



Université ABDERAHMANE MIRA BEJAÏA

Faculté des Sciences Exactes

Département de Recherche Opérationnelle

Mémoire-fin du cycle

En Vue l'obtention du diplôme de Master

Spécialité : Modélisation Mathématique et Techniques de Décision

Thème :

MODÈLE KUMARSWAMY GAMMA GÉNÉRALISÉ :
THÉORIE ET APPLICATION

Présenté par :

M_r Soudani Abdelkadir

M_r Mouhous Fahem

Soutenue publiquement le 03/07/2017 devant le jury suivant :

Présidente	Amroun Sonia	MAA	U.de Béjaïa.
Encadreur	Zougab Nabil	MCA	U.de Béjaïa.
Examinatrice	Djerroud Lamia	Docteur(LMD)	U.de Béjaïa.
Examinatrice	Ziane Yasmina	MAB	U.de Béjaïa.

Remerciements

Tout d'abord nous tenons à remercier **ELLAH** le clément qui nous a donné la force et le courage d'aller au bout de notre objectif.

Nos remerciements les plus vifs vont à l'attention de notre promoteur *D^r* N.ZOUGAB pour avoir accepté de diriger ce travail, pour ses conseils et orientations.

Nous tenons à remercier les membres du jury d'avoir accepté de juger notre travail.

Nous remercions tous nos enseignants.

Enfin nous remercions toute personne ayant contribué de près ou de loin à l'élaboration de ce travail.

Dédicace

Je dédie ce travail :

- * *À* Mes chers parents, qui, par leur dévouement, leur engagement et leur générosité, m'ont permis de suivre ce bon chemin, que **ELLAH** les protèges.

- * *À* Mes frères et sœurs : Nadjemeddine, A. Anour, A. Djalile, Nour Elhouda, Ahlem, Souaad.

- * *À* Mes bon amis : B. Ayoub , S. Amine, M.A. Ouahab, M. Salah.

- * *À* Mes oncles et toutes la famille Soudani.

- * *À* Mon binôme **FAHEM**, merci pour ton sérieux et compréhension.

Soudani Abdelkadir

Dédicace

Je dédie ce travail :

- * *A* Mes chers parents, qui, par leur dévouement, leur engagement et leur générosité, m'ont permis de suivre ce bon chemin, que **ELLAH** les protèges.

- * *A* Mes sœurs et frères : Lamia, Kahina, Tarik.

- * *A* Mes bon amis : Soufian, Amer.

- * *A* Mes oncles et toutes la famille Mouhous.

- * *A* Mon binôme **ABDELKADIR**, merci pour ton sérieux et compréhension.

Mouhous Fahem

Table des matières

Introduction générale	6
1 Modèle Kumaraswamy gamma généralisé	8
1.1 Introduction	8
1.2 Modèle gamma généralisé	8
1.3 Distribution de Kumaraswamy	9
1.4 Distribution de Kumaraswamy gamma généralisé	10
1.4.1 Fonction de répartition	10
1.4.2 Fonction de densité de probabilité	10
1.4.3 Fonction de survie	17
1.4.4 Fonction hasard	17
1.4.5 Développement de KumGG sous forme d'un mélange de GG	18
1.4.6 Moments d'ordre r	19
1.4.7 Simulation d'une variable aléatoire de loi KumGG	21
1.5 Conclusion	22
2 Estimation dans le modèle KumGG	23
2.1 Méthode du maximum de vraisemblance	23
2.1.1 Estimateurs du maximum de vraisemblance	23
2.1.2 Estimation par intervalle de confiance	25
2.2 Analyse bayésien	26
2.2.1 Principe bayésien	26

2.2.2	Choix des lois a priori	26
2.2.3	Calcul des lois a posteriori	27
2.2.4	Méthode MCMC	28
2.3	Conclusion	29
3	Application sur des données simulées	30
3.1	Introduction	30
3.2	Simulation des données	30
3.3	Estimation des paramètres	31
3.4	Résultats	31
3.5	Interprétation des résultats	35
4	Application sur des données réelles	37
4.1	Introduction	37
4.2	Modèle exponentiel généralisé	37
4.3	Première application	38
4.4	Deuxième application	39
4.5	Troisième application	40
4.6	Discussion	41
	Conclusion générale	42
		43
	Annexe	43
A.1	Développement en série entière de la fonction gamma incomplète	43
A.2	Elements de la matrice d'information du Fisher	44
A.3	Code source avec R	46
	Bibliographie	54

Introduction générale

Le modèle gamma généralisé (GG) à trois paramètres introduit par Stacy [6] est le modèle le plus populaire pour traiter des données d'analyse de survie pour plusieurs raisons : d'abord, ce modèle contient plusieurs cas particuliers (sous-modèles spéciaux) comme la distribution exponentielle, gamma, Weibull, Rayleigh, etc. Ensuite, il est approprié pour analyser des données avec différentes formes de la fonction hasard : croissante, décroissante, forme baignoire et des données unimodal. Comme troisième raison, le modèle GG trouve ses applications dans divers domaines comme l'hydrologie, l'économie, la biologie, les sciences de l'environnement, etc.

La distribution GG a été étudiée et utilisée par plusieurs auteurs. Par exemple, Nadarajah et Gupta [26] ont utilisé cette distribution pour analyser des données de sécheresse. Stacy et Mihram [7] ont proposé la méthode de maximum de vraisemblance pour estimer les paramètres du modèle GG. Cependant, on peut se heurter au problème de divergence de l'approche par maximum de vraisemblance (voir Kleiber et Kotz [3] et Hiross [11]). Pour remédier à ce problème, Gomes et al. [20] ont présenté une méthode itérative (heuristique) basée sur les tests d'ajustement pour estimer les trois paramètres de la distribution GG. Ali et al. [17] ont obtenu les distributions exactes du produit $T_1 T_2$ et du quotient T_1/T_2 , où T_1 et T_2 suivent comme lois les distributions GG ; voir aussi Malhotra et al. [14], Ortega et al. [5] pour d'autres contextes.

Le modèle de Kumaraswamy (Kum) doublement borné proposé par Kumaraswamy [23] est une loi de probabilité continue dont le support est $]0, 1[$, qui dépend de deux paramètres de forme. Ce modèle est similaire au modèle bêta, mais sa simplicité réside dans la forme simple de sa fonction de densité de probabilité et sa fonction de répartition. Dans la littérature, le modèle Kum est moins populaire que le modèle GG. Cependant, dans un article récent, Jones [16] a montré les avantages de la distribution Kum, en la comparant surtout avec le modèle bêta.

Dans ce travail nous nous sommes intéressés au modèle Kumaraswamy gamma généralisé (KumGG) à cinq paramètres proposé par de Pascoa et al. [21]. Ce modèle s'obtient, en combinant le modèle GG et le modèle Kum. Ce nouveau modèle englobe plusieurs sous-modèles

spéciaux et trouve ses applications dans divers domaines. Notre travail présenté ici est basé sur l'article de *Marcelino A.R. de Pascoa., Edwin M.M. Ortega., Gauss M. Cordeiro. The Kumaraswamy generalized gamma distribution with application in survival analysis. Statistical Methodology, 5(8) :411–433, 2011.*

Le premier objectif de ce travail est de présenter en détail le modèle KumGG. Le second objectif est de faire le point sur les méthodes d'estimation des cinq paramètres du modèle KumGG. Nous présentons la méthode du maximum de vraisemblance et l'approche bayésienne, développées toutes les deux par de Pascoa et al. [21]. Ces approches seront utilisées seulement dans le cas des données complètes.

Le troisième objectif de ce travail est d'analyser des données simulées et réelles par le modèle KumGG en estimant les paramètres de ce modèle par l'approche bayésienne, en utilisant les techniques de Monté Carlo par Chaîne de Markov (MCMC). Une comparaison au sens d'informations AIC et BIC avec le modèle exponentiel généralisé (EG) sur des données réelles sera présentée.

Nous avons structuré ce mémoire en quatre parties principales. Le premier chapitre est consacré à un rappel détaillé sur le modèle KumGG à cinq paramètres. Dans le second chapitre, on présente les différentes procédures d'estimation des paramètres du modèle KumGG 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 Kumaraswamy gamma généralisé

1.1 Introduction

Le but de ce chapitre est de présenter en détail le modèle KumGG. Dans un premier temps, nous présentons le modèle GG introduit par Stacy [6]. Dans un second temps, nous présentons le modèle Kum proposé par Kumaraswamy [23]. Enfin, nous exposons en détail le modèle KumGG développé par de Pascoa et al. [21], qui est considéré comme une combinaison des modèles GG et Kum.

1.2 Modèle gamma généralisé

Le modèle gamma généralisé (GG) à trois paramètres introduit par Stacy [6] est l'un des modèles le plus populaire pour traiter des données d'analyse de survie pour plusieurs raisons : d'abord, ce modèle contient plusieurs cas particuliers (sous-modèles spéciaux) comme la distribution exponentielle, gamma, Weibull, Rayleigh, Maxwell, etc. Soit T une variable aléatoire qui suit une loi GG de paramètres α , τ et k , notée $GG(\alpha, \tau, k)$. Alors sa fonction de répartition G est donnée par :

$$G_{\alpha, \tau, k}(t) = \frac{\gamma(k, (t/\alpha)^\tau)}{\Gamma(k)}, \quad t > 0; \quad \alpha, \tau, k > 0. \quad (1.1)$$

où $\gamma(k, x) = \int_0^x w^{k-1} e^{-w} dw$ est la fonction gamma incomplète et $\Gamma(k) = \int_0^\infty w^{k-1} e^{-w} dw$ est la fonction gamma (complète). Les propriétés de base de la distribution GG sont données par Stacy et Mihram [7] et Lawless [12] et [13]. Le modèle GG englobe plusieurs sous-modèles spéciaux présentés dans le tableau (1.1).

Distribution	τ	α	k
Gamma	1	α	k
Khi-deux	1	2	$\frac{n}{2}$
Exponentiel	1	α	1
Weibull	c	α	1
Rayleigh	2	α	$\frac{3}{2}$
Maxwell	2	α	$\frac{1}{2}$

TABLE 1.1 – Modèles spéciaux du modèle GG.

1.3 Distribution de Kumaraswamy

Le modèle de Kumaraswamy (Kum) doublement borné proposé par Kumaraswamy est une loi de probabilité continue dont le support est $]0, 1[$, qui dépend de deux paramètres de forme λ et φ . Ce modèle est similaire au modèle bêta, mais sa simplicité réside dans la forme simple de sa fonction de densité de probabilité et sa fonction de répartition ; voir Jones [16]. Soit X une variable aléatoire qui suit la loi Kum de paramètres λ et φ , notée $\text{Kum}(\lambda, \varphi)$. Alors sa fonction de répartition F est donnée par :

$$F_{\lambda, \varphi}(x) = 1 - (1 - x^\lambda)^\varphi, \quad 0 < x < 1; \quad \lambda > 0, \varphi > 0. \quad (1.2)$$

La fonction de densité de probabilité de la variable aléatoire X qui suit $\text{Kum}(\lambda, \varphi)$ est obtenue comme suit :

$$f(x) = \frac{dF_{\lambda, \varphi}(x)}{dx} = \lambda \varphi x^{\lambda-1} (1 - x^\lambda)^{\varphi-1}, \quad 0 < x < 1; \quad \lambda, \varphi > 0. \quad (1.3)$$

Dans un travail récent, Jones [16] a montré plusieurs avantages de la distribution Kum, en la comparant au modèle bêta.

1.4 Distribution de Kumaraswamy gamma généralisé

Le modèle KumGG a été proposé récemment par de Pascoa et al. [21]. Ce modèle s'obtient en combinant les modèles GG et Kum. Nous présentons les définitions et quelques propriétés du modèle KumGG.

1.4.1 Fonction de répartition

La fonction de répartition de la variable aléatoire T qui suit le modèle KumGG à cinq paramètres, s'obtient en remplaçant x dans la fonction de répartition du modèle Kum donnée par (1.2) par la fonction de répartition du modèle GG donnée par (1.1) :

$$F(t) = 1 - (1 - G_{\alpha,\tau,k}(t)^\lambda)^\varphi \quad (1.4)$$

$$= 1 - \left(1 - \{\gamma_1(k, (t/\alpha)^\tau)\}^\lambda\right)^\varphi \quad (1.5)$$

où $\gamma_1(k, x) = \frac{\gamma(k, x)}{\Gamma(k)}$.

1.4.2 Fonction de densité de probabilité

Une variable aléatoire continue T suit une loi KumGG à cinq paramètres, si elle admet pour densité de probabilité la fonction :

$$\begin{aligned} f(t) &= \frac{dF(t)}{dt} \\ &= \frac{\lambda\varphi\tau}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{\tau k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \left(\gamma_1\left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda-1} \\ &\quad \times \left(1 - \left[\gamma_1\left(k, \left(\frac{t}{\alpha}\right)^\tau\right)\right]^\lambda\right)^{\varphi-1}, \quad t > 0; \alpha, \tau, k, \lambda, \varphi > 0. \end{aligned} \quad (1.6)$$

La Figure (1.1) représente les courbes des densités de probabilité de la loi KumGG pour différentes valeurs de ces cinq paramètres.

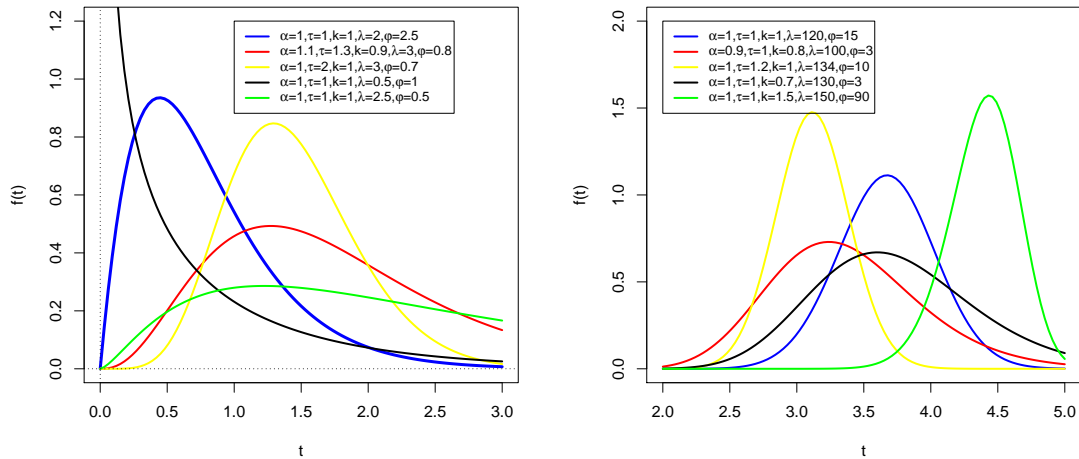


FIGURE 1.1 – Densité de probabilité de la distribution KumGG.

La distribution KumGG converge vers la distribution log-normale (LN) lorsque $k \rightarrow \infty$ et $\varphi = \lambda = 1$.

Quelques propriétés du modèle KumGG

Nous donnons quelques propriétés du modèle KumGG

- Si $T \sim \text{KumGG}(\alpha, \tau, k, \lambda, \varphi) \implies bT \sim \text{KumGG}(b\alpha, \tau, k, \lambda, \varphi)$, $b > 0$.
- Si $X \sim \text{KumGG}(1, 1, k, \lambda, \varphi) \implies T = \alpha X^{\frac{1}{\tau}} \sim \text{KumGG}(\alpha, \tau, k, \lambda, \varphi)$.
- Si $T \sim \text{KumGG}(\alpha, \tau, k, \lambda, \varphi) \implies T^m \sim \text{KumGG}(\alpha^m, \frac{\tau}{m}, k, \lambda, \varphi)$, $m \neq 0$.

Quelques modèles spéciaux du modèle KumGG

Nous donnons quelques modèles spéciaux (cas particuliers) et l'organigramme de ces modèles spéciaux.

1. Distribution exponentielle gamma généralisée

Si $\varphi = 1$, le modèle KumGG se réduit au modèle exponentiel gamma généralisé (EGG), sa fonction densité introduite par Cordeiro et al. [8] est donnée par :

$$f(t) = \frac{\lambda\tau}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{\tau k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \left(\gamma_1\left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda-1}, t > 0.$$

Si $\tau = \varphi = 1$ avec $k = 1$, le cas spécial correspond à la distribution exponentielle exponentielle (EE) donnée par [25] et [24] de fonction de densité donnée par :

$$f(t) = \frac{\lambda}{\alpha} \exp\left(-\frac{t}{\alpha}\right) \left(\gamma_1\left[1, \frac{t}{\alpha}\right]\right)^{\lambda-1}, t > 0.$$

Si $\tau = 2$ avec $k = \varphi = 1$, on obtient la distribution Rayleigh généralisée (RG) de fonction de densité donnée par :

$$f(t) = \frac{2\lambda t}{\alpha^2} \exp\left(-\frac{t}{\alpha}\right)^2 \left(\gamma_1\left[1, \left(\frac{t}{\alpha}\right)^2\right]\right)^{\lambda-1}, t > 0.$$

2. Distribution Kum-Weibull [9]

Pour $k = 1$, on obtient la distribution Kum-Weibull (KumW), sa fonction de densité est donnée par :

$$f(t) = \frac{\lambda\varphi\tau}{\alpha} \left(\frac{t}{\alpha}\right)^{\tau-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \left(\gamma_1\left[1, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda-1} \left(1 - \left[\gamma_1\left(1, \left(\frac{t}{\alpha}\right)^\tau\right)\right]^\lambda\right)^{\varphi-1}, t > 0.$$

Si $\varphi = k = 1$, le modèle KumGG se réduit au modèle exponentiel Weibull (EW), sa fonction de densité est donnée par :

$$f(t) = \frac{\lambda\tau}{\alpha} \left(\frac{t}{\alpha}\right)^{\tau-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \left(\gamma_1\left[1, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda-1}, t > 0.$$

Si $\varphi = \lambda = k = 1$, alors on obtient le modèle de Weibull (W) de fonction de densité donnée par :

$$f(t) = \frac{\tau}{\alpha} \left(\frac{t}{\alpha}\right)^{\tau-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right], t > 0.$$

Si $\tau = 2$ et $k = 1$, nous obtenons la distribution Kum-Rayleigh (KumR) de fonction de densité donnée par :

$$f(t) = \frac{2\lambda t\varphi}{\alpha^2} \exp\left[-\left(\frac{t}{\alpha}\right)^2\right] \left(\gamma_1\left[1, \left(\frac{t}{\alpha}\right)^2\right]\right)^{\lambda-1} \left(1 - \left[\gamma_1\left(1, \left(\frac{t}{\alpha}\right)^2\right)\right]^\lambda\right)^{\varphi-1}, t > 0.$$

Si $k = \tau = 1$, on obtient la distribution Kum-Exponentiel (KumE) de fonction de densité donnée par :

$$f(t) = \frac{\lambda\varphi}{\alpha} \exp\left(-\frac{t}{\alpha}\right) \left(\gamma_1\left[1, \frac{t}{\alpha}\right]\right)^{\lambda-1} \left(1 - \left[\gamma_1\left(1, \frac{t}{\alpha}\right)\right]^\lambda\right)^{\varphi-1}, t > 0.$$

Si $\varphi = \lambda = k = 1$, alors on a deux sous-modèles spéciaux : la distribution exponentielle (E) si $\tau = 1$ et le modèle de Rayleigh (R) si $\tau = 2$ de densités données respectivement par :

$$f(t) = \frac{1}{\alpha} \exp\left(-\frac{t}{\alpha}\right)$$

et

$$f(t) = \frac{2t}{\alpha^2} \exp\left[-\left(\frac{t}{\alpha}\right)^2\right], t > 0.$$

3. Distribution Kum-Gamma [9]

Pour $\tau = 1$, le modèle KumGG se réduit à Kum-Gamma à quatre paramètres (KumG), sa densité est donnée par :

$$f(t) = \frac{\lambda\varphi}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{k-1} \exp\left(-\frac{t}{\alpha}\right) \left(\gamma_1\left[k, \frac{t}{\alpha}\right]\right)^{\lambda-1} \left(1 - \left[\gamma_1\left(k, \frac{t}{\alpha}\right)\right]^\lambda\right)^{\varphi-1}, t > 0.$$

Si $\varphi = \tau = 1$, on obtient la distribution exponentielle gamma (EG) avec trois paramètres de densité donnée par :

$$f(t) = \frac{\lambda}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{k-1} \exp\left(-\frac{t}{\alpha}\right) \left(\gamma_1\left[k, \frac{t}{\alpha}\right]\right)^{\lambda-1}, t > 0.$$

Si $\varphi = \tau = \alpha = 1$, alors on obtient la distribution exponentielle gamma (EG) avec deux paramètres tel que :

$$f(t) = \frac{\lambda}{\Gamma(k)} (t)^{k-1} \exp(-t) (\gamma_1[k, t])^{\lambda-1}, t > 0.$$

De plus, si $k = 1$, nous obtenons la distribution Kum-Gamma avec un seul paramètre tel que :

$$f(t) = \lambda \exp(-t) (\gamma_1[1, t])^{\lambda-1}, t > 0.$$

Si $\varphi = \lambda = \tau = 1$, on obtient la distribution gamma à deux paramètres (G) de densité donnée par :

$$f(t) = \frac{1}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{k-1} \exp\left(-\frac{t}{\alpha}\right), t > 0.$$

De plus, si $k = 1$, on obtient la distribution exponentielle (E).

4. Distribution Kum-Khi-deux

Pour $\tau = \alpha = 2$ et $k = p/2$, on obtient la distribution Kum-Khi-deux (KumKhid) de fonction de densité donnée par :

$$f(t) = \frac{\lambda\varphi}{\Gamma(\frac{p}{2})} \left(\frac{t}{2}\right)^{\frac{p}{2}-1} \exp\left[-\left(\frac{t}{2}\right)^2\right] \left(\gamma_1\left[\frac{p}{2}, \left(\frac{t}{2}\right)^2\right]\right)^{\lambda-1} \left(1 - \left[\gamma_1\left(\frac{p}{2}, \left(\frac{t}{2}\right)^2\right)\right]^\lambda\right)^{\varphi-1}.$$

Si $\varphi = 1, \alpha = \tau = 2$ et $k = p/2$, alors on obtient la distribution exponentielle-Khi-carré (EKhic) tel que :

$$f(t) = \frac{\lambda}{\Gamma(\frac{p}{2})} \left(\frac{t}{2}\right)^{\frac{p}{2}-1} \exp\left[-\left(\frac{t}{2}\right)^2\right] \left(\gamma_1\left[\frac{p}{2}, \left(\frac{t}{2}\right)^2\right]\right)^{\lambda-1}, t > 0.$$

Si $\varphi = \lambda = 1$ et si en plus de $\alpha = \tau = 2$ et $k = p/2$, on obtient la distribution de Khi-deux très connue de densité donnée par :

$$f(t) = \frac{1}{\Gamma(\frac{p}{2})} \left(\frac{t}{2}\right)^{\frac{p}{2}-1} \exp\left[-\left(\frac{t}{2}\right)^2\right], t > 0.$$

5. Distribution Kum-Maxwell [1]

Pour $\tau = 2, \alpha = \sqrt{\theta}$ et $k = 3/2$, le modèle KumGG se réduit au modèle Kum-Maxwell (KumM) tel que :

$$f(t) = \frac{4\lambda\varphi}{\sqrt{\pi}\theta^{\frac{3}{2}}} t^2 \exp\left[-\left(\frac{t^2}{\theta}\right)\right] \left(\gamma_1\left[\frac{3}{2}, \frac{t^2}{\theta}\right]\right)^{\lambda-1} \left(1 - \left[\gamma_1\left(\frac{3}{2}, \frac{t^2}{\theta}\right)\right]^\lambda\right)^{\varphi-1}, t > 0.$$

Pour $\varphi = 1, \tau = 2, \alpha = \sqrt{\theta}$ et $k = 3/2$, on obtient la distribution exponentielle de Maxwell (EM) tel que :

$$f(t) = \frac{4\lambda}{\sqrt{\pi}\theta^{\frac{3}{2}}} t^2 \exp\left[-\left(\frac{t^2}{\theta}\right)\right] \left(\gamma_1\left[\frac{3}{2}, \frac{t^2}{\theta}\right]\right)^{\lambda-1}, t > 0.$$

Si $\varphi = \lambda = 1$ et si en plus $\tau = 2, \alpha = \sqrt{\theta}$ et $k = 3/2$, le modèle se réduit à la distribution de Maxwell (M) de fonction de densité donnée par :

$$f(t) = \frac{4\lambda}{\sqrt{\pi}\theta^{\frac{3}{2}}} t^2 \exp\left[-\left(\frac{t^2}{\theta}\right)\right], t > 0.$$

6. Distribution Kum-Nakagami [22]

Pour $\tau = 2, \alpha = \sqrt{\frac{\omega}{\mu}}$ et $k = \mu$, on obtient la distribution Kum-Nakagami (KumN) tel que sa fonction de densité est donnée par :

$$f(t) = \frac{2\mu^\mu \lambda \varphi}{\omega^\mu \Gamma(\mu)} t^{2\mu-1} \exp \left[- \left(\frac{\mu t^2}{\omega} \right) \right] \left(\gamma_1 \left[\mu, \left(\frac{\mu t^2}{\omega} \right) \right] \right)^{\lambda-1} \left(1 - \left[\gamma_1 \left(\mu, \left(\frac{\mu t^2}{\omega} \right) \right) \right]^\lambda \right)^{\varphi-1}.$$

Pour $\varphi = 1, \tau = 2, \alpha = \sqrt{\frac{\omega}{\mu}}$ et $k = \mu$ on obtient la distribution exponentielle de Nakagami (EN) tel que :

$$f(t) = \frac{2\mu^\mu \lambda \varphi}{\omega^\mu \Gamma(\mu)} t^{2\mu-1} \exp \left[- \left(\frac{\mu t^2}{\omega} \right) \right] \left(\gamma_1 \left[\mu, \left(\frac{\mu t^2}{\omega} \right) \right] \right)^{\lambda-1}, t > 0.$$

si $\varphi = \lambda = 1$, en plus de $\tau = 2, \alpha = \sqrt{\frac{\omega}{\mu}}$ et $k = \mu$, nous obtenons la distribution de Nakagami (N) tel que :

$$f(t) = \frac{2\mu^\mu \lambda \varphi}{\omega^\mu \Gamma(\mu)} t^{2\mu-1} \exp \left[- \left(\frac{\mu t^2}{\omega} \right) \right], t > 0.$$

7. Distribution Kum semi-normale généralisée

Si $\tau = 2\gamma, \alpha = 2^{\frac{1}{2\gamma}} \theta$ et $k = 1/2$, le cas spécial est désigné sous le nom de la distribution Kum semi-normale généralisée (KumSNG) tel que :

$$f(t) = \frac{2\gamma \lambda \varphi}{2^{\frac{1}{2\gamma}} \theta \sqrt{\pi}} \exp \left[- \left(\frac{t}{2^{\frac{1}{2\gamma}} \theta} \right)^{2\gamma} \right] \left(\gamma_1 \left[\frac{1}{2}, \left(\frac{t}{2^{\frac{1}{2\gamma}} \theta} \right)^{2\gamma} \right] \right)^{\lambda-1} \times \left(1 - \left[\gamma_1 \left(\frac{1}{2}, \left(\frac{t}{2^{\frac{1}{2\gamma}} \theta} \right)^{2\gamma} \right) \right]^\lambda \right)^{\varphi-1}, t > 0.$$

Pour $\varphi = 1, \tau = 2\gamma, \alpha = 2^{\frac{1}{2\gamma}} \theta$ et $k = 1/2$, on obtient la distribution exponentielle semi-normale généralisée (ESNG) tel que :

$$f(t) = \frac{2\gamma \lambda}{2^{\frac{1}{2\gamma}} \theta \sqrt{\pi}} \exp \left[- \left(\frac{t}{2^{\frac{1}{2\gamma}} \theta} \right)^{2\gamma} \right] \left(\gamma_1 \left[\frac{1}{2}, \left(\frac{t}{2^{\frac{1}{2\gamma}} \theta} \right)^{2\gamma} \right] \right)^{\lambda-1}, t > 0.$$

Pour $\alpha = \sqrt{2}\theta$ et $\tau = k = 2$, on obtient la distribution Kum-semi-normale (KumSN) tel que :

$$f(t) = \frac{\lambda \varphi}{\theta} \left(\frac{t}{\theta} \right)^3 \exp \left[- \left(\frac{t}{\sqrt{2}\theta} \right)^2 \right] \left(\gamma_1 \left[2, \left(\frac{t}{\sqrt{2}\theta} \right)^2 \right] \right)^{\lambda-1} \times \left(1 - \left[\gamma_1 \left(2, \left(\frac{t}{\sqrt{2}\theta} \right)^2 \right) \right]^\lambda \right)^{\varphi-1}, t > 0.$$

Si $\varphi = 1, \alpha = \sqrt{2}\theta$ et $\tau = 2$ et $k = 1/2$, on obtient la distribution exponentielle semi-normale (ESN) tel que :

$$f(t) = \frac{\sqrt{2}\lambda}{\theta\sqrt{\pi}} \exp \left[- \left(\frac{t}{\sqrt{2}\theta} \right)^2 \right] \left(\gamma_1 \left[\frac{1}{2}, \left(\frac{t}{\sqrt{2}\theta} \right)^2 \right] \right)^{\lambda-1}, t > 0.$$

Si $\varphi = \lambda = 1, \alpha = \sqrt{2}\theta$ et $\tau = 2$ et $k = 1/2$, on obtient la distribution semi-normale (SN) très connue de fonction de densité donnée par :

$$f(t) = \frac{\sqrt{2}}{\theta\sqrt{\pi}} \exp \left[- \left(\frac{t}{\sqrt{2}\theta} \right)^2 \right],$$

Si $\varphi = \lambda = 1$ et si en plus $\tau = 2\gamma, \alpha = 2^{\frac{1}{2\gamma}}\theta$ et $k = 1/2$, nous obtenons la distribution semi-normale généralisée (SNG) [15] tel que :

$$f(t) = \frac{2\gamma}{2^{\frac{1}{2\gamma}}\theta\sqrt{\pi}} \exp \left[- \left(\frac{t}{2^{\frac{1}{2\gamma}}\theta} \right)^{2\gamma} \right], t > 0.$$

L'organigramme des modèles spéciaux est donné par la figure (1.2).

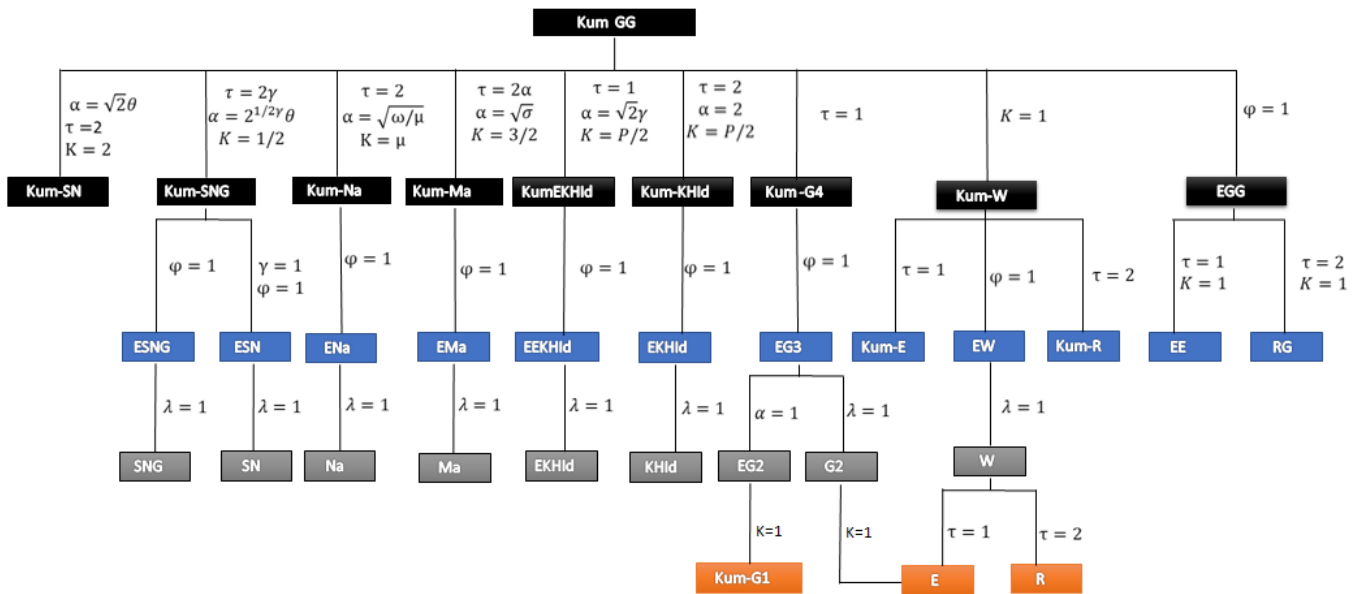


FIGURE 1.2 – Représentation des modèles spéciaux.

1.4.3 Fonction de survie

Soit T une variable aléatoire continue de loi KumGG de fonction de répartition F et de densité de probabilité f . Sa fonction de survie qui est la probabilité de survivre au-delà de t est définie par :

$$s(t) = 1 - F(t) = \left(1 - \left(\gamma_1 \left[k, \left(\frac{t}{\alpha} \right)^\tau \right] \right)^\lambda \right)^\varphi, t > 0. \quad (1.7)$$

En développant $\gamma_1(k, x)$ sous forme d'une série, la fonction survie peut s'écrire aussi comme suit :

$$s(t) = \left(1 - \left[\frac{1}{\alpha^{\tau k} \Gamma(k)} \sum_{j=0}^{\infty} \frac{(-1)^j \alpha^{(-\tau j)} t^{\tau(k+j)}}{(k+j)j!} \right]^\lambda \right)^\varphi. \quad (1.8)$$

La fonction de survie est une fonction décroissante telle que $S(0) = 1$ et $\lim_{t \rightarrow \infty} S(t) = 0$. La Figure (1.3) représente les courbes des fonctions de survie de la loi KumGG pour différentes valeurs des paramètres.

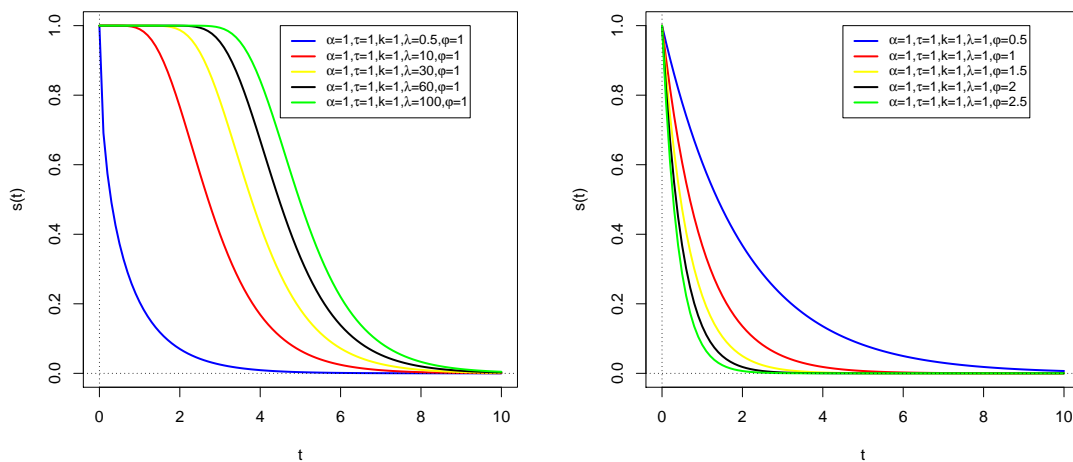


FIGURE 1.3 – Le graphe de La fonction de survie de KumGG selon les deux paramètres (λ, φ) .

1.4.4 Fonction hasard

La fonction de hasard est définie comme la probabilité conditionnelle que le phénomène se termine après une durée t sachant que l'on a atteint cette durée (taux de défaillance, taux

de décès ou risque instantané). En utilisant le théorème des probabilités conditionnelles et le modèle KumGG, la fonction hasard h est définie dans ce cas par :

$$h(t) = \frac{\lambda\varphi\tau}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{\tau k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \left(\gamma_1\left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda-1} \left(1 - \left[\gamma_1\left(k, \left(\frac{t}{\alpha}\right)^\tau\right)\right]^\lambda\right)^{-1}. \quad (1.9)$$

La Figure (1.4) représente les courbes des fonctions de hasard de la loi KumGG pour différentes valeurs des paramètres.

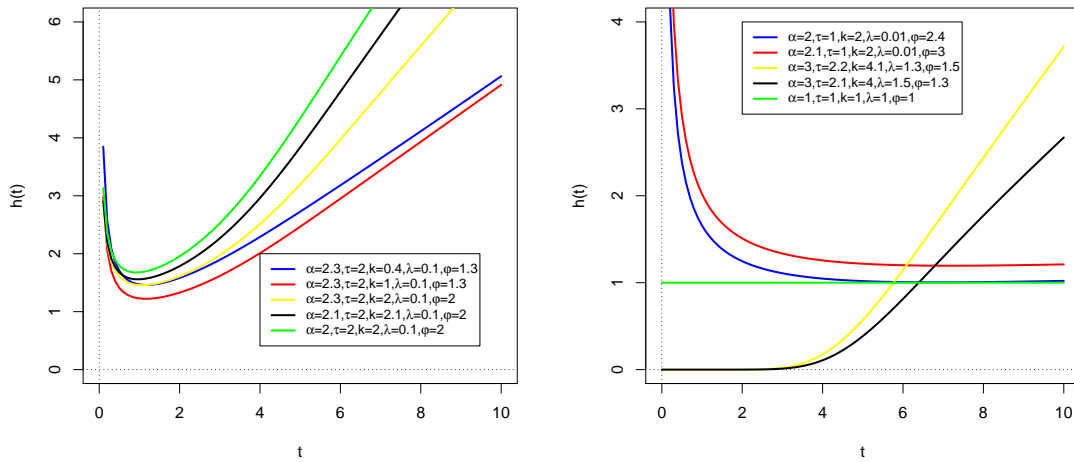


FIGURE 1.4 – Le graphe de La fonction hasard de KumGG selon les variations des paramètres.

1.4.5 Développement de KumGG sous forme d'un mélange de GG

Soit T une variable aléatoire qui suit la loi KumGG. Dans cette section, nous donnons la densité de KumGG sous forme d'un mélange de distributions GG. On a pour $|z| < 1$ et $b > 0$, $(1 - z)^{b-1}$ peut s'écrire en série entière comme suit :

$$(1 - z)^{b-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j) j!} z^j,$$

En utilisant cette série entière, la densité de T donnée par (1.6) peut être réécrite comme suit :

$$f(t) = \frac{\lambda\varphi\tau}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{\tau k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \left(\gamma_1\left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda-1} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\varphi)}{\Gamma(\varphi-j) j!} \left(\gamma_1\left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda j}. \quad (1.10)$$

En utilisant la formule (1) donnée dans l'annexe, on obtient :

$$f(t) = \frac{\lambda\varphi\tau}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{\tau k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \sum_{m,j=0}^{\infty} \frac{(-1)^j \Gamma(\varphi) S_m(\lambda)}{\Gamma(\varphi-j)j!} \left(\gamma_1 \left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda j+m}. \quad (1.11)$$

Si $\lambda j + m > 0$ on a :

$$\begin{aligned} \left(\gamma_1 \left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^{\lambda j+m} &= \left[1 - \left(1 - \gamma_1 \left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)\right]^{\lambda j+m+1-1} \\ &= \sum_{l=0}^{\infty} (-1)^l \binom{\lambda j+m}{l} \left(1 - \gamma_1 \left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^l. \end{aligned}$$

En utilisant la formule précédente on obtient :

$$\begin{aligned} f(t) &= \frac{\lambda\varphi\tau}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{\tau k-1} \exp\left[-\left(\frac{t}{\alpha}\right)^\tau\right] \sum_{j,l,m=0}^{\infty} \sum_{q=0}^l \frac{(-1)^{j+l+q} \Gamma(\varphi) S_m(\lambda)}{\Gamma(\varphi-j)j!} \\ &\quad \times \binom{\lambda j+m}{l} \binom{l}{q} \left(\gamma_1 \left[k, \left(\frac{t}{\alpha}\right)^\tau\right]\right)^q. \end{aligned} \quad (1.12)$$

En utilisant la formule (6) donnée dans l'annexe, on obtient la densité mixte suivante :

$$f(t) = \sum_{d,j,l,m=0}^{\infty} \sum_{q=0}^l w(d, j, l, q, k, m, \lambda, \varphi) g_{(\alpha, \tau, k(1+q)+d)}(t), t > 0 \quad (1.13)$$

avec

$$w(d, j, l, q, k, m, \lambda, \varphi) = \frac{(-1)^{j+l+q} \Gamma(\varphi) \Gamma[k(1+q)+d] S_m(\lambda) c_{q,d}}{\Gamma(k)^{q+1} \Gamma(\varphi-j)j!} \binom{\lambda j+m}{l} \binom{l}{q},$$

Les coefficients satisfont $\sum_{m,i=0}^{\infty} w(d, j, l, q, k, m, \lambda, \varphi) = 1$, la quantité $S_m(\lambda)$ est donnée par (2) dans l'annexe et les coefficients $c_{q,d}$ peuvent être déterminés à partir de la relation de récurrence (5) donnée aussi en annexe. L'équation (1.13) montre que la densité de KumGG peut être écrite comme un mélange de densités GG.

1.4.6 Moments d'ordre r

Considérons maintenant les différents moments de la distribution KumGG. Soit T une variable aléatoire qui suit la loi KumGG de densité f . Nous donnons deux façons pour calculer le moment d'ordre r pour la distribution KumGG. La première méthode est basée sur l'écriture de la densité KumGG sous forme d'un mélange de GG. La deuxième méthode est basée sur la définition (calcul de l'intégrale).

Première méthode

Le moment d'ordre r de la distribution GG est donnée par :

$$\mu'_{r,GG} = \frac{\alpha^r \Gamma(k+r/\beta)}{\Gamma(k)},$$

D'après l'équation (1.13), on obtient immédiatement :

$$\mu'_r = \lambda \varphi \alpha^r \Gamma(\varphi) \sum_{d,j,l,m=0}^{\infty} \sum_{q=0}^l \frac{(-1)^{j+l+q} \Gamma[k(1+q) + d + r/\tau] S_m(\lambda) c_{q,d} \binom{\lambda j + m}{l} \binom{l}{q}}{\Gamma(k)^{q+1} \Gamma(\varphi - j) j!}. \quad (1.14)$$

Cette méthode nécessite le calcul des quantités $c_{q,d}$.

Deuxième méthode

En calculant directement l'intégrale suivante :

$$\mu'_r = \int t^r f(t) dt,$$

on obtient

$$\begin{aligned} \mu'_r = \frac{\lambda \varphi \tau \alpha^{r-1}}{\Gamma(k)} \int_0^{\infty} \left(\frac{t}{\alpha}\right)^{\tau k + r - 1} \exp\left[-\left(\frac{t}{\alpha}\right)^{\tau}\right] \left(\gamma_1\left[k, \left(\frac{t}{\alpha}\right)^{\tau}\right]\right)^{\lambda-1} \\ \times \left(1 - \left[\gamma_1\left(k, \left(\frac{t}{\alpha}\right)^{\tau}\right)\right]^{\lambda}\right)^{\varphi-1} dt. \end{aligned} \quad (1.15)$$

En utilisant le changement de variable $x = \left(\frac{t}{\alpha}\right)^{\tau}$, l'équation (1.15) devient :

$$\mu'_r = \frac{\lambda \varphi \alpha^r}{\Gamma(k)} \int_0^{\infty} x^{k + \frac{r}{\tau} - 1} \exp(-x) \gamma_1(k, x)^{\lambda-1} [1 - \gamma_1(k, x)^{\lambda}]^{\varphi-1} dx, \quad (1.16)$$

Pour $\varphi > 0$ on a :

$$[1 - \gamma_1(k, x)^{\lambda}]^{\varphi-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\varphi)}{\Gamma(\varphi - j) j!} \gamma_1(k, x)^{\lambda j}, \quad (1.17)$$

On remplace (1.17) dans (1.16) on obtient :

$$\mu'_r = \frac{\lambda \varphi \alpha^r}{\Gamma(k)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\varphi)}{\Gamma(\varphi - j) j!} \int_0^{\infty} x^{k + \frac{r}{\tau} - 1} \exp(-x) \gamma_1(k, x)^{\lambda(1+j)-1} dx, \quad (1.18)$$

En utilisant le développement en série de $\gamma_1(k, x)^{\lambda(1+j)-1}$ donnée dans l'annexe et pour $\lambda > 0$, on obtient :

$$\gamma_1(k, x)^{\lambda(1+j)-1} = \sum_{l=0}^{\infty} \sum_{m=0}^l (-1)^{l+m} \binom{\lambda(1+j)-1}{l} \binom{l}{m} \gamma_1(k, x)^m, \quad (1.19)$$

On remplace (1.19) dans (1.18), on obtient :

$$\mu'_r = \frac{\lambda\varphi\alpha^r}{\Gamma(k)} \sum_{j,l=0}^{\infty} \sum_{m=0}^l v_{j,l,m} I(k, \frac{r}{\tau}, m), \quad (1.20)$$

avec

$$v_{j,l,m} = \frac{(-1)^{j+l+m} \Gamma(\varphi)}{\Gamma(\varphi-j)j!} \binom{\lambda(1+j)-1}{l} \binom{l}{m}$$

et

$$I(k, \frac{r}{\tau}, m) = \int_0^{\infty} x^{k+\frac{r}{\tau}-1} \exp(-x) \gamma_1(k, x)^m dx.$$

1.4.7 Simulation d'une variable aléatoire de loi KumGG

Dans cette section, nous présentons une méthode de simulation suivant la loi KumGG. Rappelons que dans le cas de la loi KumGG, la fonction de répartition est

$$F(t) = 1 - \left[1 - (\gamma_1(k, (t/\alpha)^\tau))^\lambda \right]^\varphi.$$

En utilisant la méthode de la fonction inverse, on peut obtenir l'équation non linéaire suivante :

$$(1 - U^{1/\varphi})^{1/\lambda} - \gamma_1(k, (t/\alpha)^\tau) = 0, \quad (1.21)$$

où U est une variable aléatoire uniforme sur $[0, 1]$. On peut simuler T , en résolvant l'équation non linéaire (1.21). Le programme **R** qui permet de simuler T de la loi KumGG est donnée dans l'annexe.

1.5 Conclusion

Dans ce chapitre, nous avons donné un rappel détaillé sur le modèle KumGG à cinq paramètres. 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 r sont donnés. Ensuite, la méthode inverse pour simuler une variable aléatoire T de loi KumGG est présenté. Dans le chapitre suivant, nous présentons les méthodes d'estimation des cinq paramètres de la distribution KumGG.

Chapitre 2

Estimation dans le modèle KumGG

Dans ce chapitre, nous présentons la méthode du maximum de vraisemblance (MV) et l'approche bayésienne pour estimer les cinq paramètres du modèle KumGG. Ces méthodes ont été développées par de Pascoa et al. [21], en particulier dans le cas des données censurées. Dans ce travail, nous présentons que le cas des données complètes.

2.1 Méthode du maximum de vraisemblance

Dans cette section, on discute les estimateurs du maximum de vraisemblance des cinq paramètres de la distribution KumGG.

2.1.1 Estimateurs du maximum de vraisemblance

Soit t_1, \dots, t_n un n observations d'un n -échantillon issu d'une variable T aléatoire de la loi KumGG de vecteur des paramètres $\theta = (\alpha, \tau, k, \lambda, \varphi)^\top$. Alors la fonction de vraisemblance notée par $L(t, \theta)$ est donnée par :

$$\begin{aligned}
L(t, \theta) &= \prod_{i=1}^n f(t_i, \theta) \\
&= \left(\frac{\lambda \varphi \tau}{\alpha \Gamma(k)} \right)^n \prod_{i=1}^n \left(\frac{t_i}{\alpha} \right)^{\tau k - 1} \exp \left[- \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\tau \right] \prod_{i=1}^n \left(\gamma_1 \left[k, \left(\frac{t_i}{\alpha} \right)^\tau \right] \right)^{\lambda - 1} \\
&\quad \times \prod_{i=1}^n \left(1 - \left[\gamma_1 \left(k, \left(\frac{t_i}{\alpha} \right)^\tau \right) \right]^\lambda \right)^{\varphi - 1}.
\end{aligned} \tag{2.1}$$

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

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

La recherche du MV peut se faire en résolvant le système d'équations :

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

Cependant, dans la pratique, on préfère remplacer ce problème (produit) par le problème équivalent pour la log-vraisemblance :

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

En utilisant l'expression (2.1) et l'équation (2.2), on obtient les équations non-linéaires suivantes

$$\left\{ \begin{array}{l}
U_\alpha(\theta) = -\frac{n\tau k}{\alpha} + \frac{\tau}{\alpha} \sum_{i=1}^n u_i - \frac{\tau}{\alpha} \sum_{i=1}^n v_i s_i + \frac{\lambda\tau(\varphi-1)}{\alpha\Gamma(k)} \sum_{i=1}^n u_i p_i = 0, \\
U_\tau(\theta) = \frac{n}{\tau} - \frac{1}{\tau} \sum_{i=1}^n u_i \log(u_i) + \frac{k}{\tau} \sum_{i=1}^n u_i \log(u_i) + \frac{1}{\tau} \sum_{i=1}^n v_i s_i \log(u_i) = 0, \\
U_k(\theta) = -n\lambda\psi(k) + \sum_{i=1}^n \log(u_i) + \sum_{i=1}^n s_i q_i + (\varphi-1)(n\lambda\psi(k)) \\
\quad \times \sum_{i=1}^n p_i \gamma_1(k, U_i) - \lambda(\varphi-1) \sum_{i=1}^n p_i q_i - \lambda\varphi \sum_{i=1}^n p_i q_i = 0, \\
U_\lambda(\theta) = \frac{n}{\lambda} + \sum_{i=1}^n \log(\gamma_1(k, U_i)) - (\varphi-1) \sum_{i=1}^n b_i (\gamma_1(k, U_i))^\lambda = 0, \\
U_\varphi(\theta) = \frac{n}{\varphi} + \sum_{i=1}^n \log(\omega_i) = 0.
\end{array} \right.$$

$$\text{où } u_i = \left(\frac{t_i}{\alpha}\right)^\tau, g_i = u_i^k \exp(-u_i), w_i = 1 - \gamma_1(k, U_i), v_i = \frac{g_i}{\Gamma(k)}, s_i = \frac{(\lambda - 1)}{\gamma_1(k, U_i)},$$

$$[\dot{\gamma}(k, U_i)]_k = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} J(u_i, k + n - 1, 1),$$

$$\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}, p_i = \frac{\gamma_1(k, U_i)^{\lambda-1}}{\omega_i}, q_i = \frac{[\dot{\gamma}(k, U_i)]_k}{\Gamma(k)}, b_i = \frac{\log(\gamma_1(k, U_i))}{\omega_i},$$

avec ψ est la fonction digamma et $J(u_i, k + n - 1, 1)$ est défini en annexe. Notons que la résolution de ce système est très difficile et n'admet pas des solutions explicites. Dans ce travail, nous utilisons les différentes fonctions (**mle**, **optim**, **nlm**, **bbsolve**) du logiciel statistique **R**. En introduisant les valeurs de vecteur initiale. $\theta^0 = (\alpha^0, \tau^0, \lambda^0, k^0, \varphi^0)^\top$, ces fonctions permet de retourner les estimateurs de $\theta = (\alpha, \tau, \lambda, k, \varphi)^\top$. Cependant, des difficultés (problème de divergence) peuvent être rencontré comme dans le cas du modèle GG.

2.1.2 Estimation par intervalle de confiance

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

$$\sqrt{n} \left[\widehat{\theta} - \theta \right] \longrightarrow \mathcal{N}_5 \left(\mathbf{0}, \mathbf{J}^{-1}(\theta) \right) \quad (2.3)$$

avec $\widehat{\theta}$ est l'estimateur MV de θ et $\mathbf{J}(\theta)$ est la matrice de l'information du Fisher donnée par :

$$\mathbf{J} = \mathbf{J}(\theta) = \begin{pmatrix} \dot{J}_{\alpha,\alpha} & \dot{J}_{\alpha,\tau} & \dot{J}_{\alpha,k} & \dot{J}_{\alpha,\lambda} & \dot{J}_{\alpha,\varphi} \\ \dot{J}_{\tau,\alpha} & \dot{J}_{\tau,\tau} & \dot{J}_{\tau,k} & \dot{J}_{\tau,\lambda} & \dot{J}_{\tau,\varphi} \\ \dot{J}_{k,\alpha} & \dot{J}_{k,\tau} & \dot{J}_{k,k} & \dot{J}_{k,\lambda} & \dot{J}_{k,\varphi} \\ \dot{J}_{\lambda,\alpha} & \dot{J}_{\lambda,\tau} & \dot{J}_{\lambda,k} & \dot{J}_{\lambda,\lambda} & \dot{J}_{\lambda,\varphi} \\ \dot{J}_{\varphi,\alpha} & \dot{J}_{\varphi,\tau} & \dot{J}_{\varphi,k} & \dot{J}_{\varphi,\lambda} & \dot{J}_{\varphi,\varphi} \end{pmatrix}.$$

Les éléments de la matrice \mathbf{J} sont donné explicitement dans l'annexe. Maintenant, pour obtenir les estimateurs par intervalle de confiance, il suffit de remplacer $\mathbf{J}^{-1}(\theta)$ par son estimateur naturel $\mathbf{J}^{-1}(\widehat{\theta})$.

2.2 Analyse bayésien

2.2.1 Principe bayésien

Le but de ce chapitre est de rappeler quelques notions sur l'analyse bayésienne. Supposant que $\mathbf{t} = (t_1, \dots, t_n)$ est le vecteur d'observations et $\theta \in \Theta$ 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 $\pi(\theta|\mathbf{t})$ du modèle est obtenue en utilisant la règle de Bayes

$$\pi(\theta|\mathbf{t}) = \frac{\pi(\mathbf{t}|\theta)\pi(\theta)}{\pi(\mathbf{t})}, \quad (2.4)$$

où $\pi(\mathbf{t}|\theta)$ est la fonction vraisemblance et $\pi(\mathbf{t}) = \int_{\Theta} \pi(\mathbf{t}|\theta)\pi(\theta)d\theta$ est la constante de normalisation (ne dépend pas de θ). En utilisant la moyenne a posteriori, l'estimateur bayésien est donné par :

$$\hat{\theta}_{bayes} = E(\theta|\mathbf{t}) = \int_{\Theta} \theta\pi(\theta|\mathbf{t})d\theta. \quad (2.5)$$

L'estimateur bayésien (2.5) est obtenu en utilisant la perte quadratique. Notons que d'autres estimateurs bayésiens peuvent être obtenue, en utilisant d'autres pertes, comme par exemple, la perte d'entropie.

2.2.2 Choix des lois a priori

Les lois a priori sur les paramètres inconnus sont données comme suit :

$$\alpha \sim \Gamma(a_1, b_1), \tau \sim \Gamma(a_2, b_2), k \sim \Gamma(a_3, b_3), \lambda \sim \Gamma(a_4, b_4), \varphi \sim \Gamma(a_5, b_5),$$

avec $\Gamma(a_i, b_i)$ est la distribution gamma de moyenne a_i/b_i , de variance a_i/b_i^2 , et de fonction de densité donnée par

$$f(v, a_i, b_i) = \frac{b_i^{a_i} v^{a_i-1} \exp(-vb_i)}{\Gamma(a_i)},$$

tel que $v > 0, a_i > 0, b_i > 0$.

2.2.3 Calcul des lois a posteriori

En utilisant le théorème de Bayes, la distribution a posteriori pour les paramètres $\alpha, \tau, k, \lambda, \varphi$ est donnée par :

$$\begin{aligned} \pi(\alpha, \tau, k, \lambda, \varphi/t) \propto & \left(\frac{\lambda \varphi \tau}{\alpha \Gamma(k)} \right)^n \prod_{i=1}^n \left(\frac{t_i}{\alpha} \right)^{\tau k - 1} \exp \left[- \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\tau \right] \prod_{i=1}^n \left(\gamma_1 \left[k, \left(\frac{t_i}{\alpha} \right)^\tau \right] \right)^{\lambda - 1} \\ & \times \prod_{i=1}^n \left(1 - \left[\gamma_1 \left(k, \left(\frac{t_i}{\alpha} \right)^\tau \right) \right]^\lambda \right)^{\varphi - 1} \times \pi(\alpha, \tau, k, \lambda, \varphi). \end{aligned} \quad (2.6)$$

Vu la complexité de la loi a posteriori (2.6), il est impossible d'obtenir directement les estimateurs bayésiens. Pour cette raison, la stratégie de calcul consiste d'abord à dériver les lois a posteriori conditionnelles des cinq paramètres et d'utiliser les méthodes de Monté Carlo par Chaîne de Markov (MCMC) proposée par Metropolis et al. [19] et Hastings [27] pour estimer les paramètres du modèle KumGG. A partir de (2.6), on peut obtenir les lois a posteriori conditionnelles des cinq paramètres comme suit :

$$\begin{aligned} \pi(\alpha/t, \tau, k, \lambda, \varphi) \propto & (\alpha^{\tau k})^{-n} \exp \left[- \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\tau \right] \prod_{i=1}^n \left(\gamma_1 \left[k, \left(\frac{t_i}{\alpha} \right)^\tau \right] \right)^{\lambda - 1} \\ & \times \prod_{i=1}^n \left(1 - \left[\gamma_1 \left(k, \left(\frac{t_i}{\alpha} \right)^\tau \right) \right]^\lambda \right)^{\varphi - 1} \times \pi(\alpha), \end{aligned}$$

$$\begin{aligned} \pi(\tau/t, \alpha, k, \lambda, \varphi) \propto & \left(\frac{\tau}{\alpha^{\tau k}} \right)^n \exp \left[- \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\tau \right] \prod_{i=1}^n t_i^{\tau k} \left(\gamma_1 \left[k, \left(\frac{t_i}{\alpha} \right)^\tau \right] \right)^{\lambda - 1} \\ & \times \prod_{i=1}^n \left(1 - \left[\gamma_1 \left(k, \left(\frac{t_i}{\alpha} \right)^\tau \right) \right]^\lambda \right)^{\varphi - 1} \times \pi(\tau), \end{aligned}$$

$$\begin{aligned} \pi(k/t, \alpha, \tau, \lambda, \varphi) \propto & (\alpha^{\tau k} \Gamma(k))^{-n} \prod_{i=1}^n t_i^{\tau k} \left(\gamma_1 \left[k, \left(\frac{t_i}{\alpha} \right)^\tau \right] \right)^{\lambda - 1} \\ & \times \prod_{i=1}^n \left(1 - \left[\gamma_1 \left(k, \left(\frac{t_i}{\alpha} \right)^\tau \right) \right]^\lambda \right)^{\varphi - 1} \times \pi(k), \end{aligned}$$

$$\pi(\lambda/t, \alpha, \tau, k, \varphi) \propto (\lambda)^n \prod_{i=1}^n \left(\gamma_1 \left[k, \left(\frac{t_i}{\alpha} \right)^\tau \right] \right)^\lambda \left(1 - \left[\gamma_1 \left(k, \left(\frac{t_i}{\alpha} \right)^\tau \right) \right]^\lambda \right)^{\varphi - 1} \times \pi(\lambda)$$

et

$$\pi(\varphi/t, \alpha, \tau, k, \lambda) \propto (\varphi)^n \prod_{i=1}^n \left(1 - \left[\gamma_1 \left(k, \left(\frac{t_i}{\alpha} \right)^\tau \right) \right]^\lambda \right)^\varphi \times \pi(\varphi).$$

2.2.4 Méthode MCMC

Notons que toutes les lois a posteriori calculées ne sont pas simulable directement. Pour cette raison, nous proposons d'utiliser les méthodes MCMC pour construire une chaîne de Markov pour le vecteur des paramètres $\theta = (\alpha, \tau, \lambda, k, \varphi)^\top$, en utilisant des valeurs initiales pour ce vecteur. Pour simuler suivant les lois a posteriori, nous utilisons l'algorithme de Metropolis-Hasting (M-H). Cet algorithme est basée sur l'utilisation d'une loi génératrice de candidat pour chaque paramètre. 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 de vecteur d'estimation des paramètres $\theta = (\alpha, \tau, \lambda, k, \varphi)^\top$. Les étapes de l'algorithme qui permet de simuler suivant les lois a posteriori sont résumées comme suit :

1 - Initialiser le vecteur des paramètres $\theta^{(1)} = (\alpha^{(1)}, \tau^{(1)}, \lambda^{(1)}, k^{(1)}, \varphi^{(1)})^\top$.

2 - Pour $m = \{2, \dots, N\}$.

(a) - Génération des lois candidates

- Générer $\alpha^c \sim N(\alpha^m, \sigma^2 = 0,0002)$; $\alpha^c = |\alpha^c|$.
- Générer $\tau^c \sim N(\tau^m, \sigma^2 = 0,0002)$; $\tau^c = |\tau^c|$.
- Générer $k^c \sim N(k^m, \sigma^2 = 0,0002)$; $k^c = |k^c|$.
- Générer $\lambda^c \sim N(\lambda^m, \sigma^2 = 0,0002)$; $\lambda^c = |\lambda^c|$.
- Générer $\varphi^c \sim N(\varphi^m, \sigma^2 = 0,0002)$; $\varphi^c = |\varphi^c|$.

(b)- Calculer les probabilités d'acceptations suivantes :

$$\bullet \rho_1 = \min \left\{ 1; \frac{\pi(\alpha^c | \tau^m, k^m \lambda^m, \varphi^m, t)}{\pi(\alpha^m | \tau^m, k^m \lambda^m, \varphi^m, t)} \right\}, \alpha^{(m)} = \begin{cases} \alpha^c & \text{si } u < \rho_1, u \sim U_{[0,1]} \\ \alpha^{(m-1)}, & \text{sinon} \end{cases}$$

$$\bullet \rho_2 = \min \left\{ 1; \frac{\pi(\tau^c | \alpha^{(m+1)}, k^m, \lambda^m, \varphi^m, t)}{\pi(\tau^m | \alpha^{(m+1)}, k^m, \lambda^m, \varphi^m, t)} \right\}, \tau^{(m)} = \begin{cases} \tau^c & \text{si } u < \rho_2, u \sim U_{[0,1]} \\ \tau^{(m-1)}, & \text{sinon} \end{cases}$$

$$\bullet \rho_3 = \min \left\{ 1; \frac{\pi(k^c | \alpha^{(m+1)}, \tau^{(m+1)}, \lambda^m, \varphi^m, t)}{\pi(k^m | \alpha^{(m+1)}, \tau^{(m+1)}, \lambda^m, \varphi^m, t)} \right\}, k^{(m)} = \begin{cases} k^c & \text{si } u < \rho_3, u \sim U_{[0,1]} \\ k^{(m-1)}, & \text{sinon} \end{cases}$$

$$\bullet \rho_4 = \min \left\{ 1; \frac{\pi(\lambda^c | \alpha^{(m+1)}, \tau^{(m+1)}, k^{(m+1)}, \varphi^m, t)}{\pi(\lambda^m | \alpha^{(m+1)}, \tau^{(m+1)}, k^{(m+1)}, \varphi^m, t)} \right\}, \lambda^{(m)} = \begin{cases} \lambda^c & \text{si } u < \rho_4, u \sim U_{[0,1]} \\ \lambda^{(m-1)}, & \text{sinon} \end{cases}$$

$$\bullet \rho_5 = \min \left\{ 1; \frac{\pi(\varphi^c | \alpha^{(m+1)}, \tau^{(m+1)}, k^{(m+1)}, \lambda^{(m+1)}, t)}{\pi(\varphi^m | \alpha^{(m+1)}, \tau^{(m+1)}, k^{(m+1)}, \lambda^{(m+1)}, t)} \right\}, \varphi^{(m)} = \begin{cases} \varphi^c & \text{si } u < \rho_5, u \sim U_{[0,1]} \\ \varphi^{(m-1)}, & \text{sinon} \end{cases}$$

3 - $m = m + 1$ et aller à 2.

4 - Calculer le vecteur d'estimation de Bayes $\hat{\theta}_{Bayes} = (\hat{\alpha}, \hat{\tau}, \hat{\lambda}, \hat{k}, \hat{\varphi})^\top$.

$$\hat{\alpha}_{Bayes} = \frac{1}{N - N_0} \sum_{m=N_0+1}^N \alpha^{(m)}, \hat{\tau}_{Bayes} = \frac{1}{N - N_0} \sum_{m=N_0+1}^N \tau^{(m)},$$

$$\hat{k}_{Bayes} = \frac{1}{N - N_0} \sum_{m=N_0+1}^N k^{(m)}, \hat{\lambda}_{Bayes} = \frac{1}{N - N_0} \sum_{m=N_0+1}^N \lambda^{(m)},$$

$$\hat{\varphi}_{Bayes} = \frac{1}{N - N_0} \sum_{m=N_0+1}^N \varphi^{(m)}.$$

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

2.3 Conclusion

Dans ce chapitre, nous avons exposé les méthodes d'estimation des paramètres du modèle KumGG, à savoir la méthode du maximum de vraisemblance 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. 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 pour le modèle KumGG en estimant ces paramètres par l'approche bayésienne.

Chapitre 3

Application sur des données simulées

3.1 Introduction

Dans ce chapitre, nous présentons les principaux résultats obtenus sur les simulations. Cette partie permet de tester l'efficacité de l'approche bayésienne dans l'estimation du modèle KumGG $(\alpha, \tau, k, \lambda, \varphi)$ sur les données simulées suivant ce modèle. Les programmes ont été réalisés à l'aide du logiciel **R**. Pour les simulations, nous procédons en deux étapes :

- Dans la première étape, nous utilisons le modèle KumGG $(\alpha, \tau, k, \lambda, \varphi)$ pour simuler les données \mathbf{t} en utilisant la relation (1.21) et en fixant le vecteur des paramètres $\theta = (\alpha, \tau, k, \lambda, \varphi)^\top$.
- La deuxième étape est consacrée à l'estimation des paramètres. Nous utilisons les observations t simulées pour estimer les paramètres avec l'approche bayésienne.

3.2 Simulation des données

Nous simulons, selon le modèle KumGG $(\alpha, \tau, k, \lambda, \varphi)$, des observations $\mathbf{t} = (t_1, \dots, t_n)$ de taille $n = (50, 100, 200)$, en fixant les paramètres de ce modèle. Nous étudions trois cas distincts selon Les paramètres fixés ont été choisis comme suit :

- 1^{er} cas : $\alpha = 1.1$, $\tau = 1.3$, $k = 0.9$, $\lambda = 3$, $\varphi = 0.8$.

- 2^{ème} cas : $\alpha = 1$, $\tau = 2$, $k = 1$, $\lambda = 3$, $\varphi = 0.7$.
- 3^{ème} cas : $\alpha = 1$, $\tau = 1$, $k = 1$, $\lambda = 0.5$, $\varphi = 1$.

3.3 Estimation des paramètres

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

- Pour l'approche bayésienne, nous utilisons les lois a priori en fixant les hyperparamètres $a_i = 0.01, b_i = 0.01, i = 1, \dots, 5$. Ensuite, nous appliquons le programme MCMC implémenté sous logiciel **R** en fixant les valeurs initiales $\theta^{(1)} = (\alpha^{(1)}, \tau^{(1)}, \lambda^{(1)}, k^{(1)}, \varphi^{(1)})^\top$. Le choix de ce dernier est très important dans le calcul des estimations bayésiennes . Notons que le nombre total d'itération est $N = 1500$ et la période de chauffage est $N_0 = 700$. Donc, nous utilisons que $N - N_0 = 800$ pour estimer le vecteur des paramètres. Nous avons utilisé $N_{sim} = 25$ réplifications.

3.4 Résultats

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

$$\bar{\alpha} = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \hat{\alpha}_i, \quad \bar{\tau} = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \hat{\tau}_i, \quad \bar{k} = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \hat{k}_i,$$

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

Les erreurs quadratiques moyennes sont données par :

$$EQM(\alpha) = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} (\hat{\alpha}_i - \alpha)^2, \quad EQM(\tau) = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} (\hat{\tau}_i - \tau)^2, \quad EQM(k) = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} (\hat{k}_i - k)^2,$$

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

où α, τ, k, λ et φ sont les valeurs fixées, $\hat{\alpha}, \hat{\tau}, \hat{\lambda}, \hat{k}$ et $\hat{\varphi}$ sont les estimateurs par l'approche bayésienne obtenus à partir de l'échantillon numéro i et $N_{sim} = 25$ est le nombre de répétition. Les figures (3.1) (3.2) (3.3) montrent les histogrammes des chaînes de Markov associées aux estimateurs des paramètres obtenues pour $n = 50$ avec une seule réplication. Notons que les résultats avec la méthode du maximum de vraisemblance n'ont pas été présentés pour cause de divergence de cette approche.

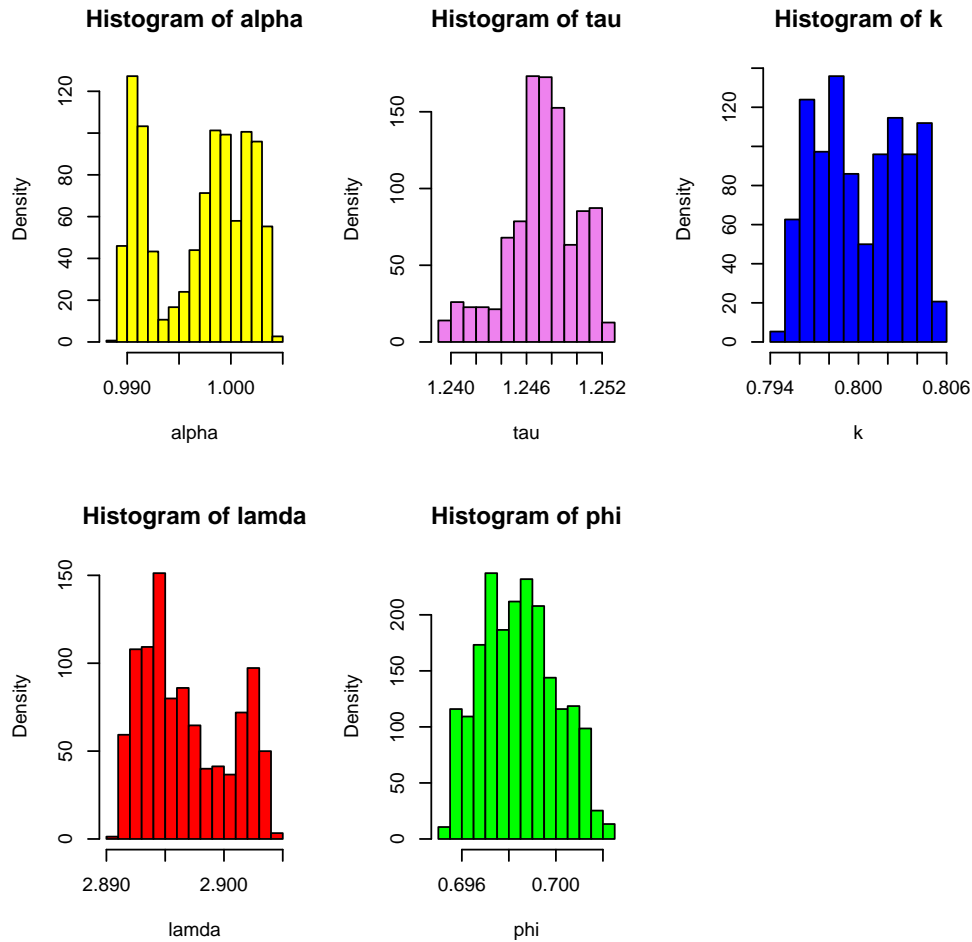


FIGURE 3.1 – Histogramme de la chaîne de Markov de 1^{er} cas.

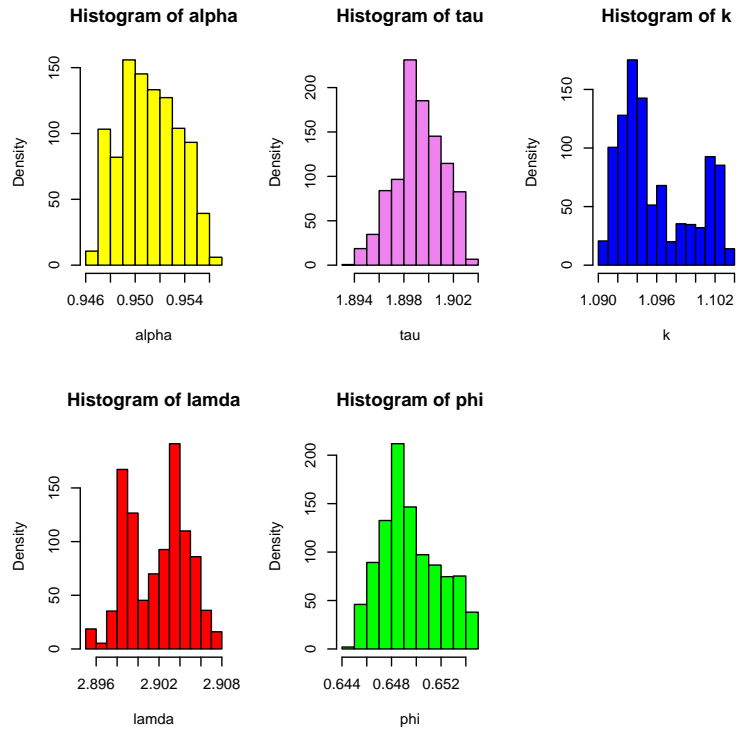


FIGURE 3.2 – Histogramme de la chaine de Markov de 2^{ème} cas.

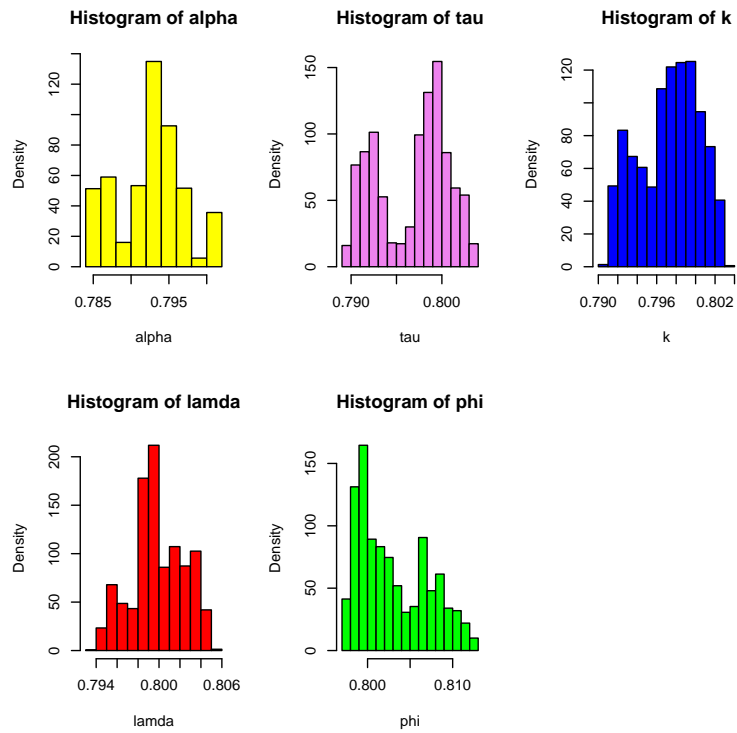


FIGURE 3.3 – Histogramme de la chaine de Markov de 3^{ème} cas.

La taille de l'échantillon		1 ^{er} cas	2 ^{ème} cas	3 ^{ème} cas
$n = 50$	$\bar{\alpha}$	0.999633	0.949136	0.800914
	EQM(α)	0.010154	0.002637	0.039679
	$\bar{\tau}$	1.259832	1.899866	0.799327
	EQM(τ)	0.001654	0.010057	0.040296
	\bar{k}	0.998894	1.098677	0.799302
	EQM(k)	0.009824	0.009779	0.040153
$n=100$	$\bar{\alpha}$	0.999224	0.951221	0.798760
	EQM(α)	0.010203	0.002416	0.040536
	$\bar{\tau}$	1.258127	1.900693	0.800474
	EQM(τ)	0.001807	0.009890	0.039841
	\bar{k}	0.998012	1.099545	0.801117
	EQM(k)	0.009653	0.009932	0.039575
$n=200$	$\bar{\alpha}$	1.000321	0.947946	0.800837
	EQM(α)	0.009989	0.002755	0.039699
	$\bar{\tau}$	1.259642	1.900418	0.802253
	EQM(τ)	0.001661	0.009942	0.039139
	\bar{k}	1.000284	1.099850	0.800383
	EQM(k)	0.010088	0.010012	0.039892
$n=200$	$\bar{\lambda}$	2.899971	2.901529	0.801027
	EQM(λ)	0.010048	0.009737	0.090663
	$\bar{\varphi}$	0.700826	0.650435	0.801003
	EQM(φ)	0.009892	0.002497	0.039627
	$\bar{\lambda}$	2.899392	2.902064	0.798303
	EQM(λ)	0.010156	0.009641	0.089016
$n=200$	$\bar{\varphi}$	0.700704	0.648231	0.799680
	EQM(φ)	0.009894	0.002713	0.040155

TABLE 3.1 – Estimation des paramètres dans les 3 cas .

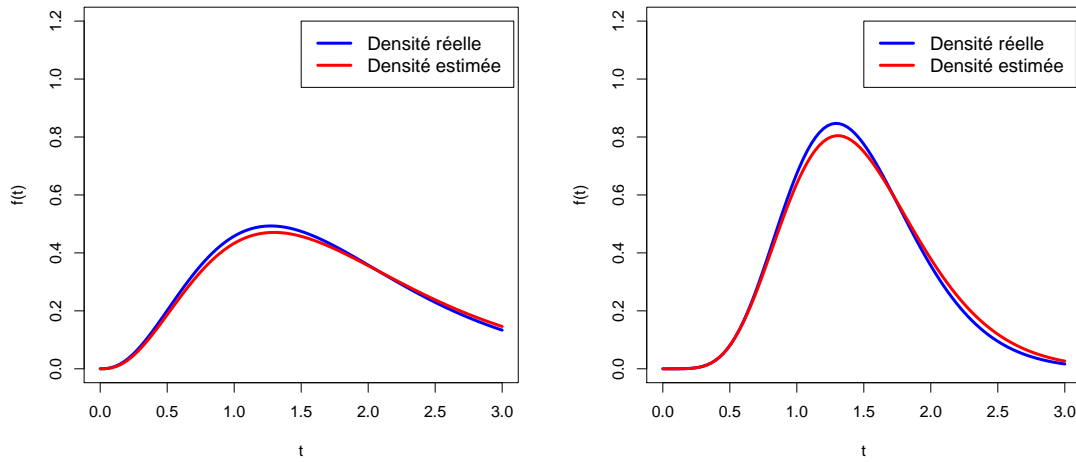


FIGURE 3.4 – Comparaison entre la vraie densité et celle estimée par Bayes.

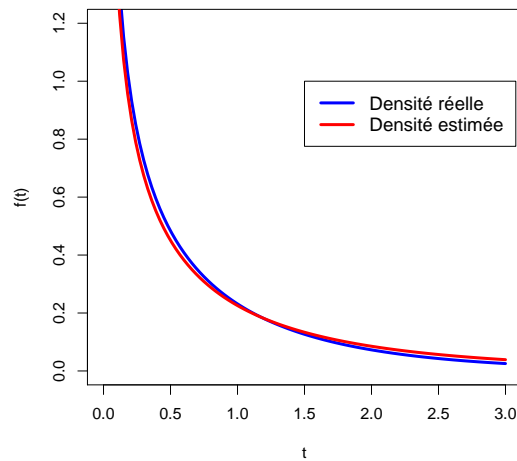


FIGURE 3.5 – Comparaison entre la vraie densité et celle estimée par Bayes.

3.5 Interprétation des résultats

La partie simulation a pour objectif de tester la convergence et l'efficacité de l'approche bayésienne dans l'estimation des paramètres α, τ, k, λ et φ . Nous avons choisi la moyenne a posteriori comme estimateur bayésien. Nous avons utilisé notre programme MCMC pour es-

timer les paramètres α, τ, k, λ et φ . Les résultats obtenus par cette méthode nous permet de constater que les estimations sont proches des valeurs attendues (vraie valeurs des paramètres). Les résultats montrent que l'erreur quadratique moyenne (EQM) diminue en augmentant la taille de l'échantillon n ; voir le tableaux (3.1). Les graphes montrent aussi que les résultats obtenus par Bayes sont satisfaisant; voir les figures (3.4) et (3.5).

Chapitre 4

Application sur des données réelles

4.1 Introduction

Nous allons appliquer le modèle KumGG sur trois jeux de données réels. Nous estimons les paramètres α, τ, k, λ et φ par l'approche bayésienne. Afin de réaliser une étude comparative avec le modèle exponentielle généralisé (EG) à deux paramètres α et λ , nous calculons les critères AIC et BIC donnés comme suit :

- Critère d'Information Akaike en anglais Akaike Information Criterion (AIC)

$$AIC = 2k - 2 \ln(L),$$

où k est le nombre des paramètres à estimer et L est la fonction de vraisemblance.

- Critère d'Information Bayésien en anglais Bayésien Information Criterion (BIC)

$$BIC = k \ln(n) - 2 \ln(L), \quad n \text{ est la taille de l'échantillon.}$$

4.2 Modèle exponentiel généralisé

Le modèle exponentiel généralisé (EG) à deux paramètres α et λ a été proposé et étudié par Kundu and Gupta [25].

La fonction de répartition F et la fonction de densité f de la variable aléatoire X qui suit une loi EG de paramètres α et λ sont données par :

$$F(x) = (1 - \exp(-\lambda x))^\alpha; \alpha, \lambda, x > 0$$

et

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

4.3 Première application

Nous considérons les données décrivant les temps entre les défaillances (milliers d'heures) des pompes à réacteurs, voir Suprawhardana et Sangadji [18] :

2.160, 0.746, 0.402, 0.954, 0.491, 6.560, 4.992, 0.347, 0.150, 0.358, 0.101, 1.359, 3.465, 1.060, 0.614, 1.921, 4.082, 0.199, 0.605, 0.273, 0.070, 0.062, 5.320.

Nous analysons les données en utilisant le modèle KumGG et le modèle EG. Nous calculons les estimateurs de Bayes (moyenne a posteriori) en utilisant le programme MCMC en utilisant le nombre total des itérations $N = 50000$, la période de chauffage $N_0 = 25000$ et les hyper-paramètres des lois a priori $a_i = b_i = 0.01, i = 1, 2, \dots, 5$ pour le modèle KumGG. En parallèle nous calculons les estimateurs de Bayes du modèle (EG) à l'aide de programme MCMC avec Gibbs en utilisant le même nombre total des itérations N , la même période de chauffage N_0 et les hyper-paramètres des lois a priori $a = b = c = d = 0$, voir le programme dans l'annexe. Les résultats obtenus des estimations bayésiennes des deux modèles sont présentés dans le tableau (4.1) et le graphe (4.1). Pour comparer les deux modèles, nous donnons les critères AIC et BIC.

Modèle	$\hat{\alpha}$	$\hat{\tau}$	\hat{k}	$\hat{\lambda}$	$\hat{\varphi}$	AIC	BIC
KumGG	4.9626	3.9978	0.0198	0.7314	0.4662	103,8573	109.5312
EG	1.9370	(-)	(-)	2.2355	(-)	121.0977	132.7752

TABLE 4.1 – Estimation des paramètres des modèles KumGG et EG avec AIC et BIC.

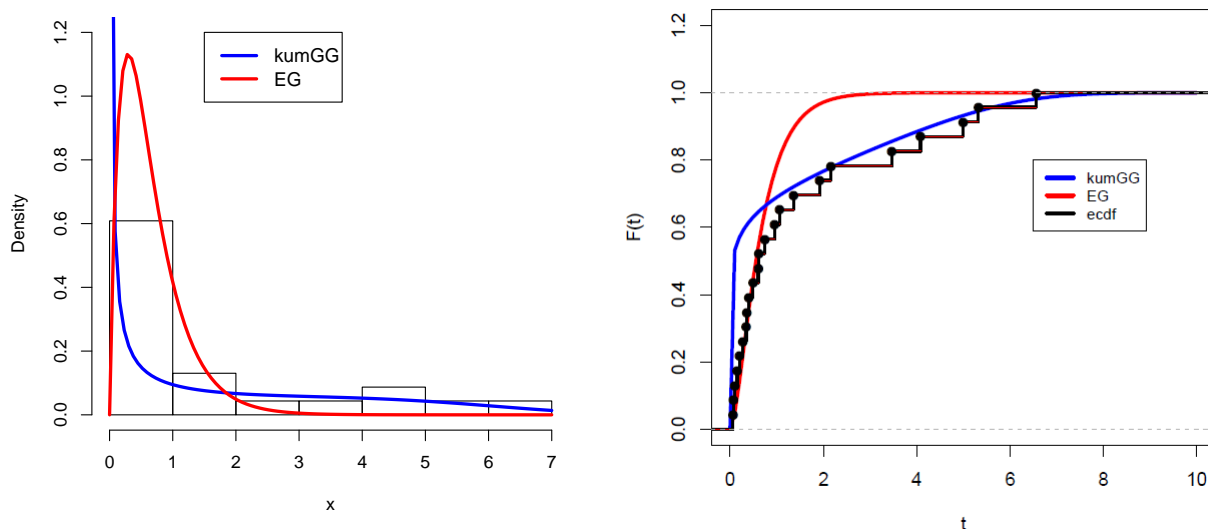


FIGURE 4.1 – [1] Fonctions de densités estimées des deux modèles KumGG et EG. [2] Fonction de répartition empirique et fonctions de répartition estimées des deux modèles KumGG et EG.

4.4 Deuxième application

La deuxième application décrit les durées de vie de 50 industriels, les appareils mettent le test de vie à l’heure zéro. Ces données ont été utilisées par Mudholkar et Srivastava [10] :

0.1, 0.2, 1, 1, 1, 1, 1, 2, 3, 6, 7, 11, 12, 18, 18, 18,18, 18, 21, 32, 36, 40, 45, 46, 47, 50,
55, 60, 63, 63, 67, 67, 67, 67, 72, 75, 79, 82, 82, 83, 84, 84, 84, 85, 85, 85, 85, 85, 86, 86

De la même manière, nous utilisons notre programme MCMC avec Gibbs pour estimer les paramètres du modèle KumGG et EG en utilisant ces données. Les résultats sont données dans le tableau (4.2) et le graphe (4.2).

Modèle	$\hat{\alpha}$	$\hat{\tau}$	\hat{k}	$\hat{\lambda}$	$\hat{\varphi}$	AIC	BIC
KumGG	84.0156	74.9973	0.0040	0.7085	0.5299	451.3275	457.0053
EG	0.7694	(-)	(-)	0.0183	(-)	483.9994	499.5623

TABLE 4.2 – Estimation des paramètres des deux modèles KumGG et EG avec AIC et BIC.

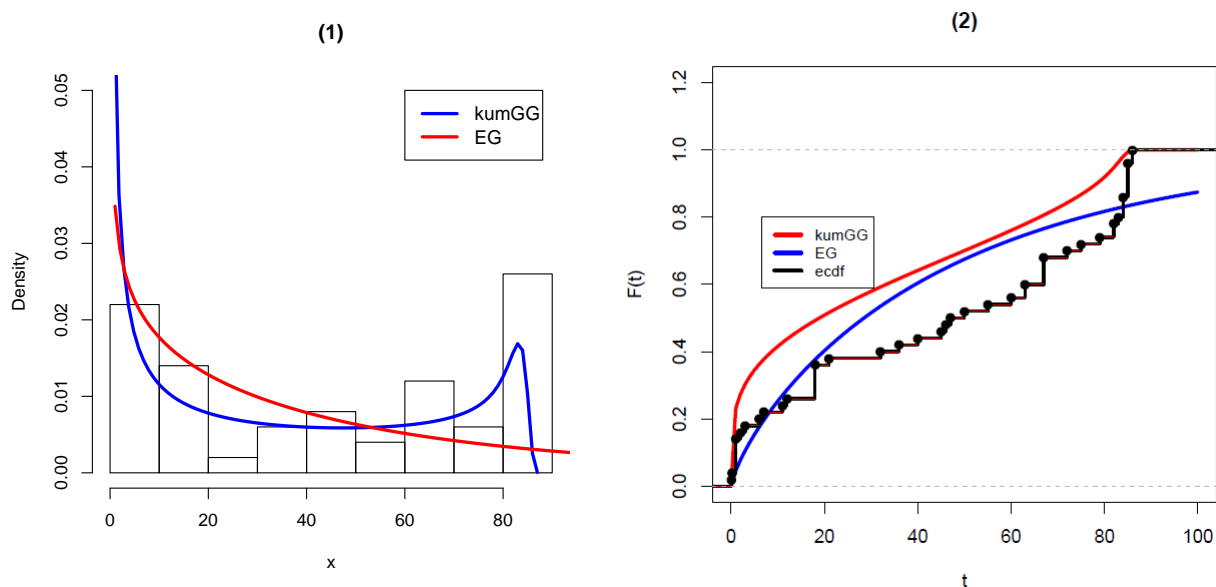


FIGURE 4.2 – [1] Fonctions de densités estimées des deux modèles KumGG et EG. [2] Fonction de répartition empirique et fonctions de répartition estimées des deux modèles KumGG et EG.

4.5 Troisième application

La troisième application porte sur les tests de l’endurance des roulements ‘à billes rigides.

Les données sont (voir Lawless [13] et Kundu and Gupta [4]) :

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.

De la même manière, nous utilisons notre programme MCMC avec Gibbs pour estimer les paramètres du modèle KumGG et EG en utilisant ces données. Les résultats sont données dans le tableau (4.3) et le graphe (4.3).

Modèle	$\hat{\alpha}$	$\hat{\tau}$	\hat{k}	$\hat{\lambda}$	$\hat{\varphi}$	AIC	BIC
KumGG	17.0501	0.6723	1.8801	1.9602	2.3202	258.6479	264.3254
EG	5.4122	(-)	(-)	0.0319	(-)	230.0022	241.6796

TABLE 4.3 – Estimation des paramètres des deux modèles KumGG et EG avec AIC et BIC.

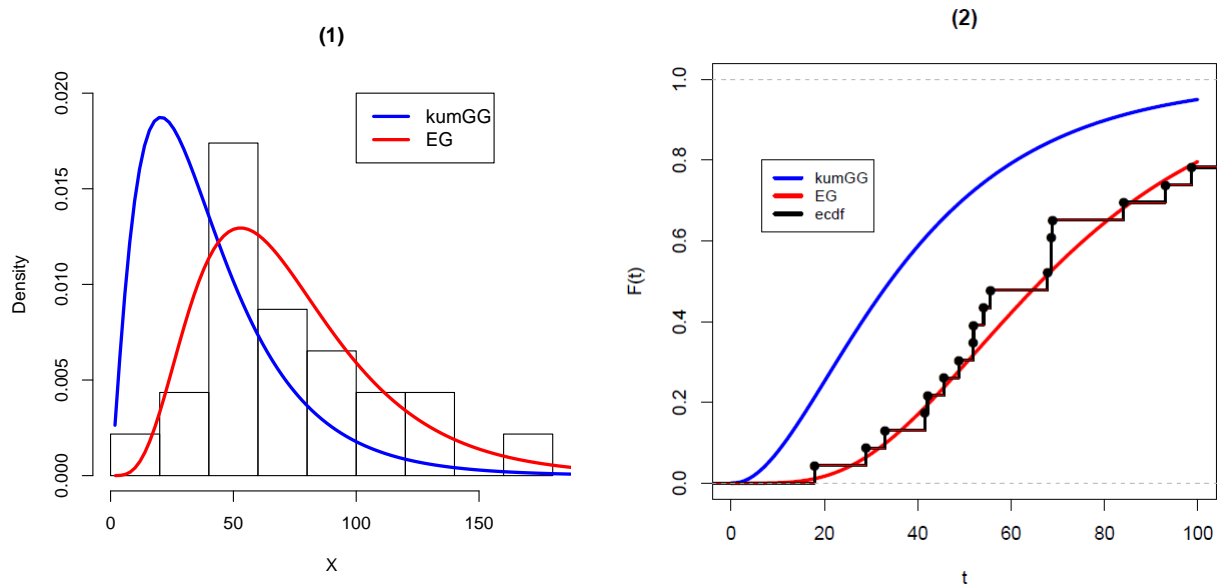


FIGURE 4.3 – [1] Fonctions de densités estimées des deux modèles KumGG et EG. [2] Fonction de répartition empirique et fonctions de répartition estimées des deux modèles KumGG et EG.

4.6 Discussion

Dans cette partie, nous avons analysé des données réelles en utilisant le modèle KumGG et le modèle EG, pour faire une comparaison, en estimant les paramètres de ces deux modèles par l'approche bayésienne. Les résultats donnés sous forme de tableaux et graphiques montrent que le modèle KumGG est très approprié pour la première et la deuxième application. Cependant, les résultats obtenus dans le cas de la troisième application par KumGG sont mauvais, et les meilleurs résultats sont obtenus en utilisant le modèle EG.

Conclusion générale

Ce travail a pour objet l'étude théorique et pratique du modèle Kumarswamy gamma généralisé (KumGG) avec cinq paramètres α, τ, k, λ et φ (KumGG($\alpha, \tau, k, \lambda, \varphi$)).

La première partie de ce travail a été consacrée au rappel sur le modèle KumGG. Nous avons présentés la forme de la fonction de répartition, de la densité de probabilité, de la fonction de survie et la fonction hasard. La méthode de simulation de la variable aléatoire T de loi KumGG($\alpha, \tau, k, \lambda, \varphi$) a été aussi présentée ainsi que leurs programmes implémentés sous logiciel **R**.

Dans la seconde partie, notre objectif était de présenter en détail les méthodes d'estimation des paramètres du modèle KumGG. Premièrement, nous avons rappelé le principe de la méthode du maximum de vraisemblance dans l'estimation de ce modèle. Dans un second temps, nous avons présenté par l'approche bayésienne. Dans cette approche, nous avons développé un algorithme MCMC pour estimer les paramètres car la loi a posteriori 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 KumGG en estimant les paramètres l'approche bayésienne. Une comparaison avec le modèle EG au sens d'AIC et BIC est présentée. Notons que les résultats avec la méthode du maximum de vraisemblance n'ont pas été présentés pour cause de divergence de cette approche.

Annexe

A.1 Développement en série entière de la fonction gamma incomplète

Nous donnons le développement en série entière de la fonction gamma incomplète. D'après Cordeiro et Castro [9], on peut écrire

$$\gamma_1 \left[k, \left(\frac{t}{\alpha} \right)^\tau \right]^{\lambda-1} = \sum_{m=0}^{\infty} S_m(\lambda) \gamma_1 \left[k, \left(\frac{t}{\alpha} \right)^\tau \right]^m \quad (1)$$

avec

$$S_m(\lambda) = \sum_{j=m}^{\infty} (-1)^{j+m} \binom{\lambda-1}{j} \binom{j}{m}. \quad (2)$$

Le développement en série entière de la fonction gamma incomplète est donnée par :

$$\gamma_1 \left[k, \left(\frac{t}{\alpha} \right)^\tau \right] = \frac{1}{\Gamma(k)} \left(\frac{t}{\alpha} \right)^{\tau k} \sum_{d=0}^{\infty} \left[- \left(\frac{t}{\alpha} \right)^\tau \right]^d \frac{1}{(k+d)d!}, \quad (3)$$

$$\left[\sum_{d=0}^{\infty} \left(\frac{t}{\alpha} \right)^{\tau d} \right]^q = \sum_{d=0}^{\infty} c_{q,d} \left(\frac{t}{\alpha} \right)^{\tau d}, \quad (4)$$

le coefficient $c_{q,d}$ pour $(d = 1, 2, \dots)$ satisfait la relation de récurrence suivante :

$$c_{q,d} = (da_0)^{-1} \sum_{p=1}^d (qp - d + p) a_p c_{q,d-p}, \quad (5)$$

$c_{q,0} = a_0^q$ et $a_p = \frac{(-1)^p}{[(k+p)p!]}$. Le coefficient $c_{q,d}$ peut être calculé en utilisant la suite suivante :

$c_{q;0}, \dots, c_{q;d-1}$, et peut être écrit explicitement en fonction des quantités a_0, \dots, a_d . D'après

l'équation (4) on obtient la forme suivante :

$$\gamma_1 \left[k, \left(\frac{t}{\alpha} \right)^\tau \right]^q = \frac{1}{\Gamma(k)^q} \left(\frac{t}{\alpha} \right)^{kq\tau} \sum_{d=0}^{\infty} c_{q,d} \left(\frac{t}{\alpha} \right)^{\tau d}, \quad (6)$$

avec le coefficient $c_{q;d}$ défini par l'équation (5).

A.2 Elements de la matrice d'information du Fisher

Les éléments de la matrice d'information du Fisher $\mathbf{J}(\theta)$ pour le vecteur des paramètres $\theta = (\alpha, \tau, k, \lambda, \varphi)$ sont donnés comme suit :

$$\begin{aligned} j_{\alpha,\alpha}(\theta) &= \frac{nk\tau}{\alpha^2} - \frac{\tau(1-\tau)}{\alpha^2} \sum_{i=1}^n u_i - \frac{\tau}{\alpha^2} \sum_{i=1}^n v_i s_i \left(-1 + \frac{\tau}{\gamma_1(k, u_i)} (\gamma_1(k, u_i)(-k + u_i) + v_i) \right) \\ &\quad + \frac{\lambda\tau(\varphi-1)}{\alpha^2} \sum_{i=1}^n v_i p_i \left(-1 + \frac{\tau}{\omega_i} (\omega_i(-k + u_i - v_i s_i) - \lambda v_i [\gamma_1(k, u_i)]^2) \right) \\ &\quad + \frac{\lambda\tau\varphi}{\alpha^2} \sum_{i=1}^n v_i p_i \left(-1 + \frac{\tau}{\omega_i} (\omega_i(-k + u_i - v_i s_i) - \lambda v_i [\gamma_1(k, u_i)]^2) \right), \end{aligned}$$

$$\begin{aligned} j_{\alpha,\tau}(\theta) &= \frac{-nk}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n u_i [1 + \log(u_i)] - \frac{1}{\alpha} \sum_{i=1}^n v_i s_i \left(1 + \frac{\log(u_i)}{\gamma_1(k, u_i)} (\gamma_1(k, u_i)(k - u_i) - v_i) \right) \\ &\quad + \frac{\lambda(\varphi-1)}{\alpha} \sum_{i=1}^n v_i p_i \left(1 + \frac{\log(u_i)}{\omega_i} (\omega_i(k - u_i + v_i p_i) + \lambda v_i [\gamma_1(k, u_i)]^2) \right), \end{aligned}$$

$$\begin{aligned} j_{\alpha,k}(\theta) &= \frac{-n\tau}{\alpha} - \frac{\tau}{\alpha} \sum_{i=1}^n v_i s_i \left[\log(u_i) - \frac{q_i}{\gamma_1(k, u_i)} \right] + \frac{\lambda\tau(\varphi-1)}{\alpha} \sum_{i=1}^n v_i p_i \\ &\quad \times \left(-\psi(k) + \frac{1}{\omega_i} (\omega_i(\log(u_i) + s_i[-\psi(k)\gamma_1(k, u_i) + q_i]) + (\lambda\omega_i p_i[-\psi(k)\gamma_1(k, u_i) + q_i])) \right), \end{aligned}$$

$$j_{\alpha,\lambda}(\theta) = \frac{-\tau}{\alpha} \sum_{i=1}^n \frac{v_i}{\gamma_1(k, u_i)} + \frac{\tau(\varphi-1)}{\alpha} \sum_{i=1}^n v_i p_i (1 + \lambda b_i (\omega_i + [\gamma_1(k, u_i)]^\lambda)),$$

$$j_{\alpha,\varphi}(\theta) = \frac{\lambda\tau}{\alpha} \sum_{i=1}^n v_i p_i, \quad J_{\tau,\varphi}(\theta) = -\frac{\lambda}{\tau} \sum_{i=1}^n v_i p_i \log(u_i)$$

$$\begin{aligned} j_{\tau,\tau}(\theta) &= \frac{-n}{\tau} - \frac{1}{\tau^2} \sum_{i=1}^n u_i [\log(u_i)]^2 + \frac{1}{\tau^2} \sum_{i=1}^n \frac{v_i s_i [\log(u_i)]^2}{\gamma_1(k, u_i)} [\gamma_1(k, u_i)(k - u_i) - v_i] \\ &\quad - \frac{\lambda(\varphi-1)}{\tau^2} \sum_{i=1}^n \frac{v_i p_i [\log(u_i)]^2}{\omega_i} (\omega_i(k - u_i + v_i s_i) + \lambda v_i [\gamma_1(k, u_i)]^2), \end{aligned}$$

$$j_{\tau,k}(\theta) = \frac{1}{\tau} \sum_{i=1}^n \log(u_i) - \frac{1}{\tau} \sum_{i=1}^n v_i s_i \log(u_i) \left[\log(u_i) - \frac{q_i}{\gamma_1(k, u_i)} \right] \\ - \frac{\lambda(\varphi - 1)}{\tau} \sum_{i=1}^n v_i p_i \log(u_i) \left(-\psi(k) + \log(u_i) \left[-\psi(k) + \frac{q_i}{\gamma_1(k, u_i)} \right] [\lambda - 1 + \lambda p_i \gamma_1(k, u_i)] \right),$$

$$j_{\tau,\lambda}(\theta) = \frac{1}{\tau} \sum_{i=1}^n \frac{v_i \log(u_i)}{\gamma_1(k, u_i)} - \frac{(\varphi - 1)}{\tau} \sum_{i=1}^n v_i p_i \log(u_i) [1 + b_i \lambda \omega (1 + p_i \gamma_1(k, u_i))],$$

$$j_{k,k}(\theta) = -n\lambda\psi'(k) + \frac{1}{\Gamma(k)} \sum_{i=1}^n s_i \left([\ddot{\gamma}(k, u_i)]_{kk} - \frac{([\dot{\gamma}(k, u_i)]_k)^2}{\gamma_1(k, u_i)} \right) \\ + n\lambda(\varphi - 1) \sum_{i=1}^n p_i \gamma_1(k, u_i) \left(\psi'(k) + \frac{\lambda\psi(k)}{\gamma_1(k, u_i)} \left[-\frac{\psi(k)\gamma_1(k, u_i)}{\omega_i} + \frac{q_i}{\omega_i} \right] \right) - \frac{\lambda(\varphi - 1)}{\Gamma(k)} \\ \times \sum_{i=1}^n p_i \left(-\psi(k)[\dot{\gamma}(k, u_i)]_k + \left(q_i(\lambda - 1) \left[-\psi(k)\gamma_1(k, u_i) + \frac{q_i}{\gamma_1(k, u_i)} \right] + \frac{[\ddot{\gamma}(k, u_i)]_{kk}}{\omega_i} \right) \right),$$

$$j_{k,\lambda}(\theta) = -n\psi(k) + \sum_{i=1}^n \frac{q_i}{\gamma_1(k, u_i)} + n\psi(k)(\varphi - 1) \sum_{i=1}^n p_i \gamma_1(k, u_i) (1 + b_i \lambda) - (\varphi - 1) \sum_{i=1}^n p_i q_i (1 + b_i \lambda),$$

$$j_{k,\varphi}(\theta) = n\lambda\psi(k) \sum_{i=1}^n p_i \gamma_1(k, u_i) - \lambda \sum_{i=1}^n p_i q_i,$$

$$j_{\lambda,\lambda}(\theta) = -\frac{n}{\lambda^2} - (\varphi - 1) \sum_{i=1}^n b_i [\gamma_1(k, u_i)]^\lambda \log[\gamma_1(k, u_i)],$$

$$j_{\lambda,\varphi}(\theta) = -\sum_{i=1}^n b_i [\gamma_1(k, u_i)]^\lambda,$$

$$j_{\varphi,\varphi}(\theta) = -\frac{n}{\varphi^2},$$

avec

$$[\dot{\gamma}(k, u_i)]_k = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} J(u_i, k + n - 1, 1); [\ddot{\gamma}(k, u_i)]_{kk} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} J(u_i, k + n - 1, 2),$$

$\psi'(\cdot)$ est la dérivée de la fonction digamma.

Les éléments $u_i, g_i, \omega_i, v_i, s_i, p_i, q_i$ et b_i sont définis dans le chapitre 2. La fonction $J(\cdot, \cdot, \cdot)$ est obtenue en calculant l'intégrale donnée par Prudnikov et al. [2].

$$J(a, p, 1) = \int_0^a x^p \log(x) dx = \frac{a^{p+1}}{(p+1)^2} [(p+1) \log(a) - 1].$$

et

$$J(a, p, 2) = \int_0^a x^p \log^2(x) dx = \frac{a^{p+1}}{(p+1)^3} [2 - (p+1) \log(a) (2 - (p+1) \log(a))].$$

A.3 Code source avec R

Les programmes informatiques ont été implémentés sous **R**. 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 T de loi KumGG $(\alpha, \tau, k, \lambda, \varphi)$. Ensuite, nous donnons les programmes **R** qui retournent les estimateurs de Bayes pour les paramètres α, τ, k, λ et φ .

```

=====
===== Fonction de densité =====
f <- function(x,alpha,taux,k,lamda,phi){
densité<-((lamda*phi*taux)/(alpha*gamma(k))*((x/alpha)^(taux*k-1))*exp(-(x/alpha)^taux)*
(pgamma(((x/alpha)^taux),k,1))^(lamda-1)*((1-(pgamma(((x/alpha)^taux),k,1))^(lamda))^(phi-1))
      }
=====
===== Fonction de survie =====
f <- function(x,alpha,taux,k,lamda,phi){
S<-(1-(pgamma(((x/alpha)^taux),k))^(lamda))^(phi)
      }
=====
===== Fonction hasard =====
f <- function(x,alpha,taux,k,lamda,phi){
densité<-((lamda*phi*taux)/(alpha*gamma(k))*((x/alpha)^(taux*k-1))*exp(-(x/alpha)^taux)*
      (pgamma(((x/alpha)^taux),k))^(lamda-1)*((1-(pgamma(((x/alpha)^taux),k))^(lamda))^(phi-1))
      }
=====
===== Simulation de x =====
simKumGG <- function(x,alpha,tau,k,lamda,phi){
IT <- function(x){
hit = pgamma((x/alpha)^tau,k,1)
u = runif(1)
uu =(1-(u^(1/phi)))^(1/lamda)
hit = uu-hit
hit
      }

```

```

val = uniroot(IT,c(0.0001,10),tol=0.001)
sim1 = val$root

}

simEch <-function(n){
x=0
for(i in 1:n){
x[i]=simKumGG(x,alpha,tau,k,lamda,phi)
}
x
}

##### Estimation des paramètres de modèle par la méthode de MCMC Gibbs #####

logposterior <- function(x,alpha,tau,k,lamda,phi){
n=length(x)
postalpha =(-n*tau*k)*log(alpha)-sum((x/alpha)^tau)+(lamda-1)*sum(log(pgamma((x/alpha)^tau,k,1)))
+(phi-1)* sum(log(1-(pgamma((x/alpha)^tau,k,1))^lamda))+dgamma(alpha,0.01,0.01,log=TRUE)
postalpha
}

logposteriortau <- function(x,alpha,tau,k,lamda,phi){
n=length(x)
posttau = n*(log(tau)-(tau*k)*log(alpha))-sum((x/alpha)^tau)+sum((tau*k)*log (x)+(lamda-1)*
log(pgamma((x/alpha)^tau,k,1)))+(phi-1)*
sum (log(1-(pgamma((x/alpha)^tau,k,1))^lamda))+dgamma(tau,0.01,0.01,log=TRUE)
posttau
}

logposteriorK <- function(x,alpha,tau,k,lamda,phi){
n=length(x)
postK = (-n)*((tau*k)*log(alpha)+log(gamma(k)))+sum((tau*k)*log (x)+(lamda-1)*
log(pgamma((x/alpha)^tau,k,1)))+(phi-1)*
sum (log(1-(pgamma((x/alpha)^tau,k,1))^lamda))+ dgamma(k,0.01,0.01,log=TRUE)
postK
}

logposteriorlamda=function(x,alpha,tau,k,lamda,phi){
n=length(x)
postlamda = n*log(lamda)+(lamda)*sum(log(pgamma((x/alpha)^tau,k,1)))+(phi-1)*
log(1-(pgamma((x/alpha)^tau,k,1))^lamda)+ dgamma(lamda,0.01,0.01,log=TRUE)
}

```



```

postlamda                                     }

logposteriorphi <- function(x,alpha,tau,k,lamda,phi){
  n=length(x)
  postlphi = n*log(phi)+(phi)*sum(1-(pgamma((x/alpha)^tau,k,1))^lamda)+
            dgamma(phi,0.01,0.01,log=TRUE)
  postlphi                                     }

MCMC_Gibbs <- function(x,alpha0,tau0,k0,lamda0,phi0,N,NO){
n=length(x)
alpha=tau=k=lamda=phi=0
s=a=R=z=d=0
r1=r2=r3=r4=r5=0
m=1
alpha[1]<-alpha0 ### Initialisation des paramètres
tau[1]  =tau0
k[1]    =k0
lamda[1]=lamda0
phi[1]  = phi0
while (m<=N){
alphac=rnorm(1,alpha[m],0.002) ### loi de proposition (candidate)
tauc=rnorm(1,tau[m],0.002)
kc=rnorm(1,k[m],0.002)
lambdac=rnorm(1,lamda[m],0.002)
phic=rnorm(1,phi[m],0.002)
alphac=abs(alphac)
tauc=abs(tauc)
kc=abs(kc)
lambdac=abs(lambdac)
phic=abs(phic)
r1[t]=min(0,logposterior(x,alphac,tau[m],k[m],lamda[m],phi[m])-
          logposterior(x,alpha[m],tau[m],k[m],lamda[m],phi[m]))
u=log(runif(1))
  if (u<r1[m]){

```

```

alpha[m+1]=alphac }
  else{
alpha[m+1]=alpha[m] }
r2[m]=min(0,logposteriortau(x,alpha[m+1],tauc,k[m],lamda[m],phi[m])-
  logposteriortau(x,alpha[m+1],tau[m],k[m],lamda[m],phi[m]))
  if (u<r2[m]){
tau[m+1]=tauc }
  else{
tau[m+1]=tau[m]}
r3[m]=min(0,logposteriortK(x,alpha[m+1],tau[m+1],kc,lamda[m],phi[m])-
  logposteriortK(x,alpha[m+1],tau[m+1],k[m],lamda[m],phi[m]))
  if (u<r3[m]){
  k[m+1]=kc }
  else{
k[m+1]=k[m]}

r4[m]=min(0,logposteriorlambda(x,alpha[m+1],tau[m+1],k[m+1],lambdac,phi[m])-
  logposteriorlambda(x,alpha[m+1],tau[m+1],k[m+1],lamda[m],phi[m]))
  if (u<r4[m]) {
  lamda[m+1]=lambdac }
  else{
lamda[m+1]=lamda[m]}
r5[m]=min(0,logposteriorphi(x,alpha[m+1],tau[m+1],k[m+1],lamda[m+1],phic)-
  logposteriorphi(x,alpha[m+1],tau[m+1],k[m+1],lamda[m+1],phi[m]))
  if (u<r5[m]){
phi[m+1]=phic}
  else{
phi[m+1]=phi[m]
  }
m=m+1}
print('alpha-chapeau =')
print(alpha)
op=par(mfrow=c(2,3))
hist(alpha,col="yellow") ### récupérer les histogrammes des chaines de Markov

```

```
print('tau-chapeau =')
print(tau)
hist(tau,col="violet")
print('K-chapeau=')
print(k)
hist(k,col="red")
print('lambda-chapeau =')
print(lamda)
hist(lamda,col="blue")
print('phi-chapeau =')
print(phi)
hist(phi,col="green")
print('EQM(alpha) =')
q=(1/N)*sum((alpha-1.1)^2) ### calculer EQM
to=(1/N)*sum((tau-1.3)^2)
kk=(1/N)*sum((k-0.9)^2)
ll=(1/N)*sum((lamda-3)^2)
phh=(1/N)*sum((phi-1.3)^2)
print(q)
print('EQM(tau) =')
print(to)
print('EQM(k) =')
print(kk)
print('EQM(lambda) =')
print(ll)
print('EQM(phi) =')
print(phh)
for(m in N0+1:N){
s=s+alpha[m]
a=a+k[m]
R=R+lamda[m]
d=d+phi[m]
z=z+tau[m]}
alpha_Bayes=(1/(N-N0)) * s
```

```

tau_bayes= (1/(N-N0)) * z
k_bayes= (1/(N-N0)) * a
lambda_bayes= (1/(N-N0)) * R
phi_bayes= (1/(N-N0)) * d

print('alpha-chapeau-Bayes est:') ### l'affichage des estimation
print(alpha_Bayes)
print('tau-chapeau-Bayes est:')
print(tau_bayes)
print('K-chapeau-Bayes est:')
print(k_bayes)
print('lambda-chapeau-Bayes est:')
print(lambda_bayes)
print('phi-chapeau-Bayes est:')
print(phi_bayes)  }

```

=====

La fonction de répartition et la fonction de densité de modèle $EG(\alpha, \lambda)$

=====

```

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
}

```

Estimation des paramètres de modèle EG par la méthode de MCMC Gibbs

```

posterior=function(lambda,alpha,X){
a=0
b=0
e=0
d=0

```

```

n=length(X)
postlambda=(lambda^(n+b-1))*exp(-lambda*(a+sum(X)))* prod((1-exp(-lambda*X))^(alpha-1))
postlambda
}

MCMC_Gibbs<-function(N,N0,lambda0,alpha0,X){
e=0
d=0
n=length(X)
lambda=0
alpha=0
ta=0
r=0
t=1
alpha[1] <- alpha0
lambda[1]<- lambda0

while (t<=N){
print(t)
lambdac=rnorm(1,lambda[t],0.01)
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}

s=R=0
print(alpha)
for(t in N0+1:N){
s=s+alpha[t]
R=R+lambda[t]
}
}

```

```
    }  
    alpha_Bayes=(1/(N-N0)) * s  
    lambda_bayes= (1/(N-N0)) * R  
    print('alpha-chapeau-Bayes est:')  
    print(alpha_Bayes)  
    print('lambda-chapeau-Bayes est:')  
    print(lambda_bayes)  
}
```

Bibliographie

- [1] A. Bekker., J.J. Roux. Reliability characteristics of the Maxwell distribution : a Bayes estimation study. *Communications in Statistics—Theory and Methods* 34 :2169–2178, 2005.
- [2] A.P. Prudnikov., Y.A. Brychkov., O.I. Marichev. *Integrals and Series*, vol. 1, Gordon and Breach Science Publishers, Amsterdam, 1986.
- [3] C. Kleiber., S. Kotz. *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley-Interscience, ISBN 0-471-15064-9 , 2003.
- [4] D. Kundu and R D. Gupta. Generalized exponential distribution : Bayesian estimations. *Computational Statistics and Data Analysis*, 52 :1873-1883, 2008.
- [5] E.M.M. Ortega., V.G. Cancho., G.A. Paula. Generalized log-gamma regression models with cure fraction, *Lifetime Data Analysis* 15 :79–106, 2009.
- [6] E.W. Stacy. A generalization of the gamma distribution, *Annals of Mathematical Statistics* 33 :1187–1192, 1962.
- [7] E.W. Stacy., G.A. Mihram. Parameter estimation for a generalized gamma distribution. *Technometrics* 7 :349–358, 1965.
- [8] G.M. Cordeiro., E.M.M. Ortega., G.O. Silva. The exponentiated generalized gamma distribution with application to lifetime data, *Journal of Statistical Computation and Simulation*, 2010.
- [9] G.M. Cordeiro., M. de Castro. A new family of generalized distributions, *Journal of Statistical Computation and Simulation*, 2011.
- [10] G.S. Mudholkar., D.K. Srivastava. Exponentiated Weibull family for analyzing bathtub failure-real data, *IEEE Transaction on Reliability* 42 :299–302, 1993.

- [11] H. Hirose. Maximum likelihood parameter estimation by model augmentation with application to the extended four-parameter generalized gamma distribution (on line) available on <http://search2.libi.kyutech.ac.jp/TR/pdf/csse1.pdf> consulted the 12/10/04.
- [12] J.F. Lawless. Inference in the generalized gamma and log-gamma distributions, *Technometrics* 22 :409–419, 1980.
- [13] J.F. Lawless. *Statistical Models and Methods for Lifetime Data*, John Wiley, New York, 2003.
- [14] J. Malhotra., A.K. Sharma., R.S. Kaler. On the performance analysis of wireless receiver using generalized-gamma fading model, *Annals of Telecommunications* 64 :147–153, 2009.
- [15] K. Cooray., M.M.A. Ananda. A generalization of the half-normal distribution with applications to lifetime data. *Communications in Statistics—Theory and Methods* 37 :1323–1337, 2008.
- [16] M.C. Jones. Kumaraswamy’s distribution : a beta-type distribution with some tractability advantages, *Statistical Methodology* 6 :70–81, 2009.
- [17] M.M. Ali ., J. Woo., S. Nadarajah. Generalized gamma variables with drought application, *Journal of the Korean Statistical Society* 37 :37–45, 2008.
- [18] M.S. Suprawhardana and P. Sangadji. Total time on test plot analysis for mechanical components of the RSG-GAS reactor. *Atom Indones*, 25(2), 155-61, 1999.
- [19] 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.
- [20] O. Gomes., C. Combes., A. Dussauchoy. Parameter estimation of the generalized gamma distribution, *Mathematics and Computers in Simulation* 79 :955–963, 2008.
- [21] de Pascoa., M.A.R., Ortega., E.M.M. and Cordeiro. G.M. The Kumaraswamy generalized gamma distribution with application in survival analysis, *Statistical Methodology*, 8 (5) :411-433, 2011.
- [22] P. M. Shankar., V.A. Dumane., J.M. Reid., V. Genis., F. Forsberg., C.W. Piccoli., B.B. Goldberg. Classification of ultrasonic b-mode images of breast masses using Nakagami distribution. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 48 :569–580, 2001.
- [23] P. Kumaraswamy. A generalized probability density function for double-bounded random processes, *Journal of Hydrology*, 46 :79-88, 1980.
- [24] R.D. Gupta., D. Kundu. Exponentiated exponential distribution : an alternative to gamma and Weibull distributions. *Biometrical Journal* 43 :117–130, 2001.

- [25] R.D. Gupta., D. Kundu. Generalized exponential distributions, Australian and New Zealand Journal of Statistics 41 :173–188, 1999.
- [26] S. Nadarajah., A.K. Gupta. A generalized gamma distribution with application to drought data, Mathematics and Computers in Simulation 74 :1–7, 2007.
- [27] W. Hastings. Monte carlo sampling methods using markov chains and their applications. Biometrika, 75 :97–109, 1970.

Résumé

Dans ce travail nous nous sommes intéressés au modèle Kumaraswamy gamma généralisé (KumGG) à cinq paramètres proposé par de Pascoa et al. Ce modèle s'obtient, en combinant le modèle gamma généralisé (GG) introduit par Stacy et le modèle Kumarswamy (Kum) introduit par Kumarswamy. Ce nouveau modèle englobe plusieurs sous-modèles spéciaux et trouve ses applications dans divers domaines comme l'hydrologie, l'économie, la biologie, les sciences de l'environnement, etc. Notre travail présenté ici est basé sur l'article de *Marcelino A.R. de Pascoa., Edwin M.M. Ortega., Gauss M. Cordeiro. The Kumaraswamy generalized gamma distribution with application in survival analysis. Statistical Methodology, 5(8) :411–433, 2011.*

Le premier objectif de ce travail est de présenter en détail le modèle KumGG. Le second objectif est de faire le point sur les méthodes d'estimation de ce modèle, afin d'analyser des données simulées et réelles par l'approche bayésienne, en utilisant les techniques de Monte Carlo par Chaîne de Markov (MCMC). Une comparaison au sens d'informations AIC et BIC avec le modèle exponentiel généralisé (EG) est présentée.

Abstract

In this work we are interested in the Kumaraswamy generalized gamma model (KumGG) with five parameters proposed by de Pascoa et al. This model is obtained by combining the generalized gamma model (GG) introduced by Stacy and the Kumarswamy model (Kum) introduced by Kumarswamy. This new model Encompasses several special sub-models and finds its applications in various fields such as Hydrology, economics, biology, environmental sciences, etc. Our work presented here is based on the article of *Marcelino A.R. de Pascoa., Edwin M.M. Ortega., Gauss M. Cordeiro. The Kumaraswamy Generalized gamma distribution with application in survival analysis. Statistical Methodology, 5 (8) : 411-433, 2011.* The first objective of this work is to present the KumGG model in detail. The second objective is to give the estimation methods of this model. Simulated and real data are analyzed by the Bayesian approach, using the techniques of Monté Carlo by Markov Chain (MCMC). A comparison in the sense of AIC and BIC with the generalized exponential model is presented. (EG).