

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

FONCTIONS DE PERTE EN ACTUARIAT

par

GEANINA CRACIUN

Département de mathématiques et de statistique

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures

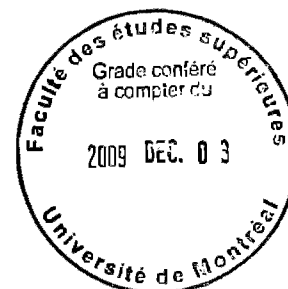
en vue de l'obtention du grade de

Maître ès sciences (M.Sc.)

en Mathématiques

Orientation actuariat

septembre 2009



Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

FONCTIONS DE PERTE EN ACTUARIAT

présenté par

GEANINA CRACIUN

a été évalué par un jury composé des personnes suivantes :

CHARLES DUGAS

(président-rapporteur)

LOUIS DORAY

(directeur de recherche)

RICHARD DUNCAN

(membre du jury)

Mémoire accepté le:

La date d'acceptation

SOMMAIRE

Ce mémoire a pour but de trouver des estimateurs pour la loi de Pareto sous différentes fonctions de perte. Pour ce faire, nous présentons les fonctions de perte LINEX, quadratique et normale reflétée modifiée, ainsi que les procédures pour calculer les estimateurs sous ces trois fonctions. Nous prenons comme estimateurs de départ les estimateurs du maximum de vraisemblance, ainsi que les estimateurs des moments. Finalement, nous terminons en comparant les estimateurs obtenus à l'aide d'échantillons de jeux de données simulées.

MOTS CLÉS EN FRANÇAIS

Non-admissibilité ; distribution de Pareto ; estimateur de Bayes ; fonction de perte LINEX et quadratique ; fonction de perte asymétrique ; méthode d'approximation Lindley pour les estimateurs de Bayes ; fonction de perte normale reflétée modifiée ; perte espérée.

SUMMARY

This report aims at finding estimators for the Pareto distribution under various loss functions. To do this, we present the LINEX, quadratic and the modified reflected normal loss functions, as well as procedures to calculate the estimates under these three loss functions. We take as starting values the maximum likelihood estimates and the moments estimators. Finally, we compare the estimators obtained with samples of simulated data.

KEYWORDS IN ENGLISH

Inadmissibility ; Pareto distribution ; Bayes estimation ; LINEX and quadratic loss functions ; Asymmetric loss functions ; Lindley's Bayes approximation form ; modified reflected normal loss function ; expected loss.

TABLE DES MATIÈRES

Sommaire	iii
Mots clés en français	iv
Summary	v
Keywords in english	vi
Liste des figures	ix
Liste des tableaux	x
Remerciements	1
Chapitre 1. Introduction	2
Chapitre 2. Théorie de la décision et fonctions de perte	6
2.1. Cadre décisionnel bayésien	6
2.1.1. Concepts et définitions de base.....	6
2.1.2. Utilité et perte.....	8
2.2. Deux critères d'optimalité : minimaxité et admissibilité	12
2.2.1. Estimateurs aléatoires.....	13
2.2.2. Minimaxité	14
2.2.3. Existence des règles minimax et stratégie maximin	14
2.2.4. Admissibilité.....	16
2.3. Analyse de la fonction de perte	19
2.3.1. Fonction de perte quadratique	19
2.3.2. Fonction de perte erreur absolue	21

2.3.3.	Perte 0-1	22
2.3.4.	Perte intrinsèque.....	23
2.3.5.	Fonctions de perte asymétriques	24
2.3.6.	Fonction de perte Linex.....	25
2.3.7.	Fonction de perte normale reflétée	26
2.3.8.	Fonction de perte normale reflétée modifiée	27
2.3.9.	Alternatives	28
2.4.	Conclusion.....	30
Chapitre 3. Estimateurs de Bayes et approximatifs appliqués au modèle de Pareto		31
3.1.	Introduction	31
3.2.	Estimation avec le critère du maximum de vraisemblance.....	32
3.3.	Estimateurs de Bayes avec β connu	35
3.4.	Estimation de Bayes avec α et β inconnus.....	40
3.4.1.	Fonction de perte quadratique	41
3.4.2.	Fonction de perte Linex.....	44
3.4.3.	Fonction de perte normale reflétée modifiée	45
Chapitre 4. Comparaisons numériques		47
4.1.	Estimateurs avec α et β inconnus	47
4.2.	Estimateurs avec β connu	55
Chapitre 5. Conclusion		63
Bibliographie		65

LISTE DES FIGURES

2.1	La fonction de perte normale reflétée modifiée (MRN)	28
2.2	Les fonctions de perte normale (MRN, SRN, Q)	29
4.1	Estimateurs pour la fonction de fiabilité.	52
4.2	Estimateurs pour le taux d'échec à $t = 1$	52
4.3	Estimateurs pour la fonction de fiabilité, $R(1)$	57
4.4	Estimateurs pour le taux d'échec à $t=1$	58

LISTE DES TABLEAUX

4.1	Estimateurs pour α	49
4.2	Estimateurs pour β	49
4.3	Estimateurs pour la fonction de fiabilité $R(1)$	50
4.4	Estimateurs pour le taux d'échec H	51
4.5	Estimateur de Bayes de α sous la fonction de perte Linex (variation d'après le paramètre a)	53
4.6	Estimateur de Bayes de β sous la fonction de perte Linex (variation d'après le paramètre a)	54
4.7	Estimateur de Bayes de la fonction de fiabilité sous la fonction de perte Linex (variation d'après le paramètre a)	54
4.8	Estimateur de Bayes du taux d'échec sous la fonction de perte Linex (variation d'après le paramètre a)	55
4.9	Estimateurs pour α	56
4.10	Estimateurs pour la fonction de fiabilité $R(1)$	57
4.11	Estimateurs de Bayes du taux d'échec $H(1)$	58
4.12	Estimateur de Bayes de α sous la fonction de perte normale réflétée.	59
4.13	Estimateur de Bayes de α sous la fonction de perte Linex (variation d'après le paramètre a)	60
4.14	Estimateur de Bayes de la fonction de fiabilité sous la fonction de perte Linex (variation d'après le paramètre a)	61

4.15	Estimateur de Bayes du taux d'échec sous la fonction de perte Linex (variation d'après le paramètre a)	62
------	--	----

REMERCIEMENTS

J'aimerais d'abord remercier mon directeur de recherche, Monsieur Louis Doray, professeur au Département de mathématiques et de statistique de l'Université de Montréal, qui a proposé et supervisé ce projet et qui m'a soutenue tout au long de ce travail. Mes remerciements vont aussi à tous les professeurs et au personnel du Département de mathématiques et de statistique pour leur travail extraordinaire, qui m'ont permis de compléter mes études. Également je voudrais remercier mon mari qui m'a encouragée à finaliser ce mémoire. Finalement, je dédie ce présent mémoire à ma mère.

Chapitre 1

INTRODUCTION

Il y a un siècle, le rôle de la statistique était de compter des choses. Les économistes ont créé un mot pour désigner la science de cette partie de l'économie politique, les dénombrements, et l'appellent statistique. Cette vocation de compter les choses dont on pouvait se gausser au siècle dernier, a revêtu quelque noblesse, depuis qu'on a reconnu combien elle est essentielle à la bonne marche des systèmes économiques, et donc des sociétés humaines. Mais en même temps, le terme statistique a progressivement désigné un domaine d'activité intellectuelle infiniment plus vaste. La première phase d'un traité élémentaire de statistique, publié il y a quelques années aux États-Unis, est celle-ci : « La statistique est un ensemble de méthodes pour prendre des décisions raisonnables en présence d'incertitudes. » (Abraham Wald (1950)). Par quel chemin a-t-elle pu passer des dénombrements à la théorie des décisions ? On convient généralement de considérer que la première pierre de la statistique mathématique fut posée par Karl Pearson, avec son célèbre mémoire de 1900 dans le *Philosophical Magazine*. En réalité, le problème de l'induction statistique avait été posé depuis longtemps, dans le mémoire du Révérend Thomas Bayes, publié après sa mort, en 1764, où il proposait une solution générale à ce problème. De grands statisticiens tels Karl Pearson, Sir Roland Aylmer Fisher et Jerzy Neyman ont échafaudé les principaux chapitres de la statistique mathématique ; dans les dernières décennies, elle est devenue indispensable dans la plupart des domaines d'activités scientifiques, mais pour réaliser l'unité de la statistique mathématique, il a fallu attendre l'oeuvre d'Abraham Wald, la théorie des fonctions de décision statistique. Le statisticien

doit, au vu des observations, prendre une décision. Un problème statistique comporte les éléments formels suivants : l'ensemble des observations, X , est l'ensemble de tous les points x , qui constituent tous les résultats d'observations possibles a priori, pour le phénomène étudié. Les diverses lois de probabilité prises en considération, F , forment un ensemble Ω . Ce sont des distributions de probabilité sur X . Les décisions possibles d forment l'ensemble \mathcal{D} . Dans certains cas, on peut être amené à prendre en considération le tirage au sort entre plusieurs décisions, avec des probabilités choisies au mieux : cela signifie qu'on prend en considération l'ensemble \mathcal{D}^* , des distributions de probabilité sur \mathcal{D} . Le problème posé est de choisir une décision d chaque fois qu'on disposera d'observations x , c'est-à-dire de choisir une manière de faire correspondre un d à chaque x ; une telle application de X dans \mathcal{D} , constitue ce que Wald appelle une fonction de décision δ ; notons Δ , l'ensemble des fonctions de décision δ . Quelles sont les considérations en vertu desquelles on peut être amené à choisir une fonction de décision plutôt qu'une autre ? Si l'on prend une décision d alors que la loi qui représente le mieux le phénomène étudié - la « vraie » loi - est F , alors on encourt certains avantages et certains désagréments. Admettons, avec Wald, que ces conséquences peuvent être représentées par un nombre, que nous appellerons la « perte » : ce nombre est donc une fonction de F et de d , que nous représenterons par $W(F, d)$ (fonction de perte).

Dans certains problèmes de contrôle de la fabrication, ce concept peut être rattaché à une perte au sens ordinaire, chiffrée en argent (ce sera quelque chose comme l'espérance des divers coûts actualisés, résultat d'une décision de rejet ou d'acceptation d'un lot de fabrication). Mais dans la plupart des problèmes statistiques, il faut admettre que la fonction de perte représente des inconvénients de toute nature. On doit choisir, non pas une décision, mais une fonction de décision : la fonction de perte n'est qu'un intermédiaire, et l'on attachera à toute fonction de décision δ , l'espérance de perte correspondante

$$r(F, \delta) = \int W(F, d) dF(x),$$

appelée fonction de risque.

La considération de la fonction de risque permet-elle de choisir une règle de décision ?

Dans ce mémoire, nous présentons différentes fonctions de perte. La fonction de perte est utilisée surtout depuis l'introduction de la philosophie de Taguchi (1985), par les statisticiens et les économistes. Les travaux de Taguchi dans le domaine du contrôle de qualité ont marqué le domaine entier. La fonction de perte proposée par Taguchi n'est autre que la mesure de qualité d'un produit. La fonction de perte établit une mesure financière de l'insatisfaction de l'utilisateur vis-à-vis des performances quand elle dévie de sa valeur visée. Donc la performance moyenne et la variance sont des mesures de qualité critiques.

Traditionnellement, la mesure de la qualité est perçue comme une fonction échelon, ce qui veut dire que le produit est soit bon, soit mauvais. Ceci suppose que le produit est uniformément bon s'il satisfait les spécifications (plus basse et plus haute spécification). Taguchi croit plutôt que le client devient de plus en plus insatisfait quand la performance est loin du but. Il a suggéré d'utiliser une fonction de perte quadratique pour représenter l'insatisfaction du client avec un produit donné. La courbe est centrée sur la valeur ciblée qui donne la meilleure performance du point de vue du client. Identifier la meilleure valeur n'est pas une tâche facile. Les cibles sont très souvent la meilleure valeur supposée par le concepteur. Ceci est donc une conception basée sur l'avis du client et non sur des spécifications faites par des ingénieurs.

En se basant sur l'article de Soliman (2001), nous présentons la fonction de perte quadratique, ses estimateurs de vraisemblance et ses estimateurs bayésiens associés, reliés à la distribution de Pareto. Ensuite nous présentons la fonction de perte Linex et le calcul de ses paramètres. En se basant sur l'article de Spring (1996) comme une réponse aux critiques visant la fonction de perte quadratique, nous présentons la fonction de perte normale reflétée, laquelle est basée sur la fonction de densité normale. (Chapitre 2)

Enfin, nous présentons un exemple numérique du calcul des estimateurs sous chaque fonction de perte, quadratique, Linex et normale reflétée. Nous simulons des données et faisons les calculs en utilisant le logiciel MatLab. En utilisant les

valeurs des estimateurs nous calculons la fonction de fiabilité et aussi, la fonction du taux d'échec. (Chapitre 3)

Ensuite, nous comparons et choisissons parmi les estimateurs obtenus, le meilleur estimateur.

Également, nous présentons graphiquement la variation des fonctions de fiabilité et aussi, des fonctions du taux d'échec. (Chapitre 4)

Chapitre 2

THÉORIE DE LA DÉCISION ET FONCTIONS DE PERTE

Dans ce chapitre, nous présentons les principes de base de la théorie de la décision. Entre autres, nous expliquons les concepts de fonction de perte et d'estimateur bayésien. Ensuite nous présentons différentes fonctions de perte et les estimateurs bayésiens associés.

2.1. CADRE DÉCISIONNEL BAYÉSIEN

Dans cette section, nous présentons également les fonctions de perte usuelles en statistique bayésienne et nous décrivons les fonctions de perte que nous proposons pour notre modèle.

2.1.1. Concepts et définitions de base

Nous tenons à mentionner que cette section est inspirée de Robert (2001). Un des buts premiers de l'inférence statistique est de fournir des outils de décision. Ces outils sont souvent basés sur l'estimation des paramètres d'un certain modèle statistique.

Cependant, il est intéressant de pouvoir analyser l'impact d'une décision sur un paramètre par rapport à sa vraie valeur qui est inconnue. La théorie de la décision nous fournit un cadre pour effectuer ce type d'analyse de validation d'un estimateur. Soit Θ l'espace des paramètres, \mathcal{D} l'espace de toutes les décisions possibles et \mathcal{S} l'espace échantillonnal. Une règle de décision $\delta(x)$ est une fonction

de $\mathcal{S} \rightarrow \mathcal{D}$ telle que $\delta : x \rightarrow d = \delta(x)$; pour chaque observation x on peut prendre une décision d .

Nous allons maintenant voir un ensemble d'axiomes qui assurent l'existence de ces fonctions pour l'ensemble des décisions.

La détermination actuelle de la fonction de perte est souvent difficile, en particulier parce que, en pratique, déterminer les conséquences de chaque action pour chaque valeur de θ est presque impossible lorsque \mathcal{D} ou Θ sont de grands ensembles ou parfois des ensembles avec un nombre infini d'éléments. De plus, dans les modèles qualitatifs, il est difficile de quantifier les conséquences de chaque décision.

Comme la détermination de la fonction de perte est subjective pour la décision, le responsable de la décision incite souvent le statisticien à utiliser les pertes classiques, sélectionnées à cause de leur simplicité. De telles pertes sont aussi nécessaires pour un traitement théorique de la déduction des procédures optimales, quand il n'y a aucune motivation pratique pour le choix d'une fonction de perte particulière. Il est toujours mieux de prendre une décision dans un temps fini en utilisant un critère approximatif plutôt qu'un temps très grand pour déterminer exactement la fonction de perte nécessaire.

Une base fondamentale de la théorie de la décision bayésienne est l'inférence statistique qui devrait commencer avec la détermination rigoureuse de trois facteurs :

- (1) la famille de loi des observations, $f(x|\theta)$, variable aléatoire continue ;
- (2) la distribution a priori des paramètres, $\pi(\theta)$;
- (3) la perte associée aux décisions, $L(\delta, \theta)$.

La distribution a priori, la perte et même quelquefois la distribution d'échantillonnage, sont tirées de considérations partiellement subjectives.

Les critiques les plus fréquentes du paradigme bayésien oublient souvent de tenir compte du problème de la construction de la fonction de perte, bien que cela puisse être au moins aussi compliqué que la détermination de la distribution a priori. En plus, l'existence d'une fonction de perte implique des informations a priori sur le modèle qu'une analyse bayésienne pourrait utiliser plus efficacement.

Lindley (1985) déclare que la perte et la distribution a priori sont difficiles à séparer et devraient être analysées simultanément. Dans certains cas, il est possible de réduire la classe de fonctions de perte acceptables par des considérations d'invariance, par exemple quand le modèle est invariant sous l'action d'un groupe de transformations. Ces arguments d'invariance sont également utilisés par les adversaires de l'approche décisionnelle, car ils doivent faire appel à des restrictions de l'espace des décisions pour pouvoir justifier leur procédure inférentielle.

Exemple : Considérons le problème de l'estimation de la moyenne θ d'un vecteur normal, $X \sim \mathcal{N}_n(\theta, \Sigma)$, où Σ est la matrice diagonale connue avec les éléments diagonaux $\sigma_i^2 (1 \leq i \leq n)$. Dans ce cas, $\mathcal{D} = \Theta = \mathbf{R}^n$, et δ_i représente une évaluation de θ_i , pour n'importe quel i , $1 \leq i \leq n$. Si aucun renseignement supplémentaire n'est disponible sur le modèle, il semble logique de choisir une fonction de perte de la forme

$$\sum_{i=1}^n L\left(\frac{\delta_i - \theta_i}{\sigma_i}\right), \quad (2.1.1)$$

où L prend son minimum à 0. Effectivement, les composantes avec une plus grande variance n'ont pas trop de poids quand les erreurs d'estimation $(\delta_i - \theta_i)$ sont normalisées par σ_i . Le choix usuel de L est la perte quadratique, c'est-à-dire l'erreur d'estimation globale est la somme des erreurs au carré. Il faudrait spécifier que les paramètres ne sont pas indépendants. Car s'ils sont indépendants, alors la minimisation de la somme des pertes

$$\sum_{i=1}^n L\left(\frac{\delta_i - \theta_i}{\sigma_i}\right)$$

donnera le même résultat que l'optimisation individuelle de chacun des paramètres et alors, les poids relatifs accordés aux erreurs n'ont aucune importance.

2.1.2. Utilité et perte

Du point de vue de la théorie de la décision, le modèle statistique implique maintenant trois espaces : \mathcal{X} , l'espace des observations, Θ , l'espace des paramètres, et \mathcal{D} , l'espace des décisions. L'inférence statistique consiste donc à prendre

une décision $d \in \mathcal{D}$ concernant $\theta \in \Theta$ au vu d'une observation $x \in \mathcal{X}$, x et θ étant reliés par la loi de la variable aléatoire $f(x|\theta)$.

La plupart du temps, la décision d sera obtenue en évaluant une fonction de θ , $h(\theta)$, aussi exactement que possible.

La théorie de la décision ajoute l'hypothèse supplémentaire que chaque action d peut être évaluée et donc conduit à une récompense r , d'utilité $U(r)$ (sous l'hypothèse de rationalité des décideurs). Nous représenterons cette utilité par $U(\theta, d)$, pour souligner l'influence de ces deux facteurs. Quand d'autres facteurs aléatoires r interviennent dans U , nous écrirons $U(\theta, d) = \mathbf{E}[U(r)]$. Donc, $U(\theta, d)$ peut être vue comme mesure de proximité entre l'estimation proposée d et la vraie valeur $h(\theta)$. Dès que la fonction d'utilité a été construite (ou approximée), nous définissons la fonction de perte correspondante :

$$L(\theta, d) = -U(\theta, d).$$

En général, la fonction de perte doit être non-négative, ce qui implique que $U(\theta, d) \leq 0$, et donc qu'il n'y a aucune décision avec une utilité infinie. L'existence d'une limite inférieure sur L peut être critiquée comme étant trop stricte.

Il peut aussi être soutenu que, du point de vue statistique, la fonction de perte L représente effectivement la perte encourue par une mauvaise évaluation de la fonction d'intérêt de θ , et donc que même la meilleure évaluation de cette fonction, lorsque θ est connue, est au mieux une perte nulle. Autrement, il y aurait une discontinuité de la fonction de perte qui pourrait même prévenir le choix d'une procédure de décision.

Évidemment, à part les cadres les plus banals, il est généralement impossible d'uniformément minimiser la fonction de perte quand θ est inconnu. Pour tirer un critère de comparaison efficace de la fonction de perte, l'approche fréquentiste propose de considérer plutôt la perte moyenne (ou le risque fréquentiste) :

$$\begin{aligned} R(\theta, \delta) &= \mathbf{E}_\theta[L(\theta, \delta(X))] \\ &= \int_{\mathcal{X}} L(\theta, \delta(x)) f(x|\theta) dx, \end{aligned}$$

où $\delta(x)$ est la règle de décision, c'est-à-dire l'allocation d'une décision à chaque résultat x (les observations x suivent la densité $f(x|\theta)$), de l'expérience au hasard. La fonction δ , de \mathcal{X} en \mathcal{D} , est nommée estimateur (alors que la valeur $\delta(x)$ est l'estimé de θ).

Quand il n'y a pas de risque de confusion, on note par \mathcal{D} , l'ensemble des estimateurs. Le paradigme fréquentiste compte sur ce critère pour comparer des estimateurs et, si possible, choisir le meilleur estimateur, le raisonnement étant que les estimateurs sont évalués sur leur performance de longue durée pour toutes les valeurs possibles du paramètre θ .

Remarquez pourtant, qu'il y a plusieurs difficultés associées à cette approche :

- (1) L'erreur (la perte) est faite en moyenne sur les différentes valeurs de x proportionnellement à la densité $f(x|\theta)$. Nous ne tirons pas parti de l'observation x autrement que par le fait qu'elle provient de la loi $f(x|\theta)$. Le critère de risque évalue des procédures sur leur performance à long terme et pas directement pour l'observation donnée, x . Une telle évaluation peut être satisfaisante pour le statisticien.
- (2) L'analyse fréquentiste du problème de décision suppose implicitement que ce problème sera rencontré à plusieurs reprises, pour l'évaluation de la fréquence, pour avoir du sens. En effet, $R(\theta, \delta)$ est approximativement la perte moyenne sur les répétitions indépendentes et identiquement distribuées de la même expérience, selon la loi des grands nombres.
- (3) Pour une procédure δ , le risque $R(\theta, \delta)$ est une fonction du paramètre θ . Donc, la comparaison des procédures de décision sera compliquée, sinon impossible, d'un point de vue fréquentiste, puisque le critère ne permet pas de discriminer entre deux fonctions de risque lorsqu'elles se croisent. Au mieux, on pourra chercher s'il existe une décision δ_0 qui minimise $R(\theta, \delta)$ uniformément en θ .

Contrairement à l'approche fréquentiste, l'approche bayésienne pour la théorie de la décision s'intègre sur l'espace Θ puisque θ est inconnue, au lieu de s'intégrer sur l'espace \mathcal{X} , comme x est connu. Elle compte sur l'espérance de perte a

posteriori :

$$\varrho(\pi, d|x) = \mathbb{E}^\pi[L(\theta, d)|x] = \int_{\Theta} L(\theta, d)\pi(\theta|x)d\theta, \quad (2.1.2)$$

où $\pi(\theta|x)$ est la loi a posteriori continue de θ donnée par l'expression : $\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta}$ et qui prend la moyenne de l'erreur (la perte) selon la distribution a posteriori du paramètre θ , conditionnellement à la valeur x . Étant donné x , l'erreur moyenne résultante de la décision d est en effet $\varrho(\pi, d|x)$. La fonction de perte a posteriori est donc une fonction de x mais cette dépendance ne dérange pas, contrairement à la dépendance fréquentiste du risque au paramètre parce que x , contrairement à θ , est connu.

Étant donné une distribution a priori $\pi(\theta)$, il est aussi possible de définir le risque intégral, qui est le risque fréquentiste moyen sur les valeurs de θ selon leur distribution a priori :

$$r(\pi, \delta) = E^*[R(\theta, \delta)] = \int_{\Theta} \int_{\mathcal{X}} L(\theta, \delta(x))f(x|\theta)dx \pi(\theta)d\theta. \quad (2.1.3)$$

L'intérêt de ce dernier concept est qu'il associe un nombre réel avec chaque estimateur, pas une fonction de θ . Ceci induit donc un ordre total sur l'ensemble des estimateurs et donc il permet une comparaison directe des estimateurs.

Ceci implique qu'en tenant compte de l'information a priori au travers de la loi a priori, l'approche bayésienne est suffisamment réductrice (dans un sens positif) pour permettre une décision effective.

De plus, les deux notions présentées sont équivalentes dans le sens où elles mènent à la même décision.

Théorème 2.1.1. *Un estimateur minimisant le risque intégral $r(\pi, \delta)$ peut être obtenu en sélectionnant pour chaque $x \in \mathcal{X}$, la valeur $\delta(x)$ qui minimise la perte a posteriori $\varrho(\pi(\theta), \delta|x)$ puisque :*

$$r(\pi, \delta) = \int_{\mathcal{X}} \varrho(\pi, \delta(x)|x)f(x)dx, \quad (2.1.4)$$

où $f(x) = \int_{\Theta} f(x|\theta)\pi(\theta)d\theta$.

DÉMONSTRATION. En utilisant la théorème de Fubini nous permettant d'interchanger l'ordre des intégrales, nous avons :

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

□

Ce résultat donne la définition suivante d'un estimateur bayésien.

Définition 2.1.1. *Un estimateur bayésien associé à une distribution a priori π et une fonction de perte L est un estimateur $\delta^\pi(x)$ qui minimise $r(\pi, \delta)$. Pour chaque $x \in \mathcal{X}$, il est donné par $\delta^\pi(x)$, le minimum d'après le d de $\varrho(\pi, d|x)$. La valeur $r(\pi) = r(\pi, \delta^\pi)$ est alors appelée le risque de Bayes.*

Le théorème précédent donne alors le moyen de déterminer des estimateurs de Bayes. Le résultat est valide pour les lois propres et les lois impropres, donc même si le risque bayésien est infini. Dans le cas d'une loi impropre, on définira l'estimateur de Bayes associé, dit estimateur de Bayes généralisé, comme celui minimisant, pour chaque x , la perte a posteriori.

Dans la sous-section suivante, nous étudierons l'optimalité de l'estimateur de Bayes du point de vue critère d'optimalité fréquentiste.

2.2. DEUX CRITÈRES D'OPTIMALITÉ : MINIMAXITÉ ET ADMISSIBILITÉ

Afin de traiter ces deux critères d'optimalité, nous allons définir les estimateurs aléatoires, ensuite nous présenterons les critères d'optimalité et nous vérifierons que l'estimateur de Bayes est optimal dans le sens de ces critères.

2.2.1. Estimateurs aléatoires

Similairement à l'étude de la fonction d'utilité, nous devons inclure, en sus des estimateurs classiques, les estimateurs aléatoires. Alors nous étendons l'espace de décision D , à l'ensemble des estimateurs aléatoires, D^* . Utiliser un estimateur aléatoire δ^* signifie que l'action est générée selon la distribution avec une densité de probabilité $\delta^*(x, \cdot)$, une fois que l'observation x a été prise. La perte d'un estimateur δ^* est alors définie comme la perte moyenne

$$L(\theta, \delta^*(x)) = \int_D L(\theta, a) \delta^*(x, a) da, \quad (2.2.1)$$

où a est un vecteur de paramètres. Cette extension est nécessaire au traitement des notions de minimaxité et d'admissibilité. Bien sûr, ces estimateurs ne sont pas utilisables, ne serait-ce que parce qu'ils contredisent le principe de vraisemblance, qui donne plusieurs réponses possibles pour la même valeur de x (et donc de $l(\theta|x)$).

Les estimateurs aléatoires sont cependant nécessaires du point de vue fréquentiste, par exemple, pour la théorie fréquentiste, puisqu'ils donnent accès à des niveaux de confiance autrement impossibles à atteindre. L'ensemble D^* apparaît alors comme une version complète de D . Cependant, la modification de l'espace de décision ne change pas la réponse bayésienne, comme montré dans le résultat ci-dessous.

Théorème 2.2.1. *Pour toute distribution π sur Θ , le risque de Bayes sur l'ensemble des estimateurs aléatoires est le même que le risque de Bayes sur l'ensemble d'estimateurs non-aléatoires :*

$$\inf_{\delta \in D} r(\pi, \delta) = \inf_{\delta^* \in D^*} r(\pi, \delta^*) = r(\pi). \quad (2.2.2)$$

DÉMONSTRATION. Pour tout $x \in X$ et pour tout $\delta^* \in D^*$, nous avons :

$$\begin{aligned} & \int_{\Theta} \int_D L(\theta, a) \delta^*(x, a) da \pi(\theta|x) d\theta \\ &= \int_D \int_{\Theta} L(\theta, a) \pi(\theta|x) d\theta \delta^*(x, a) da \end{aligned}$$

$$\begin{aligned} &\geq \int_D \inf_a \int_{\Theta} L(\theta, a) \pi(\theta|x) d\theta \delta^*(x, a) da \\ &= \varrho(\pi, \delta^\pi | x). \end{aligned}$$

□

2.2.2. Minimaxité

Le critère de minimaxité introduit ici apparait comme une assurance contre le pire cas parce qu'il vise à minimiser l'espérance de perte dans le cas le plus défavorable.

Définition 2.2.1. *Le risque minimax associé avec une fonction de perte L est la valeur*

$$\bar{R} = \inf_{\delta \in D^*} \sup_{\theta} R(\theta, \delta) = \inf_{\delta \in D^*} \sup_{\theta} E_{\theta}[L(\theta, \delta(X))] \quad (2.2.3)$$

(i.e. la borne inférieure d'après le δ qui appartient à D^* , de la borne supérieure d'après le θ de la fonction de risque) et l'estimateur minimax est tout estimateur δ_0 (éventuellement aléatoire) tel que

$$\sup_{\theta} R(\theta, \delta_0) = \bar{R}.$$

2.2.3. Existence des règles minimax et stratégie maximin

Une importante difficulté avec la minimaxité est que l'estimateur n'existe pas nécessairement. On a des conditions suffisantes pour l'existence, mais en particulier, il existe une stratégie minimax quand Θ est fini et la fonction de perte est continue. Plus généralement, l'espace de décision D a été considéré comme étant contenu dans un autre espace, ce qui rend l'ensemble des fonctions de risque dans cet espace plus grand compact. De ce point de vue et sous certaines conditions, il est possible de déterminer les estimateurs minimax quand la fonction de perte est continue. Cependant, ces extensions impliquent des techniques topologiques trop avancées pour être incluses dans cette thèse. Nous donnons alors seulement le résultat suivant (voir Blackwell et Girshick (1954) pour une preuve).

Théorème 2.2.2. *Si $D \subset \mathbb{R}^k$ est un ensemble compact et si $L(\theta, d)$ est continue et convexe sur la variable d pour tout $\theta \in \Theta$, alors il existe un estimateur minimax non-aléatoire.*

La restriction aux estimateurs non-aléatoires quand la perte est convexe vient de l'inégalité de Jensen puisque

$$L(\theta, \delta^*) = \mathbb{E}^{\delta^*} [L(\theta, \delta)] \geq L(\theta, \mathbb{E}^{\delta^*}(\delta)).$$

Ce résultat est un cas spécial du théorème de Rao-Blackwell (voir Lehmann and Casella (1998)). Le résultat suivant souligne la relation entre l'approche bayésienne et le principe minimax.

Lemme 2.2.1. *Le risque bayésien est toujours plus petit que le risque minimax :*

$$\underline{R} = \sup_{\pi} r(\pi) = \sup_{\pi} \inf_{\delta \in D} r(\pi, \delta) \leq \bar{R} = \inf_{\delta \in D} \sup_{\theta} R(\theta, \delta).$$

La première valeur est appelée risque maximin et la distribution π^* telle que $r(\pi^*) = \underline{R}$ est appelée la distribution la moins favorable, quand elle existe. Ces dernières impliquent le plus grand risque de Bayes possible et donc les distributions les moins intéressantes en terme de performance de perte s'ils ne sont pas suggérés par l'information déjà disponible. Ce résultat est logique car l'information disponible ne peut qu'améliorer les erreurs d'estimations même dans le pire des cas.

Un cas particulièrement intéressant correspond à la définition suivante.

Définition 2.2.2. *On dit que le problème d'estimation admet une solution quand*

$$\underline{R} = \bar{R}$$

et donc quand

$$\sup_{\pi} \inf_{\delta \in D} r(\pi, \delta) = \inf_{\delta \in D} \sup_{\theta} R(\theta, \delta).$$

Lorsque le problème admet une valeur, on peut obtenir des estimateurs minimax en calculant les estimateurs de Bayes associés aux distributions a priori les moins favorables. En pratique, le lemme suivant permet de donner des conditions suffisantes pour la minimaxité.

Lemme 2.2.2. *Si δ_0 un estimateur de Bayes en ce qui concerne π_0 et si $R(\theta, \delta_0) \leq r(\pi_0)$ pour chaque θ sur le support de π_0 , alors δ_0 est minimax et π_0 est la distribution la moins favorable.*

Puisque l'estimateur minimax correspond à l'estimateur de Bayes généralisé (dans le cas d'un risque bayésien infini, on définira l'estimateur de Bayes associé, dit estimateur de Bayes généralisé, comme celui minimisant, pour chaque x , la perte a posteriori), il est souvent nécessaire d'utiliser un argument limite pour établir la minimaxité, au lieu de calculer directement le risque de Bayes en utilisant le lemme précédent.

Lemme 2.2.3. *S'il existe une séquence (π_n) de distributions propres telle que l'estimateur généralisé de Bayes δ_0 satisfait*

$$R(\theta, \delta_0) \leq \lim_{n \rightarrow \infty} r(\pi_n) < +\infty$$

pour chaque $\theta \in \Theta$, alors δ_0 est minimax.

2.2.4. Admissibilité

Ce deuxième critère fréquentiste inclut un ordonnancement partiel sur D^* en comparant les risques fréquentistes des estimateurs, $R(\theta, \delta)$.

Définition 2.2.3. *Un estimateur δ_0 est inadmissible s'il existe un estimateur δ_1 qui domine δ_0 , i.e. pour tout θ ,*

$$R(\theta, \delta_0) \geq R(\theta, \delta_1),$$

et pour au moins une valeur θ_0 du paramètre,

$$R(\theta_0, \delta_0) > R(\theta_0, \delta_1).$$

Si non, δ_0 est dit admissible.

Ce critère est particulièrement intéressant pour son action réductive. En effet, théoriquement, les estimateurs inadmissibles ne devraient pas être considérés car ils peuvent être uniformément améliorés. L'admissibilité toute seule n'est pas suffisante pour valider l'utilisation d'un estimateur. Par exemple, un estimateur constant $\delta(x) = \theta_0$ est généralement admissible. D'un point de vue fréquentiste, il est donc important de trouver un estimateur qui satisfait les deux critères, minimaxité et admissibilité.

Proposition 2.2.1. *Si il existe un estimateur minimax unique, cet estimateur est admissible.*

DÉMONSTRATION. Si δ^* est le seul estimateur minimax, pour tout estimateur $\tilde{\delta} \neq \delta^*$,

$$\sup_{\theta} R(\theta, \tilde{\delta}) > \sup_{\theta} R(\theta, \delta^*).$$

Donc $\tilde{\delta}$ ne peut dominer δ^* .

□

Quand la fonction de perte L est strictement convexe, elle nous permet d'avoir la caractérisation suivante.

Proposition 2.2.2. *Si δ_0 est admissible avec un risque constant, δ_0 est l'unique estimateur minimax.*

DÉMONSTRATION. Pour $\delta_0 \in \Theta$, on a

$$\sup_{\theta} R(\theta, \delta_0) = \sup_{\theta} R(\theta_0, \delta_0).$$

Alors, s'il existe δ_1 tel que $\tilde{R} \leq \sup_{\theta} R(\theta, \delta_1) < R(\theta_0, \delta_0)$, δ_0 ne peut être admissible.

□

Nous avons montré précédemment que certains estimateurs minimax sont des estimateurs bayésiens aussi. L'admissibilité est encore plus reliée au paradigme de Bayes dans le sens où, dans la plupart des problèmes statistiques, les estimateurs de Bayes « engendrent » la classe des estimateurs admissibles. En effet, ces derniers peuvent être exprimés par des estimateurs de Bayes ou des estimateurs généralisés de Bayes ou par des limites d'estimateurs de Bayes. Nous donnons ici deux propositions importantes.

Proposition 2.2.3. *Si une distribution a priori π est strictement positive sur Θ , avec un risque de Bayes fini et la fonction de risque $R(\theta, \delta)$ est une fonction continue de θ pour chaque δ , l'estimateur de Bayes δ^π est admissible. (On va noter δ^π , l'estimateur de Bayes associé à la distribution a priori π .)*

DÉMONSTRATION. Supposons que δ^π est inadmissible et soit δ' qui domine uniformément δ^π . On a alors pour chaque θ , $R(\theta, \delta') \leq R(\theta, \delta^\pi)$ et dans un ensemble ouvert C de Θ , $R(\theta, \delta') < R(\theta, \delta^\pi)$. En intégrant cette inégalité, on obtient :

$$r(\pi, \delta') < r(\pi, \delta^\pi) = \int_{\Theta} R(\theta, \delta^\pi) \pi(\theta) d\theta,$$

ce qui est impossible, parce que le risque bayésien est toujours plus petit que le risque minimax (lemme 2.2.1.).

□

Proposition 2.2.4. *Si l'estimateur de Bayes associé à la distribution a priori π est unique, alors il est admissible.*

Même si l'estimateur de Bayes n'est pas unique, il est toujours possible de trouver au moins un estimateur de Bayes admissible.

Proposition 2.2.5. *Si un estimateur de Bayes, δ^π , associé avec une distribution a priori π , est tel que le risque de Bayes*

Nous avons montré précédemment que certains estimateurs minimax sont des estimateurs bayésiens aussi. L'admissibilité est encore plus reliée au paradigme de Bayes dans le sens où, dans la plupart des problèmes statistiques, les estimateurs de Bayes « engendrent » la classe des estimateurs admissibles. En effet, ces derniers peuvent être exprimés par des estimateurs de Bayes ou des estimateurs généralisés de Bayes ou par des limites d'estimateurs de Bayes. Nous donnons ici deux propositions importantes.

Proposition 2.2.3. *Si une distribution a priori π est strictement positive sur Θ , avec un risque de Bayes fini et la fonction de risque $R(\theta, \delta)$ est une fonction continue de θ pour chaque δ , l'estimateur de Bayes δ^π est admissible. (On va noter δ^π , l'estimateur de Bayes associé à la distribution a priori π .)*

DÉMONSTRATION. Supposons que δ^π est inadmissible et soit δ' qui domine uniformément δ^π . On a alors pour chaque θ , $R(\theta, \delta') \leq R(\theta, \delta^\pi)$ et dans un ensemble ouvert C de Θ , $R(\theta, \delta') < R(\theta, \delta^\pi)$. En intégrant cette inégalité, on obtient :

$$r(\pi, \delta') < r(\pi, \delta^\pi) = \int_{\Theta} R(\theta, \delta^\pi) \pi(\theta) d\theta,$$

ce qui est impossible, parce que le risque bayésien est toujours plus petit que le risque minimax (lemme 2.2.1.).

□

Proposition 2.2.4. *Si l'estimateur de Bayes associé à la distribution a priori π est unique, alors il est admissible.*

Même si l'estimateur de Bayes n'est pas unique, il est toujours possible de trouver au moins un estimateur de Bayes admissible.

Proposition 2.2.5. *Si un estimateur de Bayes, δ^π , associé avec une distribution a priori π , est tel que le risque de Bayes*

$$r(\pi) = \int_{\Theta} R(\theta, \delta^\pi) \pi(\theta) d\theta$$

est fini, alors δ^π est admissible.

2.3. ANALYSE DE LA FONCTION DE PERTE

La fonction de perte illustre la relation entre la baisse de performance et les coûts que celle-ci entraîne pour le client et la société en général. On considère en effet que tout écart par rapport à la valeur optimale engendre des pertes, car le produit n'offre plus les performances optimales. Cette fonction peut aussi être utilisée pour définir les spécifications et les tolérances en fonction de l'analyse des coûts.

2.3.1. Fonction de perte quadratique

Proposée par Legendre (1805) et Gauss (1810), cette perte est sans aucun doute le critère d'évaluation le plus commun.

$$L(\theta, d) = (\theta - d)^2. \quad (2.3.1)$$

Dans son article de 1810, Gauss a déjà reconnu le caractère aléatoire de la perte quadratique et le défendait pour des raisons de simplicité. De telles critiques restent valides aujourd'hui. Mais cette perte n'en demeure pas moins utilisée car elle donne en général des solutions bayésiennes acceptables, i.e. celles fournies par une inférence non-décisionnelle fondée sur la densité a posteriori.

Proposition 2.3.1. *L'estimateur de Bayes δ^π associé à la distribution a priori π et avec la perte quadratique, est donné par l'espérance a posteriori*

$$\delta^\pi = E^\pi[\theta|x] = \frac{\int_{\theta} \theta f(x|\theta) \pi(\theta) d\theta}{\int_{\theta} f(x|\theta) \pi(\theta) d\theta} = \delta^\pi(x).$$

DÉMONSTRATION. $E^\pi[(\theta - \delta)^2|x] = E^\pi[\theta^2|x] - 2\delta E^\pi[\theta|x] + \delta^2$. La perte actuelle a posteriori atteint un minimum à $\delta^\pi(x) = E^\pi[\theta|x]$. \square

Corollaire 2.3.1. *L'estimateur de Bayes δ^π associé avec π et avec la perte quadratique pondérée $L(\theta, \delta) = \omega(\theta)(\theta - \delta)^2$, où $\omega(\theta)$ est une fonction non-négative,*

est

$$\delta^\pi = \frac{E^\pi[\omega(\theta)\theta|x]}{E^\pi[\omega(\theta)|x]}.$$

Corollaire 2.3.2. *Quand $\Theta \subseteq \mathbb{R}^p$, l'estimateur de Bayes δ^π associé avec π et avec la perte quadratique pondérée $L(\theta, \delta) = (\theta - \delta)^t Q(\theta - \delta)$ est l'espérance a posteriori*

$$\delta^\pi(x) = E^\pi[\theta|x],$$

pour chaque matrice Q , de dimension $p \times p$, symétrique et positive définie.

Nous présentons un exemple du calcul de l'estimateur sous le coût

$$L(\theta, \delta) = 1 - I_\theta(\delta)$$

et sous le coût quadratique, où $I_\theta(\delta)$ est 1, si $\delta = \theta$, et 0 sinon.

Exemple 2.3.1. *Soient x_1 et x_2 deux observations de*

$$P_\theta(x = \theta - 1) = P_\theta(x = \theta + 1) = 0.5.$$

On cherche à estimer θ (donc ici $D = \Theta$) par $\delta(x_1, x_2)$ sous le coût

$$L(\theta, \delta) = 1 - I_\theta(\delta),$$

Pour $\delta_0(x_1, x_2) = \frac{x_1 + x_2}{2}$, on a

$$\begin{aligned} R(\theta, \delta_0) &= 1 - P_\theta(\delta_0(x_1, x_2) = \theta) \\ &= 1 - P_\theta(x_1 \neq x_2) = 0.5. \end{aligned}$$

Ce calcul ne montre pas que l'estimateur est toujours exact si $x_1 \neq x_2$ et toujours faux sinon. Il exprime clairement le besoin d'un critère plus adaptatif (i.e. plus conditionnel). Pour le modèle considéré, on cherche à estimer la 'valeur' de l'estimateur de θ ,

$$\delta(x_1, x_2) = \begin{cases} (x_1 + x_2)/2, & \text{si } x_1 \neq x_2 \\ x_1 + 1, & \text{sinon,} \end{cases}$$

par $\alpha(x_1, x_2)$ sous le coût quadratique (2.3.2)

$[I_\theta(\delta(x_1, x_2)) - \alpha(x_1, x_2)]^2$. La fonction α évalue d'une certaine façon la probabilité que δ prenne la vraie valeur θ .

Deux estimateurs peuvent être proposés :

(i) $\alpha_0(x_1, x_2) = 0.75$, qui est l'espérance de $I_\theta(\delta(x_1, x_2))$,

et

(ii) $\alpha_1(x_1, x_2) = \begin{cases} 1, & \text{si } x_1 \neq x_2 \\ 0.5, & \text{si } x_1 = x_2. \end{cases}$

Les risques des deux estimateurs sont alors :

$$R(\theta, \alpha_0) = E_\theta(I_\theta(\delta(x_1, x_2)) - 0.75)^2 = 0.75 - 0.75^2 = 0.1875$$

et

$$R(\theta, \alpha_1) = E_\theta(I_\theta(\delta(x_1, x_2)) - \alpha_1)^2 = 0.5^2 \frac{1}{2} = 0.125$$

Alors α_1 sera préféré à α_0 , puisque cet estimateur est plus adaptatif. Cet exemple est présenté par Berger et Wolpert (1988) comme un argument en faveur des raisonnements conditionnels.

2.3.2. Fonction de perte erreur absolue

Une solution alternative à la perte quadratique en une dimension est l'utilisation de la perte erreur absolue :

$$L(\theta, d) = |\theta - d|,$$

considérée par Laplace (1773), ou, plus généralement une fonction multilinéaire :

$$L_{k_1, k_2}(\theta, d) = \begin{cases} k_2(\theta - d), & \text{si } \theta > d \\ k_1(d - \theta), & \text{sinon} \end{cases}$$

Ces fonctions croissent moins vite que les fonctions de perte quadratique.

Huber (1964) propose un mélange entre les fonctions de pertes erreur absolue et les fonctions de perte quadratique, de manière à garder la pénalisation quadratique autour de 0.

$$L(\theta, d) = \begin{cases} (d - \theta)^2, & \text{si } |d - \theta| < k \\ 2k|d - \theta| - k^2, & \text{sinon.} \end{cases}$$

Proposition 2.3.2. *Un estimateur de Bayes associé à la distribution π et à la perte multi-linéaire, est un $(\frac{k_2}{k_1+k_2})$ fractile de $\pi(\theta|x)$.*

DÉMONSTRATION. Nous avons

$$\begin{aligned} & \mathbb{E}(L(\theta, a)|x) \\ &= k_1 \int_{-\infty}^a (a - \theta)\pi(\theta|x)d\theta + k_2 \int_a^{+\infty} (\theta - a)\pi(\theta|x)d\theta \\ &= k_1 \int_{-\infty}^a P(\theta < y|x)dy + k_2 \int_a^{+\infty} P(\theta > y|x)dy, \end{aligned}$$

par une intégration par parties. Dérivant par rapport à a , nous obtenons alors

$$k_1 P(\theta < a|x) - k_2 P(\theta > a|x) = 0,$$

soit

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

□

En particulier, si $k_1 = k_2$, c'est-à-dire en cas d'une erreur de perte absolue, l'estimateur de Bayes est la médiane postérieure, qui est l'estimateur obtenu par Laplace. Il est à noter que lorsque π a un support discontinu, la proposition précédente donne un exemple d'estimateur de Bayes multiple pour certaines valeurs de x .

2.3.3. Perte 0-1

Cette fonction de perte est utilisée surtout dans les méthodes classiques de test d'hypothèse, formalisées par Neyman et Pearson. Plus généralement, c'est un exemple typique de perte non quantitative. En effet, pour cette perte, la pénalité associée avec un estimateur δ est 0 si la réponse est correcte et 1 sinon.

Exemple 2.3.2. : Prenons le test de $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \notin \Theta_0$. Alors $\mathcal{D} = \{0, 1\}$, où 1 signifie l'acceptation de H_0 et 0 le rejet. Pour la perte 0-1

$$L(\theta, d) = \begin{cases} (1 - d), & \text{si } \theta \in \Theta_0 \\ d, & \text{sinon,} \end{cases} \quad (2.3.3)$$

le risque associé est alors

$$R = \mathbb{E}_\theta[L(\theta, \delta(x))] = \begin{cases} P_\theta(\delta(x) = 0), & \text{si } \theta \in \Theta_0 \\ P_\theta(\delta(x) = 1), & \text{sinon} \end{cases}$$

ce qui donne exactement les erreurs de type-1 et type-2 à la base de la théorie de Neyman-Pearson.

Proposition 2.3.3. L'estimateur de Bayes associé avec π et avec la perte (2.3.3) est

$$\delta^\pi(x) = \begin{cases} 1, & \text{si } P(\theta \in \Theta_0|x) > P(\theta \notin \Theta_0|x) \\ 0, & \text{sinon} \end{cases}$$

donc δ^π est égal à 1 si et seulement si $P(\theta \in \Theta_0|x) > \frac{1}{2}$.

2.3.4. Perte intrinsèque

Il arrive que certains cas soient si non-informatifs que non seulement la fonction de perte est inconnue, mais il n'y a même pas de paramétrisation naturelle. Ces cas arrivent lorsque la distribution $f(x|\theta)$ est elle-même d'intérêt, par exemple, dans les cas de prédictions. Dans ces cas non-informatifs, il paraît naturel d'utiliser des fonctions de perte qui comparent directement les distributions $f(\cdot|\theta)$ et $f(\cdot|\delta)$ associées avec le vrai paramètre θ et l'estimé δ . La fonction de perte

$$L(\theta, \delta) = d(f(\cdot|\theta), f(\cdot|\delta))$$

est en fait sans paramètres. Deux mesures de distances usuelles sont :

- (1) L'entropie : $L_e(\theta, \delta) = \mathbb{E}[\log(\frac{f(x|\theta)}{f(x|\delta)})]$ qui est aussi appelée la divergence de Kullback-Leibler et qui n'est pas une distance dans le sens mathématique, car elle n'est pas symétrique.
- (2) La distance de Hellinger : $L_H(\theta, \delta) = \frac{1}{2}\mathbb{E}_\theta[(\sqrt{\frac{f(x|\delta)}{f(x|\theta)}} - 1)^2]$.

2.3.5. Fonctions de perte asymétriques

Nous introduisons dans cette section une grande classe de fonctions de perte asymétriques. Supposons que nous devons estimer une variable $\theta \in \Theta$ continue ou discrète. Ici, Θ est soit un ensemble continu sur \mathbb{R} ou bien une collection infinie de point réels $\{\dots < \theta_0 < \theta_1 < \dots\}$. Premièrement, nous choisissons une mesure de probabilité arbitraire Π sur Θ et une fonction de perte $L(\delta, \theta)$, où δ est la règle de décision. La perte peut être une des fonctions de perte standard et symétrique; cependant la symétrie n'est pas requise pour la discussion qui suit. La mesure Π peut être interprétée comme une distribution a priori, ou a posteriori, après que l'échantillon $X = \{X_1, \dots, X_n\}$ eut été observé.

Soit la fonction de perte asymétrique W de la forme

$$W(\delta, \theta) = K_1 L(\delta, \theta) I\{\delta \leq \theta\} + K_2 L(\delta, \theta) I\{\delta > \theta\} \quad (2.3.4)$$

pour des valeurs positives K_1 et K_2 .

On peut se limiter à étudier seulement le cas où $K_1 \leq K_2$, parce que l'autre cas peut être obtenu en remplaçant θ par $-\theta$. Nous pouvons aussi supposer que $K_1 = 1$. Donc la fonction de perte W est de la forme :

$$W(\delta, \theta) = \begin{cases} L(\delta, \theta), & \text{si } \delta \leq \theta \\ (1 + \lambda)L(\delta, \theta), & \text{si } \delta > \theta \end{cases} \quad (2.3.5)$$

pour $\lambda \geq 0$, qui correspond au coût relatif additionnel de la surestimation. Soit δ^* l'estimateur minimal de risque de θ , c'est-à-dire, l'argument qui minimise la fonction $r(\delta)$:

$$\delta^* = \delta^*(\lambda) = \operatorname{argmin}_\delta r(\delta)$$

où $r(\delta) = \int W(\delta, \theta) d\Pi(\theta)$.

Si la surestimation a un coût plus élevé, on tend à sous-estimer θ . Un moyen d'exprimer ceci est de considérer une famille de mesures pondérées $\{\Pi_t\}$ obtenues

de Π en augmentant les probabilités des petites valeurs de θ ,

$$\Pi_t = \begin{cases} (1 + \lambda)\Pi / (1 + \lambda \int_{\theta \leq t} d\Pi(\theta)), & \text{si } \theta < t \\ \Pi / (1 + \lambda \int_{\theta \leq t} d\Pi(\theta)), & \text{si } \theta \geq t. \end{cases} \quad (2.3.6)$$

Pour chaque t , soit

$$\tilde{\delta}(t) = \operatorname{argmin}_{\delta} r_t(\delta),$$

c'est-à-dire, l'argument qui minimise la fonction $r_t(\delta)$, où $r_t(\delta) = \int L(\delta, \theta) d\Pi_t(\theta)$.

Si Π est une distribution a posteriori $\pi(\theta|\mathbf{X})$ sous certaines lois a priori $\pi(\theta)$, alors δ^* est la règle de Bayes. Dans ce cas, on définit une fonction a priori $\pi_t(\theta)$ similairement à (2.3.6) en assignant les plus grandes probabilités aux petites valeurs de θ . Donc Π_t dans (2.3.6) est la fonction a posteriori qui lui correspond.

2.3.6. 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 croît presque exponentiellement d'un côté de zéro et est approximativement linéaire de l'autre côté. Sous l'hypothèse que la perte minimale est obtenue pour $\tilde{u} = u$, la fonction de perte Linex pour u soit α , soit β s'exprime par

$$L(\Delta) \propto e^{a\Delta} - a\Delta - 1, \quad a \neq 0, \quad (2.3.7)$$

où $\Delta = (\tilde{u} - u)$ et \tilde{u} est un estimateur de u . Le signe et la norme de a représentent 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; (2.3.7) devient

$$E_u(L(\tilde{u} - u)) \propto e^{a\tilde{u}} E_u(e^{-au}) - a(\tilde{u} - E_u(u)) - 1, \quad (2.3.8)$$

où $E_u(\cdot)$ représente l'espérance a posteriori relative à la densité a posteriori de u . L'estimateur de Bayes \tilde{u}_L de \tilde{u} sous la fonction de perte Linex est la valeur de \tilde{u} qui minimise (2.3.8). Pour trouver l'estimateur, nous dérivons l'équation (2.3.8) par rapport à \tilde{u} et nous obtenons :

$$\frac{d}{d\tilde{u}}(E_u(L(\tilde{u} - u))) = ae^{-a\tilde{u}} E_u(e^{-au}) - a. \quad (2.3.9)$$

En égalant cette expression à 0, nous obtenons

$$ae^{-a\tilde{u}}E_u(e^{-au}) = a, \quad (2.3.10)$$

d'où

$$e^{-a\tilde{u}} = E_u(e^{-au}). \quad (2.3.11)$$

En appliquant le logarithme, nous trouvons

$$-a\tilde{u} = \log E_u(e^{-au}). \quad (2.3.12)$$

Alors, l'estimateur de Bayes \tilde{u}_L de \tilde{u} sous la fonction de perte Linex est

$$\tilde{u}_L = -\frac{1}{a} \log(E_u(e^{-au}))E_u(e^{-au}), \quad (2.3.13)$$

étant donné que $E_u(e^{-au})$ existe et est finie.

2.3.7. Fonction de perte normale reflétée

La fonction de perte exprime une perte à la société en terme d'unité monétaire. Ce concept sous-entend la notion que toute déviation de la valeur cible, implique une dégradation dans la performance du produit, et conséquemment, une perte. Spring (1993) en réponse aux critiques visant la fonction de perte quadratique, développa la fonction de perte normale reflétée, laquelle est basée sur la fonction de densité normale,

$$L_{RN}(y) = K_{MAX} \left\{ 1 - \exp \left[-\frac{1}{2} \left(\frac{y-T}{\gamma} \right)^2 \right] \right\}. \quad (2.3.14)$$

Cette fonction a trois paramètres, K_{MAX} , T et γ , où T est la cible. Pour simplification, Spring (1993) a suggéré que les paramètres K_{MAX} et γ soient définis comme suit : K_{MAX} est la perte maximale qui peut être atteinte et γ est la déviation de la valeur cible, quand la perte maximale est atteinte. Plus précisément, $T \pm \Delta$ est un intervalle, où la perte représente 99,97% de la perte maximale. Si $\gamma = \Delta/4$, le résultat suivant sera la fonction normale reflétée de Spring (1993)

$$L_{SRN}(y) = K_{MAX} \left\{ 1 - \exp \left[-8 \left(\frac{y-T}{\Delta} \right)^2 \right] \right\}. \quad (2.3.15)$$

Nous notons que K_{MAX} et γ ont des significations différentes dans le modèle de la fonction de perte normale reflétée (RN) qu'ils ont dans la fonction de perte

quadratique. (Quand on parle de la fonction de perte quadratique, on parle de la forme reparamétrisée de la fonction de perte quadratique, donnée par Taguchi (1989)

$$L_Q(y) = \frac{K_\Delta}{\Delta^2}(y - T)^2,$$

où le paramètre K_Δ représente la perte correspondante à une déviation spécifiée, Δ , par rapport à la cible, T).

2.3.8. Fonction de perte normale reflétée modifiée

Spring (1996) définit la fonction de perte normale reflétée modifiée (MRN) comme suit :

$$L_{MRN}(y) = \frac{K_\Delta}{1 - \exp[-\frac{1}{2}(\frac{\Delta}{\gamma})^2]} \left\{ 1 - \exp \left[-\frac{1}{2} \left(\frac{y - T}{\gamma} \right)^2 \right] \right\}, \quad (2.3.16)$$

où les paramètres sont K_Δ , T , Δ . Notons que la nouvelle constante, K_Δ , n'est pas la perte maximale atteinte. K_Δ est définie comme étant la valeur de la perte à une distance spécifique Δ de la cible et T est la cible.

La perte maximale et les paramètres de forme K et γ déterminent de façon complète la fonction de perte normale reflétée. La forme générale de la courbe est un simple reflet d'une fonction de densité normale (sur l'axe horizontal). Plutôt que continuer d'augmenter de manière quadratique, la courbe se stabilise à la perte maximale. Le paramètre de la perte maximale K permet à la perte maximale finie d'être analysée. Le paramètre de forme, ainsi que la perte maximale permet la modification de la fonction de perte en reflétant différentes pertes qui apparaissent dans des processus différents. Pour les processus avec de petites décroissances dans la qualité pour une large région autour de la cible, le paramètre de forme sera plus grand que pour les processus où se produisent de grands changements dans la perte pour de petites déviations de la cible. La fonction de perte normale reflétée fournit une représentation alternative de la perte associée à la caractéristique de qualité.

La figure 2.1 nous montre l'évolution de la fonction de perte normale reflétée modifiée (MRN) en fonction du paramètre γ .

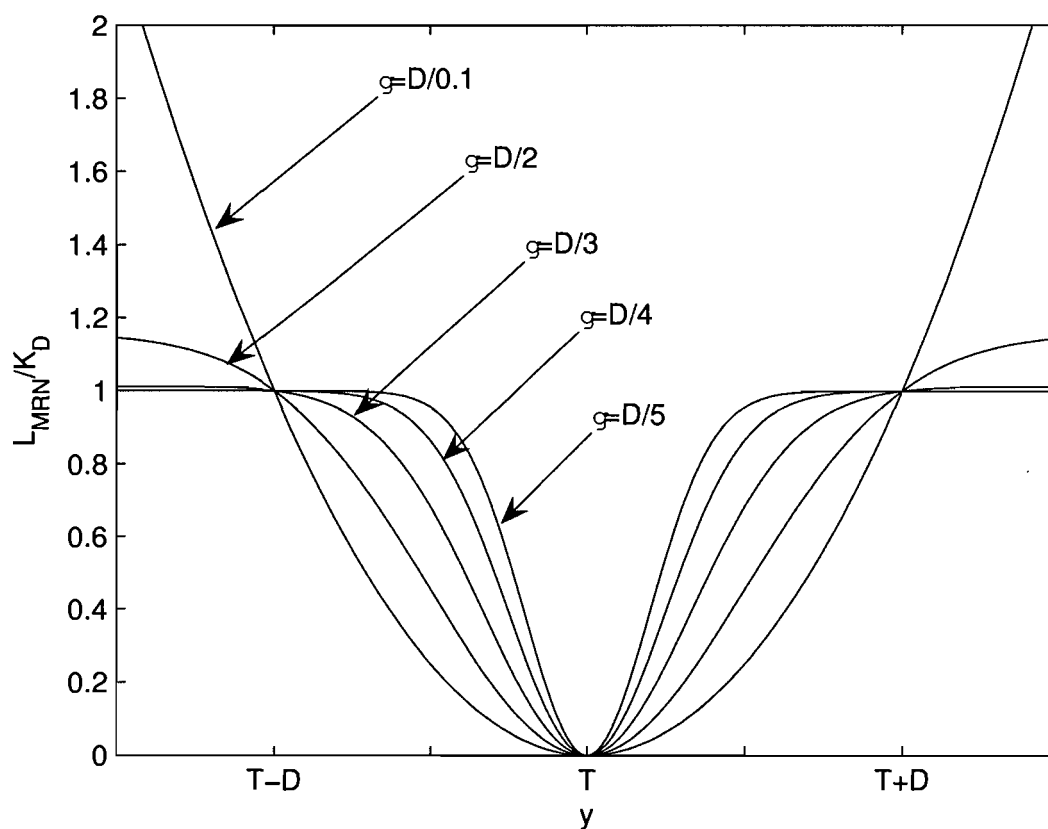


FIGURE 2.1. La fonction de perte normale reflétée modifiée (MRN)

La figure 2.2 nous montre les trois fonctions de perte, la fonction de perte normale reflétée modifiée (MRN), la fonction de perte normale reflétée de Spring (SRN) et la fonction quadratique (Q).

2.3.9. Alternatives

Il arrive parfois que le décideur, malgré une analyse approfondie du problème, soit incapable de déterminer exactement la fonction de perte. Par exemple, si le décideur est un comité formé de plusieurs experts qui arrivent à des fonctions de perte différentes, ils peuvent par la même occasion obtenir des distributions a priori différentes. Lorsque la perte ne peut pas être entièrement déterminée, nous pouvons proposer des classes de pertes paramétrisées, le décideur n'ayant plus alors qu'à choisir les paramètres en fonction de ses impératifs, en prenant par exemple :

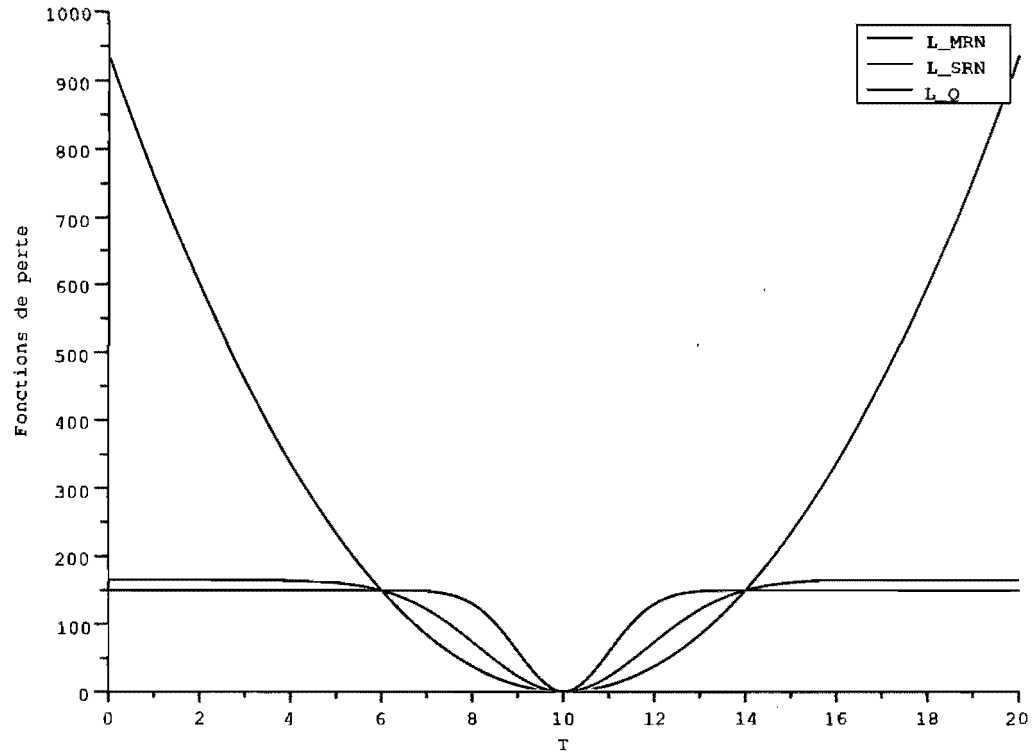


FIGURE 2.2. Les fonctions de perte normale (MRN, SRN, Q)

$$L_1(\theta, \delta) = \log[\alpha(\theta - \delta)^2 + 1],$$

$$L_2(\theta, \delta) = 1 - \exp[-c(\theta - \delta)^2].$$

Si la perte est incertaine, nous pouvons la représenter par la variable aléatoire $L(\theta, a)$, et intégrer cette dimension supplémentaire; si F est la distribution de perte, nous minimisons

$$\int_{\Theta} \int_{\Omega} L(\theta, \delta(x)) dF(x) d\pi(\theta),$$

F dépendant éventuellement de θ ou des observations x . C'est le seul cas véritablement intéressant car, sinon, la minimisation revient à utiliser la perte moyenne, $\int_{\Omega} L(\theta, \delta, \omega) dF(\omega)$. Nous pouvons aussi considérer un ensemble des pertes et chercher un estimateur performant pour toutes ces pertes. Évidemment, nous perdons l'unidimensionnalité du critère et donc l'unicité du meilleur estimateur de Bayes.

2.4. CONCLUSION

Dans ce chapitre nous avons présenté les principes de base de la théorie de la décision. Ainsi, nous avons introduit le concept de fonction d'utilité et de fonction de perte. Nous avons présenté l'approche bayésienne du calcul des estimateurs et aussi, les critères d'optimalité pour ces estimateurs. Ensuite nous avons procédé à une analyse de la fonction de perte, en présentant les fonctions suivantes : la fonction de perte quadratique, la fonction de perte erreur absolue, la perte 0-1, la perte intrinsèque, les fonctions de perte asymétriques, la fonction de perte Linex, la fonction de perte normale reflétée, ainsi que la fonction de perte normale reflétée modifiée. Évidemment, cette présentation n'est pas exhaustive, mais elle met en place l'axiomatisation du cadre inférentiel statistique, appelée « Théorie de la décision ».

Chapitre 3

ESTIMATEURS DE BAYES ET APPROXIMATIFS APPLIQUÉS AU MODÈLE DE PARETO

3.1. INTRODUCTION

Dans ce chapitre, nous présentons trois fonctions de perte et leurs estimateurs bayésiens associés, reliés à la distribution de Pareto. La distribution de Pareto joue un rôle important dans les études socio-économiques. Elle est souvent utilisée comme modèle dans les domaines d'analyse incluant la distribution de la population dans une ville, la fluctuation des prix du marché, et l'emplacement des champs de pétrole. De plus, la distribution de Pareto a trouvé des applications dans le domaine militaire. Les estimateurs et les propriétés des distributions de Pareto ont fait l'objet de discussions de la part de plusieurs auteurs. Il y a entre autres, Malik (1970), Arnold et Press (1983), Tiwari (1996), Pandey (1996), Abdel-Ghaly (1998) et Hossain (2000).

La fonction de densité de la distribution de Pareto de paramètres (α, β) , ($\alpha > 0, \beta > 0$) est donnée par :

$$f(x) = \alpha\beta^\alpha(x + \beta)^{-(\alpha+1)}, \quad x > 0 \quad (3.1.1)$$

La fonction de fiabilité est définie pour tout $t > 0$ par $R(t) = 1 - F(t)$, où $F(t) = P(T \leq t)$ représente la probabilité qu'un dispositif choisi au hasard dans la population ait une défaillance avant l'instant t . La fonction $R(t)$ représente la probabilité qu'un dispositif choisi au hasard dans la population n'ait pas de

défaillance avant l'instant t . La fonction de fiabilité de la distribution de Pareto est égale à

$$R(t) = \beta^\alpha (t + \beta)^{-\alpha}, \quad t > 0 \quad (3.1.2)$$

La fonction de taux d'échec est définie pour tout $t > 0$ par la densité conditionnelle à t étant donné la survie jusqu'à t , c'est-à-dire $H(t) = f(t)/R(t)$ et elle permet la détermination du nombre de défaillances (échecs) arrivant par unité de temps. La fonction de taux d'échec de la distribution Pareto est égale à

$$H(t) = \alpha(t + \beta)^{-1}, \quad t > 0 \quad (3.1.3)$$

où α est le paramètre de forme et β est le paramètre d'échelle. La fonction de densité Pareto (3.1.1) peut être obtenue à partir d'un mélange de la distribution exponentielle de paramètre θ , où le paramètre d'échelle θ est distribué selon une densité gamma de paramètres (α, β) .

Pour la plupart des statisticiens, intéressés principalement à contrôler le taux de variabilité, il est devenu standard d'utiliser la fonction de perte quadratique, qui est symétrique et donne la même importance aux pertes dues à la surestimation et la sous-estimation ayant les mêmes grandeurs. Cependant, cette restriction n'est pas toujours pratique. Par exemple, pour l'estimation des fonctions de fiabilité et de taux d'échec, une surestimation est très souvent plus grave qu'une sous-estimation. Dans ce cas, l'utilisation de fonctions de perte symétriques n'est pas très appropriée (voir Basu (1991)). De plus, pour l'estimation des paramètres des distributions de Pareto, la surestimation peut ne pas avoir la même importance que la sous-estimation, parce que l'inégalité du gain est estimée avec des paramètres de forme (voir Pandey (1996)). Ceci nous amène à penser que les fonctions de pertes asymétriques sont plus appropriées.

3.2. ESTIMATION AVEC LE CRITÈRE DU MAXIMUM DE VRAISSEMBLANCE

Supposons qu'un ensemble de n éléments aient des durées de vies indépendantes qui suivent une distribution de Pareto de paramètres (α, β) avec une densité (3.1.1). Les temps d'échec sont enregistrés quand ils arrivent jusqu'à ce qu'un

nombre fixé r d'éléments aient échoué. Soit $\underline{x} = (x_1, x_2, \dots, x_r)$, où x_i est le temps du $i^{\text{ème}}$ élément qui échoue. Puisque les $(n-r)$ éléments restants n'ont pas échoué encore et donc ont des temps de vies plus grands que x_r (le temps du $r^{\text{ème}}$ élément qui échoue), la fonction de vraisemblance peut être écrite sous la forme :

$$L(\alpha, \beta; \underline{x}) = \frac{n!}{(n-r)!} \prod_{i=1}^r f(x_i) \times (1 - F(x_r))^{(n-r)} \quad (3.2.1)$$

$$\begin{aligned} &= \frac{n!}{(n-r)!} \alpha^r \beta^{r\alpha} \prod_{i=1}^r (x_i + \beta)^{-(\alpha+1)} \left(\frac{\beta}{x_r + \beta} \right)^{(n-r)\alpha} \\ &= \frac{n!}{(n-r)!} \alpha^r \beta^{n\alpha} \prod_{i=1}^n (x_i + \beta)^{-(\alpha+1)} (x_r + \beta)^{-(n-r)\alpha}. \end{aligned}$$

$$L(\alpha, \beta; \underline{x}) \propto \alpha^r \times \beta^{n\alpha} e^{-T} \quad (3.2.2)$$

où $T = \alpha(n-r) \log(x_r + \beta) + (\alpha+1) \sum_{i=1}^r \log(x_i + \beta)$ et \propto est le symbole de proportionnalité. Le logarithme de la fonction de vraisemblance, la fonction de logvraisemblance, est donné par :

$$\ell(\alpha, \beta) = \log L(\alpha, \beta) = r \log \alpha + \alpha n \times \log \beta - T \quad (3.2.3)$$

En dérivant l'équation (3.2.3) par rapport à α , on obtient l'expression suivante :

$$\frac{\partial \ell(\alpha, \beta)}{\partial \alpha} = r/\alpha + n \times \log \beta - (n-r) \times \log(x_r + \beta) - \sum_{i=1}^r \log(x_i + \beta). \quad (3.2.4)$$

En égalant la dérivée à 0, on obtient :

$$r/\alpha - (n-r) \log(x_r + \beta) + n \log \beta - \sum_{i=1}^r \log(x_i + \beta) = 0. \quad (3.2.5)$$

En supposant que β est connu et égal à β_0 , l'estimateur du maximum de vraisemblance (MLE) de α , $\hat{\alpha}$, peut être calculé en utilisant l'équation

$$r/\alpha = T_1 - n \log \beta_0, \quad (3.2.6)$$

$$\text{d'où} \quad \hat{\alpha} = r/(T_1 - n \log \beta_0) \quad (3.2.7)$$

avec $T_1 = (n-r) \log(x_r + \beta_0) + \sum_{i=1}^r \log(x_i + \beta_0)$.

Si les deux paramètres α et β sont inconnus, leurs estimateurs du maximum de vraisemblance (MLE), $\hat{\alpha}$ et $\hat{\beta}$, sont obtenus en prenant l'expression (3.2.4) et en calculant les dérivées partielles de cette expression, par rapport à α et β .

Nous obtenons les équations suivantes :

$$\frac{\partial \ell}{\partial \alpha} = r/\alpha + n \log \beta - T_1 = 0 \quad (3.2.8)$$

$$\frac{\partial \ell}{\partial \beta} = \alpha n/\beta - T_2 = 0, \quad (3.2.9)$$

où $T_2 = \alpha(n-r)(x_r + \beta)^{-1} + (\alpha+1) \sum_{i=1}^r (x_i + \beta)^{-1}$. Nous allons éliminer α entre les deux équations de (3.2.8) et (3.2.9), en l'isolant dans la première équation.

Nous obtenons

$$\alpha = r/(T_1 - n \log \beta), \quad (3.2.10)$$

où $T_1 = (n-r) \log(x_r + \beta) + \sum_{i=1}^r \log(x_i + \beta)$.

Nous allons remplacer la valeur de α dans l'équation 3.2.9 avec la valeur obtenue en 3.2.10. Après simplification, nous obtenons

$$\frac{nr}{\beta T_4} - T_3 = 0 \quad (3.2.11)$$

où $T_3 = r(n - r/T_4)(x_r + \beta) + (r/T_4 + 1) \sum_{i=1}^r (x_i + \beta)^{-1}$

et $T_4 = T_1 - n \log \beta$,

qui peut être résolu en utilisant par exemple la méthode d'itération de Newton-Raphson.

La méthode de Newton-Raphson (1685), est un algorithme efficace pour trouver des approximations d'un zéro (ou racine) d'une fonction d'une variable réelle à valeurs réelles. L'algorithme consiste à linéariser une fonction f en un point et à prendre le point d'annulation de cette linéarisation comme approximation du zéro recherché. On réitère cette procédure avec l'approximation obtenue.

Appliqué à la dérivée d'une fonction, cet algorithme permet d'obtenir une évaluation des points critiques. La méthode de Newton se généralise en dimension supérieure. La raison réside en une utilisation du théorème du point fixe, qui cependant n'est pas nécessaire pour comprendre le sens du résultat.

Bien que la méthode soit très efficace, certains aspects pratiques doivent être pris en compte. Avant tout, la méthode de Newton nécessite que la dérivée soit

effectivement calculée. Dans les cas où la dérivée est seulement estimée en prenant la pente entre deux points de la fonction, la méthode prend le nom de méthode de la sécante, moins efficace (d'ordre 1,618 qui est le nombre d'or) et inférieure à d'autres algorithmes. Par ailleurs, si la valeur de départ est trop éloignée du vrai zéro, la méthode de Newton peut entrer en une boucle infinie sans produire d'approximation améliorée.

En solutionnant l'équation (3.2.10), nous trouvons l'estimateur du maximum de vraisemblance $\hat{\beta}$ de β .

Avec $\hat{\beta}$ maintenant connu, α est calculé dans l'équation suivante

$$\hat{\alpha} = r / \left\{ (n - r) \log(x_r + \hat{\beta}) - n \log \hat{\beta} \right\}. \quad (3.2.12)$$

Les estimateurs du maximum de vraisemblance (MLE) de $R(t)$ et $H(t)$ sont donnés respectivement par les équations (3.1.2) et (3.1.3) en remplaçant α et β par $\hat{\alpha}$ et $\hat{\beta}$. Ainsi on obtient :

$$\hat{R}(t) = \hat{\beta}^{\hat{\alpha}} (t + \hat{\beta})^{-\hat{\alpha}}, \quad t > 0 \quad (3.2.13)$$

$$\hat{H}(t) = \hat{\alpha} (t + \hat{\beta})^{-1}, \quad t > 0. \quad (3.2.14)$$

3.3. ESTIMATEURS DE BAYES AVEC β CONNU

Si β est connu, supposons que $\gamma(c, d)$ est la loi conjuguée a priori de α ,

$$\pi(\alpha) = \alpha^{c-1} d^c e^{-d\alpha} / \Gamma(c), \quad \alpha, c, d > 0. \quad (3.3.1)$$

(Il est nécessaire le plus souvent de faire un choix arbitraire de la loi a priori. En particulier, on utilise des lois usuelles (normale, gamma, bêta, etc.) ou bien, l'approche a priori conjuguée, qui nécessite une quantité limitée d'information. On dit qu'une loi a priori $\pi(\theta)$ est conjuguée si $f(x|\theta)$ et $\pi(\theta)$ appartiennent à la même famille de lois).

Nous connaissons du chapitre 1 la formule pour calculer la densité a posteriori

$$\pi^*(\alpha|\underline{x}) = \frac{\prod_{i=1}^n f(x_i|\alpha)\pi(\alpha)}{\int_0^\infty \prod_{i=1}^n f(x_i|\alpha)\pi(\alpha)d\alpha}.$$

En remplaçant la fonction de vraisemblance (3.2.2) et la densité a priori (3.3.1) dans la formule ci-haut, on obtient

$$\pi^*(\alpha|\underline{x}) = \frac{\alpha^r \beta^{nr} e^{-T} \alpha^{c-1} d^c e^{-d\alpha} / \Gamma(c)}{\int_0^\infty \alpha^r \beta^{nr} e^{-T} \alpha^{c-1} d^c e^{-d\alpha} / \Gamma(c) d\alpha},$$

où $T = \alpha(n-r) \log(x_r + \beta) + (\alpha+1) \sum_{i=1}^r \log(x_i + \beta)$.

Nous allons évaluer séparément l'intégrale

$$\begin{aligned} & \int_0^\infty \frac{\alpha^r \beta^{nr} e^{-T} \alpha^{c-1} d^c e^{-d\alpha} d\alpha}{\Gamma(c)} \\ &= \frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)}}{\Gamma(c)} \int_0^\infty \alpha^{r+c-1} e^{n\alpha \log \beta - (n-r)\alpha \log(x_r + \beta) - \alpha \sum_{i=1}^r \log(x_i + \beta) - d\alpha} d\alpha. \end{aligned}$$

Nous avons transformé

$$\beta^{n\alpha} = e^{\log(\beta)^{n\alpha}} = e^{n\alpha \log(\beta)}.$$

En utilisant la notation suivante,

$$w = (n-r) \log(x_r + \beta) + \sum_{i=1}^r \log(x_i + \beta) - n \log \beta,$$

l'intégrale devient

$$\frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)}}{\Gamma(c)} \int \alpha^{r+c-1} e^{-\alpha(w+d)} d\alpha$$

Nous allons faire le changement de variable suivant

$$\alpha(w+d) = y.$$

En dérivant, nous trouvons $dy = (w+d)d\alpha$.

L'intégrale deviendra

$$\begin{aligned} & \frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)}}{\Gamma(c)} \int_0^\infty y / (w+d)^{r+c-1} e^{-y} \frac{dy}{w+d} \\ &= \frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)} \Gamma(r+c)}{\Gamma(c)(w+d)^{r+c}}. \end{aligned}$$

En remplaçant le résultat de l'intégrale dans la formule initiale de la densité a posteriori, nous obtenons la densité a posteriori de α

$$\pi^*(\alpha|\underline{x}) = (w+d)^{r+c} \alpha^{r+c-1} e^{-\alpha(w+d)} / \Gamma(r+c), \quad (3.3.2)$$

où $w = (n-r) \log(x_r + \beta) + \sum_{i=1}^r \log(x_i + \beta) - n \log \beta$.

Nous calculons l'estimateur de Bayes avec la perte quadratique.

Cet estimateur $\tilde{\alpha}_Q$ est la moyenne a posteriori donnée par

$$\delta^\pi = E^\pi[\theta|x] = \frac{\int_\theta \theta f(x|\theta)\pi(\theta)d\theta}{\int_\theta f(x|\theta)\pi(\theta)d\theta}.$$

D'abord, nous calculons l'intégrale qui se trouve au numérateur

$$\int \frac{\alpha \alpha^r \beta n \alpha e^{-T} \alpha^{c-1} d^c e^{-d\alpha}}{\Gamma(c)} d\alpha,$$

où $T = \alpha(n-r) \log(x_r + \beta) + (\alpha+1) \sum_{i=1}^r \log(x_i + \beta)$.

En faisant les calculs et en notant

$$w = (n-r) \log(x_r + \beta) + \sum_{i=1}^r \log(x_i + \beta) - n \log \beta,$$

l'intégrale du numérateur est égale à

$$\begin{aligned} & \frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)}}{\Gamma(c)} \int \alpha^{r+c} e^{-\alpha(n-r) \log(x_r + \beta) - \alpha \sum_{i=1}^r \log(x_i + \beta) - n \log \beta - \alpha d} d\alpha \\ &= \frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)}}{\Gamma(c)} \int \alpha^{r+c} e^{w+d} d\alpha. \end{aligned}$$

En effectuant le changement de variable

$$y = \alpha(w+d),$$

et en calculant sa dérivée,

$$dy = (w+d)d\alpha,$$

l'intégrale devient

$$\begin{aligned} & \frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)}}{\Gamma(c)} \int \left(\frac{y}{w+d}\right)^{r+c} \frac{e^y}{w+d} d\alpha \\ &= \frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)}}{\Gamma(c)(w+d)^{r+c+1}} \int y^{r+c} e^y dy \\ &= \frac{d^c e^{-\sum_{i=1}^r \log(x_i + \beta)} \Gamma(r+c+1)}{\Gamma(c)(w+d)^{r+c+1}}. \end{aligned}$$

L'intégrale du dénominateur a déjà été calculée (dans le cadre du calcul de la densité a posteriori).

En remplaçant les deux résultats et après les simplifications, nous trouvons

$$\tilde{\alpha}_Q = \frac{\Gamma(r+c+1)}{(w+d)\Gamma(r+c)} = \frac{(r+c)\Gamma(r+c)}{(w+d)\Gamma(r+c)}.$$

Avec la perte quadratique, l'estimateur de Bayes $\tilde{\alpha}_Q$ de α est la moyenne a posteriori donnée par

$$\tilde{\alpha}_Q = (r+c)/(w+d). \quad (3.3.3)$$

Sous la fonction de perte de Linex (2.3.7), où $\Delta = \tilde{\alpha} - \alpha$, nous calculons l'estimateur de Bayes $\tilde{\alpha}_L$ de α en utilisant (2.3.13).

Ainsi nous devons déterminer la valeur de $E_u(e^{-au})$.

$$\begin{aligned} E_u(e^{-au}) &= \int_0^\infty e^{-a\alpha} \pi^*(\alpha|\underline{x}) d\alpha \\ &= \frac{1}{\Gamma(r+c)} \int_0^\infty e^{-a\alpha} (w+d)^{r+c} \alpha^{r+c-1} e^{-\alpha(w+d)} d\alpha \\ &= \frac{(w+d)^{r+c}}{\Gamma(r+c)} \int_0^\infty e^{-\alpha(a+w+d)} \alpha^{r+c-1} d\alpha. \end{aligned}$$

Nous effectuons le changement de variable

$$y = \alpha(a+w+d),$$

et nous calculons sa dérivée

$$dy = (a+w+d)d\alpha.$$

L'expression devient

$$\begin{aligned} &\frac{(w+d)^{r+c}}{\Gamma(r+c)} \int_0^\infty e^{-y} \left(\frac{y}{a+w+d}\right)^{r+c-1} \frac{dy}{a+w+d} \\ &= \frac{(w+d)^{r+c}}{\Gamma(r+c)} \times \frac{1}{(a+w+d)^{r+c-1}} \times \frac{1}{a+w+d} \int_0^\infty e^{-y} y^{r+c-1} dy \\ &= \frac{(w+d)^{r+c}}{\Gamma(r+c)} \frac{\Gamma(r+c)}{(a+w+d)^{r+c}} \end{aligned}$$

$$= \left(\frac{a + w + d}{w + d} \right)^{-r+c} = \left(1 + \frac{a}{w + d} \right)^{-r+c}.$$

Alors, l'estimateur de Bayes $\tilde{\alpha}_L$ de α est donné par

$$\tilde{\alpha}_L = [(r + c)/a] \log[1 + a/(w + d)]. \quad (3.3.4)$$

Estimation de $R(t)$ et $H(t)$:

Nous savons que la fonction de fiabilité se calcule à l'aide de la formule (3.1.2)

$$R = \left(\frac{\beta}{t + \beta} \right)^\alpha.$$

En appliquant le logarithme, nous obtenons

$$\log(R) = \alpha[\log(t + \beta) - \log \beta]^{-1}.$$

En divisant par

$$[\log(t + \beta) - \log \beta]^{-1},$$

nous obtenons

$$\alpha = [\log(t + \beta) - \log \beta] \log R(t).$$

En posant $k = [\log(t + \beta) - \log \beta]^{-1}$, l'expression devient $\alpha = \frac{1}{k} \log R(t)$.

En remplaçant $\alpha = \frac{1}{k} \log R(t)$ dans l'équation (2.3.13), nous obtenons la densité a posteriori de la fonction de fiabilité $R = R(t)$

$$\pi^*(R|\underline{x}) = \frac{[(w + d)k]^{r+c}}{\Gamma(r + c)} (-\log R)^{r+c-1} R^{k(w+d)-1}, \quad 0 < R < 1. \quad (3.3.5)$$

L'estimateur de Bayes de R sous la fonction de perte de l'erreur quadratique est obtenu de la même façon que l'estimateur $\tilde{\alpha}_Q$ de α ,

$$\tilde{R}_Q = [1 + 1/(w + d)k]^{-r(r+c)}. \quad (3.3.6)$$

Sous la fonction de perte Linex, l'estimateur de Bayes de la fonction de fiabilité R utilisant la fonction (2.3.13) est

$$\tilde{R}_L = -\frac{1}{a} \log \left[\sum_{j=1}^{\infty} \frac{(-a)^j}{j!} \left(1 + \frac{j}{(w + d)k} \right)^{-(r+c)} \right]. \quad (3.3.7)$$

Nous savons que la fonction du taux d'échec est donnée par la formule

$$H(t) = \alpha(t + \beta)^{-1}.$$

En substituant $\alpha = k^*H(t)$ dans l'équation (3.3.1), où $k^* = (t + \beta)$, nous obtenons la densité a posteriori de la fonction du taux d'échec $H = H(t)$

$$\pi^*(H|\underline{x}) = \frac{[(w+d)k^*]^{r+c}}{\Gamma(r+c)} H^{r+c-1} e^{-H(w+d)k^*}, \quad H > 0. \quad (3.3.8)$$

Sous la fonction de perte quadratique, l'estimateur de Bayes de la fonction du taux d'échec H est obtenu en substituant $\alpha = k^*H(t)$ dans l'équation (3.3.5), où $k^* = (t + \beta)$,

$$\tilde{H}_Q = (r+c)/(w+d)k^*. \quad (3.3.9)$$

Sous la fonction de perte Linex (2.3.7), l'estimateur de Bayes de H est :

$$\tilde{H}_L = [(r+c)/a] \log [1 + a/((w+d)k^*)]. \quad (3.3.10)$$

3.4. ESTIMATION DE BAYES AVEC α ET β INCONNUS

Sous l'hypothèse que les deux paramètres α et β sont inconnus, il n'y a pas de simplification possible en utilisant le principe de suffisance pour la vraisemblance correspondant à un échantillon de la densité de Pareto (α, β) (3.1.1). On considère donc la densité a priori conjointe (Gamma-Exponentielle) comme le produit de la densité conditionnelle de α , étant donné β (ayant comme valeur le conjugué de la loi Gamma a priori utilisée par Arnold et Press (1983)), et une exponentielle à un seul paramètre pour β .

Soit le conjugué de la loi Gamma a priori utilisée par Arnold et Press (1983), de la forme suivante

$$g_1(\alpha|\beta) = \beta^{-\zeta} \alpha^{\zeta-1} e^{-\alpha/\beta} / \Gamma(\zeta), \quad \alpha, \beta, \zeta > 0 \quad (3.4.1)$$

et la loi exponentielle à un seul paramètre pour β , de la forme suivante

$$g_2(\beta) = (1/\delta) e^{-\beta/\delta}. \quad (3.4.2)$$

La densité conjointe de α et β est donnée par

$$\pi(\alpha, \beta) = g_1(\alpha|\beta)g_2(\beta) = A^* \beta^{-\zeta} \alpha^{\zeta-1} \exp\{-[(\alpha/\beta) + (\beta/\delta)]\}, \quad (3.4.3)$$

où $A^* = [\delta\Gamma(\zeta)]^{-1}$.

En combinant la fonction de vraisemblance (3.2.2) et la fonction de densité a

priori (3.4.3), la densité conjointe a posteriori de α et β est

$$\pi^*(\alpha, \beta) \propto \alpha^{r+\zeta-1} \beta^{\alpha n-\zeta} \exp\{-[(\alpha/\beta) + (\beta/\delta) + T]\}, \quad \alpha > 0, \beta > 0. \quad (3.4.4)$$

3.4.1. Fonction de perte quadratique

L'estimateur de Bayes d'une fonction $U = U(\alpha, \beta)$ de paramètres inconnus α et β , sous la contrainte de la perte quadratique est la moyenne a posteriori

$$\tilde{U}_Q = E(U|\underline{x}) = \frac{\int \int U(\alpha, \beta) \pi^*(\alpha, \beta) d\alpha d\beta}{\int \int \pi^*(\alpha, \beta) d\alpha d\beta}, \quad (3.4.5)$$

où $\pi^*(\alpha, \beta)$ est la densité conjointe a posteriori de α et β . Le rapport d'intégrales dans l'équation (3.4.4) n'a pas de forme plus compacte et nous devons alors considérer la méthode d'approximation suivante.

La procédure de Lindley : la procédure de Lindley a été utilisée par plusieurs auteurs pour obtenir des estimateurs de Bayes des paramètres de quelques distributions. On peut citer par exemple Sinha (1985), Sinha et Sloan (1988), Howlader et Weiss (1988, 1989) et Al-Hussaini et Jaheen (1992, 1994).

Lindley (1980) a développé une méthode d'approximation pour l'évaluation de l'espérance de $U(\theta)$ par

$$E(U(\theta)|\underline{x}) = \frac{\int U(\theta) e^{\ell(\theta)+\rho(\theta)} d\theta}{\int e^{\ell(\theta)+\rho(\theta)} d\theta}. \quad (3.4.6)$$

qui n'est autre que l'estimation de Bayes de $U(\theta)$ sous une fonction de perte quadratique, où $\rho(\theta) = \log(p(\theta))$, $p(\theta)$ est une fonction arbitraire de θ , et $\ell(\theta)$ est le logarithme de la fonction de vraisemblance. Pour plus de détails, voir Lindley (1980).

Dans le cas de deux paramètres, $\theta = (\theta_1, \theta_2)$, l'approximation de Lindley se simplifie et s'écrit sous la formule suivante :

$$\begin{aligned} E(U(\theta)|\underline{x}) &= U(\theta) + (A/2) + \rho_1 A_{12} + \rho_2 A_{21} \\ &\quad + (1/2) [\ell_{30} B_{12} + \ell_{21} C_{12} + \ell_{21} C_{21} + \ell_{03} B_{21}], \end{aligned} \quad (3.4.7)$$

où $A = \sum_{i=1}^2 \sum_{j=1}^2 U_{ij} \sigma_{ij}$, $\ell_{\eta\epsilon} = (\partial^{\eta+\epsilon} \ell / \partial \theta_1^\eta \partial \theta_2^\epsilon)$, $\eta, \epsilon = 0, 1, 2, 3$, $\eta + \epsilon = 3$
 $\rho_i = (\partial U / \partial \theta_i)$, $U_i = (\partial U / \partial \theta_i)$, $U_{ij} = (\partial^2 U / \partial \theta_i \partial \theta_j)$, pour $i, j = 1, 2$,

et pour $i \neq j$,

$$\begin{aligned} A_{ij} &= U_i \sigma_{ii} + U_j \sigma_{ji}, & B_{ij} &= (U_i \sigma_{ii} + U_j \sigma_{ij}) \sigma_{ii}, \\ C_{ij} &= 3U_i \sigma_{ii} \sigma_{ij} + U_j (\sigma_{ii} \sigma_{jj} + 2\sigma_{ij}^2), \end{aligned}$$

où σ_{ij} est le $(i, j)^e$ élément de l'inverse de la matrice $\{-\ell_{ij}\}$, $i, j = 1, 2$, avec $\ell_{ij} = (\partial^2 \ell / \partial \theta_i \partial \theta_j)$. Toutes les fonctions dans l'équation (3.4.6) sont évaluées en utilisant les estimateurs du maximum de vraisemblance de (θ_1, θ_2) .

Dans notre cas, $(\theta_1, \theta_2) = (\alpha, \beta)$, $U(\theta) = U(\alpha, \beta)$; pour appliquer la formule de Lindley (3.4.6), nous obtenons d'abord les éléments σ_{ij} de l'inverse de $\{-\ell_{ij}\}$, $i, j = 1, 2$, donnés par

$$\sigma_{11} = M/N, \quad \sigma_{12} = \sigma_{21} = -D/N, \quad \sigma_{22} = r/\alpha^2 N,$$

où

$$\begin{aligned} M &= \alpha n / \beta^2 - \alpha(n-r)/(x_r + \beta)^2 - (\alpha + 1) \sum_{i=1}^r 1/(x_i + \beta)^2, \\ N &= (r/\alpha^2)M - D^2, \text{ et} \\ D &= (n-r)/(x_r + \beta) + \sum_{i=1}^r 1/(x_i + \beta) - n/\beta. \end{aligned}$$

À partir de la densité de l'équation (3.4.1), nous observons que

$$\rho(\alpha, \beta) = \log[\pi(\alpha, \beta)] = \text{constante} + (\zeta - 1) \log \alpha - \zeta \log \alpha - \zeta \log \beta - \alpha/\beta - \beta/\delta. \quad (3.4.8)$$

Donc,

$$\rho_1 = \partial \rho / \partial \alpha = [(\zeta - 1)/\alpha] - 1/\beta, \quad \rho_2 = \partial \rho / \partial \beta = \alpha/\beta^2 - \zeta/\beta - 1/\delta. \quad (3.4.9)$$

De plus,

$$\begin{aligned} \ell_{21} &= 0, & \ell_{12} &= b_1 - n/\beta, \\ \ell_{30} &= 2r/\alpha^3, & \ell_{03} &= 2(\alpha n/\beta^3 - b_2), \end{aligned}$$

$$\begin{aligned} \text{où } b_1 &= (n-r)/(x_r + \beta)^2 + \sum_{i=1}^r 1/(x_i + \beta)^2, \\ b_2 &= \alpha(n-r)/(x_r + \beta)^3 + (\alpha + 1) \sum_{i=1}^r 1/(x_i + \beta)^3. \end{aligned}$$

En substituant les valeurs ci-dessus dans l'équation (3.4.6), l'estimateur de Bayes obtenu en utilisant la méthode de Lindley relative à la fonction de perte quadratique d'une fonction $U(\alpha, \beta)$, notée par \tilde{U}_Q est

$$\tilde{U}_Q = E[U(\alpha, \beta)|\underline{x}] = U(\alpha, \beta) + \Phi + \Psi_1 U_1 + \Psi_2 U_2, \quad (3.4.10)$$

où $\Phi = (1/2N)[MU_{11} + D(U_{12} + U_{21}) + rU_{22}/\alpha^2]$,

$$\begin{aligned} \Psi_1 = (1/N)[M\rho_1 - D\rho_2] + (1/\alpha^2 N^2)[(b_1 - n/\beta)(rM/2 + \alpha^2 D^2) \\ + rM^2/\alpha - Dr(n\alpha/\beta^3 - b_2)], \end{aligned}$$

$$\begin{aligned} \Psi_2 = (1/\alpha^2 N)[r\rho_2 - \alpha^2 D\rho_1] + (r/\alpha^4 N^2)[r(n\alpha/\beta^3 - b_2) \\ - \alpha DM - (3/2)\alpha^2 D(b_1 - n/\beta)]. \end{aligned}$$

Toutes les fonctions à droite de l'équation (3.4.10) doivent être évaluées aux points $\hat{\alpha}$ et $\hat{\beta}$.

Cas spéciaux :

(i) En substituant $U(\alpha, \beta) = \alpha$ dans l'équation (3.4.10), on obtient l'estimateur de Bayes de α donné par :

$$\tilde{\alpha}_Q = \alpha + \Psi_1, \text{ évalué au point } (\hat{\alpha}, \hat{\beta}). \quad (3.4.11)$$

(ii) En substituant $U(\alpha, \beta) = \beta$, dans l'équation (3.4.10), on obtient l'estimateur de Bayes de β donné par :

$$\tilde{\beta}_Q = \beta + \Psi_2, \text{ évalué au point } (\hat{\alpha}, \hat{\beta}). \quad (3.4.12)$$

(iii) En substituant $U(\alpha, \beta) = R(t) = [\beta/(t + \beta)]^\alpha$ dans l'équation (3.4.10), on obtient l'estimateur de Bayes de $R = R(t)$ donné par

$$\tilde{R}_Q = R(t) \{1 + Q + (1/\alpha)[\log(R(t))]\Psi_1 + [\alpha t/(\beta^2 + \beta t)]\Psi_2\}, \quad (3.4.13)$$

évalué au point $(\hat{\alpha}, \hat{\beta})$, où

$$\begin{aligned} Q = (1/2\alpha^2 N) \{M \log(R(t)) + [2t\alpha^2 D/(\beta^2 + \beta t)][1 + \log R(t)] \\ + rt(\alpha t - t - 2\beta)/[\beta(\beta + t)]^2\}. \end{aligned} \quad (3.4.14)$$

(iv) En substituant $U(\alpha, \beta) = H(t) = \alpha/(t + \beta)^{-1}$ dans l'équation (3.4.10), on obtient l'estimateur de Bayes de $H = H(t)$ donné par

$$\tilde{H}_Q = H(t) [1 + Q_1 + (1/\alpha)\Psi_1 - \Psi_2/(t + \beta)], \quad (3.4.15)$$

évalué au point $(\hat{\alpha}, \hat{\beta})$, où $Q_1 = (1/2\alpha^2 N)[r/(t + \beta)^2 - \alpha D/(t + \beta)]$.

3.4.2. Fonction de perte Linex

L'estimateur de Bayes d'une fonction $U = U(\alpha, \beta)$ de paramètres inconnus α et β , sous la fonction de perte Linex (2.3.7) à partir de l'équation (2.3.13) est donné par

$$\tilde{U}_L = -\frac{1}{a} \log(\tilde{\Omega}_u), \quad (3.4.16)$$

où

$$\tilde{\Omega}_u = E(e^{-aU} | \underline{x}) = \frac{\int \int e^{-aU} \pi^*(\alpha, \beta) d\alpha d\beta}{\int \int \pi^*(\alpha, \beta) d\alpha d\beta}. \quad (3.4.17)$$

L'équation (3.4.16) est encore un rapport d'intégrales, nous pouvons donc appliquer la même technique (méthode de Lindley) que nous avons utilisée pour (3.4.6).

Cas particuliers :

(i) Si dans l'équation (3.4.10), $U(\alpha, \beta) = e^{-a\alpha}$, alors l'estimateur de Bayes de α relatif à la fonction de perte Linex est donné par

$$\begin{aligned} \tilde{\alpha}_Q &= -\frac{1}{a} \log [E(e^{-a\alpha})] \\ &= \alpha + \log[(1 + (a^2 M/2N) - a\Psi_1)], \end{aligned} \quad (3.4.18)$$

évalué au point $(\hat{\alpha}, \hat{\beta})$.

(ii) Si dans l'équation (3.4.10), $U(\alpha, \beta) = e^{-a\beta}$, alors l'estimateur de Bayes de β relatif à la fonction de perte Linex est donné par

$$\tilde{\beta}_L = \beta + \log[(1 + (ra^2/2\alpha^2 N) - a\Psi_2)], \quad (3.4.19)$$

évalué au point $(\hat{\alpha}, \hat{\beta})$.

(iii) Si dans l'équation (3.4.10), $U(\alpha, \beta) = e^{-aR(t)}$, alors l'estimateur de Bayes de β relatif à la fonction de perte Linex est donné par

$$\tilde{R}_L = R + \log[1 + (ah^\alpha Q_2/2N\beta^4) - ah^\alpha(\Psi_1 \log h + \alpha th\Psi_2/\beta^2)], \quad (3.4.20)$$

évalué au point $(\hat{\alpha}, \hat{\beta})$, où

$$Q_2 = M\beta^4(\alpha h^\alpha - 1)(\log h)^2 + 2th\beta D[\alpha \log(h(ah^\alpha - 1)) - 1] \\ + (rth/\alpha)[\alpha ath^{\alpha+1} - t(\alpha + 1)h + 2\beta], \text{ et}$$

$$h = \beta/(\beta + t).$$

(iv) Si dans l'équation (3.4.10), $U(\alpha, \beta) = e^{-aH(t)}$, alors l'estimateur de Bayes de $H = H(t)$ relatif à la fonction de perte Linex est donné par

$$\tilde{H}_L = H + \log[(1 + Q_3 + \alpha b_2 \Psi_2 - b_1 \Psi_1)], \quad (3.4.21)$$

évalué au point $(\hat{\alpha}, \hat{\beta})$, où

$$Q_3 = (1/2N)[aMb_2 + 2Db_2(1 - \alpha b_1) + rb_3(b_1 - 2/\alpha)], \text{ et} \\ b_i = a/(\beta + t)^i, \quad i = 1, 2, 3.$$

3.4.3. Fonction de perte normale reflétée modifiée

Nous introduisons la fonction de perte de Spring (1993) en réponse aux critiques visant la fonction de perte quadratique, fonction basée sur la fonction de densité normale.

$$L_{RN}(y) = K_{MAX} \left\{ 1 - \exp \left[-\frac{1}{2} \left(\frac{y - T}{\gamma} \right)^2 \right] \right\}. \quad (3.4.22)$$

Cette fonction a trois paramètres, K_{MAX} , T et γ , où T est la cible. Nous proposons une modification simple de la fonction de perte normale reflétée (RN), et obtiendrons la fonction de perte normale reflétée modifiée. Nous définissons la fonction de perte normale reflétée modifiée (MRN)

$$L_{MRN}(y) = \frac{K_\Delta}{1 - \exp[-\frac{1}{2}(\frac{\Delta}{\gamma})^2]} \left\{ 1 - \exp \left[-\frac{1}{2} \left(\frac{y - T}{\gamma} \right)^2 \right] \right\}, \quad (3.4.23)$$

où les paramètres sont K_Δ , T , Δ , et notons que la nouvelle constante, K_Δ , n'est pas la perte maximale atteinte.

L'estimateur de Bayes d'une fonction $U = U(\alpha, \beta)$ de paramètres inconnus α et β , sous la fonction de perte normale reflétée modifiée est encore un rapport d'intégrales; nous pouvons donc appliquer la même technique (méthode de

Lindley) que nous avons utilisée pour (3.4.2).

$$\tilde{U}_N = E(U|\underline{x}) = \frac{\int \int U(\alpha, \beta) \pi^*(\alpha, \beta) d\alpha d\beta}{\int \int \pi^*(\alpha, \beta) d\alpha d\beta}. \quad (3.4.24)$$

Spring (1993) a déduit la perte attendue sous la fonction de perte normale reflétée (RN)

$$E[L_{RN}(Y)] = K_{MAX} \left[1 - \frac{\gamma}{\sqrt{\sigma^2 + \gamma^2}} \exp\left(-\frac{(\mu - T)^2}{2(\sigma^2 + \gamma^2)}\right) \right]. \quad (3.4.25)$$

Le résultat correspondant pour la perte attendue sous la fonction de perte normale reflétée modifiée (MRN) est

$$E[L_{MRN}(Y)] = \frac{K_{\Delta}}{1 - \exp[-\frac{1}{2}(\Delta/\gamma)^2]} \left\{ 1 - \frac{\gamma}{\sqrt{\sigma^2 + \gamma^2}} \exp\left[-\frac{(\mu - T)^2}{2(\sigma^2 + \gamma^2)}\right] \right\}. \quad (3.4.26)$$

Une approximation simple de la perte attendue sous la fonction de perte normale reflétée modifiée pour n'importe quelle densité de probabilité peut être obtenue en termes des quatre premiers moments de $f(y)$, en utilisant le développement de Taylor pour $L_{MRN}(Y)$ (voir Sun (1996)).

$$E[L_{MRN}(Y)] \simeq \frac{K'}{8\gamma^2} [4\gamma^2 - (\mu - T)^2 + 22\gamma^2 - 3(\mu - T)^2\sigma^2 - 4(\mu - T)^2\alpha_3\sigma^3 - \alpha_4\sigma^4], \quad (3.4.27)$$

$$\text{où } K' = \frac{K_{\Delta}}{1 - \exp[-\frac{1}{2}(\Delta/\gamma)^2]},$$

$$\alpha_3 = E(Y - \mu)^3/\sigma^3$$

$$\text{et } \alpha_4 = E(Y - \mu)^4/\sigma^4.$$

Nous allons minimiser cette espérance dans le chapitre suivant (de façon numérique) et nous allons obtenir une valeur pour μ . En remplaçant cette valeur dans l'équation 3.4.10, nous obtiendrons l'estimateur de Bayes de α sous la fonction de perte normale reflétée modifiée.

Chapitre 4

COMPARAISONS NUMÉRIQUES

Au chapitre 3, nous avons présenté les différentes méthodes d'estimation des paramètres pour les différentes fonctions de perte proposées. Ainsi nous avons présenté la méthode des moments, la méthode du maximum de vraisemblance et la méthode du calcul des estimateurs de Bayes, pour la fonction de perte quadratique, la fonction de perte Linex, ainsi que pour la fonction de perte normale réflétée. Nous présenterons un exemple numérique pour illustrer les méthodes d'estimation présentées au chapitre 3 et enfin nous discuterons de ces méthodes.

Des comparaisons numériques sont effectuées à l'aide du logiciel MATLAB et les Tableaux 4.1 à 4.15 résument ces résultats.

4.1. ESTIMATEURS AVEC α ET β INCONNUS

D'abord nous allons analyser les estimateurs dans le cas où les paramètres α et β sont inconnus. (Tableaux 4.1 à 4.8).

Dans le tableau 4.1, nous avons mis les valeurs calculées des différents estimateurs pour le paramètre, α en utilisant pour les dix simulations des valeurs de $\alpha = 0.277$ et $\beta = 4.35$.

Dans la première colonne, nous avons la taille de l'échantillon n , dans la deuxième colonne, nous avons l'estimateur du maximum de vraisemblance pour α , $\hat{\alpha}$; dans la troisième colonne, nous avons l'écart-type pour α ; ensuite nous avons l'estimateur de Bayes sous la fonction de perte quadratique pour α , ($\tilde{\alpha}_Q$

suivi de son écart-type et de l'estimateur des moments pour α , $\bar{\alpha}$ et de son écart-type; dans les deux dernières colonnes, nous avons l'estimateur de Bayes sous la fonction LINEX ($a = -5$) pour α , $\tilde{\alpha}_L$, suivi de son écart type. (Chapitre 3)

Dans le tableau 4.2, nous avons mis les valeurs calculées des différents estimateurs pour le paramètre, β en utilisant pour les simulations des valeurs de $\beta = 4.35$.

Dans la première colonne, nous avons la taille de l'échantillon n , dans la deuxième colonne, nous avons l'estimateur du maximum de vraisemblance pour β , $\hat{\beta}$; dans la troisième colonne, nous avons l'écart-type pour β ; ensuite nous avons l'estimateur de Bayes sous la fonction de perte quadratique pour β , $(\tilde{\beta}_Q)$ suivi de son écart-type et de l'estimateur des moments pour β , $\bar{\beta}$ et de son écart-type; dans les deux dernières colonnes, nous avons l'estimateur de Bayes sous la fonction Linex ($a = -5$) pour β , $\tilde{\beta}_L$, suivi de son écart type. (Chapitre 3) Les estimés $\hat{\alpha}$ et $\hat{\beta}$ sont obtenus en résolvant le système d'équations suivant (Chapitre 2) :

$$\frac{\partial \ell}{\partial \alpha} = r/\alpha + n \log \beta - T_1 = 0; \quad (4.1.1)$$

$$\frac{\partial \ell}{\partial \beta} = \alpha n/\beta - T_2 = 0, \quad (4.1.2)$$

où $T_2 = \alpha(n-r)(x_r + \beta)^{-1} + (\alpha+1) \sum_{i=1}^r (x_i + \beta)^{-1}$

$T_1 = (n-r) \log(x_r + \beta) + \sum_{i=1}^r \log(x_i + \beta)$.

L'estimateur de Bayes sous la fonction de perte quadratique pour α et pour β , $(\tilde{\alpha}_Q, \tilde{\beta}_Q)$ se calcule en utilisant la procédure d'approximation de Lindley pour l'évaluation de l'espérance de $U(\alpha, \beta)$

$$\tilde{U}_Q = E(U|\underline{x}) = \frac{\int \int U(\alpha, \beta) \pi^*(\alpha, \beta) d\alpha d\beta}{\int \int \pi^*(\alpha, \beta) d\alpha d\beta}.$$

L'estimateur de Bayes sous la fonction Linex se calcule en utilisant la formule (2.3.13, Chapitre 2).

$$\tilde{u}_L = -\frac{1}{a} \log(E_u(e^{-au})) E_u(e^{-au}).$$

Dans le tableau 4.3, nous présentons les valeurs des estimateurs pour la fonction de fiabilité à $t = 1$ ainsi que les valeurs des écarts-types respectifs. Nous

TABLE 4.1. Estimateurs pour α .

n	$\hat{\alpha}$	$s_{\hat{\alpha}}$	$\tilde{\alpha}_Q$	$s_{\tilde{\alpha}_Q}$	$\bar{\alpha}$	$s_{\bar{\alpha}}$	$\tilde{\alpha}_L$	$s_{\tilde{\alpha}_L}$
10	0.255	0.066	0.223	0.083	2.011	0.089	0.200	0.063
20	0.218	0.065	0.211	0.097	2.030	0.056	0.211	0.065
200	0.167	0.056	0.210	0.060	2.091	0.029	0.179	0.044
250	0.174	0.064	0.179	0.053	2.066	0.049	0.190	0.045
300	0.152	0.058	0.210	0.056	2.087	0.051	0.160	0.051
350	0.160	0.049	0.167	0.041	2.033	0.050	0.162	0.049
400	0.244	0.043	0.299	0.039	2.049	0.061	0.223	0.055
450	0.164	0.044	0.166	0.037	2.027	0.045	0.159	0.047
500	0.139	0.035	0.136	0.036	2.022	0.044	0.137	0.030

TABLE 4.2. Estimateurs pour β .

n	$\hat{\beta}$	$s_{\hat{\beta}}$	$\tilde{\beta}_Q$	$s_{\tilde{\beta}_Q}$	$\bar{\beta}$	$s_{\bar{\beta}}$	$\tilde{\beta}_L$	$s_{\tilde{\beta}_L}$
10	3.988	0.311	3.543	0.321	7.890	0.501	3.733	0.233
20	4.103	0.314	4.339	0.331	8.056	0.432	3.567	0.221
200	4.254	0.325	5.223	0.290	8.987	0.503	3.182	0.241
250	5.433	0.206	6.639	0.213	9.811	0.489	3.009	0.207
300	3.654	0.242	4.205	0.211	7.897	0.476	2.473	0.187
350	3.669	0.219	4.179	0.198	8.233	0.527	2.953	0.156
400	3.251	0.189	4.476	0.187	8.879	8.953	3.457	0.151
450	4.198	0.175	4.996	0.174	9.025	9.301	2.532	0.143
500	4.365	0.157	4.541	0.157	8.011	8.117	3.005	0.135

calculons les valeurs des estimateurs en utilisant la formule de la fonction de fiabilité suivante (Chapitre 3)

$$R(t) = \beta^\alpha (t + \beta)^{-\alpha}. \quad (4.1.3)$$

La valeur théorique pour R est de 0.7184, pour $t = 1$, $\alpha = 0.277$ et $\beta = 4.35$. Dans la première colonne, nous avons l'estimateur du maximum de vraisemblance, suivi

de son écart-type, ensuite l'estimateur de Bayes sous la fonction de perte quadratique, suivi de son écart-type et dans les deux dernières nous avons l'estimateur de Bayes sous la fonction de perte LINEX ($a = -5$), suivi de son écart-type.

Nous utilisons les notations suivantes : Pour l'estimateur du maximum de vraisemblance pour α , $\hat{\alpha}$ et pour l'estimateur du maximum de vraisemblance pour β , $\hat{\beta}$; ensuite pour l'estimateur de Bayes sous la fonction de perte quadratique pour α et pour β , $(\tilde{\alpha}_Q, \tilde{\beta}_Q)$. L'estimateur des moments pour α et pour β sont notés $(\bar{\alpha}, \bar{\beta})$. L'estimateur de Bayes sous la fonction LINEX ($a = -5$) pour α est noté, $\tilde{\alpha}_L$, ainsi que l'estimateur de Bayes sous la fonction LINEX ($a = -5$) pour β , noté $\tilde{\beta}_L$. Pour les écart-types on utilise la notation suivante : Pour l'écart-type correspondant à l'estimateur du maximum de vraisemblance pour α , $s_{\hat{\alpha}}$ et pour l'écart-type correspondant à l'estimateur du maximum de vraisemblance pour β , $s_{\hat{\beta}}$. En effet, on utilise la notation s pour l'écart-type en précisant à l'indicateur le type d'estimateur.

TABLE 4.3. Estimateurs pour la fonction de fiabilité $R(1)$.

n	$\hat{R}(1)$	$s_{\hat{R}(1)}$	$\tilde{R}_Q(1)$	$s_{\tilde{R}_Q}$	$\tilde{R}_L(1)$	$s_{\tilde{R}_L}$
200	0.961	0.025	0.955	0.025	0.981	0.022
250	0.960	0.028	0.943	0.027	0.972	0.023
300	0.943	0.015	0.935	0.016	0.959	0.014
350	0.935	0.017	0.933	0.019	0.958	0.017
400	0.937	0.015	0.930	0.017	0.961	0.015
450	0.947	0.011	0.928	0.012	0.955	0.012
500	0.928	0.012	0.921	0.012	0.949	0.011

Dans le tableau 4.4, nous présentons les valeurs des estimateurs pour le taux d'échec. Nous calculons les valeurs des estimateurs en utilisant la formule suivante pour le taux d'échec (Chapitre 3), à $t = 1$,

$$H(t) = \alpha(t + \beta)^{-1}. \quad (4.1.4)$$

Dans la première colonne, nous avons l'estimateur du maximum du vraisemblance, ensuite l'estimateur de Bayes sous la fonction de perte quadratique et dans la dernière l'estimateur de Bayes sous la fonction de perte LINEX ($a = -5$).

TABLE 4.4. Estimateurs pour le taux d'échec H .

n	\hat{H}	$s_{\hat{H}}$	\tilde{H}_Q	$s_{\tilde{H}_Q}$	\tilde{H}_L	$s_{\tilde{H}_L}$
200	0.052	0.030	0.057	0.037	0.026	0.024
250	0.048	0.032	0.058	0.044	0.019	0.026
300	0.049	0.025	0.057	0.026	0.016	0.021
350	0.046	0.024	0.059	0.028	0.015	0.019
400	0.044	0.020	0.055	0.022	0.014	0.020
450	0.043	0.018	0.054	0.019	0.011	0.016
500	0.038	0.016	0.052	0.018	0.011	0.015

En se basant sur l'échantillon de taille n , nous avons calculé les estimateurs pour α et pour β . En remplaçant $(\hat{\alpha}, \hat{\beta})$, $(\tilde{\alpha}_Q, \tilde{\beta}_Q)$, $(\tilde{\alpha}_L, \tilde{\beta}_L)$, dans la formule du calcul de la fonction de fiabilité, $R(t)$, on obtient les valeurs pour $\hat{R}(1)$, $\tilde{R}_Q(1)$ et $\tilde{R}_L(1)$. On a calculé aussi les écarts-types. La figure 4.1 nous montre l'évolution de ces valeurs par rapport à la représentation graphique de la fonction de fiabilité, $R(t)$, où la valeur de t est 1 et $n = 200$.

En se basant sur l'échantillon de taille n , nous avons calculé les estimateurs pour α et pour β . En remplaçant $(\hat{\alpha}, \hat{\beta})$, $(\tilde{\alpha}_Q, \tilde{\beta}_Q)$, $(\tilde{\alpha}_L, \tilde{\beta}_L)$, dans la formule du calcul du taux d'échec, $H(1)$, on obtient les valeurs pour $\hat{H}(1)$, $\tilde{H}_Q(1)$ et $\tilde{H}_L(1)$. La figure 4.2 nous montre l'évolution de ces valeurs par rapport à la représentation graphique de la fonction du taux d'échec $H(1)$, où la valeur de t est 1 et $n = 200$.

Les tableaux 4.1 à 4.4 montrent que les estimateurs de Bayes sous la fonction de perte LINEX sont meilleurs que les estimateurs de Bayes sous la fonction quadratique et que les estimateurs du maximum de vraisemblance. Pour des petites

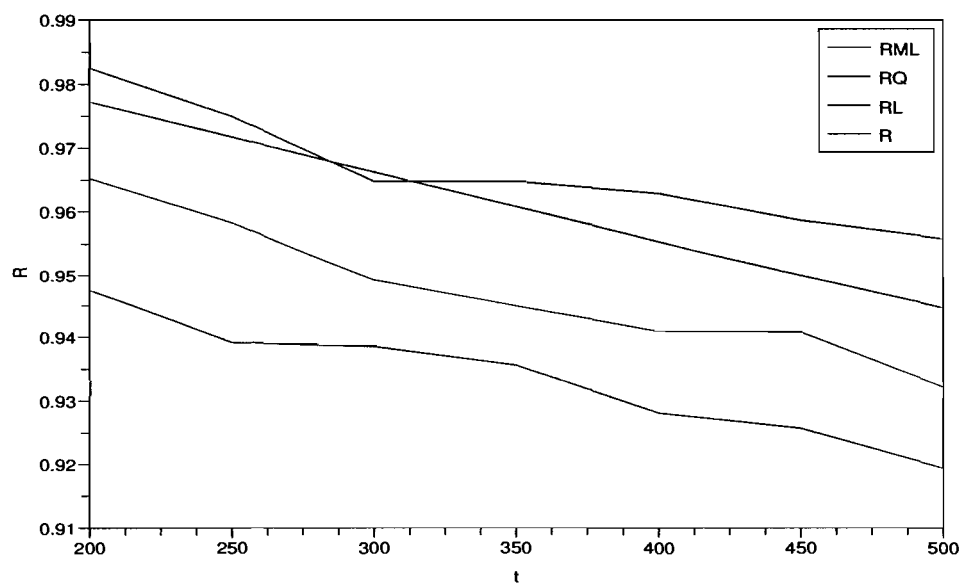


FIGURE 4.1. Estimateurs pour la fonction de fiabilité.

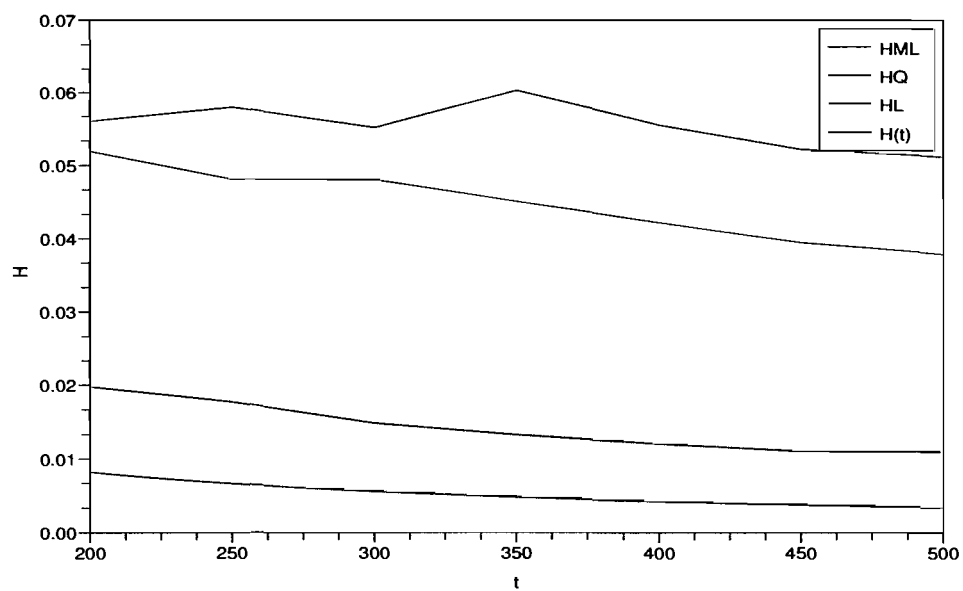


FIGURE 4.2. Estimateurs pour le taux d'échec à $t = 1$.

valeurs de n ($n < 30$), nous nous attendions à ce que les estimateurs du maximum de vraisemblance pourraient être meilleurs que les estimateurs de Bayes obtenus par l'approximation Lindley, parce que la forme de Lindley a une approximation de $\sigma(n^{-1})$.

En regardant les figures 1 et 2, nous voyons que la courbe dans le cas de l'évaluation de l'estimateur de Bayes sous la fonction de perte Linex est plus proche de la courbe réelle que les autres évaluations tant pour la fonction de fiabilité que pour la fonction de taux d'échec.

Dans le tableau 4.5, nous présentons différentes valeurs de l'estimateur de Bayes de α sous la fonction Linex, en faisant varier le paramètre a . D'après l'étude de Varian (1975) nous savons que si le paramètre a prend des valeurs petites (autour de zéro), la forme de la fonction de perte Linex s'approche de la forme de la fonction de perte quadratique, alors les estimateurs de Bayes sous la fonction de perte Linex se confondent avec les estimateurs de Bayes obtenus sous la fonction de perte quadratique. Si le paramètre a prend des valeurs différents de zéro, les estimateurs de Bayes sous la fonction de perte Linex sont différents des estimateurs de Bayes obtenus sous la fonction de perte quadratique.

TABLE 4.5. Estimateur de Bayes de α sous la fonction de perte Linex (variation d'après le paramètre a).

a	$\tilde{\alpha}_L$	$s_{\tilde{\alpha}_L}$
-20	0.534	0.033
20	0.344	0.041
-10	0.631	0.031
10	0.433	0.035
-5	0.264	0.036
5	0.459	0.042

Dans le tableau 4.6, nous présentons différentes valeurs de l'estimateur de Bayes de β sous la fonction Linex, en faisant varier le paramètre a .

TABLE 4.6. Estimateur de Bayes de β sous la fonction de perte Linex (variation d'après le paramètre a).

a	$\tilde{\beta}_L$	$s_{\tilde{\beta}_L}$
-20	7.813	0.351
20	6.251	0.338
-10	8.019	0.423
10	7.895	0.411
-5	7.632	0.379
5	2.553	0.377

Dans le tableau 4.7, nous présentons différentes valeurs de l'estimateur de Bayes de la fonction de fiabilité sous la fonction de perte Linex, en faisant varier le paramètre a .

TABLE 4.7. Estimateur de Bayes de la fonction de fiabilité sous la fonction de perte Linex (variation d'après le paramètre a).

a	\tilde{R}_L	$s_{\tilde{R}_L}$
-20	0.949	0.022
20	0.963	0.031
-10	0.959	0.035
10	0.934	0.029
-5	0.945	0.027
5	0.906	0.025

Ensuite nous présentons différentes valeurs de l'estimateur de Bayes du taux d'échec sous la fonction de perte Linex, en faisant varier le paramètre a (tableau 4.8).

En faisant varier le paramètre a , les tableaux 4.5 à 4.8 montrent que pour des valeurs de a positives, les estimateurs de Bayes de α et β sous la fonction de perte Linex sont inférieurs aux estimateurs de Bayes sous la fonction de perte Linex calculés pour des valeurs de a négatives. Étant donné que le paramètre a

TABLE 4.8. Estimateur de Bayes du taux d'échec sous la fonction de perte Linex (variation d'après le paramètre a).

a	\tilde{H}_L	$s_{\tilde{H}_L}$
-20	0.045	0.007
20	0.034	0.005
-10	0.068	0.005
10	0.054	0.006
-5	0.075	0.007
5	0.055	0.008

fait partie d'un exposant, le signe du paramètre a va influencer les valeurs des estimateurs.

4.2. ESTIMATEURS AVEC β CONNU

Nous sommes rendus à analyser les estimateurs dans le cas où β est connu, $\beta_0 = 4.35$. (Tableaux 4.9 à 4.15)

Dans le tableau 4.9 nous avons les valeurs des différents estimateurs pour α . Dans la première colonne, nous avons la taille de l'échantillon n , dans la deuxième colonne, nous avons l'estimateur du maximum de vraisemblance pour α , $\hat{\alpha}$; dans la troisième colonne nous avons l'estimateur de Bayes sous la fonction de perte quadratique, $\tilde{\alpha}_Q$ et l'estimateur de Bayes sous la fonction Linex ($a=10$), $\tilde{\alpha}_L$. (Pour avoir une meilleure idée sur les valeurs des estimateurs de Bayes sous la fonction Linex, nous faisons varier le paramètre a). L'estimateur du maximum du vraisemblance (MLE) de α , $\hat{\alpha}$, peut être calculé en utilisant l'équation

$$\hat{\alpha} = r / (T_1 - n \log \beta_0) \quad (4.2.1)$$

avec $T_1 = (n - r) \log(x_r + \beta_0) + \sum_{i=1}^r \log(x_i + \beta_0)$. Sous la fonction de perte quadratique, l'estimateur de Bayes $\tilde{\alpha}_Q$ de α est la moyenne a posteriori donnée par

$$\tilde{\alpha}_Q = (r + c) / (w + d). \quad (4.2.2)$$

(Voir Chapitre 3). L'estimateur de Bayes $\tilde{\alpha}_L$ de α sous la fonction de perte Linex ($a=10$) est donné par

$$\tilde{\alpha}_L = [(r + c)/a] \log\{1 + a/(w + d)\}. \quad (4.2.3)$$

TABLE 4.9. Estimateurs pour α .

n	$\hat{\alpha}$	$s_{\hat{\alpha}}$	$\tilde{\alpha}_Q$	$s_{\tilde{\alpha}_Q}$	$\tilde{\alpha}_L$	$s_{\tilde{\alpha}_L}$
200	0.169	0.002	0.172	0.004	0.157	0.003
250	0.155	0.002	0.189	0.003	0.152	0.004
300	0.156	0.003	0.157	0.004	0.153	0.005
350	0.154	0.003	0.171	0.004	0.153	0.003
400	0.152	0.002	0.159	0.004	0.149	0.004
450	0.140	0.003	0.154	0.004	0.137	0.004
500	0.152	0.003	0.186	0.003	0.145	0.005

Ensuite nous présentons les valeurs des estimateurs pour la fonction de fiabilité (Tableau 4.10). Nous calculons les valeurs des estimateurs en utilisant la formule de la fonction de fiabilité suivante (Chapitre 3)

$$R(t) = \beta_0^\alpha (t + \beta_0)^{-\alpha}. \quad (4.2.4)$$

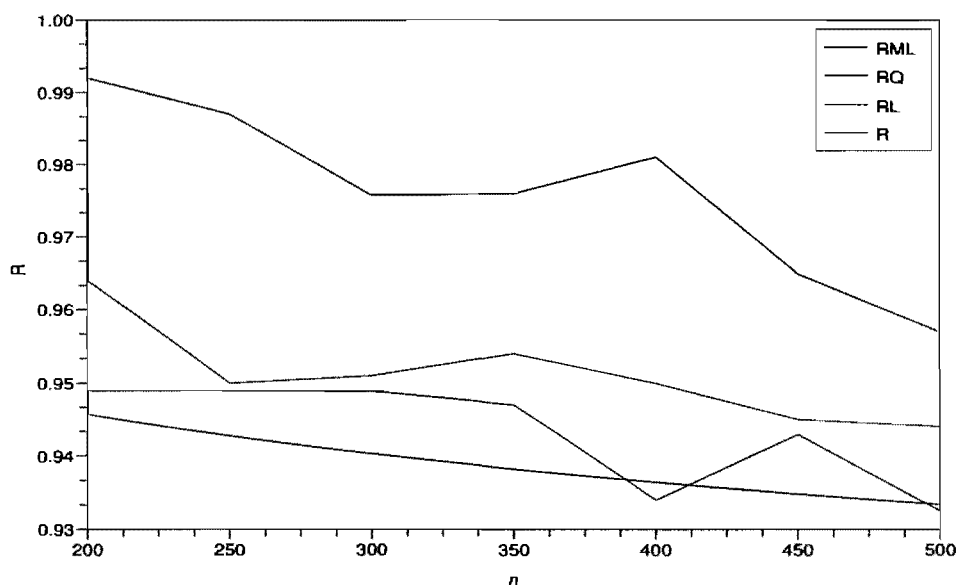
Dans la première colonne nous avons l'estimateur du maximum de vraisemblance, dans la deuxième colonne, l'estimateur de Bayes sous la fonction de perte quadratique et dans la dernière, l'estimateur de Bayes sous la fonction Linex ($a=10$).

En se basant sur l'échantillon de taille n , nous avons calculé les estimateurs pour α . En remplaçant $\hat{\alpha}$, $\tilde{\alpha}_Q$, $\tilde{\alpha}_L$, dans la formule du calcul de la fonction de fiabilité, $R(1)$, on obtient les valeurs pour $\hat{R}(1)$, $\tilde{R}_Q(1)$ et $\tilde{R}_L(1)$. La figure 4.3 nous montre l'évolution de ces valeurs par rapport à la représentation graphique de la fonction de fiabilité, $R(1)$, où la valeur de t est 1 et $n = 200$.

Dans le tableau 4.11 nous présentons les valeurs des estimateurs pour le taux

TABLE 4.10. Estimateurs pour la fonction de fiabilité $R(1)$.

n	$\hat{R}(1)$	$s_{\hat{R}}$	$\tilde{R}_Q(1)$	$s_{\tilde{R}_Q}$	$\tilde{R}_L(1)$	$s_{\tilde{R}_L}$
200	0.964	0.0028	0.949	0.0025	0.992	0.0022
250	0.950	0.0023	0.950	0.0029	0.977	0.0021
300	0.951	0.0020	0.949	0.0020	0.957	0.0019
350	0.954	0.0020	0.950	0.0022	0.966	0.0019
400	0.950	0.0019	0.935	0.0020	0.991	0.0017
450	0.920	0.0018	0.919	0.0019	0.946	0.0016
500	0.950	0.0018	0.915	0.0018	0.969	0.0015

FIGURE 4.3. Estimateurs pour la fonction de fiabilité, $R(1)$.

d'échec. Nous calculons les valeurs des estimateurs en utilisant la formule suivante pour le taux d'échec (Chapitre 3)

$$H(t) = \alpha(t + \beta)^{-1} \quad (4.2.5)$$

Dans la première colonne, nous avons la taille de l'échantillon n , dans la deuxième colonne, nous avons l'estimateur du maximum de vraisemblance ; dans la troisième colonne nous avons l'estimateur de Bayes sous la fonction de perte quadratique et dans la dernière, l'estimateur de Bayes sous la fonction de perte

Linex ($a = 10$). En se basant sur l'échantillon de taille n , nous avons calculé les

TABLE 4.11. Estimateurs de Bayes du taux d'échec $H(1)$.

n	$\hat{H}(1)$	$s_{\hat{H}}$	$\tilde{H}_Q(1)$	$s_{\tilde{H}_Q}$	$\tilde{H}_L(1)$	$s_{\tilde{H}_L}$
200	0.048	0.0027	0.067	0.0028	0.052	0.0023
250	0.046	0.0026	0.061	0.0027	0.049	0.0022
300	0.043	0.0024	0.053	0.0025	0.041	0.0022
350	0.044	0.0020	0.049	0.0023	0.040	0.0020
400	0.042	0.0019	0.046	0.0022	0.038	0.0021
450	0.043	0.0018	0.045	0.0020	0.039	0.0018
500	0.041	0.0017	0.041	0.0020	0.035	0.0016

estimateurs pour α . En remplaçant $\hat{\alpha}$, $\tilde{\alpha}_Q$, $\tilde{\alpha}_L$, dans la formule du calcul du taux d'échec, $H(t)$, on obtient les valeurs pour $\hat{H}(1)$, $\tilde{H}_Q(1)$ et $\tilde{H}_L(1)$. La figure 4.4 nous montre l'évolution de ces valeurs par rapport à la représentation graphique de la fonction du taux d'échec, $H(t)$, où la valeur de t est 1 et $n = 200$.

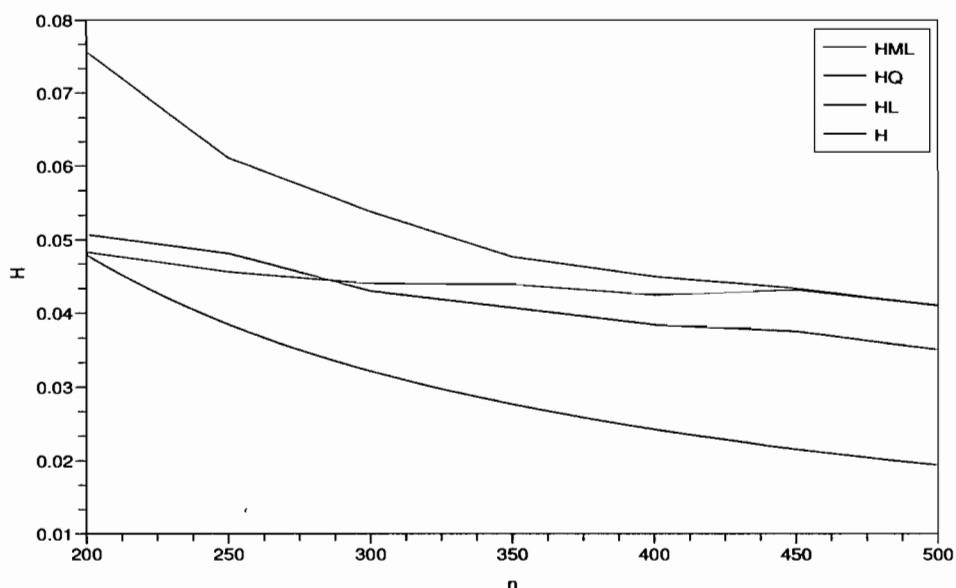


FIGURE 4.4. Estimateurs pour le taux d'échec à $t=1$.

Les tableaux 4.9 à 4.11 montrent que pour des échantillons de taille 200 à 500 les estimateurs de Bayes avec la fonction de perte Linex sont inférieurs aux estimateurs de Bayes sous la fonction de perte quadratique. Par contre, ils sont un peu plus grands que les estimateurs du maximum de vraisemblance. En regardant les figure 3 et 4, nous voyons que la courbe dans le cas de l'évaluation de l'estimateur de Bayes sous la fonction de perte Linex est plus proche de la courbe réelle que pour les autres estimateurs tant pour la fonction de fiabilité que pour la fonction de taux d'échec.

Maintenant, nous calculons les estimateurs de Bayes de α sous la fonction de perte normale reflétée. Nous minimisons la fonction de perte normale reflétée modifiée (MRN) :

$$E[L_{MRN}(Y)] = \frac{K_{\Delta}}{1 - \exp[-\frac{1}{2}(\Delta/\gamma)^2]} \left\{ 1 - \frac{\gamma}{\sqrt{\sigma^2 + \gamma^2}} \exp \left[-\frac{(\mu - T)^2}{2(\sigma^2 + \gamma^2)} \right] \right\}. \quad (4.2.6)$$

où les paramètres sont K_{Δ} , T , Δ . Pour faire le calcul, nous prenons pour les paramètres, les valeurs suivantes : $K_{\Delta} = 150$, $T = 10$, $\Delta = 4$. En prenant des échantillons de tailles de 200 à 500, nous obtenons les estimateurs présentés dans le tableau 4.12.

TABLE 4.12. Estimateur de Bayes de α sous la fonction de perte normale reflétée.

n	$\hat{\alpha}_N$	$s_{\hat{\alpha}_N}$
200	8.613	0.238
250	7.455	0.224
300	9.072	0.345
350	8.779	0.267
400	8.245	0.254
450	8.998	0.268
500	8.002	0.209

Nous observons que les estimateurs de Bayes pour α sous la fonction de perte normale reflétée sont supérieurs aux estimateurs de Bayes avec la fonction de perte

Linex, ainsi qu'aux estimateurs de Bayes sous la fonction de perte quadratique et aussi aux estimateurs du maximum de vraisemblance.

Dans le tableau 4.13, nous présentons les valeurs de l'estimateur de Bayes sous la fonction de perte Linex $\hat{\alpha}_{BL}$ de α donné par

$$\hat{\alpha}_{BL} = [(r + c)/a] \log[1 + a/(w + d)]. \quad (4.2.7)$$

Nous faisons varier les valeurs du paramètre a et nous prenons un échantillon de $n = 200$ données.

TABLE 4.13. Estimateur de Bayes de α sous la fonction de perte Linex (variation d'après le paramètre a).

a	$\tilde{\alpha}_L$	$s_{\tilde{\alpha}_L}$
-20	0.991	0.0032
20	0.172	0.0023
-10	0.423	0.0031
10	0.301	0.0028
-5	0.178	0.0033
5	0.259	0.0029

La variation du paramètre a montre que pour des valeurs de a positives, les estimateurs de Bayes de α sous la fonction de perte Linex sont inférieurs aux estimateurs de Bayes sous la fonction de perte Linex calculés pour des valeurs de a négatives.

Dans le tableau 4.14, nous présentons les valeurs d'estimateurs pour la fonction de fiabilité en faisant varier le paramètre a . L'échantillon considéré a la taille $n = 200$.

Nous calculons les valeurs d'estimateurs en utilisant la formule de la fonction de fiabilité suivante (Chapitre 3), pour $t = 1$:

$$R(t) = \beta^\alpha (t + \beta)^{-\alpha}. \quad (4.2.8)$$

Nous faisons varier les valeurs du paramètre a et nous considérons un échantillon de taille $n = 200$.

$$\hat{u}_{BL} = -\frac{1}{a} \log [E_u(e^{-au})] E_u(e^{-au}). \quad (4.2.9)$$

TABLE 4.14. Estimateur de Bayes de la fonction de fiabilité sous la fonction de perte Linex (variation d'après le paramètre a).

a	\tilde{R}_L	$s_{\tilde{R}_L}$
-20	0.993	0.0038
20	0.971	0.0040
-10	0.989	0.0037
10	0.993	0.0039
-5	0.961	0.0035
5	0.993	0.0038

En ce qui concerne l'estimateur de Bayes de la fonction de fiabilité, on observe que même si pour des valeurs du paramètre a positives, les estimateurs de Bayes de la fonction de fiabilité sous la fonction de perte Linex sont inférieurs aux estimateurs de Bayes de la fonction de fiabilité sous la fonction de perte Linex calculés pour des valeurs de a négatives, l'écart est très petit.

Nous présentons les valeurs d'estimateurs pour le taux d'échec en faisant varier le paramètre a . Nous calculons les estimateurs pour un échantillon de taille $n = 200$, en utilisant la formule suivante pour le calcul de la fonction du taux d'échec (Chapitre 3), pour $t = 1$:

$$H(t) = \alpha(t + \beta)^{-1}. \quad (4.2.10)$$

Pour les estimateurs de Bayes du taux d'échec sous la fonction de perte Linex, la tendance se maintient, en général, c'est-à-dire que pour des valeurs du paramètre a positives, les estimateurs de Bayes de la fonction du taux d'échec sous la fonction de perte Linex sont inférieurs aux estimateurs de Bayes de la fonction du taux d'échec sous la fonction de perte Linex calculés pour des valeurs de a négatives. Par contre, l'écart est proche de zéro.

TABLE 4.15. Estimateur de Bayes du taux d'échec sous la fonction de perte Linex (variation d'après le paramètre a).

a	\tilde{H}_L	$s_{\tilde{H}_L}$
-20	0.008	0.0004
20	0.009	0.0003
-10	0.005	0.0005
10	0.008	0.0004
-5	0.045	0.0004
5	0.007	0.0003

Chapitre 5

CONCLUSION

Ce mémoire visait à comparer les estimateurs obtenus sous différentes fonctions de perte. Basé sur un échantillon des temps de vie de la distribution de Pareto, les estimateurs de Bayes de deux paramètres (inconnus) de la fonction de fiabilité et du taux d'échec sont obtenus en utilisant l'approximation de Bayes dû à Lindley. On calcule les estimateurs de Bayes sous la fonction de perte quadratique, sous la fonction Linex et sous la fonction normale reflétée modifiée, ainsi que les estimateurs de maximum de vraisemblance et les estimateurs des moments. Les comparaisons sont faites entre ces estimateurs calculés en se basant sur l'étude de simulations. On étudie l'effet de fonctions de perte symétriques et asymétriques sur ces estimateurs.

Le cas de α et β inconnus

Les tableaux 4.1 à 4.8 montrent que les estimateurs de Bayes sous la fonction de perte LINEX sont meilleurs que les estimateurs de Bayes sous la fonction quadratique et que les estimateurs du maximum de vraisemblance. Pour des petites valeurs de n ($n < 30$), nous nous attendions à ce que les estimateurs du maximum de vraisemblance pourraient être meilleurs que les estimateurs de Bayes obtenus par l'approximation Lindey. En regardant les figures 4.1 et 4.2, nous voyons que la courbe dans le cas de l'évaluation de l'estimateur de Bayes sous la fonction de perte Linex est plus proche de la courbe réelle que les autres évaluations tant pour la fiabilité que pour la fonction de taux d'échec.

Le cas de β connu

Les tableaux 4.9 à 4.15 montrent que pour des échantillons tant petits que grands, les estimateurs de Bayes avec la fonction de perte Linex sont inférieurs aux estimateurs de Bayes sous la fonction quadratique et aussi aux estimateurs des moments. Par contre ils sont un peu plus grands que les estimateurs du maximum de vraisemblance. En regardant les figures 3 et 4, nous voyons que la courbe dans le cas de l'évaluation de l'estimateur de Bayes sous la fonction de perte Linex est plus proche de la courbe réelle que les autres évaluations tant pour la fiabilité que pour la fonction de taux d'échec.

D'après les résultats, nous observons que dans la plupart des situations impliquant l'évaluation de la fiabilité et l'évaluation du taux d'échec, des fonctions de perte asymétrique (Linex) sont plus appropriées que la fonction de perte quadratique. Pour les deux cas : α et β inconnus et β connu, l'influence de la variation du paramètre a , dans la fonction de perte Linex, détermine le fait qu'en général, pour des valeurs positives de a , les estimateurs de α sont inférieurs aux estimateurs de α obtenus pour des valeurs négatives de a . Le même phénomène se produit pour les estimateurs de β , de la fonction de fiabilité, R et du taux d'échec, H , sauf que l'écart entre les valeurs est très petit (voir les tableaux 4.5 à 4.8 et 4.13 à 4.15). En ce qui concerne l'estimateur de α sous la fonction de perte normale réflétée modifiée, on observe qu'il prend des valeurs plus grandes que les estimateurs calculés sous les autres fonctions de perte.

Globalement, les estimateurs de Bayes calculés sous la fonction de perte Linex sont meilleurs que les estimateurs de Bayes sous la fonction de perte quadratique, sous la fonction de perte normale réflétée modifiée, ainsi que les estimateurs du maximum de vraisemblance. Par contre, parce que les résultats des simulations sont ici loin d'être exhaustifs, il est encore difficile de choisir une fonction de perte en particulier qui donnerait les « meilleurs » estimateurs.

Ainsi, pour de futures recherches, il serait intéressant d'élargir le calcul des estimateurs pour d'autres fonctions de perte, telle que la fonction de perte 0-1 et la fonction de perte quadratique pondérée.

BIBLIOGRAPHIE

- [1] Abdel-Ghaly, A.A., Attia, A.F., Aly, H.M. Estimation of the Parameters of Pareto Distribution and the Realibility Function Using Accelered Life Testing with Censoring. *Commun.Statist.Theor.Meth.* 1998, 27(2), 469-484.
- [2] Al-Hussaini, E.K., Jaheen, Z.F. Bayesian Estimation of the Parameters, Reliability and failure Rate Functions of the Burr Type XII Failure Model. *J.Statist.Comput.Simul.* 1992, 41, 1-40.
- [3] Al-Hussaini, E.K., Jaheen, Z.F. Approximate Bayes Estimators Applied to the Burr Model. *Commun.Statist.Theor.Meth.* 1994, 23(1), 99-121.
- [4] Arnold, B.C, Press, S.J. Bayesian Inference for Pareto Populations. *J.Econometrics.* 1983, 21, 287-306.
- [5] Arnold, B.C, Press, S.J. Bayesian Estimation and Prediction for Pareto Data. *J.Amer.Stat.Assoc.* 1989, 84, 1079-1084.
- [6] Basu, A.P., Ebrahimi, N. Bayesian Approach to Life Testing and Reliability Estimation Using Asymmetric Loss Function. *J.Stat.Plann.Infer.* 1991, 29, 21-31.
- [7] Berger, J. *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlang, NY, 1985.
- [8] Berger, J., Wolpert, R. *The Likelihood Principle*.IMS Lecture Notes-Monograph Series, Hayward, California. 1988.
- [9] Canfield, R.V. A Bayesian Approach to Reliability Estimation Using Loss Function. *IEEE Trans.Reliability.* 1970, R-19, 13-16.
- [10] Devroye, L. *Non-Uniform Random Variate Generation*. New York :Springer-Verlang, 1986.

- [11] Hossain, A.M., Zimmer, W.J. Comparisons of Methods of Estimation for Pareto Distribution of the First Kind. *Commun.Statist.Theor.Meth.* 2000, 29(4), 859-878.
- [12] Howlader, H.A., Weiss, G. On Bayesian Estimation of the Cauchy Parameters. *Sankhya Series B.* 1988, 50, 350-361.
- [13] Howlader, H.A., Weiss, G. Bayes Estimators of the Reliability of the Logistic Distribution. *Commun.Statist.Theor.Meth.* 1989, 18(1), 245-259.
- [14] Lindley, D.V. Approximate Bayesian Method. *Trabajos de Estadística.* 1980, 31, 223-237.
- [15] Malik, H.J. Estimation of the Parameters of the Pareto Distribution. *Metrika.* 1970, 15, 126-132.
- [16] Morlat, G. Statistique et théorie de la décision. *Mathématiques et Sciences Humaines.* 1964, p.1-7, <http://www.numdam.org/numdam-bin/>.
- [17] Newton, I., Raphson, J. De analysi per aequationes numero terminorum infinitas, 1685, [http://fr.wikipedia.org/wiki/Méthode de Newton/](http://fr.wikipedia.org/wiki/Méthode_de_Newton/).
- [18] Pandey, B.N., Singh, B.P., Mishra, C.S. Bayes Estimation of Shape Parameter of Classical Pareto Distribution Under Linex Loss Function. *Commun.Statist.Theor.Meth.* 1996, 25(12), 3125-3145.
- [19] P.Robert, C. *The Bayesian Choice.* New York : Second Edition, Springer-Verlang, 2001.
- [20] Sinha, S.K. Bayes Estimation of Reliability Function of Normal Distribution. *IEEE Trans.Reliability.* 1985, R-34, 360-362.
- [21] Sinha, S.K., Sloan, J.A. Bayes Estimation of the Parameters and Reliability Function of the 3-Parameter Weibull Distribution. *IEEE Trans.Reliability.* 1988, R-37, 364-369.
- [22] Soliman, A.A. Linex and Quadratic Approximate Bayes Estimators Applied to the Pareto Model. *Commun.Statist.-Simula.* 2001, 30(1), 47-62.
- [23] Spring, A.F. The reflected normal loss function. *The Canadian Journal of Statistics,* Vol.21, No.3, 1993 : 321-330.
- [24] Sun, F., Laramée, J-Y., Ramberg S.J. On Spring's normal loss function, *The Canadian Journal of Statistics,* Vol.24, No.2, 1996, 241-249.

- [25] Tiwari, R.C., Yang, Y., Zalkikar, J.N. Bayes Estimation for the Pareto Failure-Model Using Gibbs Sampling. *IEEE Trans.Reliability*. 1996, R-45, 471-476.
- [26] Zellner, A. Bayesian Estimation and Prediction Using Asymmetric Loss Functions. *J.Amer.Stat.Assoc.* 1986, 81, 446-451.
- [27] Varian, H.R. *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*. Amsterdam : North-Holland, 1975, 195-208.
- [28] Wald, A. *Statistical Decision Functions*. 1950, Wiley.