METHODS OF ESTIMATING THE PARAMETERS OF THE QUASI LINDLEY DISTRIBUTION

Festus C. Opone [1]
Department of Mathematics, University of Benin, Benin City, Nigeria

Nosakhare Ekhosuehi
Department of Mathematics, University of Benin, Benin City, Nigeria

1. INTRODUCTION

The Lindley distribution introduced by Lindley (1958) have received considerable attention in developing a generalized form of the distribution. Ghitany *et al.* (2013) proposed the power Lindley distribution as an extension of the classical one-parameter Lindley distribution by considering the power transformation $X = T^{\frac{1}{\alpha}}$. Shanker *et al.* (2013) introduced a two-parameter Lindley distribution which they call the Sushila distribution. Shibu and Irshad (2016) proposed the extended new generalized Lindley distribution, a variant of this distribution is the new extended generalized Lindley distribution introduced in Maya and Irshad (2017).

Shanker and Mishra (2013) proposed the quasi Lindley distribution (QLD) with parameters (α, θ) and its probability density function is defined by

$$f(x, \alpha, \theta) = \frac{\theta(\alpha + x\theta)e^{-\theta x}}{\alpha + 1}, \qquad x > 0, \theta > 0, \alpha > -1. \tag{1}$$

The QLD which is an extension of the one-parameter Lindley distribution is a two-component mixture of exponential (θ) and a special case of gamma $(2, \theta)$ distribution. Equation (1) can also be written in the form

$$f(x, \alpha, \theta) = pf_1(x) + (1-p)f_2(x), \tag{2}$$

where $f_1(x) = \theta e^{-\theta x}$ and $f_2(x) = \theta^2 e^{-\theta x}$ are the density functions of the exponential distribution (θ) and gamma distribution $(2, \theta)$ respectively and $p = \frac{\alpha}{\alpha+1}$ is the mixing proportion.

---

[1] Corresponding Author. E-mail: festus.opone@physci.uniben.edu

The corresponding cumulative distribution function of the QLD is given by

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

Amongst the mathematical properties of the distribution studied were; the shape of density function, cumulative distribution function, failure rate function, mean residual life function, stochastic ordering and its moments with related measures. The method of moments and maximum likelihood was used in estimation of the parameters of the distribution and was applied to a real lifetime data set. It was shown that the quasi Lindley distribution can be used as an alternative model to the one-parameter Lindley distribution and the popular exponential distribution, as it provides better fit in applicability to lifetime data. For more detail on the mathematical properties of the QLD, we refer readers to Shanker and Mishra (2013).

In spite of the mathematical properties studied, Shanker and Mishra (2013) did not address the quantile function of the distribution which can be used to generate random samples from the distribution. Although, the method of moment and maximum likelihood were used in estimation of the parameters of the distribution, their work fails to examine the performance and accuracy of the parameter estimates of the distribution. These pitfalls form the basis of our study.

Motivation of this paper arose from the work of Roozegar and Nadarajah (2017), which presents meaningful criticism and useful suggestions on the generalized Lindley distribution proposed by Nedjar and Zeghdoudi (2016). The remaining sections of this paper are organized as follows: Section 2 presents the quantile function of the quasi Lindley distribution. In Section 3, the method of moment and the maximum likelihood for parameter estimation are considered and a simulation study was conducted to examine the behaviour of the estimators of each parameter. Finally, in Section 4, we examined the applicability of the QLD alongside with other related existing distributions in modeling lifetime data sets.

2.   QUANTILE FUNCTION OF THE QUASI LINDLEY DISTRIBUTION

Jodra (2010) derived a closed-form expression for the quantile function of probability distributions which are related to the Lambert W function. This expression enables us to generate random samples from the distribution through the use of inverse transform.

THEOREM 1. *For any $\theta > 0, \alpha > -1$, the quantile function of the quasi Lindley distribution X is defined by*

$$x = -\frac{1}{\theta} - \frac{\alpha}{\theta} - \frac{1}{\theta} W_{-1}\left(-\frac{(1-u)(\alpha+1)}{e^{(1+\alpha)}}\right), \tag{4}$$

where W_{-1} denotes the negative branch of the Lambert W function.

PROOF. Let $F(x)$ be the distribution function of the QLD defined in Equation (3). For any $\theta > 0, \alpha > -1$ and $0 < u < 1$, we need to set $F(x) = u$ to obtain a system of non-linear equation which is given by

$$-(1+\alpha+\theta x)e^{-\theta x} = (u-1)(\alpha+1). \tag{5}$$

Multiplying both sides of Equation (5) by $e^{-(1+\alpha)}$, we obtain

$$-(1+\alpha+\theta x)e^{-(\theta x+\alpha+1)} = -(1-u)(\alpha+1)e^{-(1+\alpha)}. \tag{6}$$

From Equation (6), we see that $-(1+\alpha+\theta x)$ is the Lambert W function of the real argument $-(1-u)(\alpha+1)e^{-(1+\alpha)}$. Thus, we have

$$W\left(-\frac{(1-u)(\alpha+1)}{e^{(1+\alpha)}}\right) = -(1+\alpha+\theta x). \tag{7}$$

Moreover, for any $\theta > 0, \alpha > -1$ and $x > 0$, it is immediate that $(1+\alpha+\theta x) > 1$ and it can also be checked that $(1-u)(\alpha+1)e^{-(1+\alpha)}\epsilon(-\frac{1}{e},0)$ since $u\epsilon(0,1)$. Therefore, by taking into account the properties of the negative branch of the Lambert W function, Equation (7) becomes

$$W_{-1}\left(-\frac{(1-u)(\alpha+1)}{e^{(1+\alpha)}}\right) = -(1+\alpha+\theta x)$$

$$x = -\frac{1}{\theta}-\frac{\alpha}{\theta}-\frac{1}{\theta}W_{-1}\left(-\frac{(1-u)(\alpha+1)}{e^{(1+\alpha)}}\right). \tag{8}$$

This completes the proof. □

TABLE 1
Some quantiles of the QLD for selected values of the parameters.

u	$(\theta=0.3, \alpha=0.1)$	$(\theta=0.1, \alpha=0.5)$	$(\theta=0.2, \alpha=2)$	$(\theta=2, \alpha=1)$
0.01	0.2722	0.2958	0.0753	0.0272
0.02	0.4667	0.5839	0.1511	0.0505
0.03	0.6302	0.8657	0.2276	0.0715
0.04	0.7759	1.1419	0.3046	0.0911
0.05	0.9098	1.4134	0.3823	0.1094
0.06	1.0351	1.6809	0.4606	0.1268
0.07	1.1540	1.9448	0.5395	0.1435
0.08	1.2678	2.2058	0.6191	0.1597
0.09	1.3776	2.4641	0.6994	0.1753
0.10	1.4840	2.7201	0.7804	0.1906

Oluyede *et al.* (2016), also suggested an explicit form of expressing the quantile function of probability distributions using numerical method. As an alternative to the inverse transform method, random samples from probability distribution are generated by solving the system of non-linear equation defined in Equation (5) using numerical method. The quantiles in Table 1 was generated by solving the system of non-linear equation given in Equation (5).

3. PARAMETER ESTIMATION

3.1. Method of maximum likelihood (MLM)

Let $x_1, x_2, \cdots, x_n$ be a random sample of size n from QLD(α, θ), then the log-likelihood estimate is defined by

$$\ell(x, \alpha, \theta) = \sum_{i=1}^{n} \log\left[\frac{\theta(\alpha + x\theta)e^{-\theta x}}{\alpha + 1}\right] \tag{9}$$

$$= n\log\theta - n\log(\alpha + 1) + \sum_{i=1}^{n} \log(\alpha + x\theta) - n\theta\bar{X}. \tag{10}$$

The associated score function is defined by

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \frac{x}{(\alpha + x\theta)} - n\bar{X} = 0, \tag{11}$$

$$\frac{\partial \ell}{\partial \alpha} = -\frac{n}{\alpha + 1} + \sum_{i=1}^{n} \frac{1}{(\alpha + x\theta)} = 0. \tag{12}$$

The maximum likelihood estimators $\hat{\theta}$ and $\hat{\alpha}$ can be achieved using the Newton Raphson iterative method.

3.2. Method of moment (MOM)

The moment estimation is a technique for constructing an estimator of a parameter that is based on matching the sample moments with the corresponding distribution (theoretical) moments.

Let $x_1, x_2, \cdots, x_n$ represent a random sample of size n drawn from a probability distribution for which we seek an unbiased estimator for the rth moment. An expression for the sample moment is given by

$$m'_r = \frac{1}{n} \sum_{r=1}^{n} x^r. \tag{13}$$

Shanker and Mishra (2013) defined the rth moment of the quasi Lindley distribution as

$$\mu_r' = \frac{\Gamma(r+1)(\alpha+r+1)}{\theta^r(\alpha+1)}, \quad r = 1, 2, 3, \ldots \tag{14}$$

Taking $r = 1, 2, 3$ and 4 in Equation (14), the first four raw moments of the QLD are obtained as

$$\mu_1' = \frac{1}{\theta}\left(\frac{\alpha+2}{\alpha+1}\right), \quad \mu_2' = \frac{2}{\theta^2}\left(\frac{\alpha+3}{\alpha+1}\right), \quad \mu_3' = \frac{6}{\theta^3}\left(\frac{\alpha+4}{\alpha+1}\right), \quad \mu_4' = \frac{24}{\theta^4}\left(\frac{\alpha+5}{\alpha+1}\right).$$

Equating Equations (13) and (14), when $r = 1$, an estimate for the parameter θ is obtained as

$$\hat{\theta} = \frac{1}{\bar{X}}\left(\frac{\alpha+2}{\alpha+1}\right). \tag{15}$$

Similarly, to obtain an estimate for the parameter α, we divide the second moment by the square of the first moment to get an expression which is a function of α only.

$$\frac{\mu_2'}{\mu_1^2} = \frac{2(\alpha+3)(\alpha+1)}{(\alpha+2)^2} = k. \tag{16}$$

Equation (16) results to a system of quadratic equation given by

$$(2-k)\alpha^2 + 4(2-k)\alpha + (6-4k) = 0. \tag{17}$$

Solving the system of equations in (17), an expression for the estimate of α is obtained as

$$\hat{\alpha} = \frac{-(4-2k) + \sqrt{4-2k}}{2-k}, \tag{18}$$

where k is obtained by replacing the first moment (μ_1') and the second moment (μ_2') by m_1' and m_2' respectively.

3.3. Interval estimates

In this subsection, the asymptotic confidence intervals for the parameters of the QLD are presented. Under the normality condition, $\hat{\eta} \sim N[\eta, I^{-1}(\eta)]$, i.e. the asymptotic distribution of an estimator $\hat{\eta}$ of a parameter η is approximately normal with mean η and variance obtained by inverting the Fisher information matrix. The Fisher information matrix is given by

$$I(\eta_k) = -E\left(\frac{\partial^2 \ell}{\partial \eta^2}\right) = -E\begin{pmatrix} \frac{\partial^2 \ell}{(\partial \theta)^2} & \frac{\partial^2 \ell}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \ell}{\partial \alpha \partial \theta} & \frac{\partial^2 \ell}{(\partial \alpha)^2} \end{pmatrix}, \quad \eta = (\theta, \alpha)^T,$$

where

$$\frac{\partial^2 \ell}{(\partial \theta)^2} = -\sum_{i=1}^{n} \frac{x^2}{(\alpha + x\theta)^2} - \frac{n}{\theta^2},$$

$$\frac{\partial^2 \ell}{\partial \theta \partial \alpha} = \frac{\partial^2 \ell}{\partial \alpha \partial \theta} = \sum_{i=1}^{n} \frac{x}{(\alpha + x\theta)^2},$$

$$\frac{\partial^2 \ell}{(\partial \alpha)^2} = \frac{n}{(\alpha+1)^2} - \sum_{i=1}^{n} \frac{1}{(\alpha + x\theta)^2}.$$

The approximate $(1 - \delta)100$ CIs for the parameters θ and α are respectively $\hat{\alpha} \pm Z_{\frac{\delta}{2}} \sqrt{\text{var}(\hat{\alpha})}$ and $\hat{\theta} \pm Z_{\frac{\delta}{2}} \sqrt{\text{var}(\hat{\theta})}$, where $\text{var}(\hat{\alpha})$ and $\text{var}(\hat{\theta})$ are the variance of α and θ which are given by the first and second diagonal element of the variance-covariance matrix $I^{-1}(\eta_k)$ and $Z_{\frac{\delta}{2}}$ is the upper$(\frac{\delta}{2})$ percentile of the standard normal distribution.

3.4. Simulation study

In this subsection, we consider two methods of parameter estimation (MLM and MOM) to investigate the performance and accuracy of the parameter estimates of the QLD. The flexibility of these methods are compared through a simulation study for different parameter values as well as different sample sizes. We generated random data from the QLD using Equation (5). The Monte Carlo simulation study is repeated 1000 times for different sample sizes $n = 60, 90, 120, 150$ and parameter values ($\theta = 0.35$, $\alpha = 0.15$), ($\theta = 1$, $\alpha = 0.2$) and ($\theta = 2$, $\alpha = 0.3$).

An algorithm for the simulation study is given by the following steps.

1. Choose the value N (i.e. number of Monte Carlo simulation);

2. choose the values $\eta_0 = (\theta_0, \alpha_0)$ corresponding to the parameters of the QLD (θ, α);

3. generate a sample of size n from QLD;

4. compute the maximum likelihood estimates $\hat{\eta}_0$ of η_0 and the moment estimate defined in Equations (15) and (18);

5. repeat steps (3-4), N-times;

6. compute the bias $= \frac{1}{N} \sum_{i=1}^{N} (\hat{\eta}_i - \eta_0)$ and the mean square error (MSE) $= \frac{1}{N} \sum_{i=1}^{N} (\hat{\eta}_i - \eta_0)^2$.

Tables 2, 3 and 4 present two methods of estimating the parameters of quasi Lindley distribution. The performance of these methods is compared based on the bias and mean square error criteria at different choice of sample size n.

TABLE 2
Monte Carlo simulation results for $\alpha = 0.2$, $\theta = 1$.

n	Methods	Bias(α)	Bias(θ)	MSE(α)	MSE(θ)
60	MLM	0.0338	0.0266	0.1010	0.0225
	MOM	0.0527	0.0397	0.3602	0.0286
90	MLM	0.0296	0.0168	0.0481	0.0119
	MOM	0.0368	0.0284	0.1206	0.0170
120	MLM	0.0256	0.0033	0.0334	0.0090
	MOM	0.0271	0.0185	0.0890	0.0134
150	MLM	0.0153	0.0011	0.0291	0.0069
	MOM	0.0188	0.0108	0.0574	0.0106

TABLE 3
Monte Carlo simulation results for $\alpha = 0.3$, $\theta = 2$.

n	Methods	Bias(α)	Bias(θ)	MSE(α)	MSE(θ)
60	MLM	0.0605	0.0447	0.2462	0.1003
	MOM	0.0661	0.0790	0.3550	0.1151
90	MLM	0.0316	0.0258	0.0809	0.0526
	MOM	0.0505	0.0450	0.1517	0.0744
120	MLM	0.0266	0.0181	0.0645	0.0409
	MOM	0.0454	0.0357	0.1422	0.0646
150	MLM	0.0243	0.0058	0.0479	0.0290
	MOM	0.0317	0.0218	0.1043	0.0487

TABLE 4
Monte Carlo simulation results for $\alpha = 0.15$, $\theta = 0.35$.

n	Methods	Bias(α)	Bias(θ)	MSE(α)	MSE(θ)
60	MLM	0.0467	0.0137	0.1628	0.0027
	MOM	0.0625	0.0145	0.5057	0.0031
90	MLM	0.0299	0.0042	0.0552	0.0015
	MOM	0.0506	0.0069	0.1114	0.0022
120	MLM	0.0263	0.0025	0.0320	0.0011
	MOM	0.0461	0.0043	0.0975	0.0016
150	MLM	0.0066	0.0004	0.0231	0.0008
	MOM	0.0067	0.0030	0.0501	0.0012

Evidently, as the sample size n increases, the values of the bias and the mean square error of the parameter estimates for both methods decreases and hence the estimation precision of the parameters increases. Although, the method of moment has a closed form expression for the parameter estimates, the method of maximum likelihood performs better for all sample size considered, therefore we recommend the method of maximum likelihood in estimating the parameters of the quasi Lindley distribution.

4. DATA ANALYSIS

In this section, we applied the quasi Lindley distribution to a real lifetime data and compare its fit with the ones attained by power Lindley distribution (PLD), Sushila distribution and Lindley distribution. Table 5 shows the waiting time (in minutes) of 100 Bank customers reported in Ghitany *et al.* (2008).

TABLE 5
Waiting time data.

0.8	0.8	1.3	1.5	1.8	1.9	1.9	2.1	2.6	2.7	2.9	3.1	3.2
3.3	3.5	3.6	4.0	4.1	4.2	4.2	4.3	4.3	4.4	4.4	4.6	4.7
4.7	4.8	4.9	4.9	5.0	5.3	5.5	5.7	6.1	6.2	6.2	6.2	6.2
6.3	6.7	6.9	7.1	7.1	7.1	7.1	7.4	7.6	7.7	8.0	8.2	8.6
8.6	8.6	8.8	8.8	8.9	8.9	9.5	9.6	9.7	9.8	10.7	10.9	11.0
11.0	11.1	11.2	11.2	11.5	11.9	12.4	12.5	12.9	13.0	13.1	13.3	13.6
13.7	13.9	14.1	15.4	15.4	17.3	17.3	18.1	18.2	18.4	18.9	19.0	19.9
20.6	21.3	21.4	21.9	23.0	27.0	31.6	33.1	38.5	-	-	-	-

The variance-covariance matrix of the data is given by

$$I^{-1}(\eta_k) \;=\; \begin{pmatrix} 0.000299 & -0.000715 \\ -0.000715 & 0.0053140 \end{pmatrix}$$

and the 95% confidence intervals for the model parameters are estimated as $\theta \epsilon (0.177291 \quad 0.245109)$ and $\alpha \epsilon (-0.222279 \quad 0.063478)$.

The parameter estimates, the Log-lik, Akaike information criterion (AIC), Kolmogorov-Smirnov $(K - S)$ and the Cramér von Mises (W^*) statistic with their corresponding p-value of the distributions for the waiting time data are shown in Table 6.

TABLE 6
Summary statistic of the results of waiting time data.

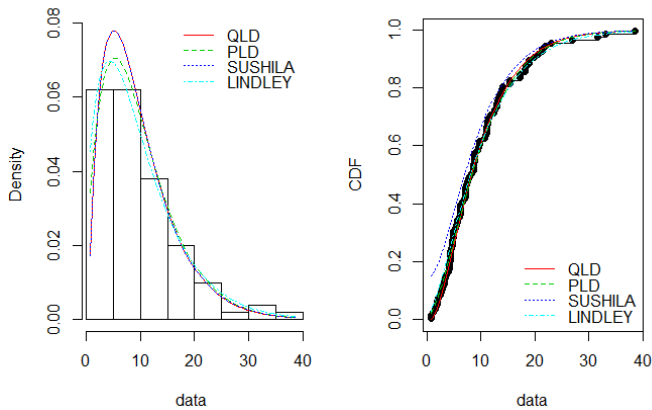| Distributions	Estimates	Log-lik	AIC	$K - S$ (p-value)	W^* (p-value)
QLD	$\alpha = -0.0791$ $\theta = 0.2112$	-316.9255	637.8511	0.0567 (0.9052)	0.0449 (0.9076)
PLD	$\alpha = 1.0830$ $\theta = 0.1531$	-318.3186	640.6372	0.0519 (0.9504)	0.0456 (0.9034)
SUSHILA	$\alpha = -0.3757$ $\theta = -0.0794$	-316.9255	637.8511	0.1528 (0.0187)	0.0614 (0.022)
LINDLEY	$\theta = 0.1866$	-319.0374	640.0748	0.0678 (0.7494)	0.0582 (0.8265)

Figure 1 – Density and cumulative distribution fit for the waiting time data.
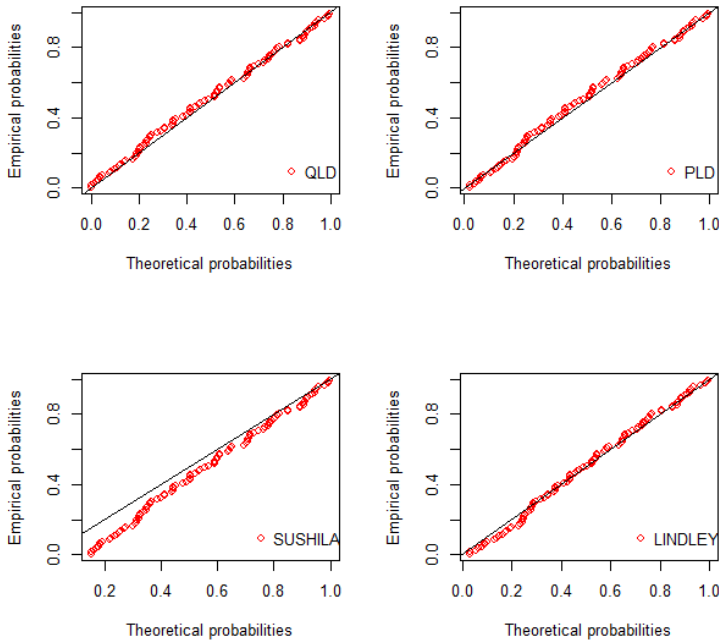
Figure 2 – Probability-Probability plot for the waiting time data.

The graphical illustration of the density and cumulative distribution fit and P-P plots of the distributions for the waiting time data is shown in Figures 1 and 2 respectively.

REFERENCES

M. GHITANY, B. ATIEH, S. NADADRAJAH (2008). *Lindley distribution and its applications*. Mathematics and Computers in Simulation, 78, pp. 493–506.

M. GHITANY, D. AL-MUTAIRI, N. BALAKRISHNAN, I. AL-ENEZI (2013). *Power Lindley distribution and associated inference*. Computational Statistics and Data Analysis, 64, pp. 20–33.

P. JODRA (2010). *Computer generation of random variables with Lindley or Poisson–Lindley distribution via the Lambert W function*. Mathematical Computations and Simulation, 81, pp. 851–859.

D. V. LINDLEY (1958). *Fiducial distributions and Bayes theorem*. Journal of the Royal Statistical Society, Series B, 20, no. 1, pp. 102–107.

R. MAYA, M. R. IRSHAD (2017). *New extended generalized Lindley distribution: Properties and applications*. Statistica, 77, pp. 33–52.

S. NEDJAR, H. ZEGHDOUDI (2016). *On gamma Lindley distribution: Properties and applications*. Journal of Computational and Applied Mathematics, 298, pp. 167–174.

B.O. OLUYEDE, S. FOYA, G. WARAHENA-LIYANAGE, S. HUANG (2016). *The log-logistic Weibull distribution with applications to lifetime data*. Austrian Journal of Statistics, 45, pp. 43–69.

R. ROOZEGAR, S. NADARAJAH (2017). *On a generalized Lindley distribution*. Statistica, 77, no. 2, pp. 149–157.

R. SHANKER, A. MISHRA (2013). *A quasi-Lindley distribution*. African Journal of Mathematics and Computer Science Research, 6, no. 4, pp. 64–71.

R. SHANKER, S. SHARMA, U. SHANKER, R. SHANKER (2013). *Sushila distribution and its application to waiting times data*. International Journal of Business Management, 3, no. 2, pp. 1–11.

D. S. SHIBU, M. R. IRSHAD (2016). *Extended new generalized Lindley distribution*. Statistica, 76, pp. 41–56.

SUMMARY

In this paper, we review the quasi Lindley distribution and established its quantile function. A simulation study is conducted to examine the bias and mean square error of the parameter estimates of the distribution through the method of moment estimation and the maximum likelihood estimation. Result obtained shows that the method of maximum likelihood is a better choice of estimation method for the parameters of the quasi Lindley distribution. Finally, an applicability of the quasi Lindley disttribution to a waiting time data set suggests that the distribution demonstrates superiority over the power Lindley distribution, Sushila distribution and the classical one-parameter Lindley distribution in terms of the maximized loglikelihood, the Akaike information criterion, the Kolmogorov-Smirnov and Cramér von Mises test statistic.

Keywords: Quasi Lindley distribution; Quantile function; Moment estimation; Maximum likelihood estimation.