academicJournals

Vol. 6(4), pp. 64-71, April 2013
DOI 10.5897/AJMCSR 12.067
ISSN 2006-9731 ©2013 Academic Journals
http://www.academicjournals.org/AJMCSR

African Journal of Mathematics and Computer Science Research

Full Length Research Paper

A quasi Lindley distribution

Rama Shanker[1]* and A. Mishra[2]

[1]College of Business and Economics, P. O. Box 3158, Asmara, Eritrea, N-E Africa.
[2]Department of Statistics, Patna University, Patna, India.

Accepted 4 March, 2013

A two-parameter Quasi Lindley distribution (QLD), of which the Lindley distribution (LD) is a particular case, has been introduced. Its moments, failure rate function, mean residual life function and stochastic orderings have been discussed. It is found that the expressions for failure rate function, mean residual life function, and stochastic orderings of the QLD shows its flexibility over Lindley distribution and exponential distribution. Although, the QLD has two parameters, the expressions for coefficients of variation, skewness, and kurtosis depend upon only one parameter. The maximum likelihood method and the method of moments have been discussed for estimating its parameters. The distribution has been fitted to some data-sets to test its goodness of fit to which earlier the Lindley distribution has been fitted by others and it is found that to almost all these data-sets the QLD provides closer fits than those by the Lindley distribution.

Key words: Lindley distribution, moments, failure rate function, mean residual life function, stochastic ordering, estimation of parameters, goodness of fit.

INTRODUCTION

Lindley (1958), introduced a one-parameter distribution, known as Lindley distribution, given by its probability density function

$$f\left(x;\theta\right)=\frac{\theta^2}{\theta+1}\left(1+x\right)e^{-\theta x}\ ;\ x>0,\theta>0 \tag{1}$$

It can be seen that this distribution is a mixture of exponential (θ) and gamma $(2,\theta)$ distributions. Its cumulative distribution function has been obtained as

$$F\left(x\right)=1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x};\quad x>0,\theta>0 \tag{2}$$

Ghitany et al. (2008a) have discussed various properties of this distribution and showed that in many ways Equation (1) provides a better model for some applications than the exponential distribution. The first four moments about origin of the Lindley distribution have been obtained by Ghitany et al. (2008a) as:

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

and its central moments have been obtained as

$$\mu_2=\frac{\theta^2+4\theta+2}{\theta^2\left(\theta+1\right)^2},\quad \mu_3=\frac{2\left(\theta^3+6\theta^2+6\theta+2\right)}{\theta^3\left(\theta+1\right)^3},\quad \mu_4=\frac{3\left(3\theta^4+24\theta^3+44\theta^2+32\theta+8\right)}{\theta^4\left(\theta+1\right)^4} \tag{4}$$

Ghitany et al. (2008a) studied the various properties of this distribution. A discrete version of this distribution has been suggested by Deniz and Ojeda (2011) having its applications in count data related to insurance. Sankaran (1970) obtained the Lindley mixture of Poisson distribution. Ghitany et al. (2008b, c) obtained size-biased and zero-truncated version of Poisson- Lindley distribution and discussed their various properties and applications. Ghitany and Al-mutairi (2009) discussed

*Corresponding author. E-mail: shankerrama2009@gmail.com.

various estimation methods for the discrete Poisson-Lindley distribution. Bakouch et al. (2012) obtained an extended Lindley distribution and discussed its various properties and applications. Mazucheli and Achcar (2011) discussed the applications of Lindley distribution to competing risks lifetime data. Ghitany et al. (2011) developed a two-parameter weighted Lindley distribution and discussed its applications to survival data. Zakerzadah and Dolati (2010) obtained a generalized Lindley distribution and discussed its various properties and applications.

In this study, a two parameter quasi Lindley distribution (QLD), of which the Lindley distribution Equation (1) is a particular case, has been suggested. Its first four moments and some of the related measures have been obtained. Its failure rate, mean residual rate and stochastic ordering have also been studied. The nature of the QLD, its distribution function and its hazard rate function has been shown graphically by drawing different graphs for different values of its parameters. Estimation of its parameters has been discussed and the distribution has been fitted to some of those data sets where the Lindley distribution has earlier been fitted by others, to test its goodness of fit.

QUASI LINDLEY DISTRIBUTION

Quasi Lindley distribution with parameters α and θ is defined by its probability density function (p.d.f)

$$f(x;\alpha,\theta)=\frac{\theta(\alpha+x\theta)}{\alpha+1}e^{-\theta x}\ ;\quad x>0,\theta>0,\alpha>-1 \tag{5}$$

It can easily be seen that at $\alpha=\theta$, the QLD Equation (5) reduces to the Lindley distribution Equation (1) and at $\alpha=0$, it reduces to the gamma distribution with parameters $(2,\theta)$. The p.d.f. Equation (5) can be shown as a mixture of exponential (θ) and gamma $(2,\theta)$ distributions as follows:

$$f(x;\alpha,\theta)=pf_1(x)+(1-p)f_2(x) \tag{6}$$

Where

$p=\frac{\alpha}{\alpha+1}$, $f_1(x)=\theta e^{-\theta x}$ and $f_2(x)=\theta^2 x e^{-\theta x}$.

The nature of QLD for different values of its parameters θ and α has been shown graphically in the Figure 1(a), (b), (c) and (d). In the figure QL (1,1) means QLD with parameters $\theta=1,\ \text{and}\ \alpha=1$. In Figure 1(d), five graphs of QLD for different values of its parameters have been combined for ready comparisons.

The first derivative of Equation (5) is $f'(x)=\frac{\theta^2 e^{-\theta x}}{\alpha+1}(1-\alpha-x\theta)$ and $f'(x)=0$ gives $x=\frac{1-\alpha}{\theta}$.

From this it follows that:

(i) for $|\alpha|<1$, $x_0=\frac{1-\alpha}{\theta}$ is the unique critical point at which $f(x)$ is maximum.

(ii) for $\alpha\geq 1$, $f'(x)\leq 0$, that is, $f(x)$ is decreasing in x.

Therefore, the mode of the QLD is given by

$$\text{Mode}=\begin{cases}\frac{1-\alpha}{\theta}, & |\alpha|<1\\ 0, & \end{cases} \tag{7}$$

The cumulative distribution function (c.d.f.) of the QLD is obtained as

$$F(x)=1-\frac{1+\alpha+\theta x}{\alpha+1}e^{-\theta x}\ ;\quad x>0,\theta>0,\alpha>-1 \tag{8}$$

The graphs of distribution function of QLD for different values of its parameters θ and α are shown in Figure 2. In Figure 2 CQL(1,1) means cumulative distribution function with parameters $\theta=1\text{and}\ \alpha=1$.

MOMENTS AND SOME RELATED MEASURES

The rth moment about origin of the QLD has been obtained as

$$\mu_r'=\frac{\Gamma(r+1)(\alpha+r+1)}{\theta^r(\alpha+1)};\ r=1,2,. \tag{9}$$

Taking $r=1,2,3$ and 4 in Equation (9), the first four moments about origin of the QLD are obtained as

$$\mu_1'=\frac{1}{\theta}\left(\frac{\alpha+2}{\alpha+1}\right),\ \mu_2'=\frac{2}{\theta^2}\left(\frac{\alpha+3}{\alpha+1}\right),\ \mu_3'=\frac{6}{\theta^3}\left(\frac{\alpha+4}{\alpha+1}\right),\ \mu_4'=\frac{24}{\theta^4}\left(\frac{\alpha+5}{\alpha+1}\right) \tag{10}$$

It can easily be verified that for $\alpha=\theta$, the moments about origin of the QLD reduce to the respective moments of the Lindley distribution. The central moments of the QLD have thus been obtained as

$$\mu_2=\frac{\alpha^2+4\alpha+2}{\theta^2(\alpha+1)^2},\ \mu_3=\frac{2(\alpha^3+6\alpha^2+6\alpha+2)}{\theta^3(\alpha+1)^3},\ \mu_4=\frac{3(3\alpha^4+24\alpha^3+44\alpha^2+32\alpha+8)}{\theta^4(\alpha+1)^4} \tag{11}$$

The coefficients of variation (γ), skewness $(\sqrt{\beta_1})$ and the

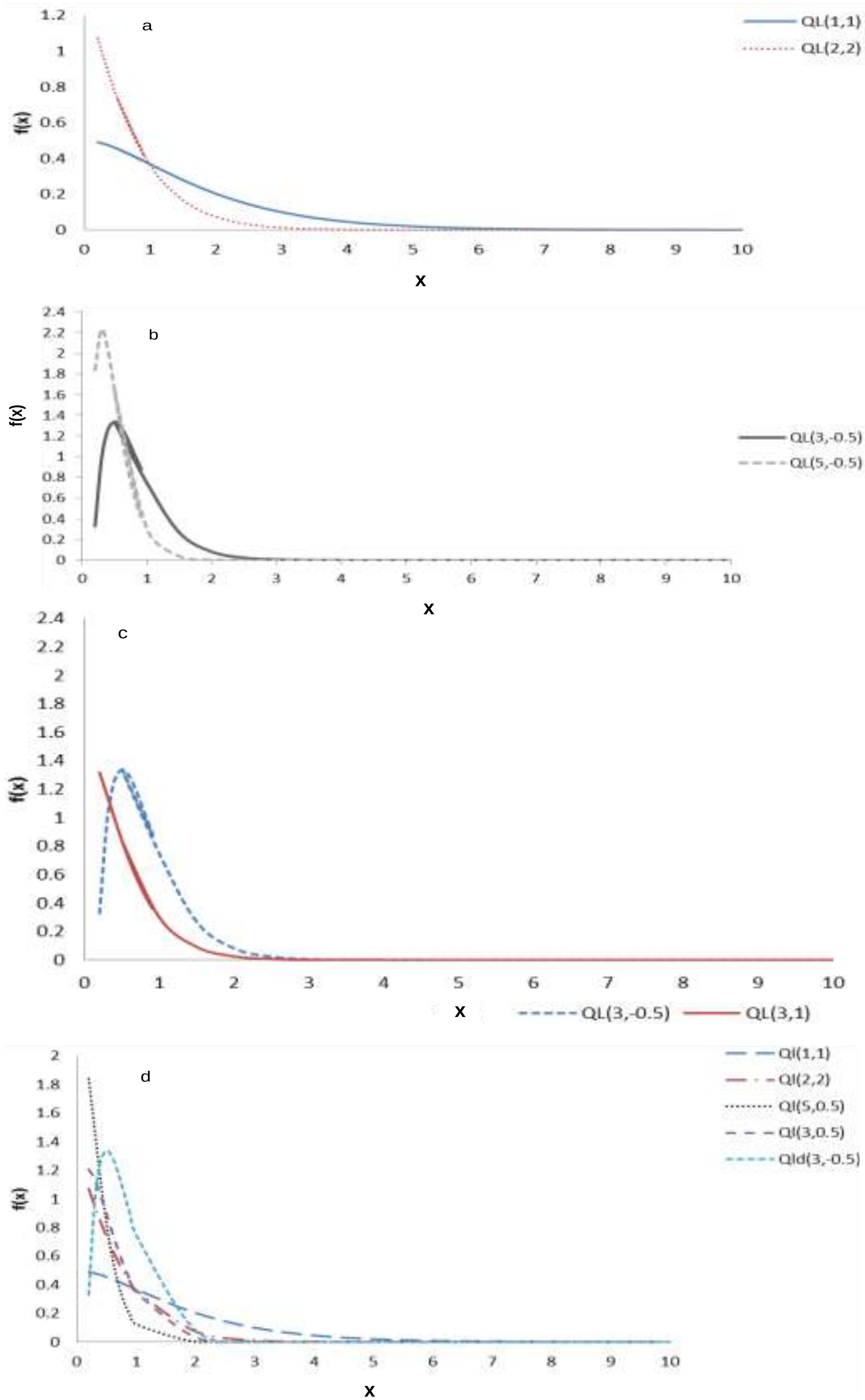

Figure 1. Plots of the density function (5) for some parameter values θ and α.

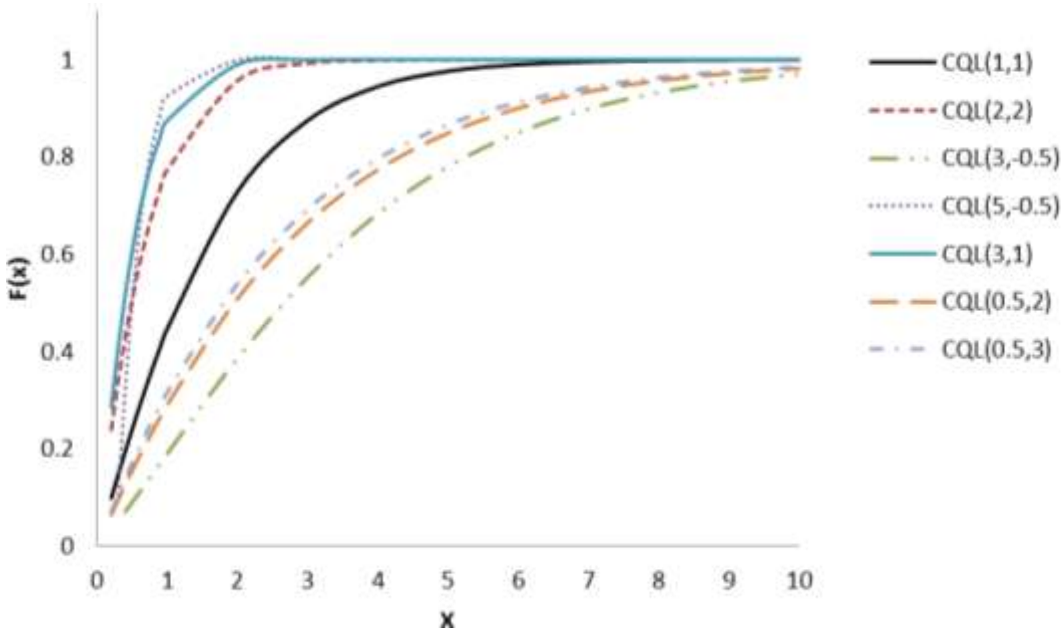

Figure 2. Plots of the distribution function (8) for some parameter values θ and α.

kurtosis (β_2) of the QLD have been obtained as

$$\gamma = \frac{\sigma}{\mu_1'} = \frac{\sqrt{\alpha^2+4\alpha+2}}{\alpha+2} \qquad (12)$$

$$\sqrt{\beta_1} = \frac{2\left(\alpha^3+6\alpha^2+6\alpha+2\right)}{\left(\alpha^2+4\alpha+2\right)^{3/2}} \qquad (13)$$

$$\beta_2 = \frac{3\left(3\alpha^4+24\alpha^3+44\alpha^2+32\alpha+8\right)}{\left(\alpha^2+4\alpha+2\right)^2} \qquad (14)$$

It is interesting to note that all these expressions are independent of the parameter θ and depend upon the parameter α only. It can also be seen that the QLD is positively skewed.

FAILURE RATE AND MEAN RESIDUAL LIFE

For a continuous distribution with p.d.f. $f(x)$ and c.d.f. $F(x)$, the failure rate function (also known as the hazard rate function) and the mean residual life function are respectively defined as

$$h(x) = \lim_{\Delta x \to 0} \frac{P\left(X < x+\Delta x \mid X > x\right)}{\Delta x} = \frac{f(x)}{1-F(x)} \qquad (15)$$

And

$$m(x) = E\left[X - x \mid X > x\right] = \frac{1}{1-F(x)} \int_x^\infty \left[1-F(t)\right] dt \qquad (16)$$

The corresponding failure rate function, $h(x)$ and the mean residual life function, $m(x)$ of the QLD are thus given by:

$$h(x) = \frac{\theta\left(\alpha + x\theta\right)}{1+\alpha+x\theta} \qquad (17)$$

And

$$m(x) = \frac{1}{\left(1+\alpha+\theta x\right)e^{-\theta x}} \int_x^\infty \left(1+\alpha+\theta t\right)e^{-\theta t} dt = \frac{2+\alpha+\theta x}{\theta\left(1+\alpha+\theta x\right)} \qquad (18)$$

It can be easily verified that $h(0) = \frac{\theta\alpha}{1+\alpha} = f(0)$ and $m(0) = \frac{2+\alpha}{\theta(1+\alpha)} = \mu_1'$. It is also obvious that $h(x)$ is an increasing function of x, α and θ, whereas $m(x)$ is a decreasing function of x, α and θ. The failure rate function and the mean residual life function of the QLD show its flexibility over Lindley distribution and exponential distribution. The nature of hazard rate function of QLD has been shown graphically in Figure 3 for different values of its parameters and it is obvious from the various graphs that the hazard rate function is

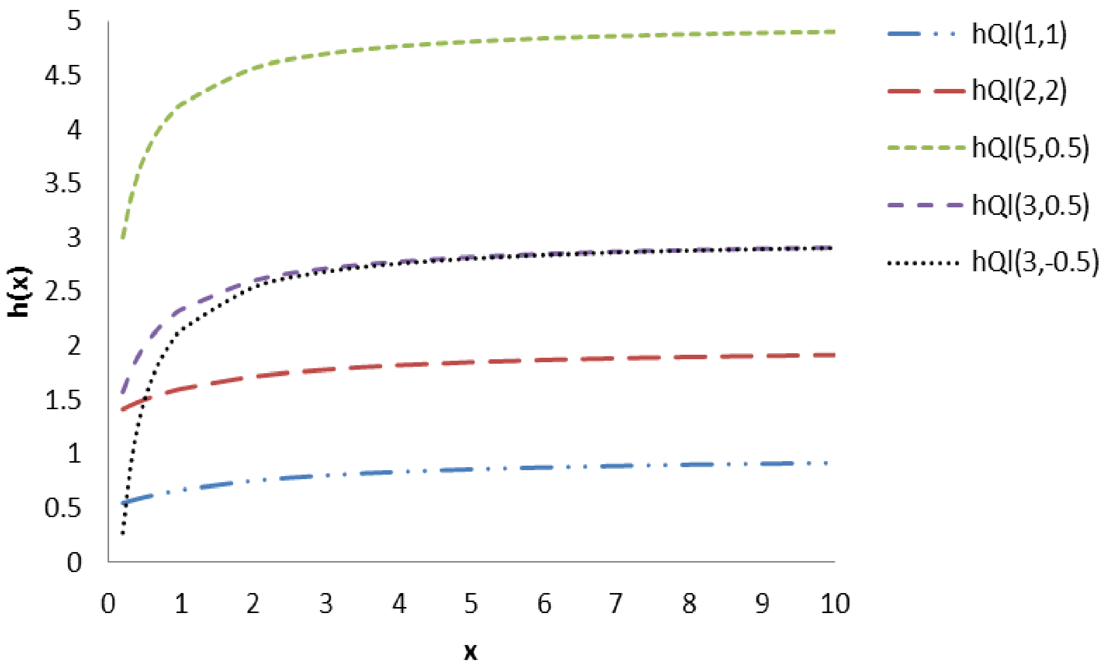

Figure 3. Plots of the hazard rate function (17) for some parameter values θ and α.

monotonically increasing for $x, \theta, \text{and}\, \alpha$. In the figure, hQ (1,1) means hazard rate function with parameters $\theta = 1 \text{and}\, \alpha = 1$.

STOCHASTIC ORDERINGS

Stochastic ordering of positive continuous random variables is an important tool for judging the comparative behaviour. A random variable X is said to be smaller than a random variable Y in the

(i) Stochastic order $(X \le_{st} Y)$ if $F_X(x) \ge F_Y(x)$ for all x

(ii) Hazard rate order $(X \le_{hr} Y)$ if $h_X(x) \ge h_Y(x)$ for all x

(iii) Mean residual life order $(X \le_{mrl} Y)$ if $m_X(x) \le m_Y(x)$ for all x

(iv) Likelihood ratio order $(X \le_{lr} Y)$ if $\frac{f_X(x)}{f_Y(x)}$ decreases in x.

The following results due to Shaked and Shanthikumar (1994) are well known for establishing stochastic ordering of distributions

$$\begin{array}{l} X \le_{lr} Y \Rightarrow X \le_{hr} Y \Rightarrow X \le_{mrl} Y \\ \qquad\qquad\quad \Downarrow \\ \qquad\qquad X \le_{st} Y \end{array} \tag{19}$$

The QLD is ordered with respect to the strongest "likelihood ratio" ordering as shown in the following theorem:

Theorem: Let $X \sim$ QLD and $Y \sim$ QLD(θ_2, α_2). If $\theta_1 = \theta_2$ and $\alpha_1 \ge \alpha_2$ (or if $\alpha_1 = \alpha_2$ and $\theta_1 \ge \theta_2$), then $X \le_{lr} Y$ and hence $X \le_{hr} Y$, $X \le_{mrl} Y$ and $X \le_{st} Y$.

Proof: We have

$$\frac{f_X(x)}{f_Y(x)} = \frac{\theta_1(\alpha_2+1)}{\theta_2(\alpha_1+1)} \frac{\alpha_1 + x\theta_1}{\alpha_2 + x\theta_2} e^{-(\theta_1-\theta_2)x} \quad ; x > 0$$

Now

$$\log \frac{f_X(x)}{f_Y(x)} = \log \frac{\theta_1(\alpha_2+1)}{\theta_2(\alpha_1+1)} + \log(\alpha_1 + x\theta_1) - \log(\alpha_2 + x\theta_2) - (\theta_1 - \theta_2)x.$$

Thus

$$\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} = \frac{\theta_1}{\alpha_1 + x\theta_1} - \frac{\theta_2}{\alpha_2 + x\theta_2} - (\theta_1 - \theta_2) = \frac{\theta_1\alpha_2 - \theta_2\alpha_1}{(\alpha_1 + x\theta_1)(\alpha_2 + x\theta_2)} - (\theta_1 - \theta_2) \tag{20}$$

Case (i) If $\theta_1 = \theta_2$ and $\alpha_1 \ge \alpha_2$, then $\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} < 0$. This means that $X \le_{lr} Y$ and hence $X \le_{hr} Y$, $X \le_{mrl} Y$ and $X \le_{st} Y$.

Case (ii) If $\alpha_1 = \alpha_2$ and $\theta_1 \geq \theta_2$, then $\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)} < 0$.

This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$. This theorem shows the flexibility of QLD over Lindley and exponential distributions.

ESTIMATION OF PARAMETERS

Maximum likelihood estimates

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from QLD Equation (5) and let f_x be the observed frequency in the sample corresponding to $X = x \quad (x = 1, 2, ..., k)$ such that $\sum_{x=1}^{k} f_x = n$, where k is the largest observed value having non-zero frequency. The likelihood function, L of the QLD Equation (5) is given by

$$L = \left(\frac{\theta}{\alpha+1}\right)^n \prod_{x=1}^{k} \left(\alpha + x\theta\right)^{f_x} e^{-n\theta\bar{X}} \tag{21}$$

and so the log likelihood function is obtained as

$$\log L = n\log\theta - n\log(\alpha+1) + \sum_{x=1}^{k} f_x \log(\alpha + x\theta) - n\theta\bar{X} \tag{22}$$

The two log likelihood equations are thus obtained as

$$\frac{\partial \log L}{\partial\theta} = \frac{n}{\theta} + \sum_{x=1}^{k}\frac{xf_x}{\alpha + x\theta} - n\bar{X} = 0 \tag{23}$$

$$\frac{\partial \log L}{\partial\alpha} = -\frac{n}{\alpha+1} + \sum_{x=1}^{k}\frac{f_x}{\alpha + x\theta} = 0 \tag{24}$$

The two Equations (23) and (24) do not seem to be solved directly. However, the Fisher's scoring method can be applied to solve these equations. We have

$$\frac{\partial^2 \log L}{\partial\theta^2} = -\frac{n}{\theta^2} - \sum_{x=1}^{k}\frac{x^2 f_x}{\left(\alpha + x\theta\right)^2} \tag{25}$$

$$\frac{\partial^2 \log L}{\partial\theta\,\partial\alpha} = -\sum_{x=1}^{k}\frac{xf_x}{\left(\alpha + x\theta\right)^2} \tag{26}$$

$$\frac{\partial^2 \log L}{\partial\alpha^2} = -\frac{n}{(\alpha+1)^2} - \sum_{x=1}^{k}\frac{f_x}{(\alpha + x\theta)^2} \tag{27}$$

The following equations for $\hat{\theta}$ and $\hat{\alpha}$ can be solved:

$$\begin{bmatrix} \frac{\partial^2 \log L}{\partial\theta^2} & \frac{\partial^2 \log L}{\partial\theta\,\partial\alpha} \\ \frac{\partial^2 \log L}{\partial\theta\,\partial\alpha} & \frac{\partial^2 \log L}{\partial\alpha^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \begin{bmatrix} \hat{\theta} - \theta_0 \\ \hat{\alpha} - \alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \log L}{\partial\theta} \\ \frac{\partial \log L}{\partial\alpha} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \tag{28}$$

where θ_0 and α_0 are the initial values of θ and α, respectively. These equations are solved iteratively till sufficiently close values of $\hat{\theta}$ and $\hat{\alpha}$ are obtained.

Estimates from moments

Using the first two moments about origin of the QLD, we have

$$\frac{\mu_2'}{\mu_1'^2} = k \text{ (say)} = \frac{2(\alpha+3)(\alpha+1)}{(\alpha+2)^2} \tag{29}$$

This gives a quadratic equation in α as

$$(2-k)\alpha^2 + 4(2-k)\alpha + (6-4k) = 0 \tag{30}$$

Replacing the first and the second moments μ_1' and μ_2' by the respective sample moments, $\bar{X}$ and m_2' an estimate of k can be obtained, using which, the Equation (30) can be solved and an estimate of α obtained. Substituting this estimate of α in the expression for the mean of the QLD, an estimate of θ can be obtained as

$$\hat{\theta} = \left(\frac{\alpha+2}{\alpha+1}\right)\frac{1}{\bar{X}} \tag{31}$$

GOODNESS OF FIT

The QLD has been fitted to a number of data- sets to which earlier the Lindley distribution has been fitted by others and it was found that to almost all these data-sets, the QLD provides closer fits than those by the Lindley distribution. Here the fittings of the QLD to two such data-sets have been presented in Tables 1 and 2.

The first data set is regarding the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by Bjerkedal (1960) and the second data set is regarding mortality grouped data for blackbird species, reported by Paranjpe and Rajarshi (1986).

The expected frequencies according to the Lindley distribution have also been given for ready comparison with those obtained by the QLD. The estimates of the parameters have been obtained by the method of moments. It can be seen that the QLD gives much closer fits than the Lindley distribution and thus provides a better alternative to the Lindley distribution.

Conclusion

In this study, we propose a two-parameter QLD, of which the one-parameter LD is a particular case. Several

Table 1. Data of survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by Bjerkedal (1960).

Survival time (In days)	Observed frequency	Expected frequency	
		Lindley distribution	QLD
0 – 80	8	16.1	10.7
80 – 160	30	21.9	26.9
160 – 240	18	15.4	17.7
240 – 320	8	9.0	9.1
320 – 400	4	5.5	4.3
400 – 480	3	1.8	1.9
480 – 560	1	2.3	1.4
Total	72	72.0	72.0
Estimates of parameters		$\hat{\theta} = 0.011$	$\hat{\alpha} = -0.2609,\ \hat{\theta} = 0.013$
χ^2		7.8878	1.1976
d.f.		3	2

Table 2. Mortality grouped data for blackbird species reported by Paranjpe and Rajarshi (1986).

Survival time (In days)	Observed frequency	Expected frequency	
		Lindley distribution	QLD
0 – 1	192	173.536	167.974
1 – 2	60	98.560	88.422
2 – 3	50	46.464	46.218
3 – 4	20	20.064	23.971
4 – 5	12	8.096	12.426
5 – 6	7	3.168	6.336
6 – 7	6	1.408	3.274
7 – 8	3	0.352	1.654
> 8	2	0.352	1.725
Total	352	352.000	352.000
Estimates of parameters		$\hat{\theta} = 0.984$	$\hat{\alpha} = 7.4910, \hat{\theta} = 0.7312$
χ^2		49.846	16.464
d.f.		4	3

properties of the QLD such as moments, skewness and kurtosis have been discussed. Various reliability properties such as failure rate function, mean residual life function, stochastic orderings have been obtained and discussed and shown that the QLD is more flexible than Lindley and exponential distributions. The density function of the QLD along with its cumulative distribution function and hazard rate function has been shown graphically for different values of its parameters for comparative study with Lindley distribution. The estimation of parameters by the method of maximum likelihood and the method of moments has been discussed. Finally, the proposed distribution has been fitted to a number of data sets relating to survival times to test its goodness of fit to which earlier the Lindley distribution has been fitted and it is found that QLD provides better fits than those by the Lindley distribution

ACKNOWLEDGEMENT

The authors express their gratitude to the referee for his valuable comments and suggestions which improved the quality of the paper.

REFERENCES

Bakouch HS, Al-Zahrani BM, Al-Shomrani AA, Marchi VAA, Louzada F (2012). An extended Lindley distribution. J. Korean Stat. Soc. 41(1):75-85.

Bjerkedal T (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. Am. J. Epidemiol. 72(1):130-148.

Deniz EG, Ojeda EC (2011). The discrete Lindley distribution-Properties and Applications. J. Stat. Comput. Simul. 81(11):1405-1416.
Ghitany ME, Atieh B, Nadarajah S (2008a). Lindley distribution and its Applications. Math. Comput. Simul. 78(4):493-506.
Ghitany ME, Al-Mutairi DK (2008b). Size-biased Poisson-Lindley distribution and its Applications. Metron- Int. J. Stat. LXVI(3):299-311.
Ghitany ME, Al-Mutairi DK (2009). Estimation methods for the discrete Poisson- Lindley distribution. J. Stat. Comput. Simul. 79(1):1-9.
Ghitany ME, Al-Mutairi DK, Nadarajah S (2008c). Zero-truncated Poisson-Lindley distribution and its Applications. Math. Comput. Simul. 79(3):279-287.
Ghitany ME, Al-qallaf F, Al-Mutairi DK, Hussain HA (2011). A two parameter weighted Lindley distribution and its applications to survival data. Math. Comput. Simul. 81(6):1190-1201.
Lindley DV (1958). Fiducial distributions and Bayes' theorem. J. Royal Stat. Soc. Series B, 20:102-107.
Mazucheli J, Achcar JA (2011). The Lindley distribution applied to competing risks lifetime data. Comput. Methods Programs Biomed. 104(2):188-192.
Paranjpe S, Rajarshi MB (1986). Modeling non-monotonic survivorship data with bath tube distributions. Ecology 67(6):1693-1695.
Sankaran M (1970). The discrete Poisson-Lindley distribution. Biometrics pp. 145-149.
Shaked M, Shanthikumar JG (1994). Stochastic Orders and Their Applications, Academic Press, New York.
Zakerzadah H, Dolati A (2010). Generalized Lindley distribution. J. Math. Ext. 3(2):13-25.