

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mouloud Mammeri de Tizi-Ouzou

Faculté des Sciences

Département de Mathématiques

Mémoire de Master

Présenté par Maazouz Hamza

En vue de l'obtention du grade de Master

Option : Probabilités et Statistique

Thème

MODÈLE EXPONENTIEL GÉNÉRALISÉ : THÉORIE ET
APPLICATION

Soutenu publiquement le /09/2016 devant le jury suivant :

Président	BOUDIBA M.AREZKI	MCA	U. de Tizi-Ouzou
Encadreur	ZOUGAB NABIL	MCA	U. de Tizi-Ouzou
Examineur	GRAÏCHE FARID	MCB	U. de Tizi-Ouzou
Examineur	YOUSFI SMAIL	MCB	U. de Tizi-Ouzou

Remerciements :

Je remercie Dieu le tout puissant qui a guidé mes pas vers les portes du savoir et qui m'a donné le courage et la volonté pour l'élaboration de ce travail.

Je remercie vivement les membres du jury pour l'honneur qu'ils me font en acceptant de juger ce travail.

Il m'est particulièrement agréable de remercier mon promoteur Monsieur Zougab Nabil, pour m'avoir proposé ce thème, pour ses orientations et sa disponibilité qui m'ont permis de mener mon travail à bien.

Enfin, j'exprime mes remerciements à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

Dédicace :

Je dédie

- A mes très chers parents.
- A mes frères : abdu allah , tahar , Radouane et sa petite famille.
- A mes soeurs : Bouchra , Chahra zed et son mari Habib.
- A toute la famille.
- A tous les enseignants et enseignantes de département de Mathématiques.
- A tous mes amis et plus précisément à :
Ouerdia ,kahina,Aldjia,Lyés , Madjid, Idir, Samir, Chafik, Mohamed, Ali et Hassene
.

Table des matières

Introduction générale	8
1 Modèle exponentiel généralisé	10
Modèle exponentiel généralisé	10
1.1 Définitions et propriétés	10
1.1.1 Fonction de répartition	10
1.1.2 Fonction de densité	11
1.1.3 Fonction de survie	13
1.1.4 Fonction hasard	14
1.1.5 Moments d'ordre k	16
1.1.6 Fonction génératrice des moments	18
1.1.7 Simulation d'une variable aléatoire de loi $EG(\alpha, \lambda)$	19
1.2 Conclusion	19
2 Estimation dans le modèle exponentiel généralisé	20
Estimation dans le modèle exponentiel généralisé	20
2.1 Méthode du maximum de vraisemblance	20
2.1.1 Estimateurs du maximum de vraisemblance	20
2.1.2 Estimateurs par intervalle de confiance	22
2.2 Méthode des moments	23

	4
2.3 Approche bayésienne	24
2.3.1 Principe bayésien	24
2.3.2 Estimateurs de Bayes	25
2.3.2.1 Choix des lois a priori	25
2.3.2.2 Calcul des lois a posteriori	25
2.4 Conclusion	27
3 Application sur des données simulées	29
Application sur des données simulées	29
3.1 Simulation des données	29
3.2 Estimation des paramètres	30
3.3 Résultats	30
3.4 Discussion	31
4 Application aux données réelles	33
Application aux données réelles	33
4.1 Applications réelles	33
4.1.1 Première application	33
4.1.2 Deuxième application	35
4.2 Discussion	36
5 Conclusion	40
Annexe	42
Bibliographie	46

Table des figures

1.1	Fonction de répartition de loi exponentielle généralisée EG(1, 1), EG(3, 1) et EG(5, 1).	11
1.2	Densité de probabilité de loi exponentielle généralisée EG(1, 1), EG(3, 1) et EG(5, 1).	12
1.3	Fonction de survie de loi exponentielle généralisée EG(1, 1), EG(3, 1) et EG(5, 1).	14
1.4	Fonction de hasard de loi exponentielle généralisée EG(1, 1), EG(3, 1) et EG(5, 1).	15
3.1	Loi a posteriori de α obtenue sur un échantillon de taille $n = 50$ simulé selon la loi exponentielle généralisée EG(2, 2).	31
3.2	Loi a posteriori de λ obtenue sur un échantillon de taille $n = 50$ simulé selon la loi exponentielle généralisée EG(2, 2).	32
4.1	Loi a posteriori de α obtenue par MCMC Gibbs à partir de la première application.	34
4.2	Loi a posteriori de λ obtenue obtenue par MCMC Gibbs à partir de la première application.	35
4.3	Les fonctions densités de loi EG avec MV (discontinu) et Bayes (continu) pour la première application.	36

4.4	La fonction de répartition empirique et les fonctions de répartition estimées avec MV (discontinu) et Bayes (continu) en utilisant le modèle EG pour la première application.	37
4.5	Loi a posteriori de α obtenue par MCMC Gibbs à partir de la deuxième application.	37
4.6	Loi a posteriori de λ obtenue obtenue par MCMC Gibbs à partir de la deuxième application.	38
4.7	Les fonctions densités de loi EG avec MV (discontinu) et Bayes (continu) pour la deuxième application.	38
4.8	La fonction de répartition empirique et les fonctions de répartition estimées avec MV (discontinu) et Bayes (continu) en utilisant le modèle EG la deuxième application.	39

Liste des tableaux

1.1	Comportement de la fonction hasard des modèles exponentiel généralisé, Gamma et Weibull.	16
3.1	Moyenne des estimateurs et les erreurs quadratiques moyennes obtenues par la méthode du maximum de vraisemblance (MV) et l'approche bayésienne (MCMC-Gibbs).	32
4.1	La distance de Kolmogorov–Smirnov (K-S) et la valeur- p dans le cas de la première application.	34
4.2	La distance de Kolmogorov–Smirnov (K-S) et la valeur- p dans le cas de la deuxième application.	36

Introduction générale

Les modèles Gamma et de Weibull à deux paramètres sont largement étudiés dans la littérature, voir par exemple, [Alexander \[1962\]](#), [Jackson \[1963\]](#), [Klinken \[1961\]](#), [Masuyama and Kuroiwa \[1952\]](#) et [Johnson et al. \[1994\]](#) pour le modèle Gamma et [Weibull \[1939\]](#), [Plait \[1962\]](#), [Johnson \[1968\]](#) et [Gorski \[1968\]](#) pour le modèle de Weibull. Les deux paramètres des modèles Gamma et de Weibull représentent les paramètres de forme et d'échelle qui jouent un rôle très important pour analyser différents types de données strictement positives. Par conséquent ces deux modèles (Gamma et Weibull) trouvent leurs applications dans divers domaines, comme par exemple, l'économie, l'industrie mécanique, la médecine, la biologie, la physique, etc.

Cependant les modèles Gamma et de Weibull possèdent des défauts majeurs. Pour le modèle Gamma, il est connu que la fonction de répartition (ou encore la fonction de survie) n'admet pas de forme explicite lorsque le paramètre de forme n'est pas un entier positif. Donc les méthodes numériques (calcul d'intégrales) sont nécessaires pour approximer la fonction de répartition, la fonction de survie ou la fonction hasard. Dans le cas du modèle de Weibull, l'inconvénient majeur est que la convergence en loi des estimateurs du maximum de vraisemblance vers la loi normale est très lente (voir [Bain \[1976\]](#)). Par conséquent, les estimateurs par intervalle de confiance sont en général mauvais. Pour ces raisons, d'autres modèles de survie ont été proposés récemment. Dans ce travail nous sommes intéressé au modèle exponentiel généralisé à deux paramètres proposé par [Gupta and Kundu \[2001b\]](#) et étudié par [Gupta and Kundu \[2001a\]](#) et [Kundu and Gupta \[2008\]](#) ;

voir aussi [Raqab and Ahsanullah \[2001\]](#) et [Raqab \[2002\]](#).

Le premier objectif de ce travail est de faire le point sur les méthodes d'estimation des deux paramètres du modèle exponentiel généralisé. Nous présentons la méthode du maximum de vraisemblance développée par [Gupta and Kundu \[2001a\]](#), la méthode des moments proposée par [Gupta and Kundu \[2001a\]](#) et l'approche bayésienne développée récemment par [Kundu and Gupta \[2008\]](#). Notons que ces approches ont été développées dans le cas des données complètes.

Le second objectif de ce travail est d'analyser des données simulées et réelles par le modèle exponentiel généralisé en estimant les paramètres de ce modèle par la méthode du maximum de vraisemblance et l'approche bayésienne.

Nous avons structuré ce mémoire en quatre parties principales. Le premier chapitre est consacré à un rappel détaillé sur le modèle exponentiel généralisé à deux paramètres.

Dans le second chapitre, on présente les différentes procédures pour estimer les deux paramètres du modèle exponentiel généralisé en décrivant leurs principes.

Dans le chapitre trois et quatre, nous présentons les résultats obtenus à partir des données simulées et réelles.

Ce mémoire se termine par une conclusion générale.

Chapitre 1

Modèle exponentiel généralisé

Le modèle exponentiel généralisé à deux paramètres α et λ a été proposé et étudié par [Gupta and Kundu \[2001b\]](#) (voir aussi [Gupta and Kundu \[1999\]](#) et [Gupta and Kundu \[2001a\]](#)).

1.1 Définitions et propriétés

Dans cette partie, nous présentons les définitions et quelques propriétés du modèle exponentiel généralisé, ainsi que quelques interprétations probabilistes de ce modèle.

1.1.1 Fonction de répartition

On dit que F est une fonction de répartition de la variable aléatoire X qui suit une loi exponentielle généralisée de paramètres $\alpha, \lambda \in \mathbb{R}_+^*$, notée $EG(\alpha, \lambda)$, si F est donnée par :

$$F(x, \alpha, \lambda) = (1 - e^{-\lambda x})^\alpha; \quad \alpha, \lambda, x > 0$$

avec α est le paramètre de forme et λ est le paramètre d'échelle. On remarque que si $\alpha = 1$, la distribution exponentielle généralisée coïncide avec la loi exponentielle de paramètre λ . La Figure ci-dessous représente les courbes de fonction de répartition de loi exponentielle

généralisée avec $\lambda = 1$ et différents valeurs de $\alpha \in \{1, 3, 5\}$.

FIGURE 1.1: *Fonction de répartition de loi exponentielle généralisée EG(1, 1), EG(3, 1) et EG(5, 1).*

1.1.2 Fonction de densité

Une variable aléatoire continue X suit une loi de exponentielle généralisée $EG(\alpha, \lambda)$, si elle admet pour densité de probabilité la fonction :

$$f(x, \alpha, \lambda) = \frac{dF(x)}{dx} = \alpha\lambda(1 - e^{-\lambda x})^{\alpha-1}e^{-\lambda x}; \quad \alpha, \lambda, x > 0.$$

Proposition 1. *La densité de la distribution exponentielle généralisée est log-convexe si $\alpha < 1$ et log-concave si $\alpha > 1$.*

Preuve 1. *Il suffit de calculer la deuxième dérivée du logarithme de la fonction de densité.*

On a :

$$\frac{\partial^2}{\partial x^2} \log f(x, \alpha, \lambda) = -(\alpha - 1) \frac{e^{-\lambda x}}{(1 - e^{-\lambda x})^2} \quad (1.1)$$

donc $\frac{\partial^2}{\partial x^2} \log f(x, \alpha, \lambda)$ est log-convexe pour $\alpha < 1$ et log-concave pour $\alpha > 1$. □

La Figure ci-dessous représente les courbes des densités de probabilités de loi exponentielle généralisée $EG(1, 1)$, $EG(3, 1)$ et $EG(5, 1)$.

FIGURE 1.2: *Densité de probabilité de loi exponentielle généralisée EG(1, 1), EG(3, 1) et EG(5, 1).*

Nous donnons quelques cas particuliers et quelques propriétés :

1. Si $\alpha = 1$, on a $EG(1, \lambda)$ est la loi exponentielle de paramètre $\lambda > 0$ de densité de probabilité donnée par :

$$f(x, \lambda) = \lambda e^{-\lambda x}; \quad \lambda, x > 0.$$

2. Si $\lambda = 1$, on a $\text{EG}(\alpha, 1) = \text{EG}(\alpha)$ de densité de probabilité donnée par :

$$f(x, \alpha) = \alpha(1 - e^{-x})^{\alpha-1}e^{-x}; \quad \alpha, x > 0.$$

3. Si $X \sim \text{EG}(\alpha)$, alors $\lambda X \sim \text{EG}(\alpha, \lambda)$.

4. Si $X \sim \text{EG}(\alpha)$, alors $Y = e^{-X} \sim \text{Beta}(1, \alpha)$.

5. Si $X_i \sim \text{EG}(\alpha_i)$ pour $i = 1, \dots, n$, alors la densité de probabilité de la variable aléatoire $X = \sum_{i=1}^n X_i$ est donnée par :

$$f_X(x) = \sum_{j=1}^{\infty} C_j f(x, \alpha^* + j)$$

où $f(x, \alpha^* + j)$ est la densité de la loi $\text{EG}(\alpha^* + j)$, $\alpha^* = \sum_{i=1}^n \alpha_i$, $C_j = \frac{C_0 \alpha^*}{(\alpha^* + j)} C_j^{(n)}$, $j = 1, \dots$, avec $C_0 = \frac{\prod_{i=1}^n \Gamma(\alpha_i + 1)}{\Gamma(\alpha^* + 1)}$, $C_j^{(k)} = \frac{(\alpha_1 + \dots + \alpha_{k-1})_j}{(\alpha_1 + \dots + \alpha_k)_j} \sum_{i=0}^j \frac{(\alpha_k)_i}{i!} C_{j-i}^{(k-1)}$ pour $k = 3, \dots, n$ et $C_j^{(2)} = \frac{(\alpha_1)_j (\alpha_2)_j}{j! (\alpha_1 + \alpha_2)_j}$. Ici $(\alpha)_j = \Gamma(\alpha + j) / \Gamma(\alpha)$. Puisque $C_j > 0$ et $\sum_{j=1}^{\infty} C_j = 1$, alors la loi de $X = \sum_{i=1}^n X_i$ est un mélange de lois EG.

1.1.3 Fonction de survie

Soit X une variable aléatoire continué de loi exponentielle généralisée de fonction de répartition F et de densité de probabilité f . Sa fonction de survie qui est la probabilité de survivre au-delà de x est définie par :

$$S(x, \alpha, \lambda) = \Pr(X > x) = 1 - F(x) = 1 - (1 - e^{-\lambda x})^\alpha; \quad \alpha, \lambda, x > 0.$$

La fonction de survie est une fonction décroissante telle que $S(0) = 1$ et $\lim_{x \rightarrow \infty} S(x) = 0$. La Figure ci-dessous représente les courbes des fonctions de survie de loi exponentielle généralisée avec $\lambda = 1$ et différentes valeurs de α .

FIGURE 1.3: *Fonction de survie de loi exponentielle généralisée EG(1, 1), EG(3, 1) et EG(5, 1).*

1.1.4 Fonction hasard

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). En utilisant le théorème des probabilités conditionnelles et le modèle exponentielle généralisée $EG(\alpha, \lambda)$, la fonction hasard h est donnée par :

$$h(x) = \frac{f(x)}{S(x)} \quad (1.2)$$

$$= \frac{\alpha \lambda (1 - e^{-\lambda x})^{\alpha-1} e^{-\lambda x}}{1 - (1 - e^{-\lambda x})^\alpha}. \quad (1.3)$$

La Figure ci-dessous représente les courbes des fonctions de hasard de loi exponentielle généralisée avec $\lambda = 1$ et différentes valeurs de α .

FIGURE 1.4: *Fonction de hasard de loi exponentielle généralisée EG(1, 1), EG(3, 1) et EG(5, 1).*

Pour la distribution exponentielle généralisée $EG(\alpha, \lambda)$, la fonction hasard qui joue un rôle important est croissante si $\alpha < 1$, décroissante si $\alpha > 1$ et constante si $\alpha = 1$. Il est intéressant aussi d'étudier les similitudes de la densité et les fonctions de distributions de la famille exponentielle généralisée avec les familles Gamma et Weibull. Par exemple, l'étude du comportement de la fonction hasard des trois distributions se résume dans le tableau 1.1.

TABLE 1.1: *Comportement de la fonction hasard des modèles exponentiel généralisé, Gamma et Weibull.*

paramètre	EG	Gamma	Weibull
$\alpha = 1$	λ	λ	λ
$\alpha > 1$	croît de 0 à λ	croît de 0 à λ	croît de 0 à ∞
$\alpha < 1$	décroît de ∞ à λ	décroît de ∞ à λ	décroît de ∞ à 0

avec la densité de la distribution Gamma :

$$f_G(x, \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad \alpha, \lambda, x > 0$$

et la densité de la distribution Weibull :

$$f_W(x, \alpha, \lambda) = \lambda \alpha (\alpha x)^{\alpha-1} e^{-\lambda x}, \quad \alpha, \lambda, x > 0$$

La fonction hasard de la distribution exponentielle généralisée se comporte comme la fonction hasard de la distribution Gamma, ce qui est tout à fait différente de la fonction hasard de la distribution Weibull.

1.1.5 Moments d'ordre k

Considérons maintenant les différents moments de la distribution exponentielle généralisé. Soit X une variable aléatoire qui suit la loi $EG(\alpha, \lambda)$. Le moment d'ordre k est donné par :

$$\begin{aligned} E(X^k) &= \int_0^\infty x^k f(x) dx \\ &= \alpha \lambda \int_0^\infty x^k (1 - e^{-\lambda x})^{\alpha-1} e^{-\lambda x} dx. \end{aligned}$$

En utilisant la représentation de la série de $(1 - e^{-\lambda x})^{\alpha-1}$ on a :

$$(1 - e^{-\lambda x})^{\alpha-1} = \sum_{i=0}^{\infty} (-1)^i C_{\alpha-1}^i e^{-i\lambda x}$$

comme $C_{\alpha-1}^i = \frac{(\alpha-1)\dots(\alpha-i)}{i!}$, alors on obtient

$$E(X^k) = \frac{\alpha\Gamma(k+1)}{\lambda^k} \sum_{i=0}^{\infty} (-1)^i C_{\alpha-1}^i \frac{1}{(i+1)^{k+1}}.$$

Donc $\forall k \geq 0$ la série est convergente implique que tous les moments d'ordre k existent.

– pour $k = 1$, on obtient l'espérance de X :

$$E(X) = \frac{\alpha}{\lambda} \sum_{i=0}^{\infty} (-1)^i C_{\alpha-1}^i \frac{1}{(i+1)^2}.$$

– pour $k = 2$ on obtient le moment d'ordre 2 de X :

$$E(X^2) = \frac{2\alpha}{\lambda^2} \sum_{i=0}^{\infty} (-1)^i C_{\alpha-1}^i \frac{1}{(i+1)^3}.$$

1.1.6 Fonction génératrice des moments

On appelle fonction génératrice des moments de X qui suit la loi $EG(\alpha, \lambda)$, la fonction $M(t)$ définie par :

$$\begin{aligned} M(t) &= E(e^{tX}) \\ &= \int_0^{\infty} e^{tX} f(x) dx \\ &= \alpha\lambda \int_0^{\infty} (1 - e^{-\lambda x})^{\alpha-1} e^{(t-\lambda)x} dx. \quad 0 < t < \lambda \end{aligned}$$

On pose $y = e^{-\lambda x}$ on obtient :

$$\begin{aligned} M(t) &= \alpha \int_0^1 (1-y)^{\alpha-1} y^{i\frac{t}{\lambda}} dy \\ &= \frac{\Gamma(\alpha+1)\Gamma(1-\frac{t}{\lambda})}{\Gamma(\alpha+1-\frac{t}{\lambda})} \end{aligned}$$

On sait que :

$$E[X^k] = \frac{\partial^k M(0)}{\partial t^k},$$

donc, pour calculer l'espérance de X , il suffit de calculer la première dérivée de $M(t)$ au point $t = 0$. On pose $t = 0$, alors

$$\begin{aligned} E(X) &= \frac{\partial M(0)}{\partial t} \\ &= \frac{\Gamma'(\alpha + 1)\Gamma(1) - \Gamma'(1)\Gamma(\alpha + 1)}{\lambda\Gamma(1)\Gamma(\alpha + 1)} \\ &= \frac{1}{\lambda}(\psi(\alpha + 1) - \psi(1)) \end{aligned}$$

avec $\psi(a) = \frac{\Gamma'(a)}{\Gamma(a)}$ est la fonction digamma (voir [Abramowitz and Stegun \[1972\]](#) et [Amos \[1983\]](#)).

1.1.7 Simulation d'une variable aléatoire de loi $EG(\alpha, \lambda)$

Dans cette section, nous présentons une méthode de simulation suivant la loi $EG(\alpha, \lambda)$. Rappelons que dans le cas de la loi $EG(\alpha, \lambda)$, la fonction de répartition $F(x, \alpha, \lambda) = (1 - e^{-\lambda x})^\alpha$; $\alpha, \lambda, x > 0$. Sa fonction inverse F^{-1} possède une forme simple analytique. Donc, on peut employer la méthode de simulation inverse (méthode populaire et simple) pour simuler X , en utilisant la relation suivante :

$$X = F^{-1}(U) = -\frac{1}{\lambda} \log \left(1 - U^{\frac{1}{\alpha}} \right) \quad (1.4)$$

avec U est une variable aléatoire uniforme sur $[0, 1]$. Le programme R qui permet de simuler X de loi $EG(\alpha, \lambda)$ est donné dans l'annexe.

1.2 Conclusion

Dans ce chapitre, nous avons donné un rappel détaillé sur le modèle exponentiel généralisé de deux paramètres α et λ . Dans un premier temps, les fonctions de répartition, de densité de probabilité, de survie et de hasard sont présentées. Dans un second temps, les calculs des moments d'ordre k et de la fonction génératrice des moments sont donnés.

Ensuite, la méthode inverse pour simuler une variable aléatoire X de loi $EG(\alpha, \lambda)$ est présenté. Dans le chapitre suivant, nous présentons les différentes méthodes d'estimation des paramètres de la distribution $EG(\alpha, \lambda)$.

Chapitre 2

Estimation dans le modèle exponentiel généralisé

Dans ce chapitre, nous présentons la méthode du maximum de vraisemblance (MV), la méthode des moments et l'approche bayésienne pour estimer les deux paramètres α et λ du modèle exponentiel généralisé $EG(\alpha, \lambda)$. Ces méthodes ont été développées respectivement par [Gupta and Kundu \[2001a\]](#), [Gupta and Kundu \[2001a\]](#) et [Kundu and Gupta \[2008\]](#).

2.1 Méthode du maximum de vraisemblance

Dans cette section, on discute les estimateurs du maximum de vraisemblance des deux paramètres α et λ de la distribution exponentielle généralisée.

2.1.1 Estimateurs du maximum de vraisemblance

Soit x_1, \dots, x_n un n observations d'un n -échantillon issu d'une variable X aléatoire de la loi $EG(\alpha, \lambda)$. Alors la fonction de vraisemblance notée par $L(x, \alpha, \lambda)$ est donnée par :

$$\begin{aligned}
L(x, \alpha, \lambda) &= \prod_{i=1}^n f(x_i, \alpha, \lambda) \\
&= \prod_{i=1}^n \alpha \lambda (1 - e^{-\lambda x_i})^{\alpha-1} e^{-\lambda x_i}.
\end{aligned} \tag{2.1}$$

La méthode de maximum de vraisemblance consiste à choisir comme estimateur $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ qui réalise un maximum de la vraisemblance, c'est-à-dire :

$$L(x, \hat{\theta}) \geq L(x, \theta), \quad \forall \theta \in \Theta.$$

La recherche du MV peut se faire en résolvant l'équation :

$$\frac{\partial L}{\partial \theta} = 0$$

qui vérifie $\frac{\partial^2 L}{\partial \theta^2} < 0$. Cependant, dans la pratique, on préfère remplacer ce problème (produit) par le problème équivalent pour la log-vraisemblance :

$$\frac{\partial \log L}{\partial \theta} = 0 \tag{2.2}$$

avec $\frac{\partial^2 \log L}{\partial \theta^2} < 0$. En utilisant l'expression (2.1) et l'équation (2.2), on obtient les équations non-linéaires suivantes

$$\begin{cases} \frac{\partial \log L(x, \alpha, \lambda)}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \log(1 - e^{-\lambda x_i}) = 0, \\ \frac{\partial \log L(x, \alpha, \lambda)}{\partial \lambda} = \frac{n}{\lambda} + (\alpha - 1) \sum_{i=1}^n \frac{x_i e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} - \sum_{i=1}^n x_i = 0. \end{cases}$$

Notons que la résolution de ce système est relativement difficile et n'admet pas des solutions explicites. Dans ce travail, nous utilisons la fonction `mle` du logiciel statistique R. En introduisant les valeurs initiales α^0 et λ^0 , la fonction `mle` permet de retourner les estimateurs de α et λ , en calculant le maximum de la log-vraisemblance de manière itérative.

Les étapes principales sont les suivantes (voir aussi l'annexe pour le programme R) :

1. Fixer les paramètres α et λ de la loi
2. Générer un n -échantillon de X à partir de la loi $\text{EG}(\alpha, \lambda)$;
3. Donner des valeurs initiales $(\alpha^{(0)}, \lambda^{(0)})$;
4. Calculer les estimateurs MV des paramètres de la loi $\text{EG}(\alpha, \lambda)$.

2.1.2 Estimateurs par intervalle de confiance

Les estimateurs par intervalle de confiance peuvent être obtenus, en utilisant le résultat de la normalité asymptotique suivant :

$$\sqrt{n} [(\hat{\alpha} - \alpha), (\hat{\lambda} - \lambda)] \longrightarrow \mathcal{N}_2(\mathbf{0}, \mathbf{I}^{-1}(\alpha, \lambda)) \quad (2.3)$$

avec $\hat{\alpha}$ et $\hat{\lambda}$ sont les estimateurs MV de α et λ et $\mathbf{I}(\alpha, \lambda)$ est la matrice de l'information du Fisher donnée par :

$$\mathbf{I}(\alpha, \lambda) = \begin{pmatrix} E\left(\frac{\partial^2 \log L}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 \log L}{\partial \alpha \partial \lambda}\right) \\ E\left(\frac{\partial^2 \log L}{\partial \lambda \partial \alpha}\right) & E\left(\frac{\partial^2 \log L}{\partial \lambda^2}\right) \end{pmatrix}$$

Les éléments de la matrice de l'information du Fisher sont données comme suit, pour $\alpha > 2$

$$E\left(\frac{\partial^2 \log L}{\partial \alpha^2}\right) = -\frac{n}{\alpha^2},$$

$$E\left(\frac{\partial^2 \log L}{\partial \alpha \partial \lambda}\right) = \frac{n}{\lambda} \left[\frac{\alpha}{\alpha - 1} (\psi(\alpha) - \psi(1)) - (\psi(\alpha + 1) - \psi(1)) \right]$$

$$E\left(\frac{\partial^2 \log L}{\partial \lambda^2}\right) = -\frac{n}{\lambda^2} \left[1 + \frac{\alpha(\alpha - 1)}{\alpha - 2} (\psi'(1) - \psi'(\alpha - 1)) + (\psi(\alpha - 1) - \psi(1))^2 \right]$$

$$- \frac{n\alpha}{\lambda^2} [(\psi'(1) - \psi(\alpha)) + (\psi(\alpha) - \psi(1))^2]$$

et pour $0 < \alpha \leq 2$

$$E\left(\frac{\partial^2 \log L}{\partial \alpha^2}\right) = -\frac{n}{\alpha^2},$$

$$E\left(\frac{\partial^2 \log L}{\partial \alpha \partial \lambda}\right) = \frac{n\alpha}{\lambda} \int_0^\infty x e^{-2x} (1 - e^{-x})^{\alpha-2} dx$$

$$E\left(\frac{\partial^2 \log L}{\partial \lambda^2}\right) = -\frac{n}{\lambda^2} - \frac{n\alpha(\alpha-1)}{\lambda^2} \int_0^\infty x^2 e^{-2x} (1 - e^{-x})^{\alpha-2} dx$$

Maintenant, pour obtenir les estimateurs par intervalle de confiance, il suffit de remplacer $\mathbf{I}^{-1}(\alpha, \lambda)$ par son estimateur naturel $\mathbf{I}^{-1}(\hat{\alpha}, \hat{\lambda})$.

2.2 Méthode des moments

Dans cette section, nous présentons la méthode des moments (MM) pour estimer les paramètres α et λ de la distribution exponentielle généralisée EG. Si $X \sim \text{EG}(\alpha, \lambda)$, alors on a :

$$\mu = E(X) = \frac{1}{\lambda} (\psi(\alpha + 1) - \psi(1)), \quad (2.4)$$

et

$$\sigma^2 = \text{Var}(X) = -\frac{1}{\lambda^2} (\psi'(\alpha + 1) - \psi'(1)). \quad (2.5)$$

voir le chapitre précédent pour plus de détail. Rappelons que ψ et ψ' désignent respectivement la fonction digamma et sa dérivée. A partir de (2.4) et (2.5), nous obtenons le coefficient de variation CV :

$$\text{CV} = \frac{\sigma}{\mu} = \frac{\sqrt{(\psi'(1) - \psi'(\alpha + 1))}}{(\psi(\alpha + 1) - \psi(1))}. \quad (2.6)$$

Le coefficient de variation CV ne dépend pas de λ . En utilisant une réalisation d'un échantillon de taille n , x_1, x_2, \dots, x_n , alors l'estimateur $\widehat{\text{CV}}$ de CV est donné par :

$$\widehat{\text{CV}} = \frac{s}{\bar{x}} = \frac{\sqrt{(\psi'(1) - \psi'(\alpha + 1))}}{(\psi(\alpha + 1) - \psi(1))}. \quad (2.7)$$

où $\bar{x} = (1/n) \sum_{i=1}^n x_i$ et $s^2 = (1/(n-1)) \sum_{i=1}^n (x_i - \bar{x})^2$ sont les estimateurs empiriques de μ et σ^2 . Maintenant, on peut utiliser (2.7) pour calculer un estimateur $\widehat{\alpha}_{MM}$ de α par la méthode des moments. Cependant, l'équation (2.7) n'admet pas de solution exacte, donc une procédure numérique (méthode itérative) est nécessaire pour approximer cette solution. En estimant α par $\widehat{\alpha}_{MM}$, on peut déduire un estimateur $\widehat{\lambda}_{MM}$ de λ par la méthode des moments, en utilisant la relation suivante :

$$\widehat{\lambda}_{MM} = \frac{(\psi(\widehat{\alpha}_{MM} + 1) - \psi(1))}{\bar{x}}. \quad (2.8)$$

Notons que la relation (2.8) est obtenue à partir de l'équation (2.4).

2.3 Approche bayésienne

2.3.1 Principe bayésien

Le but de ce chapitre est de rappeler quelques notions sur l'analyse bayésienne. Supposant que $\mathbf{x} = (x_1, \dots, x_n)$ est le vecteur d'observations et $\theta = (\alpha, \lambda) \in \Theta$ est le vecteur des paramètres à estimer. L'idée principale de l'analyse bayésienne repose sur la loi a posteriori des paramètres en considérant θ comme aléatoire. L'espace des paramètres Θ est muni d'une loi de probabilité π et nous noterons $\theta \sim \pi$. La loi π est appelée loi a priori de θ choisie en fonction des connaissances disponibles sur θ avant la prise en compte des observations. La loi a posteriori $\pi(\theta|\mathbf{x})$ du modèle est obtenue en utilisant la règle de Bayes

$$\pi(\theta|\mathbf{x}) = \frac{\pi(\mathbf{x}|\theta)\pi(\theta)}{\pi(\mathbf{x})}, \quad (2.9)$$

où $\pi(\mathbf{x}|\theta)$ est la fonction vraisemblance et $\pi(\mathbf{x}) = \int_{\Theta} \pi(\mathbf{x}|\theta)\pi(\theta)d\theta$ est la constante de normalisation (ne dépend pas de θ). En utilisant la moyenne a posteriori, l'estimateur

bayésien est donné par :

$$\hat{\theta}_{bayes} = E(\theta|\mathbf{x}) = \int_{\Theta} \theta \pi(\theta|\mathbf{x}) d\theta. \quad (2.10)$$

Notons que l'estimateur bayésien (2.10) est obtenu en utilisant la perte quadratique.

2.3.2 Estimateurs de Bayes

Dans cette section, nous considérons l'estimation bayésienne des paramètres inconnus α et λ . Les lois a priori choisies pour les paramètres α et λ sont définies et les loi a posteriori de ces deux paramètres sont évaluées.

2.3.2.1 Choix des lois a priori

Les lois a priori sur les paramètres inconnus α et λ sont données comme suit (voir [Kundu and Gupta \[2008\]](#)) :

$$\pi(\alpha) = \frac{b^a}{\Gamma(a)} \alpha^{a-1} e^{-b\alpha}; \quad \alpha > 0, \quad (2.11)$$

et

$$\pi(\lambda) = \frac{d^c}{\Gamma(c)} \lambda^{c-1} e^{-c\lambda}; \quad \lambda > 0. \quad (2.12)$$

avec a, b, c et d sont des hyper-paramètres positifs et connus. Notons que lorsque $a = b = c = d = 0$, les lois a priori de α et λ sont des lois a priori de Jeffreys.

2.3.2.2 Calcul des lois a posteriori

Soit $\mathbf{x} = (x_1, \dots, x_n)$ n observations d'une variable aléatoire X de la loi $EG(\alpha, \lambda)$. Après la spécification des lois a priori, par le théorème de Bayes, la loi a posteriori de

(α, λ) sachant les données \mathbf{x} est donnée par :

$$\begin{aligned} \pi(\alpha, \lambda | \mathbf{x}) &= \frac{L(\mathbf{x}, \alpha, \lambda) \pi(\alpha) \pi(\lambda)}{\int_0^\infty \int_0^\infty L(\mathbf{x}, \alpha, \lambda) \pi(\alpha) \pi(\lambda) d\alpha d\lambda} \\ &= \frac{\alpha^{n+a-1} \lambda^{n+d-1} e^{-\lambda(c+\sum_{i=1}^n x_i)} e^{-b\alpha} \prod_{i=1}^n (1 - e^{-\lambda x_i})^{\alpha-1}}{\int_0^\infty \int_0^\infty \alpha^{n+a-1} \lambda^{n+d-1} e^{-\lambda(c+\sum_{i=1}^n x_i)} e^{-b\alpha} \prod_{i=1}^n (1 - e^{-\lambda x_i})^{\alpha-1} d\alpha d\lambda} \end{aligned} \quad (2.13)$$

Vu la complexité de la loi a posteriori (2.13) il est impossible d'obtenir directement les estimateurs bayésien par (2.10). Pour cette raison, la stratégie de calcul consiste d'abord à dériver les lois a posteriori conditionnelles de $\alpha | \lambda, \mathbf{x}$ et $\lambda | \alpha, \mathbf{x}$ et d'utiliser les méthodes de Monté Carlo par Chaîne de Markov (MCMC) proposée par [Metropolis et al. \[1953\]](#) et [Hastings \[1970\]](#) pour estimer les paramètres α et λ .

A partir de (2.13), il clair que la loi a posteriori de α sachant λ et les données \mathbf{x} est donnée par :

$$\pi(\alpha | \lambda, \mathbf{x}) \propto \alpha^{n+a-1} e^{-\alpha(b - \sum_{i=1}^n \log(1 - e^{-\lambda x_i}))}. \quad (2.14)$$

où \propto désigne proportionnel. Donc, la loi a posteriori de $\alpha | \lambda, \mathbf{x}$ est une loi Gamma de paramètres $n + a$ et $b - \sum_{i=1}^n \log(1 - e^{-\lambda x_i})$. Nous allons maintenant dériver la densité a posteriori de $\lambda | \alpha, \mathbf{x}$. La loi a posteriori de $\lambda | \alpha, \mathbf{x}$ s'écrit comme suit :

$$\pi(\lambda | \alpha, \mathbf{x}) \propto \alpha^{n+a-1} \lambda^{n+d-1} e^{-\lambda(c+\sum_{i=1}^n x_i)} \prod_{i=1}^n (1 - e^{-\lambda x_i})^{\alpha-1}. \quad (2.15)$$

Notons que la loi de $\alpha | \lambda, \mathbf{x}$ est simulable facilement ($\alpha | \lambda, \mathbf{x} \sim \text{Gamma}$). Cependant, la loi de $\lambda | \alpha, \mathbf{x}$ n'est pas simulable directement. Pour cette raison, nous proposons d'utiliser les méthodes MCMC avec Gibbs pour construire une chaîne de Markov pour le paramètre α et λ , en utilisant des valeurs initiales $\alpha^{(1)}$ et $\lambda^{(1)}$. Pour simuler suivant la loi de $\lambda | \alpha, \mathbf{x}$, nous utilisons l'algorithme de Metropolis-Hasting (M-H) à marche aléatoire. Cet algorithme est basé sur l'utilisation d'une loi génératrice de candidats de la forme $q(\tilde{\lambda} | \lambda^{(i)}) = q(\tilde{\lambda} - \lambda^{(i)})$. Le candidat $\tilde{\lambda}$ est généré par le processus $|\tilde{\lambda} = \lambda^{(i)} + \tau^2 \epsilon|$ où τ^2 est un paramètre choisi de telle sorte à obtenir un taux d'acceptation qui est proche de 0.5 (voir [Gelman et al. \[1996\]](#)) et ϵ est une marche aléatoire choisie comme gaussienne

de moyenne 0 et de variance 1, c'est à dire $\epsilon \sim \mathcal{N}(0, 1)$. Une étape de Gibbs est utilisée pour simuler la loi de $\alpha|\lambda, \mathbf{x}$. Après une période de chauffage N_0 (burn-in period) et un nombre total d'itérations N très large, les chaînes de Markov construites convergent vers les distributions d'intérêts (lois a posteriori conditionnelles). Les itérations N_0 ne sont pas utilisées dans le calcul des estimateurs de α et λ . Les étapes de l'algorithme qui permet de simuler suivant les lois de $\alpha|\lambda, \mathbf{x}$ et $\lambda|\alpha, \mathbf{x}$ sont résumées comme suit :

Step 1- initialiser $\alpha^{(1)}$ et $\lambda^{(1)}$

Step 2- pour $t \in \{2, \dots, N\}$

(a) Générer $\tilde{\lambda} \sim \mathcal{N}(\lambda^{(t-1)}, \tau^2)$; $\tilde{\lambda} = |\tilde{\lambda}|$

(b) Calculer la probabilité d'acceptation $\rho = \min\left\{1, \frac{\pi(\tilde{\lambda}|\alpha^{(t-1)}, \mathbf{x})}{\pi(\lambda^{(t-1)}|\alpha^{(t-1)}, \mathbf{x})}\right\}$

$$\lambda^{(t)} = \begin{cases} \tilde{\lambda}, & \text{si } u < \rho, \quad u \sim U_{[0,1]} \\ \lambda^{(t-1)}, & \text{sinon} \end{cases}$$

(c) Générer $\alpha^{(t)}$ selon Gamma $\left(n + a, b - \sum_{i=1}^n \log(1 - e^{-\lambda^{(t)} x_i})\right)$

Step 3- $t = t + 1$ et aller à 2

Step 4- Calculer les estimateurs de Bayes $\hat{\alpha}_{Bayes}$ et $\hat{\lambda}_{Bayes}$

$$\hat{\alpha}_{Bayes} = \frac{1}{N - N_0} \sum_{t=N_0+1}^N \alpha^{(t)} \quad \text{et} \quad \hat{\lambda}_{Bayes} = \frac{1}{N - N_0} \sum_{t=N_0+1}^N \lambda^{(t)}.$$

Le programme R qui permet de retourner les estimateurs de Bayes est présenté dans l'annexe.

Remarque 1. Dans *Kundu and Gupta [2008]*, les auteurs ont utilisés aussi l'approche de *Lindley [1980]* pour trouver les estimateurs de Bayes. L'idée de base de l'approche de *Lindley* consiste à approximer un rapport de deux intégrales.

2.4 Conclusion

Dans ce chapitre, nous avons exposé les différentes méthodes d'estimation des paramètres α et λ du modèle exponentiel généralisé, à savoir la méthode du maximum de

vraisemblance, la méthode des moments et l'approche bayésienne. Le principe de la méthode du maximum de vraisemblance consiste à maximiser la fonction vraisemblance ou encore la fonction log-vraisemblance. La méthode des moments consiste à faire égaliser les moments théoriques et empiriques. Dans le cas de l'approche bayésienne, le calcul des estimateurs se base sur la loi a posteriori. Dans les deux chapitres suivants, nous analysons des données simulées et réelles par le modèle $EG(\alpha, \lambda)$ en estimant les paramètres α et λ par la méthode du maximum de vraisemblance et l'approche bayésienne.

Chapitre 3

Application sur des données simulées

Dans cette section, nous présentons les principaux résultats obtenus sur des simulations. Cette partie permet de tester l'efficacité de la méthodes du maximum de vraisemblance et de l'approche bayésienne dans l'estimation du modèle $EG(\alpha, \lambda)$ sur des données simulées suivant ce modèle. Elle permet aussi la comparaison des performances des deux approches. Les résultats ont été réalisées à l'aide du logiciel R un système d'analyse statistique et graphique (voir). Pour les simulations, nous procédons en deux étapes :

- Dans la première étape, nous utilisons le modèle $EG(\alpha, \lambda)$ pour simuler les données \mathbf{x} en utilisant la relation (1.4) et en fixant les paramètres α et λ .
- La deuxième étape est consacrée à l'estimation des paramètres. Nous utilisons les observations \mathbf{x} simulées pour estimer les paramètres avec la méthode du maximum de vraisemblance et l'approche bayésienne.

3.1 Simulation des données

Nous simulons, selon le modèle $EG(\alpha, \lambda)$, des observations $\mathbf{x} = (x_1, x_2, \dots, x_n)$ de taille $n = \{10, 30, 50, 80, 100, 200\}$, en fixant les paramètres de ce modèle. Les paramètres fixes

ont été choisis comme suit :

$$\alpha = 2 \quad \text{et} \quad \lambda = 2.$$

3.2 Estimation des paramètres

En utilisant nos programme implémentés sous R (voir annexe), nous avons estimés les paramètres de modèle en utilisant les observations simulées dans la première étape.

- Pour la méthode du maximum de vraisemblance, nous utilisons la méthode `mle` sous R en fixant les valeurs initiales de α et λ .
- Pour l’approche bayésienne, nous utilisons les lois a priori (2.11) et (2.12) en fixant les hyper-paramètres $a = b = c = d = 1$. En suite, nous appliquons le programme MCMC avec une étape de Gibbs implémenté sous R en fixant les valeurs initiales $\alpha^{(1)}$ et $\lambda^{(1)}$. Notons que le nombre total d’itérations est $N = 1000$ et la période de chauffage est $N_0 = 300$. Donc, nous utilisons que $N - N_0 = 700$ pour estimer les paramètres.

3.3 Résultats

Les résultats sont donnés sous forme de tableau et de graphique. Le tableau 3.1 rassemble la taille de l’échantillon, la méthode utilisée, la moyenne (Moy) pour chaque paramètre et les erreurs quadratiques moyennes (EQM). La moyenne est donnée comme suit :

$$\bar{\hat{\alpha}} = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \hat{\alpha}_i \quad \text{et} \quad \bar{\hat{\lambda}} = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \hat{\lambda}_i.$$

Les erreurs quadratiques moyennes sont données par :

$$EQM(\alpha) = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} (\hat{\alpha}_i - \alpha)^2 \quad \text{et} \quad EQM(\lambda) = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} (\hat{\lambda}_i - \lambda)^2.$$

où α et λ sont les valeurs fixées, $\hat{\alpha}_i$ et $\hat{\lambda}_i$ sont les estimateurs par maximum de vraisem-

blance ou Bayes obtenus à partir de l'échantillon numéro i et $N_{sim} = 500$ est le nombre de répétition de simulation. Les figures 3.1 et 3.2 montrent les histogrammes des lois a posteriori de α et λ obtenues en utilisant le programme MCMC avec Gibbs et un seul échantillon de taille $n = 50$ simulé selon la loi exponentielle généralisée EG(2, 2).

FIGURE 3.1: *Loi a posteriori de α obtenue sur un échantillon de taille $n = 50$ simulé selon la loi exponentielle généralisée EG(2, 2).*

FIGURE 3.2: *Loi a posteriori de λ obtenue sur un échantillon de taille $n = 50$ simulé selon la loi exponentielle généralisée EG(2, 2).*

3.4 Discussion

La partie simulation a pour objectif de tester et de comparer les performances de la méthode du maximum de vraisemblance et l'approche bayésienne dans l'estimation des paramètres α et λ du modèle EG. Nous avons choisi la moyenne a posteriori comme estimateur bayésien. Pour la méthode du maximum de vraisemblance, nous avons utilisé la fonction `mle` sous R pour trouver les estimateurs. Dans le cas de l'approche bayésienne, nous avons utilisé notre programme MCMC-Gibbs pour estimer les paramètres α et λ . Les résultats obtenus par les deux méthodes nous permet de constater que les estimations sont proches des valeurs attendues (valeurs des paramètres fixées $\alpha = 2$ et $\lambda = 2$). Les résultats montrent aussi que l'approche bayésienne est meilleure que la méthode du maximum de vraisemblance au sens de l'erreur quadratique moyenne (EQM) (voir le tableau 3.1).

n	Méthode	$\bar{\alpha}$	EQM(α)	$\bar{\lambda}$	EQM(λ)
$n = 15$	MV	2.3569	0.3111	2.2643	0.1594
	MCMC	1.9602	0.0552	1.9972	0.006
$n = 30$	MV	2.1865	0.3331	2.0955	0.2090
	MCMC	1.9989	0.0879	2.0013	0.001
$n = 50$	MV	2.0922	0.3515	2.0419	0.1942
	MCMC	1.9717	0.1177	2.0010	0.007
$n = 80$	MV	2.2020	0.3843	2.1556	0.2384
	MCMC	1.9717	0.1433	2.0002	0.001

$n = 100$	MV	2.2458	0.3840	2.1206	0.1614
	MCMC	2.0204	0.0990	1.9990	0.009
$n = 200$	MV	2.2128	0.3978	2.1034	0.1861
	MCMC	2.0052	0.0982	2.0007	0.001

TABLE 3.1: Moyenne des estimateurs et les erreurs quadratiques moyennes obtenues par la méthode du maximum de vraisemblance (MV) et l'approche bayésienne (MCMC-Gibbs).

Chapitre 4

Application aux données réelles

4.1 Applications réelles

Nous allons appliquer le modèle exponentiel généralisé EG sur deux jeux de données réels. Nous estimons les paramètres α et λ par la méthode du maximum de vraisemblance et l'approche bayésienne.

4.1.1 Première application

Nous considérons les données décrivant les durées de vie du système de conditionnement d'air d'un plane air (avion). Les données sont (voir [Linhart and Zucchini \[1986\]](#) et [Gupta and Kundu \[2001b\]](#)) :

23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95.

Nous analysons les données en utilisant le modèle exponentiel généralisé $EG(\alpha, \lambda)$. Nous appliquons la fonction `mle` pour estimer les paramètres α et λ . Les résultats obtenus sont $\hat{\alpha}_{MV} = 0.8104$ et $\hat{\lambda}_{MV} = 0.0145$.

Maintenant, nous calculons les estimateurs de Bayes (moyenne a posteriori) en utilisant le programme MCMC avec Gibbs en utilisant le nombre total des itérations $N = 5000$, la

période de chauffage $N_0 = 2500$ et les hyper-paramètres des lois a priori $a = b = c = d = 0$ (voir Kundu and Gupta [2008] pour le choix des hyper-paramètres). Les résultats obtenus (moyennes a posteriori) sont $\hat{\alpha}_{Bayes} = 0.8314$ et $\hat{\lambda}_{Bayes} = 0.0152$. Notons aussi que les lois a posteriori de α et λ sont données dans les figures 4.1 et 4.2. Nous traçons aussi la densité de loi EG en utilisant les paramètres estimés par maximum vraisemblance et Bayes (voir la figure 4.3). Pour vérifier la validité du modèle, nous calculons la distance de Kolmogorov–Smirnov (K-S) et la valeur- p entre la fonction de répartition empirique et les fonctions de répartition de $EG(\hat{\alpha}_{MV} = 0.8104, \hat{\lambda}_{MV} = 0.0145)$ et $EG(\hat{\alpha}_{Bayes} = 0.8314, \hat{\lambda}_{Bayes} = 0.0152)$. Les résultats obtenus sont donnés dans le tableau 4.1 (voir aussi la figure 4.4).

Méthode	Distance K-S	valeur- p
Maximum vraisemblance (mle)	0.1732	0.3291
Bayes (MCMC Gibbs)	0.1708	0.3450

TABLE 4.1: La distance de Kolmogorov–Smirnov (K-S) et la valeur- p dans le cas de la première application.

FIGURE 4.1: Loi a posteriori de α obtenue par MCMC Gibbs à partir de la première application.

FIGURE 4.2: Loi a posteriori de λ obtenue obtenue par MCMC Gibbs à partir de la première application.

FIGURE 4.3: Les fonctions densités de loi EG avec MV (discontinu) et Bayes (continu) pour la première application.

4.1.2 Deuxième application

La deuxième application porte sur les tests de l'endurance des roulements à billes rigides. Les données sont (voir Lawless [1982] et Kundu and Gupta [2008]) :

17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40.

FIGURE 4.4: *La fonction de répartition empirique et les fonctions de répartition estimées avec MV (discontinu) et Bayes (continu) en utilisant le modèle EG pour la première application.*

De la même manière, nous utilisons la fonction `mle` sous `R` et le programme MCMC avec Gibbs pour estimer les paramètres du modèle EG en utilisant ces données. Pour la méthode du maximum de vraisemblance avec la fonction `mle`, nous obtenons $\hat{\alpha}_{MV} = 5.2958$ et $\hat{\lambda}_{MV} = 0.0323$. Dans le cas de l'approche bayésienne, les résultats obtenus sont $\hat{\alpha}_{MV} = 5.4122$ et $\hat{\lambda}_{MV} = 0.0319$. Les lois a posteriori de α et λ sont données dans les figures 4.5 et 4.6. Nous traçons aussi les densités de probabilités avec les estimateurs de maximum de vraisemblance et Bayes (moyenne a posteriori) dans la figure 4.7. Afin de valider le modèle, nous calculons la distance de Kolmogorov–Smirnov (K-S) et la valeur- p comme dans le cas de la première application. Les résultats obtenus sont présentés dans le tableau 4.2 et dans la figure 4.8.

Méthode	Distance K-S	valeur- p
Maximum vraisemblance (mle)	0.1065	0.9566
Bayes (MCMC Gibbs)	0.1235	0.8742

TABLE 4.2: *La distance de Kolmogorov–Smirnov (K-S) et la valeur- p dans le cas de la deuxième application.*

FIGURE 4.5: *Loi a posteriori de α obtenue par MCMC Gibbs à partir de la deuxième application.*

FIGURE 4.6: *Loi a posteriori de λ obtenue par MCMC Gibbs à partir de la deuxième application.*

4.2 Discussion

Dans cette partie, nous avons analysé des données réelles en utilisant le modèle exponentiel généralisé $EG(\alpha, \lambda)$ en estimant les paramètres de ce modèle par la méthode du maximum de vraisemblance et l'approche bayésienne. Les résultats donnés sous forme de tableaux et graphiques montrent que le modèle EG est très approprié pour ces deux ensembles de données. Nous constatons aussi que les performances de la méthode du maximum de vraisemblance et l'approche bayésienne au sens de la distance K-S sont proches.

FIGURE 4.7: *Les fonctions densités de loi EG avec MV (discontinu) et Bayes (continu) pour la deuxième application.*

FIGURE 4.8: *La fonction de répartition empirique et les fonctions de répartition estimées avec MV (discontinu) et Bayes (continu) en utilisant le modèle EG la deuxième application.*

Notons aussi que les résultats obtenus dans ce travail sont très proches de ceux obtenus par [Gupta and Kundu \[2001b\]](#) et [Kundu and Gupta \[2008\]](#).

Chapitre 5

Conclusion

Ce travail a pour objet l'étude théorique et pratique du modèle exponentiel généralisé avec deux paramètres α et λ ($\text{EG}(\alpha, \lambda)$).

La première partie de ce travail a été consacrée au rappel sur le modèle exponentiel généralisé. Nous avons présentés la forme de la fonction de répartition, de la densité de probabilité, de la fonction de survie et hasard. la méthode de simulation de la variable aléatoire X de loi $\text{EG}(\alpha, \lambda)$ a été aussi présentée.

Dans la seconde partie, notre objectif était de présenter en détail les méthodes d'estimation des paramètres α et λ . Premièrement, nous avons rappelés le principe de la méthode du maximum de vraisemblance dans l'estimation de ce modèle. La fonction `mle` sous `R` a été utilisée pour estimer les paramètres α et λ car le problème n'admet pas de solution exacte. Dans un second temps, nous avons présentés l'approche bayésienne. Dans cette approche, nous avons développés un algorithme MCMC avec une étape de Gibbs pour estimer les paramètres α et λ car la loi a posteriori conjointe de $\alpha, \lambda | \mathbf{x}$ est complexe.

Ensuite, nous nous sommes intéressés aux applications sur des données simulées et

réelles. Le but de cette partie est d'appliquer le modèle exponentiel généralisé en estimant les paramètres α et λ par la méthode de maximum de vraisemblance et de l'approche bayésienne. Une comparaison au sens de l'erreur quadratique moyenne entre les deux méthodes (maximum vraisemblance et Bayes) a été réalisée.

Annexe

Code source avec R

Les programmes informatiques ont été implimentés sous R en utilisant les packages Stat4 et MASS. Dans un premier temps, nous présentons les programmes mis en place pour le calcul des fonctions de répartition, de densité, de survie et de hasard. Dans un second temps, nous présentons la méthode inverse pour simuler une variable aléatoire X de loi $EG(\alpha, \lambda)$. Ensuite, nous donnons les programmes R qui retournent les estimateurs de maximum de vraisemblance et les estimateurs de Bayes pour les paramètres α et λ .

```
#=====
##### calcul des fonctions (répartition, densité, survie, hasard)
#=====

FrEG=function(x,alpha,lambda){
fr=(1-exp(-lambda*x))^(alpha)
fr}

dEG=function(x,alpha,lambda){
den=alpha*lambda*((1-exp(-lambda*x))^(alpha-1))*exp(-lambda*x)
den}

sEG=function(x,alpha,lambda){
```

```
fs=1-(1-exp(-lambda*x))^(alpha)
fs}

hEG=function(x,alpha,lambda){
fh=dEG(x,alpha,lambda)/sEG(x,alpha,lambda)
fh}

#=====
##### simulation de X qui suit  $\{\text{rm EG}\}(\alpha, \lambda)$ 
#=====

simEG=function(alpha,lambda,n){
u=runif(n)
X=(-1/lambda)*log(1-u^(1/alpha))
X}

#=====
##### Estimation des paramètres alpha et lambda par mle et MCMC Gibbs
#=====

##### mle #####

vrais<-function (alpha,lambda, X){
n=length(X)
logvrais=n*log(alpha)+n*log(lambda)+
(alpha-1)*sum(log(1-exp(-lambda*x)))-lambda*sum(x)
return(-logvrais)
}

estimation_MV=mle(vrais,start=list(alpha=1,lambda=1))
```

```
##### MCMC avec Gibbs #####

posterior=function(lambda,alpha,X){
a=0
b=0
e=0
d=0
n=length(X)
post_lambda=(lambda^(n+b-1))*exp(-lambda*(a+sum(X)))*
prod((1-exp(-lambda*X))^(alpha-1))
postlam_lambda}

MCMC_Gibbs<-function(T,lambda0,alpha0,X){
e=0
d=0
n=length(x)
lambda=0
alpha=0
ta=0
r=0
t=1
lambda[1]<-lambda0
alpha[1]=alpha0
while (t<=T){
print(t)
lambdac=rnorm(1,lambda[t],0.008) #####loi de proposition (candidate)
```

```
lambdac=abs(lambdac)
r[t]=min(0,log(posterior(lambdac,alpha[t],X))-log(posterior(lambda[t],alpha[t],X)))
u=log(runif(1))
if (u<r[t]){
  lambda[t+1]=lambdac
  ta=ta+1}
else{
  lambda[t+1]=lambda[t]}
alpha[t+1]=rgamma(1,n+d,e-sum(log(1-exp(-lambda[t+1]*x))))
t=t+1}
lambda
alpha
result<-list(lam=lam, alp=alp, ta=ta)
}
```

Bibliographie

- M. Abramowitz and I.A Stegun. *Handbook of Mathematical Functions*. New York : Dover, 1972.
- G N. Alexander. *The use of gamma distribution in estimating regulated output from storages*. Transactions in Civil Engineering, Institute of Engineers, Australia 4, 29–34, 1962.
- D E. Amos. A portable fortran subroutine for derivatives of the psi function. *Algorithm 610, ACM Transactions on Mathematical Software*, 9(4) :494–502, 1983.
- L J. Bain. *Statistical analysis of reliability and life testing model*. Marceland Dekker Inc., New York, 1976.
- A. Gelman, G O. Roberts, and W R. Gilks. Efficient metropolis jumping rules. *Bayesian Statistics*, 5 :599–608, 1996.
- A C. Gorski. Beware of the weibull euphoria. *Transactions of IEEE–Reliability*, 17 : 202—203, 1968.
- R D. Gupta and D. Kundu. Generalized exponential distributions. *Austr. New Zealand J. Statistl*, 41 :173—188, 1999.
- R D. Gupta and D. Kundu. Generalized exponential distributions : different methods of estimation. *J. Statist. Comput. Simulation*, 69 :315—338, 2001a.

- R D. Gupta and D. Kundu. Exponentiated exponential family : An alternative to gamma and weibull distributions. *Biometrical Journal*, 43 :117—130, 2001b.
- W. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 75 :97–109, 1970.
- O A Y. Jackson. Fitting a gamma or log normal distribution to fiber diameter measurements of wooldtops. *Applied Statistics*, 1 :161—166, 1963.
- Z F. Jaheen. Empirical bayes inference for generalized exponential distribution based on records. *Commun. Statist. Theory Methods*, 33(8) :1851—1861, 2001.
- L G. Johnson. *The probabilistic basic of cumulative damage*. Transactions of the 22nd. Technical Conference of the American Society of Quality Control, 133–140, 1968.
- N L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions-1*. 2nd edition. John Wiley and Sons, New York, 1994.
- N L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions-2*. 2nd edition. John Wiley and Sons, New York, 1995.
- J V. Klinken. A method for inquiring whether the gamma distribution represents the frequency distribution of industrialaccident costs. *Actuarlele Studien*, 3 :83—92, 1961.
- D. Kundu and R D. Gupta. Generalized exponential distribution : Bayesian estimations. *Computational Statistics & Data Analysis*, 52 :1873—1883, 2008.
- J F. Lawless. *Statistical Models and Methods for Lifetime Data*. Wiley, NewYork, 1982.
- D V. Lindley. *Approximate Bayesian method*. *Trabajos Estadist.* 31, 223–237, 1980.
- H. Linhart and W. Zucchini. *Model Selection*. Wiley, NewYork, 1986.
- M. Masuyama and Y. Kuroiwa. *Tables for the likelihood solutions of gamma distribution and its medicalappl ications*. Reports of Statistical Applications Research (JUSE) 1, 18–23, 1952.

- N. Metropolis, A W. Rosenbluth, M N. Rosenbluth, A H. Teller, and E. Teller. Equations of state calculations by fast computing machine. *J. Chem. Phys*, 21 :1087–1093, 1953.
- A. Plait. The weibull distribution—with tables. *Industrial Quality Control*, 19 :17—26, 1962.
- M Z. Raqab. Inferences for generalized exponential distribution based on record statistics. *J. Statist. Plann. Inference*, 104 :339—350, 2002.
- M Z. Raqab and M. Ahsanullah. Estimation of the location and scale parameters of generalized exponential distribution based on order statistics. *J. Statist. Comput. Simulation*, 69 :109—124, 2001.
- W. Weibull. *A statistical theory of the strength of material*. Ingeniors Vetenskaps Akademiens, Stockholm 151, 1939.