

N° d'ordre:

RÉPUBLIQUE ALGERIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ MOULOUD MAMMERRI DE TIZI OUZOU
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES
LABORATOIRE LMPA



MÉMOIRE DE MASTER

Filière : Mathématiques
Spécialité : Probabilités & Statistique

Par

MARIA KHENDRICHE

ROBUSTESSE BAYÉSIENNE POUR L'ESTIMATION DU PARAMÈTRE DE LA LOI DE POISSON

Soutenue le 5 Octobre 2022 devant le jury :

HAMAZ ABDELGHANI	UMMTO	Président du jury
BELKACEM CHERIFA	UMMTO	Examineur
ATIL LYNDA	UMMTO	Encadreur

Année Universitaire : 2021/2022

*A mon très cher père (arezki khendriche),
Autant de phrases et d'expressions aussi éloquentes soit-elles ne sauraient
exprimer ma gratitude et ma reconnaissance. Tu as su m'inculquer le sens de
la responsabilité, de l'optimisme et de la confiance en soi face aux difficultés de
la vie. Tes conseils ont toujours guidé mes pas vers la réussite. Ta patience sans
fin, ta compréhension et ton encouragement sont pour moi le soutien
indispensable que tu as toujours su m'apporter. Je te dois ce que je suis
aujourd'hui et ce que je serai demain et je ferai toujours de mon mieux pour
rester ta fierté et ne jamais te décevoir.
Que Dieu, le miséricordieux, t'accueille dans son éternel paradis. . .*

REMERCIEMENTS

IL m'est agréable d'adresser mes premiers remerciements à mon promoteur Madame Atil Lynda. Je la remercie de m'avoir encadrée, orientée, aidé et conseillée.

J'adresse mes sincères remerciements à tous les professeurs et toutes les personnes qui par leurs paroles, leurs écrits, leurs conseils et leurs critiques ont guidé mes réflexions et ont accepté à me rencontrer et répondre à mes questions durant mes recherches.

Aussi, je tiens à remercier ma mère et mes frères et ma soeur (Amine, Asma, Ishak) qui ont toujours cru en moi, soutenue et encouragée à travailler et réussir, qui sans leur amour, leur confiance et leurs sacrifices je ne serais là où je suis et la personne que je suis aujourd'hui.

Mes amis Manel, Yacine, Nissa, Maroua, Wassim, Yasmine, Sarah, et Ourdia. Je ne peux trouver les mots justes et sincères pour vous exprimer mon affection et mes pensées, vous êtes pour moi des sœurs et des frères sur qui je peux compter.

Je remercie également toute l'équipe pédagogique du master probabilités et statistiques pour avoir assuré la partie théorique de ma formation, indispensable à une bonne entrée dans le monde professionnel.

Enfin, je tiens également à remercier les membres du jury qui me feront l'honneur de juger notre travail.

Tizi-Ouzou, le 15 novembre 2022.

TABLE DES MATIÈRES

TABLE DES MATIÈRES	iv
LISTE DES FIGURES	v
LISTE DES TABLEAUX	v
INTRODUCTION	1
1 OUTILS STATISTIQUES ET MÉTHODES D'ESTIMATION	3
1.1 INTRODUCTION	3
1.2 ESTIMATION CLASSIQUE	3
1.2.1 Estimateur classique	3
1.2.2 Exemples d'estimateurs classiques	3
1.2.3 Construction d'estimateurs	4
1.3 ESTIMATION BAYÉSIENNE	5
1.3.1 Une brève histoire de la Statistique Bayésienne	5
1.3.2 Estimation Bayésienne	7
1.3.3 Le modèle statistique Bayésien	7
1.3.4 Information a priori	7
1.3.5 Description du modèle Bayésien	7
1.3.6 Fonctions coût usuelles	9
1.3.7 Le coût quadratique	10
1.3.8 La fonction coût absolue	10
1.3.9 La fonction coût 0-1	11
1.3.10 La fonction coût entropique	11
2 ESTIMATION DU PARAMÈTRE DE LA LOI DE POISSON ET NOTION DE ROBUSTESSE	13
2.1 INTRODUCTION	13
2.2 LA LOI DE POISSON	13
2.3 MÉTHODES D'ESTIMATION D'UNE LOI DE POISSON	13
2.3.1 L'estimateur du maximum de vraisemblance (The Maximum Likelihood Estimator MLE)	13
2.3.2 L'estimation Bayésienne	15
2.3.3 L'estimateur Bayésien du paramètre θ associé à la fonction coût quadratique	16
2.3.4 L'estimateur Bayésien du paramètre θ associé à la fonction coût quadratique généralisée	17
2.4 MÉTHODE DE MONTE-CARLO PAR CHAÎNES DE MARKOV (MCMC) .	19
2.4.1 Algorithme	19
2.5 LA ROBUSTESSE BAYÉSIENNE	21
2.5.1 Quelques notions de base	21
3 ROBUSTESSE BAYÉSIENNE DU PARAMÈTRE DE LA LOI DE POISSON	23

3.1	INTRODUCTION	23
3.2	APPLICATION	23
	CONCLUSION GÉNÉRALE	34
	BIBLIOGRAPHIE	35

LISTE DES FIGURES

3.1	Tableau de variation de $f(\alpha)$	26
3.2	Tableau de variation de $f(\alpha)$	26
3.3	Tableau de variation de $f(\alpha)$	27
3.4	Tableau de variation de $g(\beta)$	30
3.5	Tableau de variation de $g(\beta)$	30

LISTE DES TABLEAUX

1.1	Estimateurs Bayésienne du paramètre θ sous coût quadratique pour les lois a priori conjuguées des familles exponentielles usuelles.	10
-----	--	----

INTRODUCTION GÉNÉRALE

LA statistique fait partie de ce que l'on appelle la science des données. Elle est à la fois une science, une méthode et un ensemble de techniques. Elle reste cependant un domaine des mathématiques, ayant comme outil principal les probabilités.

Dans la littérature, deux approches sont considérées ; l'approche **statistique fréquentielle** appelée communément la statistique classique et l'approche **statistique Bayésienne**. Dans l'approche statistique classique, "tout est dans les données", ces données permettent de décrire l'incertitude sur le paramètre θ . Dans ce cas de figure, θ est une constante inconnue, et la loi des observations est la pièce maîtresse dans l'inférence fréquentiste. La démarche statistique Bayésienne se distingue de celle de la statistique classique par le fait que le paramètre θ est une variable aléatoire, à laquelle on associe une distribution de probabilité appelée distribution a priori, cette loi doit être indépendante des observations.

Le choix de cette loi a priori constitue le point le plus critiqué de l'analyse Bayésienne par les non Bayésiens. Il peut exister un ensemble de lois compatibles pour le paramètre θ . Le théorème de Bayes nous permet de rassembler l'information a priori avec celle apportée par les données pour une nouvelle distribution dite la distribution a posteriori, notée $\pi(\theta|x)$ qui est le fondement de l'inférence Bayésienne.

Comme il a été énoncé précédemment, plusieurs lois a priori peuvent convenir. La robustesse Bayésienne consiste à regrouper toutes ces lois en une classe et à évaluer les changements effectués sur les quantités a posteriori lorsque la loi a priori varie dans cette classe, afin de parvenir à une inférence robuste, pour construire des estimateurs tout aussi robustes.

Le modèle considéré dans notre travail est le modèle de Poisson. L'approche statistique Bayésienne et la robustesse de ses méthodes feront l'objet de notre mémoire, sur le paramètre de la loi de Poisson.

La loi de Poisson joue un rôle très important dans l'analyse statistique des données, ce type de données provient des situations dans lesquelles il existe plusieurs possibilités de se produire en cas d'intérêt, comme le nombre de clients appelant au centre d'aide en une journée, les visiteurs d'une chaîne YouTube d'une idole, les patients atteints de Covid-19...etc. Donc, la loi de Poisson peut être utilisée pour déterminer la probabilité de multiples événements dans une période donnée.

Divers chercheurs ont développé des procédures d'inférence pour la loi de Poisson. Araveeporn (2014) a proposé des statistiques pour tester des hypothèses en utilisant la moyenne de l'estimateur du paramètre de Poisson obtenue via l'estimateur du maximum de vraisemblance, méthode de Monte-Carlo par chaînes de Markov, il a également proposé quelques approches Bayésiennes. Hassan et al(2008), ont étudié l'estimateur Bayésien et l'estimateur du maximum de vraisemblance pour une loi de Poisson zero-truncated. Tout comme pour la loi de Poisson, la technique d'estimation Bayésienne a également été appliquée sur les paramètres de la loi géométrique, la loi Binomiale, la loi de Pareto, et en utilisant la loi uniforme et la loi a priori de Jeffreys sous différentes fonctions coût, Le résultat indique principalement que l'estimation Bayésienne utilisant les différentes fonctions coût était proche des valeurs réelles, suggérant que cela pourrait être une approche utile pour estimer les paramètres d'intérêt dans diverses situations.

Les méthodes d'estimation du paramètre d'une variable aléatoire peuvent être regroupées en trois catégories : classique, Bayésien, et Bayésien Empirique (EB), qui sont toutes utiles dans différents ensembles de circonstances, les méthodes Bayésiennes fournissent des estimations du paramètre inconnu en utilisant une loi a priori informative fixe, sont meilleures que les approches classiques pour la plupart des situations impliquant une fonction coût. Par conséquent, l'estimation d'un paramètre de Poisson à l'aide de méthodes Bayésiennes sous différentes fonctions coût présente un intérêt à cette étude.

Notre mémoire est organisé comme suit : Le premier chapitre donnera un aperçu général sur l'approche statistique Bayésienne, en introduisant des méthodes de construction d'un estimateur et les fonctions coût usuelles utilisées.

Dans le second chapitre, nous introduirons la loi de Poisson, sa construction et son intérêt dans son utilisation. Divers estimateurs Bayésiens du paramètre de la loi de Poisson sont construits. La robustesse Bayésienne y est consacrée, tout en abordant la notion de "**Range**", qu'on appelle également "**oscillation**" ou "**étendue**" du coût a posteriori. Enfin le dernier chapitre mettra en oeuvre le concept de la robustesse Bayésienne. Nous y présenterons l'article intitulé "Stability of the Bayesian estimator of the Poisson mean under the in-exactly specified Gamma prior " où une analyse Bayésienne est menée sur la loi de Poisson sous la fonction coût quadratique et la loi a priori gamma. Zielinski et Meczarski ont eu pour objectif la construction d'un estimateur robuste (estimateur stable).

OUTILS STATISTIQUES ET MÉTHODES D'ESTIMATION



1.1 INTRODUCTION

Dans ce premier chapitre nous allons rappeler les principales notions de base nécessaires pour la maîtrise de la suite du mémoire. Nous introduirons la notion de l'estimateur classique et l'estimateur Bayésien toute en introduisant des méthodes de construction d'un estimateur . Mais aussi, Les lois qui interviennent dans la statistique Bayésienne . Et enfin les fonctions coût usuelles utilisées dans cette dernière .

1.2 ESTIMATION CLASSIQUE

Définition 1.1 *L'estimation consiste à utiliser un échantillon de n observations de la population pour approximer les paramètres de cette population. Il s'agit de plus de pouvoir émettre un jugement sur la qualité de l'information tirée de cet échantillon.*

1.2.1 Estimateur classique

On souhaite estimer un paramètre θ d'une population qui peut être sa moyenne $E(X)$, son écart-type $\text{VAR}(X)$, Un estimateur de θ est une statistique T (donc une fonction de (X_1, \dots, X_n)) dont la réalisation est envisagée comme une "bonne valeur" du paramètre θ . On parle d'estimation de θ associée à cet estimateur la valeur observée lors de l'expérience, c'est-à-dire la valeur prise par la fonction au point observé (x_1, \dots, x_n) .

1.2.2 Exemples d'estimateurs classiques

Estimateur de la moyenne de X

On prend en général comme estimateur de X la valeur :

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i$$

on l'appelle la moyenne empirique de X .

Estimateur de la variance de X

$\frac{n}{n-1}\delta^2$ est un estimateur consistant de $\text{VAR}(X)$ dans le cas de tirage avec remise.

Dans le cas de tirage sans remise

$$\frac{N-1}{N} \frac{n}{n-1} \delta^2$$

Qui vaut bien δ^2 lorsque $n = N$: l'échantillon est identique à la population, donc on mesure la variance réelle.

On peut remarquer que, pour N grand, le calcul avec remise et le calcul sans remise donnent des résultats presque équivalents. On prend donc en général, pour estimateur sans biais de $\text{VAR}(X)$ la valeur :

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

appelée variance empirique sans biais de X .

1.2.3 Construction d'estimateurs**Méthode des moments**

La méthode des moments est un outil d'estimation intuitif qui remonte au début des statistiques introduit en 1894 par Karl Pearson, elle consiste à estimer les paramètres recherchés en assimilant certains moments théoriques (en fonction de ces paramètres) avec leurs contreparties empiriques. L'égalisation se justifie par la loi des grands nombres Cela signifie que l'espérance mathématique peut être approchée par la moyenne empirique. Il faut donc résoudre un système d'équations.

Méthode du maximum de vraisemblance

L'estimateur du maximum de vraisemblance, comme son nom l'indique, maximise la vraisemblance définie comme suit :

Dans le cas discret considérons un échantillon $(X_1 = x_1, \dots, X_n = x_n)$ iid tel que :

$$p(x_1, \dots, x_n; \theta) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$$

$$p(x_1, \dots, x_n; \theta) = \prod_{i=1}^n \mathbb{P}(X_i = x_i)$$

car les X_i sont indépendantes

$$p(x_1, \dots, x_n; \theta) = \prod_{i=1}^n \mathbb{P}(X = x_i)$$

car les X_i sont de même loi .

Dans le cas continu iid :

$$p(x_1, \dots, x_n; \theta) = f(x_1, \dots, x_n)$$

$$p(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f_{X_i(x_i)}$$

car les X_i sont indépendantes

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

car les X_i sont de même loi .

Estimer un paramètre par la méthode du maximum de vraisemblance revient à chercher le ou les paramètres θ qui vont maximiser la quantité qui dépend des variables aléatoires de l'échantillon et du paramètre θ , c'est une méthode très utile lorsqu'on souhaite construire des tests robustes.

Remarque : La méthode des moments et la méthode du maximum de vraisemblance ne sont pas les seules méthodes qui existe mais ces deux méthodes sont les plus connues et les plus utiles dans la statistique.

1.3 ESTIMATION BAYÉSIENNE

1.3.1 Une brève histoire de la Statistique Bayésienne

Christian P. Robert a précisé dans son livre (Le choix Bayésien Principes et pratique 2005) que différents livres ont été écrits sur l'histoire de la Statistique Bayésienne, notamment Stigler (1986), Dale (1991), Lad (1996) et Hald (1998). Nous ne faisons ici que souligner quelques points forts du développement de cette discipline durant les deux cents dernières années. La formule de Bayes est apparue pour la première fois en 1761, dans le cadre de l'exemple binomial, exposé par le révérend Thomas Bayes devant la "Royal Society", est publiée de façon posthume par son ami R. Price en 1763. Pierre Simon Laplace redécouvrit ensuite cette formule de manière plus générale en 1773, sans, semble-t-il, avoir connaissance des travaux précédents de Bayes. L'utilisation du principe Bayésien devint alors courant pendant le siècle suivant, comme le rapporte Stigler (1986), mais des critiques commencèrent à émerger vers la fin du XIXème siècle, comme par exemple dans Venn (1886) ou Bertrand (1889), en particulier sur le choix de la loi a priori uniforme et des paradoxes de paramétrisation qui en résultent, voir Zabell (1989). Puis, malgré des formalisations plus poussées du paradigme Bayésien par Edgeworth et (Karl) Pearson au tournant du siècle et, plus tard, par Keynes (1921), le début du xx^{ème} siècle fut surtout marqué par, tout d'abord, Kolmogorov, qui proposa dans les années 1920 une axiomatisation de la théorie des probabilités semblant contredire le paradigme Bayésien et la notion de probabilité subjective, ensuite par Fisher qui s'éloigna de l'approche Bayésienne (Fisher, 1912) en définissant la fonction de vraisemblance (Fisher, 1922), puis en développant la Statistique fiduciaire (Fisher, 1930), et qui ne révisa jamais son opinion négative sur la Statistique

Bayésienne. Cette opposition paraît quelque peu paradoxale, car la Statistique fiduciaire tentait, en un certain sens, de surmonter la difficulté de choisir une loi a priori en la construisant à partir de la fonction de vraisemblance (Seidenfeld, 1992), dans le même esprit que les approches non informatives de Jeffreys (1939) et Bernardo (1979).

L'approche fiduciaire a été abandonnée progressivement après la mise en évidence de paradoxes fondamentaux (voir Stein, 1959, Wilkinson, 1977, et les références dans Zabell, 1992).

Le livre de Jeffreys (1939) est le premier traité moderne de Statistique Bayésienne : il couvre, en plus de la notion d'a priori non informatif, celles de loi prédictive, de facteur de Bayes et d'a priori impropre. Mais cet ouvrage publié au moment du développement par Fisher de la Statistique de la vraisemblance et des intervalles de confiance par Neyman (1934), ne rencontra pas le même succès. Les approches alternatives à la Statistique Bayésienne devinrent alors standard dans les années 1930, avec l'introduction des estimateurs du maximum de vraisemblance et le développement d'une théorie formalisée de la Statistique mathématique, pour laquelle les lois a priori n'apparaissent au mieux que comme une façon de construire des estimateurs optimaux, voir Wald (1950) ou Ibragimov et Has'minskii (Ibragimov et Has'minskii, 1981, Chapitre 6). Les tentatives d'une formalisation plus poussée de l'approche Bayésienne par Gini ou de Finetti, des années 1930 aux années 1970, ne se traduisirent pas par une plus grande popularité face à la théorie alors dominante de Neyman-Pearson, même si la communauté Bayésienne s'accroissait et produisait des traités tels que ceux de Savage (1954) et de Lindley (1965, 1971).

On peut avancer que ce n'est que très récemment que la Statistique Bayésienne a pris un nouvel élan, grâce au développement de nouveaux outils numériques qui ont toujours joué un rôle central pour le paradigme Bayésien et l'intérêt vite croissant des praticiens pour cette approche de modélisation statistique, comme souligné dans l'article de Berger (2000) sur l'état présent et futur de la Statistique Bayésienne.

La vitalité actuelle de la Statistique Bayésienne peut être mise en évidence par le pourcentage élevé d'articles Bayésiens publiés dans les revues statistiques ou concernant d'autres domaines scientifiques. Il semble donc que les praticiens de ce siècle prendront mieux en compte les avantages de la Statistique Bayésienne que leurs prédécesseurs du xx^{ème} siècle.

1.3.2 Estimation Bayésienne

Définition 1.2 *L'estimation Bayésienne s'appuie sur un théorème énoncé par le mathématicien Thomas Bayes. Ce théorème donne une méthode pour calculer la probabilité d'un phénomène grâce à la connaissance de certaines informations. L'estimation Bayésienne est donc la méthode qui s'appuie sur ce raisonnement.*

1.3.3 Le modèle statistique Bayésien

Définition 1.3 *On appelle modèle statistique Bayésien, la donnée d'un modèle classique $(\mathcal{X}, A, \{P_\theta, \theta \in \Theta\})$ où \mathcal{X} représente l'espace des observation, $f(X|\theta)$ la densité de P et d'une loi a priori noté $\pi(\theta)$. En résumé le modèle statistique Bayésien consiste donc en la donnée d'une loi a priori et de la loi des observations x_i .*

1.3.4 Information a priori

On entend par information a priori de paramètre θ , toute information disponible sur θ , en dehors de celle apportée par les observations (l'échantillon). On modélise cette information par une loi dont sa densité est notée par $\pi(\theta)$.

Remarque : Le choix de la loi a priori est très important dans la statistique Bayésienne, tout en étant très délicat, car il peut nous mener à prendre des décisions erronées :

Exemple

Nous notons le nombre annuel de sinistrés d'une police d'assurance choisie au hasard dans un portefeuille. Si celui-ci est homogène, c'est-à-dire si la sinistralité est la même pour toutes les polices, un modèle très souvent applicable est la loi de Poisson de paramètre λ . Si le portefeuille est hétérogène, une bonne description de la réalité en général est de considérer que, lors du choix d'une police, le paramètre λ est lui-même une réalisation d'une variable aléatoire de loi Gamma, qui est donc la loi a priori du paramètre.

La démarche de l'analyse Bayésienne conduit au calcul d'une loi a posteriori noté $\pi(\theta|X)$.

1.3.5 Description du modèle Bayésien

Considérons le modèle Bayésien $(\mathcal{X}, A, \{P_\theta, \theta \in \Theta\})$ où θ est le paramètre inconnu, et $\pi(\theta)$ la loi a priori.

Les lois qui interviennent dans la statistique Bayésienne sont :

Loi du couple

Cette loi est appelée loi jointe, sa densité est notée $h(\theta, x)$ on a

$$h(\theta, x) = f(X|\theta) \cdot \pi(\theta)$$

Loi marginale de X

La loi marginale de X notée $f(x)$ est calculée de la manière suivante :

$$f(x) = \int_{\Theta} h(\theta, x) d\theta$$

Loi a posteriori

C'est la loi constitutionnelle de θ sachant X, sa densité est noté $\pi(\theta|X)$, cette loi est une actualisation de la loi a priori $\pi(\theta)$ aux vu de l'observation x_i . En vertu du théorème de Bayes, on a

$$\pi(\theta|X) = \frac{f(X|\theta) \cdot \pi(\theta)}{f(X)}$$

où

$f(X|\theta)$ est la fonction de vraisemblance,

Et $f(X) = \int_{\Theta} f(X|\theta) \cdot \pi(\theta) d\theta$.

Le risque de Bayes

L'information a priori est un atout supplémentaire qui intervient dans la statistique Bayésienne, elle est mise à la disposition su statisticien via la loi a priori π_{θ} .

Définition 1.4 *considérons la moyenne du risque fréquentiste suivant la loi a priori π*

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

démonstration L'égalité découle directement du Théorème de Fubini, car, comme $L(\theta, \delta) \geq 0$

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

Définition 1.5 On définit le coût a posteriori $\varrho(\pi, \delta(x))$ comme étant la moyenne du coût par rapport à la loi a posteriori.

$$\begin{aligned}\varrho(\pi, \delta(x)) &= E^\pi[L(\pi, \delta)|x] \\ &= \int_{\Theta} L(\theta, \delta(x))\pi(\theta)d\theta\end{aligned}$$

Théorème 1.1 Un estimateur minimisant le risque intégré $r(\pi, \delta)$ est obtenu par sélection, pour chaque $x \in \mathfrak{X}$, de la valeur $\delta(x)$ qui minimise le cout moyen a posteriori, $\varrho(\pi, \delta|x)$, puisque

$$r(\pi, \delta) = \int_{\mathfrak{X}} \varrho(\pi, \delta(x)|x)f(x)dx \quad (1.1)$$

Ce résultat mène à la définition suivante d'un estimateur de Bayes.

Définition 1.6 Un estimateur de Bayes associé à une distribution a priori π et une fonction de coût L est un estimateur δ^π minimisant $r(\pi, \delta)$. Pour chaque $x \in \mathfrak{X}$, ce dernier est donné par

$$\delta^\pi(x) = \arg \min_d \varrho(\pi, d|x) \quad (1.2)$$

La valeur $r(\pi) = r(\pi, \delta^\pi)$ est alors appelée risque Bayésien.

Remarque : La statistique classique repose sur la loi des observations et la statistique Bayésienne sur la loi a posteriori, mais aussi dans la statistique Bayésienne θ désigne une variable aléatoire contrairement à la statistique classique où θ est une constante inconnue :

1.3.6 Fonctions coût usuelles

L'objectif des études inférentielles est de fournir une décision une décision au statisticiens.

Les différentes décision sont comparées au moyen d'un critère d'évolution donné par la fonction coût .

Définition 1.7 On appelle fonction coût toute fonction L :

$$\left\{ \begin{array}{l} \Theta \times \mathcal{A} \longrightarrow \mathbb{R}^+ \\ ((\theta, \delta(x)) \longrightarrow L(\theta, \delta(x))) \end{array} \right.$$

Où $\delta(x)$ est une règle de décision

$L(\theta, \delta(x))$ évalue le coût ou la perte associé a la décision $\delta(x)$ quand le paramètre vaut θ . Elle permet de calculer la perte pour une mauvaise décision.

Nous allons définir 3 fonctions coûts usuelles, elles sont appréciées pour leurs utilités et simplicités, elles nous permettent d'obtenir aisément l'estimateur Bayésien.

1.3.7 Le coût quadratique

Définition 1.8 La fonction coût quadratique est définie par :

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

une variante de cette fonction coût est le coût quadratique pondérée définie par :

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

proposition

L'estimateur de Bayes δ^π associé à la loi a priori π et à la fonction coût quadratique est la moyenne a posteriori de θ . Cet estimateur est appelé Maximum Mean Squared Error (MMSE) donx :

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

Loi de x	Loi conjuguée	Moyenne a posteriori
Normale $\mathcal{N}(\theta, \sigma^2)$	Normale $\mathcal{N}(\mu, \tau^2)$	$\frac{\mu\sigma^2 + \tau^2x}{\sigma^2 + \tau^2}$
Poisson $\mathcal{P}(\theta)$	Gamma $\mathcal{G}(\alpha, \beta)$	$\frac{\alpha + x}{\beta + 1}$
Gamma $\mathcal{G}(v, \theta)$	Gamma $\mathcal{G}(\alpha, \beta)$	$\frac{\alpha + v}{\beta + x}$
Binomiale $\mathcal{B}(n, \theta)$	Beta $\mathcal{B}_e(\alpha, \beta)$	$\frac{\alpha + x}{\alpha + \beta + n}$
Binomiale négative $\mathcal{N}_{eg}(n, \theta)$	beta $\mathcal{B}_e(\alpha, \beta)$	$\frac{\alpha + n}{\alpha + \beta + x + n}$
Multinomoiale $\mathcal{M}_{kn}; \theta_1, \dots, \theta_k$	Dirichlet $\mathcal{D}(\alpha_1, \dots, \alpha_k)$	$\frac{\alpha_i + x_i}{(\sum_j \alpha_j) + n}$
Normale $\mathcal{N}(\mu, 1/\theta)$	Gamma $\mathcal{G}(\alpha/2, \beta/2)$	$\frac{\alpha + 1}{\beta + (\mu - x)^2}$

TABLE 1.1 – Estimateurs Bayésienne du paramètre θ sous coût quadratique pour les lois a priori conjuguées des familles exponentielles usuelles.

1.3.8 La fonction coût absolue

Définition 1.9 La fonction coût absolue est définie par :

$$L(\theta, \delta) = \begin{cases} \theta - \delta & \text{si } \theta > \delta \\ \delta - \theta & \text{si } \theta \leq \delta \end{cases} \Rightarrow L(\theta, \delta) = |\theta - \delta|.$$

proposition

L'estimateur de Bayes δ^π associé à la loi a priori π et à la fonction coût absolue est la médiane de $\pi(\theta|X)$.

1.3.9 La fonction coût 0-1

Définition 1.10 La fonction coût 0-1 est définie par :

$$L(\theta, \delta) = \begin{cases} 0 & \text{si } |\theta - \delta| < \varepsilon \\ 1 & \text{si } |\theta - \delta| \geq \varepsilon \end{cases} \quad \forall \varepsilon > 0.$$

proposition

L'estimateur de Bayésien δ^π associé à la loi a priori π et à la fonction coût 0-1 est le mode de la distribution $\pi(\theta|X)$, on l'appelle également Maximum A Posteriori (MAP).

Remarque

Le MAP est utilisé lorsque le MMSE est difficile à mettre en œuvre pour des raisons de calcul d'intégrales compliquées. Cependant le MAP n'est pas particulièrement recommandé mais il se comporte asymptotiquement comme le maximum de vraisemblance puisque $\pi(\theta)$ a un poids qui diminue lorsque $n \rightarrow \infty$.

1.3.10 La fonction coût entropique

L'utilisation d'une fonction coût quadratique $L(\theta, \hat{\theta})$ pour estimer le paramètre d'échelle peut être critiquée, car elle pénalise plus fortement la sur-estimation. Une fonction de perte alternative peut être définie sur la base du nombre d'informations de Kullback-Leibler. Kullback(1959) a décrit la distance entropique comme l'information moyenne de la fonction de vraisemblance $f(t, \theta)$ par rapport à $f(t, \hat{\theta})$, où $t = (t_1, t_2, \dots, t_n)$, et par conséquent, la fonction de perte peut être définie comme suit

$$L(\theta, \hat{\theta}) = E \left[\ln \frac{f(t, \hat{\theta})}{f(t, \theta)} \right]. \tag{1.3}$$

En conséquence, elle se réduit à la forme qui suit

$$L(\theta, \hat{\theta}) \propto \left(\frac{\hat{\theta}}{\theta} \right) - \ln \left(\frac{\hat{\theta}}{\theta} \right) - 1. \tag{1.4}$$

Cette fonction coût est connue sous le nom de fonction coût d'entropie et a été introduite pour la première fois par James et Stein(1961) pour estimer la matrice de variance-covariance (c'est-à-dire de dispersion) de la loi normale multivariée. Dey et al.(1987) ont considéré cette fonction coût comme une estimation conjointe des paramètres d'échelle et de leurs réciproques, pour p lois Gamma indépendantes. Rukhin et Ananda(1990) ont considéré le problème de l'estimation de la variance d'un vecteur normal multivarié sous la fonction coût d'entropie et la fonction coût quadratique. La fonction coût (2.16) a également été utilisée par de nombreux autres auteurs (Yang (1992), Wieczorkowski et Zielinski (1992)...etc.). Calabria et Pulcini(1996) ont défini la fonction coût par Entropie Générale comme suit

$$L(\theta, \hat{\theta}) \propto \left(\frac{\hat{\theta}}{\theta}\right)^{c_1} - c_1 \ln \left(\frac{\hat{\theta}}{\theta}\right) - 1. \quad (1.5)$$

La constante c_1 impliquée dans (2.18) est son paramètre de forme. Elle reflète l'écart par rapport à la symétrie. Lorsque $c_1 > 0$, la surestimation $\hat{\theta} > \theta$ est considérée comme plus grave que la sous-estimation de même ampleur et vice versa. Il est inutile de mentionner que la fonction coût (2.18) est une généralisation de la fonction coût d'entropie (2.17). L'estimateur Bayésien θ_{EG} de θ sous la fonction coût par Entropie Générale (2.18) est donné par

$$\hat{\theta}_{EG} = [E_{\theta}(\theta^{-c_1})]^{(-1/c_1)}, \quad (1.6)$$

Sous la condition condition que l'espérance $E_{\theta}(\theta^{-c_1})$ existe et soit finie. Ici, E_{θ} désigne l'espérance par rapport à la FDP. a posteriori de θ . Notez que si l'on met $c_1 = -1$ dans (2.19), on obtient l'estimateur Bayésien sous la fonction coût quadratique L , qui associe une importance égale aux pertes pour une surestimation et une sous-estimation de même ampleur.

Remarque

Les estimateurs bayésiens donnés par les formules..., Tendent asymptotiquement vers l'estimateur du maximum de vraisemblance ceci se justifie par le fait que lorsque $n \rightarrow \infty$, la taille de l'échantillon est tellement grande que l'information obtenue de l'échantillon l'emporte et prédomine sur l'information a priori. Pourquoi se contenter d'une information a priori lorsque l'on a l'information exacte de la population.

ESTIMATION DU PARAMÈTRE DE LA LOI DE POISSON ET NOTION DE ROBUSTESSE

2

2.1 INTRODUCTION

Dans ce chapitre nous utiliserons les notions de base présentées dans le chapitre 1 pour une estimation du paramètre θ d'une loi de Poisson, la loi a priori mobilisant l'information apportée par le paramètre est la loi Gamma. Et pour finir nous introduirons la robustesse et nous présenterons quelques résultats qui seront détaillées par la suite .

2.2 LA LOI DE POISSON

La loi de Poisson est une loi de probabilité discrète. Elle décrit la probabilité qu'un événement se réalise durant un intervalle de temps donné, lorsque la probabilité de réalisation d'un événement est très faible et que le nombre d'essais est très grand.

La loi de Poisson a été introduite en 1838 par Denis Poisson (1781–1840), dans son ouvrage Recherches sur la probabilité des jugements en matière criminelle et en matière civile. Le sujet principal de cet ouvrage consiste en certaines variables aléatoires qui dénombrent, entre autres choses, le nombre d'occurrences (parfois appelées « arrivées ») qui prennent place pendant un laps de temps donné.

2.3 MÉTHODES D'ESTIMATION D'UNE LOI DE POISSON

2.3.1 L'estimateur du maximum de vraisemblance (The Maximum Likelihood Estimator MLE)

L'estimateur du maximum de vraisemblance (MLE) est une méthode simple pour construire un estimateur pour le paramètre inconnu d'une distribution de probabilité en maximisant une fonction de vraisemblance. Elle peut être appliquée à une grande variété de problèmes statistiques et fournit un estimateur raisonnable et excellent pour le paramètre lorsque la taille de l'échantillon est grande.

Considérons un échantillon (X_1, X_2, \dots, X_n) de variable aléatoire indépendantes et identiquement distribuées d'une loi de Poisson de paramètre θ . La fonction de densité de X_i ; notée par $f(x_i|\theta)$ peut être écrite comme suit :

$$f(x_i|\theta) = \frac{e^{-\theta}\theta^{x_i}}{x_i!} \quad ; \quad x_i = 0, 1, 2, \dots; \theta > 0. \quad (2.1)$$

Où θ est le taux moyen de réalisations d'événements.

La fonction de masse de probabilité conjointe ou le produit de n termes est appelée fonction de vraisemblance, qui est définie comme suit :

$$\begin{aligned} L(x|\theta) &= \prod_{i=1}^n f(x_i|\theta) = f(x_1|\theta) \cdot f(x_2|\theta) \dots f(x_n|\theta) , \\ L(x|\theta) &= \frac{e^{\theta} \theta^{x_1}}{x_1!} \frac{e^{\theta} \theta^{x_2}}{x_2!} \dots \frac{e^{\theta} \theta^{x_n}}{x_n!} , \\ L(x|\theta) &= \frac{e^{-n\theta} \theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} . \end{aligned} \quad (2.2)$$

$$\begin{aligned} \ln L(x|\theta) &= \ln \left(\frac{e^{-n\theta} \theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \right) , \\ \ln L(x|\theta) &= -n\theta + \sum_{i=1}^n x_i \ln \theta - \ln \prod_{i=1}^n x_i! . \end{aligned} \quad (2.3)$$

A partir du logarithme de la fonction de vraisemblance, le maximum de $\ln L(X|\theta)$ se produit à la même valeur de θ que le maximum de $L(X|\theta)$. Si $\ln L(X|\theta)$ est différentiable par rapport à θ , les conditions nécessaires à l'apparition d'un maximum sont vérifiées en appliquant :

$$\frac{\partial \ln L(X|\theta)}{\partial \theta} = -n + \frac{\sum_{i=1}^n x_i}{\theta} = 0. \quad (2.4)$$

Comme nous ne savons pas s'il s'agit d'une valeur maximale ou minimale, la deuxième dérivée de l'estimateur peut être utilisée pour prouver que l'estimateur est le maximum et ceci lorsque la deuxième dérivée est inférieure à θ comme suit

$$\frac{\partial^2 \ln L(X|\theta)}{\partial \theta^2} = -\frac{\sum_{i=1}^n x_i}{\theta^2} < 0. \quad (2.5)$$

En résolvant la condition nécessaire comme mentionné ci-dessus, nous pouvons obtenir le MLE de θ en tant que

$$\hat{\theta}_{MLE} = \frac{\sum_{i=1}^n x_i}{n}. \quad (2.6)$$

2.3.2 L'estimation Bayésienne

Soit X_1, X_2, \dots, X_n des variables aléatoires issues d'une loi de Poisson. La fonction de masse de probabilité des variables aléatoires est donnée par $f(x_i|\theta)$ avec la fonction de vraisemblance $L(\theta) = \prod_{i=1}^n f(x_i|\theta)$. Considérons que l'information a priori apportée par θ soit immobilisée par la distribution Gamma de paramètre α et β , où α est le paramètre de forme et β est le paramètre d'échelle. Il est donnée par :

$$\pi(\theta|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}; \alpha, \beta, \theta > 0. \quad (2.7)$$

La loi a posteriori peut être obtenue en considérant la combinaison de la fonction de vraisemblance (2.2) et de la distribution a priori (2.7) comme suit :

$$f(x|\theta)\pi(\theta) = e^{-n\theta} \frac{\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \theta^{\alpha-1} e^{-\beta\theta} \frac{\beta^\alpha}{\Gamma(\alpha)},$$

$$f(x|\theta)\pi(\theta) = \frac{\beta^\alpha e^{-(n+\beta)\theta} \theta^{\sum_{i=1}^n x_i + \alpha - 1}}{\prod_{i=1}^n x_i! \Gamma(\alpha)}. \quad (2.8)$$

$$\int_0^{+\infty} f(x|\theta)\pi(\theta)d\theta = \int_0^{+\infty} \frac{\beta^\alpha e^{-(n+\beta)\theta} \theta^{\sum_{i=1}^n x_i + \alpha - 1}}{\prod_{i=1}^n x_i! \Gamma(\alpha)} d\theta,$$

$$\int_0^{+\infty} f(x|\theta)\pi(\theta)d\theta = \frac{\beta^\alpha}{\Gamma(\alpha) \prod_{i=1}^n x_i!} \int_0^{+\infty} \theta^{\alpha-1 + \sum_{i=1}^n x_i} e^{-\theta(n+\beta)} d\theta,$$

On pose :

$$\alpha' = \alpha + \sum_{i=1}^n x_i,$$

$$\beta' = \beta + n,$$

$$\int_0^{+\infty} f(x|\theta)\pi(\theta)d\theta = \frac{\beta^\alpha}{\Gamma(\alpha) \prod_{i=1}^n x_i!} \int_0^{+\infty} \theta^{\alpha'-1} e^{-\theta\beta'} \frac{\beta'^{\alpha'} \Gamma(\alpha')}{\Gamma(\alpha') \beta'^{\alpha'}} d\theta,$$

Et donc la loi a posteriori de θ est :

$$\Pi(\theta|X) = \frac{\beta^\alpha e^{-(n+b)\theta} \theta^{\sum_{i=1}^n x_i + \alpha - 1}}{\frac{\prod_{i=1}^n x_i! \Gamma(\alpha)}{\beta^\alpha \Gamma(\alpha')}} = \frac{\beta'^{\alpha'} \theta^{\alpha' - 1} e^{-\beta'}}{\Gamma(\alpha')},$$

cela implique que la loi a posteriori peut-être écrite comme :

$$\Pi(\theta|X) = \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \theta^{\alpha' - 1} e^{-\beta'}. \quad (2.9)$$

qui est une loi Gamma de paramètres α' et β' ; avec $\alpha' = \alpha + \sum_{i=1}^n x_i$ et $\beta' = \beta + n$
donc,

$$\theta|X \sim \text{Gamma}\left(\alpha + \sum_{i=1}^n x_i, \beta + n\right). \quad (2.10)$$

2.3.3 L'estimateur Bayésien du paramètre θ associé à la fonction coût quadratique

La fonction coût quadratique est définie par :

$$L(\theta; \hat{\theta}) = (\theta - \hat{\theta})^2, \quad (2.11)$$

où θ est l'estimateur Bayésien

L'estimateur Bayésien associé à la loi a priori $\Pi(\theta|X)$ et à la fonction coût quadratique est la moyenne a posteriori de θ et peut-être écrite comme suit;

$$\hat{\theta}_{BS} = E(\theta|x),$$

Dans le cas présent ceci représente l'espérance de la loi Gamma de paramètre $(\alpha + \sum_{i=1}^n x_i, \beta + n)$

Donc

$$\hat{\theta}_{BS} = \frac{\alpha + \sum_{i=1}^n x_i}{\beta + n}. \quad (2.12)$$

2.3.4 L'estimateur Bayésien du paramètre θ associé à la fonction coût quadratique généralisée

$$L(\theta; \hat{\theta}) = \left(\frac{\theta - \hat{\theta}}{\theta} \right)^2. \quad (2.13)$$

La fonction coût donnée par la formule (2.13) est la fonction coût quadratique généralisée

$$\hat{\theta}_{QG} = \frac{E \left[\frac{1}{\hat{\theta}} \right]}{E \left[\frac{1}{\theta^2} \right]} \quad (2.14)$$

L'estimateur Bayésien du paramètre θ associé à cette fonction coût est obtenu de la manière suivante ;

Commençons par le calcul de $E \left[\frac{1}{\theta} \right]$

$$\begin{aligned} E \left[\frac{1}{\theta} \right] &= \int_0^{+\infty} \frac{1}{\theta} \pi(\theta|X) d\theta \\ &= \int_0^{+\infty} \frac{1}{\theta} \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \theta^{\alpha'-1} e^{-\beta'\theta} d\theta \\ &= \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \int_0^{+\infty} \theta^{\alpha'-2} e^{-\beta'\theta} d\theta \\ &= \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \int_0^{+\infty} \theta^{(\alpha'-1)-1} e^{-\beta'\theta} d\theta \\ &= \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \int_0^{+\infty} \theta^{a-1} e^{-\beta'\theta} d\theta \end{aligned}$$

$$a = \alpha' - 1$$

$$\begin{aligned} &\frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \frac{\Gamma(a)}{\beta'^a} \int_0^{+\infty} \frac{\beta'^a}{\Gamma(a)} \theta^{a-1} e^{-\beta'\theta} d\theta \\ &= \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \frac{\Gamma(a)}{\beta'^a} \end{aligned}$$

$$\begin{aligned} E \left[\frac{1}{\theta^2} \right] &= \int_0^{+\infty} \frac{1}{\theta^2} \pi(\theta|X) d\theta \\ &= \int_0^{+\infty} \frac{1}{\theta^2} \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \theta^{\alpha'-1} e^{-\beta'\theta} d\theta \\ &= \frac{\beta'^{\alpha'}}{\Gamma(\alpha')} \int_0^{+\infty} \theta^{(\alpha'-2)-1} e^{-\beta'\theta} d\theta \end{aligned}$$

$$a' = \alpha' - 2$$

$$\begin{aligned} & \frac{\beta'^{\alpha'} \Gamma(a')}{\Gamma(\alpha') \beta'^{a'}} \int_0^{+\infty} \frac{\beta'^{a'}}{\Gamma(a')} \theta^{a'-1} e^{-\beta'\theta} d\theta \\ &= \frac{\beta'^{\alpha'} \Gamma(a')}{\Gamma(\alpha') \beta'^{a'}} \end{aligned}$$

$$\begin{aligned} \frac{E \left[\frac{1}{\theta} \right]}{E \left[\frac{1}{\theta^2} \right]} &= \frac{\frac{\beta'^{\alpha'} \Gamma(a)}{\Gamma(\alpha') \beta'^a}}{\frac{\beta'^{\alpha'} \Gamma(a')}{\Gamma(\alpha') \beta'^{a'}}} \\ &= \frac{\Gamma(a)}{\beta' \Gamma(a')} \\ &= \frac{\Gamma(\alpha' - 1)}{\beta' \Gamma(\alpha' - 2)} \end{aligned}$$

on a

$$\begin{aligned} \Gamma(\alpha' - 1) &= \Gamma(\alpha' - 1 + 1 - 1) \\ &= (\alpha' - 2) \Gamma(\alpha' - 2) \end{aligned}$$

$$\frac{E \left[\frac{1}{\theta} \right]}{E \left[\frac{1}{\theta^2} \right]} = \frac{\alpha' - 2}{\beta'}$$

L'estimateur Bayésien de θ sous une fonction coût quadratique généralisée est exprimé comme suit ;

$$\hat{\theta}_{QG} = \frac{\alpha - 2 + \sum_{i=1}^n x_i}{n + \beta}. \quad (2.15)$$

2.4 MÉTHODE DE MONTE-CARLO PAR CHAÎNES DE MARKOV (MCMC)

L'analyse Bayésienne traite tous les paramètres comme étant des variables aléatoires, elle attribue des distributions préalables pour caractériser la connaissance des valeurs des paramètres, et utilise la loi a posteriori compte tenu des observations. Pour caractériser les connaissances sur les valeurs des paramètres, et utilise la loi a posteriori compte tenu des données observées comme base d'inférence. Souvent la loi a posteriori est une distribution assez compliquée, comportant de nombreux paramètres. Pendant de nombreuses années, ce côté rébarbatif dans le calcul des intégrales a rendu la statistique Bayésienne moins attrayante, jusqu'à l'apparition des méthodes de simulation pour générer des échantillons à partir des lois a posteriori à savoir les méthodes de Monte Carlo par les chaînes de Markov (MCMC).

L'échantillonnage de Gibbs est une méthode MCMC populaire qui génère des valeurs qui évoluent toujours vers de nouvelles valeurs, et surtout plus important, ne nécessite pas de spécification des loi proposées.

On a effectué l'échantillonnage de Gibbs au moyen d'un logiciel appelé WinBUGS (Bayesian Inference Using Gibbs Sampling) introduit par Spiegelhalter et al. Nous utilisons les échantillons MCMC du paramètre obtenus via WinBUGS pour calculer des sommaires a posteriori approximatifs comme la loi a posteriori

Soit X_1, X_2, \dots, X_n des variables aléatoires iid suivant une loi de Poisson avec le paramètre λ , et soit λ une variable aléatoire de loi Gamma avec les paramètres a et b . Les paramètres estimés sont λ , a et b .

2.4.1 Algorithme

L'algorithme d'échantillonnage de Gibbs de la Méthode de Monte-Carlo par chaînes de Markov se déroule comme suit

1. Définir les valeurs initiales $a^{(t)}$ à partir d'une loi Exponentielle avec paramètres τ et $b^{(t)}$ de la loi Gamma avec le paramètre $(0,1)$. Remarquez que a, b sont les paramètres de la loi Gamma et que les valeurs de a et b sont supérieures à θ , ce qui correspond à la loi Exponentielle et à la loi Gamma.
2. Pour $t = 1, 2, \dots, T$ mettre à jour $a^{(t)}$ et $b^{(t)}$.

3. Générer $\lambda^{(t)}$ à partir de la fonction de la loi a posteriori basée sur la loi Gamma avec les paramètres $a^{(t)}$ et $b^{(t)}$ suivants 1.
4. Tracez la densité de la fonction de la loi a posteriori.
5. Calculez la moyenne, la médiane, et l'écart type de la fonction de la loi a posteriori.

Ainsi, l'estimateur MCMC est

$$\hat{\theta}_{MCMC} = \frac{1}{T} \sum_{t=1}^T \lambda^{(t)} \quad (2.16)$$

De plus, la méthode MCMC permet d'obtenir a et b et fait une approximation comme suit :

$$\hat{a}_{MCMC} = \frac{1}{T} \sum_{t=1}^T a^{(t)} \quad \text{et} \quad \hat{b}_{MCMC} = \frac{1}{T} \sum_{t=1}^T b^{(t)}$$

L'estimateur Bayésien est utilisé \hat{a}_{MCMC} et \hat{b}_{MCMC} pour calculer les paramètres λ comme suit :

$$\hat{\theta}_{MCMC} = \frac{\left(\frac{\sum_{i=1}^n x_i}{n} + \hat{a}_{MCMC} \right) \hat{b}_{MCMC}}{n \hat{b}_{MCMC} + 1} \quad (2.17)$$

2.5 LA ROBUSTESSE BAYESIENNE

La robustesse au sens de la stabilité a été explicitement définie dans Zielinski (1977) (voir aussi Zielinski 1983) et a été implicitement utilisée, par exemple, dans le passage suivant de Sivaganesan et Berger(1989) :

“ L’objectif d’une étude de robustesse est de trouver l’étendue (RANGE) de la quantité a posteriori , $\rho(x, \pi)$, qui nous intéresse, car π varie par rapport à Γ . Ainsi nous chercherons donc

$$\underline{\rho}(x, \pi) = \inf_{\pi \in \Gamma} \rho(x, \pi) \quad \text{et} \quad \bar{\rho}(x, \pi) = \sup_{\pi \in \Gamma} \rho(x, \pi). \quad (2.18)$$

Les quantités qui seront prises en compte comprennent la moyenne a posteriori, la variance a posteriori et la probabilité a posteriori d’un ensemble. Si le RANGE de la quantité a posteriori est petit, alors on peut être assuré de la robustesse par rapport au processus, à condition que Γ soit suffisamment large pour refléter l’incertitude possible de π_0 . Si le RANGE est large, il n’y a pas de robustesse par rapport à Γ ... ”

Le RANGE est l’étendue de valeurs de la quantité d’intérêt. Nous utilisons également le terme « oscillation ». Au lieu de la moyenne a posteriori, la variance a posteriori, ou encore la probabilité a posteriori, nous avons choisi le risque a posteriori car nous estimons que le risque est le principal point d’intérêt de l’estimation Bayésienne.

2.5.1 Quelques notions de base

Les différentes approches de la robustesse Bayésienne

Il existe dans la robustesse Bayésienne trois approches principales

Robustesse par rapport à la loi a priori ou approche informelle (prior robustness)

Dans ce cas la robustesse Bayésienne, consiste à construire une classe de lois a priori Γ , et étudier par la suite les changements effectués sur les quantité a posteriori autour de cette classe. La robustesse est réalisée s’il n’y a pas un grand changement entre les moyennes a posteriori sous les lois a priori, c’est-à-dire que le choix des lois a priori n’a pas d’influence.

Robustesse par rapport à la fonction de perte (loss robustness)

En général, un groupe de décideurs pourrait avoir différentes idées, différentes évaluations des conséquences de leur actions, ce qui entraîne différentes fonctions de pertes. Dans ce cas, il peut être nécessaire d'évaluer la robustesse des méthodes Bayésiennes au choix de la fonction de perte exactement de la même manière qu'au choix de la loi a priori : si une classe de fonctions de pertes est disponible, les changements effectués sur les quantités a posteriori peuvent être examinés.

Robustesse conjointe par rapport à la fonction de perte et à la loi a priori (loss and prior robustness)

La recherche sur la sensibilité conjointement par rapport à la loi a priori et à la fonction de perte n'est pas très abondante. En effet, il est possible que le problème soit robuste seulement par rapport à la loi a priori ou seulement par rapport à la fonction de perte, mais plutôt sensible lorsque les deux éléments sont considérés conjointement. Un aperçu de ces analyses générales de sensibilité est donné dans Rios Insua et al (2000). Basu et DasGupta (1995) considèrent conjointement les lois a priori dans une classe de distributions et un nombre fini de fonctions de perte. Martin et Rios Insua (1996) ont utilisé des dérivées de Fréchet pour étudier la sensibilité locale pour les petites perturbations à la fois dans la distribution a priori et dans la fonction de perte.

Définition 2.1 *le Range du coût a posteriori de l'action a par rapport à \mathcal{Q} est donnée par*

$$\left\{ \sup_{L \in \mathcal{Q}} \rho_L(\pi, x, a) - \inf_{L \in \mathcal{Q}} \rho_L(\pi, x, a) \right\} \quad (2.19)$$

Définition 2.2 *Une action a^* est ε -posteriori loss-robust par rapport à \mathcal{Q} si, pour tout $L \in \mathcal{Q}$,*

$$\Delta(a^*) = \left| \rho_L(\pi, x, a^*) - \inf_{a \in \mathcal{A}} \rho_L(\pi, x, a) \right| < \varepsilon \quad (2.20)$$

Définition 2.3 *minimax a posteriori : L_1 est préférable à L_2 si :*

$$\sup_x \inf_a \rho_{L_1}(\pi, x, a) < \sup_x \inf_a \rho_{L_2}(\pi, x, a) \quad (2.21)$$

Nous utiliserons la notation suivante. L'espace des actions sera désigné par \mathcal{A} et l'espace des paramètres par Θ . La perte liée à la réalisation de l'action A lorsque Θ sera désignée par $L(\theta, a)$, où \mathcal{Q} est une classe de fonctions de perte. En supposant tout au long que toutes les variables aléatoires sont continues, le coût a posteriori de l'action a par rapport à l'a priori π et à l'observation x sera désignée par

$$\rho_L(\pi, x, a) = E_{\pi(\cdot|x)} L(\theta, a) = \int_{\Theta} L(\theta, a) \pi(\theta|x) \quad (2.22)$$

où $\pi(\theta|x)$, la loi a posteriori de θ étant donné x , est donnée par $\pi(\theta)f(x|\theta)/m(x)$, tandis que $\pi(\theta)$ est la loi a priori de θ ; $f(x|\theta)$ est la fonction de vraisemblance de θ et $m(x)$ est la loi de x donnée par $\int_{\Theta} f(x)\pi(\theta)d\theta$. Par souci de simplification, nous supposerons tout au long de ce document que l'information sur l'échantillon est fournie par une seule donnée x .

ROBUSTESSE BAYÉSIENNE DU PARAMÈTRE DE LA LOI DE POISSON

3

3.1 INTRODUCTION

L'estimateur Bayésien du paramètre de la loi de Poisson de paramètre λ sous la loi a priori Gamma $g(\alpha_0, \beta_0)$ est stable (robuste) dans le sens où la loi a priori est comprise dans un intervalle : $((g(\alpha, \beta_0) \text{ avec } \alpha \in [\alpha_0 - \delta, \alpha_0 + \delta]) \text{ et } ((g(\alpha_0, \beta) \text{ avec } \beta \in [\beta_0 - \delta, \beta_0 + \delta]))$ alors l'oscillation du risque a posteriori est de l'ordre $O(\delta)$. Un estimateur avec l'oscillation $O(\delta^2)$ est construit; il minimise également l'oscillation du risque a posteriori lorsque le paramètre de forme est compris dans un intervalle fini

3.2 APPLICATION

Dans l'approche qui est appliquée par Marek Męczarski et Ryszard Zieliński (1991), il est possible de déterminer laquelle des deux procédures statistiques est la plus robuste et d'indiquer de manière raisonnable le problème de la construction de la procédure la plus robuste.

considérons un échantillon (X_1, X_2, \dots, X_n) de variable aléatoire indépendantes et identiquement distribuées d'une loi de Poisson de paramètre λ

$$(X_1, X_2, \dots, X_n) \sim \mathbf{P}(\lambda) \quad (3.1)$$

Et sa loi de probabilité est :

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (3.2)$$

Maintenant on considère que la loi a priori pour λ est une loi Gamma de paramètre (α_0, β_0) .

Et de là nous obtiendrons la loi a posteriori en considérant la combinaison de la fonction de vraisemblance et de cette distribution a priori. Les résultats seront comme suit :

$$\lambda \sim g(\alpha_0, \beta_0) \quad ; \quad \pi(\lambda) = \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \lambda^{\alpha_0-1} e^{-\beta_0 \lambda} \quad (3.3)$$

tel que $\alpha_0 > 0$, $\beta_0 > 0$

$$\pi(\lambda|X) = \frac{L(x_1, x_2, \dots, x_n|\lambda)\pi(\lambda)}{\int L(x_1, x_2, \dots, x_n|\lambda)\pi(\lambda)d\lambda}$$

$$L((x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i, \lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

cela implique que la loi a posteriori s'écrit comme

$$\pi(\lambda|x) = \frac{(n + \beta_0)^{\sum_{i=1}^n x_i + \alpha_0}}{\Gamma(\sum_{i=1}^n x_i + \alpha_0)} e^{-(n+\beta_0)\lambda} \lambda^{\sum_{i=1}^n x_i + \alpha_0 - 1}$$

qui est une loi Gamma de paramètres $\alpha_0 + \sum_{i=1}^n x_i$ et $\beta_0 + n$

$$\lambda|x \sim \text{Gamma}(\sum_{i=1}^n x_i + \alpha_0, n + \beta_0) \quad (3.4)$$

Il est bien connu que l'estimateur Bayésien de λ sous la fonction coût quadratique et la loi a priori Gamma $\Gamma(\alpha_0, \beta_0)$ est l'espérance de la loi a posteriori ,il est donné par :

$$\hat{\lambda}_B = \frac{S_n + \alpha_0}{n + \beta_0} \quad (3.5)$$

$$\text{avec } S_n = \sum_{i=1}^n x_i$$

Son coût a posteriori est sa variance qui s'écrit comme :

$$\rho(\pi, \delta) = \int_0^{+\infty} (\lambda - \delta)^2 \pi(\lambda|X) d\lambda = \text{var}(\lambda|x) = \frac{S_n + \alpha_0}{(n + \beta_0)^2}$$

Supposons maintenant que le paramètre α de la loi a priori $\Gamma(\alpha_0, \beta_0)$ est compris dans un intervalle $[\alpha_1, \alpha_2]$, $\alpha \in [\alpha_1, \alpha_2]$, $0 < \alpha_1 \leq \alpha_0 \leq \alpha_2 < \infty$, plutôt que d'être égal à α_0 . Alors la distribution a posteriori de λ est $g(\alpha + S_n, \beta_0 + n)$, sous la fonction coût quadratique, le risque a posteriori de $\hat{\lambda}_B$ est égal à :

$$\frac{S_n + \alpha}{(n + \beta_0)^2} + \frac{(\alpha - \alpha_0)^2}{(n + \beta_0)^2}. \quad (3.6)$$

En effet

$$\rho(\pi, \delta) = E[(\lambda - \hat{\lambda})^2] = E[\lambda^2] - 2\hat{\lambda}E[\lambda] + \hat{\lambda}^2$$

$$\text{sachant que } E[\lambda^2] = V[\lambda] + (E[\lambda])^2$$

$$\begin{aligned} E[(\lambda - \hat{\lambda})^2] &= \frac{S_n + \alpha_0}{n + \beta_0} + \frac{(\alpha - \alpha_0)^2}{(n + \beta_0)^2} - \frac{2(\alpha_0 + S_n)(\alpha + S_n)}{(n + \beta_0)^2} + \frac{(S_n + \alpha_0)^2}{n + \beta_0} \\ &= \frac{S_n + \alpha_0}{(n + \beta_0)^2} + \left(\frac{\alpha - \alpha_0}{\beta_0 + n} \right)^2 \end{aligned}$$

Donc

$$\rho(\pi, \delta) = E[(\lambda - \hat{\lambda})^2] = \frac{S_n + \alpha}{(n + \beta_0)^2} + \frac{(\alpha - \alpha_0)^2}{(\beta_0 + n)^2} \quad (3.7)$$

L'étape suivante consiste à calculer le Range ou l'oscillation de la formule (3.7). Pour ce faire, nous allons devoir trouver le max et le min de cette formule par rapport à α , nous ferons varier la valeur de α , dans l'intervalle $[\alpha_1, \alpha_2]$, considérons le coût a posteriori comme une fonction de α , Trouvons le max(3.7) et le min(3.7), pour cela étudions

$$f(\alpha) = (\alpha + S_n) + (\alpha - \alpha_0)^2$$

$$f(\alpha) = \alpha + S_n + \alpha^2 + \alpha_0^2 - 2\alpha\alpha_0$$

$$f'(\alpha) = 1 + 2\alpha - 2\alpha_0$$

avec $f'(\alpha) = 0$ nous aurons donc :

$$2\alpha + 2\alpha_0 + 1 = 0$$

$$\alpha = \alpha_0 - \frac{1}{2}$$

Mais tout dépendra où se trouve α_1 par rapport à $\alpha_0 - \frac{1}{2}$

α	$\alpha_0 - \frac{1}{2}$	α_2	
$f'(\alpha)$	-	0	+
$f(\alpha)$			

FIGURE 3.1 – Tableau de variation de $f(\alpha)$

Dans cette situation nous considérerons alors deux cas :

- $\alpha_1 \geq \alpha_0 - \frac{1}{2}$
- $\alpha_1 < \alpha_0 - \frac{1}{2}$

Le premier cas : Quand $\alpha_1 \geq \alpha_0 - \frac{1}{2}$

α	$\alpha_0 - \frac{1}{2}$	α_1	α_2
$f'(\alpha)$	-	0	+
$f(\alpha)$			

FIGURE 3.2 – Tableau de variation de $f(\alpha)$

$$\begin{aligned}
 \max(3.7) - \min(3.7) &= \frac{\alpha_2 + S_n}{(\beta_0 + n)^2} + \frac{(\alpha_2 - \alpha_0)^2}{(\beta_0 + n)^2} - \frac{\alpha_1 + S_n}{(\beta_0 + n)^2} + \frac{(\alpha_1 - \alpha_0)^2}{(\beta_0 + n)^2} \\
 &= \frac{\alpha_2 - \alpha_1}{(\beta_0 + n)^2} + \frac{(\alpha_2 - \alpha_0 - \alpha_1 + \alpha_0)(\alpha_2 - \alpha_0 + \alpha_1 - \alpha_0)}{(\beta_0 + n)^2} \\
 &= \frac{\alpha_2 - \alpha_1}{(\beta_0 + n)^2} + \frac{(\alpha_2 - \alpha_1)(\alpha_2 + \alpha_1 - 2\alpha_0)}{\beta_0 + n)^2} \\
 &= \frac{\alpha_2 - \alpha_1}{(\beta_0 + n)^2} [1 + \alpha_1 + \alpha_2 + -2\alpha_0] \\
 &= \frac{2(\alpha_2 - \alpha_1)}{(\beta_0 + n)^2} \left(\frac{\alpha_1 + \alpha_2 + 1}{2} - \frac{\alpha}{2} \right)
 \end{aligned}$$

Donc l'oscillation est égale à :

$$\frac{1}{(\beta_0 + n)^2} 2(\alpha_2 - \alpha_1) \left(\frac{\alpha_1 + \alpha_2 + 1}{2} - \alpha_0 \right) \quad \text{si } \alpha_1 \geq \alpha_0 - \frac{1}{2} \quad (3.8)$$

Le deuxième cas : Quand $\alpha_1 < \alpha_0 - \frac{1}{2}$

α	α_1	$\alpha_0 - \frac{1}{2}$	α_2
$f'(\alpha)$	-	0	+
$f(\alpha)$			

FIGURE 3.3 – Tableau de variation de $f(\alpha)$

Le min est atteint pour $\alpha = \alpha_0 - \frac{1}{2}$, quand au max il sera atteint pour $\alpha = \alpha_1$ ou $\alpha = \alpha_2$

$$\begin{aligned} \min(3) &= \frac{1}{(\beta_0 + n)^2} \left(\alpha_0 - \frac{1}{2} + S_n + \left(\alpha_0 - \frac{1}{2} - \alpha_0 \right)^2 \right) \\ &= \frac{1}{(\beta_0 + n)^2} \left(\alpha_0 + S_n - \frac{1}{4} \right) = \frac{\alpha_0 + S_n - \frac{1}{4}}{(\beta_0 + n)^2} \end{aligned}$$

Supposons que le max est atteint pour α_1 , il est égal à :

$$\frac{\alpha_0 + S_n}{(\beta_0 + n)^2} + \frac{(\alpha_1 - \alpha_0)^2}{(\beta_0 + n)^2}$$

L'oscillation est égale à :

$$\begin{aligned} \frac{\alpha_1 + S_n}{(\beta_0 + n)^2} + \frac{(\alpha_1 - \alpha_0)^2}{(\beta_0 + n)^2} - \frac{\alpha_0 + S_n - \frac{1}{4}}{(\beta_0 + n)^2} \\ \frac{1}{(\beta_0 + n)^2} \left(\alpha_1 - \alpha_0 + \frac{1}{2} \right)^2 \end{aligned} \quad (3.9)$$

Donc l'oscillation est égale à :

$$\begin{cases} \frac{1}{(\beta_0 + n)^2} 2(\alpha_2 - \alpha_1) \left(\frac{\alpha_1 + \alpha_2 + 1}{2} - \alpha_0 \right) & \text{si } \alpha_1 \geq \alpha_0 - \frac{1}{2} \\ \frac{1}{(\beta_0 + n)^2} \max \left\{ \left(\alpha_1 - \alpha_0 + \frac{1}{2} \right)^2, \left(\alpha_2 - \alpha_0 + \frac{1}{2} \right)^2 \right\} & \text{sinon} \end{cases} \quad (3.10)$$

Dans le cas quadratique, l'oscillation ne dépend pas de S_n .

Maintenant reprenons le même travail mais en supposant que le paramètre β de la loi a priori $\Gamma(\alpha_0, \beta_0)$ est compris dans un intervalle $[\beta_1, \beta_2]$, $\beta \in [\beta_1, \beta_2]$, $0 < \beta_1 \leq \beta_0 \leq \beta_2 < \infty$, plutôt que d'être égal à β_0 . Alors la distribution a posteriori de λ est $g(\alpha_0, \beta)$, sous la fonction coût quadratique, le risque a posteriori de $\hat{\lambda}_B$ est égal à :

$$\begin{aligned} E[(\lambda - \hat{\lambda})^2] &= E[\lambda^2] - 2\hat{\lambda}E[\lambda] + \hat{\lambda}^2 \\ &= E[\lambda^2] - 2\hat{\lambda}_{\alpha_0, \beta_0} E[\lambda]_{\pi(\alpha_0, \beta)} + \hat{\lambda}^2(\alpha_0, \beta_0) \\ &= E[\lambda^2] - 2 \frac{(\alpha_0 + S_n)}{(\beta_0 + n)} \frac{(\alpha_0 + S_n)}{(\beta + n)} + \frac{(\alpha_0 + S_n)^2}{(\beta_0 + n)^2} \\ E[\lambda^2] &= V(\lambda) + (E[\lambda])^2 = \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + \frac{(\alpha_0 + S_n)^2}{(\beta + n)^2} \end{aligned} \quad (3.11)$$

$$\begin{aligned} E[(\lambda - \hat{\lambda})^2] &= \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + \frac{(\alpha_0 + S_n)^2}{(\beta + n)^2} - 2 \frac{(\alpha_0 + S_n)^2}{(\beta_0 + n)(\beta + n)} + \frac{(\alpha_0 + S_n)^2}{(\beta + n)^2} \\ &= \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + \left[\frac{(\alpha_0 + S_n)}{(\beta + n)} - \frac{(\alpha_0 + S_n)}{(\beta_0 + n)} \right]^2 \\ &= \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + (\alpha_0 + S_n)^2 \left[\frac{1}{(\beta + n)} - \frac{1}{(\beta_0 + n)} \right]^2 \\ &= \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + (\alpha_0 + S_n)^2 \left[\frac{\beta_0 + n - \beta - n}{(\beta + n)(\beta_0 + n)} \right]^2 \\ &= \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + (\alpha_0 + S_n)^2 \left[\frac{(\beta - \beta_0)^2}{(\beta + n)^2(\beta_0 + n)^2} \right] \\ E[(\lambda - \hat{\lambda})^2] &= \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + \frac{(\alpha_0 + S_n)^2(\beta - \beta_0)^2}{(\beta + n)^2(\beta_0 + n)^2} \\ &= \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + \frac{(\alpha_0 + S_n)^2}{(\beta_0 + n)^2} \left[\frac{(\beta - \beta_0)}{(\beta + n)} \right]^2 \end{aligned}$$

Donc

$$\rho(\pi, \delta) = E [(\lambda - \hat{\lambda})^2] = \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + \hat{\lambda}^2 \frac{(\beta - \beta_0)^2}{(\beta + n)^2} \quad (3.12)$$

L'étape suivante consiste à calculer le Range ou l'oscillation de la formule (3.12). Pour ce faire, nous allons devoir trouver le max et le min de cette formule par rapport à β , nous ferons varier la valeur de α , dans l'intervalle $[\beta_1, \beta_2]$, considérons le coût a posteriori comme une fonction de β , Trouvons le max(3.12) et le min(3.12), pour cela étudions

$$g(\beta) = \frac{(\alpha_0 + S_n)}{(\beta + n)^2} + \hat{\lambda}^2 \frac{(\beta - \beta_0)^2}{(\beta + n)^2} \quad (3.13)$$

$$g'(\beta) = \frac{-2(\alpha_0 + S_n)}{(\beta + n)^3} + \frac{2\hat{\lambda}^2[(\beta - \beta_0)(\beta + n) - (\beta - \beta_0)^2]}{(\beta + n)^3}$$

$$\hat{\lambda}^2[(\beta + \beta_0)(\beta + n) - (\beta - \beta_0)^2] = \alpha_0 + S_n$$

$$(\beta - \beta_0)(\beta + n) - (\beta - \beta_0)^2 = \frac{(\alpha_0 + S_n)}{\hat{\lambda}^2}$$

$$(\beta - \beta_0) = \frac{\alpha_0 + S_n}{n + \beta_0} \frac{1}{\hat{\lambda}^2}$$

Avec $g'(\beta) = 0$ nous aurons donc :

$$\beta = \beta_0 + \frac{1}{\hat{\lambda}} \quad (3.14)$$

Tout comme la cas de α tout dépendra où se trouve β_2 par rapport à $\beta_0 + \frac{1}{\hat{\lambda}}$

Considérons 2 cas :

- $\beta_1 < \beta_2 \leq \beta_0 + \frac{1}{\hat{\lambda}}$
- $\beta_1 < \beta_0 < \beta_0 + \frac{1}{\hat{\lambda}} < \beta_2$

Le premier cas : Quand $\beta_1 < \beta_2 \leq \beta_0 + \frac{1}{\hat{\lambda}}$

β	β_1	β_0	β_2	$\beta_0 + \frac{1}{\lambda}$	
$g'(\beta)$		-		0	+
$g(\beta)$					

FIGURE 3.4 – Tableau de variation de $g(\beta)$

Le minimum est atteint par β_2 et le maximum est est atteint par β_1

$$\begin{aligned} & \max(3.12) - \min(3.12) \\ & \frac{(\alpha_0 + S_n)}{(\beta_1 + n)^2} + \hat{\lambda}^2 \frac{(\beta_1 - \beta_0)^2}{(\beta_1 + n)^2} - \frac{(\alpha_0 + S_n)}{(\beta_2 + n)^2} + \hat{\lambda}^2 \frac{(\beta_2 - \beta_0)^2}{(\beta_2 + n)^2} \\ & = (\alpha_0 + S_n) \left[\frac{1}{(\beta_1 + n)^2} - \frac{1}{(\beta_2 + n)^2} \right] + \hat{\lambda}^2 \left[\frac{(\beta_1 - \beta_0)^2}{(\beta_1 + n)^2} - \frac{(\beta_2 - \beta_0)^2}{(\beta_2 + n)^2} \right] \end{aligned}$$

L'oscillation est égale à :

$$\frac{(\alpha_0 + S_n)(\beta_2 - \beta_1)(\beta_2 + \beta_1 + 2n)}{(\beta_1 + n)^2(\beta_2 + n)^2} + \hat{\lambda}^2 \left[\frac{(\beta_1 - \beta_0)^2}{(\beta_1 + n)^2} - \frac{(\beta_2 - \beta_0)^2}{(\beta_2 + n)^2} \right] \quad (3.15)$$

Le deuxième cas $\beta_1 < \beta_0 < \beta_0 + \frac{1}{\lambda} < \beta_2$

β	β_1	$\beta_0 + \frac{1}{\lambda}$	β_2
$g'(\beta)$	-	0	+
$g(\beta)$			

FIGURE 3.5 – Tableau de variation de $g(\beta)$

Le min est atteint par $\beta_0 + \frac{1}{\hat{\lambda}}$ et le max est atteint soit par β_1 soit par β_2 .

$$g\left(\beta_0 + \frac{1}{\hat{\lambda}}\right) = \frac{(\alpha_0 + S_n)}{\left(\beta_0 + \frac{1}{\hat{\lambda}} + n\right)^2} + \hat{\lambda}^2 \frac{(\beta_0 + \frac{1}{\hat{\lambda}} - \beta_0)^2}{\left(\beta_0 + \frac{1}{\hat{\lambda}} + n\right)^2} = \frac{(\alpha_0 + S_n + 1)\hat{\lambda}^2}{(\alpha_0 + S_n + 1)^2}$$

$$g\left(\beta_0 + \frac{1}{\hat{\lambda}}\right) = \frac{\hat{\lambda}^2}{(\alpha_0 + S_n + 1)}$$

$$\max g(\beta) = \max g(\beta_i) = \frac{(\alpha_0 + S_n)}{(\beta_i + n)^2} + \hat{\lambda}^2 \frac{(\beta_i - \beta_0)^2}{(\beta_i + n)^2} \quad \text{avec } i = \overline{1,2}$$

Donc l'oscillation est égale à :

$$\begin{cases} \frac{(\alpha_0 + S_n)(\beta_2 - \beta_1)(\beta_2 + \beta_1 + 2n)}{(\beta_1 + n)^2(\beta_2 + n)^2} + \hat{\lambda}^2 \left[\frac{(\beta_1 - \beta_0)^2}{(\beta_1 + n)^2} - \frac{(\beta_1 - \beta_0)^2}{(\beta_2 + n)^2} \right], & \text{si } \beta_0 + \frac{1}{\hat{\lambda}} \geq \beta_2 \\ \max \left\{ \frac{(\alpha_0 + S_n)}{(\beta_1 + n)^2} + \hat{\lambda}^2 \frac{(\beta_1 - \beta_0)^2}{(\beta_1 + n)^2}, \frac{(\alpha_0 + S_n)}{(\beta_2 + n)^2} + \hat{\lambda}^2 \frac{(\beta_2 - \beta_0)^2}{(\beta_2 + n)^2} \right\}, & \text{sinon.} \end{cases} \quad (3.16)$$

L'oscillation dépend de S_n

si $\alpha_1 = \alpha_0 - \delta$ et $\alpha_2 = \alpha_0 + \delta$ alors, pour un petit δ , l'oscillation du risque a posteriori est :

$$\begin{aligned} & \frac{1}{(\beta_0 + n)^2} 2(\alpha_0 + \delta - \alpha_0 + \delta) \left(\frac{\alpha_0 - \delta + \alpha_0 + \delta + 1}{2} - \alpha_0 \right) \\ & = \frac{1}{(\beta_0 + n)^2} 2\delta \end{aligned} \quad (3.17)$$

qui est d'ordre $O(\delta)$.

cela signifie que l'estimateur Bayésien $\hat{\lambda}_B$ est infiniment stable sous de petits changements de paramètre de forme de la loi a priori $G(\alpha_0, \beta_0)$
 Dans ce qui suit, nous allons construire un estimateur dont le risque a posteriori est d'ordre $O(\delta^2)$.

L'estimateur Bayésien est une fonction linéaire de S_n . Considérons la classe L_∞ , de toutes les fonctions linéaires de S_n et essayons de trouver dans cette classe celle avec l'oscillation uniformément la plus petite. Sans perte de généralité, nous considérerons les estimateurs :

$$\hat{\lambda}_{A,B} = \frac{A + B \sum_{i=1}^n x_i}{\beta_0 + n}, A, B \in \mathbb{R}^1. \quad (3.18)$$

Si la loi a priori est $G(\alpha, \beta_0)$ alors le risque a posteriori de $\hat{\lambda}_{A,B}$

$$\begin{aligned}
 R((\alpha, \beta_0), \hat{\lambda}_{A,B}) &= E[\lambda^2]_{\pi(\alpha, \beta_0)} - 2\hat{\lambda}_{A,B}E[\lambda]_{\pi(\alpha_0, \beta)} + \hat{\lambda}^2(A, B) \\
 &= \frac{(\alpha + S_n)}{(\beta_0 + n)^2} + \frac{(\alpha + S_n)^2}{(\beta_0 + n)^2} - 2\frac{(A + BS_n)(\alpha + S_n)}{(\beta_0 + n)^2} + \frac{(A + BS_n)^2}{(\beta_0 + n)^2} \\
 &+ \frac{(\alpha + S_n)}{(\beta_0 + n)^2} + \left[\frac{\alpha + S_n - A - BS_n}{B_0 + n} \right]^2
 \end{aligned}$$

Donc

$$R((\alpha, \beta_0), \hat{\lambda}_{A,B}) = \frac{\alpha + S_n}{(\beta_0 + n)^2} + \left[\frac{(A - \alpha) + S_n(B - 1)}{B_0 + n} \right]^2 \quad (3.19)$$

L'oscillation (3.19) peut s'écrire aussi de cette façon

$$= \frac{1}{(\beta_0 + n)^2} \cdot \phi(\alpha) \quad (3.20)$$

avec

$$\phi(\alpha) = \alpha^2 - (2[A + (B - 1) \sum_{x=1}^n x_i] - 1)\alpha = \alpha(\alpha - 2[A + (B - 1) \sum_{x=1}^n x_i] + 1). \quad (3.21)$$

Minimiser l'oscillation de $\phi(\alpha)$ revient à minimiser l'oscillation de

$$\psi(\alpha) = \alpha^2 - (2[A + (B - 1) \sum_{x=1}^n x_i] - 1)\alpha = \alpha(\alpha - 2[A + (B - 1) \sum_{x=1}^n x_i] + 1). \quad (3.22)$$

Il est évident que l'oscillation de $\psi(\alpha)$ pour $\alpha \in [\alpha_1, \alpha_2]$ atteint son minimum si les centres des intervalles $[\alpha_1, \alpha_2]$ coïncident, l'oscillation minimale est égale à $[0, 2(A + (B - 1)S_n) - 1]$

$$\left(\frac{1}{2}(\alpha_2 - \alpha_1) \right)^2. \quad (3.23)$$

et ne dépend pas d'un choix spécifique de (A,B) sauf que

$$A + (B - 1)S_n = \frac{1}{2}(\alpha_1 + \alpha_2 + 1). \quad (3.24)$$

Tout choix de (A, B) satisfaisant (3.24) nous donne le même estimateur de λ que nous appellerons $\hat{\lambda}_\infty$:

$$\hat{\lambda}_\infty = \frac{\frac{1}{2}(\alpha_1 + \alpha_2 + 1) + S_n}{\beta_0 + n}. \quad (3.25)$$

Le risque a posteriori de l'estimateur (3.25) pour $\alpha \in [\alpha_1, \alpha_2]$ est donné par la formule suivante

$$R((\alpha, \beta_0), \hat{\lambda}_\infty) = \frac{1}{(\beta_0 + n)^2} \left[\alpha + S_n + \left(\frac{\alpha_1 + \alpha_2 + 1}{2} - \alpha \right)^2 \right]. \quad (3.26)$$

Dans le cas ou $\alpha_1 = \alpha_0 - \delta$ et $\alpha_2 = \alpha_0 + \delta$, alors de (3.25) nous avons

$$\begin{aligned}\hat{\lambda}_\infty &= \frac{\frac{1}{2}(\alpha_0 - \delta + \alpha_0 + \delta + 1) + S_n}{\beta_0 + n} \\ &= \frac{\frac{1}{2}(2\alpha_0 + 1) + S_n}{\beta_0 + n} \\ \hat{\lambda}_\infty &= \frac{\frac{1}{2} + \alpha_0 + S_n}{\beta_0 + n}\end{aligned}\tag{3.27}$$

Il s'ensuit que l'estimateur de Bayes λ n'est pas la limite des estimateurs les plus stables dans L_∞

pour expliquer cela, on observe que l'oscillation (3.23) et (3.26) quand α est compris dans l'intervalle $[\alpha_0 - \delta, \alpha_0 + \delta]$ d'ordre $O(\delta^2)$. Le prix à payer pour un estimateur plus stable est que son risque a posteriori soit plus grand par rapport à $\hat{\lambda}_B$ de $\frac{1}{4}(\beta_0 + n)^{-2}$.

L'estimateur $\hat{\lambda}_\infty$ a la plus petite oscillation dans la classe L_∞ de tous les estimateurs linéaires de λ ; il semble raisonnable de restreindre la classe L_∞ de tous les estimateurs linéaires dont le risque n'est pas plus grand que le risque du meilleur estimateur (c'est à dire l'estimateur Bayésien) plus un petit positif, et pour trouver l'estimateur le plus robuste dans cette classe. nous formalisons le problème comme suit.

Soit I_ε la classe d'estimateurs $\hat{\lambda}_{A,B}$ tels que le risque a posteriori de $\hat{\lambda}_{A,B}$ n'est uniformément pas plus grand que le risque a posteriori de $\hat{\lambda}_B$. Cela nous donne la condition suivante :

$$\left(\frac{(A - \alpha_0) + (B - 1)S_n}{\beta_0 + n} \right)^2 \leq \varepsilon\tag{3.28}$$

pour tout

$$S_n \in [0, +\infty)$$

Pour remplir cette condition, B doit être égal à 1, sinon le coté gauche de (3.28) croîtrait avec S_n jusqu'à l'infini. Par conséquent la condition (3.28) prend la forme suivante

$$B = 1 \quad \text{et} \quad \left(\frac{A - \alpha_0}{\beta_0 + n} \right)^2 \leq \varepsilon\tag{3.29}$$

CONCLUSION GÉNÉRALE

Dans sa définition, la robustesse est la qualité d'une méthode capable de donner des résultats d'une exactitude et d'une précision acceptables dans des conditions diverses. L'analyse Bayésienne robuste que l'on appelle aussi l'analyse de la sensibilité Bayésienne vise à quantifier et à interpréter l'incertitude induite par la connaissance partielle de l'information partielle ou/et de la fonction perte. L'objectif de cette analyse est d'aboutir à une décision optimale sous une fonction perte et une distribution a priori spécifiées sur l'espace des paramètres. L'étude de cette sensibilité se concentre principalement sur le calcul des changements apportés sur les quantités a posteriori, lorsque la loi a priori π varie dans une classe. Megarski et Zelinski (1991) se sont intéressés à l'estimateur Bayésien du paramètre de la loi de Poisson, ils ont mesuré la sensibilité de cet estimateur, lorsque le paramètre α de la loi a priori $Gamma(\alpha, \beta)$ varie à travers $[\alpha_1, \alpha_2]$. Si $\alpha_1 = \alpha_0 - \delta$, et $\alpha_2 = \alpha_0 + \delta$, alors pour un δ petit l'oscillation du risque a posteriori est $2(\beta_0 + n)^{-2}\delta$. Par la suite, les auteurs ont pu construire un estimateur stable et qui est Bayésien (c à d le coût a posteriori reste le plus petit). Nous avons repris le travail de Megarski et Zelinski, mais cette fois-ci en faisant varier β le paramètre d'échelle de la loi a priori. Cependant pour avoir cette stabilité, δ doit être supérieur à 1. Comme perspective, il serait intéressant de prendre une autre fonction coût plus adéquate et d'essayer de construire un estimateur Bayésien stable par rapport au variatiOn de β .

BIBLIOGRAPHIE

- [1] Robert, C. (2005). *Le choix bayésien : Principes et pratique*. Springer Science Business Media.
- [2] Stigler, S. (1986). *The History of Statistics*. Belknap, Cambridge.
- [3] Dale, A. (1991). *A History of Inverse Probability*. Springer-Verlag, New York.
- [4] Lad, F. (1996). *Operational Subjective Statistical Methods : a Mathematical, Philosophical and Historical Introduction*. John Wiley, New York.
- [5] Hald, A. (1998). *An History of Mathematical Statistics*. John Wiley, New York.
- [6] Stigler, S. (1986). *The History of Statistics*. Belknap, Cambridge.
- [7] Venn, J. (1886). *The Logic of Chance*. Macmillan, London.
- [8] Bertrand, J. (1889). *Calcul des Probabilités*. Gauthier-Villars, Paris.
- [9] Zabell, S. (1989). Fisher on the history of inverse probability. *Statist. Science*.
- [10] Keynes, J. (1921). *A Treatise on Probability*. Macmillan, London.
- [11] Fisher, R. (1912). On an absolute criterion for fitting frequency curves. *Messenger of Mathematics*.
- [12] Fisher, R. (1922). On the mathematical foundations of theoretical Statistics. *Philos. Trans. Roy. Soc. London*.
- [13] Fisher, R. (1930). Inverse probability. *Proc. Cambridge Philos.*
- [14] Seidenfeld, T. (1992). R.A. Fisher's fiducial argument and Bayes' theorem. *Statist. Science*.
- [15] Bernardo, J. (1979). Reference posterior distributions for Bayesian inference (with discussion). *J. Royal Statist. Soc. Series B*.
- [16] Bernardo, J. (1979). Reference posterior distributions for Bayesian inference (with discussion). *J. Royal Statist. Soc. Series B*.

-
- [17] Stein, C. (1959). An examination of wide discrepancy between fiducial and confidence intervals. *Ann. Statist.*
- [18] Wilkinson, G. (1977). On resolving the controversy in statistical inference. *J. Royal Statist. Soc.*
- [19] Zabell, S. (1992). Fisher and the fiducial argument. *Statist. Science.*
- [20] Pearson, K. (1894). Contribution to the mathematical theory of evolution. *Proc. Trans.*
- [21] Araveeporn, A. (2014). Parameter Estimation of Poisson Distribution by Using Maximum Likelihood, Markov Chain Monte Carlo, and Bayes method. *Science Technology Asia.*
- [22] Hassan, A., Ahmad, P. B., Bhatti, M. I. (2008). On the bayes estimator of parameter and reliability function of the zero-truncated poisson distribution. *Journal of the Korean Society for Industrial and Applied Mathematics.*
- [23] Howlader, H. A., Balasooriya, U. (2003). Bayesian estimation of the distribution function of the poisson model. *Biometrical Journal : Journal of Mathematical Methods in Biosciences.*
- [24] Takada, Y. (2001). Bayes sequential estimation of Poisson mean under a LINEX loss function. *Sequential Analysis.*
- [25] Lee, C. H., Hwang, L. C. (2011). Asymptotic optimal estimation of Poisson mean under LINEX loss function. *Communications in Statistics-Theory and Methods.*
- [26] Srivastava, U. (2012). Bayesian estimation of shift point in Poisson model under asymmetric loss functions. *Pakistan Journal of Statistics and Operation Research.*
- [27] Okasha, H. M., Wang, J. (2016). E-Bayesian estimation for the geometric model based on record statistics. *Applied Mathematical Modelling.*
- [28] Han, M. (2011). E-Bayesian estimation of the reliability derived from Binomial distribution. *Applied Mathematical Modelling.*
- [29] Han, M. (2017). The E-Bayesian and hierarchical Bayesian estimations of Pareto distribution parameter under different loss functions. *Journal of Statistical Computation and Simulation.*

-
- [30] Han, M. (2019). E-Bayesian estimation of the exponentiated distribution family parameter under LINEX loss function. *Communications in Statistics-Theory and Methods*.
- [31] Hasan, M. R., Baizid, A. R. (2017). Bayesian estimation under different loss functions using gamma prior for the case of exponential distribution. *Journal of Scientific Research*.
- [32] Hassan, M.R.(March). A Bayesian Approach for Estimating the Scale Parameter of Double Exponential Distribution under Symmetric and Asymmetric Loss Functions. *International Journal of Science and Research* 8.
- [33] Naji, L. F., Rasheed, H. A. (2019). Estimate the Two Parameters of Gamma Distribution Under Entropy Loss Function. *Iraqi Journal of Science*.
- [34] Okasha, H., Mustafa, A. (2020). E-Bayesian estimation for the Weibull distribution under adaptive type-I progressive hybrid censored competing risks data. *Entropy*.
- [35] Athirakrishnan, R. B., Abdul-Sathar, E. I. (2022). E-Bayesian and hierarchical Bayesian estimation of inverse Rayleigh distribution. *American Journal of Mathematical and Management Sciences*.
- [36] Yosboonruang, N., Niwitpong, S. A., Niwitpong, S. (2019). Measuring the dispersion of rainfall using Bayesian confidence intervals for coefficient of variation of delta-lognormal distribution : a study from Thailand. *PeerJ*.
- [37] Maneerat, P., Niwitpong, S. A., Niwitpong, S. (2020). A Bayesian approach to construct confidence intervals for comparing the rainfall dispersion in Thailand. *PeerJ*.
- [38] Thangjai, W., Niwitpong, S. A., Niwitpong, S. (2021). Bayesian confidence intervals for coefficients of variation of pm10 dispersion. *Emerging Science Journal*.
- [39] Okasha, H. M. (2014). E-Bayesian estimation for the Lomax distribution based on type-II censored data. *Journal of the Egyptian Mathematical Society*.

-
- [40] Muhammad, I. J. A. Z. (2021). Bayesian Estimation of the Shape Parameter of Lomax Distribution under Uniform and Jeffery Prior with Engineering Applications. *Gazi University Journal of Science*.
- [41] Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. *The Annals of Mathematical Statistics*.
- [42] Carlin, B. P., Gelfand, A. E. (1991). A sample reuse method for accurate parametric empirical Bayes confidence intervals. *Journal of the Royal Statistical Society : Series B (Methodological)*.
- [43] Li, C. P., Hao, H. B. (2019). E-Bayesian estimation and hierarchical Bayesian estimation of Poisson distribution parameter under entropy loss function. *IJAM*.
- [44] Zhang, Y. Y., Wang, Z. Y., Duan, Z. M., Mi, W. (2019). The empirical Bayes estimators of the parameter of the Poisson distribution with a conjugate gamma prior under Stein's loss function. *Journal of Statistical Computation and Simulation*.
- [45] Kullback, S. (1959). *Information theory and statistics*. John Wiley and Sons, Inc. New York.
- [46] Stein, C., James, W. (1961, June). Estimation with quadratic loss. In *Proc. 4th Berkeley symp. mathematical statistics probability*.
- [47] SRINIVASAN, C. (1987). Simultaneous estimation of parameters under entropy LOSS. *Journal of Statistical Planning and Inference*.
- [48] Rukhin, A. L., Ananda, M. M. (1992). Risk behavior of variance estimators in multivariate normal distribution. *Statistics probability letters*.
- [49] Yang, M. C. (1992). Ridge estimation of independent Poisson means under entropy loss. *Statistics Risk Modeling*.
- [50] Wiczorkowski, R., Zieliński, R. (1992). Minimax estimation of binomial probability with entropy loss function. *Statistics Risk Modeling*.
- [51] Calabria, R., Pulcini, G. (1996). Point estimation under asymmetric loss functions for left-truncated exponential samples. *Communications in Statistics-Theory and Methods*.

-
- [52] Geman, S. (1984). Gibbs distribution, and the Bayesian restoration of images. *IEEE Proc. Pattern Analysis and Machine Intelligence*.
- [53] Gelfand, A. E., Hills, S. E., Racine-Poon, A., Smith, A. F. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*.
- [54] Casella, G., George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*.
- [55] Spiegelhalter D., Thomas A., Best N., and Lunn D., WinBUGS User Manual, Version 1.4 , MRC Biostatistics Unit, Institute of Public Health and Department of Epidemiology and Public Health, Imperial College School of Medicine, UK, 2003.
- [56] R Development Core Team, R. (2004). A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, .
- [57] Huber, P.J. (1981). *Robust Statistics* (Wiley, New York).
- [58] Zieliński, R. (1977) Robustness ; a quantitative approach, *Bull. Acad. Polon. SC ;. Ser. Math. Astr. Phys.*
- [59] Zieliński, R. (1983). Robust statistical procedures : a general approach. In *Stability Problems for Stochastic Models*. Springer, Berlin, Heidelberg..
- [60] Sivaganesan, S., Berger, J. O. (1989). Ranges of posterior measures for priors with unimodal contaminations. *The Annals of Statistics*.
- [61] Supharakonsakun, Y. (2021). Bayesian approaches for Poisson distribution parameter estimation. *Emerging Science Journal*.
- [62] Męczarski, M., Zieliński, R. (1991). Stability of the Bayesian estimator of the Poisson mean under the inexactly specified gamma prior. *Statistics probability letters*.
- [63] Singh, S. K., Singh, U., Kumar, D. (2011). Bayesian estimation of the exponentiated gamma parameter and reliability function under asymmetric loss function. *REVSTAT–Stat J*.

Résumé

Le but du travail est d'étudier la robustesse de l'estimation Bayésienne du paramètre de la distribution de Poisson. Plusieurs critères seront utilisés pour le choix adéquat de la loi a priori et de la fonction coût à utiliser.

Abstract

The purpose of the work is to study the robustness of the Bayesian estimate of the Poisson distribution parameter. Several criteria will be used for the appropriate choice of the prior distribution and the loss function to be used.