

Article

Estimation Parameter of $R = P(Y < X)$ for Length-Biased Weighted Lomax Distributions in the Presence of Outliers

Hossein Karimi and Parviz Nasiri *

Department of Statistics, Payame Noor University (PNU), 19395-4697 Tehran, Iran; hkarimi@pnu.ac.ir

* Correspondence: pnasiri45@yahoo.com

Received: 22 November 2017; Accepted: 14 February 2018; Published: 16 February 2018

Abstract: The concept of length-biased distribution is applied in expending proper models for lifetime data. The length-biased distribution is a special case of well-known weighted distribution. In this article, we introduce a length-biased weighted Lomax distribution (LBWLD) with k presence of outliers and estimate the parameter of $R = P(Y < X)$ when the random variables X and Y are independent and have LBWLD in presence of outliers and without outliers, respectively. The bias and mean square error (MSE) of the estimator are examined with simulations of numerical and bootstrap resampling. Analysis of a real data set is considered for illustrative purposes.

Keywords: Lomax distribution; length-biased weighted Lomax distribution; outliers; maximum likelihood estimation; mean squared error

1. Introduction

In recent years, the Lomax distribution (Pareto distribution II), has been applied in the field of life testing because it fits the business failure data. The Lomax distribution is presented by Lomax [1]. In 2010, Abdullah and Abdullah [2] estimated the parameters of the Lomax distribution based on generalized probability weighted moments. The Lomax distribution was applied in the theoretical background in a number of ways. Weighted distribution theory gives a unified approach to deal with model specification and data interpretation problems. Weighted distributions are frequent in studies related to reliability, survival analysis, analysis of family data, biomedicine, ecology and several other areas; Stene [3] and Oluyede and George [4]. Gupta and Tripathi [5], Patil and Rao [6] and Das and Roy [7] researched the length-biased Weighted Generalized Rayleigh distribution with its properties. Moreover, they devised a length-biased Weighted Weibull distribution [8]. Many authors have researched important findings on weighted distributions. A unified concept of weighted distribution was proposed by Rao [9] where sampling situations example of being modeled are identified through weighted distributions. In the context of reliability, the life of a component which has a random strength X and is subject to a random stress Y is described by stress-strength models. The failed component at the instant means the stress is assigned to exceed the strength, and in case $X > Y$, the component functions properly. Consequently, $R = P(Y < X)$ is a measure of component reliability. The parameter of R is the reliability parameter. This type of function can be of practical importance in many applications. For example, provided that X is the response for a control group, and Y is the treatment group, then $R = P(Y < X)$, which is a measure of the effect of the treatment. This $R = P(Y < X)$ can be applied when estimating heritability of a genetic trait. For more information on R , Halperin et al. [10], Simonoff et al. [11], Reiser and Farragi [12] and Bamber [13] are recommended. Bamber [13] provides a geometrical interpretation of $A(X, Y) = P(X < Y) + \frac{1}{2}P(X = Y)$ and reveals that $A(X, Y)$ is a practical formula in measuring the size of the difference between two populations. Weerahandi and Johnson [14] suggested inferential procedures for $P(X > Y)$ by considering

that X and Y are independent normal random variables. Nadarajah [15] examined the reliability for Laplace distribution and its generalizations, where he proposes the double Lomax distribution and introduces its appliance to the stress-strength model. The double Lomax distribution is the ratio independent and analogously classical Laplace distributions. For given random variables X and Y , the distribution ratio X/Y is of absorption in biological and physical sciences, econometrics, and ranking and selection. The archetypes in the context consist of Mendelian inheritance ratios in genetics, mass to energy ratios in nuclear physics, target to control precipitation in meteorology, inventory ratios in economics, and the stress-strength model in the realm of reliability. The function $R = P(X > Y)$ is of practical use in many cases like clinical trials, genetics, and reliability. Let X be a nonnegative random variable with its probability density function (PDF) of $f(x)$, and then the PDF of the weight random variable X_w yields through

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))}; x \geq 0,$$

where

$$E(w(x)) = \int_0^\infty w(x)f(x)dx; 0 < E(w(x)) < \infty.$$

The random variable X_w is named the weight version of X and its distribution corresponds to that of X and is named the weighted distribution with weight function $w(x)$. Consequently, at $w(x) = x$, the obtained distribution is named a length-biased distribution and the PDF of a length-biased random variable X yields

$$f(x) = \frac{xf(x)}{E(X)}; x \geq 0.$$

1.1. Lomax Distribution and Properties

A random variable X is distributed as Lomax distribution (LD) with two parameters of θ (shape parameter) and λ (scale parameter) if it is of the following PDF

$$f(x; \theta, \lambda) = \frac{\theta}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\theta+1)}; x \geq 0, \theta, \lambda > 0.$$

It is easy to show that

$$E(X) = \frac{\lambda}{\theta - 1}; \theta > 1, \quad Var(X) = \frac{\lambda^2 \theta}{(\theta - 1)^2 (\theta - 2)}; \theta > 2.$$

The corresponding cumulative distribution function (CDF) is exposed through

$$F(x; \theta, \lambda) = 1 - \left(1 + \frac{x}{\lambda}\right)^{-\theta}; \theta > 2.$$

It is assumed that (X_1, \dots, X_n) is a random sample of size n from LD, where the maximum likelihood estimators (MLEs) of the parameters of λ and θ , represented by $\hat{\lambda}$ and $\hat{\theta}$, are obtained through direct maximization of the log-likelihood function. Different log-likelihood functions with respect to parameters are achieved through normal equations

$$\frac{n}{\theta} - \sum_{i=1}^n \log\left(1 + \frac{x_i}{\lambda}\right) = 0, \quad \frac{n}{\lambda} - (\theta + 1) \sum_{i=1}^n \left(\frac{x_i}{\lambda(\lambda + x_i)}\right) = 0. \quad (1)$$

The MLEs are simultaneous solutions of the nonlinear normal Equation (1). So, by simplification of (1), we have

$$\hat{\theta}(\lambda) = \frac{n}{\sum_{i=1}^n \log\left(1 + \frac{x_i}{\lambda}\right)}, \quad \hat{\lambda}(\theta) = \frac{n}{(\theta + 1) \sum_{i=1}^n \left(\frac{x_i}{\lambda(\lambda + x_i)}\right)}.$$

The estimators cannot be written in a closed-form expression.

1.2. LBWLD and Properties

An LBWLD is derived by weights of $w(x) = x$ in the weighted Lomax distribution [16]

$$f(x; \theta, \lambda) = \frac{\theta(\theta-1)}{\lambda^2} x \left(1 + \frac{x}{\lambda}\right)^{-(\theta+1)}; \quad x \geq 0, \theta > 1, \lambda > 0.$$

It is easy to show that

$$E(X) = \frac{2\lambda}{\theta-2}; \quad \theta > 2, \quad \text{Var}(X) = \frac{2\lambda^2\theta}{(\theta-2)^2(\theta-3)}; \quad \theta > 3.$$

The corresponding CDF is

$$F(x; \theta, \lambda) = 1 - \left(1 + \frac{x}{\lambda}\right)^{-\theta} \left(1 + \frac{x\theta}{\lambda}\right),$$

where θ and λ are the shape and scale parameters, respectively.

Let (Y_1, \dots, Y_m) be a random sample of size m from LBWLD with two parameters of θ and λ . The log-likelihood function would be

$$l(\theta, \lambda) = m \log(\theta(\theta-1)) - 2m \log \lambda + \sum_{i=1}^m \log y_i - (\theta+1) \sum_{i=1}^m \log \left(1 + \frac{y_i}{\lambda}\right).$$

The partial derivatives of $l(\theta, \lambda)$ with respect to θ and λ are, respectively

$$\frac{m(2\theta-1)}{\theta(\theta-1)} - \sum_{i=1}^m \log \left(1 + \frac{y_i}{\lambda}\right) = 0, \quad \frac{2m}{\lambda} - (\theta+1) \sum_{i=1}^m \left(\frac{y_i}{\lambda(\lambda+y_i)}\right) = 0.$$

The MLE $\hat{\eta} = (\hat{\theta}, \hat{\lambda})$ of $\eta = (\theta, \lambda)$ is obtained with simultaneous solutions nonlinear system of normal equations, given by

$$\frac{m(2\theta-1)}{\theta(\theta-1)} = \sum_{i=1}^m \log \left(1 + \frac{y_i}{\lambda}\right), \quad \frac{2m}{\lambda} = (\theta+1) \sum_{i=1}^m \left(\frac{y_i}{\lambda(\lambda+y_i)}\right). \quad (2)$$

2. Joint Distribution of Random Variables (X_1, \dots, X_n) in Presence of k Outliers

According to Dixit [17], let random variables (X_1, \dots, X_n) be such that k of them (the number of outliers) are distributed as LBWLD with parameters of β (shape parameter) and λ (scale parameter) and the remaining $(n-k)$ random variables are distributed as LBWLD with parameters of α (shape parameter) and λ (scale parameter). The joint distribution of (X_1, \dots, X_n) with k outliers is expressed as

$$f(\underline{x}) = C(n, k)^{-1} \frac{(\beta(\beta-1))^k (\alpha(\alpha-1))^{n-k}}{\lambda^2} \prod_{i=1}^n x_i \left(1 + \frac{x_i}{\lambda}\right)^{-(\alpha+1)} \Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{-\beta},$$

where $\Sigma^* = \sum_{A_1=1}^{n-k+1} \sum_{A_2=A_1+1}^{n-k+2} \dots \sum_{A_k=A_{k-1}+1}^n$.

It is easy to show that marginal distribution of X is expressed as [18]

$$f(x; \alpha, \beta, \lambda) = \frac{\bar{b}\alpha(\alpha-1)}{\lambda^2} x \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)} + \frac{b\beta(\beta-1)}{\lambda^2} x \left(1 + \frac{x}{\lambda}\right)^{-(\beta+1)}; \quad x \geq 0, \alpha, \beta > 1, \lambda > 0. \quad (3)$$

The CDF is

$$F(x; \alpha, \beta, \lambda) = \bar{b} \left(1 - \left(1 + \frac{x}{\lambda}\right)^{-\alpha} \left(1 + \frac{x\alpha}{\lambda}\right)\right) + b \left(1 - \left(1 + \frac{x}{\lambda}\right)^{-\beta} \left(1 + \frac{x\beta}{\lambda}\right)\right),$$

where $b = \frac{k}{n}$ and $\bar{b} = 1 - b$.

3. Method of Maximum Likelihood Estimation (MLE)

The MLEs of the parameters are achieved through direct maximization of the log-likelihood function. Partial derivations of log-likelihood functions with respect to parameters and setting the results equal to zero will produce normal equations. The MLE is the result of solving a nonlinear system of normal equations. It is more practical to apply nonlinear optimization algorithms like the quasi-Newton algorithm to maximize the log-likelihood function in its numerical sense.

Provided that (X_1, \dots, X_n) is a random sample of size n from LBWLD with presence of k outliers, then the log-likelihood function is presented as

$$l(\alpha, \beta, \lambda) = \log L(\alpha, \beta, \lambda) = -\log C(n, k) + k \log(\beta^2 - \beta) + (n - k) \log(\alpha^2 - \alpha) - 2n \log \lambda + \sum_{i=1}^n \log x_i - (\alpha + 1) \sum_{i=1}^n \log(1 + \frac{x_i}{\lambda}) + \log \Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta}.$$

The partial derivatives of $l(\alpha, \beta, \lambda)$ with respect to α, β and λ are as

$$\frac{\partial l(\alpha, \beta, \lambda)}{\partial \alpha} = \frac{(n - k)(2\alpha - 1)}{\alpha(\alpha - 1)} - \sum_{i=1}^n \log(1 + \frac{x_i}{\lambda}) + \frac{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta} \log \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)}{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta}} = 0,$$

$$\frac{\partial l(\alpha, \beta, \lambda)}{\partial \beta} = \frac{k(2\beta - 1)}{\beta(\beta - 1)} - \frac{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta} \log \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)}{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta}} = 0,$$

$$\frac{\partial l(\alpha, \beta, \lambda)}{\partial \lambda} = -\frac{2n}{\lambda} + (\alpha + 1) \sum_{i=1}^n \frac{x_i}{\lambda(\lambda + x_i)} - \frac{(\alpha - \beta) \Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta} \sum_{r=1}^k \frac{x_{A_r}}{\lambda(\lambda + x_{A_r})}}{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta}} = 0.$$

The MLE $\hat{\eta} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda})$ of $\eta = (\alpha, \beta, \lambda)$ is obtained as simultaneous solutions in a nonlinear system of normal equations, given by

$$\frac{(n - k)(2\alpha - 1)}{\alpha(\alpha - 1)} + \frac{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta} \log \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)}{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta}} = \sum_{i=1}^n \log(1 + \frac{x_i}{\lambda}), \quad (4)$$

$$\frac{k(2\beta - 1)}{\beta(\beta - 1)} = \frac{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta} \log \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)}{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta}}, \quad (5)$$

$$\lambda(\alpha + 1) \sum_{i=1}^n \frac{x_i}{\lambda(\lambda + x_i)} - \frac{\lambda(\alpha - \beta) \Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta} \sum_{r=1}^k \frac{x_{A_r}}{\lambda(\lambda + x_{A_r})}}{\Sigma^* \prod_{r=1}^k \left(1 + \frac{x_{A_r}}{\lambda}\right)^{\alpha - \beta}} = 2n. \quad (6)$$

Maximum Likelihood Estimator of $R = P(Y < X)$

Let the random variables X and Y be independent and Y be LBWLD to parameters of (θ, λ) and the random variables (X_1, \dots, X_n) are such that k of them are distributed as LBWLD to parameters of (β, λ) , and the remaining $(n - k)$ random variables are distributed as LBWLD with parameters of (α, λ) . The parameters of (θ, α, β) are the shape parameters and parameter of λ is the scale parameter. Using the marginal distribution of X given by (3), we obtain

$$R = P(Y < X) = \int_0^\infty \int_0^x \frac{\theta(\theta-1)}{\lambda^2} y \left(1 + \frac{y}{\lambda}\right)^{-(\theta+1)} dy \left[\frac{b\alpha(\alpha-1)}{\lambda^2} x \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)} + \frac{b\beta(\beta-1)}{\lambda^2} x \left(1 + \frac{x}{\lambda}\right)^{-(\beta+1)} \right] dx =$$

$$\frac{b\theta(\theta-1)(\theta+3\beta-2)}{(\beta+\theta)(1-(\beta+\theta))(2-(\beta+\theta))} + \frac{\bar{b}\theta(\theta-1)(\theta+3\alpha-2)}{(\alpha+\theta)(1-(\alpha+\theta))(2-(\alpha+\theta))}.$$

The value of R does not depend on parameter of λ .

If we insert $k = 0$, then X and Y will be LBWLD, without outliers, with different shape parameters and identical scale parameter. So

$$R = P(Y < X) = \frac{\theta(\theta-1)(\theta+3-2)}{(\alpha+\theta)(1-(\alpha+\theta))(2-(\alpha+\theta))}.$$

Suppose (Y_1, \dots, Y_m) is a random sample for random variable Y from LBWLD to parameters of (θ, λ) . The MLEs $(\hat{\theta}, \hat{\alpha}, \hat{\beta})$ of the parameters of (θ, α, β) are yielded with simultaneous solutions in nonlinear system of normal equations, (2) and (4)–(6). If (θ, α, β) replaces $(\hat{\theta}, \hat{\alpha}, \hat{\beta})$ into R, then they will obtain the MLE \hat{R} of the parameter of R.

4. Numerical Experiments

The Equation $F(x) = u$, where u is an observation as the uniform distribution (0, 1) and $F(x)$ is the CDF of LBWLD with Presence of Outliers distributions, is applied in running simulation by generating random samples that follow LBWLD with Presence of Outliers distributions. The iterations of this simulation are 1000 per sample size, $(m, n) = (10, 10), (20, 20), (50, 50), (10, 30), (10, 50), (30, 10), (50, 10)$, in two cases of known and unknown λ , for $k = 0, 1, 2$; $\lambda = 2$; $\alpha = 7$; $\beta = 5$, and $\theta = 4$. The averages bias and the means squared error of estimations of parameters of $\alpha, \beta, \theta, \lambda$, and R were computed for these simulations. The results have been truncated after four decimal places.

4.1. Bootstrap Procedure

We consider the bootstrap resampling method for deriving the bias-corrected MLEs based on the idea of Efron [19]. Let $x = \{x_1, \dots, x_n\}$ be a random sample of size n from the random variable X with distribution function F and $\hat{\tau} = h(x)$ be an estimator of parameter $\tau = g(F)$ that is a function of F known. In this procedure, we generate a large number of pseudo-samples $x^* = \{x_{*1}, \dots, x_{*n}\}$ from the sample x and compute the corresponding bootstrap replicates $\hat{\tau}^*$, say $\hat{\tau}^* = h(x^*)$. Thus, the empirical distribution of $\hat{\tau}^*$ is used to estimate the distribution function of $\hat{\tau}$. Therefore, we generate N bootstrap samples independently from the sample x and compute the corresponding bootstrap estimates $(\hat{\tau}^{*(1)}, \dots, \hat{\tau}^{*(N)})$. The second-order bias-corrected MLEs for τ can be obtained as $\tau^B = 2\hat{\tau} - \bar{\hat{\tau}^*}$, where $\bar{\hat{\tau}^*} = \frac{1}{N} \sum_{i=1}^N \hat{\tau}^{*(i)}$ [20].

4.2. Results

Tables 1 and 2 present the values of the averages bias and the MSEs of the estimates for LBWLD with outliers in two cases of known and unknown λ and values of bootstrap method based on 500 resampling are reported within brackets. The values of MSEs of the estimations of $\alpha, \beta, \theta, \lambda$, and R often decrease while sample sizes n and m increase. Average biases and MSEs decrease by applying a bootstrap procedure. The sampling experiments and all computations were performed using R (V. 3.4.3) software.

Table 1. Average Bias and mean square errors (MSEs) for $(\hat{\theta}, \hat{\alpha}, \hat{\beta}, \hat{R})$ (λ is known, $\lambda = 2, \alpha = 7, \beta = 5, \theta = 4$).

k	(m, n)	Bias($\hat{\theta}$)	MSE($\hat{\theta}$)	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Bias(\hat{R})	MSE(\hat{R})
0	(10, 10)	0.1777 (0.1035)	0.6642 (0.5926)	0.4956 (0.1643)	2.879 (2.6043)			0.002 (0.0012)	0.0089 (0.0078)
	(20, 20)	0.0965 (0.0539)	0.2846 (0.2601)	0.2252 (0.0829)	1.1884 (1.1737)			0.0023 (0.0008)	0.004 (0.0033)
	(50, 50)	0.0257 (0.0094)	0.0903 (0.0579)	−0.02 (−0.0081)	0.4169 (0.3652)			0.0053 (0.0021)	0.0018 (0.0016)
	(10, 30)	0.2630 (0.1491)	0.9787 (0.9532)	−0.0728 (−0.0235)	0.6040 (0.6004)	-	-	0.0256 (0.0073)	0.0083 (0.0065)
	(10, 50)	0.4301 (0.1463)	0.769 (0.6011)	0.0158 (0.0074)	0.4197 (0.3891)			0.0129 (0.0049)	0.0061 (0.0057)
	(30, 10)	0.0427 (0.0197)	0.1508 (0.0982)	0.3797 (0.0683)	2.5054 (2.3406)			−0.0034 (−0.0007)	0.0068 (0.0049)
	(50, 10)	0.0143 (0.0095)	0.102 (0.007)	0.3671 (0.1105)	2.4787 (2.2834)			−0.0038 (−0.0012)	0.0054 (0.0039)
	(10, 10)	0.2798 (0.0238)	0.8891 (0.6916)	−0.7307 (−0.2837)	1.0846 (1.04613)	1.3279 (0.4927)	3.0497 (2.5729)	0.0508 (0.0216)	0.0114 (0.0089)
	(20, 20)	−0.166 (−0.0734)	0.2775 (0.2376)	−0.8895 (−0.3156)	1.0125 (0.8296)	0.9462 (0.0816)	2.0133 (2.0019)	0.0238 (0.0101)	0.0034 (0.003)
	(50, 50)	0.082 (0.0164)	0.1261 (0.1079)	−0.232 (−0.1067)	0.393 (0.2162)	0.5182 (0.2376)	2.0159 (1.9437)	0.0143 (0.0071)	0.0030 (0.0017)
	(10, 30)	0.2042 (0.0794)	0.8145 (0.7001)	−0.2779 (−0.1106)	0.5969 (0.3941)	0.6013 (0.2389)	2.7171 (2.0437)	0.0315 (0.0068)	0.0094 (0.0085)
	(10, 50)	−0.2365 (−0.0857)	0.3066 (0.2477)	−0.1874 (−0.0736)	0.2956 (0.2577)	1.5741 (0.5642)	3.0463 (2.1634)	−0.0144 (−0.0069)	0.0036 (0.0027)
1	(30, 10)	0.1639 (0.0699)	0.2575 (0.2262)	−0.5999 (−0.2638)	0.9701 (0.6957)	0.9584 (0.3766)	3.5822 (3.0421)	0.0414 (0.0183)	0.0059 (0.0051)
	(50, 10)	0.0149 (0.0057)	0.1065 (0.0891)	−0.8144 (−0.1149)	1.1065 (1.0236)	1.117 (0.2459)	2.8738 (2.5882)	0.0364 (0.0134)	0.0037 (0.0023)
	(10, 10)	0.2307 (0.0786)	0.5697 (0.5531)	−1.0046 (−0.4169)	1.6083 (1.2681)	1.5446 (0.8837)	7.1450 (5.4953)	0.0509 (0.0221)	0.0101 (0.0082)
	(20, 20)	−0.0437 (−0.0193)	0.2383 (0.2046)	−0.5288 (−0.1937)	0.6237 (0.6176)	1.7306 (0.7308)	4.6385 (3.9176)	0.0097 (0.0037)	0.0024 (0.0019)
	(50, 50)	0.1315 (0.0533)	0.109 (0.0874)	−0.4862 (−0.1303)	0.4466 (0.3583)	1.4114 (0.4692)	2.7341 (2.1346)	0.033 (0.0049)	0.0021 (0.0017)
	(10, 30)	0.2540 (0.0841)	0.8483 (0.6729)	−0.4458 (−0.1197)	0.6660 (0.642)	1.1649 (0.0974)	2.7889 (2.0858)	0.0374 (0.0173)	0.0078 (0.0055)
	(10, 50)	0.3617 (0.0519)	0.8422 (0.6503)	−0.3122 (−0.1804)	0.3586 (0.3042)	1.2256 (0.604)	2.4825 (1.9925)	0.0418 (0.0138)	0.0069 (0.0052)
	(30, 10)	0.0465 (0.0081)	0.1896 (0.1364)	−0.9052 (−0.3972)	1.8056 (1.4519)	0.5538 (0.2172)	3.0168 (2.0826)	0.0443 (0.0199)	0.0071 (0.0067)
	(50, 10)	0.087 (0.0026)	0.1125 (0.1027)	−0.8716 (−0.3304)	1.6284 (1.3726)	0.8903 (0.3371)	3.5537 (2.6735)	0.0411 (0.0153)	0.0064 (0.0049)
	(10, 10)	0.2307 (0.0786)	0.5697 (0.5531)	−1.0046 (−0.4169)	1.6083 (1.2681)	1.5446 (0.8837)	7.1450 (5.4953)	0.0509 (0.0221)	0.0101 (0.0082)
	(20, 20)	−0.0437 (−0.0193)	0.2383 (0.2046)	−0.5288 (−0.1937)	0.6237 (0.6176)	1.7306 (0.7308)	4.6385 (3.9176)	0.0097 (0.0037)	0.0024 (0.0019)
	(50, 50)	0.1315 (0.0533)	0.109 (0.0874)	−0.4862 (−0.1303)	0.4466 (0.3583)	1.4114 (0.4692)	2.7341 (2.1346)	0.033 (0.0049)	0.0021 (0.0017)
	(10, 30)	0.2540 (0.0841)	0.8483 (0.6729)	−0.4458 (−0.1197)	0.6660 (0.642)	1.1649 (0.0974)	2.7889 (2.0858)	0.0374 (0.0173)	0.0078 (0.0055)
	(10, 50)	0.3617 (0.0519)	0.8422 (0.6503)	−0.3122 (−0.1804)	0.3586 (0.3042)	1.2256 (0.604)	2.4825 (1.9925)	0.0418 (0.0138)	0.0069 (0.0052)
	(30, 10)	0.0465 (0.0081)	0.1896 (0.1364)	−0.9052 (−0.3972)	1.8056 (1.4519)	0.5538 (0.2172)	3.0168 (2.0826)	0.0443 (0.0199)	0.0071 (0.0067)
	(50, 10)	0.087 (0.0026)	0.1125 (0.1027)	−0.8716 (−0.3304)	1.6284 (1.3726)	0.8903 (0.3371)	3.5537 (2.6735)	0.0411 (0.0153)	0.0064 (0.0049)

Table 2. Average Bias and MSEs for $(\hat{\theta}, \hat{\alpha}, \hat{\beta}, \hat{\lambda}, \hat{R})$ (λ is unknown, $\lambda = 2, \alpha = 7, \beta = 5, \theta = 4$).

k	(m, n)	Bias($\hat{\theta}$)	MSE($\hat{\theta}$)	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Bias($\hat{\lambda}$)	MSE($\hat{\lambda}$)	Bias(\hat{R})	MSE(\hat{R})
0	(10, 10)	1.4208 (0.5043)	9.2244 (8.1636)	1.1415 (0.5819)	16.6559 (12.6429)			0.9217 (0.3706)	6.2925 (5.7923)	0.0770 (0.0169)	0.0381 (0.0308)
	(20, 20)	1.0802 (0.2982)	8.5436 (6.8431)	1.1970 (0.7441)	16.3828 (13.1642)			0.7274 (0.4467)	4.1251 (3.909)	0.0463 (0.0187)	0.0359 (0.0235)
	(50, 50)	0.5604 (0.2013)	2.1340 (1.8906)	1.1679 (0.4927)	10.6799 (9.699)			0.4568 (0.1692)	1.6848 (1.5617)	0.0063 (0.0023)	0.0049 (0.0043)
	(10, 30)	0.6324 (0.2964)	5.0385 (4.0052)	0.3961 (0.1928)	12.3881 (10.9871)	-	-	0.2262 (0.0672)	1.9292 (1.3863)	0.0388 (0.0213)	0.0104 (0.0076)
	(10, 50)	0.4301 (0.1627)	3.3806 (1.8726)	0.5893 (0.2213)	11.6472 (9.1627)			0.2621 (0.1168)	1.7616 (1.3607)	0.0185 (0.0069)	0.0064 (0.0053)
	(30, 10)	0.3907 (0.1927)	2.8727 (2.5102)	0.959 (0.6132)	13.3461 (11.7429)			0.3507 (0.2265)	1.7833 (1.563)	0.0023 (0.0008)	0.0042 (0.0033)
	(50, 10)	0.4464 (0.1762)	2.6139 (2.1472)	1.3108 (0.5106)	12.7951 (11.0624)			0.3811 (0.2361)	1.6407 (1.2391)	−0.006 (−0.0021)	0.0035 (0.0034)
	(10, 10)	1.4208 (0.5043)	9.2244 (8.1636)	1.1415 (0.5819)	16.6559 (12.6429)			0.9217 (0.3706)	6.2925 (5.7923)	0.0770 (0.0169)	0.0381 (0.0308)
	(20, 20)	1.0802 (0.2982)	8.5436 (6.8431)	1.1970 (0.7441)	16.3828 (13.1642)			0.7274 (0.4467)	4.1251 (3.909)	0.0463 (0.0187)	0.0359 (0.0235)
	(50, 50)	0.5604 (0.2013)	2.1340 (1.8906)	1.1679 (0.4927)	10.6799 (9.699)			0.4568 (0.1692)	1.6848 (1.5617)	0.0063 (0.0023)	0.0049 (0.0043)
	(10, 30)	0.6324 (0.2964)	5.0385 (4.0052)	0.3961 (0.1928)	12.3881 (10.9871)	-	-	0.2262 (0.0672)	1.9292 (1.3863)	0.0388 (0.0213)	0.0104 (0.0076)
	(10, 50)	0.4301 (0.1627)	3.3806 (1.8726)	0.5893 (0.2213)	11.6472 (9.1627)			0.2621 (0.1168)	1.7616 (1.3607)	0.0185 (0.0069)	0.0064 (0.0053)
	(30, 10)	0.3907 (0.1927)	2.8727 (2.5102)	0.959 (0.6132)	13.3461 (11.7429)			0.3507 (0.2265)	1.7833 (1.563)	0.0023 (0.0008)	0.0042 (0.0033)
	(50, 10)	0.4464 (0.1762)	2.6139 (2.1472)	1.3108 (0.5106)	12.7951 (11.0624)			0.3811 (0.2361)	1.6407 (1.2391)	−0.006 (−0.0021)	0.0035 (0.0034)

Table 2. Cont.

1	(10, 10)	0.2671 (0.1065)	3.6920 (2.8976)	−0.0349 (−0.0094)	12.597 (11.2067)	1.28 (0.8131)	9.476 (8.201)	0.1808 (0.1035)	1.7161 (1.6015)	0.0304 (0.0089)	0.0123 (0.0089)
	(20, 20)	−0.2886 (−0.1102)	1.4603 (1.0553)	−0.5423 (−0.3798)	5.8502 (4.8455)	0.7005 (0.4102)	7.1436 (5.9314)	−0.2061 (−0.1244)	0.9535 (0.9013)	0.0058 (0.0039)	0.0041 (0.0038)
	(50, 50)	−0.0882 (−0.0273)	1.1903 (0.9677)	0.0359 (0.0247)	5.6179 (3.7369)	0.0997 (0.0616)	4.4333 (3.195)	−0.0286 (−0.0166)	0.6767 (0.4592)	−0.0048 (−0.0032)	0.0022 (0.002)
	(10, 30)	−0.7491 (−0.2943)	1.2002 (1.1763)	−1.7413 (−1.0025)	8.0579 (7.0341)	−0.0827 (−0.0701)	3.1955 (2.4681)	−0.7284 (−0.4239)	0.8882 (0.8102)	0.0305 (0.0116)	0.0051 (0.0047)
	(10, 50)	−0.2938 (−0.1514)	1.2297 (1.1833)	−1.0995 (−0.7793)	7.6932 (6.1106)	−0.2627 (−0.1524)	1.6751 (1.6017)	−0.3839 (−0.1262)	0.9041 (0.7581)	0.0402 (0.0173)	0.0048 (0.0038)
	(30, 10)	0.4111 (0.1761)	2.7002 (2.0596)	0.8199 (0.5911)	10.4623 (9.6197)	1.8155 (1.1078)	13.3161 (11.2647)	0.2865 (0.0678)	1.4344 (1.1214)	0.0232 (0.0106)	0.0068 (0.0066)
	(50, 10)	0.4313 (0.2144)	1.9358 (1.3712)	0.7317 (0.3627)	7.3343 (5.9132)	1.6602 (1.0562)	14.4777 (13.8713)	0.3404 (0.1261)	1.0343 (0.8276)	0.0037 (0.0021)	0.0044 (0.004)
	(10, 10)	0.5306 (0.1963)	4.0137 (3.5147)	0.4682 (0.2109)	10.383 (7.7981)	1.9795 (0.8052)	12.3963 (9.6824)	0.3455 (0.2209)	1.999 (1.7013)	0.0205 (0.0109)	0.0117 (0.0077)
	(20, 20)	0.0559 (0.0346)	1.4795 (1.2006)	0.0568 (0.0366)	7.017 (5.9943)	1.046 (1.0024)	6.1927 (5.811)	−0.0907 (−0.056)	0.711 (0.4273)	0.005 (0.0026)	0.0047 (0.0038)
	(50, 50)	0.2864 (0.1806)	1.5215 (1.324)	0.6017 (0.3913)	8.5793 (7.0642)	0.4162 (0.2648)	3.7811 (2.9167)	−0.0907 (−0.0389)	0.711 (0.6988)	0.0077 (0.0048)	0.0028 (0.0019)
	(10, 30)	−0.326 (−0.1744)	2.0059 (1.5783)	−0.8033 (−0.5627)	6.4739 (6.1634)	0.0577 (0.0204)	2.6519 (2.1062)	−0.3543 (−0.1634)	0.8374 (0.7018)	0.0183 (0.0103)	0.0057 (0.0052)
	(10, 50)	−0.3079 (−0.2109)	2.5945 (2.1617)	−1.2414 (−1.1435)	5.7281 (4.6001)	−0.3951 (−0.0849)	2.4483 (1.9736)	−0.445 (−0.2634)	1.5109 (0.9274)	0.0388 (0.0095)	0.0054 (0.0046)
	(30, 10)	0.3205 (0.1932)	2.9004 (2.4107)	0.7782 (−0.2353)	15.3845 (12.3712)	1.6217 (0.5743)	12.8923 (10.4129)	0.2998 (−0.1406)	1.3399 (1.0377)	0.0072 (0.0054)	0.0076 (0.0051)
	(50, 10)	0.7885 (0.4462)	1.9817 (1.6433)	1.6293 (0.8724)	14.3568 (11.762)	2.8004 (1.0079)	17.1023 (14.3726)	0.6473 (0.3709)	1.3393 (1.1611)	−0.0035 (−0.0025)	0.0065 (0.0047)

5. Application

In order to illustrate the usefulness of the proposed model, we consider data reported by King et al. [21]. A laboratory investigator interested in the relationship between diet and the development of tumors divided rats into two groups and fed them with saturated fat and unsaturated fat diets, respectively. The data represent the tumor-free time, the time from injection to the time that a tumor develops or to the end of the study. Group One contains $m = 30$ data (Unsaturated Fat) and Group Two $n = 25$ data (Saturated Fat). The data are as follows.

Unsaturated Fat

112, 68, 84, 109, 153, 143, 60, 70, 98, 164, 63, 63, 77, 91, 91, 66, 70, 77, 63, 66, 66, 94, 101, 105, 108, 112, 115, 126, 161, 178.

Saturated Fat

124, 58, 56, 68, 79, 89, 10786, 142, 110, 96, 142, 86, 75, 117, 98, 105, 126, 43, 46, 81, 133, 165, 170, 200, 200.

The required numerical evaluations were carried out using the Package of R software. We fitted length-biased weighted Lomax distributions with presence of outliers for $k = 0, 1, 2$ to this data. These three distributions were fitted into the subject data using maximum likelihood estimation. The MLEs, estimated standard errors (SE) of the parameters and the corresponding log-likelihood values ($-2\text{Log}(-L)$), Akaike information criteria (AIC; [22]), and P-values for Kolmogorov-Smirnov test (K-S P-value) are displayed in Table 3.

The model with a minimum of AIC value was chosen as the best model to fit the data. From Table 3 we conclude that the LBWLD with presence of $k = 1$ outliers is better than others.

Table 3. Maximum likelihood estimators (MLEs), standard errors (SE), Likelihood, Akaike information criterions (AIC) and K-S P-value values.

k	$\hat{\lambda}$	SE($\hat{\lambda}$)	$\hat{\theta}$	SE($\hat{\theta}$)	$\hat{\alpha}$	SE($\hat{\alpha}$)	$\hat{\beta}$	SE($\hat{\beta}$)	\hat{R}	−2Log−L	AIC	K-S p-Value
0	115.8826	42.9286	4.3728	1.1209	2.738	0.54891	-	-	0.7031	317.0236	323.0236	0.2896
1	113.89	47.3094	3.6107	1.104	4.1799	0.5221	1.3303	0.3213	0.4549	312.6703	320.6703	0.2879
2	92.4893	35.5477	3.7304	0.9435	4.1878	0.52552	1.4047	0.3861	0.4858	312.8561	320.8561	0.2838

6. Conclusions

This paper proposed a length-biased weighted Lomax distribution with presence of outliers. The MLES of the parameters have been investigated. The problem of estimating $R = P(Y < X)$ parameters for length-biased weighted Lomax distributions with Presence of Outliers have been addressed. The simulation experiment was carried out for $k = 0, 1, 2$; $\lambda = 2$; $\alpha = 7$; $\beta = 5$; and $\theta = 4$, and the results were given in Tables 1 and 2. The average bias and the MSE of the estimates based on 1000 replications and bootstrap method based on 500 resampling were reported. The results indicated that the MSEs of the estimations decrease when sample sizes increase and the average biases and MSEs of the estimations decrease when applying the bootstrap method. The usefulness of the distributions was illustrated in the analysis of real data. The results indicated that the LBWLD in presence of $k = 1$ outliers may be used for a wider range of statistical applications.

Author Contributions: Both authors performed experiments, contributed to the software and the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lomax, K.S. Business Failures: Another Example of the Analysis of Failure Data. *J. Am. Stat. Assoc.* **1954**, *49*, 847–852. [\[CrossRef\]](#)
2. Abd-Elfattah, A.M.; Alharbey, A.H. Estimation of Lomax Parameters Based on Generalized Probability Weighted Moment. *J. King Abdulaziz Univ. Sci.* **2010**, *22*, 171–184. [\[CrossRef\]](#)
3. Stene, J. Probability Distributions Arising from the Ascertainment and the Analysis of Data on Human Families and Other Groups. *Stat. Distrib. Sci. Work Appl. Phys. Soc. Life Sci.* **1981**, *6*, 233–244.
4. Oluyede, B.O.; George, E.O. On Stochastic Inequalities and Comparisons of Reliability Measures for Weighted Distributions. *Math. Probl. Eng.* **2002**, *8*, 1–13. [\[CrossRef\]](#)
5. Gupta, A.K.; Tripathi, R.C. Weighted Bivariate Logarithmic Series Distributions. *Commun. Stat. Theory Methods* **1996**, *25*, 1099–1117. [\[CrossRef\]](#)
6. Patill, G.P.; Rao, C.R.; Ratnaparkhi, M.V. On Discrete Weighted Distribution and their Use in Model Choice for Observed Data. *Commun. Stat. Theory Methods* **1986**, *15*, 907–918. [\[CrossRef\]](#)
7. Das, K.K.; Roy, T.D. Applicability of Length Biased Weighted Generalized Rayleigh Distribution. *Adv. Appl. Sci. Res.* **2011**, *2*, 320–327.
8. Das, K.K.; Roy, T.D. On Some Length-Biased Weighted Weibull Distribution. *Adv. Appl. Sci. Res.* **2011**, *2*, 465–475.
9. Rao, C.R. On Discrete Distributions Arising out of Methods of Ascertainment. *Indian J. Stat. Ser. A* **1965**, *27*, 311–324.
10. Halperin, M.; Gilbert, P.R.; Lachin, J.M. Distribution Free Confidence Intervals for $\Pr(X_1 < X_2)$. *Biometrics* **1987**, *43*, 71–80.
11. Simonoff, J.S.; Hochberg, Y.; Reiser, B. Alternative Estimation Procedures for $\Pr(X < Y)$ in Categorized Data. *Biometrics* **1986**, *42*, 895–907. [\[PubMed\]](#)
12. Reiser, B.; Farragi, D. Confidence Bounds for $\Pr(a'x > b'y)$. *Statistics* **1994**, *25*, 107–111. [\[CrossRef\]](#)
13. Bamber, D. The Area above the Ordinal Dominance Graph and the Area below the Receiver Operating Characteristic Graphs. *J. Math. Psychol.* **1975**, *12*, 387–415. [\[CrossRef\]](#)
14. Weerahandi, S.; Johnson, R.A. Testing Reliability in a Stress-Strength Model When X and Y are Normally Distributed. *Technometrics* **1992**, *34*, 83–91. [\[CrossRef\]](#)
15. Nadarajah, S. Reliability of Laplace Distributions. *Math. Problems Eng.* **2004**, *2*, 169–183. [\[CrossRef\]](#)

16. Dixit, U.J.; Nasiri, P.F. Estimation of Parameters of the Exponential Distribution in the Presence of Outliers Generated from Uniform Distribution. Available online: ftp://luna.sta.uniroma1.it/RePEc/articoli/2001-LIX-3_4-13.pdf (accessed on 15 February 2018).
17. Dixit, U.J. Characterization of the Gamma Distribution in the Presence of k Outliers. *Bull Bombay Math Colloq.* **1987**, *4*, 54–59.
18. Ahmad, A.; Ahmad, S.P.; Ahmed, A. Length-Biased Weighted Lomax Distribution: Statistical Properties and Application. *Pak. J. Stat. Oper.* **2016**, *12*, 1178. [[CrossRef](#)]
19. Efron, B. *The Jackknife, the Bootstrap and other Resampling Plans*. CBMS-NSF Regional Conference Series in Applied Mathematics; SIAM: Philadelphia, PA, USA, 1982.
20. Wang, M.; Wang, W. Bias-Corrected Maximum Likelihood Estimation of the Parameters of the Weighted Lindley Distribution. *Commun. Stat. Simul. Comput.* **2017**, *46*, 530–545. [[CrossRef](#)]
21. King, M.; Bailey, D.M.; Gibson, D.G.; Pitha, J.V.; McCay, P.B. Incidence and Growth of Mammary Tumors Induced by 7, 12-Dimethylbenz[a]anthracene as Related to the Dietary Content of Fat and Antioxidant. *J. Nat. Cancer Inst.* **1979**, *63*, 656–664. [[CrossRef](#)]
22. Akaike, H. On entropy maximization principle. In *Applications of Statistics*; Krishnaiah, P.R., Ed.; Elsevier Science Ltd.: North-Holland, Amsterdam, 1977; pp. 27–41.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).