COMPARISON OF MAXIMUM LIKELIHOOD ESTIMATION AND BAYESIAN ESTIMATION ON EXPONENTIATED POWER LOMAX DISTRIBUTION

S. Amirtha Rani Jagulin[1] and A. Venmani[2]

•

[1,2] Department of Mathematics and Statistics, College of Science and Humanities,
SRM Institute of Science and Technology,
Kattankulathur, Tamil Nadu, India
as4371@srmist.edu.in[1], venmania@srmist.edu.in[2]

Abstract

The aim of the paper is to apply the Bayesian estimation under squared error loss function to the Exponentiated Power Lomax (EPOLO) distribution to estimate the parameters and then compare with maximum likelihood estimation. The reliability of the distribution is analyzed by computing survival and hazard function for the exponentiated power lomax distribution. The mean square error will help to compare the different estimates like Bayesian and maximum likelihood estimation to decide the best one. Bayesian estimation and the maximum likelihood estimation are discussed for the distribution. A simulation study is done using the R programming software to generate random values and estimate the parameters for different n = 15, 30, 50, 100 and parameter values taken as 0.5 and 1.5. At n=100 and c= 0.5 the mean square error of bayes and mle are same and survival and hazard function mean square error are decreases when the sample size increases, this indicates the distribution is fitted good in all areas.

Keywords: Exponentiated Power Lomax (EPOLO) distribution, Bayesian estimation, Reliability, Mean Square Error (MSE), Monte Carlo simulation.

1. Introduction

According to El-Monsef [12], the EPOLO distribution performs superior fits than several well-known distributions to the data such as the number of ball bearing revolutions, the lung cancer patient's tumor size, and the total COVID-19 deaths in Egypt. The EPOLO (Exponentiated Power Lomax) distribution is an extension of the POLO (Power Lomax) distribution. By exponentiating the POLO cumulative distribution function to positive real power c we get this distribution. The CDF of the EPOLO distribution is given by

$$F(x) = (1-\gamma^{\alpha}\left(\gamma+x^{\beta}\right)-\alpha)^{c} \quad x>0; \alpha,\beta,\gamma,c\ >\ 0 \tag{1}$$

The PDF of EPOLO distribution is given by

$$f(x) =\ \alpha\beta c\gamma^{\alpha}x^{\beta-1}\left(\ \gamma+x^{\beta}\right)^{-\alpha-1}(1-\gamma^{\alpha}(\gamma+x^{\beta})^{-\alpha})^{c-1} \quad x\ >\ 0;\ \alpha,\beta,\gamma,c\ >\ 0 \tag{2}$$

When $c=\beta=1$, it will result to Lomax distribution and if $c = 1$, then (2) will leads to POLO distribution. Lomax [10] has proposed by Lomax distribution is a mixture of gamma and exponential distributions. It is also denoted as Pareto II distribution. Applying the Lomax distribution to simulate lifetime data was common. It is helpful for income modeling, business failure data, and biological issues. The random variable X follows a L(α,γ) distribution, the cumulative distribution function (CDF) is given as

$$F(x) = 1 - (\frac{x}{\gamma} + 1)^{-\alpha} \qquad x > 0; \alpha, \gamma > 0 \tag{3}$$

The probability density function (PDF) is given as

$$f(x) = \frac{\alpha}{\gamma}\left(\frac{x}{\gamma} + 1\right)^{-\alpha-1} x > 0; \alpha, \gamma > 0 \tag{4}$$

Lingappaiah [9] created different ways of estimation procedures for Lomax distribution. Myhre and Saunders [13] were described right censored data were subjected to Lomax distribution. According to Hassan [5], Lomax distribution can be utilized for dependability modeling and life testing. Balakrishnan and Ahsanullah [3] investigated the record moments and distribution properties of the Lomax distribution. Kilany [7] discussed the weighted Lomax distribution. Lemonte and Cordeiro [8] used the Lomax distribution to study its expansion. This paper is divided as follows: section 2 provides the information about the Bayesian analysis with prior, likelihood, posterior function and also the parameter estimation of the parameter c. Section 3, derivation of posterior of the maximum likelihood function using the likelihood function, section 4 provides the knowledge about the reliability measures, survival and hazard function. Section 5, describes the quantile function to generate random numbers. In section 6 and 7 a simulation study is applied and analyzed for EPOLO distribution, section 8 gives the conclusion about the result.

2. Bayesian estimation

The Reverend Thomas Bayes (1701–1761), who used a subjective method to quantify probability, is the topic of the philosophy known as Bayesian. The most famous accomplishment of Bayes was never published. Richard Price updated his notes and published them after his death (1763). The statistician using Bayesian, he is free to interpret probability as a frequency and a level of belief or a function that calculated with the mathematical principles of probability, depending on which interpretation best fits the task. The Bayesian technique gives the option of adding earlier knowledge of the pertinent parameters. Mahmoud et al. [11] and Tierney [17] introduces the Bayesian concept and describes how it is computationally implemented using MCMC algorithms. Amal s.Hassan et al [2] and Shrestha and Kumar [15] analyzed Bayesian estimation takes parameters as probabilistic variables with random variables. The Bayesian is helpful in analysis since it can take the prior knowledge into account.
In this EPOLO distribution, we have four parameters, but estimating the four parameters analytically is not possible, so we have estimated the shape parameter c alone here for the better understanding of Bayesian analysis by Hesham and Soha [6]. We assume the prior for the parameter c as a gamma distribution with the pdf.

$$\pi(c) \propto c^{a-1} e^{-bc}\, c > 0, \qquad a, b > 0 \tag{5}$$

The likelihood functions of EPOLO for the random samples X_1, X_2… X_n of independent and identically distributed random variables is given by

$$L\,(\alpha,\beta,\lambda,c) = \,(\alpha\beta c\gamma^{\alpha}\,)^{n}\prod_{i=1}^{n} x^{\beta-1}\,\left(\gamma+x^{\beta}\right)^{-\alpha-1}\,\left(1-\gamma^{\alpha}\left(\gamma+x^{\beta}\right)^{-\alpha}\right)^{c-1}\; x\,>\,0; \quad \alpha,\beta,\gamma,c\,>\,0 \tag{6}$$

The likelihood only for the shape parameter c is taken as

$$L(\frac{x}{c}) \;= c^{n}\prod_{i=1}^{n}(1-\gamma^{\alpha}(\gamma+x^{\beta})^{-\alpha})^{c} \tag{7}$$

Using the bayes theorem the posterior is given by

$$\pi(\frac{c}{x}) \;= \frac{L(\frac{x}{c})\,\pi(c)}{\int_0^{\infty} L(\frac{x}{c})\,\pi(c)dc} \tag{8}$$

The posterior density of the shape parameter c can be obtained as

$$\pi(\frac{c}{x}) \;= \frac{[b-w]^{n+a}}{\Gamma(n+a)}\; c^{n+a-1}exp\{-c(b\,-\,w)\} \tag{9}$$

Where $w = \sum_{i=1}^{n} log(1-\gamma^{\alpha}\,(\gamma+x_i^{\beta})^{-\alpha}$

The Bayes estimate of $\hat{c}$, Sowbhagya [16] and Nada S. Karam [14] provides a knowledge on the squared error loss function, is derived as

$$\hat{c} \;= \int_{-\infty}^{\infty} c\,f(\frac{c}{x})dc \tag{10}$$

$$\hat{c} \;= \frac{n+a}{b+w} \tag{11}$$

3. Maximum likelihood estimation

The likelihood function is obtained by applying the product to the density function of the given distribution over i = 1 to n. That is,

$$L(\alpha,\beta,\gamma,c) = (\alpha\beta c\gamma^{\alpha}\,)^{n}\prod_{i=1}^{n} x^{\beta-1}\;(\gamma+x^{\beta})^{-\alpha-1}\;(1-\gamma^{\alpha}(\gamma+x^{\beta})^{-\alpha})^{c-1} \tag{12}$$

The log-likelihood function for the parameters α, β, λ, and c is

$$\ln L(\alpha,\beta,\gamma,c) = n\,log\,(\alpha\beta c\,\gamma^{\alpha}) + (c-1)\sum_{i=1}^{n}\log\left(1-\gamma^{\alpha}\left(\gamma+x_i^{\beta}\right)^{-\alpha}\right) - (\alpha+1)\sum_{i=1}^{n}\log\,(x_i^{\beta}+\gamma) + (\beta-1)\sum_{i=1}^{n}log\,(x_i) \tag{13}$$

To find the maximum likelihood estimator of the parameter $\hat{c}$ we equate the log-likelihood to zero, then we get,

$$n\,c + \sum_{i=1}^{n} log(1-\gamma^{\alpha}\,(\gamma+x_i^{\beta})^{-\alpha} = 0 \tag{14}$$

Taking $w = \sum_{i=1}^{n} log(1-\gamma^{\alpha}\,(\gamma+x_i^{\beta})^{-\alpha}$

$$\hat{c} = \frac{n}{-w} \tag{15}$$

4. Reliability Measures

Survival and hazard functions referred to the cumulative distribution function (1) of EPOLO distribution are obtained. The examination of organism or technological unit breakdowns that take place after a certain time is aided by the survival function. The hazard rate is used to track the specific unit over the course of its lifespan distribution. The likelihood to fail or die, depending on the age attained, is measured by the hazard rate (HR), which is a critical factor in categorizing lifespan distributions. According to Arun Kumar Rao [1] and Hare and Sharma [4] the hazard rates are often monotonic or non-monotonic. The possibility of being alive for a specific period of time or the likelihood that an important event will take place before a certain amount of time (t) is the definition of the survival function (x). The survival mechanism is expressed as

$$S\,(x) = 1-\left(1-\gamma^{\alpha}\left(\gamma+x^{\beta}\right)^{-\alpha}\right)^{c} \quad x > 0;\ \alpha,\beta,\gamma,c > 0 \tag{16}$$

Where $S(0) = 1$ and $\lim_{x\to\infty} S(x) = 0$.

To find out MSE for the survival rate, we need to calculate

$$\widehat{S\,(x)} = \int_0^{\infty} S\,(x)\,\pi\,(\frac{c}{x})\,dc \tag{17}$$

$$\widehat{S\,(x)} = \int_0^{\infty} 1-(1-\gamma^{\alpha}(\gamma+x^{\beta})^{-\alpha})^{c}\,\frac{[b-w]^{n+a}}{\Gamma(n+a)}\ c^{n+a-1}exp\{-c(b-w)\}\ dc \tag{18}$$

The failure rate, also known as the hazard function h(x), is the ratio of the likelihood that an event will occur in a given amount of time, t, to the likelihood that it will pass off successfully. When expressing the hazard function it is given as,

$$h(x) = \frac{f(x)}{S(x)} = \frac{\alpha\beta c\gamma^{\alpha} x^{\beta-1} \left(\gamma + x^{\beta}\right)^{-\alpha-1} (1 - \gamma^{\alpha} (\gamma + x^{\beta})^{-\alpha})^{C-1}}{1 - (1-\gamma(\gamma+x^{\beta})^{-\alpha})^{C}} \quad x > 0; \tag{19}$$

$$\alpha,\beta,\gamma,c > 0$$

To calculate the MSE of the hazard function

$$\widehat{h(x)} = \int_0^{\infty} h(x)\, \pi\, (\frac{c}{x})\, dc \tag{20}$$

$$\widehat{h(x)} = \int_0^{\infty} \frac{\alpha\beta c\gamma^{\alpha} x^{\beta-1} (\gamma + x^{\beta})^{-\alpha-1} (1-\gamma^{\alpha} (\gamma + x^{\beta})^{-\alpha})^{C-1}}{1 - (1-\gamma(\gamma+x^{\beta})^{-\alpha})^{C}} \frac{[b-w]^{n+a}}{\Gamma(n+a)} c^{n+a-1} exp\{-c(b - w)\}\ dc \tag{21}$$

After finding the $\widehat{S(x)}$ *and* $\widehat{h(x)}$ using the formula then MSE is calculated to compare the errors.

5. Quantile function

Using the c.d.f, the quantile function can be found by solving the equation F(x) = u, 0< u < 1. Assume that X is a random variable with an EPOLO distribution. The equation F(Q(u)) = u defines the quantile function Q(u) and represents it.

$$F(x) = \left(1-\gamma^{\alpha}\left(\gamma + x^{\beta}\right)^{-\alpha}\right)^{c} x > 0;\ \alpha,\beta,\gamma,c > 0 \tag{24}$$

$$u^{1/c} = 1-\gamma^{\alpha}(\gamma + x^{\beta})^{-\alpha} \tag{25}$$

$$[(\frac{1-u^{\frac{1}{c}}}{\gamma^{\alpha}})^{-1/\alpha} - \gamma]^{1/\beta} = x \tag{26}$$

6. Monte Carlo simulation study

The effectiveness of the parameters is tested using a simulation study, which is conducted for the shape parameter c of the EPOLO distribution. Mean Square Error (MSE) using the two familiar estimation maximum likelihood estimation and Bayesian estimation under squared error loss function was calculated. The following algorithm is proposed by Shrestha and Kumar [12] for the simulation studies.
To simulate the random samples of x which follows the EPOLO distribution, the equation F(x) = u is used, where u (0, 1) follows uniform distribution and F(x) is the CDF of the EPOLO distribution. Fix the values of the parameters α, β, γ, c. Here we fix other parameters to estimate the shape parameter c.

- Determine the different sample sizes
- Calculate $\hat{c}$ for each of the n samples
- Calculate the MSE for the parameter *c* using the below formula

The average squared difference between the estimator and parameter, measured by the MSE, is a reliable indicator of an estimator's performance. Generally speaking, any growing function of the absolute distance $|\hat{c} - c|$ would be used to assess an estimator's quality. But compared to other distance metrics, MSE offers at least two advantages: First, it can be analyzed, and second, it has the interpretation.

$$MSE(c) = \sum_{i=1}^{n} \frac{(\hat{c} - c)^2}{n}$$

A simulation were conducted for samples n=15, 30, 50, 100 and repeated for 1,000 times with the parameter c values 0.5, 1.5.

7. Analyses

According to Venables et al. [18], all the analysis is done using the R software. R software is free source software helps a lot in statistical analysis.

Table 1: *MSE of Maximum likelihood estimate of the shape parameter c and MSE of Bayesian estimate of the shape parameter under SELF loss function is given below*

n	c	MSE (MLE)	MSE (BAYES)
15	0.5	0.0229	0.0165
	1.5	0.2322	0.1087
30	0.5	0.0093	0.0081
	1.5	0.086	0.0635
50	0.5	0.0059	0.0053
	1.5	0.0489	0.0403
100	0.5	0.0021	0.0021
	1.5	0.034	0.0263

Table 2: *MSE of survival and hazard function of c along with the Bayesian posterior is given below*

n	c	survival	hazard
15	0.5	0.031	1.03e-04
	1.5	0.00267	5.77e-07
30	0.5	0.0092	2.11e-05
	1.5	0.00076	5.02e-08
50	0.5	0.0045	7.28e-06
	1.5	0.00027	8.43e-09
100	0.5	0.0012	9.58e-07
	1.5	5.49e-05	3.72e-10

8. Conclusion

In this article, we looked at the Bayesian estimates for the exponentiated power Lomax distribution's of shape parameter c. Based on the findings in Tables 1 and 2, we note the following:

- For the shape parameter c, the MSE of the Bayesian and maximum likelihood estimate have been determined. In that MSE of Bayesian is less than MLE method that indicates the Bayesian estimation is a good estimator for estimating the parameters of any distribution. As sample size increases both will be same.
- The survival and hazard function using the Bayesian posterior is calculated and it indicates the distribution is of good fit. The error decreases as the sample size increases for survival and hazard functions.

References

[1] Arun Kumar Rao. (2016). Estimation of Reliability Function of Lomax Distribution via Bayesian Approach. International Journal of Mathematics Trends and Technology (IJMTT) – Volume 40 Number

[2] Amal S Hassan., Said g., and Nassr. (2018). Power lomax poisson distribution: properties and estimation. journal of data science, 105-128.

[3] Balakrishnan, N., and Ahsanullah, M. (1994). Relations for single and product moments of record values from lomax distribution. Sankhya: indian j statist. Ser b, 56:140-146.

[4] Hare Krishna., and Ranjeet Sharma. (2007). Estimation of reliability characteristics of general system configuration. Chaudhary Charan Singh University, Meerut, India, IJQRM.

[5] Hassan, A.S., and Al-ghamdi, A.S. (2009). Optimum step stress accelerated life testing for lomax distribution. J appl sci res,5:2153-2164

[6] Hesham mohamed reyad., and Soha othman ahmed. (2016). Bayesian and e-bayesian estimation for the kumaraswamy distribution based on type-ii censoring. international journal of advanced mathematical sciences, 4 (1).

[7] Kilany, N.M. (2016). Weighted lomax distribution. Springerplus, 5(1):1862.

[8] Lemonte, A.J., and Cordeiro, G.M. (2013). An extended lomax distribution. Statistics, 47(4):800-816.

[9] Lingappaiah, G. S. (1986). On the pareto distribution of the second kind (lomax distribution). Revista de mathemtica e estatistica, 4:63–68.

[10] Lomax, K.S. (1954). Business failures: another example of the analysis of failure data. J am statist assoc, 45:21–29.

[11] Mahmoud, M. A., Soliman, A. A., Abd ellah, A. H., and El-sagheer, R. M. (2013). Mcmc technique to study the bayesian estimation using record values from the lomax distribution. International journal of computer applications, 73(5).

[12] Mohamed mohamed ezzat abd el-monsef. (2020). The exponentiated power lomax distribution and its applications, Qual Reliab Engng Int, 1–24.

[13] Myhre, J., and saunders, S. (1982). Screen testing and conditional probability of survival. Lect notes-monogr series, 166–178.

[14] Nada S. Karam. (2014). Bayesian analysis of five exponentiated distributions under different priors and loss functions, iraqi journal of science, vol 55, no.3b, pp: 1353-1369.

[15] Shrestha, S. K., and kumar, V. (2014). Bayesian analysis of extended lomax distribution. International journal of mathematics trends and technology, 7(1), 33-41.

[16] Sowbhagya S prabhu. (2020). E-bayesian estimation of the shape parameter of lomax model under symmetric and asymmetric loss functions, International Journal of Statistics and Applied Mathematics, 5(6): 142-146

[17] Tierney, l. (1994). Markov chains for exploring posterior distributions. The annals of statistics, 1701-1728.

[18] Venables, W. N., Smith D. M., and the R core team. (2022). An introduction to r notes on r: a programming environment for data analysis and graphics version 4.2.1.