MCMC Technique to Study the Bayesian Estimation Using Record Values from the Lomax Distribution

M. A. W. Mahmoud
Mathematics Department,
Faculty of Science, Al-Azhar University,
Nasr City (11884), Cairo, Egypt.

A. A. Soliman
Mathematics Department,
Faculty of Science,
Islamic University,Madinah, Saudi Arabia.

A. H. Abd Ellah
Mathematics Department,
Faculty of Science, Sohag University,
Sohag 82524, Egypt

R. M. El-Sagheer
Mathematics Department,
Faculty of Science, Al-Azhar University,
Nasr City (11884), Cairo, Egypt.

ABSTRACT

In this paper, the Bayes estimators of the unknown parameters of the Lomax distribution under the assumptions of gamma priors on both the shape and scale parameters are considered. The Bayes estimators cannot be obtained in explicit forms. So we propose Markov Chain Monte Carlo (MCMC) techniques to generate samples from the posterior distributions and in turn computing the Bayes estimators. Point estimation and confidence intervals based on maximum likelihood and bootstrap methods are also proposed. The approximate Bayes estimators obtained under the assumptions of non-informative priors, are compared with the maximum likelihood estimators using Monte Carlo simulations. One real data set has been analyzed for illustrative purposes.

General Terms:

Computer Science Algorithms

Keywords:

Lomax distribution, Bayesian and non-Bayesian estimations, Gibbs and Metropolis sampler, Bootstrap methods.ifx

1. INTRODUCTION

Record values and the associated statistics are of interest and importance in many areas of real life applications involving datd relating to meteorlogy, sport, economics and lifetesting. Many authors have studied records and associated statistics. Among them are Resnick [25], Nagaraja [21], Ahsanulla [3, 4], Arnold et al. [6,7], Raqab and Ahsanulla [24], Raqab [23] and Abd Ellah [1, 2].

The Lomax distribution belongs to the class of decreasing failure rate distributions see also Chahkandi and Ganjali, [9] Sometimes it is called Pareto distribution of the second kind or Pareto Type-II distribution. It was introduced by Lomax [18] as a model for business failure data. For its applications as lifetime distribution and extensions, we refer to Marshall and Olkin [19] Bryson [8] has argued that Lomax distributions provide a very good alternative to common lifetime distributions like exponential, Weibull, or gamma distributions when the experimenter presumes that the population distribution may be heavy-tailed. Details on Pareto distributions as well as areas of application can be found in Arnold [5], Lomax distribution can be considered as a mixture of the exponential gamma distribution. Lomax distribution includes increasing and decreasing hazard rates as well. Lomax distribution has been shown to be useful for modeling and analizing the life time data in medical and biological sciences, engineering, etc. So, it has been received the greatest attention from theoretical and applied statisticians primarily due to its use in reliability and lifetesting studies. Many statistical methodes have been developed for this distribution, for a review of Lomax distribution see Habibullh and Ahsanullah [15], Upadhyay and Peshwani [29] and Abd Ellah [1, 2] and references of them. Agreat deal of research has been done on estimating the parameters of a Lomax using both classical and Bayesian techniques.

Therefore, the purpose of this paper is to develops the Bayes estimates and Markov Chain Monte Carlo (MCMC) techniques to compute the credible intervals and bootstrap confidence intervals of the unknown parameters of Lomax distribution under the upper record values.

Let $X_1, X_2, X_3, \cdots$ be a sequence of independent and identically distributed random variables with *cdf* $F(x)$ and *pdf* $f(x)$. Set $Y_n = \max(X_1, X_2, X_3, \cdots, X_n)$, $n \geq 1$, X_j is said to be an upper record and is denoted by $X_{U(j)}$ if $Y_j > Y_{j-1}$, $j > 1$. Let $X_{U(1)}, X_{U(2)}, X_{U(3)}, \cdots, X_{U(n)}$ be the first upper record valuse of size n arising from a sequence $\{X_i\}$ of i.i.d Lomax variables with the probability density function pdf

$$f(x) = \alpha\beta^{\alpha}(x+\beta)^{-(\alpha+1)}, x \geq 0, \alpha, \beta > 0. \quad (1)$$

and cumulative distribution function cdf

$$F(x) = 1 - \beta^{\alpha}(x+\beta)^{-\alpha}, x \geq 0, \alpha, \beta > 0, \quad (2)$$

where β is the scale parameter and α is the shape parameter.

The rest of the paper is organized as follows. In Section 2, give a brief description of Markov chain Monte Carlo (MCMC). MLE and parametric bootstrap confidence interval are discussed in Section 3 and 4. Section 5 describes Bayes estimates and construction of credible intervals using the MCMC techniques. Section 6 contains the analysis of a real life data set to illustrate our proposed method. A simulation studies are reported in order to give an assessment of the performance of the different estimation methods in Section 7. Finally we conclude with some comments in Section 8.

2. MARKOV CHAIN MONTE CARLO TECHNIQUES

Markov chain Monte Carlo (MCMC) methods use computer simulation of Markov chains in the parameter space Gilks et al. [14] and Gamerman [11]. The Markov chains are defined in such a way that the posterior distribution in the given statistical inference problem is the asymptotic distribution. This allows to use ergodic averages to approximate the desired posterior expectations. Several standard approaches to define such Markov chains exist, including Gibbs sampling, Metropolis-Hastings and reversible jump, see for example Metropolis et al. [20] and Hastings [17] Using these algorithms it is possible to implement posterior simulation in essentially any problem which allow pointwise evaluation of the prior distribution and likelihood function.

2.1 Gibbs sampler

The Gibbs sampling algorithm is one of the simplest Markov Chain Monte Carlo algorithms. It was introduced by Geman and Geman [13]. The paper by Gelfand and Smith [12] helped to demonstrate the value of the Gibbs algorithm for a range of problems in Bayesian analysis. The scheme can best be described in the following steps:
Algorithm 1

1. Choose an arbitrary starting point $\Phi^{(0)} = (\Phi_1^{(0)}, \cdots, \Phi_d^{(0)})$ for which $g\left(\Phi^{(0)}\right) > 0$.

2. Generate $\Phi_1^{(t)}$ from conditional distribution $g(\Phi_1 \mid \Phi_2^{(t-1)}, \Phi_3^{(t-1)}, \cdots, \Phi_d^{(t-1)})$.

3. Generate $\Phi_2^{(t)}$ from conditional distribution $g(\Phi_2 \mid \Phi_1^{(t)}, \Phi_3^{(t-1)}, \cdots, \Phi_1^{(t-1)})$.

4. Finally, generate $\Phi_d^{(t)}$ from conditional distribution $g(\Phi_d \mid \Phi_1^{(t)}, \Phi_2^{(t)}, \Phi_3^{(t)}, \cdots, \Phi_{d-1}^{(t-1)})$.

5. Repeat steps 2-4.

Often we treat the initially generated values as burn-in values which are to be discarded. Usually, one will discard the first 1000 value or so.

2.2 The Metropolis-Hastings algorithm

Suppose that our goal is to draw samples from some distribution $f\left(\Phi|x\right) = \nu g\left(\Phi\right)$, where ν is the normalizing constant which may not be known or very difficult to compute. The Metropolis-Hastings (M-H) algorithm provides a way of sampling from $f\left(\Phi|x\right)$ without requiring us to know ν. Let $q\left(\Phi^{(b)}|\Phi^{(a)}\right)$ be an arbitrary transition kernel, that is the probability of moving or jumping from current state $\Phi^{(a)}$ to $\Phi^{(b)}$. This is sometimes called the proposal distribution. The following algorithm will generate a sequence of values $\Phi^{(1)}, \Phi^{(2)} \cdots$, which form a Markov Chain with stationary distribution given by $f\left(\Phi|x\right)$.
Algorithm 2

1. Choose an arbitrary starting point $\Phi^{(0)}$ for which $f\left(\Phi^{(0)}|x\right) > 0$.

2. At time t, sample a candidate point or proposal Φ^* from the proposal distribution $q\left(\Phi^*|\Phi^{(t-1)}\right)$.

3. Calculate the acceptance probability

$$\rho\left(\Phi^{(t-1)}, \Phi^*\right) = \min\left[1, \frac{f\left(\Phi^*|x\right) q\left(\Phi^{(t-1)}|\Phi^*\right)}{f\left(\Phi^{(t-1)}|x\right) q\left(\Phi^*|\Phi^{(t-1)}\right)}\right]. \quad (3)$$

4. Generate $U \sim U\left(0,1\right)$.

5. If $U \leq \rho\left(\theta^{(t-1)}, \Phi^*\right)$ accept the proposal and set $\Phi^{(t)} = \Phi^*$. Otherwise, reject the proposal and set $\Phi^{(t)} = \Phi^{(t-1)}$.

6. Repeat steps 2-5.

If the proposal distribution is symmetric, so $q\left(\Phi|\Psi\right) = q\left(\Psi|\Phi\right)$ for all possible Ψ and Φ then, in particular, we have $q\left(\Phi^{(t-1)}|\Phi^*\right) = q\left(\Phi^*|\Phi^{(t-1)}\right)$, so that the acceptance probability (3) is given by

$$\rho\left(\Phi^{(t-1)}, \Phi^*\right) = \min\left[1, \frac{f\left(\Phi^*|x\right)}{f\left(\Phi^{(t-1)}|x\right)}\right]. \quad (4)$$

3. ESTIMATION OF THE PARAMETERS

This section, estimate α and β, by considering the maximum likelihood and compute the observed Fisher information based on the likelihood equations. These will enable us to develop pivotal quantities based on the limiting normal distribution, the resulting pivotal quantities can be used to develop interval estimates.

3.1 Maximum likelihood estimation (MLE)

Suppose that $\underline{x} = x_{U(1)}, x_{U(2)}, ..., x_{U(n)}$ be the first upper record values of size n from Lomax (α, β). The likelihood function for observed record $\underline{x}$ given by see Arnold et al. [6]

$$\ell(\alpha, \beta|\underline{x}) = f(x_{u(n)}) \prod_{i=1}^{n-1} \frac{f(x_{u(i)})}{1 - F(x_{u(i)})}, \quad (5)$$

where $f(.)$ and $F(.)$ are given, respectively, by (1) and (2) substituting $f(.)$ and $F(.)$ in the likelihood function obtain

$$\ell(\alpha, \beta|\underline{x}) = \alpha^n \beta^\alpha (x_{u(n)} + \beta)^{-\alpha} \prod_{i=1}^{n} (x_{u(i)} + \beta)^{-1}. \quad (6)$$

The log-likelihood function may then be written as

$$\begin{aligned} L(\alpha, \beta|\underline{x}) &= \log \ell(\alpha, \beta|\underline{x}) \\ &= n \log \alpha + \alpha \log \beta - \alpha \log\left(x_{u(n)} + \beta\right) \\ &- \sum_{i=1}^{n} \log\left(x_{u(i)} + \beta\right), \end{aligned} \quad (7)$$

Upon differentiating (7) with respect to α, and β, and equating each result to zero, two equations must be simultaneously satisfied to obtain MLE of $\widehat{\alpha}$ and $\widehat{\beta}$, The maximum likelihood equations of α, and β can be obtained as the solution of

$$\frac{\partial L(\alpha, \beta|\underline{x})}{\partial \alpha} = \frac{n}{\alpha} + \log \beta - \log\left(x_{u(n)} + \beta\right), \quad (8)$$

and

$$\frac{\partial L(\alpha, \beta|\underline{x})}{\partial \beta} = \frac{\alpha}{\beta} - \frac{\alpha}{\left(x_{u(n)} + \beta\right)} - \sum_{i=1}^{n} \frac{1}{\left(x_{u(i)} + \beta\right)}. \quad (9)$$

Solving $\frac{\partial L(\alpha, \beta|\underline{x})}{\partial \alpha} = 0$ for α gives, from (8)

$$\widehat{\alpha} = \frac{n}{\log\left(x_{u(n)} + \beta\right) - \log \beta} \quad (10)$$

Using (10) in (9) we obtain

$$\begin{aligned} &\frac{n}{\beta\left[\log\left(x_{u(n)} + \beta\right) - \log \beta\right]} - \sum_{i=1}^{n} \frac{1}{\left(x_{u(i)} + \beta\right)} \\ &- \frac{n}{\left(x_{u(n)} + \beta\right)\left[\log\left(x_{u(n)} + \beta\right) - \log \beta\right]} = 0. \end{aligned} \quad (11)$$

Since (11) cannot be solved analytically some numerical methods such as Newton-Raphson iteration scheme must be employed.to solve (11) and get the MLE, $\widehat{\beta}$, and hence $\widehat{\alpha}$, by using the equation

$$\widehat{\alpha} = \frac{n}{\log\left(x_{u(n)} + \widehat{\beta}\right) - \log \widehat{\beta}} \quad (12)$$

3.2 Observed Fisher information

The asymptotic variances and covariances of the MLE for parameters α, and β are given by elements of the inverse of the Fisher information matrix

$$\mathbf{I}_{ij} = E\left[-\frac{\partial^2 L}{\partial \alpha \partial \beta}\right];\ i,\ j = 1, 2. \quad (13)$$

Unfortunately, the exact mathematical expressions for the above expectations are very difficult to obtain. Therefore, we give the approximate (observed) asymptotic varaince-covariance matrix for the MLE, which is obtained by dropping the expectation operator E

$$\begin{bmatrix} -\dfrac{\partial^2 L(\alpha,\beta|\underline{x})}{\partial\alpha^2} & -\dfrac{\partial^2 L(\alpha,\beta|\underline{x})}{\partial\alpha\partial\beta} \\ -\dfrac{\partial^2 L(\alpha,\beta|\underline{x})}{\partial\beta\partial\alpha} & -\dfrac{\partial^2 L(\alpha,\beta|\underline{x})}{\partial\beta^2} \end{bmatrix}^{-1}_{(\widehat{\alpha},\widehat{\beta})} = \begin{bmatrix} var(\widehat{\alpha}) & cov(\widehat{\alpha},\widehat{\beta}) \\ cov(\widehat{\beta},\widehat{\alpha}) & var(\widehat{\beta}) \end{bmatrix},$$

with

$$\frac{\partial^2 L(\alpha,\beta|\underline{x})}{\partial\alpha^2} = -\frac{n}{\alpha^2}, \tag{14}$$

$$\frac{\partial^2 L(\alpha,\beta|\underline{x})}{\partial\alpha\partial\beta} = \frac{\partial^2 L(\alpha,\beta|\underline{x})}{\partial\beta\partial\alpha} = \frac{1}{\beta} - \frac{1}{\left(x_{u(n)}+\beta\right)}, \tag{15}$$

$$\frac{\partial^2 L(\alpha,\beta|\underline{x})}{\partial\beta^2} = \frac{-\alpha}{\beta^2} + \frac{1}{\left(x_{u(n)}+\beta\right)^2} + \sum_{i=1}^{n} \frac{1}{\left(x_{u(i)}+\beta\right)^2}. \tag{16}$$

The asymptotic normality of the MLE can be used to compute the approximate confidence intervals for parameters α, and β. Therefore, $(1-\gamma)100\%$ confidence intervals for parameters α, and β become

$$\widehat{\alpha} \pm Z_{\gamma/2}\sqrt{var(\widehat{\alpha})} \quad \text{and} \quad \widehat{\beta} \pm Z_{\gamma/2}\sqrt{var(\widehat{\beta})}, \tag{17}$$

where $Z_{\gamma/2}$ is the percentile of the standard normal distribution with right-tail probability $\gamma/2$.

4. BOOTSTRAP CONFIDENCE INTERVALS

This subsection, propose to use confidence intervals based on the parameteric bootstrap methods (i) percentile bootstrap method (Boot-p) based on the idea of Efron [10]. (ii) bootstrap-t method (Boot-t) based on the idea of Hall [16]. The algorithms for estimating the confidenc eintervals using both methods are illustrated as follows.

4.1 Percentile bootstrap method

Algorithm 3

1. From the original data $\underline{x} = x_{U(1)}, x_{U(2)}, ..., x_{U(n)}$ compute the ML estimates of the parameters $\widehat{\alpha}$ and $\widehat{\beta}$ by equations (11) and (12).
2. Use $\widehat{\alpha}$ and $\widehat{\beta}$ to generate a bootstrap sample $\underline{x}^* = x^*_{U(1)}, x^*_{U(2)}, ..., x^*_{U(n)}$.
3. As in step 1, based on $\underline{x}^*$ compute the bootstrap sample estimates of α and β, say $\widehat{\alpha}^*$and $\widehat{\beta}^*$.
4. Repeat steps 2-3 N times representing N bootstrap MLE's of (α,β) based on N different bootstrap samples.
5. Arrange all $\widehat{\alpha}^{*\prime}s$ and $\widehat{\beta}^{*\prime}s$, in an ascending order to obtain the bootstrap sample $(\varphi_l^{[1]}, \varphi_l^{[2]}, ..., \varphi_l^{[N]})$, $l = 1, 2$ (where $\varphi_1 \equiv \widehat{\alpha}^*, \varphi_2 \equiv \widehat{\beta}^*$).

Let $G(z) = P(\varphi_l \leq z)$ be the cumulative distribution function of φ_1.Define $\varphi_{lboot} = G^{-1}(z)$ for given z. The approximate bootstrap $100(1-\gamma)\%$ confidence interval of φ_l is given by

$$[\varphi_{lboot}(\frac{\gamma}{2}), \varphi_{lboot}(\frac{1-\gamma}{2})].$$

4.2 Bootstrap-t method

Algorithm 4

1. From the original data $\underline{x} = x_{U(1)}, x_{U(2)}, ..., x_{U(n)}$ compute the ML estimates of the parameters: $\widehat{\alpha}$ and $\widehat{\beta}$ by equations (11) and (12).

2. Using $\widehat{\alpha}$ and $\widehat{\beta}$ generate abootstrap sample $\{x_1^*, x_2^*, ..., x_n^*\}$. Based on these data, compute the bootstrap estimate of α and β using (11 and 12), say $\widehat{\alpha}^*$and $\widehat{\beta}^*$ and following statistics

$$T_1^* = \frac{\sqrt{n}(\widehat{\alpha}^* - \widehat{\alpha})}{\sqrt{Var(\widehat{\alpha}^*)}} \text{ and } T_2^* = \frac{\sqrt{n}(\widehat{\beta}^* - \widehat{\beta})}{\sqrt{Var(\widehat{\beta}^*)}}$$

where $Var(\widehat{\alpha}^*)$ and $Var(\widehat{\beta}^*)$ are obtained using the Fisher information matrix.

3. Repeat step 2, N boot times.

4. For the T_1^* and T_2^* values obtained in step 2, determine the upper and lower bounds of the $100(1-\gamma)\%$ confidence interval of α and β as follows: let $H(x) = P(T_i^* \leq x), i = 1, 2$ be the cumulative distribution function of T_1^* and T_2^*. For a given x, define

$$\widehat{\alpha}_{Boot-t}(x) = \widehat{\alpha} + n^{-1/2}\sqrt{Var(\widehat{\alpha})}H^{-1}(x)$$
and
$$\widehat{\beta}_{Boot-t}(x) = \widehat{\beta} + n^{-1/2}\sqrt{Var(\widehat{\beta})}H^{-1}(x).$$

Here also, $Var(\widehat{\alpha})$ and $Var(\widehat{\beta})$ can be computed as same as computing the $Var(\widehat{\alpha}^*)$ and $Var(\widehat{\beta}^*)$. The approximate $100(1-\gamma)\%$ confidence interval of α and β are given by

$$\left[\widehat{\alpha}_{Boot-t}(\frac{\gamma}{2}), \widehat{\alpha}_{Boot-t}(1-\frac{\gamma}{2})\right]$$
and
$$\left[\widehat{\beta}_{Boot-t}(\frac{\gamma}{2}), \widehat{\beta}_{Boot-t}(1-\frac{\gamma}{2})\right].$$

5. BAYES ESTIMATION USING MCMC

The use of Bayesian MCMC to obtain inferences for population parameters is attractive because it: (i) allows a full Bayesian analysis of the data including appropriate descriptions of the uncertainty in any function of the parameters, (ii) it allows a wide range of descriptions of the joint distribution of uncertainties in the data including the magnitude of historical peaks and exceedance thresholds, and (iii) the numerical procedures are relative straightforward without significant increase in complexity or computational effort with additional descriptions of data uncertainty, nor does it require that the parameter space be represented by a grid for numerical integration. The Bayesian approach is introduced and its computational implementation with MCMC algorithms is described. Gibbs sampling procedure and the MH method are used to generate samples from the posterior density function and in turn compute the Bayes point estimates and also construct the corresponding credible intervals based on the generated posterior samples. Considering model (1), assume the following gamma prior densities for α and β as

$$\pi_1(\alpha|a,b) = \begin{cases} \dfrac{b^a}{\Gamma(a)}\alpha^{a-1}\exp(-b\alpha) & \text{if } \alpha > 0 \\ 0 & \text{if } \alpha \leq 0. \end{cases} \tag{18}$$

and

$$\pi_2(\beta|c,d) = \begin{cases} \frac{d^c}{\Gamma(c)}\beta^{c-1}\exp(-d\beta) & \text{if } \beta > 0 \\ 0 & \text{if } \beta \le 0. \end{cases} \quad (19)$$

Multiplying $\pi_1(\alpha|a,b)$ by $\pi_2(\beta|c,d)$ we obtain the joint prior density of α and β; given by

$$\pi(\alpha,\beta) = \frac{b^a d^c}{\Gamma(a)\Gamma(c)}\alpha^{a-1}\beta^{c-1}\exp(-b\alpha - d\beta) \quad (20)$$

Based on the likelihood function of the observed sample is same as (6) and the joint prior in (20), the joint posterior density of α and β given the data is

$$\pi^*(\alpha,\beta|\underline{x}) = \frac{\ell(\alpha,\beta|\underline{x}) \times \pi(\alpha,\beta)}{\int_0^\infty\int_0^\infty \ell(\alpha,\beta|\underline{x}) \times \pi(\alpha,\beta)d\alpha d\beta}, \quad (21)$$

therefore, the Bayes estimate of any function of α and β say $g(\alpha,\beta)$, under squared error loss function is

$$\begin{aligned}\tilde{g}(\alpha,\beta) &= E_{\alpha,\beta|data}(g(\alpha,\beta)) \\ &= \frac{\int_0^\infty\int_0^\infty g(\alpha,\beta)\ell(\alpha,\beta|\underline{x})\pi(\alpha,\beta)d\alpha d\beta}{\int_0^\infty\int_0^\infty \ell(\alpha,\beta|\underline{x})\pi(\alpha,\beta)d\alpha d\beta}. \quad (22)\end{aligned}$$

Generally, the ratio of two integrals given by (22) can not be obtained in a closed form. In this case, use the MCMC method to generate samples from the posterior distributions and then compute the Bayes estimator of $g(\alpha,\beta)$ under the squared errors loss (SEL) function. Therefore, we propose the approaches of MCMC technique to approximate (22) . See, for example, (Robert and Casella [27]) and Recently, (Rezaei et al. [26], Soliman et al. [28].

5.1 MCMC approach

Consider the MCMC method to generate samples from the posterior distributions and then compute the Bayes estimates of the parameters α and β under the squared errors loss (SEL) function. A wide variety of MCMC schemes are available, and it can be difficult to choose among them. An important subclass of MCMC methods are Gibbs sampling and more general Metropolis-within-Gibbs samplers. The advantage of using the MCMC method over the MLE method is that we can always obtain a reasonable interval estimate of the parameters by constructing the probability intervals based on the empirical posterior distribution. This is often unavailable in maximum likelihood estimation. Indeed, the MCMC samples may be used to completely summarize the posterior uncertainty about the parameters and , through a kernel estimate of the posterior distribution. This is also true of any function of the parameters.
The expression for the joint posterior can be obtained up to proportionality by multiplying the likelihood with the joint prior and this can be written as

$$\begin{aligned}&\pi^*(\alpha,\beta) \propto \alpha^{n+a-1}\beta^{\alpha+c-1}\times \\ &\exp[-(b\alpha + d\beta + \alpha\log(x_{u(n)}+\beta)]\prod_{i=1}^{n}\left(x_{u(i)}+\beta\right)^{-1} \quad (23)\end{aligned}$$

from (23) it is clear that the posterior density function of α given β is proportional to

$$\pi_1^*(\alpha|\beta) \propto \alpha^{n+a-1}\exp -\alpha\left[b + \log(x_{u(n)}+\beta) - \log\beta\right]. \quad (24)$$

Therefore, the posterior density function of α given β, is gamma with parameters $(n+a)$ and $\left(b + \log(x_{u(n)}+\beta) - \log\beta\right)$ and, therefore, samples of α can be easily generated using any gamma generating routine.

The posterior density function of β given α can be written as

$$\begin{aligned}&\pi_2^*(\beta|\alpha) \propto \beta^{\alpha+c-1}\times \\ &\exp\left[-d\beta - \alpha\log(x_{u(n)}+\beta) - \sum_{i=1}^{n}\log(x_{u(i)}+\beta)\right] \quad (25)\end{aligned}$$

The posterior distribution of β given α Eq. (25) cannot be reduced analytically to well known distributions and therefore it is not possible to sample directly by standard methods, but the plot of it (see Figure. 1) show that it is similar to normal distribution. So to generate random numbers from this distribution, we use the Metropolis-Hastings method with normal proposal distribution. The choice of the hyperparameters (a,b,c,d) which make (25) close to the proposal distribution and obviously more convergence of the MCMC iteration. We propose the following MCMC algorithm to draw samples from the posterior density functions; and in turn compute the Bayes estimates and also, construct the corresponding credible intervals.

Algorithm 5

1. $\beta_0 = \widehat{\beta}$, $M = nburn$.

2. Generate α_1 from gamma distribution $\pi_1^*(\alpha|\beta)$.

3. Generate β_1 from $\pi_2^*(\beta|\alpha)$ using (MH) algorithm in subsection (2.2).

4. Compute $\alpha^{(t)}$ and $.\beta^{(t)}$.

5. Repeat steps 2-4 N times.

6. Obtain the Bayes estimates of α and β with respect to the SEL function as

$$\widehat{E}(\alpha|data) = \frac{1}{N-M}\sum_{i=M+1}^{N}\alpha_i. \quad ,.$$

$$\widehat{E}(\beta|data) = \frac{1}{N-M}\sum_{i=M+1}^{N}\beta_i. \quad .$$

7. To compute the credible intervals of α and β, $\alpha_1,...,\alpha_N$ order and $\beta_1,...,\beta_N$ as $\alpha_{(1)} < ... < \alpha_{(N)}$ and $\beta_{(1)} < ... < \beta_{(N)}$. Then the $100(1-\gamma)\%$ symmetric credible intervals of α and β become

$$(\alpha_{(N\frac{\gamma}{2})}, \alpha_{(N(1-\frac{\gamma}{2}))}) \text{ and } (\beta_{(N\frac{\gamma}{2})}, \beta_{(N(1-\frac{\gamma}{2}))}).$$

6. ILLUSTRATIVE EXAMPLE

To illustrate the inferential procedures developed in the preceding sections, we choose the real data set which was also used in Lawless (1982-pp 185). These data are from (Nelson [22]) concerning the data on time to breakdown of an insulating fluid between electrodes at a voltage of 34 k.v. (minutes). The 19 times to breakdown are

0.96	4.15	0.19	0.78	8.01	31.75	7.35
6.50	8.27	33.91	32.52	3.16	4.85	2.78
4.67	1.31	12.06	36.71	72.89		

Therefore, we observe the upper record values from the observed data as follows: 0.96, 4.15, 8.01, 31.75, 33.91, 36.71, 72.89. Amodel suggested by engineering considerations is that, for a fixed voltage level, time to breakdown has a Lomax distribution. Based on these seven upper record values, we compute the approximate MLEs, Bootstrap (Boot-p, Boot-t) and Bayes estimates of α and β using MCMC method, since we do not have any prior information availiable, we used noninformative priors $(a = b = c = d = 0)$ on both α and β.The density function of $\pi_2^*(\beta|\alpha)$ as given in (25) is plotted Figure 1**.** It can be approximated by normal distribution function as mentioned in Subsection 5.1. Also the 95%, approximate maximum likelIhood estimation (AMLE) confidence intervals, Bootstrap confidence intervals and approximate credible intervals based on the MCMC samples, the results are given in Table 1. The plot of histogram of α and β generated by MCMC method are given in Figures 2 and 3. This was done with 1000 bootstrap sample and 10 000 MCMC sample and discard the first 1000 values as 'burn-in'.

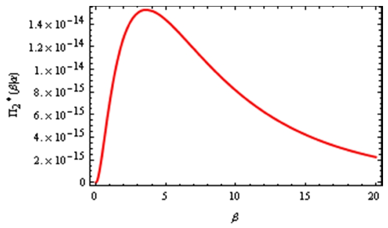

Fig. 1. Posterior density function of β given α

Table 1. Results obtained by MLE, Bootstrap and MCMC method of α and β.

Method	(α, β)	Point	Interval	Length
MLEs	α	3.0448	[-1.6138, 7.7034]	9.3172
	β	8.1311	[-19.686, 35.949]	55.6348
Boot-p	α	2.9237	[1.28610, 6.4212]	5.1351
	β	9.2354	[2.0133, 18.9971]	16.9838
Boot-t	α	2.8397	[3.36570, 9.6159]	6.2502
	β	8.8782	[2.7398, 19.6096]	16.8698
MCMC	α	2.8864	[0.76820, 6.9523]	6.1841
	β	8.2531	[0.1675, 26.9346]	26.7671

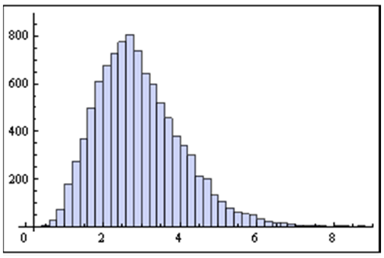

Fig. 2. Histogram of Alpha generated by MCMC method

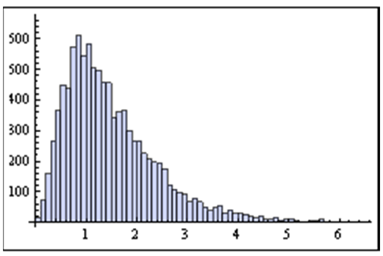

Fig. 3. Histogram of Beta generated by MCMC method

7. SIMULATION STUDY

To evaluate the behavior of the proposed methods, we simulated 1000 upper record samples from a Lomax distribution.
Use different sample of sizes (n), different hyperparameters (a, b, c, d) and three sets of parameter values $(\alpha, \beta) = (2.2)$, $(1, 2)$, $(1, 1)$ mainly to compare the MLEs and different Bayes estimators and also to explore their effects on different parameter values. First, use the noninformative gamma priors for both the parameters, that is, when the hyperparameters are 0. Call it prior 0: $a = b = c = d = 0$. Note that as the hyperparameters go to 0, the prior density becomes inversely proportional to its argument and also becomes improper. This density is commonly used as an improper prior for parameters in the range of 0 to infinity, and this prior is not specifically related to the gamma density. For computing Bayes estimators, other than prior 0, also used informative prior, including prior 1, $a = 2$, $b = 1$, $c = 2$ and $d = 1$, we used the squared error loss function to compute the Bayes estimates. Also computed the Bayes estimates and 95% credible intervals based on 10 000 MCMC samples and discard the first 1000 values as 'burn-in'. Average Bayes estimates, mean squared errors (MSEs) and coverage percentages. For comparison purposes, we also compute the MLEs and the 95% confidence intervals based on the observed Fisher information matrix. Finally, use the same 1000 replicates to compute different estimates Tables 2-4 report the results based on MLEs and the Bayes estimators (using both the Gibbs sampling procedure) using non-informative prior and informative prior on both α and β.

Table 2. Average values of the different estimators, the corresponding MSEs and coverage percentages when $(\alpha, \beta) = (2, 2)$.

n	MLE		MCMC(Prior 0)		MCMC(Prior 1)	
	α	β	α	β	α	β
5	1.8676	2.3165	1.7207	2.3148	1.7668	2.1863
	0.5432	0.3713	0.5348	0.3051	0.2149	0.1801
	0.951	0.970	0.955	0.997	0.985	0.998
7	1.9983	2.4212	1.8743	2.4492	1.8693	2.1855
	0.5287	0.3326	0.5171	0.3024	0.2102	0.1778
	0.960	0.970	0.975	0.985	0.975	0.987
9	2.0783	2.3297	1.9918	2.4593	1.9553	2.1544
	0.5106	0.3205	0.4888	0.3005	0.2069	0.1730
	0.935	0.980	0.965	0.986	0.990	0.985
12	2.1487	2.3616	2.0917	2.546	2.0441	2.1774
	0.5007	0.3093	0.4639	0.2972	0.2018	0.1696
	0.970	0.975	0.970	0.998	0.990	0.977
15	2.0757	2.1450	2.0448	2.4396	2.0301	2.1337
	0.4432	0.2456	0.4430	0.2330	0.2000	0.1675
	0.955	0.939	0.965	0.999	0.970	0.993
18	2.0913	2.2884	2.0616	2.5471	2.0471	2.1752
	0.3029	0.2303	0.2774	0.2259	0.1869	0.1380
	0.945	0.960	0.945	0.997	0.950	0.999
20	2.0999	2.2256	2.0734	2.4950	2.0652	2.1500
	0.2444	0.2044	0.2288	0.2001	0.1595	0.1345
	0.965	0.965	0.970	0.999	0.981	0.977
23	2.1479	2.3117	2.1252	2.6028	2.1044	2.1957
	0.2396	0.1998	0.2144	0.1833	0.1455	0.1335
	0.975	0.980	0.975	0.963	0.980	0.999
25	2.0209	2.0786	2.0076	2.4299	2.0125	2.0993
	0.1960	0.1658	0.1864	0.1534	0.1104	0.1035
	0.980	0.940	0.980	0.997	0.982	0.973

Note: The first figure represents the average estimates,with the corresponding MSEs and coverage percentages reported below it in parentheses.

8. CONCLUSION

In this paper consider the Bayes estimation of the unknown parameters of the Lomax distribution when the data are upper record values. We assume the gamma priors on the unknown parameters and provide the Bayes estimators under the assumptions of squared error loss functions. It is observed that the Bayes estimators can not be obtained in explicit forms and they can be obtained using the numerical integration. Because of that we have used MCMC technique to generate posterior sample. we observe the following.

(i) From the results obtained in Tables 2 and 4. It can be seen that the performance of the Bayes estimators with respect to the noninformative prior (prior 0) is quite close to that of the MLEs.

(ii) Tables 2-4 report the results based on noninformative prior (prior 0) and informative prior, (prior 1) also in these case the results based on using the Gibbs sampling procedure are quite similar in nature when comparing the Bayes estimators based on informative prior clearly shows that the Bayes estimators based on prior 1 perform better than the MLEs, in terms of MSEs.

(iii) From Tables 2-4, it is clear that the Bayes estimators based on informative prior perform much better than noninformative prior and the MLEs in terms of MSEs.

Table 3. Average values of the different estimators, the corresponding MSEs and coverage percentages when $(\alpha, \beta) = (1, 2)$.

n	MLE		MCMC(Prior 0)		MCMC(Prior 1)	
	α	β	α	β	α	β
5	0.9803	2.2879	0.8836	2.1301	1.1004	2.279
	0.1920	0.2534	0.1433	0.2215	0.1117	0.1787
	0.935	0.985	0.955	0.999	0.996	0.995
7	1.0592	2.2741	0.9907	2.2714	1.1493	2.2668
	0.1453	0.2384	0.1141	0.2156	0.1106	0.1584
	0.965	0.990	0.975	0.986	0.995	0.988
9	1.0759	2.1478	1.0293	2.2829	1.1531	2.2360
	0.1316	0.2249	0.1108	0.2092	0.1073	0.1231
	0.960	0.995	0.980	0.966	0.990	0.993
12	1.0558	2.1032	1.0236	2.2877	1.1238	2.2016
	0.1142	0.2193	0.1060	0.2031	0.0939	0.1177
	0.980	0.990	0.981	0.995	0.994	0.974
15	1.0687	2.3804	1.0369	2.3984	1.1135	2.227
	0.0964	0.1871	0.0863	0.1750	0.0818	0.1011
	0.975	0.990	0.970	0.988	0.965	0.979
18	1.023	2.2201	0.9993	2.3238	1.0678	2.1784
	0.0860	0.1618	0.0759	0.1591	0.0729	0.1004
	0.970	0.985	0.971	0.966	0.955	0.964
20	1.0174	2.3071	0.9961	2.3551	1.0593	2.202
	0.0727	0.1498	0.0651	0.1385	0.0612	0.0966
	0.967	0.958	0.963	0.982	0.947	0.961
23	1.009	2.1160	0.9899	2.1267	1.0429	2.3568
	0.0621	0.1263	0.0541	0.1180	0.0498	0.0913
	0.958	0.964	0.947	0.963	0.934	0.967
25	0.9989	1.9931	0.9958	2.167	0.9799	2.13
	0.0492	0.1040	0.0349	0.1010	0.0325	0.0869
	0.964	0.975	0.947	0.981	0.968	0.992

Note: The first figure represents the average estimates,with the corresponding MSEs and coverage percentages reported below it in parentheses.

Table 4. Average values of the different estimators, the correspondingMSEs and coverage percentages when $(\alpha, \beta) = (1, 1)$.

n	MLE		MCMC(Prior 0)		MCMC(Prior 1)	
	α	β	α	β	α	β
5	1.1804	1.1832	1.1168	1.1987	1.0987	1.1592
	0.2165	0.2207	0.1978	0.2094	0.1694	0.1968
	0.925	0.995	0.945	0.960	0.962	0.974
7	1.0902	1.0398	1.0561	1.0356	1.0275	1.1005
	0.2077	0.2171	0.1969	0.2022	0.1670	0.1893
	0.965	0.995	0.975	0.998	0.985	0.999
9	1.1776	1.8380	1.1396	1.0661	1.2678	1.1074
	0.2017	0.2024	0.1888	0.1985	0.1506	0.1754
	0.985	0.995	0.965	0.958	0.965	0.991
12	1.1023	1.687	1.0791	1.9906	1.1873	1.0587
	0.1835	0.1928	0.1636	0.1759	0.1479	0.1616
	0.960	0.995	0.955	0.974	0.950	0.988
15	1.1076	1.2317	1.0874	1.0586	1.1748	1.0663
	0.1552	0.1847	0.1508	0.1681	0.1260	0.1387
	0.965	0.985	0.955	0.949	0.955	0.963
18	1.0670	1.3539	1.0502	1.3696	1.1277	1.0213
	0.1264	0.1679	0.1103	0.1543	0.1091	0.1135
	0.965	0.985	0.940	0.966	0.945	0.978
20	1.0586	1.3334	1.0438	1.2404	1.1151	1.2933
	0.0973	0.1360	0.0618	0.1282	0.0599	0.0995
	0.960	0.985	0.975	0.996	0.970	0.999
23	1.0782	1.1933	1.0647	1.0137	1.1255	1.026
	0.0668	0.1102	0.0525	0.1052	0.0449	0.0891
	0.961	0.990	0.950	0.963	0.945	0.973
25	1.0141	1.2294	1.0513	1.2461	1.0101	1.0042
	0.0391	0.0926	0.0385	0.0899	0.0356	0.0688
	0.956	0.968	0.964	0.978	0.963	0.986

Note: The first figure represents the average estimates,with the corresponding MSEs and coverage percentages reported below it in parentheses.

Acknowledgement
The authors would like to express their thanks to the editor and referees for their useful comments and suggestions on the original version of this manuscript.

9. REFERENCES

[1] A. H. Abd Ellah, *Bayesian one sample prediction bounds for the Lomax distribution*, Indian J. Pure Appl. Math. **42** (2011), no. 5, 335–356.

[2] A. H. Abd Ellah, *Comparison of estimates using record statstics from Lomax model : Bayesian and non Bayesian approaches*, J. Stat. Res. Train. Center **3** (2006), 139-158.

[3] M. Ahsanullah, *On the record valuse from univariate distributions*, National Institute of Standards and Technology Journal of Research, Special Publications **12** (1993), 1-6.

[4] M. Ahsanullah, *Introduction to Record Statistics*, NOVA Science Publishers Inc., Hunting ton, New York, 1995.

[5] B. C. Arnold, *Pareto Distributions. In: Statistical Distributions in Scientific Work*, International Co-operative Publishing House, Burtonsville, MD, 1983.

[6] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja, *Records*, Wiley, New York, 1998.

[7] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja, *A first Course in Statistics*, Wiley, New York, 1992.

[8] M. C. Bryson, *Heavy-tailed distributions: properties and tests*, Technometrics **16** (1974), no. 1, 61–68.

[9] M. Chahkandi and M. Ganjali, *On some lifetime distributions with decreasing failure rate*, Comput. Statist. Data Anal. **53** (2009), no. 12, 4433–4440.

[10] B. Efron, *The bootstrap and other resampling plans*, In: CBMS-NSF Regional Conference Seriesin Applied Mathematics, SIAM, Philadelphia, PA, 1982.

[11] D. Gamerman, *Markov chain Monte Carlo, Stochastic Simulation for Bayesian Inference*, Chapman and Hall, London, 1997.

[12] A. E. Gelfand and A. F. M. Smith, *Sampling based approach to calculating marginal densities*, J. Amer. Statist. Asso. **85** (1990), 398–409.

[13] S. Geman and D. Geman, *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*, IEEE Trans. Pattern Anal. Mach. Intell. **6** (1984), 721–741.

[14] W. R. Gilks, S. Richardson and D. J. Spiegelhalter, *Markov chain Monte Carlo in Practices*, Chapman and Hall, London, 1996.

[15] M. Habibullah, and M. Ahsanullah, *Estimation of parameters of a Pareto distribution by generalized order statstics*, Comm. Statist. Theory Methods **29** (2000), 1597-1609.

[16] P. Hall, *Theoretical comparison of bootstrap confidence intervals*, Ann. Stat. **16** (1988), 927-953.

[17] W. K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika **57** (1970), 97–109.

[18] K. S. Lomax, *Business failure: Another example of the analysis of the failure data*, JASA **49** (1954), 847-852.

[19] A. W. Marshall and I. Olkin, *Life Distributions. Structure of Nonparametric, Semiparametric, and Parametric Families*, Springer, New York, NY, 2007.

[20] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *Equations of state calculations by fast computing machines*, J. Chem. Phy. **21** (1953), 1087–1091.

[21] H. N. Nagaraja, *Record values and related statistics-a review*, Comm. Statist. Theory Methods **17** (1988), 2223-2238.

[22] W. B. Nelson, *Applied Life Data Analysis*, Wiley, New York, 1982.

[23] M. Z. Raqab, *Inferences for generalized exponential distribution based on record Statistics*, J. Statist. Plann. Inference **52** (2002), 339-350.

[24] M. Z. Raqab and M. Ahsanulla, *Estimation of the location and scale parameters of the generalized exponential distribution based on order Statistics*, J. Statis Comput Simul. **69** (2001), 109-124.

[25] S. I. Resnick, *Extreme values, Regular variation, and point processes*, Springer-Verlag,New York, 1987.

[26] S. Rezaei, R. Tahmasbi and M. Mahmoodi, *Estimation of P[Y¡X] for generalized Pareto distribution,* J. Statist. Plann. Inference **140** (2010), 480-494.

[27] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Second edition, Springer, New York, 2004.

[28] A. A. Soliman, A. H. Abd Ellah, N. A. Abou-Elheggag, and A. A. Modhesh, *Bayesian inference and prediction of Burr type XII distribution for progressive first-failure censored sampling,* Intel. Info. Mana. **3** (2011), 175-185.

[29] S. K. Upadhyay and M. Peshwani, *Choice between Weibull and lognormal models: a smulation based Bayesian study*, Comm. Statist. Theory Methods **32** (2003), 381-405.