Mathematics and Statistics 10(3): 615-623, 2022
DOI: 10.13189/ms.2022.100318 http://www.hrpub.org

Likelihood and Bayesian Inference in the Lomax Distribution under Progressive Censoring

A. Baklizi*, A. Saadati Nik, A. Asgharzadeh

[1]Department of Mathematics, Statistics and Physics, Qatar University, Doha, Qatar
[2]Department of Statistics, University of Mazandaran, Babolsar, Iran

Received February 18, 2022; Revised April 27, 2022; Accepted May 23, 2022

Cite This Paper in the following Citation Styles

(a): *[1] A. Baklizi, A. Saadati Nik, A. Asgharzadeh , "Likelihood and Bayesian Inference in the Lomax Distribution under Progressive Censoring," Mathematics and Statistics, Vol. 10, No. 3, pp. 615 - 623, 2022. DOI: 10.13189/ms.2022.100318.*

(b): *A. Baklizi, A. Saadati Nik, A. Asgharzadeh (2022). Likelihood and Bayesian Inference in the Lomax Distribution under Progressive Censoring. Mathematics and Statistics, 10(3), 615 - 623. DOI: 10.13189/ms.2022.100318.*

Abstract The Lomax distribution has been used as a statistical model in several fields, especially for business failure data and reliability engineering. Accurate parameter estimation is very important because it is the base for most inferences from this model. In this paper, we shall study this problem in detail. We developed several points and interval estimators for the parameters of this model assuming the data are type II progressively censored. Specifically, we derive the maximum likelihood estimator and the associated Wald interval. Bayesian point and interval estimators were considered. Since they can't be obtained in a closed form, we used a Markov chains Monte Carlo technique, the so called the Metropolis – Hastings algorithm to obtain approximate Bayes estimators and credible intervals. The asymptotic approximation of Lindley to the Bayes estimator is obtained for the present problem. Moreover, we obtained the least squares and the weighted least squares estimators for the parameters of the Lomax model. Simulation techniques were used to investigate and compare the performance of the various estimators and intervals developed in this paper. We found that the Lindley's approximation to the Bayes estimator has the least mean squared error among all estimators and that the Bayes interval obtained using the Metropolis – Hastings to have better overall performance than the Wald intervals in terms of coverage probabilities and expected interval lengths. Therefore, Bayesian techniques are recommended for inference in this model. An example of real data on total rain volume is given to illustrate the application of the methods developed in this paper.

Keywords Bayesian Inference, Likelihood Inference, Lomax Distribution, Metropolis-Hastings Algorithm, Point Estimator, Progressive Censoring

Mathematics Subject Classification: 62F10; 62F15; 62N01; 62N02

1. Introduction

The Lomax distribution was introduced by Lomax [22] as a model for business failure data. It provides an alternative model when the experimenter assumes that the population is heavy-tailed [8]. It is found to be useful in several areas including business, economics, actuarial science, queuing theory and internet traffic modeling. It belongs to the class of decreasing failure rate distributions as discussed by Chahkandi and Ganjali [9].

The Lomax distribution is related to several important distributions in the literature. With a suitable choice of parameters, it is a special case of the Pareto Type II distribution. Similarly, it can be shown to be a special case of the generalized Pareto distribution, the Beta prime distribution, and the q-exponential distribution among several other distributions, see [16]. The Lomax distribution has been extended in various ways in the literature. For example, some variants of the Lomax distribution may have more than the basic two parameters. For example, the beta exponentiated Lomax distribution contains five parameters [25]. There are many other variants of the Lomax distribution including the Exponential Lomax [12], the Gamma Lomax [10], the Poisson Lomax [3], Weibull Lomax [30], Weighted Lomax [17]. There are several other variants that arise from time to time to provide suitable models for the phenomena under study.

The applications and some properties of the Lomax distribution can be found in [4] and [16]. The pdf and cdf of the underlying life time distribution are given respectively by

$$f(x,\theta,\beta)=\frac{\beta\,\theta}{(1+\beta x)^{\theta+1}},\ \ x>0,\beta>0,\theta>0, \quad (1)$$

$$F(x,\theta,\beta)=1-\frac{1}{(1+\beta x)^{\theta}},\ \ x>0,\beta>0,\theta>0. \quad (2)$$

Various authors have considered inference for this distribution with progressively censored data in the literature. Cramer et al. [11] considered estimation based on competing risks data from Lomax distribution. Likelihood and Bayesian estimation of the scale parameter were considered by Asgharzadeh and Valiollahi [5]. Likelihood estimation using Newton-Raphson and the EM algorithm was considered by Helu et al. [15]. Likelihood and Bayesian point estimation of the stress-strength reliability in this model were considered by Al-Zahrani and Al-Sobhi [2]. A similar problem for the log-logistic distribution with adaptive progressively censored data was considered by Sewailem and Baklizi [26], while Sewailem and Baklizi [27] considered modified profile likelihood estimation of the scale parameter of the Lomax distribution.

In this paper, we will extend the work of Helu et al. [15] on maximum likelihood estimation and the work of Al-Zahrani and Al-Sobhi [2] on Bayesian point estimation to include approximate Bayesian estimation based on the Lindley approximation and Metropolis-Hastings algorithm, least squares and weighted least squares estimators. We included a bias and mean squared error comparison of the various estimators as well. Furthermore, we considered likelihood and Bayesian interval estimation for the parameters of this distribution using the Metropolis-Hastings algorithm. The likelihood and Bayesian intervals were compared in terms of their coverage probabilities and expected lengths using a simulation study.

The organization of the paper is as follows. In Section 2 we will consider likelihood inference for the parameters. Least squares and weighted least squares estimators are derived in Section 3. Bayesian inference is considered in Section 4, while in Section 5 we describe a simulation study designed to investigate and compare the performance of the likelihood and Bayesian inference procedures. The results and findings are given in Section 6. An illustrative example based on a real data set is given in the final Section.

2. Likelihood Inference with Progressive Censoring

Censoring is used to reduce the time and cost of the life experiment. In a progressively type II censored sampling scheme with n units and m failures [6], a group of n independent and identical units are put on test. At the i^{th} observed failure time, a predetermined number R_i of surviving units is randomly removed from the experiment. This manner will continue until the time of the last failure. This censoring scheme is denoted by $(R_1,R_2,\ldots,R_m)$ where $m+\sum_{i=1}^{m}R_i=n$. Consider a progressive type II censoring plan $(R_1,R_2,\ldots,R_m)$ for n tested units with m failures. Let $Y_{1:m:n}\le Y_{2:m:n}\le\cdots\le Y_{m:m:n}$ be the progressive type II censored sample obtained from this plan and $y_1\le y_2\le\cdots\le y_m$ represent the observed progressive type II censored sample. The likelihood function is given by [6]

$$L(\theta,\beta)=C\prod_{i=1}^{m}f\,(y_i,\theta,\beta)\,(1-F(y_i,\theta,\beta))^{R_i}, \quad (3)$$

where $C=n(n-1-R_1)(n-2-R_1-R_2)\ldots(n-m+1-R_1\ldots-R_{m-1})$. For the Lomax distribution, it will be

$$\begin{aligned}L(\theta,\beta)&=C\prod_{i=1}^{m}\frac{\beta\theta}{(1+\beta y_i)^{\theta+1}}\left(\frac{1}{(1+\beta y_i)^{\theta}}\right)^{R_i}\\&=C\beta^m\,\theta^m\prod_{i=1}^{m}(1+\beta y_i)^{-(\theta(R_i+1)+1)}.\end{aligned} \quad (4)$$

We derive the log-likelihood function as

$$l(\theta,\beta)=\ln L(\theta,\beta)=m\ln\theta+m\ln\beta-\sum_{i=1}^{m}(\theta(R_i+1)+1)\ln(1+\beta y_i). \quad (5)$$

The maximum likelihood estimators (MLEs) of parameters are obtained by solving the likelihood equations with respect to θ and β. We have

$$\frac{\partial l(\theta,\beta)}{\partial\theta}=\frac{m}{\theta}-\sum_{i=1}^{m}(R_i+1)\ln(1+\beta y_i)=0. \quad (6)$$

From (6), the MLE of θ, say $\hat{\theta}$ is

$$\hat{\theta}=\frac{m}{\sum_{i=1}^{m}(R_i+1)\ln(1+\beta y_i)}.$$

Substituting $\hat{\theta}$ into (5), the MLE of β is the solution of the equation

$$\begin{aligned}h(\beta)=\frac{\partial l(\hat{\theta},\beta)}{\partial\beta}&=\frac{m}{\beta}\\&-\frac{m}{\sum_{i=1}^{m}(R_i+1)\ln(1+\beta y_i)}\sum_{i=1}^{m}\frac{(R_i+1)y_i}{1+\beta y_i}\\&-\sum_{i=1}^{m}\frac{y_i}{1+\beta y_i}=0.\end{aligned}$$

These equations can’t be solved analytically for the MLE $\hat{\beta}$ and we have to apply some numerical technique to solve the likelihood equation and compute the estimate $\hat{\beta}$. It is easy to show that the MLE $\hat{\beta}$ can be derived as a fixed-point solution of the equation $\Lambda(\beta)=\beta$, where

$$\Lambda(\beta)=\frac{m}{\frac{m}{\sum_{i=1}^{m}(R_i+1)\ln(1+\beta y_i)}\sum_{i=1}^{m}\frac{(R_i+1)y_i}{1+\beta y_i}+\sum_{i=1}^{m}\frac{y_i}{1+\beta y_i}}.$$

A simple iterative technique $\lambda(\beta^{(j)})=\beta^{(j)}$, where $\beta^{(j)}$ is the jth iterative, can be used to solve $\Lambda(\beta)=\beta$. The codes of R software (R Core Development Team) can also be used to solve $\Lambda(\beta)=\beta$.

To find the approximate confidence intervals for $S(t)$ for large m, we need to find the observed Fisher Information matrix of parameters θ and β therefore we find the matrix of minus the second partial derivatives of the log-likelihood function as follows:

$$J(\theta,\beta)=\begin{bmatrix}-\frac{\partial l(\theta,\beta)}{\partial\theta^2} & -\frac{\partial l(\theta,\beta)}{\partial\theta\partial\beta}\\ -\frac{\partial l(\theta,\beta)}{\partial\beta\partial\theta} & -\frac{\partial l(\theta,\beta)}{\partial\beta^2}\end{bmatrix}, \qquad (7)$$

where $\frac{\partial l(\theta,\beta)}{\partial\theta^2}=-\frac{m}{\theta^2}$, $\frac{\partial l(\theta,\beta)}{\partial\theta\partial\beta}=-\sum_{i=1}^{m}\frac{(R_i+1)y_i}{1+\beta y_i}$, $\frac{\partial l(\theta,\beta)}{\partial\beta^2}=-\frac{m}{\beta^2}+\sum_{i=1}^{m}\frac{(\theta(R_i+1)+1)y_i^2}{(1+\beta y_i)^2}$. Then the observed Fisher information matrix is given by $J(\hat{\theta},\hat{\beta})$ is is the variance-covariance matrix of the MLE. Therefore the approximate confidence intervals for θ and β are given respectively by

$$(\hat{\theta}-z_{\alpha/2}\sqrt{V\hat{a}r\,(\hat{\theta})},\hat{\theta}+z_{\alpha/2}\sqrt{V\hat{a}r\,(\hat{\theta})}), \qquad (8)$$

$$(\hat{\theta}-z_{\alpha/2}\sqrt{V\hat{a}r\,(\hat{\theta})},\hat{\theta}+z_{\alpha/2}\sqrt{V\hat{a}r\,(\hat{\theta})}), \qquad (9)$$

where z_q is the q-th percentile of the standard normal distribution. Since θ and β are positive parameters and the lower bound of the above confidence intervals can be less than zero, the modified confidence intervals can be suggested as follows:

$$(\max\{0,\hat{\theta}-z_{\alpha/2}\sqrt{V\hat{a}r\,(\hat{\theta})}\},\hat{\theta}+z_{\alpha/2}\sqrt{V\hat{a}r\,(\hat{\theta})}), \qquad (10)$$

$$(\max\{0,\hat{\beta}-z_{\alpha/2}\sqrt{V\hat{a}r\,(\hat{\beta})}\},\hat{\beta}+z_{\alpha/2}\sqrt{V\hat{a}r\,(\hat{\beta})}). \qquad (11)$$

To find an approximate confidence interval for $S(t)$, we need to use the multivariate delta method given in the following theorem [19];

Theorem 1: The Multivariate Delta Method

Suppose that $\left(\sqrt{n}(U_n-\xi),\sqrt{n}(V_n-\tau)\right)\xrightarrow{L}N(\mathbf{0},\mathbf{\Sigma})$, where $N(\mathbf{0},\mathbf{\Sigma})$ denotes the bivariate normal distribution with mean $(0,0)$ and covariance matrix $\mathbf{\Sigma}$. Let g be a real valued function that admits a Taylor expansion at the point (ξ,τ), then the distribution of $\sqrt{n}\big(g(U_n,V_n)-g(\xi,\tau)\big)\xrightarrow{L}N(0,\upsilon)$, where $\upsilon=G^t\,\Sigma^{-1}\,G$ and $G=(\frac{\partial g(\xi,\tau)}{\partial\xi},\frac{\partial g(\xi,\tau)}{\partial\tau})$.

Proof: The proof can be found in [19].

Applying the multivariate delta method with $U_n=\hat{\theta}$, $V_n=\hat{\beta}$, $\xi=\theta$, $\tau=\beta$, $\mathbf{\Sigma}=\boldsymbol{J}$, $g(\xi,\tau)=S(t)$, $g(U_n,V_n)=\hat{S}(t)\ v=var\left(\hat{S}(t)\right)$, then the multivariate delta method gives

$$\upsilon=\mathrm{var}(S\hat{(t)})\approx G^t\,J^{-1}(\theta,\beta)\,G$$

where $G=(\frac{\partial S(t)}{\partial\theta},\frac{\partial S(t)}{\partial\beta})$, $\frac{\partial S(t)}{\partial\theta}=-(1+\beta x)^{-\theta}\ln(1+\beta x)$ and $\frac{\partial S(t)}{\partial\beta}=-\theta x(1+\beta x)^{-\theta-1}$.

The approximate estimate of the variance of $S\hat{(t)}$ is obtained by replacing (θ,β) by their MLEs as follows;

$$V\hat{a}r\,(S\hat{(t)})\approx[G^t\,J^{-1}(\theta,\beta)\,G]_{(\hat{\theta},\hat{\beta})}, \qquad (12)$$

Which can be used to find the approximate confidence interval for $S(t)$ given by

$$(S\hat{(t)}-z_{\alpha/2}\sqrt{V\hat{a}r\,(S\hat{(t)})},S\hat{(t)}+z_{\alpha/2}\sqrt{V\hat{a}r\,(S\hat{(t)})}). \qquad (13)$$

3. Least Squares and Weighted Least Squares Methods

Least squares (LS) and weighted least squares (WLS) methods were proposed by Swain et al. [29] to estimate the parameters of beta distribution. For the case of progressive censoring, we need the following theorem [1];

Theorem 2: Let $Y_{1:m:n},\dots,Y_{m:m:n}$ be a progressively type-II censored random sample of size m from the a continuous distribution with pdf and cdf $f(x)$ and $F(x)$ respectively, then, we have

$$E[F(Y_{i:m:n})]=1-\prod_{j=m-i+1}^{m}S_j,\quad i=1,\dots,m,$$

$$Var[F(Y_{i:m:n})]=\left(\prod_{j=m-i+1}^{m}S_j\right)\left(\prod_{j=m-i+1}^{m}L_j-\prod_{j=m-i+1}^{m}S_j\right),\quad i=1,\dots,m,$$

where, for $j=1,\dots,m$,

$$A_j=j+\sum_{q=m-j+1}^{m}R_q,\quad S_j=\frac{A_j}{1+A_j},\quad D_j=\frac{1}{(A_j+1)(A_j+2)},$$
$$\text{and } L_j=S_j+D_j.$$

Proof: See [1]

The LS estimates (LSEs) $\hat{\theta}_{LS}$ and $\hat{\beta}_{LS}$ of θ and β can be obtained by minimizing the following quantity with respect to θ and β.

$$\eta_1(\theta,\beta) = \sum_{i=1}^{m}[F(y_{i:m:n},\theta,\beta) - E[F(Y_{i:m:n})]]^2 = \sum_{i=1}^{m}\left[\frac{1}{(1+\beta y_{i:m:n})^{\theta}} - \prod_{j=m-i+1}^{m} S_j\right]^2.$$

The WLS estimates (WLSEs) $\hat{\theta}_{WLS}$ and $\hat{\beta}_{WLS}$ of θ and β can be obtained by minimizing $\eta_2(\theta,\beta)$ with respect to θ and β:

$$\eta_2(\theta,\beta) = \sum_{i=1}^{m}\omega_i\,[F(y_{i:m:n},\theta,\beta) - E[F(Y_{i:m:n})]]^2 = \sum_{i=1}^{m}\omega_i\left[\frac{1}{(1+\beta y_{i:m:n})^{\theta}} - \prod_{j=m-i+1}^{m} S_j\right]^2,$$

where ω_i is the weight factor given by

$$\omega_i = \frac{1}{Var[F(Y_{i:m:n})]}.$$

The last two quantities can be minimized by solving the equations $\partial\eta_1/\partial\theta = 0$, $\partial\eta_1/\partial\beta = 0$ and $\partial\eta_2/\partial\theta = 0$, $\partial\eta_2/\partial\beta = 0$, with respect to θ and β.

4. Bayesian Inference

For Bayesian inference, a joint prior for (θ,β) is needed. By assuming independent gamma prior distributions for θ and β, we obtain the following joint prior for (θ,β):

$$\pi(\theta,\beta) \propto \theta^{a_1-1}\,\beta^{a_2-1}\,e^{-(b_1\theta+b_2\beta)}, \qquad \theta>0, \beta>0,$$

where the hyper-parameters a_1, b_1, a_2, b_2 are positive constants. For $a_1 = b_1 = a_2 = b_2 = 0$, the joint prior reduces to

$$\pi(\theta,\beta) \propto \frac{1}{\theta\beta}, \;\; \theta>0, \beta>0,$$

which is an improper joint prior. Independent gamma prior distributions were used by [2] and they obtained the Bayes estimators of the stress-strength $R = P_r(Y < X)$ for Lomax distribution under symmetric and asymmetric balanced loss functions.

In this section, we shall investigate the Bayesian inference methods under above joint prior to assuming the squared error loss function (SELF). By combining the likelihood function and joint prior, the joint posterior density will be

$$\pi(\theta,\beta|\mathbf{y}) \propto \theta^{m+a_1-1}\,\beta^{m+a_2-1}\,e^{-(b_1\theta+b_2\beta)}\prod_{i=1}^{m}(1+\beta y_i)^{-(\theta(R_i+1)+1)},$$

where $\mathbf{y} = (y_1,\ldots,y_m)$. We have the following theorem whose proof can be found in most tests on Bayesian statistics.

Theorem 3: The Bayes estimates of a function $h(\theta,\beta)$ based on the SELF is given by the posterior expectation of this function

$$h(\hat{\theta,\beta})_B = \mathbb{E}(h(\theta,\beta)|\mathbf{y}) = \frac{\int_0^{\infty}\int_0^{\infty} h\,(\theta,\beta)\pi(\theta,\beta|\mathbf{y})\,d\theta\,d\beta}{\int_0^{\infty}\int_0^{\infty}\pi\,(\theta,\beta|\mathbf{y})\,d\theta\,d\beta}.$$

It is not possible to compute the Bayes estimator in (16) analytically, therefore we use the Lindley's approximation and the Metropolis-Hastings algorithm, see Hamada et al. [14]. The Metropolis-Hasting algorithm is illustrated below; we will use it to obtain random observations from a target posterior distribution assuming lognormal candidate distributions for both parameters θ and β.

4.1. The Metropolis Algorithm

- Initialize $j = 0, \theta^{(j)} = 0, \beta^{(j)} = 1$,
- Set $j = 1$,
- Draw θ^* and β^* from a bivariate gamma candidate distribution with independent components,
- Compute the acceptance probability $r = \min(1, \frac{\pi(\theta^*,\beta^*)}{\pi(\theta^{(j-1)},\beta^{(j-1)})}\frac{f(\theta^{(j-1)},\beta^{(j-1)})}{f(\theta^*,\beta^*)})$, where $f(u,v)$ is the pdf of the gamma candidate distribution.
- Draw u from a uniform (0,1) distribution,
- If $u \leq r$, set $(\theta^{(j)},\beta^{(j)}) = (\theta^*, \beta^*)$. Otherwise, set $(\theta^{(j)},\beta^{(j)}) = (\theta^{(j-1)},\beta^{(j-1)})$,
- Set $j = j + 1$,
- 8- Repeat steps 3 to 7 for $N = 11000$ times,
- Calculate the Bayes estimates of θ and β using $\hat{\theta}_{BMH} = \frac{1}{N-M}\sum_{i=M+1}^{N}\theta^{(i)}$ and $\hat{\beta}_{BMH} = \frac{1}{N-M}\sum_{i=M+1}^{N}\beta^{(i)}$, where M is the number of burn-in samples,
- Substitute $\theta^{(i)}$ and $\beta^{(i)}$ in $S(t) = 1 - \frac{1}{(1+\beta t)^{\theta}}$, to compute $S^{(1)}(t), S^{(2)}(t), \dots, S^{(N)}(t)$,
- A Bayesian estimate of $S(t)$ can be found as $\hat{S}_{BMH}(t) = \frac{\sum_{i=M+1}^{N} S^{(i)}(t)}{N-M}$,
- Denote the ordered values of $\theta^{(M+1)}, \theta^{(M+2)}, \dots, \theta^{(N)}$ and $\beta^{(M+1)}, \beta^{(M+2)}, \dots, \beta^{(N)}$ by $\theta_1, \dots, \theta_{N-M}$ and $\beta_1, \dots, \beta_{N-M}$. The $\%100(1-\alpha)$ symmetric credible intervals of θ and β calculated as $(\theta_{\left((N-M)(\frac{\alpha}{2})\right)}, \theta_{\left((N-M)(1-\frac{\alpha}{2})\right)})$ and $(\beta_{\left((N-M)(\frac{\alpha}{2})\right)}, \beta_{\left((N-M)(1-\frac{\alpha}{2})\right)})$,
- Order $\boldsymbol{S^{(M+1)}(t), S^{(M+2)}(t), \dots, S^{(N)}(t)}$ as $\boldsymbol{S_1(t) \leq S_2(t) \leq \cdots \leq S_{N-M}(t)}$. Then calculate the $\boldsymbol{\%100(1-\alpha)}$ symmetric credible interval for $\boldsymbol{S(t)}$ as $(\boldsymbol{S_{\left((N-M)(\frac{\alpha}{2})\right)}(t), S_{\left((N-M)(\frac{1-\alpha}{2})\right)}(t)})$.

4.2. Lindley's Approximation Approach

The Lindley's approximation was originally proposed by Lindley [20] to approximate the ratio of integral in the posterior expectation. This method has been used in the literature to approximate the Bayes estimator, see, for example, Lindley [21] and Press [23]. Based on Lindley's approximation, the approximate Bayes estimates of θ and β under SELF are

$$\hat{\theta}_{BL} = \hat{\theta} + \frac{1}{2}[2\rho_\theta\,\sigma_{\theta\theta} + 2\rho_\beta\,\sigma_{\theta\beta} + \\ +\sigma_{\theta\theta}\left(L_{\theta\theta\theta}\sigma_{\theta\theta} + 2L_{\theta\beta\theta}\sigma_{\theta\beta} + L_{\beta\beta\theta}\sigma_{\beta\beta}\right) + \\ +\sigma_{\beta\theta}(L_{\theta\theta\beta}\sigma_{\theta\theta} + 2L_{\theta\beta\beta}\sigma_{\theta\beta} + L_{\beta\beta\beta}\sigma_{\beta\beta})],$$

$$\hat{\beta}_{BL} = \hat{\beta} + \frac{1}{2}[2\rho_\theta\,\sigma_{\beta\theta} + 2\rho_\beta\,\sigma_{\beta\beta} + \\ +\sigma_{\theta\beta}\left(L_{\theta\theta\theta}\sigma_{\theta\theta} + 2L_{\theta\beta\theta}\sigma_{\theta\beta} + L_{\beta\beta\theta}\sigma_{\beta\beta}\right) + \\ +\sigma_{\beta\beta}(L_{\theta\theta\beta}\sigma_{\theta\theta} + 2L_{\theta\beta\beta}\sigma_{\theta\beta} + L_{\beta\beta\beta}\sigma_{\beta\beta})],$$

respectively. Here $\hat{\theta}$ and $\hat{\beta}$ are the MLEs of θ and β, respectively. The details are described in the appendix given at the end of the paper.

5. A Simulation Study

A simulation study is designed to investigate the performance of the estimators and intervals. The sample size (n), the number of failures (m) and the censoring scheme R are given in the following table. We also consider two sets of the true parameters values, $(\theta, \beta) = (0.5,0.5)$ and $(1.0,1.5)$. For each case we generated 3000 progressively type II censored samples using the procedure described by [7]. For the Metropolis-Hastings algorithm we used 11000 iterations, we ignored the first 1000 iterations as the Burn-in period and calculated the Bayes estimator and the Bayesian interval using the remaining 10000 iterations. For computing Bayes estimates, we use the informative prior: $a_1 = a_2 = 2$, $b_1 = b_2 = 0.5$. We calculated the biases and mean squared errors of the estimators. We also calculated the average length and the coverage probability for each interval.

Table 1. Progressive Censoring schemes

n	M	Censoring Schemes
40	30	$(0^{29},10), (5,0^{28},5), (10,0^{29})$
60	40	$(0^{39},20), (10,0^{28},10), (20,0^{39})$
100	60	$(0^{59},40), (20,0^{58},20), (40,0^{59})$

The results of the simulation study are given in Tables 2, 3 and 4. In the tables the bias of an estimator is denoted by $b(.)$, the MSE by $MS(.)$, the coverage probability of an interval for a parameter by $CP(.)$ and the expected length by $EL(.)$. Based on the tables, it appears that the MSE of all estimators of all parameters is the largest when the all the censoring occurs at the last failure, the special case corresponding to Type II censoring. It is smallest when all the censoring occurs at the first failure. The bias has almost a similar pattern in most cases. It appears also that the Lindley approximation of the Bayes estimator has generally the least MSE among all estimators. However, the least squares and the weighted least squares estimators appear to have the least bias among all estimators. The Metropolis-Hastings approximation appears to have the worst performance in terms of Bias and MSE in most cases. For interval estimation, it appears that the coverage probabilities of the Bayesian intervals obtained from the Metropolis-Hastings algorithm attain the nominal coverage probabilities in all cases considered. On the other hand, the Wald intervals for the scale parameter appear to be anti-conservative. Comparison of interval expected lengths will be limited to the shape parameter because Wald intervals for the scale parameter are highly anti-conservative and therefore invalid. It appears that the Bayesian intervals for the shape parameter to be shorter than the Wald intervals in almost all situations under study. Both intervals attain the nominal coverage probability in almost all cases. Therefore, we recommend the use of the

Bayesian intervals in all cases as they are generally shorter and attain the nominal coverage probabilities.

We observed that the sample size or the censoring configuration does not affect the relative performance of the likelihood and Bayes inference procedures. However, for much larger sample sizes, we expect that the MLE and Bayes estimators will have similar performance as anticipated from the large sample properties of both estimators [19].

Table 2. Biases and MSEs of the point estimators for $\theta=0.5$, $\beta=0.5$.

			θ					β				
(n, m)	Scheme		MLE	LSE	WLSE	Lindley	Metropolis	MLE	LSE	WLSE	Lindley	Metropolis
(40, 30)	$(0^{29}, 10)$	Bias	0.281	0.147	0.186	0.302	0.149	0.031	0.202	0.190	0.022	0.378
		MSE	2.981	2.310	4.735	3.219	0.556	0.187	0.374	0.390	0.171	0.439
	$(5, 0^{28}, 5)$	Bias	0.147	0.058	0.075	0.119	0.077	0.034	0.183	0.160	0.039	0.415
		MSE	0.801	0.779	1.030	0.268	0.251	0.165	0.310	0.297	0.154	0.479
	$(10, 0^{29})$	Bias	0.071	0.058	0.045	0.090	0.031	0.053	0.195	0.146	0.035	0.418
		MSE	0.051	0.903	0.288	0.228	0.029	0.153	0.319	0.243	0.136	0.459
(60, 40)	$(0^{39}, 20)$	Bias	0.237	0.091	0.119	0.221	0.130	0.015	0.144	0.138	0.019	0.292
		MSE	1.856	0.486	1.004	1.541	0.251	0.119	0.221	0.235	0.138	0.291
	$(10, 0^{38}, 10)$	Bias	0.102	0.030	0.036	0.082	0.064	0.024	0.143	0.125	0.031	0.303
		MSE	0.228	0.094	0.081	0.183	0.059	0.105	0.192	0.179	0.105	0.288
	$(20, 0^{39})$	Bias	0.054	0.033	0.039	0.052	0.026	0.021	0.120	0.080	0.034	0.328
		MSE	0.030	0.082	0.092	0.030	0.023	0.088	0.165	0.124	0.110	0.289
	$(0^{59}, 40)$	Bias	0.121	0.066	0.095	0.117	0.107	0.018	0.104	0.099	0.015	0.192
		MSE	0.374	0.245	0.484	0.322	0.111	0.091	0.142	0.159	0.085	0.168
	$(20, 0^{58}, 20)$	Bias	0.055	0.020	0.027	0.066	0.041	0.013	0.088	0.076	0.006	0.191
		MSE	0.033	0.030	0.034	0.054	0.029	0.063	0.100	0.098	0.067	0.139
	$(40, 0^{59})$	Bias	0.029	0.013	0.017	0.032	0.011	0.023	0.082	0.055	0.015	0.212
		MSE	0.013	0.018	0.015	0.015	0.011	0.056	0.091	0.072	0.052	0.140

Table 3. Biases and MSEs of the point estimators for $\theta=1.0$, $\beta=1.5$.

			θ					β				
(n, m)	Scheme		MLE	LSE	WLSE	Lindley	Metropolis	MLE	LSE	WLSE	Lindley	Metropolis
(40, 30)	$(0^{29}, 10)$	Bias	0.154	-0.035	-0.030	0.175	1.059	-0.194	0.116	0.069	-0.398	0.234
		MSE	0.326	0.272	0.266	0.336	7.772	1.000	1.117	1.122	0.813	1.021
	$(5, 0^{28}, 5)$	Bias	0.175	-0.006	0.016	0.168	0.708	-0.147	0.146	0.095	-0.274	0.286
		MSE	0.316	0.252	0.260	0.303	4.931	0.897	1.018	0.969	0.660	0.907
	$(10, 0^{29})$	Bias	0.194	0.017	0.052	0.180	0.325	-0.105	0.131	0.084	-0.209	0.436
		MSE	0.280	0.219	0.211	0.256	1.939	0.721	0.989	0.864	0.516	0.980
(60, 40)	$(0^{39}, 20)$	Bias	0.168	0.018	0.024	0.166	0.939	-0.212	0.060	-0.005	-0.364	0.172
		MSE	0.355	0.305	0.309	0.348	5.219	0.982	1.089	1.094	0.773	0.907
	$(10, 0^{38}, 10)$	Bias	0.177	0.032	0.053	0.178	0.581	-0.122	0.135	0.102	-0.251	0.266
		MSE	0.302	0.248	0.265	0.308	3.083	0.808	0.953	0.962	0.609	0.831
	$(20, 0^{39})$	Bias	0.154	0.040	0.058	0.154	0.213	-0.032	0.164	0.123	-0.143	0.360
		MSE	0.212	0.207	0.179	0.200	0.659	0.591	0.822	0.733	0.433	0.832
(100, 60)	$(0^{59}, 40)$	Bias	0.150	0.021	0.033	0.164	0.632	-0.128	0.119	0.059	-0.278	0.156
		MSE	0.319	0.279	0.305	0.341	1.903	0.875	0.998	1.017	0.706	0.747
	$(20, 0^{58}, 20)$	Bias	0.171	0.043	0.059	0.174	0.402	-0.062	0.154	0.113	-0.182	0.183
		MSE	0.265	0.223	0.236	0.278	0.933	0.664	0.826	0.816	0.512	0.628
	$(40, 0^{59})$	Bias	0.120	0.047	0.067	0.101	0.142	-0.017	0.164	0.106	-0.055	0.243
		MSE	0.131	0.154	0.138	0.120	0.139	0.433	0.682	0.561	0.334	0.528

Table 4. Simulated Coverage Probabilities Expected Lengths of Wald and Bayes Intervals

			$\theta = 0.5,\ \beta = 0.5$				$\theta = 1.0,\ \beta = 1.5$			
(n, m)	scheme	CI	$EL(\beta)$	$EL(\theta)$	$CP(\beta)$	$CP(\theta)$	$EL(\beta)$	$EL(\theta)$	$CP(\beta)$	$CP(\theta)$
(40, 30)	$(0^{29}, 10)$	Wald	1.236	1.296	0.841	0.956	4.444	4.587	0.934	0.948
		Bayes	2.084	1.047	0.897	0.911	3.976	3.268	0.824	0.828
	$(5, 0^{28}, 5)$	Wald	1.237	0.956	0.845	0.957	3.946	3.501	0.898	0.947
		Bayes	2.121	0.788	0.904	0.929	3.991	2.730	0.886	0.888
	$(10, 0^{29})$	Wald	1.189	0.681	0.855	0.961	3.634	2.314	0.872	0.954
		Bayes	2.084	0.592	0.920	0.943	4.060	1.877	0.937	0.938
(60, 40)	$(0^{39}, 20)$	Wald	1.167	1.178	0.866	0.952	4.338	4.618	0.956	0.922
		Bayes	1.836	1.077	0.908	0.919	3.747	3.372	0.821	0.831
	$(10, 0^{38}, 10)$	Wald	1.111	0.856	0.869	0.959	3.808	3.275	0.905	0.945
		Bayes	1.793	0.743	0.920	0.933	3.778	2.690	0.897	0.898
	$(20, 0^{39})$	Wald	1.089	0.561	0.876	0.966	3.264	1.854	0.871	0.956
		Bayes	1.767	0.510	0.917	0.943	3.581	1.593	0.942	0.948
(100, 60)	$(0^{59}, 40)$	Wald	1.046	0.991	0.882	0.949	3.997	4.308	0.958	0.915
		Bayes	1.421	1.001	0.925	0.935	3.536	3.217	0.871	0.873
	$(20, 0^{58}, 20)$	Wald	0.951	0.613	0.899	0.961	3.344	2.710	0.917	0.942
		Bayes	1.309	0.602	0.945	0.954	3.329	2.455	0.924	0.923
	$(40, 0^{59})$	Wald	0.888	0.419	0.896	0.954	2.684	1.353	0.885	0.959
		Bayes	1.257	0.396	0.935	0.936	2.883	1.253	0.958	0.954

6. An Example

To illustrate the use of the methods studied in this paper, we use the dataset given in Simpson [27]. The data includes measurements of the total rain volume in South Florida from cloud base following seeding penetration by the aircraft. The data contains 26 observations from seeded clouds and 26 observations from control clouds. This data was further discussed and analyzed in Giles et al. [13]. The suitability of the Lomax distribution to this data was checked by Helu et al. [15]. Here we shall consider the subset of measurements in the control group and impose the progressive type II censoring scheme $R = (3, 0^{18}, 3)$. The resulting data set is as follows:

Scheme	Censored data									
$(3,0^{18},3)$	0	17.3	21.7	24.4	26.1	26.3	28.6	29.0	36.6	41.1
	47.3	69.5	81.2	87.0	95.0	147.8	163.0	244.3	321.2	345.5

The maximum likelihood, LSE, WLSE and the Bayes estimates were obtained using the results obtained in this paper. These estimates and CI are presented in Tables 5 and 6. For computing Bayes estimates, since we don't have any prior information, we used zero values of the hyper-parameters on θ and β, i.e., $a_1 = a_2 = 0,\ b_1 = b_2 = 0$.

Table 5. Point estimates of the Lomax parameters based on the data set

	Parameter				
Estimator	θ	B	$S(58.7401)$	$S(26.84343)$	$S(7.276598)$
MLE	1.4882	0.0080	0.5633	0.7483	0.9191
LSE	1.1584	0.0112	0.5567	0.7374	0.9132
WLSE	1.4076	0.0081	0.5782	0.7581	0.9225
Lindley's approximation	1.4882	0.0080	0.5636	0.7486	0.9192
Metropolis-Hastings	1.6382	0.0084	0.5745	0.7533	0.9199

The results for point estimation appear similar between the MLE and the Bayesian estimators. The variance-covariance matrix of the MLE is given by

$$J^{-1} = \begin{bmatrix} 1.097705 & -0.0080054 \\ -0.0080054 & 0.00006493 \end{bmatrix}$$

This was used to construct likelihood based intervals (Wald intervals) for the parameters and the reliability function as explained earlier. The Bayesian intervals are obtained using the Metropolis-Hastings algorithm. We obtained the following results

Table 6. Interval estimates of the Lomax parameters based on the data set

	Parameter				
Interval	*Θ*	*B*	*S*(58.7401)	*S*(26.84343)	*S*(7.276598)
Likelihood	(0.0000, 3.5413)	(0.0000, 0.0238)	(0.3866, 0.7401)	(0.6070, 0.8898)	(0.8611, 0.9772)
Bayes	(0.7555, 3.2911)	(0.0026, 0.0189)	(0.4208, 0.7200)	(0.6330, 0.8520)	(0.8699, 0.9557)

Note that the Bayesian intervals are narrower than the corresponding Wald intervals. Since we found in the simulation study that they attain the nominal coverage probabilities, they are considered better than the corresponding likelihood intervals.

Appendix

An asymptotic expansion for evaluating the ratio of integrals of the form:

$$u^* = E[u(\theta,\beta)|\mathbf{y}] = \frac{\int\int u(\theta,\beta)exp[l(\theta,\beta|\mathbf{y}) + \rho(\theta,\beta)]}{\int\int exp[l(\theta,\beta|\mathbf{y}) + \rho(\theta,\beta)]}\mathrm{d}\theta\mathrm{d}\beta,$$

where $u(\theta,\beta)$ is a function of θ and β only, and $l(\theta,\beta|\mathbf{y})$ is the log-likelihood function of Eq. and $\rho(\theta,\beta) = \log[\pi(\theta,\beta)]$, has been proposed by Lindley . Utilizing Lindley's method, we can approximate u^* as

$$\begin{aligned} u^* = {} & u + \tfrac{1}{2}[(u_{\theta\theta} + 2u_\theta\rho_\theta)\sigma_{\theta\theta} + (u_{\theta\beta} + 2u_\theta\rho_\beta)\sigma_{\theta\beta} + (u_{\beta\theta} + 2u_\beta\rho_\theta)\sigma_{\beta\theta} \\ & + (u_{\beta\beta} + 2u_\beta\rho_\beta)\sigma_{\beta\beta} + (u_\theta\sigma_{\theta\theta} + u_\beta\sigma_{\theta\beta})(L_{\theta\theta\theta}\sigma_{\theta\theta} + L_{\theta\beta\theta}\sigma_{\theta\beta} + L_{\beta\theta\theta}\sigma_{\beta\theta} + L_{\beta\beta\theta}\sigma_{\beta\beta}) \\ & + (u_\theta\sigma_{\beta\theta} + u_\beta\sigma_{\beta\beta})(L_{\theta\theta\beta}\sigma_{\theta\theta} + L_{\theta\beta\beta}\sigma_{\theta\beta} + L_{\beta\theta\beta}\sigma_{\beta\theta} + L_{\beta\beta\beta}\sigma_{\beta\beta})]. \end{aligned} \tag{17}$$

All functions of the right-hand side of Eq. (17) are to be evaluated at the MLE $\hat{u}$ of u. In our setup, we have $\pi(\theta,\beta) \propto \theta^{a_1-1}\,\beta^{a_2-1}\,e^{-(b_1\theta+b_2\beta)}$ and

$$\rho(\theta,\beta) = \log[\pi(\theta,\beta)] = (a_1-1)\log\theta + (a_2-1)\log\beta - b_1\theta - b_2\beta.$$

This turns out that

$$\rho_\theta = \frac{d\rho}{d\theta} = \frac{a_1-1}{\theta} - b_1,\ \rho_\beta = \frac{a_2-1}{\beta} - b_2,\ u_\theta = \frac{du}{d\theta},\ u_\beta = \frac{du}{d\beta}$$

$$u_{\theta\beta} = \frac{d^2u}{d\theta d\beta},\ u_{\beta\theta} = \frac{d^2u}{d\beta d\theta},\ u_{\theta\theta} = \frac{d^2u}{d\theta d\theta},\ u_{\beta\beta} = \frac{d^2u}{d\beta d\beta}.$$

For $u(\theta,\beta) = \theta,\quad u_\theta = 1,\ u_{\theta\theta} = 0 = u_\beta = u_{\theta\beta} = u_{\beta\theta} = u_{\beta\beta}$, the Bayes estimate of θ under SELF loss function is

$$\begin{aligned} \hat{\theta}_{BL} = \hat{\theta} + {} & \frac{1}{2}[2\rho_\theta\,\sigma_{\theta\theta} + 2\rho_\beta\,\sigma_{\theta\beta} + \sigma_{\theta\theta}(L_{\theta\theta\theta}\sigma_{\theta\theta} + L_{\theta\beta\theta}\sigma_{\theta\beta} + L_{\beta\theta\theta}\sigma_{\beta\theta} + L_{\beta\beta\theta}\sigma_{\beta\beta}) \\ & + \sigma_{\beta\theta}(L_{\theta\theta\beta}\sigma_{\theta\theta} + L_{\theta\beta\beta}\sigma_{\theta\beta} + L_{\beta\theta\beta}\sigma_{\beta\theta} + L_{\beta\beta\beta}\sigma_{\beta\beta}], \end{aligned}$$

and when $u(\theta,\beta) = \beta,\quad u_\beta = 1,\quad u_{\beta\beta} = 0 = u_\theta = u_{\theta\theta} = u_{\beta\theta} = u_{\theta\beta}$, we obtain the Bayes estimate of β as

$$\begin{aligned} \hat{\beta}_{BL} = \hat{\beta} + {} & \frac{1}{2}[2\rho_\theta\,\sigma_{\beta\theta} + 2\rho_\beta\,\sigma_{\beta\beta} + \sigma_{\theta\beta}(L_{\theta\theta\theta}\sigma_{\theta\theta} + L_{\theta\beta\theta}\sigma_{\theta\beta} + L_{\beta\theta\theta}\sigma_{\beta\theta} + L_{\beta\beta\theta}\sigma_{\beta\beta}) \\ & + \sigma_{\beta\beta}(L_{\theta\theta\beta}\sigma_{\theta\theta} + L_{\theta\beta\beta}\sigma_{\theta\beta} + L_{\beta\theta\beta}\sigma_{\beta\theta} + L_{\beta\beta\beta}\sigma_{\beta\beta}]. \end{aligned}$$

In the above expressions $\sigma_{ij} = (i,j)$-th element in the inverse of the negative Hessian matrix, $i,j = \theta,\beta$, and L_{ijk} implies the term obtained from differentiating $l = \log L$ with respect to i,j and k.

REFERENCES

[1] Aggarwala, R., Balakrishnan, N., "Some properties of progressive censored order statistics from arbitrary and uniform distributions with applications to inference and simulation," Journal of Statistical Planning and Inference, vol. 70, no. 1, pp. 35 – 49, 1998.

[2] Al-Zahrani, B., Al-Sobhi, M., "On Parameters Estimation of Lomax Distribution under General Progressive Censoring," Journal of Quality and Reliability Engineering, vol. 2013, Article ID 431541, 7 pages, 2013. https://doi.org/10.1155/2013/431541.

[3] Al-Zahrani, B. and Sagor, H., "*The Poisson-Lomax distribution*," Revista Colombiana de Estadistica, vol. 37, no. 1, pp. 225 – 245, 2014.

[4] Arnold, B. C., "*The Pareto distributions*," International Co-operative Publishing House; 1983.

[5] Asgharzadeh, A., Valiollahi, R., "Estimation of the scale parameter of the Lomax distribution under progressive censoring," International Journal of Statistics and Economics, vol. 6, no. S11, pp. 37 – 48, 2011.

[6] Balakrishnan, N., Aggarwala, R., "Progressive censoring-theory, methods and applications," Birkhäuser; 2000.

[7] Balakrishnan, N., Sandhu, A., "A simple simulation algorithm for generating progressive Type-II censored sample," The American Statistician, vol. 49, no. 2, pp. 229–230, 1995.

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

[9] Chahkandi, M., Ganjali, M., "On some lifetime distributions with decreasing failure rate," Computational Statistics and Data Analysis, vol. 53, no. 12, pp. 4433–4440, 2009.

[10] Cordiero, Gauss, M., Ortiga, Edwin, M.M., and Popovic, Bozidar, V., "*The Gamma-Lomax Distribution*," Journal of Statistical Computation and simulation, vol. 85, no. 2, pp. 305 – 319, 2015.

[11] Cramer, E., Schmiedt, A. B., "Progressively type-II censored competing risks data from Lomax distributions," Computational Statistics and Data Analysis. Vol. 55, 1285–1303, 2011.

[12] El-Bassiouni, A.H., Abdo, N.F. and Shahen, H.S., "*Exponential Lomax distribution*," International Journal of Computer applications, vol. 121, no. 13, pp. 24 – 29, 2015.

[13] Giles, D., Feng, H., Gadwin, R., "On the bias of the maximum likelihood estimator for the two parameter Lomax distribution," Communications in Statistics, Theory and Methods, vol. 42, no. 11, pp. 1934–1950, 2013.

[14] Hamada, M. S., "*Bayesian Reliability*," Springer, 2008.

[15] Helu, A., Samawi, H., Raqab, M. Z., "Estimation on Lomax progressive censoring using the EM algorithm," *Journal of Statistical Computation and Simulation*, vol. 85, no. 5, pp. 1035–1052, 2005.

[16] Johnson, N., Kotz, S., Balakrishnan, N. *Continuous univariate distribution*. Vol. 1. New York: John Wiley and Sons; 1994.

[17] Kilani, N.M, "Weighted Lomax distribution," SpringerPlus, vol. 5: 1862, DOI 10.1186/s40064-016-3489-2, 2016.

[18] Kundu, D., Joarder, A., "Analysis of Type -II progressively hybrid censored data," Computational Statistics and Data Analysis, vol. 50, no. 10, pp. 2509–2528, 2006.

[19] Lehmann, E. L., "*Elements of Large Sample Theory*," Springer, 1999.

[20] Lindley, D. V., "Fiducial distributions and Bayes' theorem," Journal of the Royal Statistical Society, Series B, vol. 20, no. 1, pp. 102-107, 1958.

[21] Lindley, D. V., "Approximate Bayesian methods," Trabajos de Stadistca, vol. 31, no. 1, pp. 223-237, 1980.

[22] Lomax, K., "Business failures; another example of the analysis of failure data," Journal of the American Statistical Association. Vol. 49, no. 268, pp. 847–852, 1954.

[23] Press, S. J., "The Subjectivity of Scientists and the Bayesian Approach," Wiley, 2001.

[24] R Core Team, R: A language and environment for statistical computing, r foundation for statistical computing, Vienna, https://www.R-project.org, 2020.

[25] Rajab, M., Aleem, M., Nawaz, T. and Daniyal, M., "*On five parameter Beta Lomax distribution*," Journal of Statistics, vol. 20, pp. 102 – 118, 2013.

[26] Sewailem, Maha, and Baklizi, A., "Inference for the log-logistic distribution based on an adaptive progressive type-II censoring scheme," Cogent Mathematics & Statistics, vol. 6, no. 1, DOI: 10.1080/25742558.2019.1684228, 2019.

[27] Sewailem, Maisoun, and Baklizi, A., "Modified profile likelihood estimation in the Lomax distribution," Mathematics and Statistics, vol. 10, no. 2, pp. 383 – 389, 2022. DOI: 10.13189/ms.2022.100213.

[28] Simpson, J. "Use of the gamma distribution in single-cloud rainfall analysis," Monthly Weather Review, vol. 100, pp. 309–312, 1972.

[29] Swain, J. J., Venkatraman, S., Wilson, J.R., "Least-squares estimation of distribution function in Johnson's translation system," Journal of Statistical Computation & Simulation, vol. 29, pp. 271-297, 1988.

[30] Tahir, M., Cordiero, Gauss, M., Mansoor, M. and Zubair, M., "*The Weibull-Lomax distribution: properties and applications*," Hacettepe Journal of Mathematics and Statistics, vol. 44, no. 2, pp. 461 – 80, 2015.