© Statistical Research and Training Center

مجله‌ی پژوهش‌های آماری ایران
پژوهشکده‌ی آمار
دوره‌ی ۱۳، شماره‌ی ۲، پاییز و زمستان ۱۳۹۵، صص ۱۸۱–۱۹۵

J. Statist. Res. Iran **13** (2016): 181–195
DOI: 10.18869/acadpub.jsri.13.2.181

Bayesian Estimation of Parameters in the Exponentiated Gumbel Distribution

Gholamhossein Gholami

Urmia University

Received: 4/28/2016 Approved: 4/8/2017

Abstract. The Exponentiated Gumbel (EG) distribution has been proposed to capture some aspects of the data that the Gumbel distribution fails to specify. In this paper, we estimate the EG's parameters in the Bayesian framework. We consider a 2-level hierarchical structure for prior distribution. As the posterior distributions do not admit a closed form, we do an approximated inference by using Gibbs and Metropolis-Hastings algorithm.

Keywords. Bayesian inference; exponentiated distributions; Gumbel distribution; Gibbs Sampler; Monte Carlo Markov Chain (MCMC) method; Metropolis-Hastings algorithm.

MSC 2010: 62F15; 65C05.

1 Introduction

Recently, some generalizations of several well-known distributions have been made available through using exponentiation. Let $G(x)$ be a baseline continuous cumulative distribution function (cdf), then the cdf of the exponentiated $G(x)$ distribution can be defined in two ways as follows:

$$F_1(x;\alpha) = \big[G(x)\big]^{\alpha}, \tag{1}$$

$$F_2(x;\alpha) = 1 - \big[1 - G(x)\big]^{\alpha}, \tag{2}$$

where $\alpha > 0$ is an extra shape parameter. The baseline distribution is an special case and can be obtained when $\alpha = 1$. The advantage of introducing this second parameter is that expecting the exponentiated distributions can explain various aspects of the data which the original function $(G(x))$ fails to specify. Based on this idea, Gupta and Kundu (2001) provided the exponentiated exponential distribution as a generalization of the exponential distribution. Another exponentiated distribution is the EG distribution which was introduced by Nadarajah and Kotz (2006). This distribution has many applications in weather modeling, frequency analysis of river floods, and wind speed modeling.

The CDF of Gumbel distribution is

$$G(x;\lambda) = \exp\{-\exp(-\lambda x)\}, \qquad x \in \mathbb{R}, \quad \lambda > 0,$$

by introducing α and using (1), the cdf and the density function of the EG distribution can be defined as

$$F(x;\lambda,\alpha) = \big[\exp\{-\exp(-\lambda x)\}\big]^{\alpha}, \qquad \lambda > 0, \quad \alpha > 0, \tag{3}$$

$$f(x;\lambda,\alpha) = \alpha\lambda \exp(-\lambda x)\big[\exp\{-\exp(-\lambda x)\}\big]^{\alpha}, \qquad \lambda > 0, \quad \alpha > 0, \tag{4}$$

receptively.

Maximum likelihood estimators of the parameters of this distribution were studied and analyzed by Gupta and Kundu (2001). The ML estimator for α is given by

$$\hat{\alpha}_{\mathrm{ML}} = \frac{n}{\sum_{i=1}^{n} \exp(-\lambda x_i)}$$

and for λ is the solution of the following equation,

$$\frac{1}{\lambda} + \frac{\sum_{i=1}^{n} x_i \exp(-\lambda x_i)}{\sum_{i=1}^{n} \exp(-\lambda x_i)} - \bar{x} = 0.$$

As this equation cannot be solved analytically, one can use a numerical method, such as the Newton-Raphson method, to find an approximative

© 2016, SRTC Iran

solution.

In this paper, we will obtain the estimators of these parameters within the Bayesian framework. By adopting some prior densities, the posterior densities will be calculated. The rest of the paper is organized as follows. Section 2 provides a Bayesian framework for estimating the parameters. Section 3 presents a simulation study along a real data set to compare ML and proposed estimates. Section 4 concludes the paper.

2 Bayesian Framework for Inference

In order to make statistical inferences based on the Bayesian framework, after assuming a prior density for the parameters, $\pi(\theta)$, and combining this distribution with the information brought by the data which is quantified by the likelihood function ($L(\theta|x)$), the posterior function of the parameters can be determined as follows

$$\pi(\theta|x) \propto L(\theta|x) \cdot \pi(\theta). \tag{5}$$

The remaining of the inference process is fulfilled based on the obtained posterior distribution. According to the assumed loss function, various aspects of the posterior distribution, such as the mean, median, etc can be used to estimate the parameters. See Robert (2007), Gelman et al. (2001) or Congdon (2007) for more details.

2.1 Selection of the Prior Density

Let

$$\theta \sim \pi(\theta|\eta_k), \quad \eta_{k-j} \sim \pi(\eta_{k-j}|\eta_{k-j-1}), \qquad j = 0, \ldots, k-1. \tag{6}$$

We call such a structure for prior distribution a *k-level hierarchical structure*. The parameters of the prior (η_j, $j = 1, 2, \ldots, k-1$) are called hyperparameters and η_0 is a constant-known value. Theoretically, the number of the levels of this structure can be infinite, but, in practice, a high number of the levels will lead to additional complexity of the model structure. The importance of the hyperparameters decreases by increasing the number of their levels. Therefore, regarding the importance and sensitivity of the resulted estimators, one can choose appropriate number of levels.

In terms of the data, as the only important quantity is θ, by integrating out the hyperparameters, if possible, we will obtain a marginal prior density

for θ which is not a function of any unknown quantity, that is

$$\begin{aligned} \pi(\theta|\eta_0) &= \int \pi(\theta, \eta_1, \ldots, \eta_k|\eta_0)\, d_{\eta_1} \ldots d_{\eta_k} \\ &= \int \pi(\theta|\eta_k)\Pi_{j=0}^{k-1}\pi(\eta_{k-j}|\eta_{k-j-1})\, d_{\eta_{k-j}}. \end{aligned} \tag{7}$$

This method is called the *hierarchical method for determination of prior density* Gholami (2008).

In the EG distribution case, we let

$$\alpha, \lambda \sim \exp(\eta), \qquad \eta \sim \exp(b)$$

where b is a constant and known value. Furthermore, we suppose that α and λ are apriori independent. By integrating out η, the marginal prior densities will be

$$\begin{aligned} \pi(\alpha|b) &= \int \pi(\alpha|\eta)\, \pi(\eta|b)\, d\eta \\ &= \int \eta \exp(-\eta \cdot \alpha)\, b \exp(-b \cdot \eta) d_\eta \\ &= \frac{b}{(\alpha + b)^2}, \qquad \alpha > 0, \\ \pi(\lambda|b) &= \frac{b}{(\lambda + b)^2}, \qquad \lambda > 0. \end{aligned}$$

The main issue here is the determination of the constant b. Since this value is independent of the data in hand, according to the concept of non-informative prior density, we choose it in such a way that it has the minimum possible effect on the result of the inference. This can be achieved through letting b to tend to infinity. Then, the prior density will tend to zero which is flat over the positive real numbers. See Robert (2007) for more details.

Figure 1 demonstrates the effect of size of b in its different values. As shown in the figure, the greater values result in flatter priors.

Then the posterior will be

© 2016, SRTC Iran

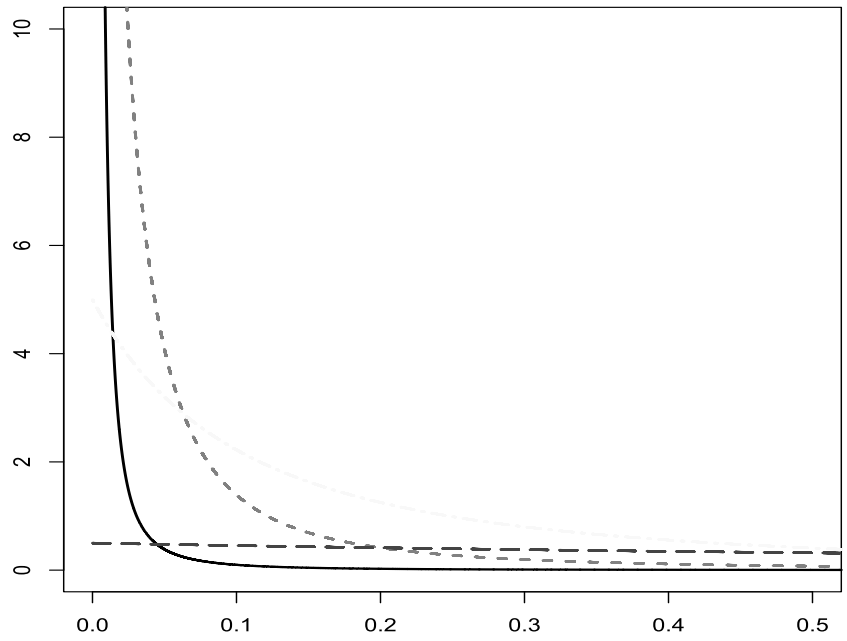

Figure 1. Plot of marginal prior for different values of b. The solid curve corresponds to $b = 0.001$, the dashed curve corresponds to $b = 0.02$, the two-dashed curve corresponds to $b = 0.2$, and the long-dashed curve corresponds to $b = 2$.

$$\begin{aligned}\pi(\alpha,\lambda|x) &\propto \pi(\alpha,\lambda|b)L(\alpha,\lambda|x)\\ &= \pi(\alpha|b)\pi(\lambda|b)L(\alpha,\lambda|x)\\ &\propto \frac{1}{(\alpha+b)^2}\cdot\frac{1}{(\lambda+b)^2}\alpha^n\lambda^n\exp(-\lambda n\bar{x})\left\{\exp\left(-\sum_{i=1}^{n}e^{-\lambda x_i}\right)\right\}^{\alpha}.\end{aligned}$$

This distribution does not entail a closed form, so it cannot be used in the rest of the inference directly. For the rest of the analysis, we will obtain a random sample from this distribution by using Gibbs Sampler. This sample provides a suitable basis for an approximate analysis. Gibbs Sampler draws the sample of the distribution of the parameters through sequential sampling

of full conditional distributions, $\pi(\boldsymbol{\theta}|-\theta_j)$, (i.e. conditional posterior density of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ given all the parameters except θ_j). A two-stage Gibbs Sampler can be considered as follows

Step 0. Choose an arbitrary starting point $\lambda^{(0)}$, and set $t = 0$.

Step 1. Generate $\alpha^{(t+1)} \sim \pi(\alpha|\lambda^{(t)}, \boldsymbol{x})$.

Step 2. Generate $\lambda^{(t+1)} \sim \pi(\lambda|\alpha^{(t+1)}, \boldsymbol{x})$.

Step 3. Set $t = t + 1$, and go to **Step 1**.

See Geman and Geman (1984) or Robert and Casella (2007) for more details.

With respect to the obtained likelihood function and the prior density, full conditional distributions of α and λ can be computed as the following

$$\begin{aligned}
\pi(\alpha|-\alpha) &\propto \frac{1}{(\alpha+b)^2}\alpha^n \left\{\exp\left(-\sum_{i=1}^{n} e^{-\lambda x_i}\right)\right\}^{\alpha}; \\
\pi(\lambda|-\lambda) &\propto \frac{1}{(\lambda+b)^2}\lambda^n \exp(-\lambda n\bar{x}) \left\{\exp\left(-\sum_{i=1}^{n} e^{-\lambda x_i}\right)\right\}^{\alpha}.
\end{aligned}$$

Since these distributions do not have a closed and standard form, it is not possible to sample from it by using direct methods. Metropolis-Hastings algorithm can be used to generate a sample from such distributions. This algorithm draws a sample from $\pi(\alpha|\lambda, \boldsymbol{x})$ as follows

Step 0. Choose an arbitrary starting point $\alpha^{(0)}$ and set $t = 0$.

Step 1. Generate a candidate points α^* from $q(\alpha|\alpha^{(t)}, \psi)$ and u from $\mathcal{U}(0, 1)$.

Step 2. Set $\alpha^{(t+1)} = \alpha^*$ if $u \leqslant \rho(\alpha^{(t)}, \alpha^*)$ and $\alpha^{(t+1)} = \alpha^{(t)}$ otherwise, when the acceptance probability is given by

$$\begin{aligned}
\rho(\alpha^{(t)}, \alpha^*) &= \min\{1, A\}, \\
A &= \frac{\pi(\alpha^*|\lambda, \boldsymbol{x})q(\alpha^{(t)}|\alpha^*, \psi)}{\pi(\alpha^{(t)}|\lambda, \boldsymbol{x})q(\alpha^*|\alpha^{(t)}, \psi)},
\end{aligned}$$

© 2016, SRTC Iran

where A is the acceptance rate.

Step 3. Set $t = t + 1$, and go to **Step 1**.

The density $\pi(\alpha|\lambda, \boldsymbol{x})$ is called the *target distribution* and $q(\alpha|\alpha^{(t)}, \psi)$ is termed as the *proposal density*, *candidate generating density* or the *instrumental distribution*. Theoretically, the only condition in choosing this distribution is that its support should include the support of the target density and its tails should be thicker than the target distribution. This distribution controls the algorithm efficiency. An ideal choice would lead to a small correlation in the produced chain (Roberts and Rosenthal, 2004). To choose an appropriate proposal, one can adopt to a parametric family, then try to find the best value for this parameter. This extra parameter, ψ, is called the tuning parameter. See Robert and Casella (2007) for more details.

In this paper, we will use a truncated normal distribution,

$$q(x|\mu, \sigma^2) = \frac{1}{\Phi(\frac{\mu}{\sigma})} \cdot \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\mathbb{I}_{(x>0)},$$

as the proposal density where $\phi(\cdot)$ and $\Phi(\cdot)$ are and cdf of standard normal distribution, receptively. We will set $\mu = \theta^{(t)}$ (current status of the chain). This algorithm can be considered as a kind of *random walk Metropolis-Hastings* algorithm. The parameter σ^2 is the tuning parameter, which controls the step sizes in walking through the parameter space. This parameter affects the efficiency of the algorithm by two means. Small values allow that the chain, obtained from the algorithm, to have a great capability of mixing (high acceptance rate), while it may not have the adequate power to explore the whole parameter space. Roberts and Rosenthal (2001) showed that the optimized value for this parameters is a value that the acceptance rate for that is approximately 0.44.

The acceptance rates for α and λ are,

$$A_\alpha = \left(\frac{\alpha^{(t)}+b}{\alpha^*+b}\right)^2 \left(\frac{\alpha^*}{\alpha^{(t)}}\right)^n \left\{\exp\left(-\sum_{i=1}^{n} e^{-\lambda x_i}\right)\right\}^{\alpha^*-\alpha^{(t)}} \frac{\Phi(\frac{\alpha^{(t)}}{\sigma_\alpha})}{\Phi(\frac{\alpha^*}{\sigma_\alpha})},$$

$$A_\lambda = \left(\frac{\lambda^{(t)}+b}{\lambda^*+b}\right)^2 \left(\frac{\lambda^*}{\lambda^{(t)}}\right)^n \exp\left\{-n\bar{x}(\lambda^*-\lambda^{(t)})\right\} \times \left[\exp\left\{-\sum_{i=1}^{n}(e^{-\lambda^* x_i}-e^{-\lambda^{(t)}x_i})\right\}\right]^\alpha \frac{\Phi(\frac{\lambda^{(t)}}{\sigma_\lambda})}{\Phi(\frac{\lambda^*}{\sigma_\lambda})},$$

respectively.

3 Simulation Study

To evaluate the performance of the proposed methods, we considered 15 simulated data sets with different sizes and parameters. The estimates are found from iterations of the algorithm with runs of length $100,000$ following burn-ins of $70,000$, while the constant value b is set to 1e+300. The tuning parameters are set such that the acceptance rates for the chains would be around 0.44. Furthermore, to compare the proposed and Maximum Likelihood (ML) methods, we replicated the experiment over m=100 times and obtained the average of the estimates, bias and mean squared error (MSE) of the estimators. The ML estimates were obtained by the Newton-Raphson method while the initial values were set to be the real values. The MLE part was carried out using Mathematica while the proposed methods was programed in R. Table 1 summarizes the outputs for the whole simulations. As seen in the table, by increasing the sample size, n, the average of bias and MSE decrease for both estimates. This table also shows that the proposed Bayesian estimator performs better than MLE in average. Furthermore, for $\alpha = 4$, $\lambda = 2$ the variability of both estimates increases but the proposed estimate is still better.

Figures 2, 3 show the trajectory of the MCMC samples and convergences of the estimators for $n = 10, 50$ and 100 when the true values of them are $\alpha = 1$ and $\lambda = 2$. These figures demonstrate that the produced chains of M-H algorithm mix very well as the chain explores almost the whole parameter space.

© 2016, SRTC Iran

Table 1. The output summaries (estimates, Bias and MSE of the estimators) of the ML for different sample sizes and different parameter values.

α , λ	n	$\hat{\alpha}_{\mathrm{ML}}$, $\hat{\alpha}_{\mathrm{Bayes.}}$	$\hat{\lambda}_{\mathrm{ML}}$, $\hat{\lambda}_{\mathrm{Bayes.}}$	(Bias) MSE($\hat{\alpha}_{\mathrm{ML}}$)	(Bias) MSE($\hat{\alpha}_{\mathrm{Bayes.}}$)	(Bias) MSE($\hat{\lambda}_{\mathrm{ML}}$)	(Bias) MSE($\hat{\lambda}_{\mathrm{Bayes.}}$)
	10	0.1914 , 0.2756	0.3639 , 0.3160	(-0.0086) 0.0173	(0.0059) 0.0240	(0.0639) 0.0194	(0.0159) 0.0079
0.2 , 0.3	50	0.1976 , 0.2168	0.3085 , 0.3008	(-0.0024) 0.0026	(0.0068) 0.0026	(0.0085) 0.0011	(0.0008) 0.0009
	100	0.1968 , 0.2058	0.3056 , 0.3022	(-0.0032) 0.0013	(0.0058) 0.0014	(0.0056) 0.0007	(0.0022) 0.0007
	10	0.2058 , 0.2746	3.4220 , 3.1086	(0.0059) 0.0177	(0.0746) 0.0234	(0.4220) 1.1732	(0.1086) 0.5985
0.2 , 3.0	50	0.1920 , 0.2042	3.1255 , 3.0826	(-0.0089) 0.0028	(0.0043) 0.0027	(0.1254) 0.1306	(0.0826) 0.1103
	100	0.1926 , 0.2024	3.0459 , 3.0288	(-0.0035) 0.0013	(0.0024) 0.0012	(0.0458) 0.0561	(0.0288) 0.0532
	10	1.1884 , 1.3016	2.3150 , 2.3901	(0.1884) 0.4623	(0.3016) 0.7871	(0.3150) 0.5559	(0.3901) 0.6450
1.0 , 2.0	50	1.0312 , 1.0182	2.0552 , 2.0299	(0.0312) 0.0256	(0.0182) 0.0251	(0.0552) 0.0577	(0.0299) 0.0576
	100	1.0266 , 1.0337	2.0621 , 2.0674	(0.0266) 0.0134	(0.0337) 0.0132	(0.0621) 0.0310	(0.0621) 0.0310
	10	3.8627 , 3.0003	0.2287 , 0.1966	(0.8627) 3.2749	(0.0002) 0.9807	(0.0287) 0.0040	(-0.0034) 0.0022
3.0 , 0.2	50	3.1513 , 3.0297	0.2048 , 0.1998	(0.1513) 0.2894	(0.0297) 0.2424	(0.0048) 0.0005	(-0.0002) 0.0005
	100	3.0664 , 3.0479	0.2018, 0.2000	(0.0669) 0.1131	(0.0479) 0.1567	(0.0018) 0.0003	(-0.0017) 0.0002
	10	5.6403 , 4.1914	2.2769 , 1.9277	(1.6403) 18.0370	(0.1924) 4.1653	(0.2769) 0.3712	(-0.0415) 0.1959
4.0 , 2.0	50	4.2609 , 4.1100	2.0305 , 1.9842	(0.2609) 0.5780	(0.1100) 0.5062	(0.0305) 0.0499	(-0.0158) 0.0511
	100	4.0695 , 4.0064	1.9978 , 1.9759	(0.0695) 0.2761	(0.0064) 0.2642	(-0.0022) 0.0256	(-0.0241) 0.0278

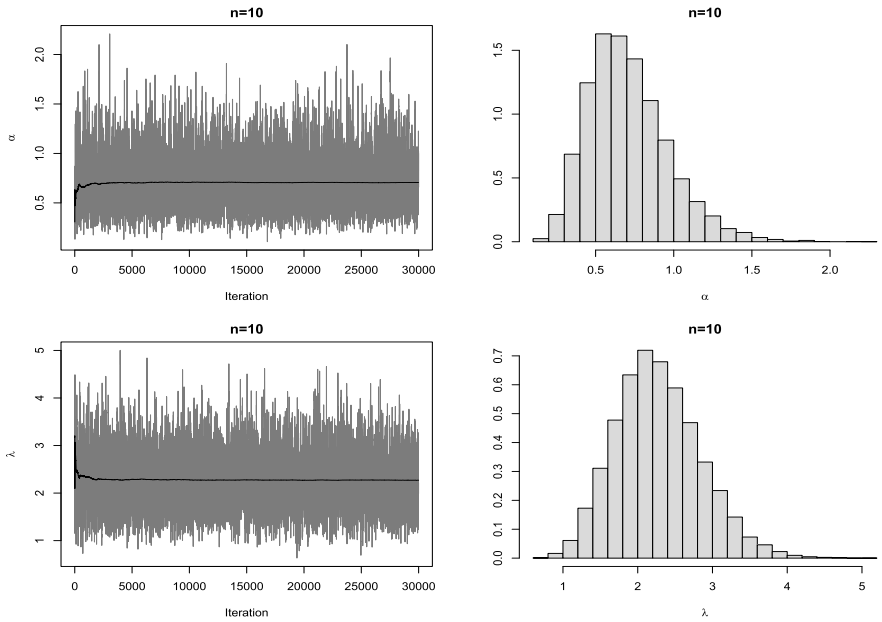

Figure 2. Evolution of the MCMC samples with convergence of the MCMC estimators (*left column*) along with the corresponding histograms (*right column*) for the produced chains, when $n = 10$. True values of the parameters are $\alpha = 1$ and $\lambda = 2$.

4 Fitting of Real Data

As mentioned before, the EG distribution has many applications in climate modeling. The quantity of *significant wave height* (Hs) is interested in oceanography. This is defined as the mean wave height of the highest third of the waves at a given location. This type of data which originates from buoy measurements were calculated as the mean of the highest 1/3 of all of wave highest during 20-minute sampling period Persson and Rydén (2010). It is well known that extreme-value distribution for this type of data is the EG distribution, therefore, we try to fit them the EG distribution. The data were taken from Persson and Rydén (2010), it consists of 21 yearly maxima of Hs (m), reported in Table 2. They are also available online from National Data Buoy Center (NDBC).

© 2016, SRTC Iran

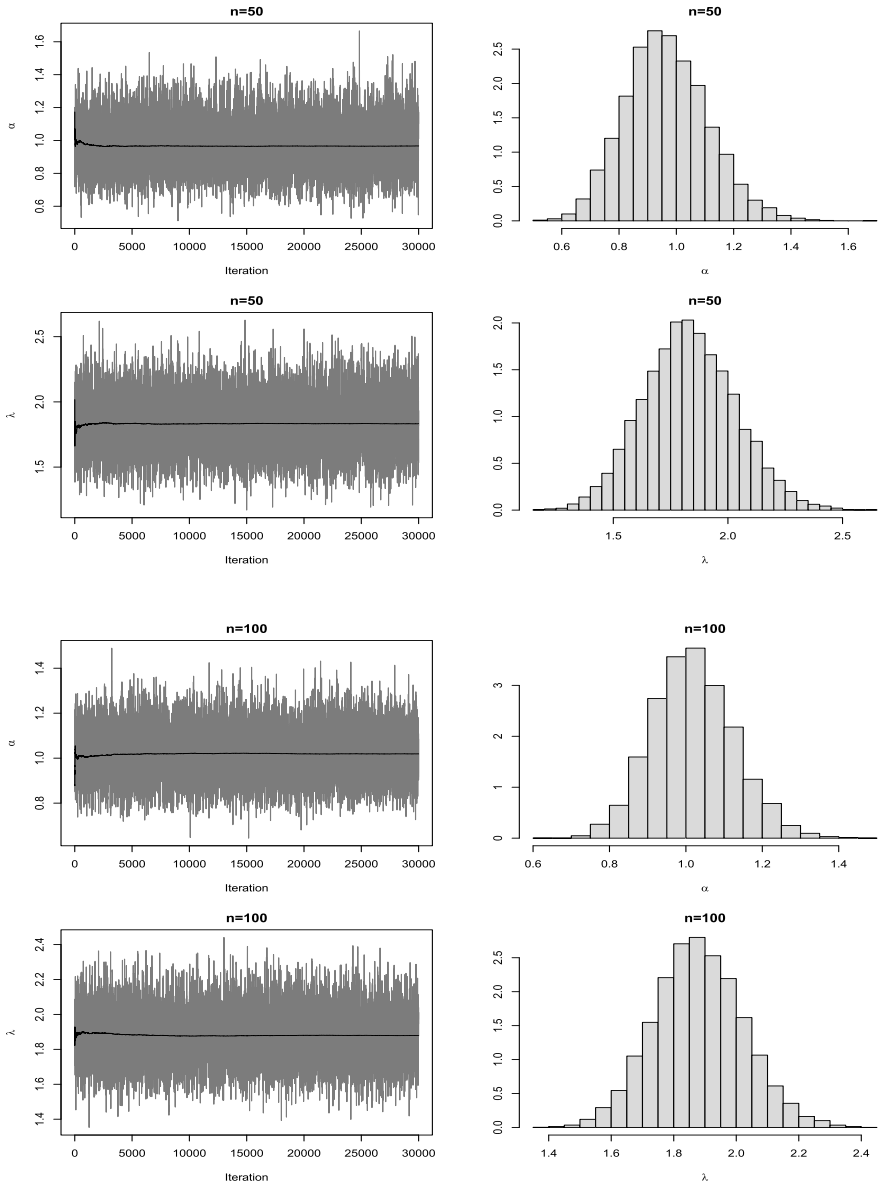

Figure 3. Evolution of the MCMC samples with convergence of the MCMC estimators (*left column*) along with the corresponding histograms (*right column*) for the produced chains, when $n = 50$ and 100. True values of the parameters are $\alpha = 1$ and $\lambda = 2$.

Table 2. Yearly maxima of Hs (m) (Buoy 46006, 41 N, 137 W)

12.90	8.80	11.80	12.70	11.70	9.10	8.40
9.60	7.20	9.80	10.80	10.10	11.20	9.56
8.25	12.47	16.32	14.65	12.78	14.23	11.21

We fit the EG distribution to Buoy data and parameters are estimated by the MLE and the proposed methods. To compare the estimators, we used the Anderson-Darling (AD) and Kolmogorov-Smirnov (KS) distances. The results are reported in Table 3. As seen in the table the estimations are quite competitive. The MLE fits better if one considers the AD distance while the KS distance shows the superiority of Bayesian estimation to MLE. The comparison of these estimates can be examined by quantile and empirical versus estimated CDF plots. A quantile plot graphs the sorted values of the observed quantile against the quantile predicted by the fitted distribution. The ith predicted quantile, y_i, is determined by $F(y_i) = (i-0.375)/(n+0.25)$, where F is the CDF of the EG distribution (Nadarajah, 2006). These plots are shown in Figure 4. Again, the figure shows that both estimates fit very well to the data.

Figure 5 shows the convergence of the proposed estimators. This figure demonstrates that the corresponding chain for the shape parameter, α, produce a skewed posterior distribution and then its variance is not small.

Table 3. ML and Bayesian estimates of the parameters along with AD and KS distances.

$\hat{\alpha}_{\mathrm{ML}}$	$\hat{\alpha}_{\mathrm{Bayes.}}$	$\hat{\lambda}_{\mathrm{ML}}$	$\hat{\lambda}_{\mathrm{Bayes.}}$	$\mathrm{AD}_{\mathrm{ML}}$	$\mathrm{AD}_{\mathrm{Bayes.}}$	$\mathrm{KS}_{\mathrm{ML}}$	$\mathrm{KS}_{\mathrm{Bayes.}}$
167.975	167.911	0.511	0.506	0.17279	0.17384	0.10053	0.083396

5 Conclusion

In this paper, we proposed a Bayesian method to estimate the parameters of the EG distribution. We compared the proposed and maximum likelihood estimates through a simulation study and a real data. We showed that, in average, the performance of the proposed method is better than the ML estimate. For the real data, we compared the estimates by quantile, CDF and empirical plots, Anderson-Darling, and Kolmogrov-smirnove distance. These measures showed that both estimates fit well for the data and they are very competitive.

© 2016, SRTC Iran

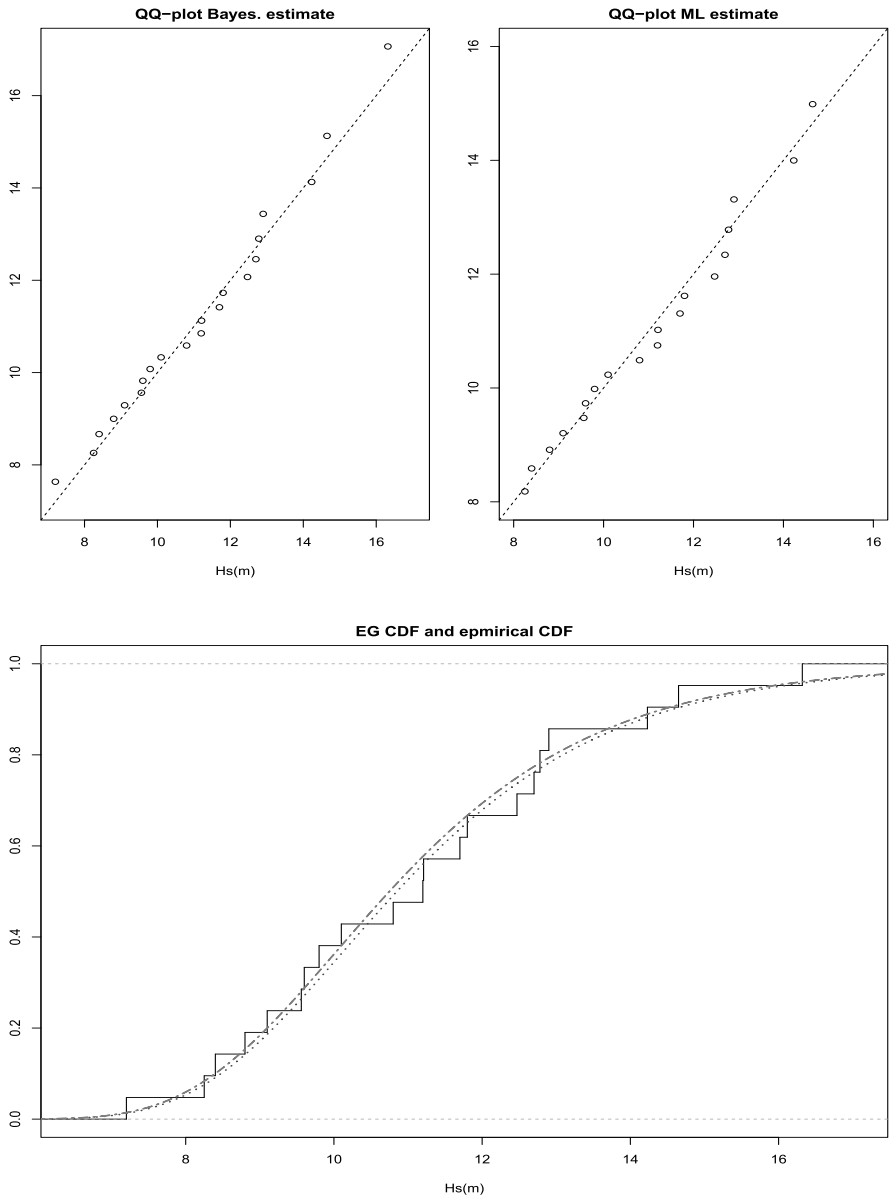

Figure 4. *Top:* The quantile plots of the Bayesian (*right*) and the ML (*left*) estimates. *Bottom:* Empirical and fitted CDFs for the Bayesian (*dotted*) and the ML (*dashed*) estimates.

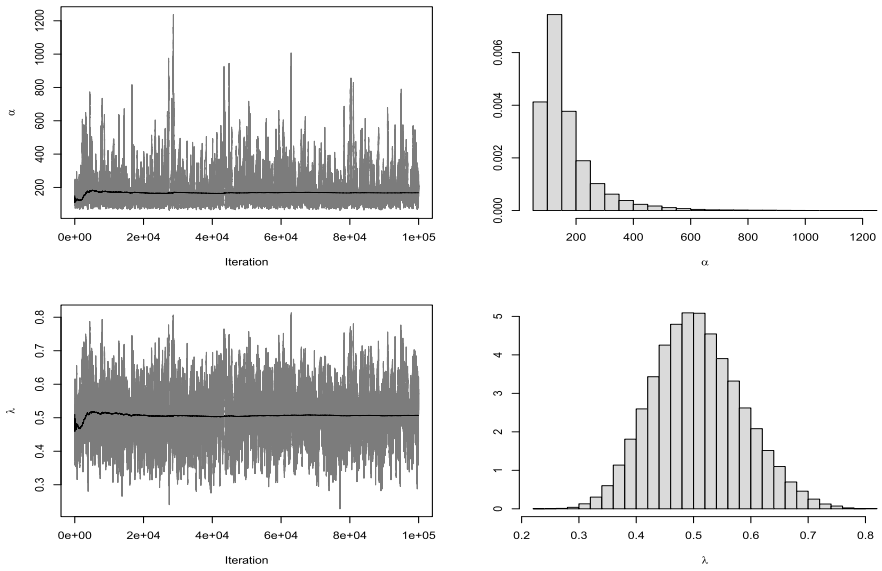

Figure 5. Evolution of the MCMC samples with convergence of the MCMC estimators (*left column*) along with the corresponding histograms (*right column*) for the produced chains of the real data.

Acknowledgment

I would like to thank the Editor and two anonymous referees for their helpful comments and considerations which improved the paper.

References

Congdon, P. (2007). *Bayesian Statistical Modelling.* Wiley, New York.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2001). *Bayesian Data Analysis.* Chapman and Hall, New York, second edition.

Geman, S. and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Trans. Pattern Anal. Mach. Intell.*, **6**, 721-741.

Gholami, G. (2008). *Le Problème du Changement de Régime: Une Approche Bayésienne.* PhD thesis, Université paris Dauphine CEREMADE, Paris.

© 2016, SRTC Iran

Gupta, R.D. and Kundu, D. (2001). Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions. *Biometrical Journal*, **43**, 117-130.

Persson, K. and Rydén, J. (2010). Exponentiated Gumbel Distribution for Estimation of Return Levels of Significant Wave Height. *Journal of Environmental Statistics*, **1**, 1-12.

Nadarajah, S. and Kotz, S. (2006).The Exponentiated Type Distributions. *Acta Applicandae Mathematica*, **92**, 97-111.

Nadarajah, S. (2006). The Exponentiated Gumbel Distribution with Climate Application. *Environmetrics*, **17**, 13-23.

Robert, C.P. (2007). *The Bayesian Choice*. Springer-Verlag, New York, second edition.

Robert, C.P. and Casella, G. (2007). *Monte Carlo Statistical Methods*. Springer-Verlag, New York, second edition.

Roberts, G.O. and Rosenthal, J.S. (2001). Optimal Scaling for Various Metropolis-Hastings Algorithms. *Statistical Science*, **16**, 351-367.

Roberts, G.O. and Rosenthal, J.S. (2004). General State Space Markov Chains and MCMC Algorithms. *Probability Surveys*, **1**, 20-71.

Gholamhossein Gholami
Department of Mathematics,
Faculty of Scinces,
Urmia University,
Urmia Iran.
email: *gh.gholami@urmia.ac.ir*