



Faculté des Sciences Exactes et Informatique
Département de Mathématiques

Mémoire de fin d'études

Présenté pour l'obtention du diplôme de

Master

Spécialité : Mathématiques.

Option : Probabilités et Statistique.

Thème

*Estimation bayésienne : Simulation
numériques*

Présenté par :

BOUKERRA AMMAR

Devant le jury composé de :

Abdi Zineb	M.A.A Université de Jijel	Président
Boudjerda Khawla	M.C.B Université de Jijel	Encadreur
Ghaouil Djawida	M.A.A Université de Jijel	Examineur

Promotion **2020/2021**

Remerciements

Tout d'abord, nous remercions **Allah** le tout puissant pour son aide et pour nous avoir guidé pour mener à bien ce travail

La première personne que nous tenons à remercier est sont **les parents**

La deuxième personne que nous tenons à remercier est notre encadreur **Mme**

Boudjerda Khawla

Nos vifs remerciements vont à tous enseignants qui nous suivis nos cinq années d'études à l'université

Enfin, nous remercions toutes les personnes qui auraient contribué d'une manière ou d'une autre à la réalisation de ce travail.

Table des matières

1	Fondement de la statistique bayésienne	1
1.1	Introduction	1
1.2	Espace de la théorie de décision	2
1.3	Distribution a posteriori	4
1.4	Choix de la distribution a priori	5
1.4.1	Distribution a priori conjuguée	5
1.4.2	Distribution a priori impropre	6
1.4.3	Distribution a priori de Jeffrey	6
1.5	Fonctions de perte	7
1.5.1	Fonction de perte quadratique	7
1.5.2	Fonction de perte 0 – 1	8
1.5.3	Fonction de perte Linex	9
1.5.4	Fonction de perte de DeGroot	9
1.5.5	Fonction de perte d'entropie	10
1.6	Risque de Bayes	10
2	Méthodes numériques utiles dans le cadre bayésien	12
2.1	La Méthode de Lindley	12
2.2	La Méthode de Kadane et Tierney	13
2.3	Méthodes MCMC	14
2.3.1	Algorithme de Metropolis-Hastings	15
2.3.2	Echantillonneur de Gibbs	18
2.4	Méthodes PMC	19

2.4.1	Introduction	19
2.4.2	Méthode d'échantillonnage préférentiel	19
2.4.3	Algorithme PMC général	20
3	Applications	22
3.1	Modèle	22
3.1.1	La fonction de vraisemblance	24
3.2	Estimation Bayésienne sous la fonction de perte quadratique	26
3.3	Estimation Bayésienne sous la fonction de perte Linex	27
3.3.1	Simulation	28
	Bibliographie	40

Introduction générale

La statistique bayésienne est une approche statistique fondée sur l'inférence bayésienne, où la probabilité exprime un degré de croyance en un événement. Le degré de croyance peut être basé sur des connaissances a priori, telles que les résultats d'expérience antérieurs, ou sur des croyances personnelles concernant l'évènement. La perspective bayésienne diffère d'un certain nombre d'autres interprétations de la probabilité, comme l'interprétation fréquentiste qui considère la probabilité comme la limite de la fréquence relative d'un événement après de nombreux essais.

Les méthodes statistiques bayésiennes reposent sur le théorème de Bayes pour calculer et mettre à jour les probabilités après l'obtention de nouvelles données. Le théorème de Bayes décrit la probabilité conditionnelle d'un événement basée sur des informations ou des croyances antérieures sur l'évènement

Dans ce travail, on s'intéresse à l'estimation bayésienne des paramètres de la loi de Weibull à deux paramètres en utilisant un plan des données complètes et une loi a priori conjuguée naturelle sur les paramètres, les fonctions de perte utilisées sont une fonction de perte quadratique, puis une fonction de perte asymétrique celle de Linex.

Ce travail est réparti en trois chapitres : dans le premier chapitre on donne quelques notions de la statistique bayésienne, le deuxième chapitre présente les différentes méthodes numériques utiles dans le cadre de la statistique bayésienne telles que la méthode de Lindley, les méthodes MCMC et les méthodes PMC. Le dernier chapitre est une application sur la loi de Weibull à deux paramètres, et on a utilisé quelques méthodes numériques cités dans le deuxième chapitre pour calculer les estimateurs bayésiens des paramètres, ainsi que leurs risques a posteriori.

Fondement de la statistique bayésienne

1.1 Introduction

La statistique est un art interdisciplinaire de la quantification sous incertitudes utilisé par les physiciens, les économistes, les ingénieurs, les biologistes, les assureurs, les psy-chologues, les météorologues, etc. Tous les praticiens soucieux de bâtir, sur des fondations solides, un pont entre théorie et données expérimentales. Depuis un siècle, la statistique s'est considérablement développée, initiant une révolution dans les modes de pensée, car elle porte un langage de représentation du monde et de ses incertitudes.

C'est aujourd'hui une science mathématique dont l'objectif est de décrire ce qui s'est produit et de faire des projections quant à ce qu'il peut advenir dans le futur. Parfois, la situation peut être simplement décrite par quelques représentations graphiques d'analyse élémentaire des données. Bien souvent, le problème est beaucoup plus compliqué car de multiples facteurs d'influence doivent être pris en compte.

Schématiquement, on construit deux ensembles avec ces facteurs. Un premier paquet contient les facteurs dits explicatifs, bien identifiés, ceux dont on souhaite étudier l'influence en détail. En ce qui concerne le second paquet de facteurs, on ne sait, ou on ne veut pas, représenter leurs effets perturbateurs au cas par cas et, de ce fait, le jargon des modélisateurs le baptise sous le terme bruit, décrit alors de façon plus grossière par ses caractéristiques statistiques générales. Dans tous les cas, l'étude de la variabilité est au centre des débats : il s'agit d'abord de caractériser l'influence des facteurs identifiés et ensuite de représenter et d'évaluer le bruit résiduel dû à ces autres facteurs non pris en compte dans l'analyse de façon explicite. Dans une telle situation, le statisticien classique utilise à la fois un raisonnement déterministe par

l'absurde, afin de proposer des valeurs acceptables pour les paramètres décrivant les effets des facteurs explicatifs et un raisonnement probabiliste, pour traduire la variabilité des résultats observés due au bruit. Ce mode de pensée s'appuie sur l'hypothèse de la réalité objective des paramètres ainsi que sur l'interprétation de la probabilité comme limite des fréquences des résultats observés.

Par contre, le statisticien bayésien utilise le même cadre de pensée pour traiter par le pari probabiliste l'interaction de ces deux niveaux d'incertitudes : ignorance quant aux valeurs possibles des paramètres et l'aléatoire des bruits entachant les résultats expérimentaux. Choisir la piste bayésienne paraîtra à certains inutilement trop sophistiquée si on se limite aux modèles élémentaires (binomial, normal, etc.), pour ces cas d'école simples, l'approche fréquentiste est facile (nombreux logiciels), et offre au praticien des résultats souvent très proches de ceux que donnerait une analyse bayésienne avec une distribution a priori peu informative. Mais pour peu que l'analyste souhaite prendre à bras le corps des problèmes plus proches de son réel quotidien, apparaissent variables multiples, données manquantes, effets aléatoires, grandeurs latentes..., la structure des modèles de la vie scientifique moderne se présente sous une forme où des couches successives de conditionnement s'emboîtent, et pour lesquels l'approche bayésienne affirme sa véritable pertinence.

1.2 Espace de la théorie de décision

Nous abrègerons ici, la théorie bayésienne de la décision statistique, qui distingue trois espaces.

Le premier espace est celui des états, c'est-à-dire les valeurs vraies ou mesurandes relatives au système. Ces valeurs ne sont généralement pas connues directement, mais à travers le processus de mesure. Les états seront notés θ et leur espace Θ .

Le deuxième espace est celui des observations ou des résultats de mesure. Ces résultats seront notés x et leur espace \mathcal{X} .

Le troisième espace est celui des stratégies.

Théorème de Bayes :

Le but étant de trouver une probabilité sur les états, alors que l'on ne dispose au départ que d'une famille de probabilités sur les observations nous conduit à raisonner dans l'espace produit des observations.

$$\mathcal{X} \times \Theta$$

La théorie générale de l'intégration montre que la donnée d'une probabilité de transition $P(x|\theta)$ de Θ vers \mathcal{X} ne suffit pas à obtenir une probabilité sur le produit. Il faut encore une probabilité dite a priori $\pi(\theta)$. Dans ce cas, la probabilité conjointe sur le produit est définie par

$$P(x, \theta) = P(x|\theta)\pi(\theta)$$

La probabilité conjointe ainsi construite induit une marginale. En particulier, la marginale sur l'espace des observations sera nommée simplement marginale et sera définie par

$$M(x) = \int_{\Theta} P(x|\theta)\pi(\theta)d\theta$$

Supposons maintenant que la mesure donne l'observation x . Il est possible de considérer la partie de $\mathcal{X} \times \Theta$, et considérer la probabilité conditionnelle induite par la conjointe sur cette partie.

Alors, cette probabilité conditionnelle s'écrit :

$$\pi(\theta|x) = \frac{P(x, \theta)}{M(x)} = \frac{P(x|\theta)\pi(\theta)}{M(x)} = \frac{P(x|\theta)\pi(\theta)}{\int_{\Theta} P(x|\theta)\pi(\theta)d\theta}.$$

Cette probabilité conditionnelle est la probabilité a posteriori sur les états étant donné l'observation x .

Si le modèle est donné par une fonction de vraisemblance $L(x, \theta)$ et $\pi(\theta)$ par une densité $f(\theta)$, on obtient la formule classique

$$f(\theta|x) = \frac{\mathcal{L}(\theta|x)f(\theta)}{\int_{\theta} \mathcal{L}(\theta|x)f(\theta)d\theta}$$

Pour remonter de l'observation à l'état, la tentation est grande de lire cette fonction avec x constant (le résultat effectif de la mesure) et θ variable. Selon l'approche fréquentiste, le meilleur choix de l'état consiste à prendre celui qui maximise la fonction \mathcal{L} .

C'est le principe du maximum de vraisemblance

$$\hat{\theta} = \arg \min \mathcal{L}(\theta|x)$$

1.3 Distribution a posteriori

La distribution a posteriori [9] est la quantité la plus importante dans l'inférence bayésienne. Il contient toutes les informations disponibles sur le paramètre θ inconnu après avoir observé les données $X = x$.

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

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

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

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

Le dénominateur peut-être écrit comme :

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

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

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

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

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

La moyenne a posteriori $E(\theta|x)$ est l'espérance de la distribution a posteriori

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

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

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

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

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

1.4 Choix de la distribution a priori

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

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

1.4.1 Distribution a priori conjuguée

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

Définition 1.4.1. (*La distribution a priori conjuguée*)

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

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

Le tableau suivant donne quelques exemples des distributions a priori avec la fonction de vraisemblance correspondante :

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

Tab. 1.1 - **Quelques distributions a priori pour différentes fonctions de vraisemblance.**

1.4.2 Distribution a priori impropre

La distribution a priori a une influence sur la distribution a posteriori. Si on veut minimiser l'influence de la distribution a priori, il est courant de spécifier une a priori "vague", par exemple avec une très grande variance.

Dans la limite cela peut conduire à une distribution préalable mauvaise avec une fonction de densité qui n'intègre pas. En raison de la constante de normalisation manquante, ces fonctions de densité sont généralement spécifiées en utilisant le signe de proportionnalité " \propto ".

Si on utilise l'a priori impropre, alors il est nécessaire de vérifier que au moins la distribution a priori est impropre. Si ce n'est pas le cas, alors l'a priori impropre peut être utilisée dans l'analyse bayésienne.

Définition 1.4.2. (*Distribution a priori impropre*)

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

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

Ou

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

Pour un paramètre θ continue ou discret, respectivement.

1.4.3 Distribution a priori de Jeffrey

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

manquante.

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

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

$$\begin{aligned}\pi_\phi(\phi) &= \pi_\theta \{h^{-1}(\phi)\} \left| \frac{dh^{-1}(\phi)}{d\phi} \right| \\ &\propto \left| \frac{dh^{-1}(\phi)}{d\phi} \right|.\end{aligned}$$

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

Définition 1.4.3. (*La distribution a priori de Jeffrey*)

Soit X une variable aléatoire avec une fonction de vraisemblance $L(\underline{x}, \theta)$ où θ est le paramètre inconnu.

L'a priori de Jeffrey est définie comme suit

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

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

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

1.5 Fonctions de perte

1.5.1 Fonction de perte quadratique

La fonction de perte quadratique est la fonction définie par :

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

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

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

$\omega(\theta)$ est une fonction de θ .

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

$$\delta^\pi(x) = E_{\pi(\cdot|x)}(\theta)$$

1.5.2 Fonction de perte 0 – 1

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

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

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

On définit donc de la manière suivante :

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

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

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

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

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

Si X n'appartient pas dans W , on ne rejette pas, on prend la décision $\delta(x) = 1$, la décision est bonne $L(\theta, \delta(x)) = 0$, Le coût s'écrit donc :

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

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

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

On retrouve le risque de première espèce.

1.5.3 Fonction de perte Linex

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

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

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

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

Où $E_{\theta}(\cdot)$ représente l'espérance a posteriori relative à la densité a posteriori de θ . L'estimateur de Bayes $\delta_{\pi}(x)$ qui minimise (*). Pour trouver l'estimateur, nous dérivons l'équation (*) par rapport à $\delta(x)$, nous obtenons

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

En égalant cette expression à zéro, nous obtenons

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

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

$$\begin{aligned} \delta(x) &= -\frac{1}{r} \ln(E_{\theta}(e^{-r\theta})) \\ &= -\frac{1}{r} \ln \frac{\int e^{-r\theta} \pi(\theta|x) d\theta}{\int \pi(\theta|x) d\theta}. \end{aligned}$$

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

1.5.4 Fonction de perte de DeGroot

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

la fonction de perte de DeGroot définie par :

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

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

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

1.5.5 Fonction de perte d'entropie

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

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

qui a minimum lorsque $d = \theta$.

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

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

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

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

1.6 Risque de Bayes

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

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

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

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

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

Il s'agit d'une fonction de x .

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

Méthodes numériques utiles dans le cadre bayésien

Dans ce chapitre, on va citer quelques principes des méthodes numériques utiles dans le cadre bayésien, et spécialement les méthodes MCMC et PMC.

2.1 La Méthode de Lindley

Lindley (1980) a développé la procédure d'approximation [11], pour les intégrales de la forme

$$\frac{\int w(\theta) \exp\{l(\theta)\} d\theta}{\int v(\theta) \exp\{l(\theta)\} d\theta} \quad (2.1)$$

avec $\theta = (\theta_1, \theta_2, \dots, \theta_m)$, $l(\theta) = \log\{L(\theta|x)\}$ est le logarithme de fonction de vraisemblance, et $w(\theta)$, $v(\theta)$ sont des fonctions arbitraires de θ .

si $v(\theta)$ est la densité à priori de θ et $w(\theta) = \Phi(\theta)v(\theta)$, l'équation(2.1) donne l'espérance à posteriori de $\Phi(\theta)$; telle que :

$$E(\Phi(\theta)|x) = \frac{\int \Phi(\theta) \exp\{l(\theta) + p(\theta)\} d\theta}{\int \exp\{l(\theta) + p(\theta)\} d\theta} = \frac{\int \Phi(\theta) \exp\{\Lambda(\theta)\} d\theta}{\int \exp\{\Lambda(\theta)\} d\theta} \quad (2.2)$$

avec $p(\theta) = \log v(\theta)$ et $\Lambda(\theta) = \log\{\pi(\theta|x)\} = l(\theta) + p(\theta)$ est le logarithme de la distribution a posteriori de θ , il est évident que le maximum de $\Lambda(\theta) = \log\{\pi(\theta|x)\}$ nous donne le mode à postérieur de θ . Lindley (1980) a obtenu l'expression requise pour $E(\Phi(\theta)|x)$.

Prenons $S_t = \Phi(\theta)$, En utilisant la méthode de Lindley, l'estimateur bayésien pour S_t est :

$$S_t^* = \Phi(\hat{\theta}) + \frac{1}{2} \sum \Phi_{ij}(\hat{\theta})\tau_{ij} + \frac{1}{2} \sum \Lambda_{ijk}(\hat{\theta})\tau_{ij}\tau_{kl}, \quad (2.3)$$

Avec :

$$\Phi_{ij} = \frac{d^2\Phi}{d\theta_i d\theta_j}, \quad \Lambda_{ijk} = \frac{d^3\Phi}{d\theta_i d\theta_j d\theta_k}, \dots etc \quad (2.4)$$

les τ_{ij} sont les (i,j) ième éléments de l'inverse de la matrice Hessienne au signe négative. la matrice des second dérivées pour $\Lambda : \{\tau_{ij}\} = \{-\Lambda_{ij}\}^{-1}$, $\hat{\theta}$ est le mode a posteriori.

2.2 La Méthode de Kadane et Tierney

Le rapprochement de Lindley [11] exige l'évaluation des dérivés du tiers de la fonction vraisemblance ou de la densité a posteriori qui peut-être très fastidieux et exige une grande précision de calcul. Tierney et Kadane (1986) ont donné une méthode alternative d'évaluation du rapport des intégrales de la forme de l'équation (2.1) Soit, les deux expressions :

$$l = \frac{L(x; \theta) + \log(p(\theta))}{n}, l^* = \frac{\log \Phi(\theta) + \log v(\theta) + L(x; \theta)}{n} \quad (2.5)$$

Donc l'équation (2.1) prend cette forme :

$$E(\Phi(\theta)|x) = \frac{\int_{\Theta} \exp\{nl^*\} d\theta}{\int_{\Theta} \exp\{nl\} d\theta} \quad (2.6)$$

Pour plus de détails, voir Tierney et Kadane [1]. Alors que Lindley [2] élargit à la fois le numérateur et le dénominateur de l'équation (2.2) au sujet d'un point commun (le mode a posteriori), Tierney et Kadane (1986) ont développé chaque intégrale séparément sur le point qui maximise l'intégrale. Cette méthode ne nécessite que les dérivées premières et deuxièmes de

la densité a posteriori. D'après Tierney et Kadane (1986), l'estimateur de Bayes dans l'équation (2.6), dans le cas multi-paramétrique, prend la forme :

$$\hat{E}(\Phi(\theta)|x) = \left(\frac{|\Sigma^*|}{|\Sigma|} \right)^{\frac{1}{2}} \exp \left[n \left\{ l^*(\hat{\theta}^*) - l(\hat{\theta}) \right\} \right] \dots\dots\dots(**)$$

ou $\hat{\theta}^*$ et $\hat{\theta}$ maximisent l^* et l respectivement, et Σ^* et Σ sont les inverses de l'Hessiennes de l^* et l au signe négative pour $\hat{\theta}^*$ et $\hat{\theta}$ respectivement.

donc Σ :

$$\Sigma = \begin{pmatrix} \frac{-d^2l}{d\alpha^2} & \frac{-d^2l}{d\alpha d\beta} \\ \frac{-d^2l}{d\alpha d\beta} & \frac{-d^2l}{d\beta^2} \end{pmatrix} \tag{2.7}$$

2.3 Méthodes MCMC

Les méthodes MCMC [10] sont publiée en 1953 par Metropolis et ses coauteurs [3], elle est étendue en 1970 par Hastings[4]. En 1984, les frères Geman proposent l'échantillonneur de Gibbs pour la restauration bayésienne d'images, Cet échantillonneur est développé par Gelfand et Smith[5]. C'est au début des années 90, après le développement et la démocratisation de l'outil informatique, que ces outils rencontrent un important succès [6].

En statistique bayésienne, θ est considéré comme un vecteur aléatoire de densité $\pi(\theta)$ (loi a priori sur θ). L'estimation bayésienne est basée sur le calcul de la loi a posteriori de θ dont la densité est notée $\pi(\theta|y)$. Les estimateurs bayésiens sont de la forme $E_\pi(h(\theta))$ où E_π est l'opérateur des espérances mathématiques pour la loi $\pi(\theta|y)$.

$$I_h = E_\pi(h(\theta)) = \int_{\Theta} h(\theta)\pi(\theta|y)d\theta. \tag{2.8}$$

Souvent, le calcul explicite, de la densité $\pi(\theta|y)$ n'est pas envisageable ; ainsi de nombreuses méthodes d'approximation ont été proposées et la plus utilisé est celle de MCMC dont son principe est de construire une chaîne de Markov ergodique de réalisations $\theta^{(1)}, \dots, \theta^{(t)}, \dots$

, qui a pour distribution stationnaire $\pi(\theta|y)$; ainsi l'estimateur qui sera obtenu est \widehat{I}_h :

$$\widehat{I}_h = \frac{1}{T} \sum_{t=1}^T h(\theta^{(t)}). \quad (2.9)$$

La convergence de \widehat{I}_h vers I_h est assurée par l'ergodicité de la chaîne de Markov. Parmi les méthodes MCMC, l'algorithme de Metropolis-Hastings (Hastings 1970) et l'algorithme de Gibbs (Geman et Geman 1984) sont les plus utilisées et ont donné lieu à de nombreux algorithmes dérivés.

2.3.1 Algorithme de Metropolis-Hastings

Principe

Le principe est de construire une chaîne de Markov ergodique de loi stationnaire $\pi(\theta|y)$ partant de $\theta^{(0)}$, on génère une chaîne $\theta^{(t)}$ à partir d'un noyau de transition de loi stationnaire $\pi(\theta|y)$, qui garantit de plus la convergence en loi vers $\pi(\theta|y)$. Pour T "assez grand" on peut considérer $\theta^{(T)}$ comme distribué suivant $\pi(\theta|y)$ et obtenir ainsi un échantillon $\theta^{(T)}$ qui est effectivement distribué suivant $\pi(\theta|y)$.

Algorithme 1

L'algorithme de Metropolis-Hastings consiste à simuler un échantillon selon une distribution objectif $\pi(\theta|y)$, à partir d'une distribution conditionnelle $q(y|\theta)$. Pour la mise en œuvre de l'algorithme, on doit pouvoir simuler facilement à partir de la distribution $q(\cdot|\theta)$, et sa densité $q(y|\theta)$ doit être disponible analytiquement (au moins à une constante multiplicative près) ou à défaut, le rapport $\frac{\pi(\theta|y)}{q(y|\theta)}$ doit être fini.

- **Étape 1** Pour $t = 0$, on initialise $\theta^{(0)}$ le premier élément de la chaîne ;
- **Étape 2** Pour $t \in \{1 \dots T\}$;
 - On génère $u \sim U_{[0,1]}$;
 - On génère y_t selon $q(\cdot|\theta^{(t-1)})$;

- On calcul $\rho(\theta^{(t)}, y_t) = \text{Min} \left\{ 1, \frac{\pi(y_t) q(\theta^{(t)}|y_t)}{\pi(\theta^{(t)}) q(y_t|\theta^{(t)})} \right\}$

- Puis, on prend

$$\theta^{(t+1)} = \begin{cases} y_t, & \text{avec la probabilité } \rho(\theta^{(t)}, y_t) \text{ si } u < \rho(\theta^{(t)}, y_t) \\ \theta^{(t)}, & \text{avec la probabilité } 1 - \rho(\theta^{(t)}, y_t) \text{ sinon} \end{cases}$$

- **Étape 3** $t=t+1$ et retourner à l'étape 2;

- **Étape 4** finalement calculer l'estimateur de Bayes $\hat{I}_h = \frac{1}{T} \sum_{t=1}^T h(\theta^{(t)})$.

Algorithme 2 (Metropolis-Hastings indépendant)

L'algorithme repose sur l'utilisation de la loi instrumentale $q(y|\theta^{(t)})$ indépendamment de $\theta^{(t)}$, c'est une généralisation de l'algorithme d'acceptation-rejet.

1. Générer y_t selon $q(y)$;
2. Prendre

$$\theta^{(t+1)} = \begin{cases} y_t, & \text{avec la probabilité } \text{Min} \left\{ 1, \frac{\pi(y_t) q(\theta^{(t)})}{\pi(\theta^{(t)}) q(y_t)} \right\} \\ \theta^{(t)}, & \text{sinon} \end{cases}$$

Algorithme 3 (Metropolis-Hastings à marche aléatoire)

Cette variante prend en compte la valeur précédemment générée pour simuler la suivante, la loi instrumentale $q(y|\theta)$ sera ainsi écrite $q(y - \theta) = q(\theta - y)$, c'est-à-dire que y_t peut s'écrire $\theta^{(t)} + \epsilon_t$, tel que ϵ_t étant une perturbation aléatoire de q , indépendante de $\theta^{(t)}$, ainsi la chaîne de Markov générée associée à q est une marche aléatoire.

le principe de cet algorithme est le suivant :

1. Étant donné $\theta^{(t)}$

2. Générer $y_t \sim q(y - \theta^{(t)})$.

3. On calcul $\rho(\theta^{(t)}, y_t) = \text{Min} \left\{ 1, \frac{\pi(y_t) q(\theta^{(t)} - y_t)}{\pi(\theta^{(t)}) q(y_t - \theta^{(t)})} \right\}$

4. Poser

$$\theta^{(t+1)} = \begin{cases} y_t, & \text{avec la probabilité } \rho(\theta^{(t)}, y_t) \\ \theta^{(t)}, & \text{avec la probabilité } 1 - \rho(\theta^{(t)}, y_t) \end{cases}$$

Algorithme 4 (Metropolis-Hastings à une variable à la fois)

Quand le paramètre à simuler est de dimension grande, on est contraint de trouver une densité instrumentale multidimensionnelle engendrant une chaîne ayant le comportement d'une chaîne de Markov. Pour cela on peut utiliser un algorithme dit à une variable à la fois.

Le principe est de simuler les composantes $(\theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_d^{(t)})$ une par une. À chaque itération de l'algorithme, on fait évoluer d composantes $\theta_i^{(t)}$ en utilisant d étapes de l'algorithme Metropolis-Hastings, ce qui signifie que pour obtenir le nouveau vecteur $\theta^{(t+1)}$ il faudra utiliser d densités instrumentales $q_i(\cdot|\cdot)$.

Pour simuler une composante $\theta_i^{(t)}$ il faut utiliser la loi instrumentale $q_i(y_i|\theta_{-i}^{(t+1)})$ et la loi cible $\pi_i(\theta_i|\theta_{-i}^{(t+1)})$, où : $\theta_{-i}^{(t+1)} = (\theta_{-i}^{(t+1)}, \dots, \theta_{i-1}^{(t+1)}, \theta_{i+1}^{(t+1)}, \dots, \theta_d^{(t+1)})$ (tel que $i = 1, \dots, d$).

On aura ainsi la probabilité d'acceptation de l'étape i :

$$\rho_i(\theta_i^{(t)}, \theta_{-i}^{(t+1)}, y_i) = \text{Min} \left\{ 1, \frac{\pi_i(y_i|\theta_{-i}^{(t+1)}) q_i(\theta_i^{(t)}|y_i, \theta_{-i}^{(t+1)})}{\pi_i(\theta_i^{(t+1)}|\theta_{-i}^{(t+1)}) q_i(y_i|\theta_i^{(t)}, \theta_{-i}^{(t+1)})} \right\}$$

1. MH $(\pi_1(\theta_1|\theta_{-1}^{(t+1)}), q_1)$
2. MH $(\pi_2(\theta_2|\theta_{-2}^{(t+1)}), q_2)$
3. ...
4. MH $(\pi_d(\theta_d|\theta_{-d}^{(t+1)}), q_d)$

Propriétés

Les algorithmes de MH ne génèrent pas d'échantillon indépendant et identiquement distribués, en particulier parce que la probabilité d'acceptation de y_t dépend de $\theta^{(t)}$.

Remarque

Les algorithmes de Metropolis-Hastings peuvent théoriquement être utilisés pour simuler un

vecteur aléatoire de dimension p en utilisant une densité instrumentale multidimensionnelle. Mais quand p est grand, ce choix est rarement fait en pratique car la convergence d'un tel algorithme serait extrêmement lente. En effet, plus la dimension de l'espace des paramètres est grande, et plus la proportion de candidats rejetés est importante. On préférera utiliser l'algorithme de Gibbs comme cadre général.

2.3.2 Echantillonneur de Gibbs

Plus généralement, si l'on peut écrire $\theta_1, \dots, \theta_p$ une partition du vecteur des paramètres θ pour le modèle étudié, et si l'on peut spécifier complètement les lois de comportement a posteriori conditionnelles alors on peut facilement simuler, pas par pas, des réalisations conditionnelles de $\theta_1, \dots, \theta_p$. En itérant le procédé un grand nombre de fois, la chaîne de Markov produite par ces simulations répétées a pour distribution stationnaire $\pi(\theta|y)$.

Algorithme 5

1. Initialiser $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_p^{(0)})$

2. A l'itération i : simuler

$$\theta_1^{(i+1)} \sim \pi_1(\theta_1 | \theta_2^{(i)}, \dots, \theta_p^{(i)})$$

$$\theta_2^{(i+1)} \sim \pi_2(\theta_2 | \theta_1^{(i+1)}, \dots, \theta_p^{(i)})$$

$$\theta_p^{(i+1)} \sim \pi_p(\theta_p | \theta_1^{(i+1)}, \dots, \theta_{p-1}^{(i+1)})$$

3. $i \rightarrow i + 1$ et aller en 2.

Les densités conditionnelles π_i sont appelées conditionnelles complètes.

Propriétés :

- Taux d'acceptation égal à 1. Toutes les valeurs simulées sont acceptées.
- Nécessite de connaître les lois conditionnelles de π , d'où une connaissance préalable de certaines propriétés probabilistes ou analytiques de π .

Difficultés des méthodes MCMC

La méthode MCMC peut présenter des difficultés importantes [7] à savoir :

- La convergence : c'est-à-dire l'atteinte de l'équilibre de la chaîne de Markov peut être très lente, surtout lorsque $\pi(\theta|y)$ est difficile à approximer par une densité de transition q facilement

simulable.

-Par ailleurs, même si il ya la convergence, détecter le "temps de chauffe" est loin d'être évident.

-MCMC est couteuse en temps de calcul.

2.4 Méthodes PMC

2.4.1 Introduction

Population Monte Carlo est le très récent et puissant algorithme introduit initialement par Cappé et al. en (2004) pour traiter le problème de reconstruction de signal dans un canal-ionique ; puis amélioré par Douc et al. en (2005) et très récemment Celeux et al. en (2006) ont également mis à profit cette méthode pour étudier des modèles à données manquantes [7].

PMC est une amélioration du schéma d'échantillonnage préférentiel classique en lui introduisant une dimension itérative assurant l'adaptation de la loi instrumentale $q(y|\theta)$ à la densité cible $\pi(\theta|y)$. Cette adaptation est implémentée de façon séquentielle [7] : au pas t de l'algorithme, M réalisations de θ sont simulées à partir d'une densité instrumentale courante dépendante des M réalisations du pas $t - 1$. Ainsi l'estimateur \hat{I}_h^T asymptotiquement non biaisé de (3.8) sera comme suit :

$$\hat{I}_h^T = \frac{1}{T} \sum_{t=1}^T \sum_{j=1}^M w_t^{(j)} h(\theta_t^{(j)}), \quad (2.10)$$

où $w_t^{(j)}$ est le poids normalisé de la particule j à la date t obtenu à partir de l'algorithme qui sera introduit plus tard.

2.4.2 Méthode d'échantillonnage préférentiel

Les algorithmes d'échantillonnage préférentiel, ou d'importance, sont des méthodes de simulation de Monte-Carlo qui, à partir d'un échantillon de variables aléatoires indépendantes et identiquement distribuées $\theta_1, \dots, \theta_M$ provenant d'une densité instrumentale ou d'importance $q(\theta)$, [7] proposent d'estimer l'intégrale :

$$I_h = \int h(\theta) \pi(\theta|y) d\theta; \quad (2.11)$$

par

$$\hat{I}_h = \sum_{i=1}^M w_i h(\theta_i); \quad (2.12)$$

où

$$w_i = \frac{r_i}{\sum_{j=1}^M r_j} \quad \text{et} \quad r_i = \frac{\pi(y|\theta_i)\pi(\theta_i)}{q(\theta_i)}. \quad (2.13)$$

Où $\pi(y|\theta_i)$ est la vraisemblance.

Ce type de méthodes peut apparaître fragiles dans des cas pratiques, l'estimateur \hat{I}_h étant rendu fortement dépendant des poids normalisés w_i .

Tout l'art de l'échantillonnage préférentiel consiste donc à choisir une densité q qui doit être proche de la loi cible.

Différentes méthodes de construction de q ont été proposées. On peut les séparer en deux catégories à savoir :

1. Les méthodes statiques, où $q(\theta)$ est déterminée de façon unique en préalable à la simulation.
2. Les méthodes adaptatives, c'est la méthode PMC.

2.4.3 Algorithme PMC général

Fondamentalement, [8] l'algorithme PMC général consiste en l'itération de deux étapes :

1. D'abord, on simule une population de n variables aléatoires (parfois dites particules) suivant une loi choisie a priori.
2. On calcule le poids de chaque particule suivant (2.13) et on retire un échantillon dans la population générée proportionnellement aux poids calculés.

Soit q_{it} ($i=1,\dots,M$ et $t=1,\dots,T$) une distribution d'importance. La représentation algorithmique qu'on va présenter correspond à ces deux étapes.

principe d'algorithme PMC général :

1. **Étape 0** : Choix de $(\theta_0^{(1)}, \dots, \theta_0^{(M)})$;
2. **Étape t** ($t=1,\dots,T$) :

Pour $i=1,\dots,M$:

- Générer $\theta_t^{(i)} \sim q_{it}(\theta)$ et calculer $r_t^{(i)} = \frac{\pi(y|\theta_t^{(i)})\pi(\theta_t^{(i)})}{q_{it}(\theta_t^{(i)})}$;
- Calculer $\omega_t^{(i)} = r_t^{(i)} / \sum_{k=1}^M r_t^{(k)}$ et ré-échantillonner les $\theta_t^{(i)}$ en utilisant les poids $\omega_t^{(i)}$;
- Construire $q_{i(t+1)}$ à partir de l'échantillon courant ;

3. Après avoir itéré T fois l'algorithme, un estimateur asymptotiquement non biaisé de l'intégrale (2.8) est donné dans la formule (2.10).

Applications

Introduction

La loi de Weibull nommée d'après walodi weibull en 1951, est une loi de probabilité continue, elle est un cas spécial de loi d'extremum généralisée au même titre que la loi de Gumbel ou la loi de Fréchet.

La distribution de Weibull est souvent utilisée dans le domaine de l'analyse de la durée de vie, grâce à sa flexibilité : comme dit précédemment, elle permet de représenter au moins approximativement une infinité de loi de probabilité.

3.1 Modèle

Dans cette partie nous présentons les cinq fonctions équivalentes de la loi de Weibull à deux paramètres α et β

Fonction de densité f :

Une variable aléatoire continue X suit une loi de Weibull de paramètres α et β , si elle admet pour densité de probabilité la fonction :

$$f(x; \alpha, \beta) = \left(\frac{\beta}{\alpha}\right) \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta}.$$

Où

- $\alpha > 0$: le paramètre d'échelle.
- $\beta > 0$: le paramètre de la forme.

Fonction de répartition F :

On dit que F est une fonction de répartition de la variable aléatoire X qui suit une loi de Weibull à deux paramètres α, β , si F est donnée par :

$$F(x; \alpha, \beta) = \left(1 - e^{-\left(\frac{x}{\alpha}\right)^\beta} \right), \quad \alpha, \beta, x > 0$$

Fonction de fiabilité R :

Soit X une variable aléatoire continue de loi de Weibull de fonction de répartition F et de densité de probabilité f . Sa fonction de fiabilité est définie par :

$$R(x) = 1 - F(x) = e^{-\left(\frac{x}{\alpha}\right)^\beta}$$

Fonction de taux de panne (hasard) h :

La fonction de hasard est définie comme la probabilité conditionnelle que le phénomène se termine après une durée x sachant que l'on a atteint cette durée (taux de panne, taux de défaillance, taux de décès ou risque instantané). La fonction de taux de hasard h est calculée à l'aide de la formule suivante :

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{R(x)},$$

donc :

$$h(x) = \frac{\left(\frac{\beta}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta}}{e^{-\left(\frac{x}{\alpha}\right)^\beta}}$$

Fonction de taux de hasard cumulé H :

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

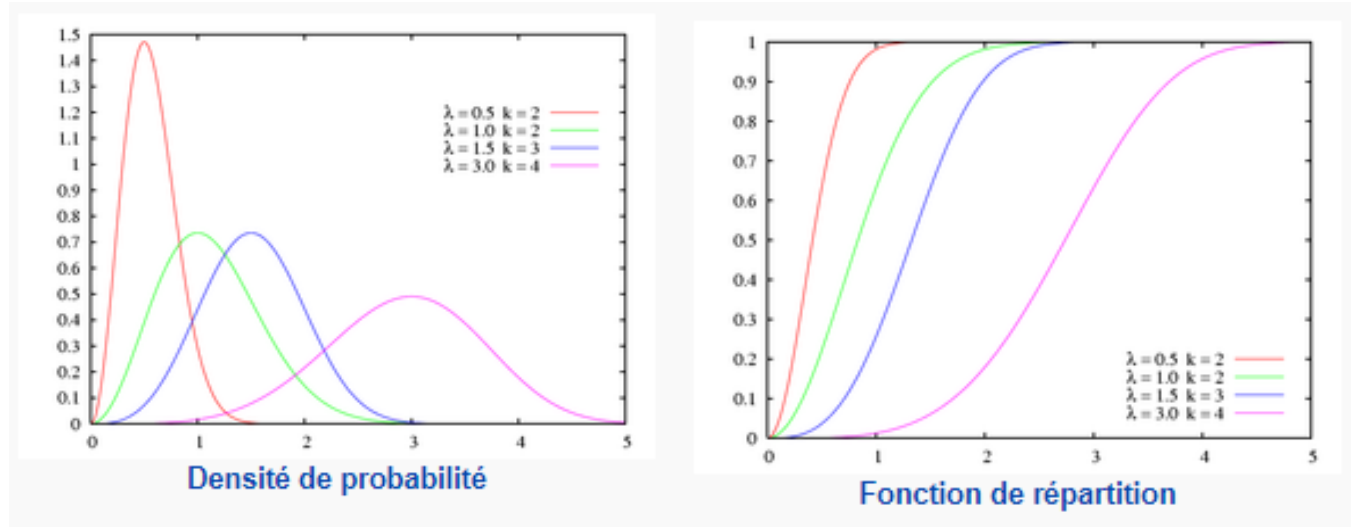
$$H(x) = \int_0^x h(u) du = -\ln(R(x)) = -\ln\left(e^{-\left(\frac{x}{\alpha}\right)^\beta}\right)$$

Moment d'ordre k :

Le moment d'ordre k de la distribution de Weibull est donné par la formule suivante : On considère la loi a priori des paramètres α et β est donnée par la formule suivante :

$$\begin{aligned} E(x^k) &= \int_0^x t^k f(t; \alpha, \beta) dt \\ &= \int_0^x t^k \left(\frac{\beta}{\alpha}\right) \left(\frac{t}{\alpha}\right)^{\beta-1} e^{-\left(\frac{t}{\alpha}\right)^\beta} dt \\ &= \frac{1}{\beta^2} \Gamma\left(1 + \frac{k}{\alpha}\right) \end{aligned}$$

-La représentation graphique de la fonction de densité f et de répartition F de la loi de Weibull (α, β) est donnée par les graphes suivants :



3.1.1 La fonction de vraisemblance

Dans cette section, on va calculer la fonction de vraisemblance de la loi de Weibull à deux paramètres α et β ensuite on calcule la loi à posteriori $\pi(\alpha, \beta | \underline{x})$

$$\begin{aligned}
 L(\underline{x}; \alpha, \beta) &= \prod_{i=1}^n f(x_i; \alpha, \beta) \\
 &= \prod_{i=1}^n \left(\frac{\beta}{\alpha}\right) \left(\frac{x_i}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x_i}{\alpha}\right)^\beta} \\
 &= \left(\frac{\beta}{\alpha}\right)^n \left(\frac{1}{\alpha}\right)^{\beta-1} \prod_{i=1}^n x_i^{\beta-1} e^{-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta} \\
 &= \left(\frac{\beta^n}{\alpha^n}\right) \left(\frac{1}{\alpha^{n\beta-n}}\right) \prod_{i=1}^n x_i^{\beta-1} e^{-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta} \\
 &= \left(\frac{\beta^n}{\alpha^{n\beta}}\right) \prod_{i=1}^n x_i^{\beta-1} e^{-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta}.
 \end{aligned}$$

- On considère loi a priori des paramètres α et β est donnée par la formule suivante :

$$\pi(\alpha, \beta) = \beta^{-a} \alpha^{-b} e^{-\frac{c}{\alpha}}, \quad a > 1 \wedge b, c > 0$$

- L'a posteriori est calculée à l'aide de la formule :

$$\pi(\alpha, \beta | \underline{x}) = \frac{L(\underline{x}; \alpha, \beta) \pi(\alpha, \beta)}{\iint_{\Theta} L(\underline{x}; \alpha, \beta) \pi(\alpha, \beta) d\alpha d\beta} = \frac{L(\underline{x}; \alpha, \beta) \pi(\alpha, \beta)}{m(x)}$$

Où

$$m(x) = \iint_{\Theta} L(\underline{x}; \alpha, \beta) \pi(\alpha, \beta) d\alpha d\beta \quad \text{est dite loi marginale}$$

d'où

$$\begin{aligned} \pi(\alpha, \beta | \underline{x}) &\propto L(\underline{x}; \alpha, \beta) \pi(\alpha, \beta) \\ \pi(\alpha, \beta | \underline{x}) &\propto \beta^n \alpha^{-n\beta} \prod_{i=1}^n x_i^{\beta-1} e^{-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta} \beta^{-a} \alpha^{-b} e^{-\frac{c}{\alpha}} \\ \pi(\alpha, \beta | \underline{x}) &\propto \beta^{n-a} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha}}. \end{aligned}$$

dans cette partie on considère la fonction de perte quadratique, Linex. Le tableau suivant présente les deux fonctions de perte et l'expression de l'estimateur bayésien avec le risque a posteriori correspondant,

Fonction de perte	L'expression	L'estimateur Bayésien	Le risque a posteriori
Quadratique	$L(\theta, \delta) = (\theta - \delta)^2$	$\hat{\delta}_Q = E_\pi(\theta x)$	$E_\pi((\theta - \hat{\delta}_Q)^2)$
Linex	$L(\theta, \delta) = e^{r\Delta} - r\Delta - 1$	$\hat{\delta}_L = -\frac{1}{r} \ln(E_\pi(e^{-r\theta}))$	$r(\hat{\delta}_Q - \hat{\delta}_L)$

Tab. 1.2 Les deux fonctions de perte et l'expression de l'estimateur bayésien avec le risque a posteriori correspondant.

3.2 Estimation Bayésienne sous la fonction de perte quadratique

On a l'estimateur Bayésien des paramètres α et β de la loi de Weibull sous la fonction de perte quadratique est la moyenne a posteriori d'où :

$$\begin{aligned}
 \hat{\alpha}_Q = E_\pi(\alpha|\underline{x}) &= \frac{\iint \alpha \pi(\alpha, \beta|\underline{x}) d\alpha d\beta}{\iint \pi(\alpha, \beta|\underline{x}) d\alpha d\beta} \\
 &= \frac{\iint \alpha \beta^{n-a} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha})} d\alpha d\beta}{\iint \beta^{n-a} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha})} d\alpha d\beta} \\
 &= \frac{\iint \beta^{n-a} \alpha^{-n\beta-b+1} \prod_{i=1}^n x_i^{\beta-1} e^{(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha})} d\alpha d\beta}{\iint \beta^{n-a} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha})} d\alpha d\beta}.
 \end{aligned}$$

► Le risque a posteriori :

$$\begin{aligned}
 PR(\hat{\alpha}_Q) &= E_\pi((\alpha - \hat{\alpha}_Q)^2) \\
 &= E_\pi(\hat{\alpha}_Q^2 - 2\alpha\hat{\alpha}_Q + \alpha^2) \\
 &= \hat{\alpha}_Q^2 - 2\hat{\alpha}_Q E_\pi(\alpha) + E_\pi(\alpha^2).
 \end{aligned}$$

$$\begin{aligned}
 \hat{\beta}_Q = E_\pi(\beta|\underline{x}) &= \frac{\iint \beta \pi(\alpha, \beta|\underline{x}) d\beta d\alpha}{\iint \pi(\alpha, \beta|\underline{x}) d\beta d\alpha} \\
 &= \frac{\iint \beta \beta^{n-a} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha})} d\beta d\alpha}{\iint \beta^{n-a} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha})} d\beta d\alpha} \\
 &= \frac{\iint \beta^{n-a+1} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha})} d\beta d\alpha}{\iint \beta^{n-a} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha})} d\beta d\alpha}.
 \end{aligned}$$

► Le risque a posteriori :

$$\begin{aligned}
 PR(\widehat{\beta}_Q) &= E_\pi((\beta - \widehat{\beta}_Q)^2) \\
 &= E_\pi(\beta^2 - 2\beta\widehat{\beta}_Q + \widehat{\beta}_Q^2) \\
 &= E_\pi(\beta^2) - 2\widehat{\beta}_Q E_\pi(\beta) + \widehat{\beta}_Q^2.
 \end{aligned}$$

3.3 Estimation Bayésienne sous la fonction de perte Linex

On considère maintenant la fonction de perte asymétrique Linex, les estimateurs Bayésienne des paramètres α et β sous cette fonction de perte sont les suivant :

$$\begin{aligned}
 \widehat{\alpha}_L &= -\frac{1}{r} \ln[E_\pi(e^{-r\alpha})], r \neq 0 \\
 &= -\frac{1}{r} \ln\left[\frac{\iint e^{-r\alpha} \pi(\alpha, \beta | \underline{x}) d\alpha d\beta}{\iint \pi(\alpha, \beta | \underline{x}) d\alpha d\beta}\right] \\
 &= -\frac{1}{r} \ln\left[\frac{\iint e^{-r\alpha} \beta^{n-r} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{\left(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha}\right)} d\alpha d\beta}{\iint \beta^{n-r} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{\left(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha}\right)} d\alpha d\beta}\right] \\
 &= -\frac{1}{r} \ln\left[\frac{\iint \beta^{n-r} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{\left(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha} - r\alpha\right)} d\alpha d\beta}{\iint \beta^{n-r} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{\left(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha}\right)} d\alpha d\beta}\right].
 \end{aligned}$$

► Le risque a posteriori :

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

Où : $\widehat{\alpha}_Q$ et $\widehat{\alpha}_L$ sont les estimateurs Bayésiens de paramètre α sous les fonctions de perte quadratique et Linex respectivement.

$$\begin{aligned}
\widehat{\beta}_L &= -\frac{1}{r} \ln[E_\pi(e^{-r\beta})], r \neq 0 \\
&= -\frac{1}{r} \ln\left[\frac{\iint e^{-r\beta} \pi(\alpha, \beta | \underline{x}) d\alpha d\beta}{\iint \pi(\alpha, \beta | \underline{x}) d\beta d\alpha}\right] \\
&= -\frac{1}{r} \ln\left[\frac{\iint e^{-r\beta} \beta^{n-r} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{\left(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha}\right)} d\beta d\alpha}{\iint \beta^{n-r} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{\left(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha}\right)} d\beta d\alpha}\right] \\
&= -\frac{1}{r} \ln\left[\frac{\iint \beta^{n-r} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{\left(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha} - r\beta\right)} d\beta d\alpha}{\iint \beta^{n-r} \alpha^{-n\beta-b} \prod_{i=1}^n x_i^{\beta-1} e^{\left(-\frac{1}{\alpha^\beta} \sum_{i=1}^n x_i^\beta - \frac{c}{\alpha}\right)} d\beta d\alpha}\right].
\end{aligned}$$

► Le risque a posteriori :

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

• On peut pas calculer l'expression analytique de ces estimateurs, c'est pour ça, on utilise des méthodes numériques pour trouver les valeurs des estimateurs Bayésiens des paramètres α et β ainsi que leurs risque a posteriori.

3.3.1 Simulation

-Dans cette partie, on va générer un n-échantillon de la loi de Weibull à deux paramètres α et β .

On prend des différentes tailles d'échantillon.

On suppose que : $a = 2, b = c = 1$ (les paramètres de la loi a priori)

On suppose aussi que : $\alpha = 1$ et $\beta = 2$.

et on utilise les différentes méthodes numériques pour calculer les estimateurs Bayésien des

paramètres ainsi que leurs risques a posteriori sous différentes fonctions de pertes (quadratique et Linex).

On va commencer par une des méthodes MCMC en particulier l'algorithme de Metropolis-Hastings.

Le tableau suivante représente les estimateurs Bayésien des paramètres α et β sous la fonction de perte quadratique (avec leurs erreurs a posteriori) obtenus en utilisant l'algorithme de Metropolis-Hastings.

On pose : $\alpha = 1, \beta = 2$

n	$\hat{\alpha}_Q(PR(\hat{\alpha}_Q))$	$\hat{\beta}_Q(PR(\hat{\beta}_Q))$
n=10	1.0779 (0.0060)	2.0348(0.0012)
n=20	1.0443(0.0019)	2.0299 (0.00089)
n=50	1.0251(0.00063)	2.0162(0.00026)
n=100	1.2161(0.6143)	2.4261(0.4768)

Les estimateurs bayesiens des paramètres α et β (avec leurs erreurs a posteriori) sous la fonction de perte Linex (avec $r = -2$) sont donnés dans le tableau suivant :

n	$\hat{\alpha}_L(PR(\hat{\alpha}_L))$	$\hat{\beta}_L(PR(\hat{\beta}_L))$
n=10	1.0544 (-0.0471)	2.0536 (0.0375)
n=20	1.0299(-0.0288)	2.0250(-0.0097)
n=50	1.0317(0.01317)	2.0235(0.0146)
n=100	1.2333(0.0343)	2.8798(0.9073)

Remarque :

On remarque que on a des bons estimateurs des paramètres de la loi de Weibull sous la fonction de perte quadratique et Linex avec différentes tailles d'échantillons.

En utilisant l'algorithme de PMC générale, on obtient les résultats données dans les tableaux ci-dessus :

1 - Sous la fonction de perte Quadratique :

n	$\hat{\alpha}_Q(PR(\hat{\alpha}_Q))$	$\hat{\beta}_Q(PR(\hat{\beta}_Q))$
n=10	1.0549 (0.0035)	2.0032(1.08 * 10 ⁻⁵)
n=20	1.1605(0.0257)	1.8723(0.0162)
n=50	0.9012(0.0097)	2.0897(0.0080)
n=100	1.0084(7.11 * 10 ⁻⁵)	2.1185(0.01406)

2 - Sous la fonction de perte Linex : (r=1)

n	$\hat{\alpha}_L(PR(\hat{\alpha}_L))$	$\hat{\beta}_L(PR(\hat{\beta}_L))$
n=10	0.5936 (0.4658)	1.0725 (0.9307)
n=20	0.6352(0.5253)	1.0387(0.8336)
n=50	0.5398(0.3614)	1.1330(0.9566)
n=100	0.06003(0.4080)	1.1642(0.9543)

Discussion :

En utilisant les méthodes MCMC (en particulier l'algorithme de Metropolis-Hastings) on obtient des bons estimateurs des paramètres de la loi de Weibull sous la fonction de perte quadratique et Linex avec différentes tailles d'échantillons (le risque a posteriori tend vers 0).

Mais, par les méthodes PMC, on obtient aussi des bons estimateurs des paramètres sous la fonction de perte quadratique (petit risque a posteriori) mais ce n'est pas le cas pour la fonction de perte Linex.

Conclusion et Perspectives

Dans cet travail, nous nous sommes intéressés à une distribution de Weibull à deux paramètres pour une estimation bayésienne des paramètres en utilisant un plan des données complètes et une loi a priori conjuguée naturelle sur les paramètres et sous deux fonctions de perte : la fonction de perte quadratique et la fonction de perte Linex. Une étude par simulation a été réalisée et les méthodes MCMC et PMC nous donnent de bons estimateurs surtout sous la fonction de perte quadratique avec différentes tailles d'échantillon.

En perspectives, ce travail peut-être élargi pour des données censurées où progressivement censurée et sous, aussi, on peut refaire la même étude mais en utilisant balanced loss functions.

Résumé

Ce travail est dédié à l'étude d'estimation des paramètres. le modèle auquel on s'intéresse est le modèle de Weibull à deux paramètres. l'approche utilisée est une approche bayésienne avec une fonction de perte symétrique (la fonction de perte quadratique), puis une fonction de perte asymétrique dont la fonction de perte Linex. en utilisant des données complètes et une loi a priori conjuguée naturelle pour les paramètres. l'expression des estimateurs bayésiens reste sous forme d'intégrales, c'est pourquoi, nous utilisons les méthodes de Monte-Carlo (MCMC) et les méthodes PMC.

Ces méthodes numériques nous permis de trouver la valeur des chaque estimateurs ainsi que son risque a posteriori.

Mots clé :

bayésien - densité a posteriori - MCMC - PMC - la loi de Weibull.

Abstract

this work is dedicated to the statistical estimation of the parameters. we consider the Weibull model with two parameters. we study the estimation problem by applying a bayesian approach using quadratic loss function, then the asymmetric loss function (Linex loss function). we use completed data and conjugate prior. The bayesian estimators is given in integral form to which we apply simulation techniques suchas MCMC and PMC methods.

this numerical methods given the values of each estimators and his posterior risk.

Key-Words :

bayesian - posterior density - MCMC - PMC - Weibull model.

Annexe 1

Algorithme de Metropolis-Hastings

```
N=1000
H1=numeric(N)
H2=numeric(N)
n=10;alpha=1;beta=2
t=numeric(N)
for(k in 1:N){
vec=rweibull(n,alpha,beta)
vec=sort(vec)
t=vec
a=2;b=1;c=1;aa=4
f = function(x,y){(y(n-a)) * (x(-n*y-b)) * prod(t(y-1)) * exp((-x(1/y)) * (sum(t)y) - (c/x))}
q = function(x,y){(x(aa/2-1)) * exp(-x/2) * (y(aa/2-1)) * exp(-y/2)}
M=500;ind=N*2
X=matrix(rep(0,ind),ncol=2,nrow=M)
Y=numeric(2)
X[1,1]=1;X[1,2]=1
for(i in 2:M){
Y=rchisq(1,2)+c(1,2)
val=(f(Y[1],Y[2])*q(X[i-1,1],X[i-1,2]))/(q(Y[1],Y[2])*f(X[i-1,1],X[i-1,2]));
alpha0=min(1,val)
u=runif(1)
```

```
if(u<alpha0)
X[i,1]=Y[1];X[i,2]=Y[2]
else
X[i,1]=X[i-1,1];X[i,2]=X[i-1,2]
} H1[k]=mean(X[,1])
H2[k]=mean(X[,2])
alphaQ=mean(H1)
betaQ=mean(H2)
alphaQ;betaQ
PRalphaQ=(alphaQ-alpha)^2
PRbetaQ=(betaQ-beta)^2
PRalphaQ;PRbetaQ
r=-2
alphaL=(-1/r)*(log(mean(exp(-r*X[, 1]))))
betaL=(-1/r)*(log(mean(exp(-r*X[, 2]))))
alphaL;betaL
PRalphaL = r*(alphaQ - alphaL)
PRbetaL = r*(betaQ - betaL)
PRalphaL;PRbetaL
```


Algorithme de PMC général

```

M=200
T=15;n=10;alpha=1;beta=2
vec=rweibull(n,alpha,beta)
t=sort(vec)
t
X=matrix(0,ncol=M, nrow=T+1)
Y= matrix(0,ncol=M, nrow=T+1)
r=matrix(0,ncol=M, nrow=T+1)
w=matrix(0,ncol=M, nrow=T+1)
histX=0
histY=0
vrais=0
X[1,]=rchisq(M,1)
Y[1,]=rchisq(M,1)
t<-1
while (t <= T){
X[t+1,]=rweibull(M,alpha,beta)
Y[t+1,]=rweibull(M,alpha,beta)
for(i in 1 :M){
vrais[i]=prod(dweibull(vec,X[t,],alpha,beta))
r[t+1,]=(vrais * dgamma(X[t + 1, ], alpha, beta))/(dchisq(vec, X[t, ], alpha, beta))
r[t+1,]=(r[t+1,]/sum(r[t+1,]))
X[t+1,]=sample(X[t+1,], M, replace = TRUE, prob =r[t+1,])
Y[t+1,]=sample(Y[t+1,], M, replace = TRUE, prob =r[t+1,])
histX=histX+sum(r[t+1,]*X[t+1,])
histY=histY+sum(r[t+1,]*Y[t+1,])
t=t+1
}

alphaQ=mean(X[1,])
betaQ=mean(X[2,])

```

alphaQ

betaQ

PR1=(alphaQ-alpha)²

PR2=(betaQ-beta)²

PR1;PR2

r=1

alphaL=(-1/r)*(log(mean(exp(-r*X[1,]))))

betaL=(-1/r)*(log(mean(exp(-r*X[2,]))))

PR3=r*(alphaQ-alphaL)

PR4=r*(betaQ-betaL)

PR3;PR4

Annexe 2

Risque fréquentiste

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

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

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

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

Définition

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

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

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

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

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

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

$$R(\theta_1, \delta_1) < R(\theta_1, \delta_2) \quad \text{et} \quad R(\theta_2, \delta_1) > R(\theta_2, \delta_2).$$

Méthode de rejet-acceptation

Si l'on désire générer des valeurs d'une variable aléatoire avec une densité $f(x)$, alors que ses propriétés ne sont pas connues, et sa fonction de répartition associée n'est pas inversible d'une manière explicite.

La méthode de rejet-acceptation est utilisée s'il existe une fonction $g(x)$ ayant un domaine identique à celui de f et une constante C positive [19], telle que :

$$f(x) \leq Cg(x), \quad \forall x \in \Theta.$$

La méthode se résume comme suit :

1. Générer un nombre y selon g ;
2. Générer un nombre u selon $U_{[0, 1]}$;
3. Si u vérifie la condition $u \leq \frac{f(y)}{Cg(x)}$ alors :

On accepte la valeur générée ;

sinon on rejette u et y .

C doit vérifier l'équation : $C \geq f(x)/g(x), \quad \forall x \in \Theta.$

La valeur optimale de C est donc :

$$C = \max_{\Theta} f(x)/g(x).$$

Le taux d'acceptation sera alors défini par :

$$\rho = f(x)/Cg(x).$$

Remarquons que si $g(x)$ est proche de zéro, alors C devient très grand et le taux d'acceptation moyen diminue. L'efficacité de l'algorithme dépend donc de l'adéquation entre f et g .

Modèle

Fonction de densité f :

Une variable aléatoire continue X suit une loi de Weibull de paramètres α et β , si elle admet pour densité de probabilité la fonction :

$$f(x; \alpha, \beta) = \left(\frac{\beta}{\alpha}\right) \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}}.$$

Où

- $\alpha > 0$: le paramètre d'échelle.
- $\beta > 0$: le paramètre de la forme.

Fonction de répartition F :

On dit que F est une fonction de répartition de la variable aléatoire X qui suit une loi de Weibull à deux paramètres α, β , si F est donnée par :

$$F(x; \alpha, \beta) = \left(1 - e^{-\left(\frac{x}{\alpha}\right)^\beta} \right), \quad \alpha, \beta, x > 0$$

Fonction de fiabilité R :

Soit X une variable aléatoire continue de loi de Weibull de fonction de répartition F et de densité de probabilité f . Sa fonction de fiabilité est définie par :

$$R(x) = 1 - F(x) = e^{-\left(\frac{x}{\alpha}\right)^\beta}$$

Fonction de taux de panne (hasard) h :

La fonction de hasard est définie comme la probabilité conditionnelle que le phénomène se termine après une durée x sachant que l'on a atteint cette durée (taux de panne, taux de défaillance, taux de décès ou risque instantané). La fonction de taux de hasard h est calculée à l'aide de la formule suivante :

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{R(x)},$$

donc :

$$h(x) = \frac{\left(\frac{\beta}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta}}{e^{-\left(\frac{x}{\alpha}\right)^\beta}}$$

Fonction de taux de hasard cumulé H :

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

$$H(x) = \int_0^x h(u) du = -\ln(R(x)) = -\ln\left(e^{-\left(\frac{x}{\alpha}\right)^\beta}\right).$$

Bibliographie

- [1] **Tierney, L., Kadane, J.B. (1986).** Accurate approximations for posterior moments and marginal densities, *J. Amer. Statist. Assoc.*, 81, 82-86.
- [2] **Lindley, D.V, (1980).** Approximate Bayesian methods . *Trabajos Estadist. Investigacion Oper.* 31, 232-245.
- [3] **N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller.** Equation of state calculations by fast computing machines. *Chemical Physics*, 21(6) :1087–1092, June 1953.
- [4] **W.K. Hastings.** Monte Carlo Sampling methods using Markov chains and their applications, volume 57. *Biometrika*, April 1970. pages 97-109.
- [5] **S. Geman and D. Geman.** Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. 6 :721–741, 1984.
- [6] **L. Tierney.** Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics*, 1994.
- [7] **N. Bousquet.** Analyse bayésienne de la durée de vie de composants industriels. Décembre 2006.
- [8] **A. Guillin, J.M. Marin, and C.P. Robert.** Estimation bayésienne approximative par échantillonnage préférentiel. rapport de recherche, Université Paris IX Dauphine, 2005.
- [9] **Boudjerda. K** ,”Etude de l’estimateur de Bayes sous différentes fonctions de perte” . Thèse de doctorat en Mathématiques, Université Badji Mokhtar Annaba, 2016/2017.
- [10] **Belaïd. N et Djerroud. L** ,”Les méthodes de Monte Carlo : (MCMC et PMC). Applications” . mémoire master Recherche Opérationnelle, Université A. Mira- Béjaïa, Juin 2013.

-
- [11] **Talhi. H** ,”L’estimation Bayesienne en Fiabilite en Presence de Donnees Censurees” .
Thèse de doctorat en Mathématiques, Université Badji Mokhtar Annaba, 2013/2014.