



Faculté des Sciences Exactes et Informatique

Département de Mathématiques

N° d'ordre :

N° de séries :

Mémoire de fin d'études

Présenté pour l'obtention du diplôme de

Master

Spécialité : Mathématiques.

Option : Probabilités et Statistique.

Thème

Estimation Bayésienne sous les données hybrides

Présenté par :

- LABSI Bochra.

Devant le jury :

Président : DJERIDI Zohra M.C.B Université de Jijel

Encadreur : BOUJERDA Khawla M.C.B Université de Jijel

Examineur : ABDI Zeyneb M.A.A Université de Jijel

Promotion **2021/2022**

✧ *Remerciements* ✧

Avant tout je remercie **Allah** de m'avoir donné le courage, la santé et la volonté de compléter
ce mémoire

En particulier je remercie mon encadreur de mémoire **Mme Boudjerda Khawla** pour
tous ses conseils lors de la rédaction de ce travail.

Un grand merci également aux membres du jury, la présidente **Mme Djeridi Zohra** et
aussi l'examinatrice **Mme Abdi Zeyneb**, qui ont bien
voulu évaluer ce mémoire

Enfin, je tiens à remercier tous ceux qui m'ont aidé de près ou de loin à la réalisation de
ce mémoire.

Dédicace

Je dédie ce travail

*À mes chères parents
qui m'ont toujours soutenu*

À ma soeur

imane

À mes chers frères

Mehdi et Mouad et Islame

*À mes meilleurs amies et particulièrement **Hadjer et Ibtesame***

Et à tout mes amies de promotion de 2^{me} année Master probabilités et statistiques

À tout les membre de ma famille

Et toutes personne qui occupe une place dans mon coeur

Merci♥Bochra

Table des matières

Résumé	v
Abstract	vi
Introduction générale	1
1 Analyse Bayésienne	3
Introduction	3
1.1 Théorème de Bayes	3
1.2 Choix de la loi a priori	5
1.2.1 Loi conjuguée	5
1.2.2 Loi non informative	6
1.2.3 Loi a priori impropre	6
1.3 Fonction de coût (de pertes)	6
1.3.1 Fonction de perte quadratique	6
1.3.2 Fonction de perte Linex	8
1.3.3 Fonction de perte $0 - 1$	8
1.3.4 Fonction de perte absolue	9
1.4 Fonction de risque	10
1.4.1 Risque fréquentiste	10
1.4.2 Risque de Bayes	11
1.4.3 Risque a posteriori	11
1.5 Estimateurs de Bayes	11

1.6	Méthodes numériques utiles dans le cadre bayésien	12
1.6.1	La méthode de Monte-Carlo	12
1.6.2	Méthodes MCMC	13
1.6.3	Algorithme de Metropolis-Hastings	14
1.6.4	Echantillonneur de Gibbs	18
2	Plans de données	22
	Introduction	22
2.1	Les données censurées	23
2.1.1	Censure à droite	24
2.1.2	Censure à gauche	24
2.1.3	Censure par intervalle	25
2.2	Données progressivement censurées	27
2.3	Données hybrides de type I	29
3	Application sous R	32
	Introduction	32
3.1	Modèle	32
3.2	Estimation par la méthode de maximum vraisemblance	35
3.3	Estimation Bayésienne	37
3.3.1	Estimation Bayésienne sous la fonction de perte quadratique	38
3.3.2	Estimation Bayésienne sous la fonction de perte Linex	40
3.4	Application	41
	Conclusion	45
	Annexes	46
	Bibliographie	51

Liste des tableaux

1.1	Lois a priori conjuguées naturelles pour quelques familles exponentielles usuelles [5].	5
1.2	Les fonctions de perte et l'expression des estimateurs bayésiens avec les risques à posteriori correspondantes [3]	12
3.1	Estimateurs de α , β et γ (erreur quadratique) par la méthode du maximum de vraisemblance.	42
3.2	Estimateurs de α , β et γ sous la fonction de perte quadratique (le risque a posteriori).	43
3.3	Estimateurs de α , β et γ sous la fonction de perte Linex (le risque a posteriori) .	44

Table des figures

3.1	Graphe de la fonction de fiabilité $R(x)$ avec différent valeurs de α	34
3.2	Graphe de la fonction de taux de panne $h(x)$ avec différentes valeurs de α	35

Résumé

Ce mémoire est dédié à l'étude de l'estimation des paramètres d'un modèle mixte, le modèle mixte est la distribution de probabilité d'un variable aléatoire qui est la dérivée d'une collection d'autres variables aléatoires comme suit : premièrement, une variable aléatoire est sélectionnée au hasard dans la collection en fonction de probabilité en fonction de probabilité des donnée, puis de la valeur de la variable aléatoire sélectionne. Le modèle auquel on s'intéresse est le modèle de Rayleigh-Pareto.

Dans ce mémoire, nous exposons l'étude d'estimateurs par la méthode de maximum de vraisemblance des paramètres de la loi de Rayleigh-Pareto, et aussi l'étude de l'estimateurs de Bayes des paramètres sous différentes fonctions de pertes et en utilisant des données hybrides. La loi a priori utilisée dans ce travail est une mixture entre loi a priori conjuguée naturelle et loi a priori non informatives d'expression des estimateurs Bayésiens reste sous forme intégrales c'est pourquoi, nous utilisons les méthodes de simulations de Monte Carlo (MCMC) et en particulier l'algorithme de Metropolis-Hastings.

Mots clés :

Rayleigh-Pareto - Bayésiens - MCMC - Metropolis-Hastings.

Abstract

This memory is dedicated to the statistical estimation of parameters of mixture distribution, mixture distribution is the probability distribution of a random variable that is derived from a collection of other random variables as follows : first, a random variable is selected by chance from the collection according to given probabilities of selection, and then the value of the selected random variable is realized. The mixture distribution that we are interested in is the Rayleigh-Pareto.

We shall especially, study the maximum likelihood estimators of the parameters of Rayleigh-Pareto model, and the Bayesian estimators of the parameters under different loss functions and by using hybrid data. The prior law used in this work is the mixture between : non informative and natural conjugate the obtained Bayesian estimator is given in integral form to which we apply simulation techniques such as the Markov chain Monte Carlo (MCMC) method and in particular the Metropolis-Hastings algorithm to derive numerical values of these estimators.

Key-Words

Rayleigh-Pareto - Bayesian - MCMC - Metropolis-Hastings.

Introduction générale

L'objet principale de la statistique est de faire, à partir d'un phénomène aléatoire, une inférence au sujet de la loi générant ces observations, en vue, d'analyser le phénomène et de prévoir un évènement futur. Donc, l'objet principale de la statistique est de mener, grâce à l'observation d'un phénomène aléatoire, une inférence sur la distribution probabiliste à l'origine de ce phénomène, c'est-à-dire, de fournir une analyse d'un phénomène passé, ou une prédiction d'un phénomène à venir de nature similaire. Parmi les approches les plus utiles dans la statistique est l'approche Bayésienne qui englobe l'information dont on dispose a priori par un retour d'expérience, sur un phénomène étudié en considérant le paramètre θ non pas comme inconnu, mais comme une variable aléatoire possédant une distribution appelée a priori. cette approche consiste à combiner une distribution a priori, selon un ensemble de croyances représentatives avec les données observées pour obtenir une distribution appelée loi a posteriori sur la quelle on pratique l'inférence statistique. De plus, l'estimation Bayésienne est dans un cadre probabiliste et tient automatiquement compte de l'incertitude des paramètres dans les prévisions. Le développement d'outils informatique puissants, contribue à la popularité croissante de cette méthode d'estimation. Dans ce présent mémoire, nous nous intéressons à l'estimation des paramètres de la loi de Rayleigh-Pareto en utilisant des données hybrides, notre travail est organisé en trois chapitre :

- Le premier chapitre présente les notions fondamentaux de la statistique Bayésienne et les méthodes numériques utiles dans le cadre Bayésien.
- Le chapitre deux présente les différents types des données censurées.
- Le dernier chapitre présente l'estimation des paramètres de la loi de Rayleigh-Pareto en utilisant des données hybrides, par la méthode du maximum de vraisemblance puis par la méthode

Bayésienne sous deux fonctions de perte ; la première symétrique qui est la fonction de perte quadratique et la deuxième a symétrique celle de Linex.

Analyse Bayésienne

Introduction

L'analyse Bayésienne, créée par Pierre Simon de Laplace et Thomas Bayes (1774), est la première étape d'étudier d'une situation et d'identifier une incertitude portée sur un paramètre inconnu θ puis de quantifier, celle-ci, à travers une représentation probabiliste en utilisant le calcul des probabilités.

L'incertitude sur θ est modélisée sous la forme d'une distribution, dite a priori, qui apporte une information sur θ pris comme étant une variable aléatoire contrairement à l'analyse fréquentiste qui la considère comme une constante.

Cette a priori est actualisée en extrayant de l'information contenu dans les observations de \mathcal{X} , pour obtenir une autre distribution dite distribution a posteriori. Donc, l'analyse Bayésienne est une approche d'analyse statistique qui est basée sur le théorème de Bayes, selon la quelle la probabilité a posteriori est proportionnelle à la probabilité a priori du paramètre multipliée par vraisemblance.

1.1 Théorème de Bayes

Le théorème de Bayes est énoncé mathématiquement comme l'équation suivante :

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

où A et B sont des évènements et $P(B) \neq 0$.

Dans le cas continue, le théorème de Bayes s'écrit sous la forme :

$$g(y|x) = \frac{f(x|y)g(y)}{\int f(x|y)g(y)dy}.$$

Bien que ce théorème d'inversion soit naturel d'un point de vue probabiliste, Bayes et Laplace sont allés plus loin et ont considéré que l'incertitude sur le paramètre θ d'un modèle peut être décrite par une distribution de probabilité π sur Θ appelée distribution a priori. L'inférence est alors, fondée sur la distribution de θ conditionnelle à x , $\pi(\theta|x)$, appelée distribution a posteriori et définie par :

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta} \quad (1.1)$$

Notons que $\pi(\theta|x)$ est ainsi proportionnelle à la distribution de x conditionnellement à θ , qui est aussi la vraisemblance, multipliée par la distribution a priori de θ . (Il semble que la généralité de (1.1) n'ait pas été perçue par Bayes, mais par Laplace, qui la développera plus avant). La contribution principale apportée par un modèle statistique Bayésien est donc de considérer en sous une distribution aléatoire pour les paramètres [5].

Exemple 1.1.1. On prend $\theta = p$, si $x \sim \mathcal{B}(n, p)$, nous considérant la loi a priori de type Bêta $p \sim \mathcal{Be}(\alpha, \beta)$, avec $\alpha = \beta = 1$, dans le cas particulier de Bayes,

$$f(x|p) = C_n^x p^x (1-p)^{n-x}.$$

avec : $C_n^x = \frac{n!}{x!(n-x)!}$.

La loi a priori est :

$$\pi(p) = \frac{1}{\mathcal{B}(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1},$$

où $\mathcal{B}(\alpha, \beta)$ est la fonction bêta.

La distribution jointe de (x, p) est alors ;

$$\varphi(x, p) = \frac{C_n^x}{\mathcal{B}(\alpha, \beta)} p^{\alpha+x-1} (1-p)^{\beta+n-x-1}.$$

La distribution marginale de x est :

$$\begin{aligned} m(x) &= \frac{C_n^x}{\mathcal{B}(\alpha, \beta)} \mathcal{B}(\alpha+x, n-x+\beta) \\ &= C_n^x \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+x)\Gamma(n-x+\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(n-x+\beta)}. \end{aligned}$$

où $\Gamma(x)$ est la fonction gamma. Puisque la distribution a posteriori de p est :

$$\pi(p|x) = \frac{p^{\alpha+x-1}(1-p)^{\beta+n-x-1}}{B(\alpha+x, n-x+\beta)},$$

qui est une loi de $\mathbf{Be}(\alpha+x, \beta+n-x)$.

1.2 Choix de la loi a priori

On distingue trois types essentiels de familles de la loi a priori

1.2.1 Loi conjuguée

Ce type des lois a priori est utilisé quand l'information a priori disponible sur le modèle est trop vague ou peu faible. Une des difficultés de l'approche Bayésienne est le calcul de la loi a posteriori. Ce calcul est facilité lorsque loi a priori et loi a posteriori ont la même forme. Dans ce cas, on parle de loi a priori conjuguée.

Définition 1.2.1. Une famille \mathcal{F} de distributions de probabilité sur Θ est dite conjuguée par une fonction de vraisemblance $f(x|\theta)$ si, pour tout $\pi \in \mathcal{F}$, la distribution a posteriori $\pi(\cdot|x)$ appartient également à \mathcal{F} .

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
Normale $N(\theta, \sigma^2)$	Normale $N(\mu, \tau^2)$	$N(\rho(\sigma^2\mu + \tau^2x), \rho\sigma^2\tau^2)$ $\rho^{-1} = \sigma^2 + \tau^2$
Poisson $P(\theta)$	Gamma $G(\alpha, \beta)$	$G(\alpha+x, \beta+1)$
Binomiale $B(n, \theta)$	Bêta $Be(\alpha, \beta)$	$G(\alpha+x, \beta+1)$
Gamma $G(\nu, \theta)$	Gamma $G(\alpha, \beta)$	$G(\alpha+\nu, \beta+x)$
Binomiale négative $Neg(m, \theta)$	Bêta $Be(\alpha, \beta)$	$Be(\alpha+m, \beta+x)$
Normale $N(\mu, \frac{1}{\theta})$	Gamma $G_a(\alpha, \beta)$	$G(\alpha+0.5, \beta+(\mu-x)^2/2)$

TABLE 1.1 – Lois a priori conjuguées naturelles pour quelques familles exponentielles usuelles [5].

Remarque 1.2.1. Une loi conjuguée peut être déterminée en considérant la forme de la vraisemblance $f(x|\theta)$ et en prenant une loi a priori de la même forme que cette dernière vue comme une fonction du paramètre. Les lois a priori conjuguées obtenues par ce procédé sont dites naturelles [16].

1.2.2 Loi non informative

Une loi non informative est une loi qui porte une information sur le paramètre à estimer dont le poids dans l'inférence est réduit. Certains auteurs la définissent également comme une loi a priori qui ne contient aucune information sur Θ ou encore comme une loi qui ne donne pas d'avantage de poids à telle ou telle valeur du paramètre.

Par exemple, supposons un ensemble fini de taille q , une loi a priori non informative peut être une loi de la forme :

$$P(\theta_i) = 1/q.$$

On a équiprobabilité, les valeurs possibles de θ , se voit, attribuer le même poids.

1.2.3 Loi a priori impropre

La loi a priori peut être impropre i.e. $\int_{\Theta} \pi(\theta)d\theta = \infty$. Ce choix de type de loi n'a donc plus d'intérêt que calculatoire et s'interprète difficilement. La construction de lois non informatives peut conduire à des lois a priori de ce type.

1.3 Fonction de coût (de pertes)

Définition 1.3.1. – On appelle fonction de coût, toute fonction L de $\Theta \times \mathcal{A}$ dans \mathbb{R} .

$L(\theta, a)$ évalue le coût d'une décision a quand le paramètre vaut θ . Elle permet donc, en quelque sorte, de quantifier la perte encourue par une mauvaise décision, une mauvaise évaluation de θ . Il s'agit d'une fonction de θ . Un coût négatif correspond à un gain [17].

1.3.1 Fonction de perte quadratique

La fonction proposée par Legendre (1805) et Gauss (1810), cette perte est sans aucun doute le critère d'évaluation le plus commun, elle s'écrit sous la formule suivante :

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

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

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

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

Proposition 1.3.1. *L'estimateur de Bayes δ^π associé à la loi a priori π et avec la perte quadratique, est donné par l'espérance a posteriori :*

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

telle que Θ l'espace des paramètres

Preuve

Comme :

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

Exemple 1.3.1. (Suite de l'Exemple 1.2.1) *calcul de l'estimateur par la fonction de perte quadratique on a :*

$$\begin{aligned} \pi(p|x) &= \frac{p^{\alpha+x-1}(1-p)^{\beta+n-x-1}}{\mathcal{B}(\alpha+x, n-x+\beta)} \\ &\propto p^{\alpha+x-1}(1-p)^{\beta+n-x-1}. \end{aligned}$$

Donc :

$$\begin{aligned} \hat{p}_Q &= E_\pi(p|x) \\ &= \int p\pi(\lambda|x)d\lambda \\ &= \frac{\alpha+x}{\alpha+\beta+n}. \end{aligned}$$

1.3.2 Fonction de perte Linex

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

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

où : $\Delta = (\delta(x) - \theta)$ et $\delta(x)$ 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 approcher de zéro, la perte Linex est approximativement la fonction de perte quadratique :

$$E_{\theta}(L(\delta(x) - \theta)) \propto e^{r\delta(x)} E_{\theta}(e^{-r\theta}) - r(\delta(x) - E_{\theta}(\theta) - 1) \quad (1.2)$$

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 (1.2). Pour trouver l'estimateur, nous dérivons l'équation (1.2) par rapport à $\delta(x)$, nous obtenons

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

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 :

$$\widehat{\delta}_L(x) = -\frac{1}{r} \ln(E_{\theta}(e^{-r\theta}));$$

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

1.3.3 Fonction de perte 0 – 1

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

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

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 $H_1 : \{\theta \in \Theta_1\}$.

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 à 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 : $L(\theta, \delta(x)) = 1$ ($x \in W$).

1.3.4 Fonction de perte absolue

La fonction de perte absolue est définie par :

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

considérée par Laplace (1773), ou, plus généralement une fonction multi-linéaire :

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

De telles fonctions croissent plus lentement que la perte quadratique.

Proposition 1.3.2. *L'estimateur de Bayes associé à la loi a priori θ et à la fonction de coût linéaire par morceaux est le rapport $(k_2/(k_1 + k_2))$ de $\pi(\theta|x)$.*

Preuve. On a

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

est obtenue par une intégration par parties.

Et on dérive par rapport à d , on obtient :

$$k_1 P^\pi(\theta < d|x) + k_2 P^\pi(\theta > d|x) = 0.$$

Et on a donc,

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

donne le cas où $k_1 = k_2$ l'estimateur de Bayes est la médiane a posteriori.

1.4 Fonction de risque

1.4.1 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. Classifier 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 1.4.1. 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 :

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

1.4.2 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 :

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

1.4.3 Risque a posteriori

Définition 1.4.2. Une fois données la loi a priori $\pi(\theta)$ sur le paramètre θ et la fonction de perte $l(\theta, \delta)$, le risque a posteriori qui note $\rho(\pi, \delta|x)$ est défini par :

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

Ainsi, le problème change selon les données ; ceci dû à la non existence d'un ordre total sur les estimateurs [12].

1.5 Estimateurs de Bayes

Soit une fonction de perte $L(\theta, \delta)$, et une loi a priori π , pour trouver l'estimateur de Bayes $\delta^\pi(x)$ on va applique l'expression suivant :

$$\delta^\pi(x) = \min_{\delta} \mathbb{E}[L(\theta, \delta)|x],$$

Cette estimateur sera déterminé analytiquement ou numériquement selon la fonction de perte L et la loi a posteriori et sa complexité. Les solutions associées à des coûts classiques sont formellement connues et correspondent aux caractéristiques usuelles d'une distribution (moyenne, médiane, fractiles,...).

Par exemple, l'estimateur de Bayes associé au coût quadratique est la moyenne a posteriori.

Lemme 1.5.1. *Soit $f(x|\theta) = h(x)e^{\theta \cdot x - \psi(\theta)}$ une distribution d'une famille exponentielle. Pour toute loi a priori π , la moyenne a posteriori de θ est donnée par*

$$\delta^\pi(x) = \nabla \log m_\pi(x) - \nabla \log h(x),$$

où ∇ est l'opérateur gradient et m_π est la loi marginale associée à π .

Preuve. L'espérance a posteriori est donnée par

$$\begin{aligned} \mathbb{E}^\pi[\theta_i|x] &= \frac{\int_{\Theta} \theta_i h(x) e^{\theta \cdot x - \psi(\theta)} \pi(\theta) d\theta}{m_\pi(x)} \\ &= \left(\frac{\partial}{\partial x_i} \int_{\Theta} h(x) e^{\theta \cdot x - \psi(\theta)} \pi(\theta) d\theta \right) \frac{1}{m_\pi(x)} - \left(\frac{\partial}{\partial x_i} h(x) \right) \frac{1}{h(x)} \\ &= \frac{\partial}{\partial x_i} [\log m_\pi(x) - \log h(x)]. \end{aligned}$$

Le tableau ci-dessus présente les trois fonctions de perte et l'expression de l'estimateur Bayésien avec le risque a posteriori correspondant, sous différentes fonctions de perte :

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

TABLE 1.2 – Les fonctions de perte et l'expression des estimateurs bayésiens avec les risques à posteriori correspondantes [3]

1.6 Méthodes numériques utiles dans le cadre bayésien

1.6.1 La méthode de Monte-Carlo

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

fonction quel conque et $f(x)$ une fonction de densité de support $\mathcal{X} \subset \mathbb{R}^m$. On cherche à calculer :

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

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

$$\bar{h}_n = \frac{1}{n} \sum_{j=1}^n h(x_j).$$

Propriétés

1- Par la loi des grand nombre, avec probabilité 1

$$\bar{h}_n \longrightarrow E(h).$$

2- Avec la condition que $E(h^2)$ soit finie

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

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

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

Ces techniques seront sur tout utiles dans les problème de grandes dimensions où les méthodes numériques traditionnel les perdent de leur efficacité. La difficulté sera de trouver une façon de générer efficacement un échantillon de variables i.i.d de densité f .

1.6.2 Méthodes MCMC

Les méthodes de Monte-Carlo par Chaîne de Markov permettent d'élargir grandement l'éventail des distributions pouvant être simulés numériquement. Elles sont relativement simples à implémenter et ne requièrent souvent que la connaissance de la fonction de densité cible à une constante près, ce qui les rend intéressantes dans de nombreuses situations.

Cependant, une implémentation naïve peut mener à des temps de calcul très longs, puis que la convergence de ces méthodes est relativement lente lorsqu'elles ne sont pas bien calibrées à une situation donnée.

Nous verrons d'abord, sommairement, les justifications théoriques de ces méthodes, puis nous

les illustrerons par la présentation de l'algorithme original de Metropolis-Hastings, pour ensuite voir les améliorations successives que peut apporter leurs particularités et leur validité théorique.

L'idée de base est de simuler une distribution de densité f en utilisant une chaîne de Markov ergodique X_t dont la distribution stationnaire est f . Le théorème ergodique garantit alors la convergence en loi de X_t vers une variable de densité f et par conséquent, pour presque toute valeur initiale X_0 :

$$\frac{1}{n} \sum_{i=1}^n h(x_i) \rightarrow E_f(h).$$

Nous appellerons MCMC toute méthode permettant de simuler une distribution en utilisant une chaîne de Markov ergodique ayant celle-ci comme distribution stationnaire. Pour construire un tel algorithme, il faut donc déterminer un ensemble de probabilités de transition P approprié, c'est-à-dire irréductible, ergodique et ayant la bonne distribution stationnaire [6].

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

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

alors la distribution stationnaire de la chaîne de Markov est π .

Démonstration

On aura la stationnarité si $\forall x, y \in S$

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

Or, sous l'hypothèse de réversibilité

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

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

1.6.3 Algorithme de Metropolis-Hastings

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

effectuées [6].

Algorithme MH

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

2. On pose :

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

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

Où : les seuils α doivent avoir la forme générale :

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

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

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

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

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

On dit que q est la densité de proposition des candidats, ou densité instrumentale, et α est les probabilités d'acceptation de ces dernières. Voyons maintenant les propriétés théoriques de cette procédure.

Propriétés[3]

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

$$P(x, y) = q(x, y)\alpha(x, y), \quad \text{si } x \neq y.$$

$$P(x, y) = 1 - \int P(x, y)dy, \quad \text{autrement.}$$

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

$$\begin{aligned} \pi(x) &= \pi(x)q(x, y)\alpha(x, y) \\ &= \frac{\pi(x)q(x, y)s(x, y)}{1 + \frac{\pi(x)q(x, y)}{\pi(y)q(x, y)}} \\ &= \frac{\pi(x)\pi(y)q(x, y)q(y, x)s(x, y)}{\pi(y)q(y, x) + \pi(x)q(x, y)} \\ &= \frac{\pi(y)q(y, x)s(y, x)}{1 + \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}} \\ &= \pi(y)q(y, x)\alpha(y, x) = \pi(y)P(y, x). \end{aligned}$$

Maintenant, il faut s'assurer que la chaîne converge bien vers sa distribution stationnaire, c'est-à-dire qu'elle est ergodique. Or, ceci est facilement vérifié la plupart du temps. Par exemple, si $q(x, y)$ est positive pour toute paire (x, y) appartenant au support de π , $P(x, y)$ aussi sera toujours positive. Ainsi, à partir d'une valeur X_t donnée, toute valeur Y_{t+1} sera atteignable en une seule étape avec une probabilité positive. La chaîne est donc irréductible. La chaîne sera aussi apériodique des moments qu'il existe au moins une paire (x, y) tel que $\alpha(x, y) < 1$ car on aura alors $P(x, y) > 0$.

Cela sera pratiquement toujours vrai et on conclut donc que la chaîne est effectivement ergodique.[6]

Exemple 1.6.1. *Modèle Binomial*

Nous allons utiliser la méthode MCMC pour estimer le paramètre de la loi binomiale de densité :

$$f(x|n, p) = C_n^x p^x (1 - p)^{n-x} ; \quad x = 0, \dots, n$$

Algorithme

L'algorithme utilisé est celui de Metropolis-Hasting. Pour $n=100$, $x=39$ et $p=0,4$ on a :

$$\pi(p|x) \propto p^{39}(1 - p)^{61};$$

- La loi de proposition est une loi uniforme ;
- À l'étape t générer $p' \sim \mathcal{U}[0, (p^{(t-1)}+1)]$
- Calculer la probabilité d'acceptation $\alpha(p(t), (p')) = \text{Min}[1, \frac{(p')^{39}(1-(p'))^{61} * \text{dunif}(p^{(t)}, 0, (p'+1))}{(p^t)^{39}(1-p^{(t)})^{61} * \text{dunif}(p', 0, (p^{(t)}+1))}]$
- Générer $u \sim \mathcal{U}[0, 1]$ et prendre :

$$p^{(t+1)} \begin{cases} (p') & \text{avec la probabilité } \alpha(p(t), (p')) & \alpha(p(t), (p')) & \text{si } u < \alpha(p(t), (p')) \\ p^{(t)} & \text{avec la probabilité } 1 - \alpha(p(t), (p')) & \text{sinon} \end{cases}$$

Résultats

L'estimateur du paramètre p obtenu avec une période de chauffe de 1000 sur 10000 itérations partant de $p_0 = 1$ est :

$$p = 0,3903357.$$

(Code source R du problème est)

```
f = fonction(t){
z = t^(39) * (1 - t)^(61)
return(z)
}
binomial = fonction(T){
p = array(dim = c(T))
p[1] = 1
for(i in 1 : T){
yi = runif(1, 0, (p[i] + 1)) loi de proposition
d = f(yi) * dunif(p[i], 0, (yi + 1))
e = f(p[i]) * dunif(yi, 0, (p[i] + 1))
k = d/e
prob = min(1, k)
u = runif(1)
if(u < prob){
p[i + 1] = yi
} else{
p[i + 1] = p[i]
}
}
return(mean(p))
}
```

1.6.4 Echantillonneur de Gibbs

Cette méthode a été utilisée par Geman (1984), dans le cadre de la restauration d'images et une forme particulière de la méthode de Monte-Carlo par chaîne de Markov (MCMC) la plus simple. Le principe repose toujours sur la décomposition du problème général (simuler selon certaine loi) en une série de problèmes fondamentaux (simuler selon des lois conditionnelles). Considérons la densité $f(x, y_1, \dots, y_p)$. Intéressons nous à la loi marginale :

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

En particulier, nous voulons obtenir l'espérance mathématique de celui-ci et sa variance. Nous nous plaçons dans une situation où l'intégrale dans le cas où les intégrations impliquées dans le calcul de la marginale sont complexe et difficile à effectuer numériquement. Cependant nous supposons que la densité conditionnelle est disponibles.

L'échantillonneur de Gibbs nous permettra de générer x suivant $f(x)$ sans utiliser son expression directement ce qui est difficile à manipuler, mais utiliser la densités conditionnelles [18].

Faites donc un échantillon (x_1, \dots, x_m) suffisamment grand pour que nous puissions utiliser la loi des grands nombres pour approximer la moyenne la variance et d'autres caractéristiques :

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

Principe de L'échantillonneur de Gibbs :

On considère le cas de base $f(x, y)$, supposons que $f(x|y)$ et $f(y|x)$ soient disponibles. On peut donc générer ce qu'on appelle une suite de Gibbs en : partant de la valeur x_0 , on génère y_0 avec $P(\cdot|x_0)$, puis x_1 avec $P(\cdot|y_0)$ puis y_1 avec $P(\cdot|x_1)$ et ainsi de suite.

Après k itérations de ce schéma, il vient une séquence : $(x_0, y_0, x_1, y_1, \dots, x_k, y_k)$. Pour k suffisamment grand, x_k est une réalisation de X .

Dans le cadre bayésien, tant que la loi conditionnelle $P(\theta_i|\theta_j; x); j \neq i$ peut être exprimée, l'algorithme de Gibbs peut obtenir la réalisation du paramètre $\theta = (\theta_1, \dots, \theta_k)$ avant la loi a posteriori $P(\theta|x)$, l'échantillonnage de Gibbs consiste à partir du vecteur initial $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})$,

Algorithme d'échantillonnage de Gibbs

simuler

$$\theta_1^{(p+1)} = P(\theta_1 | \theta_2^{(p)}, \theta_3^{(p)}, \dots, \theta_k^{(p)}; x)$$

$$\theta_2^{(p+1)} = P(\theta_2 | \theta_1^{(p)}, \theta_3^{(p)}, \dots, \theta_k^{(p)}; x)$$

.

.

.

$$\theta_k^{(p+1)} = P(\theta_k | \theta_1^{(p)}, \theta_2^{(p)}, \dots, \theta_{k-1}^{(p)}; x)$$

Exemple 1.6.2. *Modèle Gaussien (voir [1])*

On veut estimer la moyenne et l'écart type d'une loi Normale par l'approche Bayésienne à partir d'un échantillon (X_1, X_2, \dots, X_n) distribué selon cette loi. $X_i \sim \mathcal{N}(m, \sigma^2)$.

On suppose que la loi a priori de m est gaussienne de moyenne μ_0 et de variance σ_0^2 et la loi a priori de σ est une inverse gamma de paramètre α et β , c'est-à-dire $m \sim \mathcal{N}(\mu_0, \sigma_0^2)$ et $\sigma^2 \sim \mathcal{IG}(\alpha, \beta)$.

On peut montrer en utilisant la formule de Bayes que la loi a posteriori de m est une loi normale et celle de σ^2 une inverse gamma et leurs paramètres sont donnés respectivement par :

- La moyenne :

$$m | \sigma^2, x \sim \mathcal{N}(M, \Sigma^2);$$

avec

$$M = \frac{\frac{\sigma}{n}}{\frac{\sigma^2}{n} + \sigma_0^2} \mu_0 + \frac{\frac{\sigma_0^2}{n}}{\frac{\sigma^2}{n} + \sigma_0^2} \bar{x};$$

et

$$\Sigma^2 = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}};$$

où $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

Algorithme

L'algorithme utilisé est celui d'échantillonneur de Gibbs, il est appliqué sur un échantillon X généré selon la loi normale centrée réduite :

- Générer $X \sim \mathcal{N}(0, 1)$;
- Initialiser les paramètres $\mu_0, \sigma_0^2, \alpha, \beta$;

- À une itération t :
- Générer $m(t) \sim \mathcal{N}(M, \Sigma^2)$;
- Générer $\sigma^{2(t+1)} \sim \mathcal{IG}(\frac{n}{2} + \alpha, \frac{1}{2} \sum_{i=1}^n (x_i - m^t)^2 + \beta)$.

Résultats

Les estimateurs de la moyenne et de la variance obtenus en générant 500 variables aléatoires $X \sim \mathcal{N}(0, 1)$, avec un nombre de période de chauffe de 3000 sur 10000 itérations sont :

m	σ^2
0.0363304	1.009834

(Code source R de problème précédent)

```
library(MCMCpack)
gibbs = fonction(T)
{
x = rnorm(500, 0, 1)
m0 = 0.01
sig0 = 1
a0 = 1
b0 = 0.5
tau = 0
m = 0
tau[1] = 1
n = length(x)
for(i in 1 : T)
{
sig = sqrt((tau[i] * sig0)/(tau[i] + n * sig0))
mea = (tau[i]/(tau[i] + n * sig0)) * m0 + (n * sig0/(tau[i] + n * sig0)) * mean(x)
m[i] = rnorm(1, mea, sig)
a = a0 + (n/2)
b = b0 + 0.5 * sum((x - mu[i])^2)
tau[i + 1] = rinvgamma(1, a, b)
}
```

```
print(mean(m))  
print(mean(tau))  
}
```

Plans de données

Introduction

L'ensemble des données complètes signifie qu'on a les résultats analytiques pour tous les contributeurs, donc, on parle des données complètes lorsqu'on travaille sur tous les observations de l'échantillon sur lequel on fait notre étude.

Dans ce chapitre on va introduire la notion des données censurées qui sont : les données censurées, les données progressivement censurées et les données hybrides.

La fonction de vraisemblance lorsque les données sont complètes et iid se calcule à l'aide de la formule suivante

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

Exemple 2.0.1. On considère que X est une v.a.r de la loi exponentielle de paramètre λ , la loi a priori de λ est $\pi(\lambda) = \frac{1}{\lambda}$, $\lambda > 0$. Les données sont supposées des données complètes.

- La fonction de vraisemblance :

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

- L'estimateur du maximum de vraisemblance :

$$\begin{aligned}
 l(\underline{x}; \lambda) &= \ln(L(\underline{x}; \lambda)) \\
 &= \ln(\lambda^n e^{-\lambda \sum_{i=1}^n x_i}) \\
 &= \ln \lambda^n + \ln e^{-\lambda \sum_{i=1}^n x_i} \\
 &= n \ln \lambda - \lambda \sum_{i=1}^n x_i \\
 &= n \ln \lambda - n\lambda \bar{x}.
 \end{aligned}$$

- On dérive par rapport à λ :

$$\begin{aligned}
 \frac{\partial l(\underline{x}; \lambda)}{\partial \lambda} &= \frac{n}{\lambda} - n\bar{x} \\
 &= 0,
 \end{aligned}$$

donc on trouve :

$$\hat{\lambda}_{mle} = \frac{1}{\bar{x}}.$$

- La densité a posteriori avec la loi a priori :

$$\begin{aligned}
 \pi(\lambda|x) &= \frac{L(x, \lambda)\pi(\lambda)}{\int_0^{+\infty} L(x, \lambda)\pi(\lambda)} \\
 &= \frac{\lambda^{n-1} e^{-\lambda \sum_{i=1}^n x_i}}{\int_0^{+\infty} \lambda^{n-1} e^{-\lambda \sum_{i=1}^n x_i} d\lambda} \\
 &\propto \lambda^{n-1} e^{-\lambda \sum_{i=1}^n x_i} \\
 &\propto G(n, \sum_{i=1}^n x_i).
 \end{aligned}$$

- Estimateur Bayésien de λ sous la fonction de perte quadratique :

$$\begin{aligned}
 \hat{\lambda}_Q &= E_\pi(\lambda|x) \\
 &= \int_0^{+\infty} \lambda \pi(\lambda|x) d\lambda \\
 &= \frac{n}{\sum_{i=1}^n x_i}.
 \end{aligned}$$

2.1 Les données censurées

Une caractéristique importante de l'analyse de la survie est la présence des données censurées. Les données censurées sont des observations pour lesquelles la valeur exacte d'un évé-

ment n'est pas toujours connue. Cependant, nous disposons, tout de même, d'une information partielle permettant de fixer une borne inférieure (censure à droite) ou une borne supérieure (censure à gauche) une borne inférieure et supérieure (censure par intervalle) [11].

La censure est le phénomène le plus couramment rencontré lors du recueil de survie pour l'individu i , considérons :

X_i son temps de survie ;

C_i son temps de censure ;

T_i la durée réellement observée ;

Il existe trois catégories de censure qu'on nomme censure à droite, censure à gauche et censure par intervalle.

2.1.1 Censure à droite

La durée de vie est dite censuré à droite si l'individu n'a pas subi l'évènement à sa dernière observation. En présence de censure à droite, les durées de vie ne sont pas toutes observées ; pour certaines d'entre elles, on sait seulement qu'elles sont supérieures à une certaine valeur connue [10].

$$T = X \wedge C = \min(X, C)$$

Ce modèle de censure est applicable lorsque l'évènement considéré est le décès du patient et que la date de fin d'étude est prédéterminée.

2.1.2 Censure à gauche

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

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

$$\delta = I_{\{X_i \geq C\}}.$$

2.1.3 Censure par intervalle

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

Par exemple, si l'on s'intéresse à l'âge de survenue d'une maladie et que le sujet i est diagnostiqué malade au cours d'une visite, on sait seulement que $T_i \in [M_i, N_i]$ où M_i est l'âge à la visite de diagnostic et N_i est l'âge à la visite précédente.

1- La censure de type I (censure fixé)

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

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

Ce mécanisme de censure est fréquemment rencontré dans les applications industrielles. Par exemple, on peut tester la durée de vie de n objets identiques (ampoules) sur un intervalle d'observations fixé $[0, u]$.

2- La censure de type II (censure aléatoire)

On décide d'observer les durées de survie des n patients jusqu'à ce que r d'entre eux soient décédés et d'arrêter l'étude à ce moment là.

Si $X_{(i)}$ un statistique d'ordre $X_{(1)}, X_{(2)}, \dots, X_{(n)}$, donc $X_{(1)}$ est plus petite que $X_{(2)}$ et ... ainsi de suite jusqu'à $n^{iem} X_{(n)}$. La date de censure est donc $X_{(r)}$ et on observe les variables suivantes :

$$\begin{aligned} T_{(1)} &= X_{(1)} \\ &\cdot \\ &\cdot \\ &\cdot \\ T_{(r)} &= X_{(r)} \\ T_{(r+1)} &= X_{(r)} \\ &\cdot \\ &\cdot \\ &\cdot \\ T_{(n)} &= X_{(r)} \end{aligned}$$

Dans cas des données censurées, la fonction de vraisemblance est donnée par la formule suivante :

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

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

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

$$T_i = X_i \wedge C_i$$

L'information disponible peut être résumé par :

- La durée réellement observé T_i .
- Un individu $\delta_i = I_{\{X_i \leq C_i\}}$ noté δ_i avec
- $\delta_i = 1$ si l'événement est observé (d'où $T_i = X_i$). On observe les vraies durées ou les durées complètes.
- $\delta_i = 0$ si l'individu est censuré (d'où $T_i = C_i$). On observe des durées incomplètes (censurées).

Exemple 2.1.1. Suite de l'Exemple 2.0.1, on suppose que les données sont censurées.

- La fonction de vraisemblance :

$$\begin{aligned} L(\underline{x}; \lambda) &= \frac{n!}{(n-m)!} \prod_{i=1}^m f(x_i, \lambda) [1 - F(x_m; \lambda)]^{n-m} \\ &= \frac{n!}{(n-m)!} \prod_{i=1}^m \lambda e^{-\lambda x_i} [e^{-\lambda x_m}]^{n-m} \\ &= \frac{n!}{(n-m)!} \lambda^m e^{-\lambda \sum_{i=1}^m x_i} [e^{-\lambda x_m}]^{n-m}. \end{aligned}$$

- L'estimateur de maximum de vraisemblance :

$$\begin{aligned} l(\underline{x}; \lambda) &= \ln(L(\underline{x}; \lambda)) \\ &= \ln \left(\frac{n!}{(n-m)!} \lambda^m e^{-\lambda \sum_{i=1}^m x_i} [e^{-\lambda x_m}]^{n-m} \right) \\ &= \ln \frac{n!}{(n-m)!} + \ln \lambda^m + \ln e^{-\lambda \sum_{i=1}^m x_i} + \ln [e^{-\lambda x_m}]^{n-m} \\ &= \ln n! - \ln(n-m)! + m \ln \lambda - \lambda m \bar{x} - (n-m) \lambda x_m. \end{aligned}$$

- On dérive par rapport à λ :

$$\begin{aligned} \frac{\partial l(\underline{x}; \lambda)}{\partial \lambda} &= \frac{m}{\lambda} - m \bar{x} - (n-m) x_m - m \\ &= 0, \end{aligned}$$

donc on trouve :

$$\widehat{\lambda}_{mle} = \frac{m}{m\bar{x} + (n - m)x_m}.$$

- La densité a posteriori pour une a priori :

$$\begin{aligned} \pi(\lambda|x) &= \frac{L(x, \lambda)\pi(\lambda)}{\int_0^{+\infty} L(x, \lambda)\pi(\lambda)d\lambda} \\ &= \frac{\frac{n!}{(n-m)!} \lambda^{m-1} e^{-\lambda \sum_{i=1}^m x_i} e^{-\lambda x_m(n-m)}}{\frac{n!}{(n-m)!} \int_0^{+\infty} \lambda^{m-1} e^{-\lambda \sum_{i=1}^m x_i} e^{-\lambda x_m(n-m)} d\lambda} \\ &\propto \lambda^{m-1} e^{-\lambda(\sum_{i=1}^m x_i + (n-m)x_m)} \\ &\propto G\left(m, \sum_{i=1}^m x_i + (n - m)x_m\right). \end{aligned}$$

- Estimateur Bayésien de λ sous la fonction de perte quadratique :

$$\begin{aligned} \widehat{\lambda}_Q &= E_\pi(\lambda|x) \\ &= \frac{\int_0^{+\infty} \lambda \pi(\lambda|x) d\lambda}{\int_0^{+\infty} \pi(\lambda|x) d\lambda} \\ &= \frac{m}{\sum_{i=1}^m x_i + (n - m)x_m}. \end{aligned}$$

2.2 Données progressivement censurées

Les données progressivement censurées peuvent être décrites de la manière suivante : supposons que n unités indépendantes soient testées et que la quantité de censure $R = (R_1, R_2, \dots, R_m)$ ait été déterminée. Lors du premier défaillance appelé $X_{1:m:n}, R_1$ unités sont éliminées au hasard de $n - 1$ unités restantes. Au second défaillance, les $X_{2:m:n}, R_2$ unités sont éliminés au hasard de $n - R_1 - 2$ unités restantes le processus continue ainsi de suite, jusqu'au la m^{ime} défaillance, en disant $X_{m:m:n}$, tout ce qui reste $R = n - R_1 - R_2 - \dots - R_{m-1}$ sont tous éliminés. Aussi puis $X_{1:m:n} < X_{2:m:n} < \dots < X_{m:m:n}$ sont appelés les statistiques d'ordre progressivement censurées [9].

Remarque 2.2.1. On utilise l'algorithme de Balakrishnan et Sandhu pour générer les données progressivement censurées d'une distribution connue avec les cinq étapes suivantes :

- 1- Générer m variables identiquement indépendantes distribuées (iid) (u_1, u_2, \dots, u_m) a partir de la distribution uniforme $U(0; 1)$.
- 2- Soit $z_i = -\log(1 - u_i)$, z_i sont identiquement indépendantes distribuées de la distribution

exponentielle standard.

3- En donnant les censures $R = (R_1, R_2, \dots, R_m)$, soit $y_1 = \frac{z_1}{m_1}$ pour $i = 1 \dots n$

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

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

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

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

La vraisemblance dans le cas d'un plan d'expérience avec des données progressivement censurées, est donnée comme suit :

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

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

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

Exemple 2.2.1. Suite de l'Exemple 2.0.1, dans cette partie les données sont supposées progressivement censurées.

- La fonction de vraisemblance :

$$\begin{aligned} L(\underline{x}; \lambda) &= A \prod_{i=1}^m f(x_i; \lambda) [1 - F(x_i; \lambda)]^{R_i} \\ &= A \prod_{i=1}^m \lambda e^{-\lambda x_i} [e^{-\lambda x_i}]^{R_i} \\ &= A \prod_{i=1}^m \lambda e^{-\lambda x_i} e^{-\lambda x_i R_i} \\ &= A \lambda^m e^{-(1 + \sum_{i=1}^m R_i) \lambda \sum_{i=1}^m x_i}. \end{aligned}$$

- L'estimateur du maximum de vraisemblance :

$$\begin{aligned} l(\underline{x}; \lambda) &= \ln(L(\underline{x}; \lambda)) \\ &= \ln \left(A \lambda^m e^{-(1 + \sum_{i=1}^m R_i) \lambda \sum_{i=1}^m x_i} \right). \\ &= \ln A + m \ln \lambda + \ln e^{-(1 + \sum_{i=1}^m R_i) \sum_{i=1}^m x_i \lambda} \\ &= \ln A + m \ln \lambda - (1 + \sum_{i=1}^m R_i) m \bar{x} \lambda. \end{aligned}$$

- On dérive par rapport à λ :

$$\begin{aligned}\frac{\partial l(\underline{x}; \lambda)}{\partial \lambda} &= \frac{m}{\lambda} - (1 - \sum_{i=1}^m R_i)m\bar{x} \\ &= 0,\end{aligned}$$

donc on trouve :

$$\hat{\lambda}_{mle} = \frac{1}{(1 - \sum_{i=1}^m R_i)\bar{x}}.$$

- La densité a posteriori avec la loi a priori :

$$\begin{aligned}\pi(\lambda|x) &= \frac{L(x, \lambda)\pi(\lambda)}{\int_0^{+\infty} L(x, \lambda)\pi(\lambda)d\lambda} \\ &= \frac{A\lambda^{m-1}e^{-(1+\sum_{i=1}^m R_i)\lambda \sum_{i=1}^m x_i}}{\int_0^{+\infty} A\lambda^{m-1}e^{-(1+\sum_{i=1}^m R_i)\lambda \sum_{i=1}^m x_i}d\lambda} \\ &\propto \lambda^{m-1}e^{-(1+\sum_{i=1}^m R_i)\lambda \sum_{i=1}^m x_i} \\ &\propto G\left(m, (1 + \sum_{i=1}^m R_i) \sum_{i=1}^m x_i\right)\end{aligned}$$

- Estimateur Bayésien de λ sous la fonction de perte quadratique :

$$\begin{aligned}\hat{\lambda}_Q &= E_{\pi}(\lambda|x) \\ &= \int_0^{+\infty} \lambda\pi(\lambda|x)d\lambda \\ &= \frac{m}{(1 + \sum_{i=1}^m R_i) \sum_{i=1}^m X_i}.\end{aligned}$$

2.3 Données hybrides de type I

Les données hybride peuvent être décrites de la manière suivante :

Supposons qu' on a un n échantillon de variables identiquement distribuées indépendantes (iid) d'une expérience, m est un entier positif prédéterminé qui satisfait $m < n$ et t_1, t_2 sont les temps de censuré qui sont déterminé à l'avance et satisfaitent $0 < t_1 < t_2$. Dans le temps 0, les moments d'échec sont respectivement X_1, X_2, \dots, X_n .

On note :

$$\delta_i = \begin{cases} 1, & X_i \leq t_1 \\ 0, & X_i > t_1 \end{cases} \quad \text{et :} \quad \eta_i = \begin{cases} 1, & X_i \leq t_2 \\ 0, & X_i > t_2 \end{cases} \quad (i=1,2,\dots,n)$$

$$\sum_{i=1}^n \delta_i = m_1 \quad \text{et} \quad \sum_{i=1}^n \eta_i = m_2.$$

Tout d'abord le nombre des échantillons échoués avant t_1 est m_1 , si $m_1 > m$ le test sera arrêté à l'instant t_1 et les $n - m_1$ composantes défaillance seront retirées du teste où $0 < X_{1:n}, \leq X_{2:n} \leq \dots, \leq X_{m_1:n} \leq t_1 \leq X_{m_1+1:n}$ sont les moments de défaillance séquentiels. si $m_1 < m$ le teste est continu et s'arrête à l'instant t_2 . A cet instant le nombre d'échantillons ratés est m_2 et $0 < X_{1:n}, \leq X_{2:n} \leq \dots, \leq X_{m_2:n} \leq t_2$ sont les moments de défaillance séquentiels. L'avidement m_1 et m_2 sont des variables aléatoires [2].

Les deux cas ci-dessus sont, respectivement, notés Cas 1 et Cas 2 ainsi les données d'observation suivantes sont obtenues :

$$k = \begin{cases} m_1 & , \text{cas 1} \\ m_2 & , \text{cas 2} \end{cases} \quad \text{et} \quad t = \begin{cases} t_1 & , \text{cas 1} \\ t_2 & , \text{cas 2} \end{cases}$$

Où : **cas 1** est $m_1 \leq m$ et **cas 2** est $m_1 > m$.

La vraisemblance associée est donnée par :

$$L(x_i, \theta) = \prod_{i=1}^k f(x_i; \theta) [1 - F(t; \theta)]^{n-k}.$$

Exemple 2.3.1. Suite de l'Exemple 2.0.1, on suppose qu'on a des données hybrides.

- La fonction vraisemblance :

$$\begin{aligned} L(\underline{x}; \theta) &= \prod_{i=1}^k f(x_i; \theta) [1 - F(t; \theta)]^{n-k} \\ &= \prod_{i=1}^k \lambda e^{-\lambda x_i} [e^{-\lambda t}]^{n-k} \\ &= \lambda^k e^{-\lambda \sum_{i=1}^k x_i} [e^{-\lambda t}]^{n-k}. \end{aligned}$$

- L'estimateur du maximum de vraisemblance :

$$\begin{aligned}
 l(\underline{x}; \lambda) &= \ln(L(\underline{x}; \lambda)) \\
 &= \ln(\lambda^k e^{-\lambda \sum_{i=1}^k x_i} [e^{-\lambda t}]^{n-k}) \\
 &= \ln \lambda^k + \ln e^{-\lambda \sum_{i=1}^k x_i} + \ln [e^{-\lambda t}]^{n-k} \\
 &= k \ln \lambda - \lambda \sum_{i=1}^k x_i - (n-k)\lambda t.
 \end{aligned}$$

- On dérive par rapport à λ :

$$\begin{aligned}
 \frac{\partial l(\underline{x}; \lambda)}{\partial \lambda} &= \frac{k}{\lambda} - k\bar{x} - (n-k)t \\
 &= 0,
 \end{aligned}$$

- On trouve :

$$\hat{\lambda}_{mle} = \frac{k}{k\bar{x} + (n-k)t}.$$

- La densité a posteriori avec la loi a priori :

$$\begin{aligned}
 \pi(\lambda|x) &= \frac{L(x, \lambda)\pi(\lambda)}{\int_0^{+\infty} L(x, \lambda)\pi(\lambda)d\lambda} \\
 &\propto L(x, \lambda)\pi(\lambda). \\
 &\propto \lambda^{k-1} e^{-(\sum_{i=1}^k x_i + (n-k)t)\lambda} \\
 &\propto G\left(k, \sum_{i=1}^k x_i + (n-k)t\right).
 \end{aligned}$$

- Estimateur Bayésien de λ sous la fonction de perte quadratique :

$$\begin{aligned}
 \hat{\lambda}_Q &= E_\pi(\lambda|x) \\
 &= \frac{\int_0^{+\infty} \lambda \pi(\lambda|x) d\lambda}{\int_0^{+\infty} \pi(\lambda|x) d\lambda} \\
 &= \frac{k}{\sum_{i=1}^k x_i + (n-k)t}.
 \end{aligned}$$

On remarque qu'il y a une différence entre les estimation dans le cas fréquentiste et le cas bayésienne.

Application sous R

Introduction

Dans le contexte de l'analyse de survie, la distribution de Rayleigh-Pareto notée (RP) joue un rôle très important dans la modélisation de certains phénomènes aléatoires. Plusieurs auteurs ont travaillé sur cette distribution dont al-kadim et Mohamed (2018), Zeghdoudi H et al (2020),...etc.

Dans ce chapitre, nous nous intéressons à l'estimation des paramètres de la loi de Rayleigh-Pareto en utilisant des données hybrides de type I. Dans la première partie, on utilise une approche classique ; celle du maximum de vraisemblance pour estimer les paramètres, et dans la deuxième partie on utilise l'approche Bayésienne avec une loi a priori qui est une mixture entre deux lois a priori conjuguées naturelles et une loi a priori vague (non informative) pour les paramètres sous deux types de fonctions de pertes : une fonction de perte symétrique qui est la fonction de perte quadratique puis une fonction de perte asymétrique celle de Linex.

3.1 Modèle

On considère la loi de Rayleigh-Pareto à trois paramètres α , β et γ . Dans cette section, on va citer les fonctions caractéristiques de cette distribution :

Fonction de densité f

La fonction de densité de probabilité de la loi de Rayleigh-Pareto à trois paramètres α et β et γ est définie par :

$$f(x; \alpha, \beta, \gamma) = \frac{\alpha}{2\beta^2\gamma} \left(\frac{x}{\gamma}\right)^{\alpha-1} \exp\left(-\frac{1}{2\beta^2} \left(\frac{x}{\gamma}\right)^\alpha\right), \quad x > 0.$$

lorsque :

- $\alpha=1$: la distribution RP se déduit à la distribution exponentielle avec le paramètre $\lambda = \frac{\alpha}{2\beta^2\gamma}$.
- $\beta=\sqrt{\frac{1}{2}}$: la distribution RP se déduit à la distribution de Weibul à deux paramètre $W(x; \alpha, \gamma)$.

Fonction de répartition F

La fonction de répartition (F.r ou cdf en anglais pour "cumulative distribution fonction") est donnée par :

$$F(x; \alpha, \beta, \gamma) = \int_0^{\frac{1}{1-G(x)}} g(t) dt,$$

où :

- $G(x; \alpha, \gamma) = 1 - \exp\left(-\frac{x}{\gamma}\right)^\alpha$ est la fonction de répartition de la loi de Pareto.
- $g(x, \beta) = \frac{x}{\beta^2} \exp\left(-\frac{x^2}{2\beta^2}\right)$ est la fonction de densité de la loi de Rayleigh à deux paramètres.

Donc :

$$F(x; \alpha, \beta, \gamma) = 1 - \exp\left(-\frac{1}{2\beta^2} \left(\frac{x}{\gamma}\right)^\alpha\right); \quad \alpha, \beta, \gamma > 0.$$

Fonction de fiabilité R (Reliability function)

La fonction de fiabilité de la loi de Rayleigh-Pareto est définie par :

$$\begin{aligned} R(x; \alpha, \beta, \gamma) &= 1 - F(x; \alpha, \beta, \gamma) \\ &= \exp\left(-\frac{1}{2\beta^2} \left(\frac{x}{\gamma}\right)^\alpha\right), \quad x > 0 \end{aligned}$$

La représentation graphique de la fonction de fiabilité pour différentes valeurs de α et pour les valeurs $\beta = 0.25$, $\gamma = 1.5$, est donnée dans la Figure 3.1.

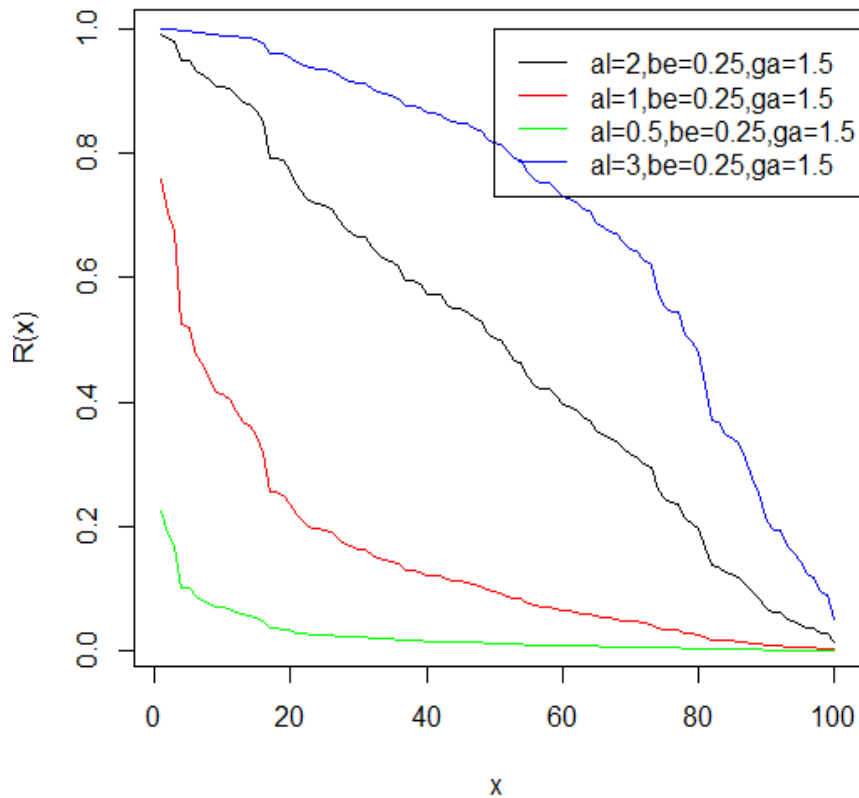


FIGURE 3.1 – Graphe de la fonction de fiabilité $R(x)$ avec différentes valeurs de α .

Fonction de taux de panne h (hasard)

La fonction de taux de hasard h de la loi de Rayleigh-Pareto est calculée à l'aide de la formule suivante :

$$\begin{aligned} h(x; \alpha, \beta, \gamma) &= \frac{f(x; \alpha, \beta, \gamma)}{1 - F(x; \alpha, \beta, \gamma)} \\ &= \frac{f(x; \alpha, \beta, \gamma)}{R(x; \alpha, \beta, \gamma)} \end{aligned}$$

Donc :

$$h(x; \alpha, \beta, \gamma) = \frac{\alpha}{2\beta^2\gamma} \left(\frac{x}{\gamma}\right)^{\alpha-1}, \quad x > 0.$$

La Figure 3.2 représente la fonction de hasard pour différentes valeurs de α et pour $\beta = 0.25$, $\gamma = 1.5$.

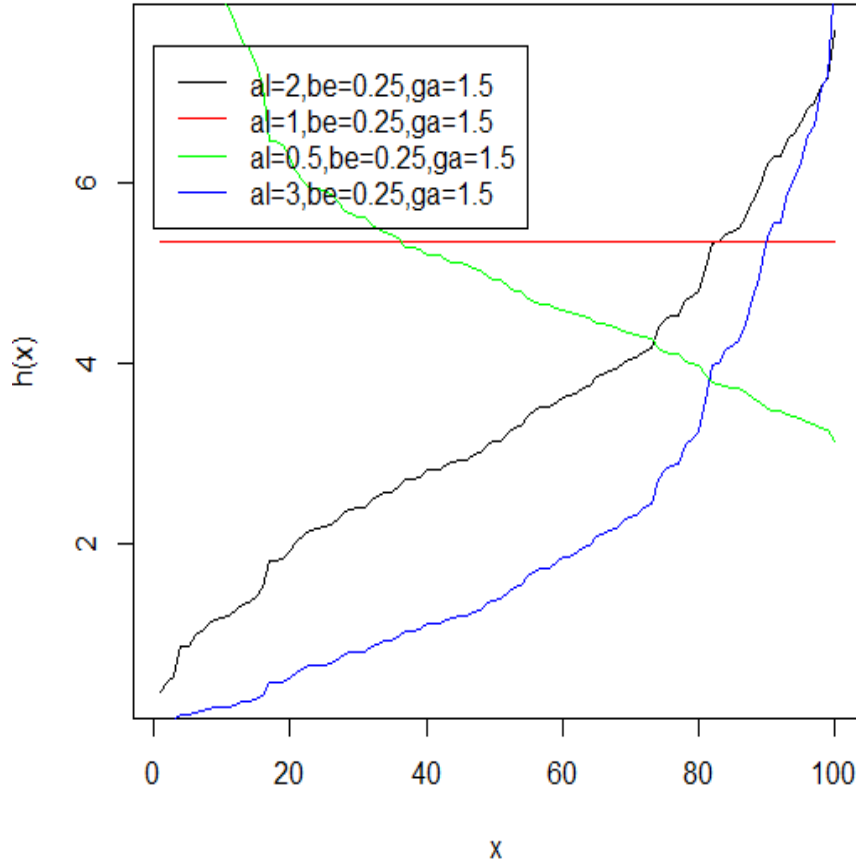


FIGURE 3.2 – Graphe de la fonction de taux de panne $h(x)$ avec différentes valeurs de α

3.2 Estimation par la méthode de maximum vraisemblance

On considère, dans cette section, que des données sont hybrides de type I. Dans ce cas, la fonction de vraisemblance se calcule à l'aide de la formule suivante :

$$L(\underline{x}; \alpha, \beta, \gamma) = \prod_{i=1}^k f(x_i; \alpha, \beta, \gamma) [1 - F(t; \alpha, \beta, \gamma)]^{n-k}$$

$$\text{avec : } k = \begin{cases} m_1 & , \text{cas 1} \\ m_2 & , \text{cas 2} \end{cases} \quad \text{et : } t = \begin{cases} t_1 & , \text{cas 1} \\ t_2 & , \text{cas 2} \end{cases}$$

$$\begin{aligned}
 L(\underline{x}; \alpha, \beta, \gamma) &= \prod_{i=1}^k \frac{\alpha}{2\beta^2\gamma} \left(\frac{x_i}{\gamma}\right)^{\alpha-1} \exp\left(-\frac{1}{2\beta^2} \left(\frac{x_i}{\gamma}\right)^\alpha\right) \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} \\
 &= k \ln \alpha - k \ln(2\beta^2\gamma^\alpha) + (\alpha - 1) \sum_{i=1}^k \ln x_i - \frac{1}{2\beta^2\gamma^\alpha} \sum_{i=1}^k x_i^\alpha - (n - k) \left[\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right].
 \end{aligned}$$

- Prenons le logarithme de la fonction de vraisemblance, on obtient :

$$\begin{aligned}
 l(\underline{x}; \alpha, \beta, \gamma) &= \ln(L(\underline{x}; \alpha, \beta, \gamma)) \\
 &= k \ln \left(\frac{\alpha}{2\beta^2\gamma^\alpha}\right) + (\alpha - 1) \ln \prod_{i=1}^k x_i - \frac{1}{2\beta^2\gamma^\alpha} \sum_{i=1}^k x_i^\alpha \\
 &\quad + (n - k) \left[-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right] \\
 &= k \ln \alpha - k \ln(2\beta^2\gamma^\alpha) + (\alpha - 1) \sum_{i=1}^k \ln x_i - \frac{1}{2\beta^2\gamma^\alpha} \sum_{i=1}^k x_i^\alpha \\
 &\quad - (n - k) \left[\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right].
 \end{aligned}$$

- On dérive par rapport à α , β et γ :

$$\begin{aligned}
 \frac{\partial l(\underline{x}; \alpha, \beta, \gamma)}{\partial \alpha} &= \frac{k}{\alpha} - k \ln \gamma + \sum_{i=1}^k \ln x_i + \frac{\ln \gamma}{2\beta^2\gamma^\alpha} \sum_{i=1}^k x_i^\alpha - \frac{1}{2\beta^2\gamma} \sum_{i=1}^k \ln x_i (x_i)^\alpha \\
 &\quad - (n - k) \left[\frac{1}{2\beta^2} \ln \left(\frac{k}{\gamma}\right) \left(\frac{t}{\gamma}\right)^\alpha\right]. \\
 \frac{\partial l(\underline{x}; \alpha, \beta, \gamma)}{\partial \beta} &= \frac{-2k}{\beta} + \frac{1}{\beta^3\gamma^\alpha} \sum_{i=1}^k x_i^\alpha + (n - k) \left[\frac{1}{\beta^3} \left(\frac{t}{\gamma}\right)^\alpha\right]. \\
 \frac{\partial l(\underline{x}; \alpha, \beta, \gamma)}{\partial \gamma} &= -k \frac{\alpha}{\gamma} + \frac{\alpha\gamma^{\alpha-1}}{2\beta^2\gamma^{2\alpha}} \sum_{i=1}^k x_i^\alpha + (n - k) \left(\frac{1}{2\beta^2} \frac{\alpha\gamma^{\alpha-1}t^\alpha}{\gamma^{2\alpha}}\right).
 \end{aligned}$$

$\hat{\alpha}_{MV}$, $\hat{\beta}_{MV}$ et $\hat{\gamma}_{MV}$ sont la solution du système d'équations suivant :

$$\begin{cases} \frac{\partial l(\underline{x}; \alpha, \beta, \gamma)}{\partial \alpha} = 0. \\ \frac{\partial l(\underline{x}; \alpha, \beta, \gamma)}{\partial \beta} = 0. \\ \frac{\partial l(\underline{x}; \alpha, \beta, \gamma)}{\partial \gamma} = 0. \end{cases}$$

- On ne peut pas obtenir les expressions analytiques explicites des estimateurs; $\hat{\alpha}_{MV}$, $\hat{\beta}_{MV}$ et $\hat{\gamma}_{MV}$ donc en utilise les méthodes numériques en particulier le package (BBsolve, optim) du R.

Package BBsolve

C'est un package qui présente des performances très élevées pour les systèmes d'équations non-linéaires, son code R est :

$$mle = BBsolve(par = value; fn = g)$par;$$

où value : est l'ensemble des valeurs initiales.

g : présente les dérivées premières de la fonction du maximum de vraisemblance .

3.3 Estimation Bayésienne

On considère les lois a priori pour les paramètres de la loi RP suivantes :

$$\alpha \sim \mathcal{G}(c, d) \Rightarrow \pi(\alpha) = \frac{d^c}{\Gamma(c)} \alpha^{c-1} \exp(-d\alpha).$$

$$\gamma \sim \mathcal{G}(a, b) \Rightarrow \pi(\gamma) = \frac{b^a}{\Gamma(a)} \gamma^{a-1} \exp(-b\gamma),$$

et :

β est une loi a priori impropre, donc : $\pi(\beta) = \frac{1}{\beta}$.

La loi a priori de triplet (α, β, γ) est :

$$\begin{aligned} \pi(\alpha, \beta, \gamma) &= \pi(\alpha)\pi(\beta)\pi(\gamma) \\ &= \frac{b^a d^c}{\Gamma(a)\Gamma(c)} \frac{a^{c-1} \gamma^{a-1}}{\beta} \exp(-d\alpha - b\gamma). \end{aligned}$$

La probabilité a posteriori est calculée à l'aide de la formule suivant :

$$\begin{aligned} \pi(\alpha, \beta, \gamma | \underline{x}) &= \frac{L(\underline{x}; \alpha, \beta, \gamma) \pi(\alpha, \beta, \gamma)}{\iiint_{\Theta} L(\underline{x}; \alpha, \beta, \gamma) \pi(\alpha, \beta, \gamma) d\alpha d\beta d\gamma} \\ &= \frac{L(\underline{x}; \alpha, \beta, \gamma) \pi(\alpha, \beta, \gamma)}{m(x)}. \end{aligned}$$

avec :

$m(x) = \iiint_{\Theta} L(\underline{x}; \alpha, \beta, \gamma) \pi(\alpha, \beta, \gamma) d\alpha d\beta d\gamma$ c'est loi marginale de x .

D'où la loi a posteriori pour un écha \underline{x} est :

$$\begin{aligned} \pi(\alpha, \beta, \gamma | \underline{x}) &= \frac{\left(\frac{\alpha}{2\beta^2\gamma^\alpha}\right)^k \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\frac{1}{2\beta^2\gamma^\alpha} \sum_{i=1}^k x_i^\alpha\right) \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} \\ &\quad \times \frac{b^a d^c}{\Gamma(a)\Gamma(c)} \frac{\alpha^{c-1} \gamma^{a-1}}{\beta} \exp(-d\alpha - b\gamma)}{\iiint_{\Theta} \left(\frac{\alpha}{2\beta^2\gamma^\alpha}\right)^k \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\frac{1}{2\beta^2\gamma^\alpha} \sum_{i=1}^k x_i^\alpha\right) \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} \\ &\quad \times \frac{b^a d^c}{\Gamma(a)\Gamma(c)} \frac{\alpha^{c-1} \gamma^{a-1}}{\beta} \exp(-d\alpha - b\gamma) d\alpha d\beta d\gamma} \end{aligned}$$

$$\begin{aligned}
 &= \frac{\frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right) \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k}}{\iint\int_{\Theta} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right) \\
 &\quad \times \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} d\alpha d\beta d\gamma} \\
 &= A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right) \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k}.
 \end{aligned}$$

avec :

$$\begin{aligned}
 A^{-1} &= \iiint_{\Theta} \left(\frac{\alpha}{2\beta^2 \gamma^\alpha}\right)^k \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha\right) \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} \\
 &\quad \times \frac{b^a d^c}{\Gamma(a)\Gamma(c)} \frac{\alpha^{c-1} \gamma^{a-1}}{\beta} \exp(-d\alpha - b\gamma) d\alpha d\beta d\gamma.
 \end{aligned}$$

3.3.1 Estimation Bayésienne sous la fonction de perte quadratique

Sous la fonction de perte quadratique, l'estimateur de Bayes est la moyenne a posteriori :

$$\begin{aligned}
 \hat{\alpha}_Q &= E_\pi(\alpha|\underline{x}) = \iiint_{\Theta} \alpha \pi(\alpha, \beta, \gamma|\underline{x}) d\alpha d\beta d\gamma \\
 &= \iiint_{\Theta} \alpha A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(\left(-\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right)\right) \\
 &\quad \times \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} d\alpha d\beta d\gamma \\
 &= A^{-1} \iiint_{\Theta} \frac{\alpha^{k+c}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right)\right) \\
 &\quad \times \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} d\alpha d\beta d\gamma.
 \end{aligned}$$

► Le risque a posteriori :

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

$$\begin{aligned}
 \widehat{\beta}_Q &= E_\pi(\beta|\underline{x}) = \iiint_{\Theta} \beta \pi(\alpha, \beta, \gamma|\underline{x}) \, d\alpha \, d\beta \, d\gamma \\
 &= \iiint_{\Theta} \beta A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right)\right) \\
 &\quad \times \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} \, d\alpha \, d\beta \, d\gamma \\
 &= A^{-1} \iiint_{\Theta} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right)\right) \\
 &\quad \times \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} \, d\alpha \, d\beta \, d\gamma.
 \end{aligned}$$

► Le risque a posteriori :

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

$$\begin{aligned}
 \widehat{\gamma}_Q &= E_\pi(\gamma|\underline{x}) = \iiint_{\Theta} \gamma \pi(\alpha, \beta, \gamma|\underline{x}) \, d\alpha \, d\beta \, d\gamma \\
 &= \iiint_{\Theta} \gamma A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right)\right) \\
 &\quad \times \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} \, d\alpha \, d\beta \, d\gamma \\
 &= A^{-1} \iiint_{\Theta} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp\left(-\left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma\right)\right) \\
 &\quad \times \left[\exp\left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma}\right)^\alpha\right)\right]^{n-k} \, d\alpha \, d\beta \, d\gamma.
 \end{aligned}$$

► Le risque a posteriori :

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

3.3.2 Estimation Bayésienne sous la fonction de perte Linex

Les estimateurs Bayésiennes des paramètres α , β et γ sous la fonction de perte Linex sont les suivants :

$$\begin{aligned}
\hat{\alpha}_L &= -\frac{1}{r} \ln[E_\pi(e^{-r\alpha})], \quad r \neq 0 \\
&= -\frac{1}{r} \ln \left[\iiint_{\Theta} e^{-r\alpha} A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp \left(- \left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma \right) \right) \right. \\
&\quad \left. \times \left[\exp \left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma} \right)^\alpha \right) \right]^{n-k} d\alpha d\beta d\gamma \right] \\
&= -\frac{1}{r} \ln \left[\iiint_{\Theta} A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp \left(- \left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma + r\alpha \right) \right) \right. \\
&\quad \left. \times \left[\exp \left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma} \right)^\alpha \right) \right]^{n-k} d\alpha d\beta d\gamma \right].
\end{aligned}$$

► Le risque a posteriori :

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

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

$$\begin{aligned}
\hat{\beta}_L &= -\frac{1}{r} \ln[E_\pi(e^{-r\beta})], \quad r \neq 0 \\
&= -\frac{1}{r} \ln \left[\iiint_{\Theta} e^{-r\beta} A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp \left(- \left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma \right) \right) \right. \\
&\quad \left. \times \left[\exp \left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma} \right)^\alpha \right) \right]^{n-k} d\alpha d\beta d\gamma \right] \\
&= -\frac{1}{r} \ln \left[\iiint_{\Theta} A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp \left(- \left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma + r\beta \right) \right) \right. \\
&\quad \left. \times \left[\exp \left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma} \right)^\alpha \right) \right]^{n-k} d\alpha d\beta d\gamma \right].
\end{aligned}$$

► Le risque a posteriori :

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

$$\begin{aligned} \hat{\gamma}_L &= -\frac{1}{r} \ln[E_\pi(e^{-r\gamma})], \quad r \neq 0 \\ &= -\frac{1}{r} \ln \left[\iiint_{\Theta} e^{-r\gamma} A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp \left(- \left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma \right) \right) \right. \\ &\quad \left. \times \left[\exp \left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma} \right)^\alpha \right) \right]^{n-k} d\alpha d\beta d\gamma \right] \\ &= -\frac{1}{r} \ln \left[\iiint_{\Theta} A^{-1} \frac{\alpha^{k+c-1}}{2^k \beta^{2k+1} \gamma^{k\alpha-a+1}} \prod_{i=1}^k x_i^{\alpha-1} \exp \left(- \left(\frac{1}{2\beta^2 \gamma^\alpha} \sum_{i=1}^k x_i^\alpha + d\alpha + b\gamma + r\gamma \right) \right) \right. \\ &\quad \left. \times \left[\exp \left(-\frac{1}{2\beta^2} \left(\frac{t}{\gamma} \right)^\alpha \right) \right]^{n-k} d\alpha d\beta d\gamma \right]. \end{aligned}$$

► Le risque a posteriori :

$$PR(\hat{\gamma}_L) = r(\hat{\gamma}_Q - \hat{\gamma}_L).$$

• Nous ne pouvons pas calculer les expressions analytiques pour ces estimateurs, pour cela, nous utilisons des méthodes numériques pour trouver les valeurs des estimateurs bayésiens des paramètres α , β et γ et leurs risques a posteriori.

3.4 Application

Dans cette section, on va générer $N=1000$ échantillons de différentes tailles n ($n = 20, n = 30, n = 50$) de la loi de Rayleigh-Pareto, dont ses paramètres initiales sont $\alpha = 2$, $\beta = 1$ et $\gamma = 1.5$, et on fait une estimation des paramètres en utilisant un plan des données hybrides de type I.

Dans la première partie, on a appliqué l'estimation du maximum de vraisemblance sur les paramètres, et on a obtenu un système d'équation non linéaire, pour cela on a utilisé le package BBSolve du R pour le résoudre [voir le programme 1 dans l'annexe].

Les estimateurs du maximum de vraisemblance avec les erreurs quadratique pour différent taille de n sont dans le tableau suivant :

n	m	t_1	t_2	$\hat{\alpha}_{mle}$ (EQ)	$\hat{\beta}_{mle}$ (EQ)	$\hat{\gamma}_{mle}$ (EQ)
20	12	11	19	1.9848 (0.0002)	0.9965 (1.17×10^{-5})	1.6968 (0.0387)
20	13	13	21	2.1311 (0.0172)	0.7090 (0.0846)	1.5965 (0.0093)
30	17	11	19	2.0278 (0.0008)	1.0430 (0.0018)	1.6474 (0.0217)
30	19	13	21	2.0511 (0.0026)	1.0211 (0.0004)	1.5395 (0.0015)
50	29	11	19	2.1043 (0.0108)	0.9425 (0.0033)	1.5029 (8.2×10^{-6})
50	32	13	21	2.0862 (0.0074)	0.9473 (0.0027)	1.5570 (0.0032)

TABLE 3.1 – Estimateurs de α , β et γ (erreur quadratique) par la méthode du maximum de vraisemblance.

Dans la deuxième partie, on a appliqué l'estimation Bayésienne sur les paramètres et aussi on n'a pas pu obtenir des formules analytiques explicites des paramètres pour cela on a utilisé l'algorithme de Metropolis-Hastings pour trouver les valeurs numériques des paramètres. Le tableau ci-dessus présente les estimateurs Bayésiens des paramètres α , β et γ avec ses risques a posteriori sous la fonction de perte quadratique et pour différentes tailles de n et m [voir le programme 2, sous R, dans l'annexe].

n	m	t_1	t_2	$\hat{\alpha}_Q$ (PR)	$\hat{\beta}_Q$ (PR)	$\hat{\gamma}_Q$ (PR)
20	12	11	19	1.9881 (0.0244)	1.0121 (0.0004)	1.5001 (0.0064)
20	13	13	21	1.9973 (0.0314)	1.0273 (0.0014)	1.5123 (0.0089)
30	17	11	19	1.9532 (0.0502)	1.0032 (0.0002)	1.4782 (0.0127)
30	19	13	21	2.0017 (0.0060)	1.0077 (7.8898)	1.5047 (0.0015)
50	29	11	19	1.5809 (0.8357)	0.8329 (0.1957)	1.2069 (0.4527)
50	32	13	21	2.0067 (0.0047)	1.0107 (0.0007)	1.5087 (0.0017)

TABLE 3.2 – Estimateurs de α , β et γ sous la fonction de perte quadratique (le risque a posteriori).

Le tableau suivant présente les estimateurs de Bayes des paramètres α , β et γ de la loi de Rayleigh-Pareto sous la fonction de perte Linex, avec différentes tailles de n et m . Les résultats sont obtenus en utilisant l'algorithme de Metropolis-Hastings sous R [voir le programme 2 dans l'annexe].

n	m	t_1	t_2	$\hat{\alpha}_L$ (PR)	$\hat{\beta}_L$ (PR)	$\hat{\gamma}_L$ (PR)
20	12	11	19	1.9711 (0.0170)	1.0120 (0.0001)	1.4964 (0.0037)
20	13	13	21	1.9754 (0.0219)	1.0270 (0.0003)	1.5072 (0.0051)
30	17	11	19	1.9205 (0.0327)	1.0031 (0.0001)	1.4711 (0.0071)
30	19	13	21	1.9974 (0.0043)	1.0077 (9.1301×10^{-6})	1.5038 (0.0009)
50	29	11	19	1.1614 (0.4194)	0.7320 (0.1008)	0.9770 (0.2298)
50	32	13	21	2.0035 (0.0032)	1.0104 (0.0002)	1.5078 (0.0009)

TABLE 3.3 – Estimateurs de α , β et γ sous la fonction de perte Linex (le risque a posteriori)

Discussion

On remarque que les trois estimateurs $\hat{\alpha}$, $\hat{\beta}$ et $\hat{\gamma}$ avec l'estimation du maximum de vraisemblance ou l'estimation Bayésienne (sous la fonction de perte quadratique et linex) sont toutes proches de la valeur initiale α_0 , β_0 et γ_0 que nous avons choisi et en plus, on remarque que l'erreur quartique, en générale, sous la fonction de perte linex est inférieur par rapport aux les deux autres estimateurs.

Conclusion

L'objectif de ce travail est d'estimer les paramètres de la loi de Rayleigh-Pareto en utilisant des données hybrides et avec deux méthodes d'estimation : la première est la méthode classique du maximum de vraisemblance et la deuxième est par la méthode d'estimation Bayésienne sous différentes fonctions de coûts. On conclut que les méthodes numériques surtout les MCMC, sont efficaces pour trouver les estimateurs malgré la complexité du modèle statistique.

Une des perspectives de ce travail est d'estimer les caractéristiques (la fonction de fiabilité, la fonction de taux de panne) de cette distribution par les deux approches d'estimations (fréquentiste et bayésienne).

Annexes

Programme 1

Estimation des paramètres par la méthode de maximum de vraisemblance

```
rm(list = ls())
N = 1000
H1 = numeric(N)
H2 = numeric(N)
H3 = numeric(N)
n = 50; m = 30; t1 = 13; t2 = 21
alpha = 2; beta = 1; gamma = 1.5
# générer la loi de RP par la méthode inverse
u = runif(n, 0, 1)
vec = (-2 * (beta^2) * (gamma^alpha) * log(1 - u))^(1/alpha)
x = sort(vec)
# Détermine la valeur de k et t
delta = numeric(n); neta = numeric(n)
for(i in 1 : n){
  if(x[i] < t1){delta[i] = 1}else{delta[i] = 0}
  if(x[i] < t2){neta[i] = 1}else{neta[i] = 0}
}
m1 = sum(delta)
m2 = sum(neta)
```

```

if(m1 > m){k = m1}else{k = m2}
if(m1 > m){t = t1}else{t = t2}
t; k
for(j in 1 : N){
#Q représente les paramètres de la loi
#dd représente les fonctions de la dérivée de la log vraisemblance
g < -function(Q){
dd < -rep(NA, length(Q))
dd[1] = (k/Q[1]) - (k * log(Q[3])) + sum(log(x)) + (log(Q[3])/(2 * (Q[2]^2) * (Q[3]^Q[1]))) *
sum(x^Q[1]) - (1/(2 * (Q[2]^2) * Q[3])) * sum(log(x)^(x^Q[1])) - (n - k) * ((1/(2 * (Q[2]^2))) *
log(k/Q[3]) * ((t/Q[3])^Q[1]))
dd[2] = ((-2 * k)/Q[2]) + (1/((Q[2]^3) * (Q[3]^Q[1]))) * sum(x^Q[1]) + (n - k) * ((1/(Q[2]^3)) *
((t/Q[3])^Q[1]))
dd[3] = (-k * (Q[1]/Q[3])) + ((Q[1] * (Q[3]^(Q[1] - 1)))/(2 * (Q[2]^2) * (Q[3]^(2 * Q[1])))) *
sum(x^Q[1]) + (n - k) * ((1/(2 * (Q[2]^2))) * (Q[1] * (Q[3]^(Q[1] - 1)) * (t^Q[1]))/(Q[3]^(2 * Q[1])))
}
library(BB)
v = BBSolve(par = c(2, 1, 1.5), fn = g)par
H1[j] = v[1]
H2[j] = v[2]
H3[j] = v[3]
}
alphamle = mean(H1)
betamle = mean(H2)
gammamle = mean(H3)
alphamle
betamle
gammamle
ralpha = (alphamle - alpha)^2
rbeta = (betamle - beta)^2
rgamma = (gammamle - gamma)^2
ralpha
rbeta

```


rgamma

Programme 2

Estimation Bayésienne de la loi de RP par la méthode de Metropolis-Hastings

```
N = 1000
H1 = numeric(N)
H2 = numeric(N)
H3 = numeric(N)
n = 20; m = 12; t1 = 11; t2 = 19
alpha = 2; beta = 1; gamma = 1.5
vec = numeric(m)
for(j in 1 : N){
  u = runif(n, 0, 1)
  vec = (-2 * (beta^2) * (gamma^alpha) * log(1 - u))^(1/alpha)
  s = sort(vec)
  # Détermine la valeur de k et t
  delta = numeric(n); neta = numeric(n)
  for(i in 1 : n){
    if(s[i] < t1){delta[i] = 1}else{delta[i] = 0}
    if(s[i] < t2){neta[i] = 1}else{neta[i] = 0}
  }
  m1 = sum(delta)
  m2 = sum(neta)
  if(m1 > m){k = m1}else{k = m2}

  if(m1 > m){t = t1}else{t = t2}
  t; k
  a = 1; b = 2; c = 1; d = 2; aa = 4;
  f = function(x, y, z){(x^(k+c-1))/((2^k)*(y^(2*k+1))*(z^(k*x-a+1)))*prod(s^(x-1))*
  exp(-1/(2*(y^2)*(z*x)))*(sum(s^x)+d*x+b*z)*(exp((-1/(2*(y^2))))*((t/z)^x))^(n-k)}
```

```

q <- function(x, y, z){(x^(aa/2 - 1) * exp(-x/2)) * (y^(aa/2 - 1) * exp(-y/2)) * (z^(aa/2 - 1) * exp(-z/2))};
M = 500; ind = N * 3;
X = matrix(rep(0, ind), ncol = 3, nrow = M);
Y = numeric(3);
Y
X[1, 1] = 1; X[1, 2] = 1; X[1, 3] = 1;
for(i in 2 : M)
{
Y = rchisq(1, 2) + c(2, 1, 1.5);
val = (f(Y[1], Y[2], Y[3]) * q(X[i - 1, 1], X[i - 1, 2], X[i - 1, 3])) / (q(Y[1], Y[2], Y[3]) * f(X[i - 1, 1], X[i - 1, 2], X[i - 1, 3]));
alpha0 = min(1, val); u = runif(1)
if(u < alpha0)
{X[i, 1] < -Y[1]; X[i, 2] < -Y[2]; X[i, 3] < -Y[3]}
else
{X[i, 1] < -X[i - 1, 1]; X[i, 2] < -X[i - 1, 2]; X[i, 3] < -X[i - 1, 3];}
}
H1[k] = mean(X[, 1])
H2[k] = mean(X[, 2])
H3[k] = mean(X[, 3])
}
# Affichage
# Sous la fonction de perte quadratique
alphaQ = mean(X[, 1])
betaQ = mean(X[, 2])
gammaQ = mean(X[, 3])
alphaQ; betaQ; gammaQ
PR1 = mean((X[, 1] - alpha)^2)
PR2 = mean((X[, 2] - beta)^2)
PR3 = mean((X[, 3] - gamma)^2)
PR1; PR2; PR3
# Sous la fonction de perte Linex

```

$$r = 1$$

$$\alpha L = (-1/r) * \log(\text{mean}(\exp(-r * X[, 1])))$$

$$\beta L = (-1/r) * \log(\text{mean}(\exp(-r * X[, 2])))$$

$$\gamma L = (-1/r) * \log(\text{mean}(\exp(-r * X[, 3])))$$

$$\alpha L; \beta L; \gamma L$$

$$PRL1 = r * (\alpha Q - \alpha L)$$

$$PRL2 = r * (\beta Q - \beta L)$$

$$PRL3 = r * (\gamma Q - \gamma L)$$

$$PRL1; PRL2; PRL3$$

Bibliographie

- [1] **Belaïd, N ; Djerroud, L.**(2013). Les méthodes de Monte Carlo : (MCMC et PMC). Applications. Mémoire de master. Université A. Mira- Béjaia.
- [2] **Bing Long.**(2021). Estimation and prediction for the Rayleigh distribution based on double Type-I hybrid censored data. Communications in Statistics - Simulation and Computation.
- [3] **Boudjerda, K.**(2017). Etude de l'estimateur de Bayes sous différentes fonctions de perte. THESE de Doctorat en Mathématique. Université Badji Mokhtar Annaba.
- [4] **Chadli, A ; Kermoune, S.** Reliability Estimation in a Rayleigh Pareto Model with Progressively Type-II Right Censored Data. Dakistan jornal of statistics and Operation Research. 729-743.
- [5] **Christian P. Robert.**(2006). Le choix bayésien Principes et pratique. ISBN-10 : 2-287-25173-1. Springer Paris Berlin Heidelberg New York.
- [6] **Grenon-Godbout, N.**(2015). Méthodes MCMC : amélioration d'un algorithme d'adaptation régionale et applications à la climatologie.
- [7] **Kermoune, S ; Chadli, A.** (2021). BAYESIAN INFERENCE FOR RAYLEIGH PARETO DISTRIBUTION UNDER PROGRESSIVELY TYPE-II RIGHT CENSORED DATA. Advances in Mathematics : Scientific Journal, no.3, 1787–1805.
- [8] **KHRIBI, L.**(2007), COMME EXIGENCE PARTIELLE DE LA MAÎTRISE EN MATHÉMATIQUES. Mémoire de master. UNIVERSITÉ DU QUÉBEC À MONTRÉAL.
- [9] **Krishna, H ; Kumar, K.**(2011). Reliability estimation in Lindley distribution with progressively type II right censored sample. Elsevier B.V. All rights reserved. 281-294.

- [10] **Philippe, S.P.**(2021). Introduction à l'analyse des durées de survie. Université Paul Sabatier-Toulouse III.
- [11] **RABHI, Y.**(2006). MODELES DE SURVIE AVEC UN POINT DE RUPTURE. UNIVERSITÉ DU QUÉBEC À MONTRÉAL.
- [12] **Rousseau, J.**(2010). Statistique Bayésienne Notes de cours.
- [13] **Shababa, B.**(2011). Bayesian analysis. STATS 225.
- [14] **Touraine, C.**(2013). Modèles illness-death pour données censurées par intervalle : Application à l'étude de la démence. THESE de Doctorat en Mathématique. L'UNIVERSITÉ BORDEAUX 2.
- [15] **Varadhan, R and Gilbert, P. D.** (2009). BB An R Package for Solving a Large System of Non linear Equations and Optimizing a High-Dimensional Non linear Objective Function, Journal of Statistical Software, Volume 32, Issue 4.
- [16] <http://web.univ-ubs.fr/lmam/gouno/BAYES/COURS/Cours3.pdf>.
- [17] <http://web.univ-ubs.fr/lmba/gouno/BAYES/COURS/Cours1.pdf>.
- [18] <http://web.univ-ubs.fr/lmam/gouno/BAYES/COURS/Cours6.pdf>.