Adıyaman University Journal of Science
https://dergipark.org.tr/en/pub/adyujsci
DergiPark AKADEMİK
ISSN 2147-1630
e-ISSN 2146-586X

Parameter Estimation Procedures for Log Exponential-Power Distribution with Real Data Applications

Mustafa Ç. KORKMAZ[1], Kadir KARAKAYA[2*], Yunus AKDOĞAN[3]

[1]*Artvin Çoruh University, Faculty of Education, Department of Measurement and Evaluation, Artvin, Türkiye*
mustafacagataykorkmaz@gmail.com, ORCID: 0000-0003-3302-0705
[2]*Selçuk University, Science Faculty, Department of Statistics, Konya, Türkiye*
kkarakaya@selcuk.edu.tr, ORCID: 0000-0002-0781-3587
[3]*Selçuk University, Science Faculty, Department of Statistics, Konya, Türkiye*
yakdogan@selcuk.edu.tr, ORCID: 0000-0003-3520-7493

Received: 14.02.2022 **Accepted:** 28.09.2022 **Published:** 30.12.2022

Abstract

In this study, some estimation techniques are investigated to estimate two parameters of the log exponential-power distribution. The maximum likelihood, quantile, least squares, weighted least squares, Anderson-Darling, and Cramer-von Mises estimation methods are studied in detail. The efficiency of these estimators is validated through Monte Carlo simulation experiments. Also, four real data applications are performed and Kolmogorov-Smirnov statistic results for all estimators are presented.

Keywords: Point estimation; Log exponential-power distribution; Maximum likelihood estimators; Practical data application.

* Corresponding Author DOI: 10.37094/adyujsci.1073616

Gerçek Veri Uygulamaları ile Log Exponential-Power Dağılımı için Parametre Tahmin Prosedürleri

Öz

Bu makalede, log exponential-power dağılımının iki parametresini tahmin etmek için çeşitli tahmin yöntemleri araştırılmıştır. En çok olabilirlik, kuantil, en küçük kareler, ağırlıklandırılmış en küçük kareler, Anderson-Darling ve Cramer-von Mises tahmin yöntemleri detaylı olarak incelenmiştir. Bu tahmin edicilerin performanslarını değerlendirmek için Monte Carlo simülasyon deneyleri yapılmıştır. Ayrıca dört gerçek veri uygulaması gerçekleştirilmiş ve tüm tahmin ediciler Kolmogorov-Smirnov istatistiği sonuçları sunulmuştur.

Anahtar Kelimeler: Nokta tahmini; Log exponential-power dağılımı; En çok olabilirlik tahmini; Gerçek veri uygulaması.

1. Introduction

The log exponential-power (LEP) model was introduced by [1], applying the $X = \exp(-T)$ transformation where distribution of the T random variable is exponential-power distribution suggested by [2]. Korkmaz et al. [1] investigates some of the mathematical features of the LEP distribution such as moments, coefficient of skewness and kurtosis, order statistics, quantiles, entropies, etc. The LEP distribution is a unit distribution and is an alternative to distributions such as Beta and Kumaraswamy [3]. The probability density function (pdf) and cumulative distribution function of $LEP(\alpha, \beta)$ distribution are presented as:

$$f(x,\alpha,\beta) = \frac{\alpha\beta}{x} e^{\alpha(-\log x)^{\beta}} \left(-\log x\right)^{\beta-1} e^{1-\exp\left\{\alpha(-\log x)^{\beta}\right\}}, \quad x \in (0,1) \tag{1}$$

and

$$F(x,\alpha,\beta) = e^{1-\exp\left\{\alpha(-\log x)^{\beta}\right\}}, \quad x \in (0,1), \tag{2}$$

where $\alpha > 0$ and $\beta > 0$ are model parameters. The following is the LEP model's hazard rate function (hrf):

$$h(x,\alpha,\beta) = \frac{\alpha\beta}{x\left(e^{\exp\left\{\alpha(-\log x)^{\beta}\right\}-1} - 1\right)} e^{\alpha(-\log x)^{\beta}} \left(-\log x\right)^{\beta-1}, \quad x \in (0,1). \tag{3}$$

It is presented in [1] that the hrf of the LEP distribution has unimodal, increasing, U-shaped, bathtub or N-shaped. Figure 1 shows the hrf and pdf plots of the LEP model for different parameter options.

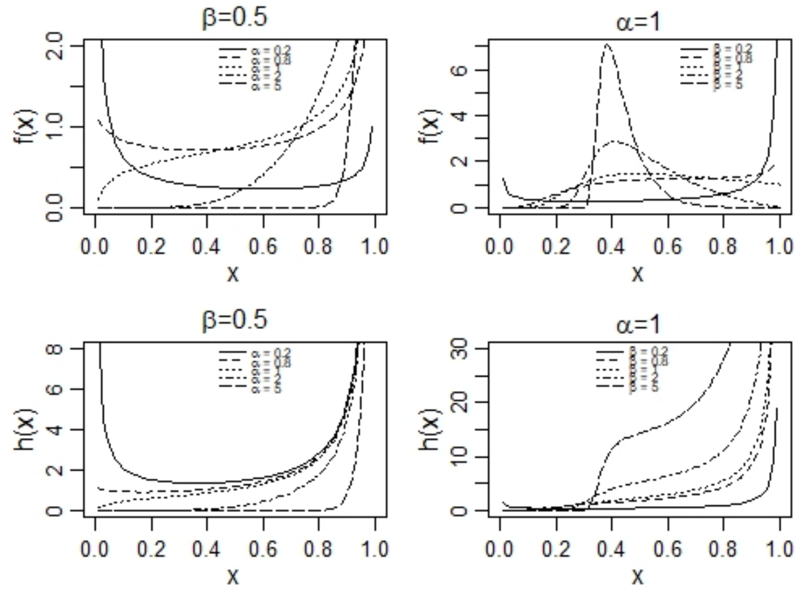

Figure 1: The pdf and hrf plots of the LEP model for some parameter values

Other issues studied in the study of [1] can be listed as follows: Only the maximum likelihood estimator approach was studied in estimate procedure. The LEP distribution was also used to develop a novel quantile regression model. The regression model parameters were also estimated using the maximum likelihood technique. In some scenarios, extensive simulation studies were conducted for the estimation of two distribution parameters.

The goal of this work is to use simulations to evaluate six alternative estimators for the parameters of the LEP distribution. The maximum likelihood estimators (MLE), least squares estimators (LSE), quantile estimators (QE), weighted least squares estimators (WLSE), Anderson-Darling estimators (AD), and Cramer-von Mises estimators (CVM) are examined for point estimation. The remainder of this paper is structured as follows: Six estimate procedures are presented in Section 2. In Section 3, simulation experiment is conducted to assess the performance of these estimators using some criteria such as bias and mean square error (MSE). Four practical data applications are considered in Section 4. Lastly, there are some concluding remarks in Section 5.

2. Different Parameter Estimation Procedures

In this section, six different estimate procedures are discussed for estimating the unknown parameters of the LEP distribution. These methods and details are described in the subsections below.

2.1. Maximum likelihood estimation

In this subsection, the MLE of the LEP distribution is provided. Let $\mathbf{\Theta} = (\alpha, \beta)^T$ be the parameter vector and $x_1, x_2, \ldots, x_n$ random sample from the LEP distribution. The log-likelihood function is presented as

$$\ell(\mathbf{\Theta}) = n + n\log\alpha + n\log\beta + \alpha\sum_{i=1}^{n}(-\log x_i)^{\beta} + (\beta-1)\sum_{i=1}^{n}\log(-\log x_i) - \sum_{i=1}^{n} e^{\alpha(-\log x_i)^{\beta}}. \quad (4)$$

Then, differentiating (4), for the MLEs, say $\hat{\alpha}$ and $\hat{\beta}$, the normal equations are obtained by

$$\frac{\partial\ell(\mathbf{\Theta})}{\partial\alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n}(-\log x_i)^{\beta} - \sum_{i=1}^{n}(-\log x_i)^{\beta} e^{\alpha(-\log x_i)^{\beta}} = 0$$

and

$$\frac{\partial\ell(\mathbf{\Theta})}{\partial\beta} = \frac{n}{\beta} + \alpha\sum_{i=1}^{n}\log(-\log x_i)(-\log x_i)^{\beta} - \alpha\sum_{i=1}^{n}\log(-\log x_i)(-\log x_i)^{\beta} e^{\alpha(-\log x_i)^{\beta}} = 0.$$

The numerical technique should be required to acquire the $\hat{\alpha}$ and $\hat{\beta}$. The quasi-Newton or Newton-Raphson algorithms can be utilized for this aim.

2.2. Quantile estimation

The quantile estimation method also known as the percentile estimation method is introduced by [4]. Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the order statistics of random sample of size n from the LEP distribution and let $q_i = \frac{i}{n+1}$ for $i = 1, 2 \ldots, n$ is the estimation of the $F(x_{(i)}, \alpha, \beta)$. Then, by minimizing the function

$$QE(\mathbf{\Theta}) = \sum_{i=1}^{n}\left(x_{(i)} - \exp\left(-\frac{\log(1-\log q_i)}{\alpha}\right)^{1/\beta}\right)^2,$$

according to the model parameters, the QE of the α and β parameters, say $\hat{\alpha}_{QE}$ and $\hat{\beta}_{QE}$, can be determined.

2.3. Least squares estimation

The LSE of α and β, say $\hat{\alpha}_{LSE}$ and $\hat{\beta}_{LSE}$, can be obtained by minimizing the function

$$LSE(\Theta)=\sum_{i=1}^{n}\left(\exp\left(1-\exp\left(\alpha\left(-\log\left(x_{(i)}\right)\right)^{\beta}\right)\right)-\frac{i}{n+1}\right)^{2}.$$

2.4. Weighted least squares estimation

The weighted least squares estimation method is the weighted version of the LSE method. By minimizing the function

$$WLSE(\Theta)=\sum_{i=1}^{n}\frac{(n+2)(n+1)^{2}}{i(n-i+1)}\left(\exp\left(1-\exp\left(\alpha\left(-\log\left(x_{(i)}\right)\right)^{\beta}\right)\right)-\frac{i}{n+1}\right)^{2},$$

the WLSEs of α and β, say $\hat{\alpha}_{WLSE}$ and $\hat{\beta}_{WLSE}$ can be derived.

2.5. Anderson-Darling estimation

The ADE of α and β, say $\hat{\alpha}_{AD}$ and $\hat{\beta}_{AD}$, are determined by minimizing the function

$$\begin{aligned}AD(\Theta)=-n-\sum_{i=1}^{n}\frac{2i-1}{n}&\left\{\log\left(\exp\left(1-\exp\left(\alpha\left(-\log\left(x_{(i)}\right)\right)^{\beta}\right)\right)\right)\right.\\&+\log\left(1-\exp\left(1-\exp\left(\alpha\left(-\log\left(x_{(i)}\right)\right)^{\beta}\right)\right)\right)\\&\left.+\log\left(\exp\left(1-\exp\left(\alpha\left(-\log\left(x_{(i)}\right)\right)^{\beta}\right)\right)\right)+\log\left(1-\exp\left(1-\exp\left(\alpha\left(-\log\left(x_{(i)}\right)\right)^{\beta}\right)\right)\right)\right\}.\end{aligned}$$

2.6. Cramer-von Mises estimation

The CVME of α and β, say $\hat{\alpha}_{CVM}$ and $\hat{\beta}_{CVM}$, are determined by minimizing the function

$$CVM(\Theta)=\frac{1}{12n}+\sum_{i=1}^{n}\left[\exp\left(1-\exp\left(\alpha\left(-\log\left(x_{(i)}\right)\right)^{\beta}\right)\right)-\frac{2i-1}{2n}\right]^{2}.$$

All equations belonging to the above estimation procedures have no explicit solutions. As a result, they must be solved numerically using well-known software such as R , Matlab, and S-Plus utilizing Newton-Raphson and quasi-Newton techniques. Numerical optimization approaches will be used to solve these functions.

3. Simulation Procedure

In this section, the empirical results have been given to see the performance of the pointed out estimators. *N=1000* samples size *n=20, 25, … ,1000* are generated from a random variable using the LEP distribution. The data generation from the LEP distribution is conducted as follows:

If the random variable U follows the standard uniform distribution, then $X \sim \exp\left\{-\left(\frac{\log\left(1-\log\left(U\right)\right)}{\alpha}\right)^\wedge\left(1/\beta\right)\right\}$ follows the LEP distribution. The true values of the first and second Monte Carlo simulation studies are $\boldsymbol{\Theta}=(0.5,0.5)$ and $\boldsymbol{\Theta}=(2,2)$ respectively. Moreover, the empirical mean, bias, and MSE belonging to related estimators have been obtained for comparisons between estimation methods. By setting $\varepsilon=\alpha$ or β, the related bias and MSE are calculated by

$$Bias_{\varepsilon}(n)=\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon_i-\hat{\varepsilon}_i\right),\quad MSE_{\varepsilon}(n)=\frac{1}{N}\sum_{i=1}^{N}\left(\varepsilon_i-\hat{\varepsilon}_i\right)^2, \tag{5}$$

respectively. Figures 2 and 3 present the results of simulation studies. Figures 2 and 3 show that as the sample size increases, the empirical means go to the true parameter values, all estimators are asymptotically unbiased, and all MSEs go to zero. Simultaneously, as the sample size increases, the empirical results get closer.

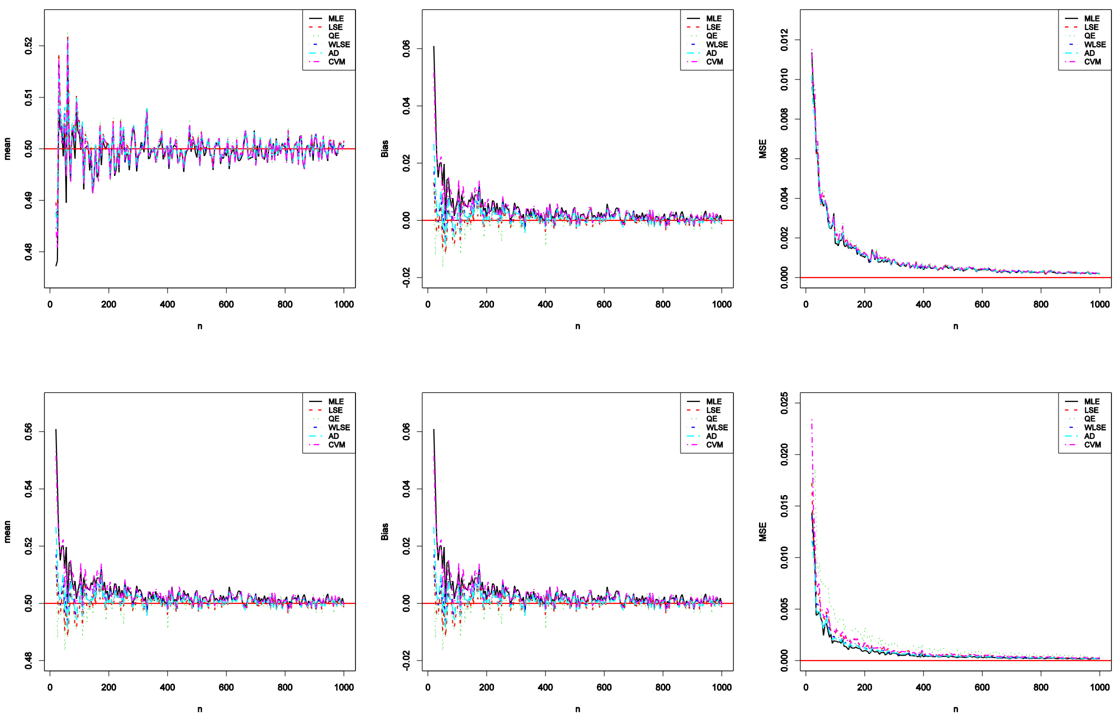

Figure 2: The α (top) and β (bottom) parameters results for $\boldsymbol{\Theta}=(0.5,0.5)$

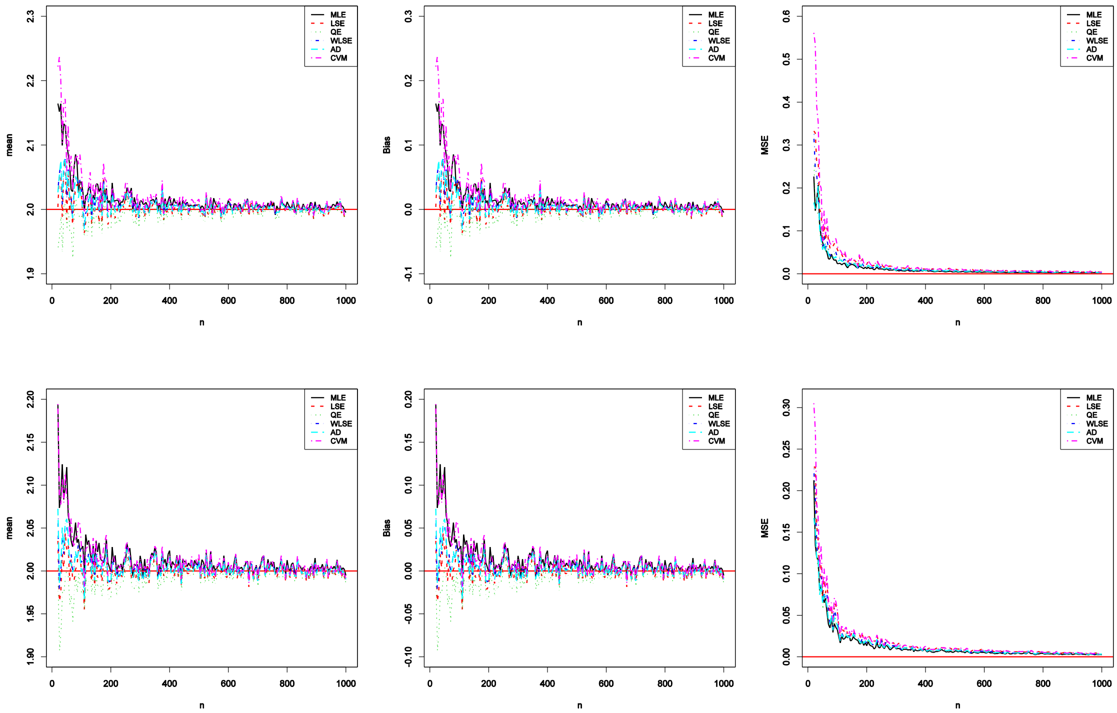

Figure 3: The α (top) and β (bottom) parameters' results for $\boldsymbol{\Theta} = (2, 2)$

4. Real Data Applications

In this section, four practical data applications for the LEP distribution are addressed. The LEP distribution is fitted to four practical datasets by estimating the parameters using the six estimators discussed in the previous sections. The MLE, LSE, QE, WLSE, AD, and CVM the parameters α and β of LEP distribution are achieved by BFGS algorithm and reported in Table 1. Table 1 also shows the Kolmogorov-Smirnov statistics (KS) and related p values for all estimators.

The first data is taken from [5] and represents the flood levels for the Susquehanna River at Harrisburg, Pennsylvania. The data are 0.654, 0.613, 0.315, 0.449, 0.297, 0.402, 0.379, 0.423, 0.379, 0.3235, 0.269, 0.740, 0.418, 0.412, 0.494, 0.416, 0.338, 0.392, 0.484 and 0.265. Recently, this data has been examined in [6, 7].

The second data set comes from [8] and show the strengths of 1.5 cm glass fibers, which were first measured by researchers at the UK National Physical Laboratory. The data are 0.17, 0.13, 0.16, 0.14, 0.20, 0.15, 0.13, 0.11, 0.15, 0.12, 0.12, 0.15, 0.12, 0.16, 0.21, 0.20, 0.23, 0.16, 0.12, 0.10, 0.32, 0.33, 0.33, 0.36, 0.38, 0.20 and 0.26.

The third data set is taken from [9] and consist of 48 rock samples from a petroleum reservoir. The detailed information about data given in [9]. Recently, this data has been examined in [10, 11]. The data are 0.0903296, 0.2036540, 0.2043140, 0.2808870, 0.1976530, 0.3286410, 0.1486220, 0.1623940, 0.2627270, 0.1794550, 0.3266350, 0.2300810, 0.1833120, 0.1509440, 0.2000710, 0.1918020, 0.1541920, 0.4641250, 0.1170630, 0.1481410, 0.1448100, 0.1330830, 0.2760160, 0.4204770, 0.1224170, 0.2285950, 0.1138520, 0.2252140, 0.1769690, 0.2007440, 0.1670450, 0.2316230, 0.2910290, 0.3412730, 0.4387120, 0.2626510, 0.1896510, 0.1725670, 0.2400770, 0.3116460, 0.1635860, 0.1824530, 0.1641270, 0.1534810, 0.1618650, 0.2760160, 0.2538320 and 0.2004470.

The fourth data set comes from [9] and show total milk production in the first birth of 107 cows from SINDI race. Recently, this data has been examined in [12, 13]. The data are 0.4365, 0.4260, 0.5140, 0.6907, 0.7471, 0.2605, 0.6196, 0.8781, 0.4990, 0.6058 ,0.6891 ,0.5770, 0.5394, 0.1479, 0.2356, 0.6012, 0.1525, 0.5483, 0.6927, 0.7261, 0.3323, 0.0671, 0.2361, 0.4800, 0.5707, 0.7131, 0.5853, 0.6768 ,0.5350, 0.4151 ,0.6789, 0.4576, 0.3259, 0.2303, 0.7687, 0.4371 ,0.3383, 0.6114, 0.3480, 0.4564, 0.7804, 0.3406, 0.4823, 0.5912 ,0.5744, 0.5481, 0.1131, 0.7290, 0.0168, 0.5529 ,0.4530, 0.3891, 0.4752, 0.3134, 0.3175 ,0.1167, 0.6750 ,0.5113,0.5447, 0.4143, 0.5627, 0.5150, 0.0776, 0.3945 ,0.4553, 0.4470, 0.5285, 0.5232, 0.6465, 0.0650, 0.8492, 0.8147, 0.3627, 0.3906, 0.4438, 0.4612, 0.3188, 0.2160, 0.6707, 0.6220, 0.5629, 0.4675, 0.6844, 0.3413,0.4332, 0.0854, 0.3821, 0.4694, 0.3635, 0.4111, 0.5349, 0.3751, 0.1546, 0.4517 ,0.2681, 0.4049, 0.5553, 0.5878 ,0.4741 ,0.3598, 0.7629, 0.5941, 0.6174, 0.6860, 0.0609, 0.6488 and 0.2747.

Table 1: Parameter estimation and KS results for four data sets based on the six estimators

		Parameter		**KS**	
Data	**Estimators**	α	β	Statistics	p-values
First	**MLE**	0.6593	2.9191	0.1366	0.8494
	LSE	0.6573	2.7806	0.1360	0.8529
	QE	0.2719	2.1158	0.5264	0.0001
	WLSE	0.6507	2.6582	0.1390	0.8339
	AD	0.6581	2.8151	0.1359	0.8535
	CVM	0.6679	3.0180	0.1317	0.8783
Second	**MLE**	0.0413	4.3744	0.1120	0.8870
	LSE	0.0584	3.7762	0.1063	0.9202
	QE	0.0330	4.3681	0.1699	0.4164
	WLSE	0.0566	3.8574	0.1044	0.9296
	AD	0.0530	3.9659	0.1050	0.9271
	CVM	0.0512	3.9894	0.0990	0.9537
Third	**MLE**	0.0938	3.5262	0.0966	0.7607
	LSE	0.0757	3.9984	0.0655	0.9859
	QE	0.0466	3.5564	0.4307	0.0000
	WLSE	0.0782	3.9342	0.0664	0.9837
	AD	0.0816	3.8366	0.0659	0.9850
	CVM	0.0709	4.1299	0.0633	0.9906

Fourth	**MLE**	0.6396	0.9669	0.1746	0.0029
	LSE	0.8227	1.6112	0.1075	0.1685
	QE	0.6623	1.0339	0.1437	0.0240
	WLSE	0.8150	1.6911	0.1116	0.1389
	AD	0.6815	1.1128	0.1280	0.0597
	CVM	0.8281	1.6353	0.1107	0.1451

5. Concluding Remarks

The LEP distribution proposed by [1] is explored in this work with relation to several point estimations. The two unknown parameters of the LEP distribution are estimated using six estimators. For two different parameter values and sample sizes, extensive Monte Carlo simulations are conducted. It is observed that when the size of the sample increases, the mean of all estimators neared the real parameter value, their biases and MSEs decreased and approached zero. In addition, the estimations and KS results for all estimators for four real data are examined.

Conflicts of Interest

No conflict of interest was declared by the authors.

References

[1] Korkmaz, M.Ç., Altun, E., Alizadeh, M., El-Morshedy, M., *The Log Exponential-Power Distribution: Properties, Estimations and Quantile Regression Model*, Mathematics, 9 (21), 2634, 2021.

[2] Smith, R.M., Bain, L.J., *An exponential power life-testing distribution*, Communication in Statistics Theory Methods, 4, 469–481, 1975.

[3] Kumaraswamy, P., *A generalized probability density function for double-bounded random processes*, Journal of Hydrology, 46, 79–88, 1980.

[4] Kao, J.H., *Computer methods for estimating Weibull parameters in reliability studies*, IRE Transactions on Reliability and Quality Control, 13, 15-22, 1958.

[5] Dumonceaux, R., Antle, C.E., *Discrimination between the log-normal and the Weibull distributions*, Technometrics, 15 (4), 923-926, 1973.

[6] Balakrishnan, N., Cohen, A.C., *Order Statistics & Inference: Estimation Methods,* Elsevier, Amsterdam, The Netherlands, 2014.

[7] Alizadeh, M., Altun, E., Cordeiro, G.M., Rasekhi, M., *The odd power Cauchy family of distributions: Properties, regression models and applications*, Journal of Statistical Computation and Simulation, 88, 785–807, 2018.

[8] Elgarhy, M., *Exponentiated generalized Kumaraswamy distribution with applications*, Annals of Data Science, 5 (2), 273-292, 2018.

[9] Cordeiro, G.M., dos Santos Brito, R., *The beta power distribution*, Brazilian Journal Of Probability And Statistics, 26 (1), 88-112, 2012.

[10] Opone, F., Iwerumor, B., *A new Marshall-Olkin extended family of distributions with bounded support*, Gazi University Journal of Science, 34 (3), 899-914, 2021.

[11] Balogun, O.S., Iqbal, M.Z., Arshad, M.Z., Afify, A.Z., Oguntunde, P.E., *A new generalization of Lehmann type-II distribution: Theory, simulation, and applications to survival and failure rate data*, Scientific African, 12, e00790, 2021.

[12] Saraçoğlu, B., Tanış, C., *A new statistical distribution: cubic rank transmuted Kumaraswamy distribution and its properties*, Journal of the National Science Foundation of Sri Lanka, 46 (4), 505-518, 2018.

[13] Jamal, F., Chesneau, C., *A new family of polyno-expo-trigonometric distributions with applications*, Infinite Dimensional Analysis, Quantum Probability and Related Topics, 22 (04), 1950027, 2019.