab MZ, Madi MT,Bayesian prediction of the total time on test using doubly censored Rayleigh data, *journal of statistical computation and simulation***72** (2002),781–789

Raqab MZ,Madi MT,Bayesian analysis for the exponentiated Rayleigh distribution , *International journal of statistics* (2009),269–288

Rayleigh J.W.S,On the resultats of a large number of vibrations of the some pitch and arbitrary phase, *Philisophical Magazine, 5th Series* (),73–78

Robert .C, L'analyse statistique Bayesienne, *Economica* (1992).

Sajid A,Mixture of the inverse Rayleigh distribtion properties and estimation in a Bayesian frame work, *Applied mathematical modelling***39** (2015),515–530

Seki T, Yokoyama S. (1993). Simple and robust estimation of the Weibull parameters. Microelectron Reliab ;33(1) :45-52.

Shalaby O.A (1993)The Bayes risk for the doubly truncated Weibull distribution. Microelectron. Reliab, 33, 2189-2192.

Shalaby O.A et El-Youcef (1993). Bayesian analysis of the parameters of doubly truncated Weibull distribution, Microelectron. Reliab, 33, 1199-1211.

SANKU D, TANUJIT D,Statistical inderence for the Rayleigh distribution under progressively type II censoring with binomial removal, *Applied mathematical modeling* **38** (2014),974–982

TAHANI AA,Estimation of the unknown parameters for the compound Rayleigh distribution based on progressive first failure censored sampling, *Open journal of statistics* (2011),161–171

Varadhan, R and Gilbert, P. D (2009). BB : An R Package for Solving a Large System of Nonlinear Equations and for Optimizing a High-Dimensional Nonlinear Objective Function. Journal of Statistical Software, Volume 32, Issue 4.

VARIAN HR,A Bayesain approach to real estate assement , *Amsterdam, North Holland* (1975),195–208

WU SJ, CHEN DH, CHEN ST,Bayesian inference for the Rayleigh distribution under progressive censored sample, *applied stochastic models in business and industry* **22** (2006),269–279

ZEBLEIN J,ZELEN M,Statistics investigation of the fatigue life of the deep groove ball bearing , *journal of the national buteau of standards* **57** (1965),273–316

Zhang,T. and Xie,M. (2011) On the upper truncated Weibull distribution and its reliability implicarions, Reliability Engineering and system safety, 96, 194-200.

Zutter BR, Oderwald RG, Murphy PA, Farrar Jr. RM. (1086). Characterizing diameter distributions with modified data types and forms of the Weibull distribution.

Forest Science, 32 :37-48.