ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2020, Vol. 41, No. 3, pp. 308–319. © Pleiades Publishing, Ltd., 2020.

The Length-Biased Weighted Lindley Distribution with Applications

Yupapin Atikankul[1*], Ampai Thongteeraparp[1], Winai Bodhisuwan[1***], and A. Volodin[2****]**

(Submitted by A. M. Elizarov)

[1] *Department of Statistics, Kasetsart University, Bangkok, 10903 Thailand*

[2] *Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, S4S 0A2 Canada*

Received October 10, 2019; revised October 25, 2019; accepted November 23, 2019

Abstract—In this paper, we propose a new length-biased distribution, which is a special case of weighted distributions. We derive some mathematical properties of the proposed distribution, including moment generating function, characteristic function and moments, and discuss parameter estimation by the method of moments and maximum likelihood estimation. We assess estimators via simulation, and show the potential of the proposed distribution by fitting it with some real-life data sets.

DOI: 10.1134/S199508022003004X

Keywords and phrases: *length-biased distribution, new weighted Lindley distribution, parameter estimation.*

1. INTRODUCTION

When observations are recorded from random process, the probability of recorded observations are not equal. Although it is incorrect to use the original distribution to describe these observations, weighted distributions can be applied in this situation. The concept of weighted distributions was proposed by Fisher [1]. Later, Rao [2] introduced discrete weighted distributions for sampling with probabilities of selection depend on the proportion of their units.

If X is a non-negative random variable with the probability density function (pdf) $f(x)$, then the corresponding weighted distribution is

$$f_w(x) = \frac{w(x)f(x)}{\mathrm{E}(w(X))},$$

where $w(x)$ is a non-negative weighted function and $\mathrm{E}(w(X)) < \infty$.

Patil and Ord [3] presented a size-biased distribution that is a special case of the weighted distribution. The weighted function of a size-biased distribution is $w(x) = x^c$, where $c = 1$ is named length biased and $c = 2$ is named as area biased distributions. Thus, the pdf of the length-biased distribution is given by

$$f_l(x) = \frac{xf(x)}{E(X)}. \tag{1}$$

*E-mail: yupapin.a@ku.th

**E-mail: fsciamu@ku.ac.th

***E-mail: fsciwnb@ku.ac.th

****E-mail: andrei@uregina.ca

Many length-biased distributions have been proposed in the literature; we review some here. Das and Roy [4] studied the length-biased weighted generalized Rayleigh distribution for the monthly mean minimum temperature data They also studied the length-biased weighted Weibull distribution in [5] for rainfall data in India. Ratnaparkhi and Naik—Nimbalkar [6] analysed oil field exploration data by the length-biased lognormal distribution. Khadim and Hussein [7] introduced the length-biased weighted Exponential and Rayleigh distributions for industrial data. Nanuwong and Bodhisuwan [8] applied the length biased beta-Pareto distribution for fire insurance data in Norway. Seenoi et al. [9] analysed data on the distance between cracks in a pipe by the length-biased exponentiated inverted Weibull distribution. Modi and Gill [10] studied the length-biased weighted Maxwell distribution and its properties. Ahmad et al. [11] proposed the length-biased weighted Lomax distribution for remission times of bladder cancer patients. Saghir et al. [12] introduced the length-biased weighted exponentiated inverted Weibull distribution for lifetime data. Ayesha [13] studied the length-biased (size-biased) Lindley distribution and its properties. Rather and Subramanian [14] investigated the length-biased Sushila distribution for waiting times of bank customers. Numerous length-biased distributions have been established for lifetime data; they provide a better fit than the original distributions.

The Lindley distribution was investigated by Lindley [15] in the context of Bayesian statistics. Ghitany et al. [16] found that the Lindley distribution is more flexible than the exponential distribution. In recent years, generalizations of the Lindley distribution have been widely used for lifetime data analysis. Asgharzadeh et al. [17] introduced a two parameter generalization of the Lindley distribution, called the new weighted Lindley (NWL) distribution. It is a mixture of the weighted exponential and weighted gamma distributions and a negative mixture of two Lindley distributions with different parameters. The Lindley and weighted Lindley distributions are included as sub-models. Moreover, the Lindley distribution and various generalizations of the Lindley distribution were compared to the NWL distribution. It has provided a better fit than competitive distributions.

In this paper, we propose a new length-biased distribution, called the length-biased weigthed Lindley (LBWL) distribution. It has been developed from the NWL distribution. The rest of this paper is organized as follows, In Section 2 the proposed distribution is introduced. In Section 3 reliability measures and mathematical properties are derived. Parameter estimation is obtained in Section 4. In Section 5 the performance of parameter estimation procedures are evaluated by a simulation study. In Section 6 the proposed distribution is applied to two real-life data sets and compared with other distributions. Conclusions are discussed in Section 7.

2. THE LENGTH-BIASED WEIGHTED LINDLEY DISTRIBUTION

In this section, the probability density function (pdf) and the cumulative distribution function (cdf) of the LBWL distribution are presented.

Definition 1. *A random variable X is said to have the length-biased weighted Lindley (LBWL) distribution, if the pdf of X is*

$$f_l(x) = \frac{\theta^3(1+\alpha)^3}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}x(1+x)(1-e^{-\theta\alpha x})e^{-\theta x}, \tag{2}$$

for $x > 0, \theta > 0, \alpha > 0$.

Note that it is a length-biased distribution to the NWL distribution. Really, the pdf of the NWL distribution is written as

$$f(x) = \frac{\theta^2(1+\alpha)^2}{\theta\alpha(1+\alpha)+\alpha(2+\alpha)}(1+x)(1-e^{-\theta\alpha x})e^{-\theta x}, \tag{3}$$

for $x > 0, \theta > 0, \alpha > 0$. The mean of the NWL distribution is

$$E(X) = \frac{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}{\theta(1+\alpha)[\theta\alpha(1+\alpha)+\alpha(2+\alpha)]}. \tag{4}$$

The LBWL distribution can be obtained by substituting (3) and (4) in (1).

Figure 1 shows various pdf plots of the LBWL distribution with some parameter values. The location mode of the LBWL distribution decreases as θ and α increase.

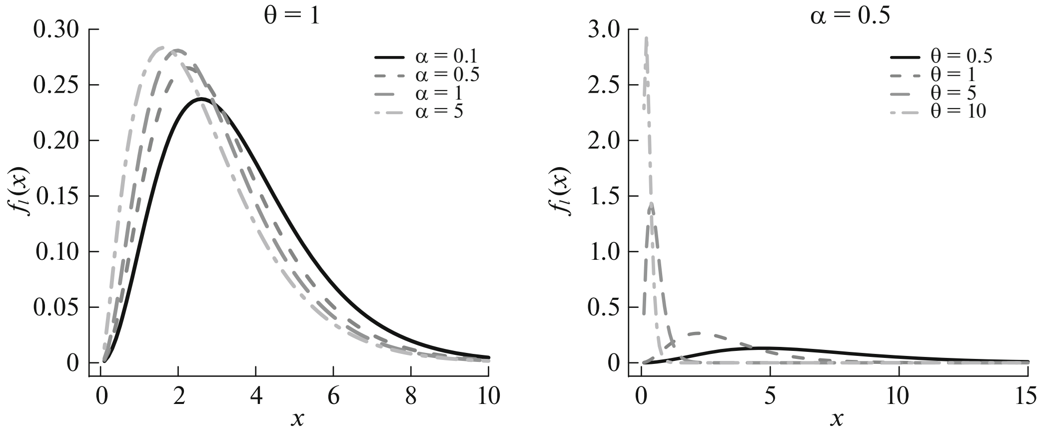

Fig. 1. Some pdf plots of the LBWL distribution with different parameter values.

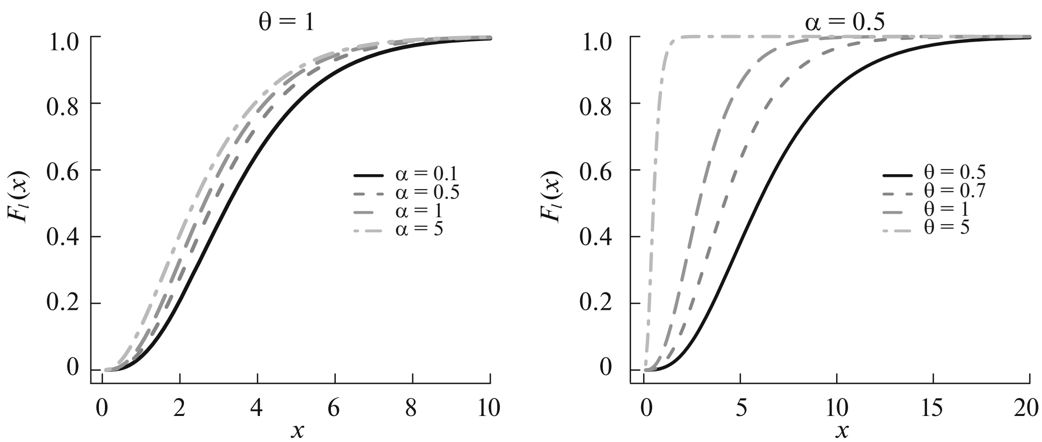

Fig. 2. Some cdf plots of the LBWL distribution with different parameter values.

The corresponding cumulative distribution function of X is obtained by

$$F_l(x) = 1 - \frac{[(\alpha+1)((\alpha+1)^2(\theta x(\theta x+\theta+2)+\theta+2)e^{\alpha\theta x} - \theta(x((\alpha+1)\theta(x+1)+2)+1)) - 2]e^{-(\alpha+1)\theta x}}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2},$$

for $x > 0, \theta > 0, \alpha > 0$.

The cumulative distribution function (cdf) plots of the NWL distribution for different parameter values are displayed in Fig. 2.

3. RELIABILITY MEASURES AND MATHEMATICAL PROPERTIES

In this section, we present reliability measures and mathematical properties of the LBWL distribution, such as the survival function, hazard function, moment generating function, characteristic function and moments.

3.1. Reliability Measures

Here, the survival function, known as the reliability function; and the hazard function, known as the failure rate of the LBWL distribution are presented.

Since the survival function can been obtained by $S_l(x) = 1 - F_l(x)$, the survival function of X is

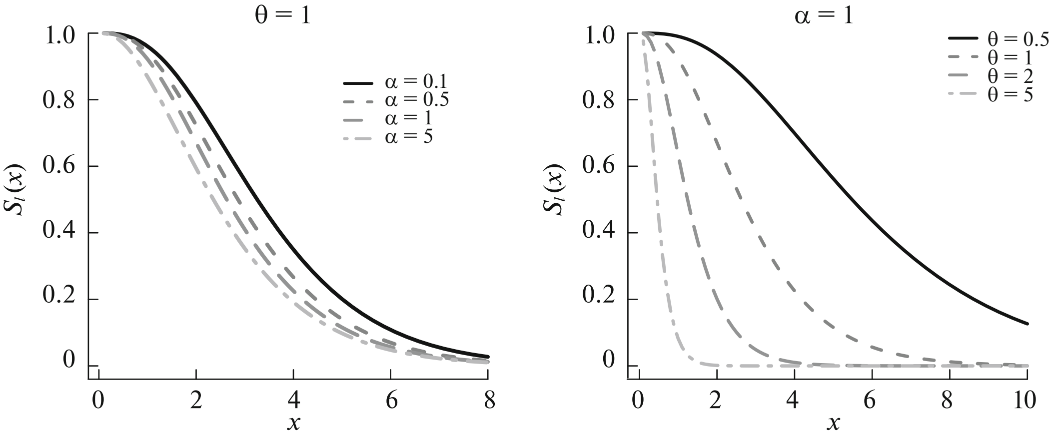

Fig. 3. Some survival function plots of the LBWL distribution with different parameter values.

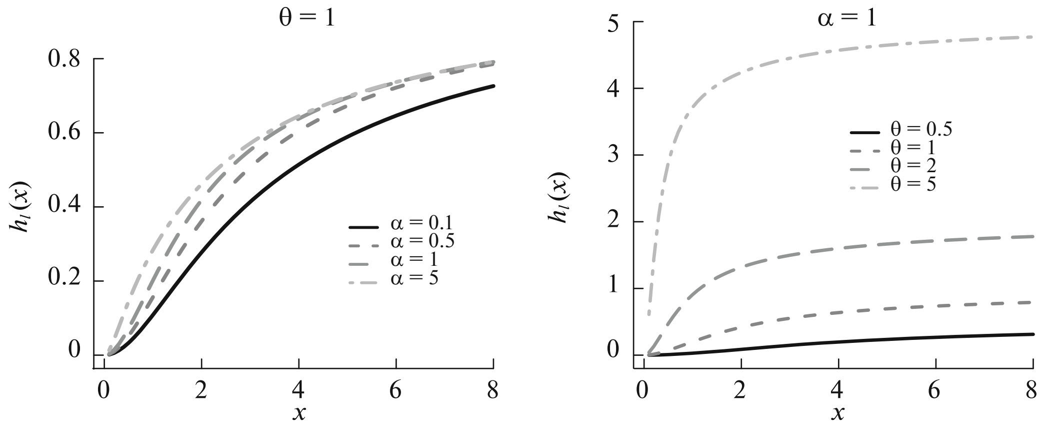

Fig. 4. Some hazard function plots of the LBWL distribution with different parameter values.

$$S_l(x)$$
$$= \frac{[(\alpha+1)((\alpha+1)^2(\theta x(\theta x+\theta+2)+\theta+2)e^{\alpha\theta x}-\theta(x((\alpha+1)\theta(x+1)+2)+1))-2]e^{-(\alpha+1)\theta x}}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}.$$

The hazard function is given by $h_l(x) = f_l(x)/S_l(x)$. Hence, the hazard function of X is

$$h_l(x) = \frac{\theta^3(1+\alpha)^3 x(1+x)(1-e^{-\theta\alpha x})}{[(\alpha+1)((\alpha+1)^2(\theta x(\theta x+\theta+2)+\theta+2)e^{\alpha\theta x}-\theta(x((\alpha+1)\theta(x+1)+2)+1))-2]e^{-\alpha\theta x}}.$$

To study the shape of hazard function; let $\eta(x) = -\dfrac{d\log f(x)}{dx}$, then

$$\frac{d\eta(x)}{dx} = \frac{\alpha^2\theta^2 e^{\alpha\theta x}}{(e^{\alpha\theta x}-1)^2} + \frac{2x^2+2x+1}{x^2(x+1)^2} > 0.$$

Thus, $h(x)$ is an increasing function.

The survival function plot of the LBWL distribution is shown in Fig. 3, whereas Fig. 4 displays the hazard function plot. Fig. 4 indicates that the hazard function is increasing as θ and α increase.

3.2. Moment Generating Function, Characteristic Function and Moments

In this subsection, we present the moment generating function, characteristic function and moments. They can be used for estimation.

Theorem 1. *Let X be a LBWL random variable with parameters (θ, α). Then the moment generating function of X is*

$$M(t) = \frac{\theta^3(1+\alpha)^3}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}\Big[\frac{\theta-t+2}{(\theta-t)^3} - \frac{\theta(1+\alpha)-t+2}{(\theta(1+\alpha)-t)^3}\Big].$$

Proof. Let $X \sim \mathrm{LBWL}(\theta, \alpha)$, then the moment generating function of X is obtained by

$$M(t) = E(e^{tX}) = \int_0^\infty e^{tx}\frac{\theta^3(1+\alpha)^3}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}x(1+x)(1-e^{-\theta\alpha x})e^{-\theta x}dx$$

$$= \frac{\theta^3(1+\alpha)^3}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}\int_0^\infty x(1+x)(1-e^{-\theta\alpha x})e^{(t-\theta)x}dx$$

$$= \frac{\theta^3(1+\alpha)^3}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}\left[\frac{\theta-t+2}{(\theta-t)^3} - \frac{\theta(1+\alpha)-t+2}{(\theta(1+\alpha)-t)^3}\right].$$

□

Similarly, the characteristic function of X is

$$\varphi(t) = \frac{\theta^3(1+\alpha)^3}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}\left[\frac{\theta-it+2}{(\theta-it)^3} - \frac{\theta(1+\alpha)-it+2}{(\theta(1+\alpha)-it)^3}\right].$$

The k^{th}moment about the origin of the LBWL distribution is written as

$$\mathrm{E}(X^k) = \frac{(k+1)!\left[(1+\alpha)^{k+3}(\theta+k+2)-\theta(1+\alpha)-k-2\right]}{\theta^k(1+\alpha)^k\left[(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2\right]}.$$

Thus, the first four moments about the origin for X are

$$\mathrm{E}(X) = \mu = \frac{2\left[(1+\alpha)^4(\theta+3)-\theta(1+\alpha)-3\right]}{\theta(1+\alpha)\left[(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2\right]},$$

$$\mathrm{E}(X^2) = \mu_2' = \frac{6\left[(1+\alpha)^5(\theta+4)-\theta(1+\alpha)-4\right]}{\theta^2(1+\alpha)^2\left[(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2\right]},$$

$$\mathrm{E}(X^3) = \mu_3' = \frac{24\left[(1+\alpha)^6(\theta+5)-\theta(1+\alpha)-5\right]}{\theta^3(1+\alpha)^3\left[(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2\right]},$$

$$\mathrm{E}(X^4) = \mu_4' = \frac{120\left[(1+\alpha)^7(\theta+6)-\theta(1+\alpha)-6\right]}{\theta^4(1+\alpha)^4\left[(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2\right]}.$$

The k^{th} moment about the mean of X is $\mu_k = E(X-\mu)^k = \sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j}\mu_j'\mu^{k-j}$. Therefore, the variance of the LBWL distribution is

$$V(X) = \frac{2(3\delta((\alpha+1)^5(\theta+4)-(\alpha+1)\theta-4)-2((\alpha+1)^4(-(\theta+3))+(\alpha+1)\theta+3)^2)}{(\alpha+1)^2\theta^2\delta^2},$$

where $\delta = (1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2$.

Figure 5 shows plots of the mean and variance of the proposed distribution against the parameters (θ, α). As θ and α increase, the mean and variance decrease.

Figure 6 displays the skewness and kurtosis plots of the LBWL distribution. The skewness and kurtosis are increasing as θ and α increase.

4. PARAMETER ESTIMATION

The parameter estimators of the LBWL distribution are obtained in this section. In this paper, two parameter estimation methods including the method of moments (MoM) and maximum likelihood estimation (MLE), have been used.

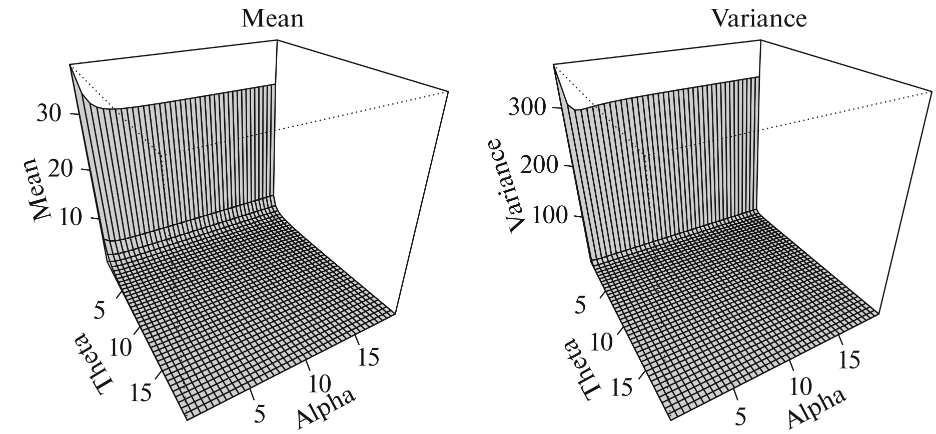

Fig. 5. Mean and variance of the LBWL distribution.

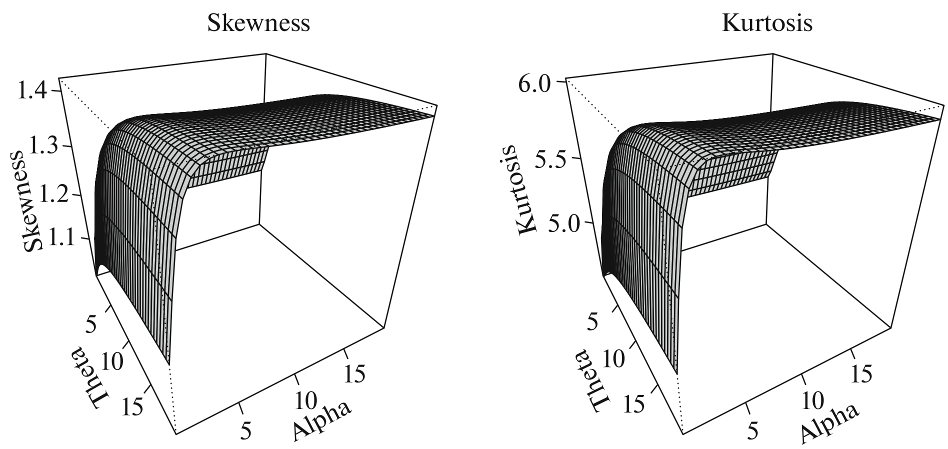

Fig. 6. Skewness and kurtosis of the LBWL distribution.

4.1. Method of Moments

Let $X_1, X_2, \ldots, X_n$ be a random sample from a LBWL distribution with parameters (θ, α). The estimates from the MoM are obtained by solving the first two population moments equal to first two sample moments. Hence, we get

$$\frac{\sum_{i=1}^{n} x_i}{n} = \frac{2\left[(1+\alpha)^4(\theta+3) - \theta(1+\alpha) - 3\right]}{\theta(1+\alpha)\left[(1+\alpha)^3(\theta+2) - (1+\alpha)\theta - 2\right]}$$

and

$$\frac{\sum_{i=1}^{n} x_i^2}{n} = \frac{6\left[(1+\alpha)^5(\theta+4) - \theta(1+\alpha) - 4\right]}{\theta^2(1+\alpha)^2\left[(1+\alpha)^3(\theta+2) - (1+\alpha)\theta - 2\right]}.$$

Since the above equations are too complicated; in this paper we have applied the gmm package [22] from the R statistical software [19] to solve them.

4.2. Maximum Likelihood Estimation

Let $X_1, X_2, \ldots, X_n$ be a random sample from a LBWL distribution with parameters (θ, α). The likelihood function of the LBNL distribution is

$$L(\theta,\alpha)=\prod_{i=1}^{n}\frac{\theta^3(1+\alpha)^3}{(1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2}x(1+x)(1-e^{-\theta\alpha x})e^{-\theta x}.$$

The log-likelihood function is given by

$$\log L(\theta,\alpha)=3n\log(\theta)+3n\log(1+\alpha)-n\log\big((1+\alpha)^3(\theta+2)-(1+\alpha)\theta-2\big)$$
$$+\sum_{i=1}^{n}\log(x_i)+\sum_{i=1}^{n}\log(1+x_i)+\sum_{i=1}^{n}\log(1-e^{-\alpha\theta x_i})-\theta\sum_{i=1}^{n}x.$$

The score functions are

$$\frac{\partial\log L(\theta,\alpha)}{\partial\theta}=\frac{3n}{\theta}-\frac{n\left((\alpha+1)^3-\alpha-1\right)}{(\alpha+1)^3(\theta+2)-(\alpha+1)\theta-2}-\sum_{i=1}^{n}\frac{x_i\left(\mathrm{e}^{\alpha x_i\theta}-\alpha-1\right)}{\mathrm{e}^{\alpha x_i\theta}-1}$$

and

$$\frac{\partial\log L(\theta,\alpha)}{\partial\alpha}=\frac{3n}{\alpha+1}-n\left(\frac{3(\theta+2)(1+\alpha)^2-\theta}{(\theta+2)(\alpha+1)^3-\theta(\alpha+1)-2}\right)+\sum_{i=1}^{n}\frac{\theta x_i}{\mathrm{e}^{\alpha\theta x_i}-1}.$$

They can be solved by numerical methods. In this paper, we have used optimx package [18] from the R statistical software [19] to obtain the maximum likelihood estimates of the LBWL distribution.

As $n\to\infty$, the distribution of $\sqrt{n}(\hat{\theta}-\theta,\hat{\alpha}-\alpha)$ is an asymptotically bivariate normal distribution with zero mean. The variances and covariances of maximum likelihood estimators are calculated by the elements of the inverse of the Fisher information matrix. Although the Fisher information matrix is complicated to obtain, we can replace the Fisher information matrix with the observed information matrix. The observed information matrix of the maximum likelihood estimators of the parameters are defined as

$$J(\theta,\alpha)=\begin{pmatrix}J_{11} & J_{12}\\ J_{21} & J_{22}\end{pmatrix}.$$

Thus, the asymptotic variances and covariances of the maximum likelihood estimators of the parameters are expressed as

$$V(\hat{\theta})=\frac{J_{22}}{\Delta},\quad \mathrm{Cov}(\hat{\theta},\hat{\alpha})=\frac{-J_{12}}{\Delta}\quad\text{and}\quad V(\hat{\alpha})=\frac{J_{11}}{\Delta},$$

where $J_{11}=-\frac{\partial^2\log L(\theta,\alpha)}{\partial\theta^2}$, $J_{12}=J_{21}=-\frac{\partial^2\log L(\theta,\alpha)}{\partial\theta\partial\alpha}$, $J_{22}=-\frac{\partial^2\log L(\theta,\alpha)}{\partial\alpha^2}$ and where Δ is the determinant of matrix J.

The second partial derivative of the log-likelihood with respect to each parameters are

$$\frac{\partial^2\log L(\theta,\alpha)}{\partial\theta^2}=-\frac{3n}{\theta^2}+\frac{n\left((\alpha+1)^3-\alpha-1\right)^2}{((\alpha+1)^3(\theta+2)-(\alpha+1)\theta-2)^2}-\sum_{i=1}^{n}\frac{\alpha^2x_i^2\mathrm{e}^{\alpha x_i\theta}}{\left(\mathrm{e}^{\alpha x_i\theta}-1\right)^2},$$

$$\frac{\partial^2\log L(\theta,\alpha)}{\partial\alpha^2}=-\frac{3n}{(\alpha+1)^2}+\frac{n(3(\theta+2)(1+\alpha)^2-\theta)^2}{((\theta+2)(\alpha+1)^3-\theta(\alpha+1)-2)^2}$$
$$-\frac{6(\theta+2)n(\alpha+1)}{(\theta+2)(\alpha+1)^3-\theta(\alpha+1)-2}-\sum_{i=1}^{n}\frac{\theta^2x_i^2\mathrm{e}^{\alpha\theta x_i}}{\left(\mathrm{e}^{\alpha\theta x_i}-1\right)^2},$$

$$\frac{\partial^2\log L(\theta,\alpha)}{\partial\theta\partial\alpha}=-\frac{2n(2\alpha+3)}{\left((\theta+2)\alpha^2+(3\theta+6)\alpha+2\theta+6\right)^2}-\sum_{i=1}^{n}\frac{x\left((\alpha\theta x_i-1)\,\mathrm{e}^{\alpha\theta x_i}+1\right)}{\left(\mathrm{e}^{\alpha\theta x_i}-1\right)^2}.$$

The asymptotic $100(1-\nu)\%$ confidence intervals (CI) for the parameters with the significance level ν, are given as

$$\hat{\theta} \pm z_{\nu/2}\sqrt{\widehat{V(\hat{\theta})}} \quad \text{and} \quad \hat{\alpha} \pm z_{\nu/2}\sqrt{\widehat{V(\hat{\alpha})}},$$

where $\widehat{V(\hat{\theta})}$ and $\widehat{V(\hat{\alpha})}$ are the maximum likelihood estimators of $V(\hat{\theta})$ and $V(\hat{\alpha})$, respectively ; $z_{\nu/2}$ is the upper $\nu/2$ quantile of the standard normal distribution.

5. SIMULATION STUDY

To evaluate the performance of the two parameter estimation procedures, the MoM and MLE, we generate random samples from the LBWL distribution by using quantile function. We have considered $\theta = 0.5, 1$ and $\alpha = 0.5, 1$ for the different sample sizes, $n = 20, 50, 100$, and 200. The simulation study has been processed for 1,000 iterations. The following measures have been calculated to assess the performance of the parameter estimation methods.

1. Average root mean square error (RMSE) $\sqrt{\dfrac{\sum_{i=1}^{1,000}(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})^2}{1,000}}$.

2. Average bias $\dfrac{\sum_{i=1}^{1,000}(\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta})}{1,000}$.

The same measure have beem applied for the paramter α. The simulated results are shown in Tables 1 and 2.

Tables 1 and 2 show good performance of the MLE for both parameters, θ and α, while the performance of the MoM is quite poor for α.

Table 1. Average RMSE (average bias) of the simulated MoM and MLE estimates for θ

n	θ	α			
		0.5		1	
		MOM	MLE	MOM	MLE
20	0.5	0.0900 (0.0055)	0.1005 (0.0334)	0.1030 (0.0282)	0.1183 (0.0545)
50		0.0624 (−0.0114)	0.0775 (0.0243)	0.0648 (0.0082)	0.0894 (0.0385)
100		0.0520 (−0.0185)	0.0678 (0.0217)	0.0490 (−0.0004)	0.0735 (0.0295)
200		0.0447 (−0.0237)	0.0592 (0.0206)	0.0387 (−0.0046)	0.0592 (0.0225)
20	1	0.2059 (0.0569)	0.2047 (0.0676)	0.2445 (0.1047)	0.2439 (0.1118)
50		0.1411 (0.0223)	0.1578 (0.0492)	0.1631 (0.0542)	0.1830 (0.0796)
100		0.1166 (0.0110)	0.1386 (0.044)	0.1257 (0.0345)	0.1497 (0.0596)
200		0.0933 (0.0045)	0.1200 (0.0417)	0.0964 (0.0202)	0.1208 (0.0457)

Table 2. Average RMSE (average bias) of the simulated MoM and MLE estimates for α

n	θ	α			
		0.5		1	
		MoM	MLE	MoM	MLE
20	0.5	27.8374 (4.7637)	0.0985 (0.0312)	58.2676 (11.9018)	0.4581 (−0.4463)
50		6.5318 (1.5031)	0.0510 (−0.0003)	36.9753 (4.9648)	0.5375 (−0.5364)
100		8.0247 (1.381)	0.0656 (0.0194)	13.8029 (2.6145)	0.4759 (−0.4715)
200		3.0277 (1.073)	0.0583 (0.0203)	9.8293 (1.733)	0.4812 (−0.4781)
20	1	786.3471 (52.493)	0.5960 (0.5648)	189.6021 (47.5014)	0.2390 (0.1116)
50		97.1535 (9.9133)	0.3666 (0.3609)	240.1302 (30.827)	0.1288 (−0.0723)
100		33.3622 (3.4095)	0.5647 (0.5484)	90.0789 (11.5268)	0.1442 (0.0572)
200		11.1476 (0.9822)	0.5564 (0.5450)	23.1019 (2.6924)	0.0860 (−0.0032)

Table 3. Distance between cracks in a pipe

30.94	18.51	16.62	51.56	22.85	22.38	19.08	49.56
17.12	10.67	25.43	10.24	27.47	14.7	14.1	29.93
27.98	36.02	19.4	14.97	22.57	12.26	18.14	18.84

Table 4. Endurance of deep groove ball bearings

17.88	28.92	33.00	41.52	42.12	45.60	48.8	51.84
51.96	54.12	55.56	67.80	68.44	68.64	68.88	84.12
93.12	98.64	105.12	105.84	127.92	128.04	173.4	

6. APPLICATION

In this section, the proposed distribution is compared with three distributions, the Lindley distribution [15], the size-biased Lindley distribution [13] and the new weighted Lindley distribution [17]. Two real-life data sets have been considered in this paper. The first data set is the distance between cracks in a pipe, taken from [21]. The second data set is the number of million revolutions before failure for each one of the 23 ball bearings in endurance test of deep groove ball bearings [23]. These data sets are presented in Tables 3 and 4, respectively.

The following figures are the total time on test (TTT). They have been applied to show the shape of the hazard function of both data sets.

Figure 7 shows the TTT plots for two data sets. The hazard functions of both data sets have an increasing shape. Therefore, both data sets can been described by the LBWL distribution.

To carry out the model selection, we consider the smallest negative log-likelihood, smallest Akaike Information Criterion (AIC), smallest Bayesian Information Criterion (BIC) and largest p-value from the Anderson–Darling (AD) test and Kolmogorov–Smirnov (KS) test. The results are shown in Table 5 and Table 6 for the first and second data sets, respectively.

Table 5 shows that the LBWL distribution has the largest p-value based on the AD and KS tests and the smallest negative log-likelihood, AIC and BIC. Among all these distributions, the Lindley distribution is the least efficient distribution.

Table 5. Parameter estimates, confidence interval, negative log-likelihood, AIC, BIC, AD statistics and KS statistics of all fitted distributions for distance between cracks in a pipe data

Distribution	Estimates (se)	95%CI	Negative log-likelihood	AIC	BIC	AD statistics (p-value)	KS statistics (p-value)
Lindley	$\hat{\theta} = 0.084$ (0.015)	(0.06, 0.107)	92.883	187.767	188.945	1.967 (0.096)	0.24 (0.106)
SBL	$\hat{\theta} = 0.128$ (0.012)	(0.098, 0.158)	89.398	180.796	181.974	0.973 (0.371)	0.162 (0.506)
NWL	$\hat{\theta} = 0.128$ (0.044) $\hat{\alpha} = 0.005$ (0.656)	(0.041, 0.214) (−1.282, 1.291)	89.398	182.796	185.152	0.973 (0.371)	0.162 (0.506)
LBWL	$\hat{\theta} = 0.172$ (0.073) $\hat{\alpha} = 0.001$ (0.832)	(0.028, 0.316) (−1.63, 1.632)	87.793	179.586	181.942	0.548 (0.697)	0.112 (0.890)

Table 6. Parameter estimates, confidence interval, negative log-likelihood, AIC, BIC, AD statistics and KS statistics of all fitted distributions for data on the endurance of deep groove ball bearings

Distribution	Estimates (se)	95%CI	Negative log-likelihood	AIC	BIC	AD statistics (p-value)	KS statistics (p-value)
Lindley	$\hat{\theta} = 0.027$ (0.004)	(0.019, 0.035)	115.736	233.471	234.607	0.932 (0.394)	0.193 (0.318)
SBL	$\hat{\theta} = 0.041$ (0.005)	(0.032, 0.051)	113.578	229.156	230.292	0.311 (0.929)	0.121 (0.849)
NWL	$\hat{\theta} = 0.041$ (0.018) $\hat{\alpha} = 0.009$ (0.870)	(0.005, 0.077) (0, 1.714)	113.578	231.156	233.427	0.311 (0.929)	0.121 (0.849)
LBWL	$\hat{\theta} = 0.049$ (0.033) $\hat{\alpha} = 0.3$ (2.4)	(0, 0.113) (0, 5.003)	113.03	230.059	232.33	0.206 (0.989)	0.12 (0.856)

Table 6 indicates that the LBWL distribution also has the largest p-value based on the AD and KS tests and the smallest negative log-likelihood, whereas the SBL distribution has the smallest AIC and BIC because it only has one parameter. The Lindley distribution provides a worst fit than others.

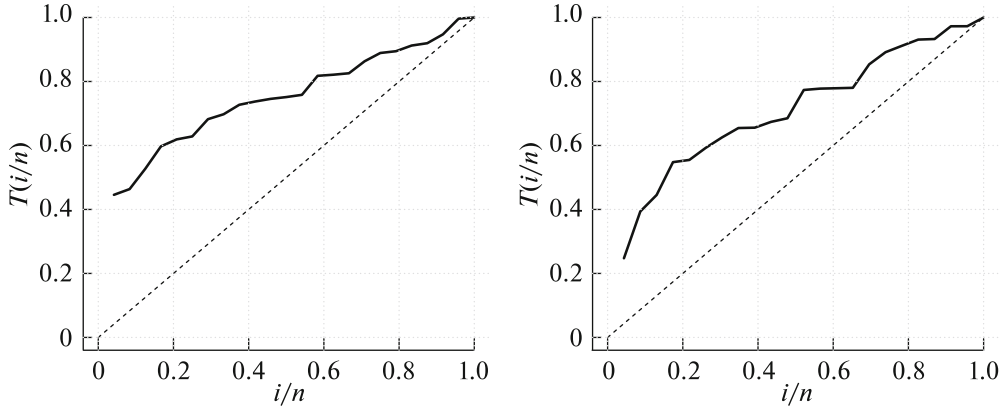

Fig. 7. TTT plots of distance between cracks in a pipe (left) and endurance of deep groove ball bearings (right).

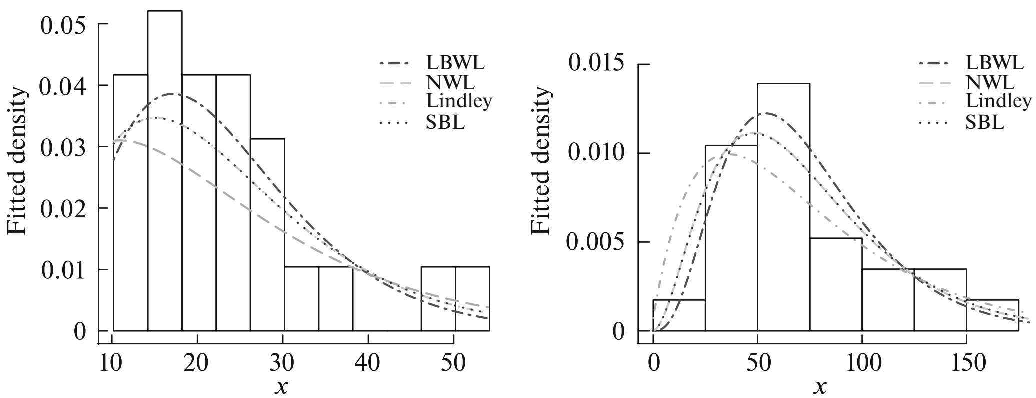

Fig. 8. Histogram and fitted pdfs of the distance between cracks in a pipe (left) and endurance of deep groove ball bearings (right).

The empirical histogram and fitted density of four distributions are displayed in Figure 8. The fitted density of LBWL distribution is closest to the empirical histogram for both data sets.

In addition, when $\alpha \to 0$, the NWL distribution becomes the SBL distribution. For these two data sets, $\hat{\alpha}$ of the NWL distribution is close to zero; therefore, the performance of the NWL and SBL distributions are similar.

7. CONCLUSION

In this paper, we introduce a new length-biased distribution, which we call the length biased weighted Lindley distribution. Some important mathematical properties including moment generating function, characteristic function and moments have been investigated. The estimation of parameters has been discussed by the method of moments and maximum likelihood estimation. The variance and covariance matrix of the maximum likelihood estimates is applied for constructing the aymptotic confidence intervals. Finally, some practical data sets were considered to illustrate the efficacy of the LBWL distribution. The criteria for the selected model are the negative log-likelihood, AIC, BIC and p-value based on the Anderson—Darling and Kolmogorov—Smirnov tests.

The results indicate that the LBWL distribution provides a satisfactory fit in both data sets. Thus, it is an alternative distribution for lifetime data.

ACKNOWLEDGMENTS

The authors thank Kasetsart University and Rajamangala University of Technology Phra Nakhon. Further, the authors are grateful to the reviewers for suggestions of the manuscript.

REFERENCES

1. R. A. Fisher, "The effect of methods of ascertainment upon the estimation of frequencies," Ann. Eugenics **6**, 13–25 (1934).
2. C. R. Rao, "On discrete distributions arising out of methods of ascertainment," Sankhyā: Indian J. Stat., Ser. A, 311–324 (1965).
3. G. P. Patil and J. K. Ord, "On size-biased sampling and related form-invariant weighted distributions," Sankhyā: Indian J. Stat., Ser. B, 48–61 (1976).
4. K. K. Das and T. D. Roy, "Applicability of length biased weighted generalized Rayleigh distribution," Adv. Appl. Sci. Res. **2**, 320–327 (2011).
5. K. K. Das and T. D. Roy, "On some length-biased weighted Weibull distribution," Adv. Appl. Sci. Res. **2**, 465–475 (2011).
6. M. V. Ratnaparkhi and U. V. Naik-Nimbalkar, "The length-biased lognormal distribution and its application in the analysis of data from oil field exploration studies," J. Mod. Appl. Stat. Methods **11**, 255–260 (2012).
7. A. K. Al-Khadim and A. N. Hussein, "New proposed length biased weighted exponential and Rayleigh distribution with application," Math. Theory Modeling **4**, 2224–2235 (2014).
8. N. Nanuwong and W. Bodhisuwan, "Length biased beta-pareto distribution and its structural properties with application," J. Math. Stat. **10**, 49 (2014).
9. P. Seenoi, T. Supapakorn, and W. Bodhisuwan, "The length-biased exponentiated inverted Weibull distribution," Int. J. Pure Appl. Math. **92**, 191–206 (2014).
10. K. Modi and V. Gill, "Length-biased weighted Maxwell distribution," Pakistan J. Stat. Oper. Res. **11**, 465–472 (2015).
11. A. Ahmad, S. P. Ahmad, and A. Ahmed, "Length-biased weighted Lomax distribution: statistical properties and application," Pakistan J.Stat. Oper. Res. **12**, 245–255 (2016).
12. A. Saghir, S. Tazeem, and I. Ahmad, "The length-biased weighted exponentiated inverted Weibull distribution," Cogent Math. **3**, 1267299 (2016).
13. A. Ayesha, "Size biased lindley distribution and its properties a special case of weighted distribution," Appl. Math. **8**, 808–819 (2017).
14. A. A. Rather and C. Subramanian, "Length biased Sushila distribution," Universal Rev. **7**, 1010–1023 (2018).
15. D. V. Lindley, "Fiducial distributions and Bayes' theorem," J. R. Stat. Soc., Ser. B: Methodol., 102–107 (1958).
16. M. E. Ghitany, B. Atieh, and S. Nadarajah, "Lindley distribution and its application," Math. Comput. Simul. **78**, 493–506 (2008).
17. A. Asgharzadeh, H. S. Bakouch, S. Nadarajah, and F. Sharafi, "A new weighted Lindley distribution with application," Brazil. J. Prob. Stat. **30**, 1–27 (2016).
18. J. C. Nash and R. Varadhan, "Unifying optimization algorithms to aid software system users: optimx for R," J. Stat. Software **43** (9), 1–14 (2011).
19. R Core Team, *R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing* (Vienna, Austria, 2018). https://www.R-project.org/. Accessed 2019.
20. E. L. Lehmann and G. E. Casella, *Theory of Point Estimation* (Springer Science, New York, 2006).
21. D. P. Murthy, M. Bulmer, and J. A. Eccleston, "Weibull model selection for reliability modelling," Reliab. Eng. Syst. Safety **86**, 257–267 (2004).
22. P. Chaussé, "Computing generalized method of moments and generalized empirical likelihood with R," J. Stat. Software **34** (11), 1–35 (2010).
23. J. F. E. Lawless, *Statistical Models and Methods for Lifetime Data* (Wiley, New Jersey, 2011).