

Generalized exponential distribution: A Bayesian approach using MCMC methods

Jorge Alberto Achcar^{a,b}, Fernando Antônio Moala^{a*} and Juliana Boleta^b

^aDepartamento de Matemática, Estatística e Computação, UNESP, Universidade Estadual Paulista, Presidente Prudente, SP, Brazil

^bDepartamento de Medicina Social, FMRP, Universidade de São Paulo, Avenida Bandeirantes, 3900, CEP: 14048-900, Ribeirão Preto, SP, Brasil

CHRONICLE

Article history:

Received July 6 2014

Received in Revised Format

August 5 2014

Accepted August 6 2014

Available online

August 6 2014

Keywords:

Generalized exponential
distribution

Non-informative priors

Bayesian analysis

MCMC methods

ABSTRACT

The generalized exponential distribution could be a good option to analyse lifetime data, as an alternative for the use of standard existing lifetime distributions as exponential, Weibull or gamma distributions. Assuming different non-informative prior distributions for the parameters of the model, we introduce a Bayesian analysis using Markov Chain Monte Carlo (MCMC) methods. Some numerical illustrations considering simulated and real lifetime data are presented to illustrate the proposed methodology, especially the effects of different priors on the posterior summaries of interest.

© 2015 Growing Science Ltd. All rights reserved

1. Introduction

A generalized exponential distribution (see Gupta & Kundu, 1999) can be a good alternative for the use of the popular gamma or Weibull distributions to analyse lifetime data (see also, Raqab, 2002; Raqab & Ahsanullah, 2001; Zheng, 2002; Sarhan, 2007; Gupta & Kundu, 2001, 2007). The generalized exponential distribution with two parameters has density given by,

$$f(t; \alpha, \lambda) = \alpha\lambda[1 - \exp(-\lambda t)]^{\alpha-1} \exp(-\lambda t), \quad (1)$$

where $t > 0$; $\alpha > 0$ and $\lambda > 0$ are respectively, shape and scale parameters. Let us denote this model as $GE(\alpha, \lambda)$. The density function given in Eq. (1) has great flexibility of fitting depending on the shape parameter α : if $\alpha < 1$, we have a decreasing function and if $\alpha > 1$, we have a unimodal function with mode given by $\lambda^{-1} \log \alpha$. Observe that if $\alpha = 1$, we have an exponential distribution with parameter λ . The survival and hazard function associated with Eq. (1) are given Eq. (2) and Eq. (3), respectively, as follows,

$$S(t; \alpha, \lambda) = P(T > t) = 1 - [1 - \exp(-\lambda t)]^\alpha, \quad (2)$$

and

* Corresponding author.

E-mail: femoala@fct.unesp.br (F. A. Moala)

$$h(t; \alpha, \lambda) = \frac{f(t; \alpha, \lambda)}{S(t; \alpha, \lambda)} = \frac{\alpha \lambda [1 - \exp(-\lambda t)]^{\alpha-1} \exp(-\lambda t)}{1 - [1 - \exp(-\lambda t)]^\alpha}. \quad (3)$$

Observe that the hazard function $h(t; \alpha, \lambda)$ has a non-decreasing trend from 0 to λ when $\alpha > 1$; a non-increasing trend from ∞ to 1 when $\alpha < 1$ and constant with $\alpha = 1$. This behavior of the hazard function given in Eq. (3) is similar to the behavior of the hazard function of a gamma distribution. Also observe that the median lifetime obtained from $S(t; \alpha, \lambda) = 1/2$ is given by,

$$t_{\text{med}} = -\frac{1}{\lambda} \log \left(1 - \left(\frac{1}{2} \right)^{1/\alpha} \right). \quad (4)$$

The moment generating function for a random variable T with a generalized exponential distribution and density of Eq. (1) is as follows, (see Gupta & Kundu, 2008)

$$M(s) = E(e^{sT}) = \frac{\Gamma(\alpha + 1) \Gamma\left(1 - \frac{s}{\lambda}\right)}{\Gamma\left(\alpha - \frac{s}{\lambda} + 1\right)}. \quad (5)$$

From Eq. (5), we find all moments of interest. The mean and variance are given, respectively, by

$$E(T) = \frac{1}{\lambda} [\psi(\alpha + 1) - \psi(1)], \quad \text{var}(T) = \frac{1}{\lambda^2} [\psi'(1) - \psi'(\alpha + 1)], \quad (6)$$

where $\psi(\cdot)$ is a digamma function given by $\psi(x) = \frac{d}{dx} \log(\Gamma(x)) = \frac{\Gamma'(x)}{\Gamma(x)}$ and $\Gamma(x)$ is a gamma function. In this paper, we develop a Bayesian analysis for the generalized exponential distribution using Markov Chain Monte Carlo (MCMC) methods (see for example, Gelfand & Smith, 1990; Chib & Greenberg, 1995) to obtain the posterior summaries of interest.

The paper is organized as follows: in Section 2, we introduce the likelihood function; in Section 3, we present a Bayesian analysis considering different non-informative priors for the parameters; in Section 4, we present inference for the survival function at a specified time; in Section 5, we introduce two numerical illustrations; finally, in Section 6, we present some conclusions.

2. The Likelihood Function

Suppose we have identically distributed lifetimes $\mathbf{t} = (t_1, \dots, t_n)'$ from a $GE(\alpha, \lambda)$ distribution. The likelihood function in the parameters α and λ , based on \mathbf{t} is then

$$L(\alpha, \lambda | \mathbf{t}) = \alpha^n \lambda^n \left[\prod_{i=1}^n [1 - \exp(-\lambda t_i)]^{\alpha-1} \right] \exp\left(-\lambda \sum_{i=1}^n t_i\right). \quad (7)$$

The logarithm of the likelihood function given in Eq. (7) is given by,

$$l(\alpha, \lambda) = \log L(\alpha, \lambda | \mathbf{t}) = n \log(\alpha) + n \log(\lambda) - \lambda \sum_{i=1}^n t_i + (\alpha - 1) \sum_{i=1}^n \log[1 - e^{-\lambda t_i}]. \quad (8)$$

The maximum likelihood estimators (MLE) for α and λ are obtained from $\partial l / \partial \alpha = 0$ and $\partial l / \partial \lambda = 0$, where,

$$\frac{\partial}{\partial \alpha} l(\alpha, \lambda) = \frac{n}{\alpha} + \sum_{i=1}^n \log[1 - e^{-\lambda t_i}] = 0 \quad (9)$$

$$\frac{\partial}{\partial \lambda} l(\alpha, \lambda) = \frac{n}{\lambda} - \sum_{i=1}^n t_i + (\alpha - 1) \sum_{i=1}^n \frac{t_i e^{-\lambda t_i}}{1 - e^{-\lambda t_i}} = 0$$

From Eq. (9), we find the MLE for α given by,

$$\hat{\alpha} = -\frac{n}{\sum_{i=1}^n \log[1 - e^{-\hat{\lambda}t_i}]} \tag{10}$$

The MLE for λ is obtained by solving the nonlinear equation,

$$\frac{n}{\hat{\lambda}} - n\bar{t} + (\hat{\alpha} - 1) \sum_{i=1}^n \frac{t_i e^{-\hat{\lambda}t_i}}{1 - e^{-\hat{\lambda}t_i}} = 0 \tag{11}$$

where $\hat{\alpha}$ is given by Eq. (10) and $n\bar{t} = \sum_{i=1}^n t_i$. Observe that we need to use an iterative method to find the MLE $\hat{\lambda}$ from Eq. (11). The second derivatives of $l(\alpha, \lambda)$ are given, respectively, by

$$\begin{aligned} \frac{\partial^2 l}{\partial \alpha^2} &= -\frac{n}{\alpha^2} \\ \frac{\partial^2 l}{\partial \lambda^2} &= -\frac{n}{\lambda^2} - (\alpha - 1) \sum_{i=1}^n \frac{t_i^2 e^{-\lambda t_i}}{(1 - e^{-\lambda t_i})^2} \\ \frac{\partial^2 l}{\partial \alpha \partial \lambda} &= \sum_{i=1}^n \frac{t_i e^{-\lambda t_i}}{1 - e^{-\lambda t_i}} \end{aligned} \tag{12}$$

Hypotheses tests and confidence intervals for α and λ can be obtained using the asymptotical normal distribution for $\hat{\alpha}$ and $\hat{\lambda}$, that is

$$(\hat{\alpha}, \hat{\lambda}) \sim N((\alpha, \lambda), I_0^{-1}), \tag{13}$$

where I_0 is the observed Fisher information matrix given by,

$$I_0 = \begin{pmatrix} -\frac{\partial^2 l}{\partial \alpha^2} & -\frac{\partial^2 l}{\partial \alpha \partial \lambda} \\ -\frac{\partial^2 l}{\partial \alpha \partial \lambda} & -\frac{\partial^2 l}{\partial \lambda^2} \end{pmatrix} \tag{14}$$

3. A Bayesian Analysis

For a Bayesian analysis of the GE(α, λ) distribution, we assume different prior distributions for α and λ . The Jeffreys non-informative prior (see for example, Box & Tiao, 1973) for α and λ is given by

$$\Pi_1(\alpha, \lambda) \propto [\det I(\alpha, \lambda)]^{1/2}, \tag{15}$$

where $I(\alpha, \lambda)$ is the Fisher information matrix given by,

$$I(\alpha, \lambda) = \begin{pmatrix} E\left(-\frac{\partial^2 l}{\partial \alpha^2}\right) & E\left(-\frac{\partial^2 l}{\partial \alpha \partial \lambda}\right) \\ E\left(-\frac{\partial^2 l}{\partial \alpha \partial \lambda}\right) & E\left(-\frac{\partial^2 l}{\partial \lambda^2}\right) \end{pmatrix} \tag{16}$$

where,

$$\begin{aligned} E\left[-\frac{\partial^2 l}{\partial \alpha^2}\right] &= \frac{n}{\alpha^2}, & E\left[-\frac{\partial^2 l}{\partial \lambda \partial \alpha}\right] &= \frac{n}{\lambda} \left\{ [\psi(\alpha+1) - \psi(1)] - \frac{\alpha}{\alpha-1} [\psi(\alpha) - \psi(1)] \right\} \\ E\left[-\frac{\partial^2 l}{\partial \lambda^2}\right] &= \frac{n}{\lambda^2} \left\{ 1 + \frac{\alpha(\alpha-1)}{\alpha-2} \left[[\psi'(1) - \psi'(\alpha-1)] + [\psi(\alpha-1) - \psi(1)]^2 \right] - \alpha \left[\psi'(1) - \psi'(\alpha) + [\psi(\alpha) - \psi(1)]^2 \right] \right\} \end{aligned} \tag{17}$$

(see Gupta & Kundu, 1999). Let us denote the prior Eq. (15) as “Jeffreys1”. A possible simplification is to consider a non-informative prior from $\Pi(\alpha, \lambda) = \Pi(\lambda|\alpha)\Pi_0(\alpha)$. Using the Jeffreys’rule, we have,

$$\pi_2(\alpha, \lambda) = \left[E\left(-\frac{\partial^2 l}{\partial \lambda^2}\right) \right]^{-\frac{1}{2}} \pi_0(\alpha) \quad (18)$$

where $E\left(-\frac{\partial^2 l}{\partial \lambda^2}\right)$ is given in Eq. (17) and $\pi_0(\alpha)$ is a non-informative prior given by $\pi_0(\alpha) \propto 1/\alpha$, $\alpha > 0$. In this way,

$$\pi_2(\alpha, \lambda) = \frac{1}{\alpha\lambda} A^{1/2}(\alpha) \quad (19)$$

$$\text{where } A(\alpha) = 1 + \frac{\alpha(\alpha-1)}{\alpha-2} [\psi'(1) - \psi'(\alpha-1)] + [\psi(\alpha-1) - \psi(1)]^2 - \alpha[\psi'(1) - \psi'(\alpha) + [\psi(\alpha) - \psi(1)]^2].$$

Let us denote the prior Eq. (19) as “Jeffreys2”. A third non-informative prior distribution is assumed considering independence between α and λ , that is,

$$\pi_3(\alpha, \lambda) \propto \frac{1}{\alpha\lambda}, \quad (20)$$

where $\alpha > 0$ and $\lambda > 0$. Let us denote the prior (20) as “Jeffreys3”. Assuming dependence between the random quantities α and λ , we could assume a bivariate prior distribution for α and λ derived from copula functions (see for example, Nelsen, 1999; Trivedi & Zimmer, 2007). A special case is given by the Farlie-Gumbel-Morgenstern Copula (see Morgenstern, 1956) given by,

$$c(u, v) = uv[1 + \delta(1-u)(1-v)], \quad (21)$$

where $u = F_1(\alpha)$ (marginal distribution function for α) and $v = F_2(\lambda)$ (marginal distribution function for λ). The joint distribution function for α and λ is given (from (21)) by,

$$F(\alpha, \lambda) = c(F_1(\alpha), F_2(\lambda)) = F_1(\alpha)F_2(\lambda)[1 + \delta(1-F_1(\alpha))(1-F_2(\lambda))], \quad (22)$$

where the parameter δ is associated to the dependence between α and λ . If $\delta = 0$, we have independence between α and λ . The joint prior density for α and λ , obtained from $\partial^2 F(\alpha, \lambda)/\partial\alpha\partial\lambda$, is given by,

$$\Pi_4(\alpha, \lambda|\delta) = f_1(\alpha)f_2(\lambda) + \delta f_1(\alpha)f_2(\lambda)[1 - 2F_1(\alpha)][1 - F_2(\lambda)], \quad (23)$$

where $f_1(\alpha)$ and $f_2(\lambda)$ are the marginal densities for α and λ . Assuming marginal exponential distributions with known hyperparameters a_1 and a_2 , we have, conditional on the hyperparameter δ ,

$$\pi_4(\alpha, \lambda | \delta) \propto \frac{1}{a_1 a_2} \exp\left(-\frac{\alpha}{a_1} - \frac{\lambda}{a_2}\right) \left[1 + \delta(2e^{-\frac{\alpha}{a_1}} - 1)(2e^{-\frac{\lambda}{a_2}} - 1) \right]. \quad (24)$$

Let us denote Eq. (24) as “Farlie-Gumbel“ or “Copula prior”. The specification of Eq. (24) must be completed by a prior distribution for δ , $\Pi(\delta)$. A suggestion for this prior could be a uniform $U[-1, 1]$ distribution on the interval $[-1, 1]$. Other prior specifications also could be considered, as independent informative Gamma distributions, that is,

$$\Pi_\alpha(\alpha) \sim \text{Gamma}(a_\alpha, b_\alpha) \quad \Pi_\lambda(\lambda) \sim \text{Gamma}(a_\lambda, b_\lambda) \quad , \quad (25)$$

where a_α , b_α , a_λ and b_λ are known hyperparameters and $\text{Gamma}(a,b)$ denotes a gamma distribution with mean a/b and variance a/b^2 .

3.1 The joint posterior distribution for α and λ assuming the Farlie-Gumbel-Morgenstern prior

Assuming the ‘‘Farlie-Gumbel-Morgenstern’’ prior $\Pi_4(\alpha, \lambda|\delta)$ introduced in Eq. (24), the statistical model and prior model together form an ordered structure in which the distribution of the data is written conditionally on the parameters (α, λ) as $f(t|\alpha, \lambda)$; the distribution of (α, λ) is written conditionally on the hyperparameter δ as $\Pi_4(\alpha, \lambda|\delta)$ and is completed by the distribution of δ , $\Pi(\delta)$. The full joint distribution of all random quantities in the problem is hierarchically written as,

$$f(t|\alpha, \lambda, \delta) = f(t|\alpha, \lambda)\Pi_4(\alpha, \lambda|\delta)\Pi(\delta) \tag{26}$$

In a three-level hierarchy with the form Eq. (26), inference about (α, λ) and δ is simply obtained through their joint posterior distribution,

$$\Pi(\alpha, \lambda, \delta|t) \propto L(\alpha, \lambda|t)\Pi_4(\alpha, \lambda|\delta)\Pi(\delta), \tag{27}$$

and inference about (α, λ) and δ are given by their marginal posteriors. Therefore, the joint posterior for the GE(α, λ) distribution parameters (α, λ) and the hiperparameter δ is given by

$$\Pi(\alpha, \lambda, \delta | t) \propto \frac{\alpha^n \lambda^n}{a_1 a_2} \left[\prod_{i=1}^n [1 - \exp(-\lambda t_i)]^{\alpha-1} \right] \exp\left(-\frac{\alpha}{a_1} - \lambda \left(\frac{1}{a_2} + \sum_{i=1}^n t_i\right)\right) \left[1 + \delta(2e^{-\frac{\alpha}{a_1}} - 1)(2e^{-\frac{\lambda}{a_2}} - 1) \right] \pi(\delta) \tag{28}$$

The next step is to specify a prior distribution for the parameter δ and a convenient prior proposed by Box and Tiao (1973) for correlation parameter in bivariate normal data is given by,

$$\Pi(\delta) \propto (1 - \delta^2)^c, \tag{29}$$

for an appropriated choice of ‘‘c’’. If we do not have any information from previous studies, a common choice is $c = 0$, that is, $\pi(\delta) \propto 1$. We decided to use $c = -1/2$, based on MCMC convergence rate and mathematics convenience, providing the joint posterior

$$\Pi(\alpha, \lambda, \delta | t) \propto \alpha^n \lambda^n \left[\prod_{i=1}^n [1 - \exp(-\lambda t_i)]^{\alpha-1} \right] \exp\left(-\frac{\alpha}{a_1} - \lambda \left(\frac{1}{a_2} + \sum_{i=1}^n t_i\right)\right) \left[1 + \delta(2e^{-\frac{\alpha}{a_1}} - 1)(2e^{-\frac{\lambda}{a_2}} - 1) \right] / \sqrt{1 - \delta^2} \tag{30}$$

In our study, the aim is to specify the posterior marginal distributions of the parameters α and λ . Thus, to estimate the parameters of interest, we use the marginal posterior distributions by integrating out the nuisance parameter δ from the joint density Eq. (30). Solving the integral above for the parameter δ and considering the integral results,

$$\int_{-1}^1 \frac{1 + \delta(2e^{-\frac{\alpha}{a_1}} - 1)(2e^{-\frac{\lambda}{a_2}} - 1)}{\sqrt{1 - \delta^2}} d\delta = \int_{-1}^1 \frac{1}{\sqrt{1 - \delta^2}} d\delta + (2e^{-\frac{\alpha}{a_1}} - 1)(2e^{-\frac{\lambda}{a_2}} - 1) \int_{-1}^1 \frac{\delta}{\sqrt{1 - \delta^2}} d\delta \text{ but } \int_{-1}^1 \frac{1}{\sqrt{1 - \delta^2}} d\delta = \pi$$

and $\int_{-1}^1 \frac{\delta}{\sqrt{1 - \delta^2}} d\delta = 0$; in this way, we have the joint posterior distribution for α and λ given by

$$\Pi(\alpha, \lambda | t) \propto \alpha^n \lambda^n \left[\prod_{i=1}^n [1 - \exp(-\lambda t_i)]^{\alpha-1} \right] \exp\left(-\frac{\alpha}{a_1} - \lambda \left(\frac{1}{a_2} + \sum_{i=1}^n t_i\right)\right), \tag{31}$$

which is the same posterior obtained if we have used $\pi(\delta) \propto 1$. Samples of the joint posterior distributions of Eq. (30) can be simulated by using MCMC methods. In this way, we simulate α from the conditional posterior distribution $\Pi(\alpha|\lambda, t)$ and λ from the conditional distribution $\Pi(\lambda|\alpha, t)$ using the Metropolis-Hastings algorithm (see for example, Chib and Greenberg, 1995).

4. Survival function S

In applications, we usually have interest in the survival function S , given by Eq. (2), that is, if X represents the lifetime of a patient under a given treatment then S represents the probability of this patient "survives" for at least a specified time t (or longer). Next, we compare the posterior densities of the survival function $S(t)$ by using the priors discussed in this paper. We note that Eq. (2) is a function of α and λ and hence it is a parameter itself with a posterior distribution $\Pi(S | t)$. To derive this

posterior we first transform (α, λ) to (S, W) where $W = \alpha$ and $S = S_0 = 1 - [1 - \exp(-\lambda t_0)]^\alpha$. To derive this posterior we first transform (α, λ) to (S, W) where $W = \alpha$ and $S = S_0$. As Jeffreys and gamma priors are invariant to 1-1 transformation then the posterior distribution for the new parameters can be derived by the Jacobian transformation, that is,

$$\Pi(S, W | t) \propto \Pi(\alpha, \lambda | t) |J| \quad (32)$$

for $\alpha = W$ and $\lambda = -\frac{1}{t_0} \log(1 - (1 - S)^{1/W})$. The Jacobian matrix J of this transformation with respect to the parameters W and S is given by:

$$J = \begin{pmatrix} 0 & 1 \\ -\frac{1}{t_0 W} \left(\frac{(1-S)^{1/W-1}}{1-(1-W)^{1/W}} \right) & \frac{\partial \lambda}{\partial W} \end{pmatrix}, \quad (33)$$

where $\frac{\partial \lambda}{\partial W}$ is the derivative of $\lambda = -\frac{1}{t_0} \log(1 - (1 - S)^{1/W})$ with respect to W .

Note that although the prior distribution of Eq. (24) belongs to the copula Farlie-Gumbel family the prior resulting from transformation S and W does not belong to the same family, that is, this prior is not invariant under nonlinear transformations. Marginal posterior of S could be obtained by integrating $\Pi(S, W | t)$ with respect to the auxiliary variable W , but this is complicated. We prefer the MCMC approach to get this posterior density. From Eq. (2), the MLE of survival function $S = S(t_0)$ is given by $\hat{S} = 1 - [1 - \exp(-\hat{\lambda} t_0)]^{\hat{\alpha}}$ where $\hat{\alpha}$ and $\hat{\lambda}$ are the MLE of α and λ , respectively.

5. Numerical Illustrations

In this section, we introduce two examples to illustrate the proposed method.

5.1 Simulated data

First of all, we present and discuss Bayesian inferences based on simulated samples of size $n = 5, 25$ and 50 generated from the GE distribution with parameters $\alpha = 1.5$ and $\lambda = 3.5$. The data set are given in Table 1.

Table 1

Random samples of 5, 25 and 50 observations from GE $f(x|\alpha, \lambda) = f(x | 1.5, 3.5)$

0.1294211	0.326627	0.288136	0.282337	0.155691
0.3328912	0.554069	0.6331775	0.1971299	0.4895222
0.4033831	0.2053582	0.1210208	0.3974832	0.1764692
0.043011	1.1469335	1.3916718	0.3462968	0.2938541
0.0908279	0.4975098	1.8754314	0.2369877	0.1198675
0.0320994	0.744982	0.020457	0.1040829	0.5089536
0.0872934	0.1195056	0.044733	0.097142	0.4601925
0.5188672	0.561651	0.346562	2.4052356	1.0941212
0.2344864	0.6629314	0.4760794	0.1884854	0.4192122
0.0781923	0.1524062	0.2121481	0.0503638	0.572726
0.3789628	0.2159233	0.0170265	1.3900206	0.7825544
0.3015115	0.1670891	0.1216162	0.9102928	0.3033985
0.1272576	0.5586492	0.167599	0.3013364	0.1478698
0.0347033	0.2230028	0.3697116	0.1550103	0.1998605
0.1243817	0.7130682	0.0123715	0.1804198	0.1135168
0.2553322	0.0556948	0.0847322	0.2095519	0.5097356

In Fig. 1, we have contour plots for the likelihood function (7) for α and λ considering the simulated data sets of Table 1 and the parameterizations (α, λ) , $(\varphi_1, \varphi_2) = (\log\alpha, \log\lambda)$ and $(\psi_1, \psi_2) = (1/\alpha, 1/\lambda)$, respectively.

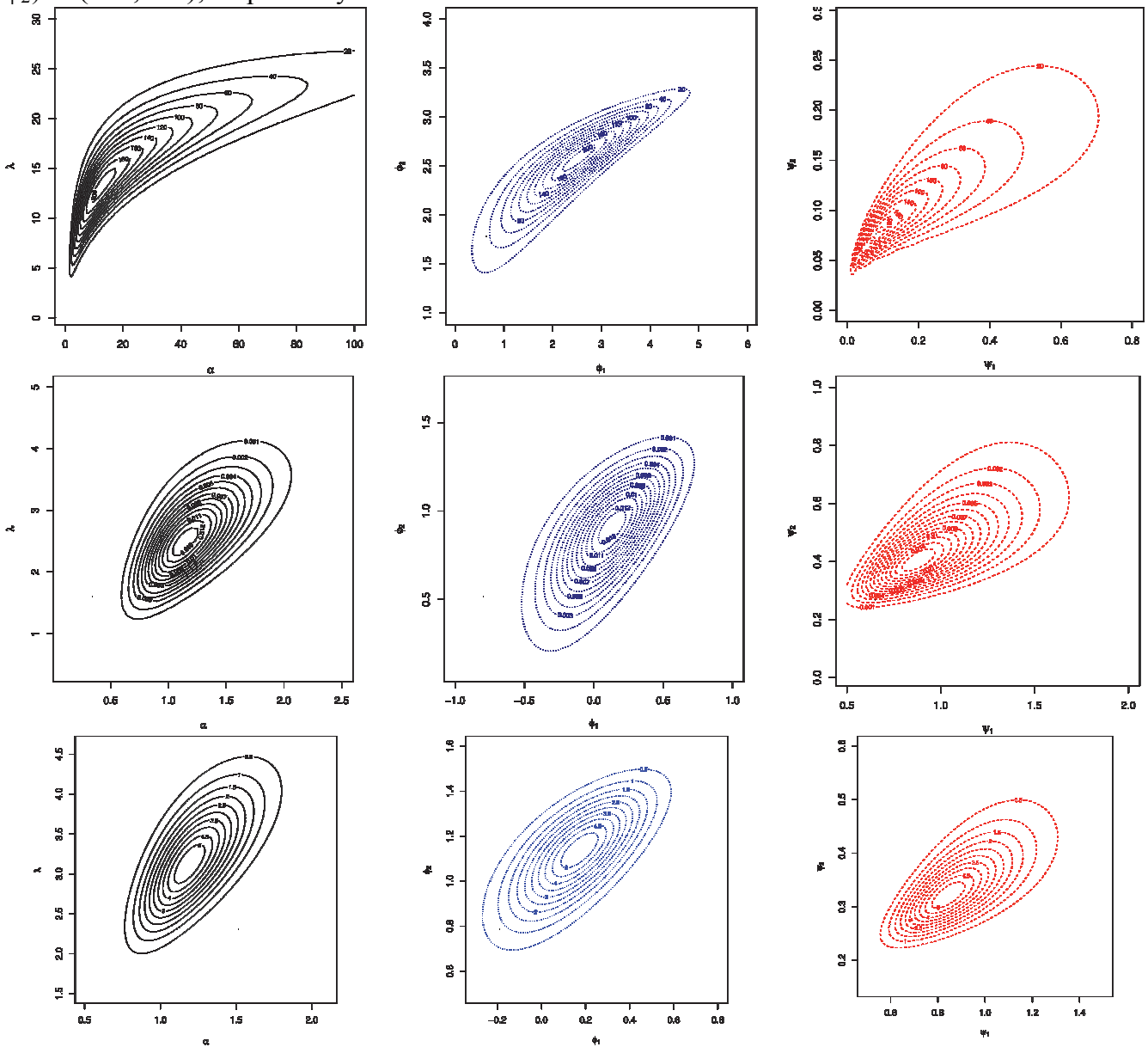


Fig. 1. Contours of likelihood on different parameterizations (α, λ) , $(\varphi_1, \varphi_2)=(\log \alpha, \log \lambda)$ and $(\psi_1, \psi_2)=(1/\alpha, 1/\lambda)$ for sample sizes $n = 5$, $n = 25$ and $n = 50$.

From the plots of Fig. 1, we observe that the contour plots are strongly affected by the different parameterizations, especially for small sample sizes as the case of $n = 5$ observations. In this case, we need to be careful to assume classical asymptotical inference results. For comparison of the different priors proposed in this paper, the joint posterior contours are given in Fig. 2, Fig. 3 and Fig.4 considering the data sets introduced in Table 1 with $n = 5$, 25 and 50, respectively. Observe that the contour plots are similar considering the different non-informative prior distributions for α and λ , specially for large sample sizes ($n = 25$ and $n = 50$). For small sample sizes ($n = 5$), we observe that the choice of non-informative priors for α and λ could be an important issue in the Bayesian analysis of $GE(\alpha, \lambda)$ distributions. We also need to appeal to numerical procedures to extract characteristics of marginal posterior distributions such as Bayes estimator, mode and credible intervals. We can then use MCMC algorithm to obtain a sample of values of α and λ from the joint posterior.

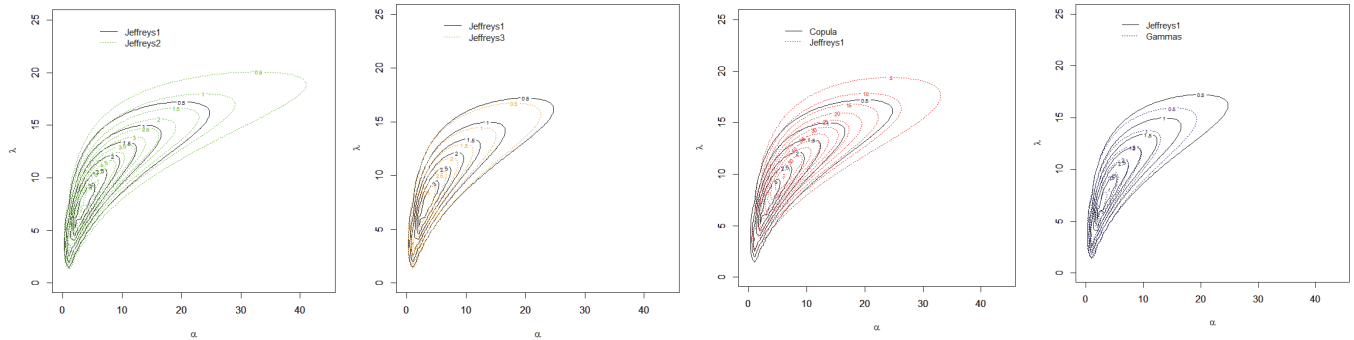


Fig. 2. Contours of posterior on (α, λ) for sample size $n = 5$

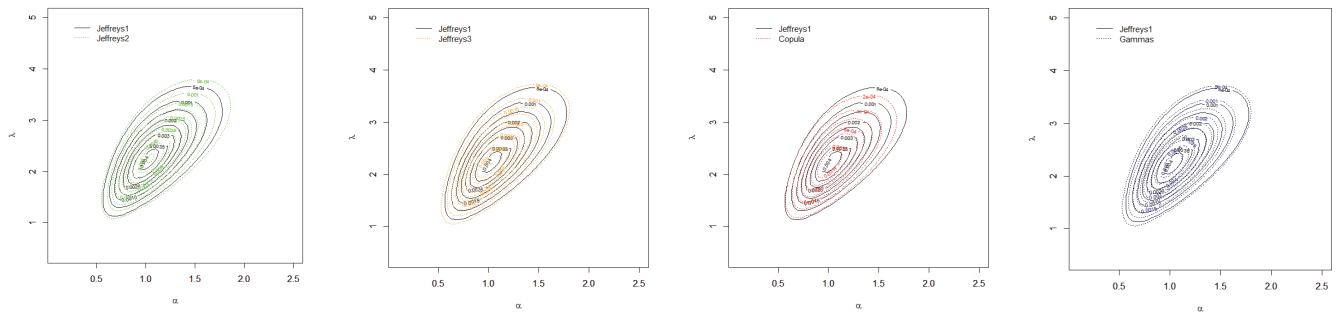


Fig. 3. Contours of posterior on (α, λ) for sample size $n = 25$

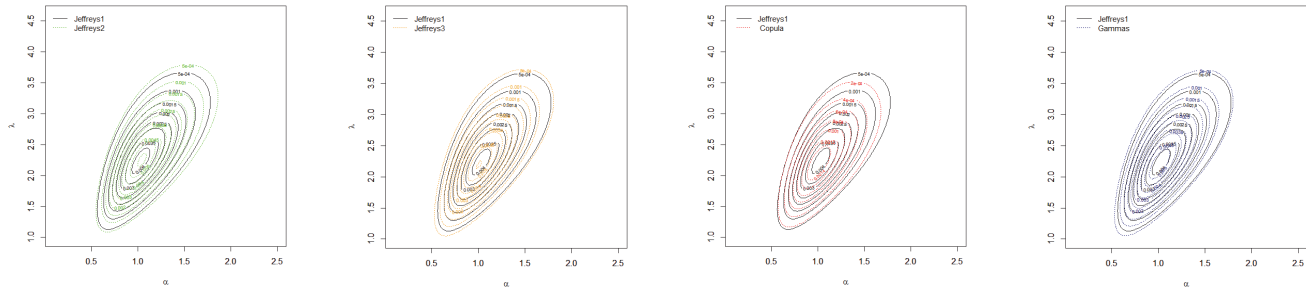


Fig. 4. Contours of posterior on (α, λ) for sample size $n = 50$

The chain is run for 25,000 iterations with a burn-in period of size 5,000. The MCMC traceplots for the posterior distribution are shown in Fig. 5, Fig. 6 and Fig. 7, and the resulting marginal distributions are plotted in Fig. 8, Fig. 9 and Fig. 10. It is important to point out that we have used the hyperparameter values $a_\alpha = b_\alpha = a_\lambda = b_\lambda = 0.01$ for the gamma priors (25) and $a_1 = 1$ and $a_2 = 1$ for the “Farlie-Gumbel” prior Eq. (28). We also have used the prior (31) for the parameter δ .

The MCMC plots suggest we have achieved convergence and the algorithm also showed a rate of acceptance around 25-35%.

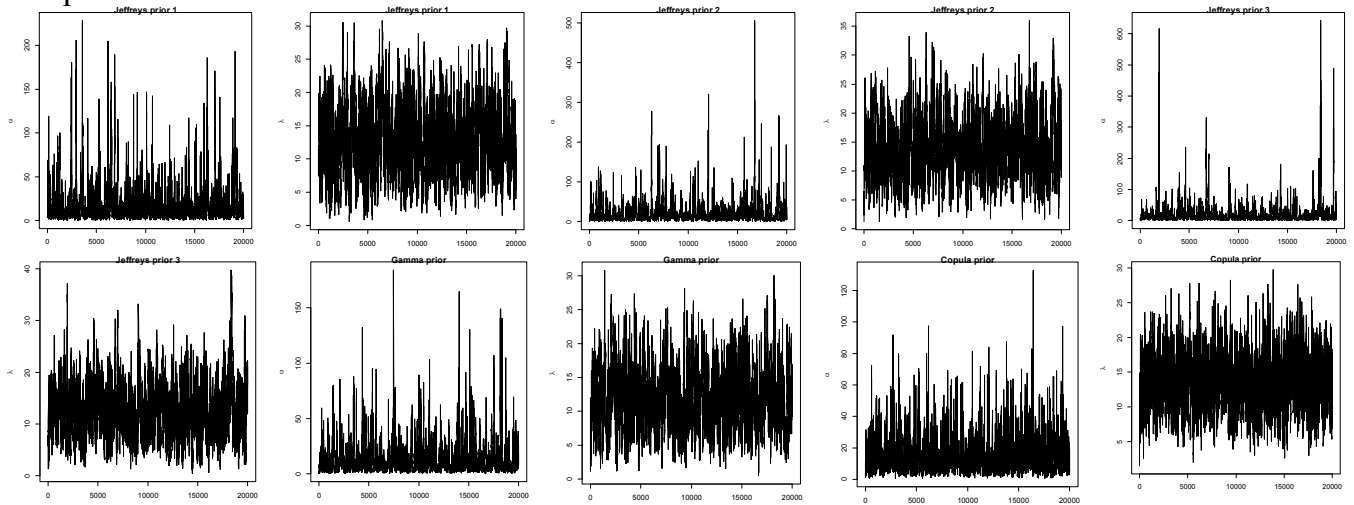


Fig. 5. The MCMC output of posterior on (α, λ) for sample size $n = 5$

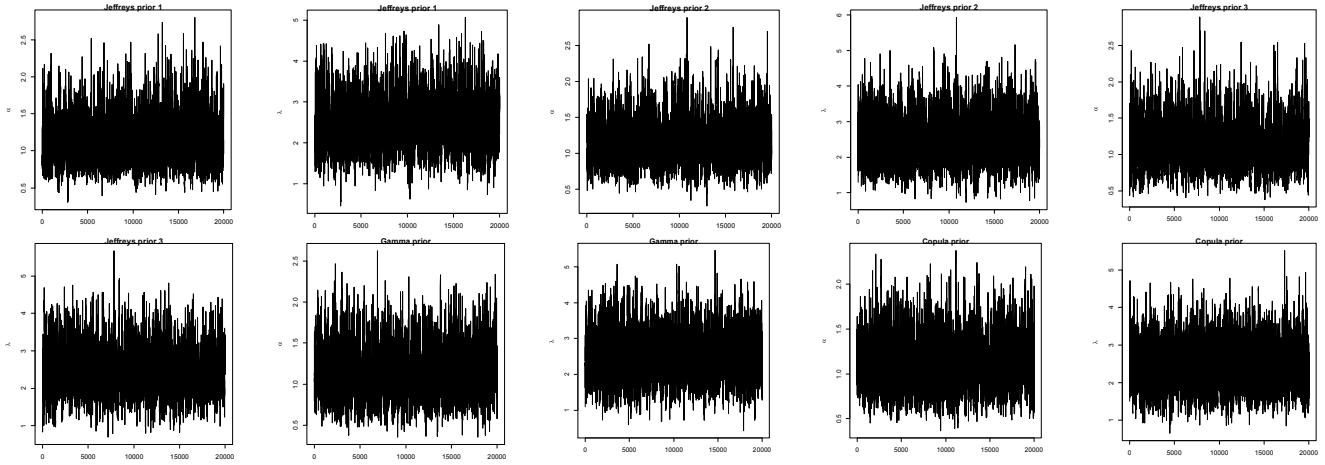


Fig. 6. The MCMC output of posterior on (α, λ) for sample size $n = 25$

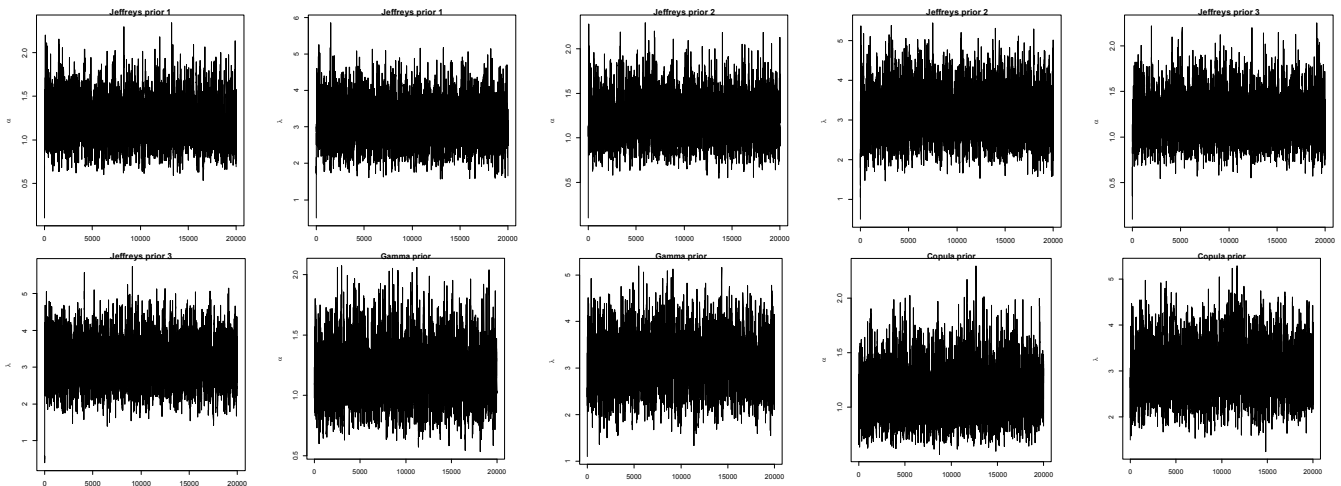


Fig. 7. The MCMC output of posterior on (α, λ) for sample size $n = 50$

Now we examine the performance of the priors by considering several point estimates for parameters α and λ . The maximum likelihood estimate (MLE) is also evaluated (see Tables 2 and 3). From the results of Tables 2 and 3, we observe that the Bayesian posterior means for α and λ are very different from the values $\alpha = 1.5$ and $\lambda = 3.5$ used to simulate the data sets considering a small sample size ($n = 5$). Also observe that in this case, the estimated variances are very large.

Table 2

Posterior mean and variance for the parameter α using the data of Table 1

	Jeffreys1	Jeffreys2	Jeffreys3	Copula	Gammas	MLE
$n = 5$	15.69 (397.11)	20.00 (703.64)	17.54 (1274.7)	13.90 (112.62)	12.56 (194.07)	12.66
$n = 25$	1.14 (0.10)	1.15 (0.10)	1.13 (0.10)	1.10 (0.08)	1.13 (0.10)	1.14
$n = 50$	1.20 (0.05)	1.21 (0.06)	1.19 (0.05)	1.14 (0.04)	1.18 (0.05)	1.19

(The values between parentheses express the posterior variance)

Table 3

Posterior mean and variance for the parameter λ using the data of Table 1

	Jeffreys1	Jeffreys2	Jeffreys3	Copula	Gammas	MLE
$n = 5$	12.36 (25.63)	13.61 (26.80)	12.50 (28.45)	12.93 (15.70)	11.63 (22.80)	13.24
$n = 25$	2.46 (0.39)	2.47 (0.41)	2.43 (0.40)	2.40 (0.36)	2.44 (0.40)	2.48
$n = 50$	3.11 (0.32)	3.14 (0.34)	3.11 (0.32)	3.02 (0.27)	3.07 (0.30)	3.121

(The values between parentheses express the posterior variance)

Obviously, other posterior summaries can be evaluated as well. For example, one may want to derive the posterior intervals for comparison. The 95% posterior intervals for each parameter α and λ obtained using the different non-informative priors are displayed in Tables 4 and 5.

Table 4
95% posterior credible intervals for the parameter α using the data of Table 1

	Jeffreys1	Jeffreys2	Jeffreys3	Copula	Gammas	Confidence Interval
n = 5	(1.23, 69.11)	(1.79, 89.45)	(1.11, 77.02)	(2.38, 41.02)	(1.13, 52.09)	(-12.58, 37.91)
n = 25	(0.64, 1.86)	(0.64, 1.84)	(0.61, 1.84)	(0.63, 1.74)	(0.61, 1.83)	(0.54, 1.74)
n = 50	(0.79, 1.71)	(0.79, 1.71)	(0.80, 1.71)	(0.78, 1.61)	(0.78, 1.69)	(0.74, 1.63)

Table 5
95% posterior credible intervals for the parameter λ using the data of Table 1

	Jeffreys1	Jeffreys2	Jeffreys3	Copula	Gammas	Confidence Interval
n = 5	(3.71, 23.68)	(4.76, 24.68)	(3.68, 24.09)	(5.70, 21.34)	(3.53, 21.83)	(3.11, 23.36)
n = 25	(1.36, 3.78)	(1.35, 3.77)	(1.32, 3.80)	(1.36, 3.68)	(1.30, 3.75)	(1.25, 3.71)
n = 50	(2.08, 4.27)	(2.07, 4.37)	(2.11, 4.28)	(2.04, 4.14)	(2.06, 4.18)	(2.04, 4.20)

From the results of Tables 4 and Table 5, we observe that the 95% credible intervals for α and λ are very large considering a small sample size ($n = 5$ observations). We also observe that with the different non-informative priors for α and λ , we have similar results. The comparison of the marginal posterior densities is given in Figures 8, 9 and 10 assuming the different prior distributions.

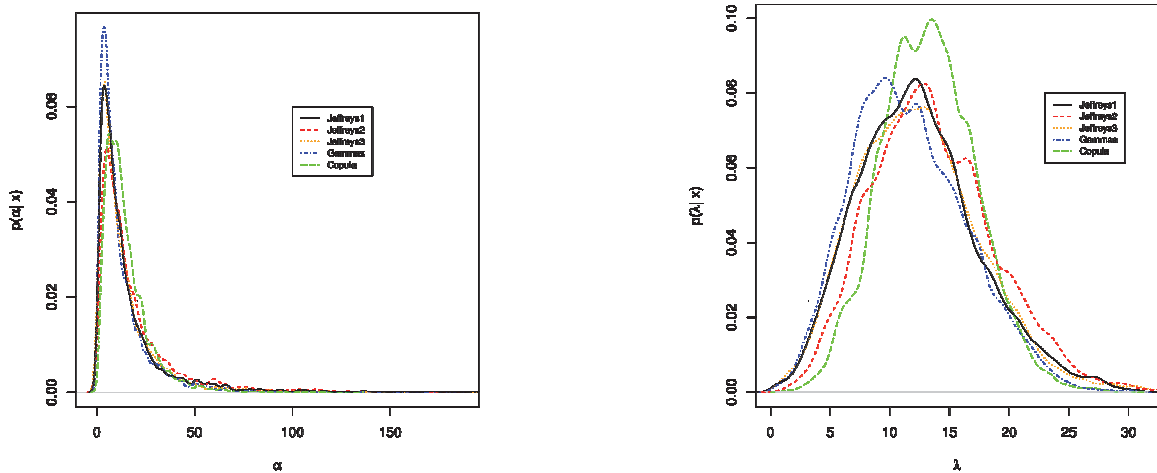


Fig. 8. Marginal posterior densities of parameters α and λ for $n = 5$ and simulated data with $(\alpha, \lambda)=(1.5, 3.5)$

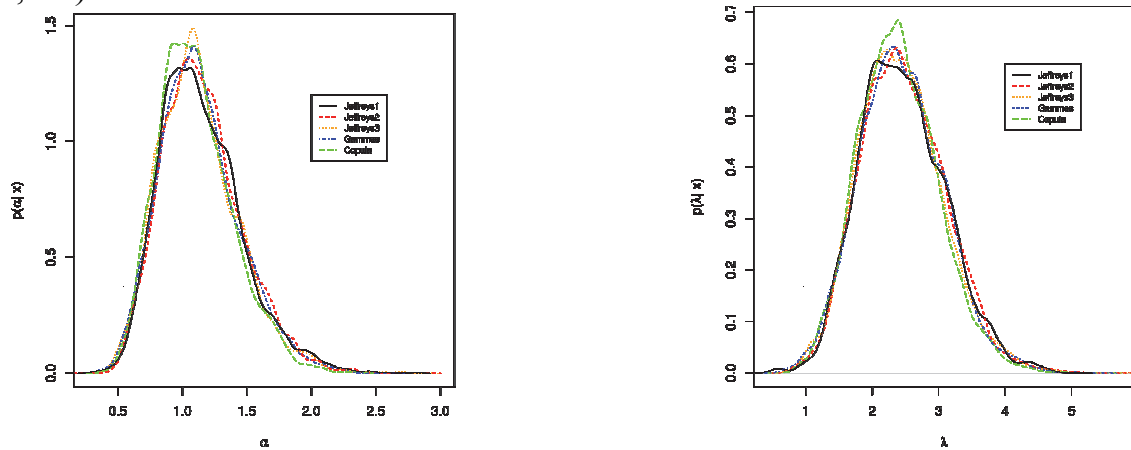


Fig. 9. Marginal posterior densities of parameters α and λ for $n = 25$ and simulated data with $(\alpha, \lambda)=(1.5, 3.5)$

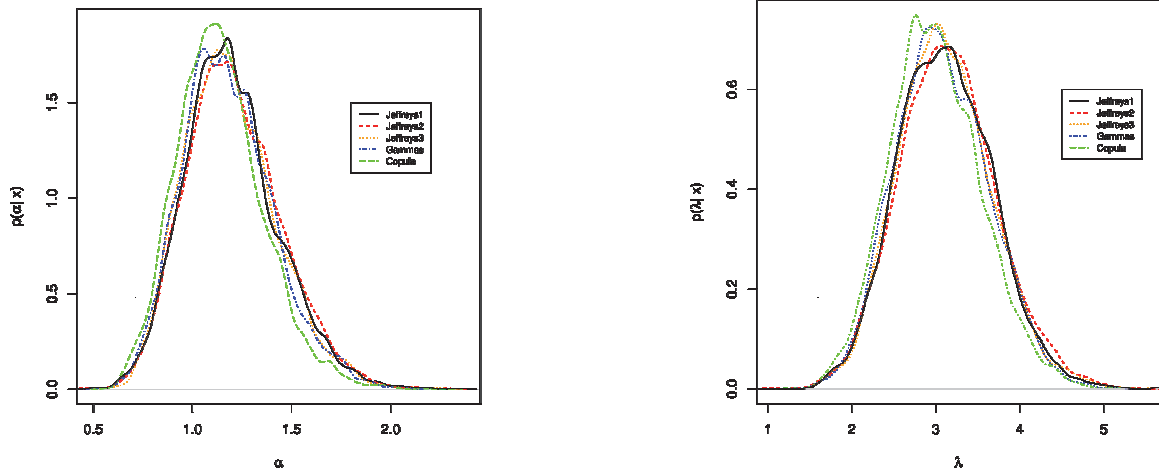


Fig. 10. Marginal posterior densities of parameters α and λ for $n = 50$ and simulated data with $(\alpha, \lambda)=(1.5, 3.5)$

From the results of Figs. (8-10), we observe that the marginal posterior densities for α and λ become more similar assuming the different non-informative priors for α and λ , as the data sample size increases ($n = 25$ and $n = 50$). Assuming $t = 0.2$, the survival function given the parameter values $\alpha = 1.5$ and $\lambda = 3.5$ is given by $S(0.2) = 0.64$. Assuming the three simulated data sets of Table 1, we have in Figs. (11-13), the marginal posterior distributions for $S(0.2)$ obtained from the 20,000 simulated Gibbs samples for the joint posterior distribution of α and λ , considering the different non-informative priors introduced in Section 3. From the plots of Figs. (11-13), we observe that the choice of non-informative priors for the parameters α and λ of the GE(α, λ) distribution could be very important to obtain accurate Bayesian inferences for the survival function, since with the use of the “Farlie-Gumbel (copula)” prior we have more accurate Bayesian inferences for $S(0.2)$. In Tables 6 and Table 7, we have the posterior summaries of interest for $S(0.2)$ assuming the different non-informative priors for α and λ . We also observe more accurate Bayesian inferences for the survival function $S(0.2)$, assuming the “Farlie-Gumbel” prior distribution for the parameters α and λ .

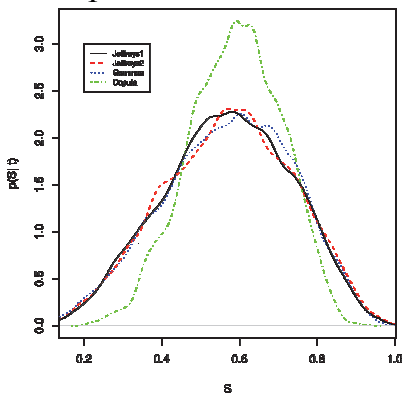


Fig. 11. Posterior densities of survival function S for $n = 5$

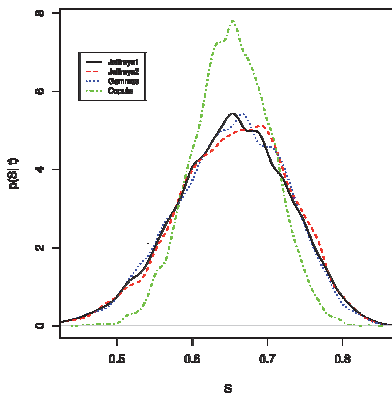


Fig. 12. Posterior densities of survival function S for $n = 25$

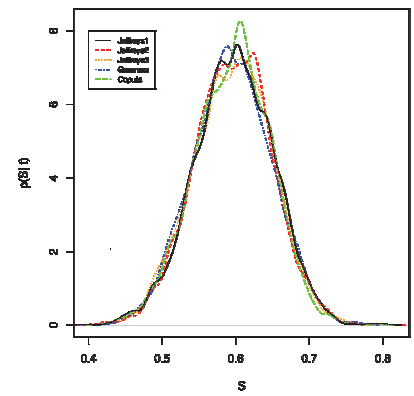


Fig. 13. Posterior densities of survival function S for $n = 50$

Table 6

Posterior mean and variance for the parameter S

	Jeffreys1	Jeffreys2	Jeffreys3	Copula	Gammas	Confidence Interval
$n = 5$	0.57 (0.027)	0.57 (0.027)	0.57 (0.026)	0.59 (0.014)	0.57 (0.027)	0.60
$n = 25$	0.65 (0.006)	0.65 (0.006)	0.65 (0.005)	0.65 (0.003)	0.65 (0.005)	0.66
$n = 50$	0.60 (0.003)	0.60 (0.003)	0.60 (0.003)	0.59 (0.001)	0.59 (0.003)	0.60

(The values between parentheses express the posterior variance)

Table 7

95% posterior intervals for the parameter S

	Jeffreys1	Jeffreys2	Jeffreys3	Copula	Gammas
n = 5	(0.25, 0.86)	(0.24, 0.87)	(0.24, 0.86)	(0.35, 0.80)	(0.23, 0.86)
n = 25	(0.50, 0.79)	(0.50, 0.79)	(0.49, 0.79)	(0.55, 0.75)	(0.50, 0.78)
n = 50	(0.49, 0.70)	(0.49, 0.70)	(0.49, 0.70)	(0.52, 0.66)	(0.49, 0.70)

(The values between parentheses express the posterior variance)

5.2 An example with literature data

In this section, let us consider a data set related to the lifetime of components (data set introduced in Lawless, 1982, page 228): 17.88; 28.92; 33.0; 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; 8.64; 105.12; 105.84; 127.92; 128.04 and 173.40. Let us denote this data as "Lawless data". This data set represents the numbers (in million) of cycles until failure of the component. Let us assume a generalized exponential distribution with density (1) to analyse the data. The maximum likelihood estimators for α and λ (see (10) and (11)) are given, respectively, by where $\hat{\alpha} = 5.2836$ and $\hat{\lambda} = 0.0323$, with 95% confidence interval (1.2714; 9.2653) and (0.0197; 0.0448) for α and λ , respectively. For a Bayesian analysis of the data let us assume the prior distribution (15), (19), (20) and (24) and the Gamma priors (25) for α and λ assuming the hyperparameter values $a_\alpha = b_\alpha = a_\lambda = b_\lambda = 0.01$. Using the software R we first simulated 5,000 Gibbs samples ("burn-in-samples") for the joint posterior distribution for α and λ that were discarded to eliminate the effect of the initial values for the random quantities α and λ used in the iterative simulation method; after this "burn-in-period", we simulated other 20,000 Gibbs samples. The convergence of the Gibbs sampling algorithm was monitored from traceplots of the simulated samples. The posterior summaries of interest considering the different prior distributions are given in Tables 8 and 9. Observe that we have used $a_1 = 1$ and $a_2 = 2$ for the copula prior (28) and the prior (31) for the parameter δ .

Table 8Posterior summaries for α (Lawless data)

Priors	Posterior mean	Posterior variance	95% posterior interval
Jeffreys1	5.451	4.856	(2.286, 10.833)
Jeffreys2	5.614	5.340	(2.481, 11.085)
Jeffreys3	5.289	4.470	(2.202, 10.602)
Copula	5.229	3.540	(2.365, 9.701)
Gammas	5.307	4.575	(2.312, 10.596)

Table 9Posterior summaries for λ (Lawless data)

Priors	Posterior mean	Posterior variance	95% posterior interval
Jeffreys1	0.0321	4.118436e-05	(0.0202, 0.0453)
Jeffreys2	0.0323	4.146039e-05	(0.0211, 0.0459)
Jeffreys3	0.0316	4.09938e-05	(0.0195, 0.0443)
Copula	0.0316	3.499097e-05	(0.0205, 0.0438)
Gammas	0.0317	4.10683e-05	(0.0203, 0.0449)

From the obtained inference results of Table 9, we observe similar results. The comparison of the marginal posterior densities is given in Fig. 14.

Table 10
Posterior summaries for S (Lawless data)

Priors	Posterior mean	Posterior variance	95% posterior interval
Jeffreys1	0.6797	0.0055	(0.529, 0.815)
Jeffreys2	0.6867	0.0055	(0.533, 0.821)
Jeffreys3	0.6791	0.0057	(0.522, 0.816)
Copula	0.6854	0.0029	(0.573, 0.786)
Gammas	0.6759	0.0060	(0.515, 0.816)

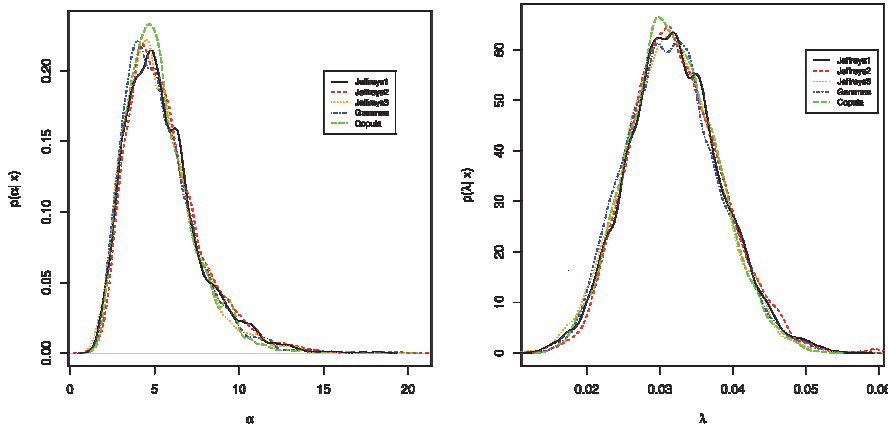


Fig. 14. Marginal posterior densities of parameters α and λ (Lawless data)

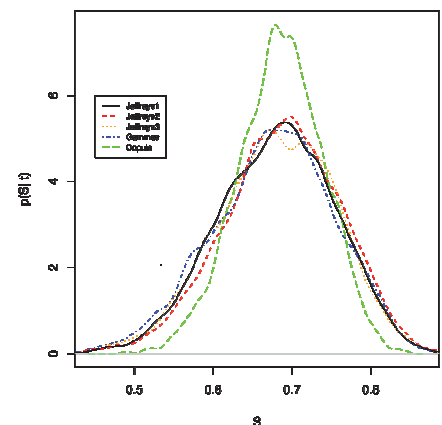


Fig. 15. Posterior density of survival for $t = 50$ (Lawless data)

In Fig. 16, we have the plots of the empirical and fitted survival functions modeled by the generalized exponential for $\hat{\alpha} = 5.23$ and $\hat{\lambda} = 0.031$ considering distribution with density given in Eq. (1) and assuming the Bayesian estimators b the copula prior distribution for α and λ (see Tables 8 and 9). From Fig. 16, we observe a good fit of the generalized exponential distribution for the data (Lawless data).

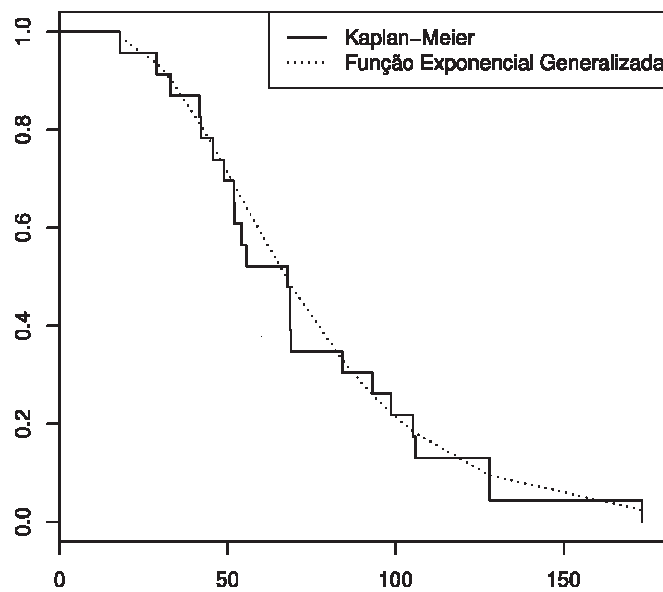


Fig. 16. Empirical and fitted survival for $t = 50$ (Lawless data)

6. Conclusion

The use of the generalized exponential distribution with density (1) could be a good alternative to analyse lifetime data, in comparison to the popular gamma distribution. Observe that the survival function (see (2)) for the $GE(\alpha, \lambda)$ distribution presents a closed analytical form and the survival function for the gamma distribution presents an incomplete gamma function. In this way, since lifetimes of medical or engineering applications present censored data, the use of $GE(\alpha, \lambda)$ can have good computational advantages. Accurate inference for the parameters of the $GE(\alpha, \lambda)$ could be obtained using MCMC methods for a Bayesian analysis of the model. In this way, we need to choose an appropriate prior distribution for the parameters of the model, especially in the situations where we do not have expert opinion to build our prior. From the results of this paper we observe that the use of copula prior distributions for the random quantities α and λ could improve our inference results, especially considering the survival function at specified time t . These results are of great interest in applications.

References

- Box, G. E., & Tiao, G. C. (2011). *Bayesian inference in statistical analysis* (Vol. 40). John Wiley & Sons.
- Chib, S., & Greenberg, E. (1995). Understanding The Metropolis-Hasting Algorithm. *The American Statistician*, 49, 327-335.
- Gelfand, A. E., & Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410), 398-409.
- Gupta, R. D., & Kundu, D. (1999). Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2), 173-188.
- Gupta, R. D., & Kundu, D. (2001). Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation*, 69(4), 315-337.
- Gupta, R. D., & Kundu, D. (2007). Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical Planning and Inference*, 137(11), 3537-3547.
- Kundu, D., & Gupta, R. D. (2008). Generalized exponential distribution: Bayesian estimations. *Computational Statistics & Data Analysis*, 52(4), 1873-1883.
- Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen. *Mitteilungsblatt Math Statistik*, 8, 234-235.
- Nelsen, R. B. (1999). *An introduction to copulas*. Springer.
- Raqab, M. Z. (2002). Inferences for generalized exponential distribution based on record statistics. *Journal of statistical planning and inference*, 104(2), 339-350.
- Raqab, M. M., & Ahsanullah, M. (2001). Estimation of the location and scale parameters of generalized exponential distribution based on order statistics. *Journal of Statistical Computation and Simulation*, 69(2), 109-123.
- Sarhan, A. M. (2007). Analysis of incomplete, censored data in competing risks models with generalized exponential distributions. *Reliability, IEEE Transactions on*, 56(1), 132-138.
- Trivedi, P. K., & Zimmer, D. M. (2007). *Copula modeling: an introduction for practitioners*. Now Publishers Inc.
- Zheng, G. (2002). On the Fisher information matrix in type II censored data from the exponentiated exponential family. *Biometrical Journal*, 44(3), 353-357.