Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.3, No.13, 2013

A new generalized Lindley distribution

Ibrahim Elbatal[1] Faton Merovci[2*] M. Elgarhy[3]

1. Institute of Statistical Studies and Research, Departmentof Mathematical Statistics, Cairo University.
2. Department of Mathematics, University of Prishtina "Hasan Prishtina" Republic of Kosovo
3. Institute of Statistical Studies and Research, Departmentof Mathematical Statistics, Cairo University.

[1]E-mail: i_elbatal@staff.cu.edu.eg
*[2] E-mail of the corresponding author: fmerovci@yahoo.com
[3]E-mail: m_elgarhy85@yahoo.com

Abstract

In this paper, we present a new class of distributions called New Generalized Lindley Distribution(NGLD). This class of distributions contains several distributions such as gamma, exponential and Lindley as special cases. The hazard function, reverse hazard function, moments and moment generating function and inequality measures are are obtained. Moreover, we discuss the maximum likelihood estimation of this distribution. The usefulness of the new model is illustrated by means of two real data sets. We hope that the new distribution proposed here will serve as an alternative model to other models available in the literature for modelling positive real data in many areas.

Keywords: Generalized Lindley Distribution; Gamma distribution, Maximum likelihood estimation; Moment generating function.

1 Introduction and Motivation

In many applied sciences such as medicine, engineering and finance, amongst others, modeling and analyzing lifetime data are crucial. Several lifetime distributions have been used to model such kinds of data. For instance, the exponential, Weibull, gamma, Rayleigh distributions and their generalizations (see, e.g., Gupta and Kundu, [10]). Each distribution has its own characteristics due specifically to the shape of the failure rate function which may be only monotonically decreasing or increasing or constant in its behavior, as well as non-monotone, being bathtub shaped or even unimodal. Here we consider the Lindley distribution which was introduced by Lindley [13]. Let the life time random variable X has a Lindley distribution with parameter θ, the probability density function (pdf) of X is given by

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

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

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

where θ is scale parameter. The Lindley distribution is important for studying stress–strength reliability modeling. Besides, some researchers have proposed new classes of distributions based on modifications of the Lindley distribution, including also their properties. The main idea is always directed by embedding former distributions to more flexible structures. Sankaran [16] introduced the discrete Poisson–Lindley distribution by combining the Poisson and Lindley distributions. Ghitany et al. [5] 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. A discrete version of this distribution has been suggested by Deniz and Ojeda [3] having its applications in count data related to insurance. Ghitany et al. [7, 8] obtained size-biased and zero-truncated version of Poisson- Lindley distribution and discussed their various properties and applications. Ghitany and Al-Mutairi [6] discussed as various estimation methods for the discrete Poisson- Lindley distribution. Bakouch et al. [1] obtained an extended Lindley distribution and discussed its various properties and applications. Mazucheli and Achcar [14] discussed the applications of Lindley distribution to competing risks lifetime data. Rama and

Mishra [15] studied quasi Lindley distribution. Ghitany et al. [9] developed a two-parameter weighted Lindley distribution and discussed its applications to survival data. Zakerzadah and Dolati [18] obtained a generalized Lindley distribution and discussed its various properties and applications.

This paper offers new distribution with three parameter called generalizes the Lindley distribution, this distribution includes as special cases the ordinary exponential and gamma distributions. The procedure used here is based on certain mixtures of the gamma distributions. The study examines various properties of the new model. The rest of the paper is organized as follows: Various statistical properties includes moment, generating function and inequality measures of the NGL distribution are explored in Section 2. The distribution of the order statistics is expressed in Section 3. We provide the regression based method of least squares and weighted least squares estimators in Section 4. Maximum likelihood estimates of the parameters index to the distribution are discussed in Section 5. Section 6 provides applications to real data sets. Section 7 ends with some conclusions

2 Statistical Properties and Reliability Measures

In this section, we investigate the basic statistical properties, in particular, r_{th} moment, moment generating function and inequality measures for the NGL distribution.

2.1 Density. survival and failure rate functions

The new generalized Lindley distribution is denoted as $NGLD(\alpha,\beta,\theta)$. This generalized model is obtained from a mixture of the gamma (α,θ) and gamma (β,θ) distributions as follows:

$$\begin{aligned} f(x,\theta,\alpha,\beta) &= pf_1(x,\alpha,\theta)+(1-p)f_2(x,\beta,\theta) \\ &= \frac{1}{1+\theta}\left[\frac{\theta^{\alpha+1}x^{\alpha-1}}{\Gamma(\alpha)}+\frac{\theta^{\beta}x^{\beta-1}}{\Gamma(\beta)}\right]e^{-\theta x};\alpha,\theta>0,x>0. \quad (3) \end{aligned}$$

where

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

The corresponding cumulative distribution function (cdf) of generalized Lindley is given by

$$F(x,\theta,\alpha,\beta)=\frac{1}{1+\theta}\left[\frac{\theta\gamma(\alpha,\theta x)}{\Gamma(\alpha)}+\frac{\gamma(\beta,\theta x)}{\Gamma(\beta)}\right], \quad (4)$$

where

$$\gamma(s,t)=\int_0^t x^{s-1}e^{-x}dx$$

is called lower incomplete gamma. Also the upper incomplete gamma is given by

$$\Gamma(\alpha,t)=\int_t^\infty x^{\alpha-1}e^{-x}dx$$

For more details about the definition of incomplete gamma, see Wall [20]. Figures 0 and 1 illustrates some of the possible shapes of the pdf and cdf of the NGL distribution for selected values of the parameters θ,α and $\beta,$ respectively.

The survival function associated with (4) is given by

$$\overline{F}(x,\theta,\alpha,\beta)=1-F(x,\theta,\alpha,\beta)=1-\frac{1}{1+\theta}\left[\frac{\theta\gamma(\alpha,\theta x)}{\Gamma(\alpha)}+\frac{\gamma(\beta,\theta x)}{\Gamma(\beta)}\right], \quad (5)$$

From (3), (4) and (5), the failure (or hazard) rate function (hf) and reverse hazard functions (rhf) of generalized Lindley distribution are given by

$$h(x)=\frac{f(x,\theta,\alpha,\beta)}{\overline{F}(x,\theta,\alpha,\beta)}=\frac{\frac{1}{1+\theta}\left[\frac{\theta^{\alpha+1}x^{\alpha-1}}{\Gamma(\alpha)}+\frac{\theta^{\beta}x^{\beta-1}}{\Gamma(\beta)}\right]e^{-\theta x}}{1-\frac{1}{1+\theta}\left[\frac{\theta\gamma(\alpha,\theta x)}{\Gamma(\alpha)}+\frac{\gamma(\beta,\theta x)}{\Gamma(\beta)}\right]},\ (6)$$

and

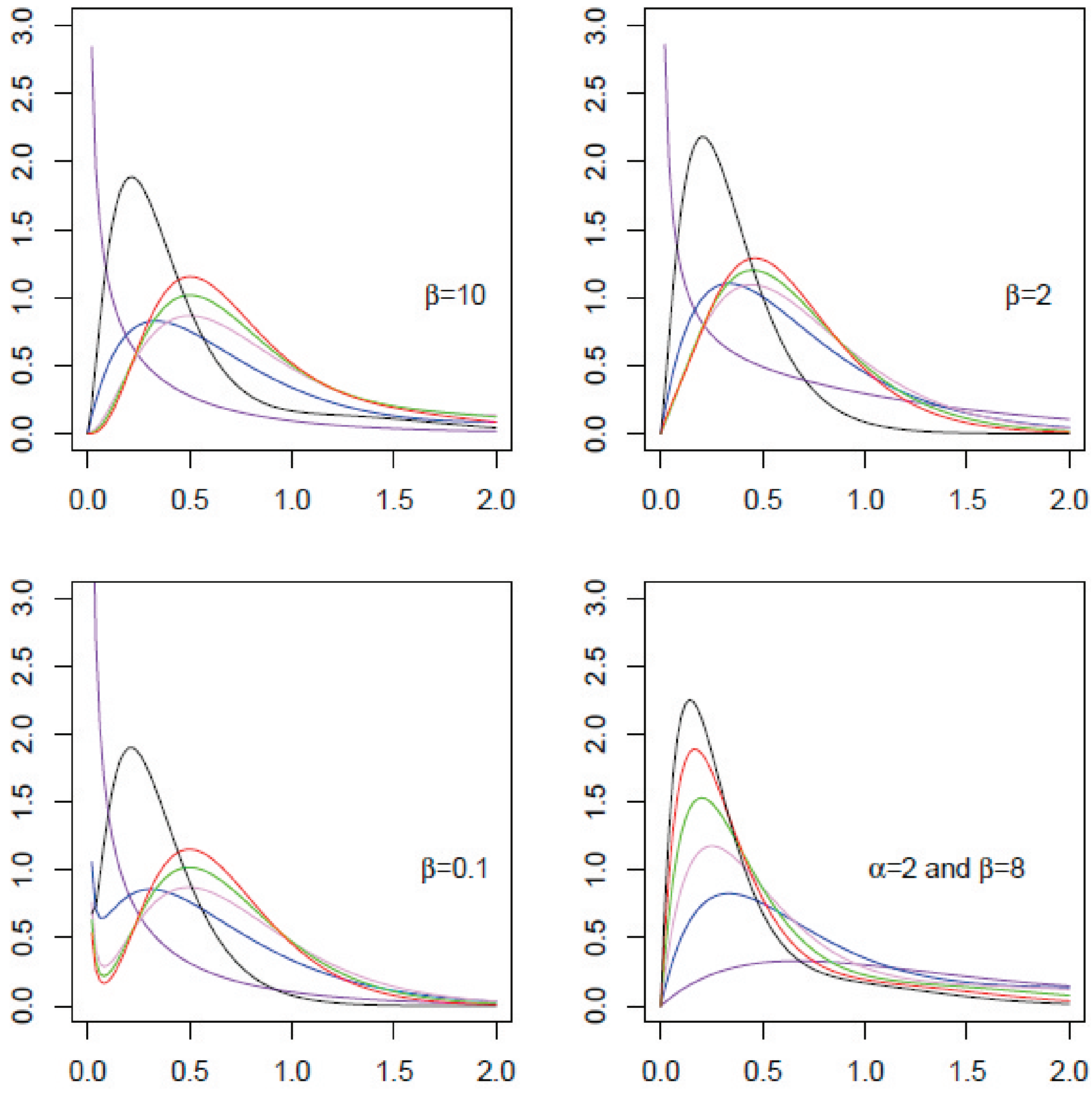

Figure 1: The pdf's of various NGL distributions for values of parameters: $\theta=1.5,3,4,5,6,7;\quad \alpha=0.5,2,3,3.5,4,2.5;$ with color shapes purple, blue, plum, green, red, black and darkcyan, respectively.

ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.3, No.13, 2013

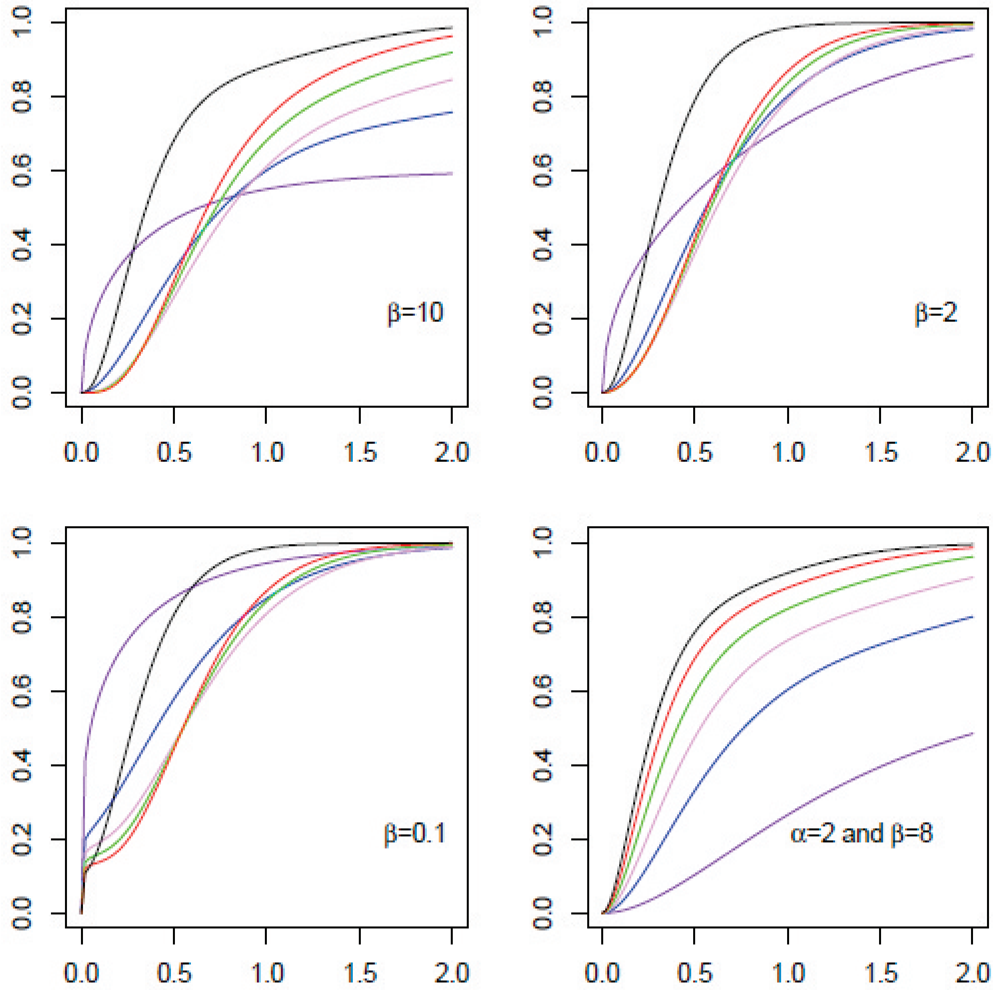

Figure 2: The cdf's of various NGL distributions for values of parameters: $\theta=1,2,3,4,5,6; a=1,2,3,4,5,6$ with color shapes red, green,plum,darcyan, black and purple, respectively.

$$\tau(x)=\frac{f(x,\theta,\alpha,\beta)}{F(x,\theta,\alpha,\beta)}=\frac{\frac{1}{1+\theta}\left[\frac{\theta^{\alpha+1}x^{\alpha-1}}{\Gamma(\alpha)}+\frac{\theta^{\beta}x^{\beta-1}}{\Gamma(\beta)}\right]e^{-\theta x}}{\frac{1}{1+\theta}\left[\frac{\theta\gamma(\alpha,\theta x)}{\Gamma(\alpha)}+\frac{\gamma(\beta,\theta x)}{\Gamma(\beta)}\right]}. \quad (7)$$

respectively.

Figure 3 illustrates some of the possible shapes of the hazard function of the NGL distribution for selected values of the parameters θ,α and β, respectively.

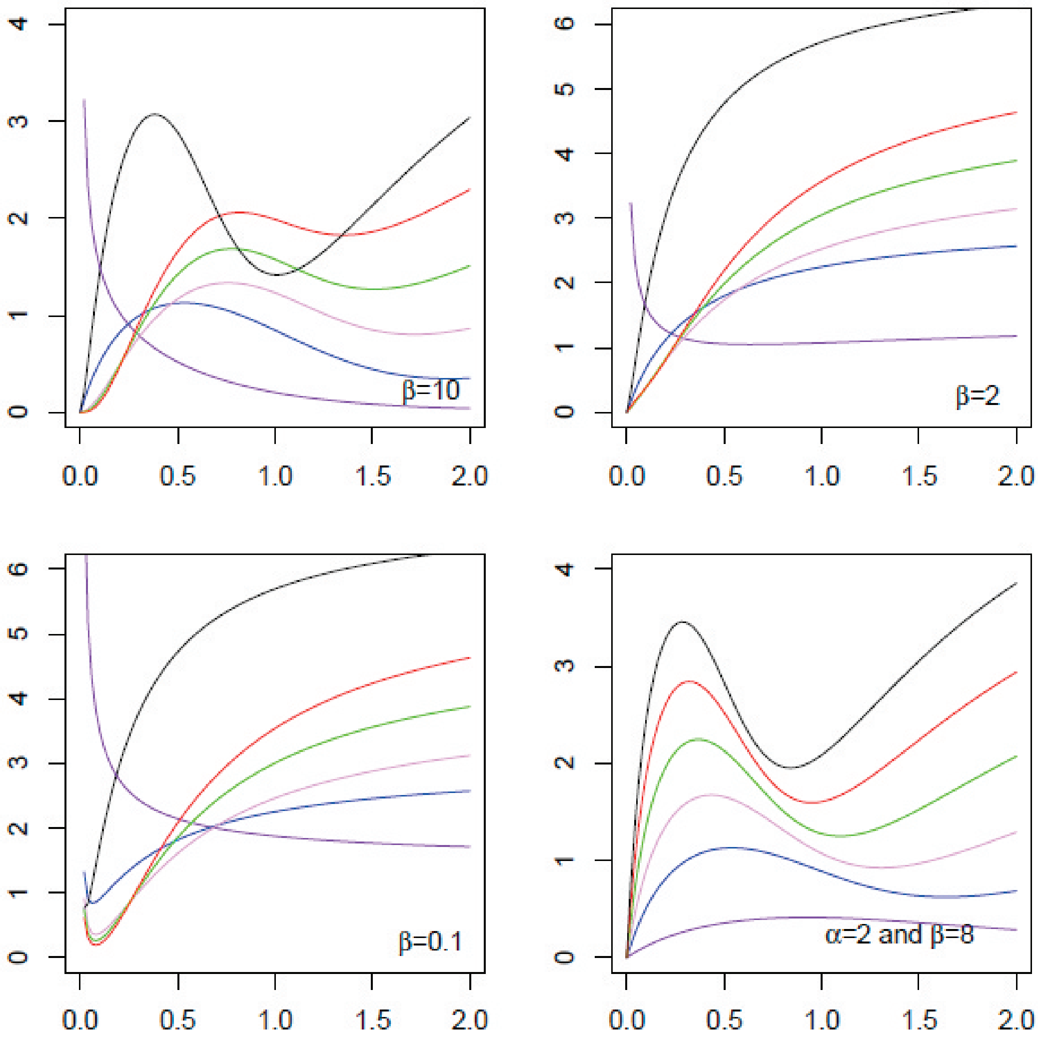

Figure 3: The hazard's of various NGL distributions for values of parameters: $\theta = 1.5,3,4,5,6,7;$ $\alpha = 0.5,2,3,3.5,4,2.5;$ with color shapes purple, blue, plum, green, red,blackand darkcyan, respectively.

The following are special cases of the generalized Lindley distribution , $GLD(\alpha,\beta,\theta)$.

1. If $\alpha = 1$ and $\beta = 2$, we get the Lindley distribution .
2. For $\alpha = \beta = \lambda$, we get the Gamma distribution with parameter (θ,λ).
3. If $\alpha = \beta = 1$, we get the exponential distribution with parameter θ.

2.2 Moments

Many of the interesting characteristics and features of a distribution can be studied through its moments. Moments are necessary and important in any statistical analysis, especially in applications. It can be used to study the most important features and characteristics of a distribution (e.g., tendency, dispersion, skewness and kurtosis).

Theorem 2.1. If X has $GL(\phi,x)$ $,\phi = (\alpha,\theta,\beta)$ then the r_{th} moment of X is given by the following

$$\mu_r^{'}(x) = \frac{1}{1+\theta}\left[\frac{\Gamma(r+\alpha)}{\theta^{r-1}\Gamma(\alpha)} + \frac{\Gamma(r+\beta)}{\theta^{r}\Gamma(\beta)}\right]. \qquad (8)$$

Proof:

Let X be a random variable following the GL distribution with parameters θ, α and β. Expressions for mathematical expectation,

variance and the r_{th} moment on the origin of X can be obtained using the well-known formula

$$\mu_r'(x) = E(X^r) = \int_0^\infty x^r f(x,\phi)dx$$

$$= \frac{1}{1+\theta}\left[\frac{\theta^{\alpha+1}}{\Gamma(\alpha)}\int_0^\infty x^{r+\alpha-1}e^{-\theta x}dx + \frac{\theta^\beta}{\Gamma(\beta)}\int_0^\infty x^{r+\beta-1}e^{-\theta x}dx\right]$$

$$= \frac{1}{1+\theta}\left[\frac{\Gamma(r+\alpha)}{\theta^{r-1}\Gamma(\alpha)} + \frac{\Gamma(r+\beta)}{\theta^r\Gamma(\beta)}\right]. \quad (9)$$

Which completes the proof .

Based on the first four moments of the GL distribution, the measures of skewness $A(\varphi)$ and kurtosis $k(\varphi)$ of the GL distribution can obtained as

$$A(\varphi) = \frac{\mu_3(\theta) - 3\mu_1(\theta)\mu_2(\theta) + 2\mu_1^3(\theta)}{\left[\mu_2(\theta) - \mu_1^2(\theta)\right]^{\frac{3}{2}}},$$

and

$$k(\varphi) = \frac{\mu_4(\theta) - 4\mu_1(\theta)\mu_3(\theta) + 6\mu_1^2(\theta)\mu_2(\theta) - 3\mu_1^4(\theta)}{\left[\mu_2(\theta) - \mu_1^2(\theta)\right]^2}.$$

2.3 Moment generating function

In this subsection we derived the moment generating function of GL distribution.

Theorem (2.2): If X has GL distribution, then the moment generating function $M_X(t)$ has the following form

$$M_X(t) = \frac{1}{1+\theta}\left[\frac{\theta^{\alpha+1}}{(\theta-t)^\alpha} + \frac{\theta^\beta}{(\theta-t)^\beta}\right]. \quad (10)$$

Proof.

We start with the well known definition of the moment generating function given by

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} f_{GL}(x,\phi)dx$$

$$= \frac{1}{1+\theta}\left[\frac{\theta^{\alpha+1}}{\Gamma(\alpha)}\int_0^\infty x^{\alpha-1}e^{-(\theta-t)x}dx + \frac{\theta^\beta}{\Gamma(\beta)}\int_0^\infty x^{\beta-1}e^{-(\theta-t)x}dx\right]$$

$$= \frac{1}{1+\theta}\left[\frac{\theta^{\alpha+1}}{(\theta-t)^\alpha} + \frac{\theta^\beta}{(\theta-t)^\beta}\right] \quad (11)$$

Which completes the proof.

In the same way, the characteristic function of the GL distribution becomes $\varphi_X(t) = M_X(it)$ where $i = \sqrt{-1}$ is the unit imaginary number.

2.4 Inequality Measures

Lorenz and Bonferroni curves are the most widely used inequality measures in income and wealth distribution (Kleiber, 2004). Zenga curve was presented by Zenga [19]. In this section, we will derive Lorenz, Bonferroni and Zenga curves for the GL distribution. The Lorenz, Bonferroni and Zenga curves are defined by

$$L_F(x)=\frac{\int_0^t xf(x)dx}{E(X)}=\frac{\left[\frac{\gamma(\alpha+1,\theta t)}{\Gamma(\alpha)}+\frac{\gamma(\beta+1,\theta t)}{\theta\Gamma(\beta)}\right]}{\left[\alpha+\frac{\beta}{\theta}\right]}. \quad (12)$$

$$B_F(x)=\frac{\int_0^t xf(x)dx}{E(X)F(x)}=\frac{L_F(x)}{F(x)}$$
$$=\frac{(1+\theta)\left[\frac{\gamma(\alpha+1,\theta t)}{\Gamma(\alpha)}+\frac{\gamma(\beta+1,\theta t)}{\theta\Gamma(\beta)}\right]}{\left[\alpha+\frac{\beta}{\theta}\right]\left[\frac{\theta\gamma(\alpha,\theta x)}{\Gamma(\alpha)}+\frac{\gamma(\beta,\theta x)}{\Gamma(\beta)}\right]}, \quad (13)$$

and

$$A_F(x)=1-\frac{\mu^-(x)}{\mu^+(x)}, \quad (14)$$

where

$$\mu^-(x)=\frac{\int_0^t xf(x)dx}{E(X)}=\frac{\left[\frac{\gamma(\alpha+1,\theta t)}{\Gamma(\alpha)}+\frac{\gamma(\beta+1,\theta t)}{\theta\Gamma(\beta)}\right]}{\left[\alpha+\frac{\beta}{\theta}\right]}$$

and

$$\mu^+(x)=\frac{\int_t^\infty xf(x)dx}{1-F(x)}$$
$$=\frac{\frac{1}{(1+\theta)}\left[\frac{\Gamma(\alpha+1,\theta t)}{\Gamma(\alpha)}+\frac{\Gamma(\beta+1,\theta t)}{\theta\Gamma(\beta)}\right]}{1-\frac{1}{1+\theta}\left[\frac{\theta\gamma(\alpha,\theta x)}{\Gamma(\alpha)}+\frac{\gamma(\beta,\theta x)}{\Gamma(\beta)}\right]}.$$

respectively.

The mean residual life (mrl) function computes the expected remaining survival time of a subject given survival up to time x. We have already defined the mrl as the expectation of the remaining survival time given survival up to time x(see Frank Guess and Frank Proschan [4].

3 Distribution of the order statistics

In this section, we derive closed form expressions for the pdfs of the r_{th} order statistic of the GL distribution, also, the measures of skewness and kurtosis of the distribution of the r_{th} order statistic in a sample of size n for different choices of $n;r$ are presented in this section. Let $X_1, X_2,..., X_n$ be a simple random sample from GL distribution with pdf and cdf given by (??) and (4), respectively.

Let $X_1, X_2,..., X_n$ denote the order statistics obtained from this sample. We now give the probability density function of $X_{r:n}$, say $f_{r:n}(x,\phi)$ and the moments of $X_{r:n}$, $r = 1,2,..., n$. Therefore, the measures of skewness and kurtosis of the distribution of the $X_{r:n}$ are presented. The probability density function of $X_{r:n}$ is given by

$$f_{r:n}(x,\Phi) = \frac{1}{B(r, n-r+1)}\left[F(x,\phi))\right]^{r-1}\left[1-F(x,\phi))\right]^{n-r} f(x,\phi)) \qquad (15)$$

where $f(x,\phi))$ and $F(x,\phi))$ are the pdf and cdf of the GL distribution given by (3) and (4), respectively, and B (.,.) is the beta function, since $0 < F(x,\phi)) < 1$, for $x > 0$, by using the binomial series expansion of $\left[1-F(x,\phi))\right]^{n-r}$, given by

$$\left[1-F(x,\phi))\right]^{n-r} = \sum_{j=0}^{n-r}(-1)^j \binom{n-r}{j}\left[F(x,\phi))\right]^j, \qquad (16)$$

we have

$$f_{r:n}(x,\phi)) = \sum_{j=0}^{n-r}(-1)^j \binom{n-r}{j}\left[F(x,\Phi)\right]^{r+j-1} f(x,\phi)), \qquad (17)$$

substituting from (3) and (4) into (17), we can ex press the k_{th} ordinary moment of the r_{th} order statistics $X_{r:n}$ say $E(X^k_{r:n})$ as a liner combination of the k_{th} moments of the GL distribution with different shape parameters. Therefore, the measures of skewness and kurtosis of the distribution of $X_{r:n}$ can be calculated.

4 Least Squares and Weighted Least Squares Estimators

In this section we provide the regression based method estimators of the unknown parameters of the GL distribution, which was originally suggested by Swain, Venkatraman and Wilson [17] to estimate the parameters of beta distributions. It can be used some other cases also. Suppose $Y_1,..., Y_n$ is a random sample of size n from a distribution function $G(.)$ and suppose $Y_{(i)}$; $i = 1,2,..., n$ denotes the ordered sample. The proposed method uses the distribution of $G(Y_{(i)})$. For a sample of size n, we have

$$E\left(G(Y_{(j)})\right) = \frac{j}{n+1}, V\left(G(Y_{(j)})\right) = \frac{j(n-j+1)}{(n+1)^2(n+2)}$$

$$\text{and} Cov\left(G(Y_{(j)}), G(Y_{(k)})\right) = \frac{j(n-k+1)}{(n+1)^2(n+2)}; \text{for } j < k,$$

see Johnson, Kotz and Balakrishnan [11]. Using the expectations and the variances, two variants of the least squares methods can be used.

Method 1 (Least Squares Estimators) . Obtain the estimators by minimizing

$$\sum_{j=1}^{n}\left(G(Y_{(j)} - \frac{j}{n+1}\right)^2,$$

with respect to the unknown parameters. Therefore in case of GL distribution the least squares estimators of $\alpha, \theta,$ and β , say $\hat{\alpha}_{LSE}$, $\hat{\theta}_{LSE}$ and $\hat{\beta}_{LSE}$ respectively, can be obtained by minimizing

$$\sum_{j=1}^{n}\left[\frac{1}{1+\theta}\left[\frac{\theta\gamma(\alpha,\theta x_{(j)})}{\Gamma(\alpha)} + \frac{\gamma(\beta,\theta x_{(j)})}{\Gamma(\beta)}\right] - \frac{j}{n+1}\right]^2$$

with respect to α, θ, and β.

Method 2 (Weighted Least Squares Estimators). The weighted least squares estimators can be obtained by minimizing

$$\sum_{j=1}^{n} w_j \left(G(Y_{(j)} - \frac{j}{n+1} \right)^2,$$

with respect to the unknown parameters, where

$$w_j = \frac{1}{V\left(G(Y_{(j)})\right)} = \frac{(n+1)^2(n+2)}{j(n-j+1)}.$$

Therefore, in case of GL distribution the weighted least squares estimators of α, θ, and β, say $\widehat{\alpha}_{WLSE}, \widehat{\theta}_{WLSE}$, and $\widehat{\beta}_{WLSE}$ respectively, can be obtained by minimizing

$$\sum_{j=1}^{n} w_j \left[\frac{1}{1+\theta} \left[\frac{\theta\gamma(\alpha, \theta x_{(j)})}{\Gamma(\alpha)} + \frac{\gamma(\beta, \theta x_{(j)})}{\Gamma(\beta)} \right] - \frac{j}{n+1} \right]^2$$

with respect to the unknown parameters only.

5 Maximum Likelihood Estimators

In this section we consider the maximum likelihood estimators (MLE's) of GL distribution. Let $\Phi = (\alpha, \theta, \beta)^T$, in order to estimate the parameters α, θ, and β of GL distribution, let $x_1, \ldots, x_n$ be a random sample of size n from $GL(\alpha, \theta, \beta, x)$ then the likelihood function can be written as

$$L(\alpha, \theta, \beta, x_{(i)}) = \prod_{i=1}^{n} \frac{1}{1+\theta} \left[\frac{\theta^{\alpha+1} x_{(i)}^{\alpha-1}}{\Gamma(\alpha)} + \frac{\theta^{\beta} x_{(i)}^{\beta-1}}{\Gamma(\beta)} \right] e^{-\theta x_{(i)}}$$

$$= \left(\frac{1}{1+\theta} \right)^n e^{-\theta \sum_{i=1}^{n} x_{(i)}} \left(\Gamma(\alpha)\Gamma(\beta) \right)^{-n}$$

$$\times \prod_{i=1}^{n} \left(\Gamma(\beta)\theta^{\alpha+1} x_{(i)}^{\alpha-1} + \Gamma(\alpha)\theta^{\beta} x_{(i)}^{\beta-1} \right) \qquad (18)$$

By accumulation taking logarithm of equation (18) , and the log- likelihood function can be written as

$$\log L = -n\log(1+\theta) - \theta \sum_{i=1}^{n} x_i - n\log\Gamma(\alpha) - n\log\Gamma(\beta)$$

$$+ \sum_{i=1}^{n} \log\left(\Gamma(\beta)\theta^{\alpha+1} x_{(i)}^{\alpha-1} + \Gamma(\alpha)\theta^{\beta} x_{(i)}^{\beta-1} \right) \qquad (19)$$

Differentiating $\log L$ with respect to each parameter α, θ, and β and setting the result equals to zero, we obtain maximum likelihood estimates. The partial derivatives of $\log L$ with respect to each parameter or the score function is given by

$$U_n(\Phi) = \left(\frac{\partial \log L}{\partial \theta}, \frac{\partial \log L}{\partial \alpha},, \frac{\partial \log L}{\partial \beta} \right)$$

where

$$\frac{\partial \log L}{\partial \theta} = \frac{-n}{(1+\theta)} - \sum_{i=1}^{n} x_i$$

$$+\sum_{i=1}^{n}\frac{(\alpha+1)\Gamma(\beta)\theta^{\alpha}x_{(i)}^{\alpha-1}+\beta\Gamma(\alpha)\theta^{\beta-1}x_{(i)}^{\beta-1}}{\left(\Gamma(\beta)\theta^{\alpha+1}x_{(i)}^{\alpha-1}+\Gamma(\alpha)\theta^{\beta}x_{(i)}^{\beta-1}\right)}=0,$$

$$\frac{\partial\log L}{\partial\alpha}=-n\psi(\alpha)+\sum_{i=1}^{n}\frac{\theta^{\alpha+1}x_{(i)}^{\alpha-1}\log(\theta x_i)+\Gamma^{'}(\alpha)\theta^{\beta}x_{(i)}^{\beta-1}}{\left(\Gamma(\beta)\theta^{\alpha+1}x_{(i)}^{\alpha-1}+\Gamma(\alpha)\theta^{\beta}x_{(i)}^{\beta-1}\right)}=0, \tag{20}$$

and

$$\frac{\partial\log L}{\partial\beta}=-n\psi(\beta)\sum_{i=1}^{n}\frac{\theta^{\alpha+1}x_{(i)}^{\alpha-1}\Gamma^{'}(\beta)+\theta^{\beta}x_{(i)}^{\beta-1}\log(\theta x_i)}{\left(\Gamma(\beta)\theta^{\alpha+1}x_{(i)}^{\alpha-1}+\Gamma(\alpha)\theta^{\beta}x_{(i)}^{\beta-1}\right)}$$
$$=0. \tag{21}$$

where $\psi(.)$ is the digamma function. By solving this nonlinear system of equations (20) - (21), these solutions will yield the ML estimators for $\widehat{\theta}$, $\widehat{\alpha}$ and $\widehat{\beta}$. For the three parameters generalized Lindley distribution $GL(\alpha,\theta,\beta,x)$ pdf all the second order derivatives exist. Thus we have the inverse dispersion matrix is given by

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\theta}\\ \hat{\beta}\end{pmatrix}\sim N\left[\begin{pmatrix}\alpha\\ \theta\\ \beta\end{pmatrix},\begin{pmatrix}\hat{V}_{\alpha\alpha} & \hat{V}_{\alpha\theta} & \hat{V}_{\alpha\beta}\\ \hat{V}_{\theta\alpha} & \hat{V}_{\theta\theta} & \hat{V}_{\theta\beta}\\ \hat{V}_{\beta\alpha} & \hat{V}_{\beta\theta} & \hat{V}_{\beta\beta}\end{pmatrix}\right]. \tag{22}$$

$$V^{-1}=-E\begin{bmatrix}V_{\alpha\alpha} & V_{\alpha\theta} & V_{\alpha\beta}\\ V_{\theta\alpha} & V_{\theta\theta} & V_{\theta\beta}\\ V_{\beta\alpha} & V_{\beta\theta} & V_{\beta\beta}\end{bmatrix} \tag{23}$$

where

$$V_{\alpha\alpha}=\frac{\partial^2 L}{\partial\alpha^2},V_{\theta\theta}=\frac{\partial^2 L}{\partial\theta^2},V_{\beta\beta}=\frac{\partial^2 L}{\partial\beta^2},$$
$$V_{\alpha\theta}=\frac{\partial^2 L}{\partial\alpha\partial\theta},V_{\beta\alpha}=\frac{\partial^2 L}{\partial\alpha\partial\beta},V_{\beta\theta}=\frac{\partial^2 L}{\partial\beta\partial\theta}$$

$$V_{\alpha\alpha}=-n\psi^{'}(\alpha)+\sum_{i=1}^{n}(A_i+B_i)$$

$$A_i=\frac{\Gamma(\beta)(\ln(\theta))^2\theta^{\alpha+1}x_i^{\alpha-1}+2\Gamma(\beta)\theta^{\alpha+1}\ln(\theta)x_i^{\alpha-1}\ln(x_i)+\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}(\ln(x_i))^2}{\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}+\Gamma(\alpha)\theta^{\beta}x_i^{\beta-1}}$$
$$+\frac{\psi^{'}(\alpha)\Gamma(\alpha)\theta^{\beta}x_i^{\beta-1}+(\Psi(\alpha))^2\Gamma(\alpha)\theta^{\beta}x_i^{\beta-1}}{\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}+\Gamma(\alpha)\theta^{\beta}x_i^{\beta-1}}$$

$$B_i=\frac{\left(\Gamma(\beta)\theta^{\alpha+1}\ln(\theta)x_i^{\alpha-1}+\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}\ln(x_i)+\Psi(\alpha)\Gamma(\alpha)\theta^{\beta}x_i^{\beta-1}\right)^2}{\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}+\Gamma(\alpha)\theta^{\beta}x_i^{\beta-1}\right)^2}$$

$$V_{\alpha\theta}=\sum_{i=1}^{n}\frac{\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}(C_i+D_i)}{\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}+\Gamma(\alpha)\theta^{\beta}x_i^{\beta-1}\right)^2\theta}$$

$$C_i = \ln(\theta)\alpha\Gamma(\alpha)\theta^\beta x_i^{\beta-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\ln(\theta) + \Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}$$

$$D_i = \Gamma(\alpha)\theta^\beta x_i^{\beta-1} + \ln(x_i)\alpha\Gamma(\alpha)\theta^\beta x_i^{\beta-1} + \ln(x_i)\Gamma(\alpha)\theta^\beta x_i^{\beta-1}$$

$$+\psi(\alpha)\Gamma(\alpha)\theta^\beta \beta x_i^{\beta-1} - \ln(\theta)\Gamma(\alpha)\theta^\beta\beta x_i^{\beta-1} - \ln(x_i)\Gamma(\alpha)\theta^\beta\beta x_i^{\beta-1} - \Psi(\alpha)\Gamma(\alpha)\theta^\beta x_i^{\beta-1}\alpha - \Psi(\alpha)\Gamma(\alpha)\theta^\beta x_i^{\beta-1}$$

$$V_{\alpha\beta} = \sum_i^n \frac{\psi(\beta)\Gamma(\beta)\theta^{\alpha+1}\ln(\theta x_i)x_i^{\alpha-1} + \psi(\alpha)\Gamma(\alpha)\theta^\beta x_i^{\beta-1}\ln(\theta x_i)}{\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}}$$

$$-\sum_i^n \frac{\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}\ln(\theta x_i) + \Psi(\alpha)\Gamma(\alpha)\theta^\beta x_i^{\beta-1}\right)\left(\Psi(\beta)\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\ln(\theta x_i)\right)}{\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\right)^2}$$

$$V_{\theta\beta} = \sum_{i=1}^n \frac{\psi(\beta)\Gamma(\beta)\theta^{\alpha+1}(\alpha+1)x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta\beta x_i^{\beta-1}\ln(\theta x_i) + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}}{\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\right)\theta}$$

$$-\sum_{i=1}^n \frac{\left(\theta^{\alpha+1}\Gamma(\beta)(\alpha+1)x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta\beta x_i^{\beta-1}\right)\left(\psi(\beta)\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\ln(\theta x_i)\right)}{\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\right)^2\theta}$$

$$V_{\theta\theta} = \frac{n}{(1+\theta)^2} + \sum_{i=1}^n \frac{\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}\alpha^2 + \Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}\alpha + \Gamma(\alpha)\theta^\beta\beta^2 x_i^{\beta-1} - \Gamma(\alpha)\theta^\beta\beta x_i^{\beta-1}}{\theta^2\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\right)}$$

$$-\sum_{i=1}^n \frac{\left(\theta^{\alpha+1}\Gamma(\beta)(\alpha+1)x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta\beta x_i^{\beta-1}\right)^2}{\theta^2\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\right)^2}$$

$$V_{\beta\beta} = -n\psi'(\beta) + \sum_{i=1}^n \frac{\psi'(\beta)\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + (\psi(\beta))^2\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1}}{\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}}$$

$$+\sum_{i=1}^n \frac{\Gamma(\alpha)\theta^\beta(\ln(\theta))^2 x_i^{\beta-1} + 2\Gamma(\alpha)\theta^\beta x_i^{\beta-1}\ln(x_i)\ln(\theta) + (\ln(x_i))^2\Gamma(\alpha)\theta^\beta x_i^{\beta-1}}{\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}}$$

$$-\sum_{i=1}^n \frac{\left(\Psi(\beta)\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\ln(\theta x_i)\right)^2}{\left(\Gamma(\beta)\theta^{\alpha+1}x_i^{\alpha-1} + \Gamma(\alpha)\theta^\beta x_i^{\beta-1}\right)^2}$$

By solving this inverse dispersion matrix these solutions will yield asymptotic variance and covariances of these ML estimators for $\overline{\theta}$,$\overline{\alpha}$ and $\overline{\beta}$. Using (22), we approximate $100(1-\gamma)\%$ confidence intervals for $\alpha, \beta,$ and $\theta,$ are determined respectively as

$$\hat{\alpha} \pm z_{\frac{\gamma}{2}}\sqrt{\hat{V}_{\alpha\alpha}}, \hat{\theta} \pm z_{\frac{\gamma}{2}}\sqrt{\hat{V}_{\theta\theta}} \text{ and } \hat{\beta} \pm z_{\frac{\gamma}{2}}\sqrt{\hat{V}_{\beta\beta}}$$

where z_γ is the upper $100\gamma_{the}$ percentile of the standard normal distribution.

Using R we can easily compute the Hessian matrix and its inverse and hence the standard errors and asymptotic confidence intervals.

We can compute the maximized unrestricted and restricted log-likelihood functions to construct the likelihood ratio (LR) test statistic for testing on some the new generalized Lindley sub-models. For example, we can use the LR test statistic to check whether the new generalized Lindley distribution for a given data set is statistically *superior* to the Lindley distribution. In any case, hypothesis tests of the type $H_0 : \varphi = \varphi_0$ versus $H_0 : \varphi \neq \varphi_0$ can be performed using a LR test. In this case, the LR test statistic for testing H_0 versus H_1 is

$\omega = 2(\ell(\hat{\varphi};x) - \ell(\hat{\varphi}_0;x))$, where $\hat{\varphi}$ and $\hat{\varphi}_0$ are the MLEs under H_1 and H_0, respectively. The statistic ω is asymptotically (as $n \to \infty$) distributed as χ_k^2, where k is the length of the parameter vector φ of interest. The LR test rejects H_0 if $\omega > \chi_{k;\gamma}^2$, where $\chi_{k;\gamma}^2$ denotes the upper $100\gamma\%$ quantile of the χ_k^2 distribution.

6 Applications

In this section, we use two real data sets to show that the beta Lindley distribution can be a better model than one based on the Lindley distribution.

Data set 1: The data set given in Table 1 represents an uncensored data set corresponding to remission times (in months) of a random sample of 128 bladder cancer patients reported in Lee and Wang [12]:

Table 1: The remission times (in months) of bladder cancer patients

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	0.20	2.23
0.52	4.98	6.97	9.02	13.29	0.40	2.26	3.57	5.06	7.09
0.22	13.80	25.74	0.50	2.46	3.64	5.09	7.26	9.47	14.24
0.82	0.51	2.54	3.70	5.17	7.28	9.74	14.76	26.31	0.81
0.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64	3.88	5.32
0.39	10.34	14.83	34.26	0.90	2.69	4.18	5.34	7.59	10.66
0.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01
0.19	2.75	4.26	5.41	7.63	17.12	46.12	1.26	2.83	4.33
0.66	11.25	17.14	79.05	1.35	2.87	5.62	7.87	11.64	17.36
0.40	3.02	4.34	5.71	7.93	11.79	18.10	1.46	4.40	5.85
0.26	11.98	19.13	1.76	3.25	4.50	6.25	8.37	12.02	2.02
0.31	4.51	6.54	8.53	12.03	20.28	2.02	3.36	6.76	12.07
0.73	2.07	3.36	6.93	8.65	12.63	22.69	5.49		

Data set 2: The following data represent the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by Bjerkedal [2]. The data are as follows:

0.1, 0.33, 0.44, 0.56, 0.59, 0.72, 0.74, 0.77, 0.92, 0.93, 0.96, 1, 1, 1.02, 1.05, 1.07, 1.07, 1.08, 1.08, 1.08, 1.09, 1.12, 1.13, 1.15, 1.16, 1.2, 1.21, 1.22, 1.22, 1.24, 1.3, 1.34, 1.36, 1.39, 1.44, 1.46, 1.53, 1.59, 1.6, 1.63, 1.63, 1.68, 1.71, 1.72, 1.76, 1.83, 1.95, 1.96, 1.97, 2.02, 2.13, 2.15, 2.16, 2.22, 2.3, 2.31, 2.4, 2.45, 2.51, 2.53, 2.54, 2.54, 2.78, 2.93, 3.27, 3.42, 3.47, 3.61, 4.02, 4.32, 4.58, 5.55

Table 2: The ML estimates, standard error and Log-likelihood for data set 1

Model	ML Estimates	Standard Error	-LL
Lindley	$\hat{\theta} = 0.196$	0.012	**419.529**
NGLD	$\hat{\theta} = 0.18$	0.035	**412.750**
	$\hat{\alpha} = 4.679$	1.308	
	$\hat{\beta} = 1.324$	0.171	

The variance covariance matrix $I(\hat{\lambda})^{-1}$ of the MLEs under the new generalized Lindley distribution for data set 1 is computed as

$$\begin{pmatrix} 0.001 & 0.031 & 0.005 \\ 0.031 & 1.711 & 0.140 \\ 0.005 & 0.140 & 0.029 \end{pmatrix}$$

Thus, the variances of the MLE of θ, α and β is $var(\hat{\theta}) = 0.001, var(\hat{\alpha}) = 1.711$ and $var(\hat{\beta}) = 0.0295$. Therefore, 95% confidence intervals for θ, α and β are $[0.113, 0.252], [2.115, 7.243]$ and $[0.987, 1.661]$ respectively.

In order to compare the two distribution models, we consider criteria like -2ℓ, AIC (Akaike information criterion), AICC (corrected Akaike information criterion), BIC(Bayesian information criterion) and K-S(Kolmogorov-Smirnov test) for the data set. The better distribution corresponds to smaller -2ℓ, AIC and AICC values:‘

$$AIC = 2k - 2\ell, \quad AICC = AIC + \frac{2k(k+1)}{n-k-1},$$

$$BIC = k * log(n) - 2\ell \text{ and } K - S = \sup_x |F_n(x) - F(x)|$$

where $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{x_i \le x}$ is empirical distribution function, $F(x)$ is comulative distribution function, k is the number of parameters in the statistical model, n the sample size and ℓ is the maximized value of the log-likelihood function under the considered model.

Table 3: The AIC, AICC, BIC and K-S of the models based on data set 1

Model	-2LL	AIC	AICC	BIC	K-S
Lindley	839.04	841.06	841.091	843.892	0.074
NGLD	825.501	831.501	831.694	840.057	0.116

The LR test statistic to test the hypotheses $H_0 : \alpha = \beta = 1$ versus $H_1 : \alpha \neq 1 \vee \beta \neq 1$ for data set 1 is $\omega = 13.539 > 5.991 = \chi^2_{2;0.05}$, so we reject the null hypothesis.

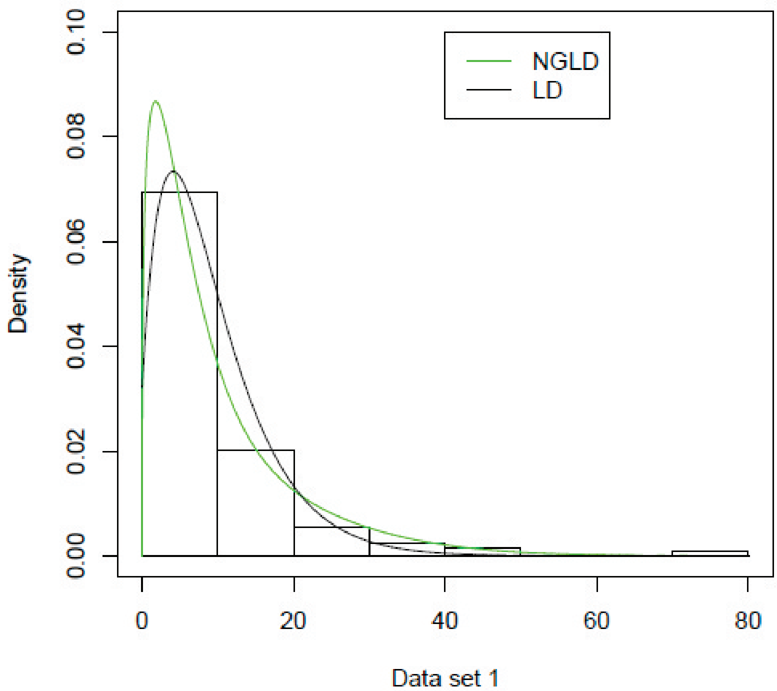

Figure 4: Estimated densities of the models for data set 1.

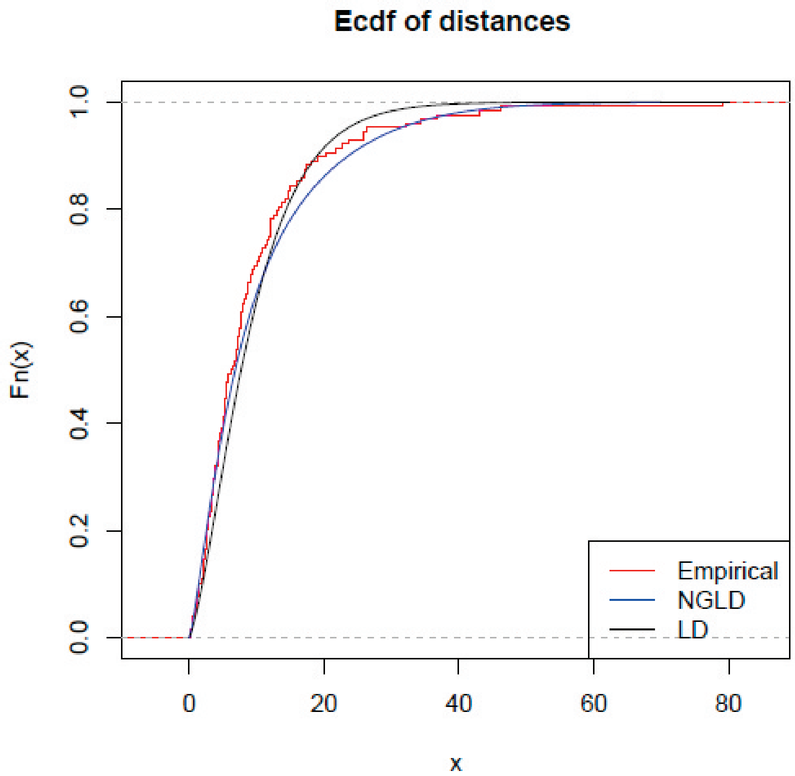

Figure 5: Estimated cumulative densities of the models for data set 1.

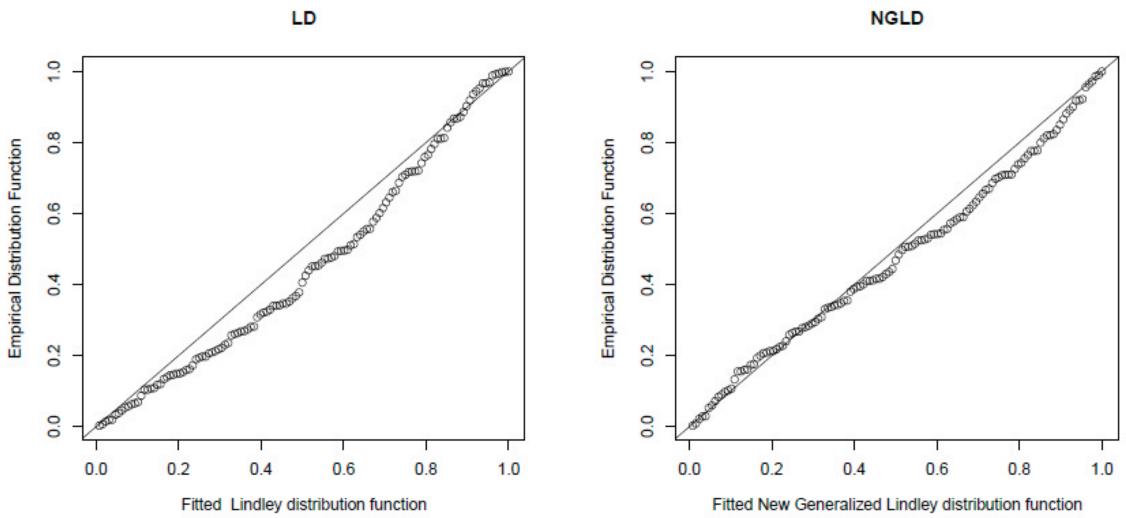

Figure 6: P-P plots for fitted Lindley and the NGLD for data set 1.

Table 4: The ML estimates, standard error and Log-likelihood for data set 2

Model	ML Estimates	St. Error	$-\ell$
Lindley	$\hat{\theta}=0.868$	0.076	106.928
NGLD	$\hat{\theta}=1.861$	0.489	94.182
	$\hat{\alpha}=3.585$	1.238	
	$\hat{\beta}=2.737$	0.554	

The variance covariance matrix $I(\hat{\lambda})^{-1}$ of the MLEs under the beta Lindley distribution for data set is computed as

$$\begin{pmatrix} 0.239 & 0.569 & -0.001 \\ 0.569 & 1.532 & -0.154 \\ -0.001 & -0.154 & 0.307 \end{pmatrix}$$

Thus, the variances of the MLE of θ,α and β is $var(\hat{\theta})=0.239, var(\hat{\alpha})=0.239$ and $var(\hat{\beta})=0.307.$ Therefore, 95% confidence intervals for θ,α and β are $[0.901,2.819],[1.158,6.011]$ and $[1.651,3.823]$ respectively.

Table 5: The AIC, AICC, BIC and K-S of the models based on data set 2

Model	-2ℓ	AIC	AICC	BIC	K-S
Lindley.	213.857	215.857	215.942	218.133	0.232
NGLD	188.364	194.364	194.722	201.194	0.075

The LR test statistic to test the hypotheses $H_0:\alpha=\beta=1$ versus $H_1:\alpha\neq 1\vee\beta\neq 1$ for data set 2 is $\omega=25.493>5.991=\chi^2_{2;0.05}$, so we reject the null hypothesis. Tables 2 and 4 shows parameter MLEs to each one of the two fitted distributions for data set 1 and 2, Tables 3 and 5 shows the values of $-2\log(L)$, AIC, AICC, BIC and K-S values. The values in Tables 3 and 5, indicate that the new generalized Lindley distribution is a strong competitor to other distribution used here for fitting data set 1 and data set 2. A density plot compares the fitted densities of the models with the empirical histogram of the observed data (Fig. 3 and 5). The fitted density for the new generalized Lindley model is closer to the empirical histogram than the fits of the Lindley model.

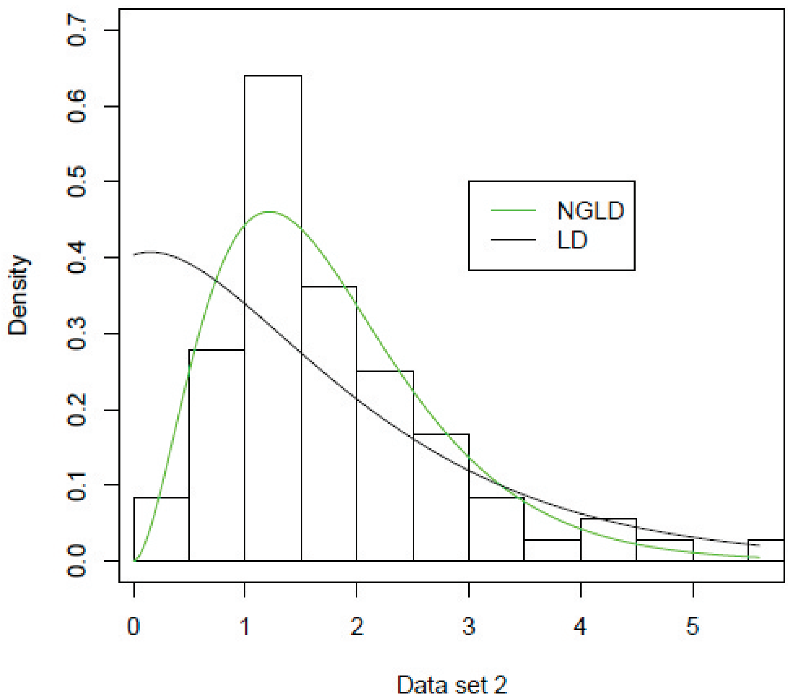

Figure 7: Estimated densities of the models for data set 2.

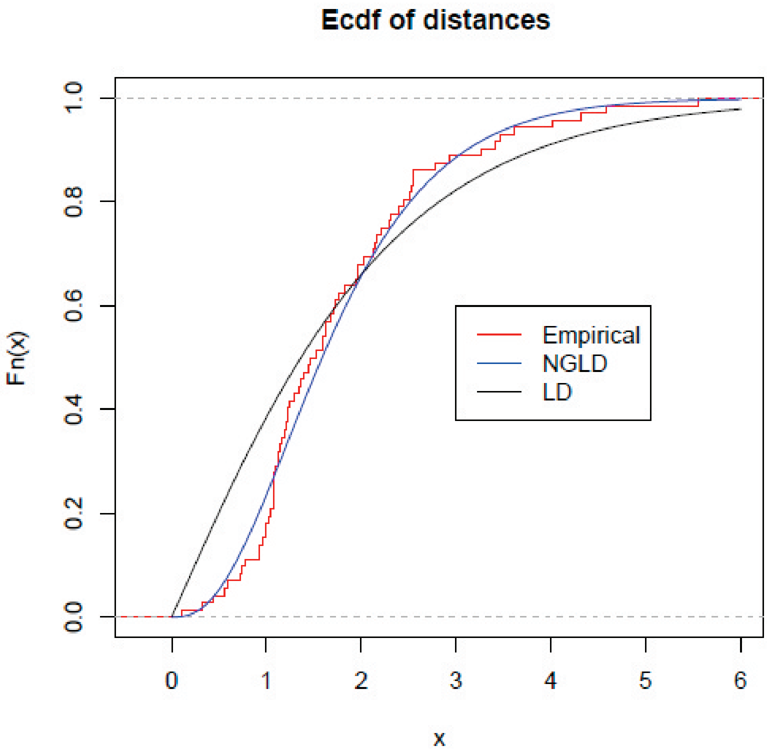

Figure 8: Estimated cumulative densities of the models for data set 2.

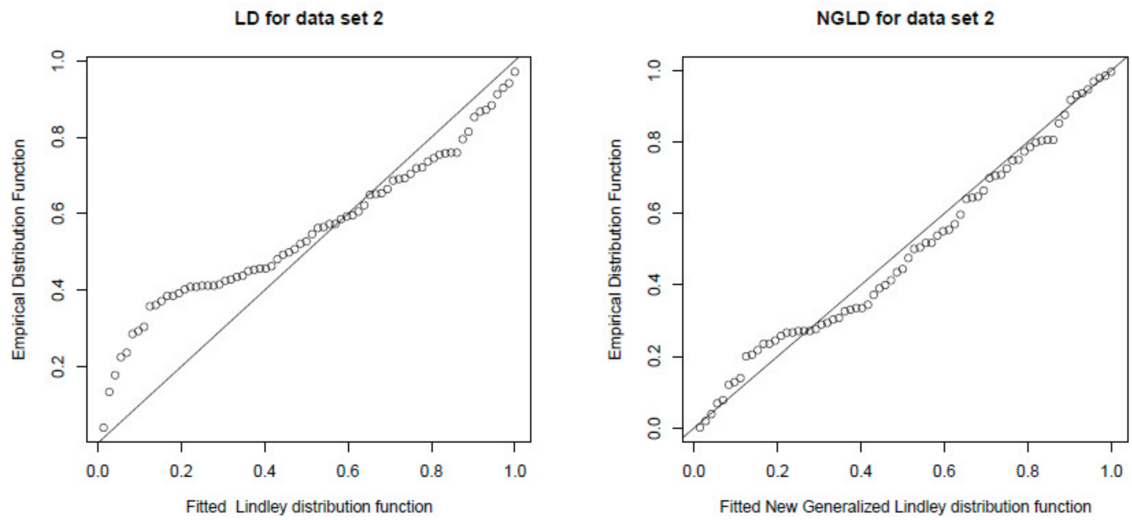

Figure 9: P-P plots for fitted Lindley and the NGLD for data set 2.

7 Conclusion

Here, we propose a new model, the so-called the new generalized Lindley distribution which extends the Lindley distribution in the analysis of data with real support. An obvious reason for generalizing a standard distribution is because the generalized form provides larger flexibility in modelling real data. We derive expansions for the moments and for the moment generating function. The estimation of parameters is approached by the method of maximum likelihood, also the information matrix is derived. We consider the likelihood ratio statistic to compare the model with its baseline model. Two applications of the new generalized Lindley distribution to real data show that the new distribution can be used quite effectively to provide better fits than the Lindley distribution.

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 Pies infected with Different Doses of Virulent Tubercle Bacilli. American Journal of Hygiene, 72(1), 130-48.

Deniz EG, Ojeda EC (2011). The discrete Lindley distribution Properties and Applications. J. Stat. Comput. Simul.81(11), 1405- 1416.

Frank Guess and Frank Proschan. Mean residual life: theory and applications. (1985). Technical Report 85-178, North Carolina State University and Florida State University, Tallahassee,Florida, June 1985. Ghitany ME, Atieh B, and Nadarajah S (2008a). Lindley distribution and its Applications. Math. Comput. Simul. 78(4):493-506.

Ghitany ME, and Al-Mutairi DK (2009). Estimation methods for the discrete Poisson- Lindley distribution. J. Stat. Comput. Simul. 79(1):1-9.

Ghitany ME, and 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, and 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, and Hussain H A (2011). A two parameter weighted Lindley distribution and its applications to survival data. Math. Comput. Simul. 81(6), 1190-1201.

Gupta, R.D., and Kundu. D (1999). Generalized exponential distributions. Australian and New Zealand Journal of Statistics. 41, 173 –188.

Johnson, N., Kotz, S. and Balakrishnan,N. (1995). Continuous Univariate Distribution, V(2), 2nd edition, New York, Wiley.

Lee E.T. and Wang J.W.,(2003).Statistical Methods for Survival Data Analysis, 3rd ed.,Wiley, NewYork.
Lindley DV (1958). Fiducial distributions and Bayes' theorem. J. Royal Stat. Soc. Series B, 20:102-107.

Mazucheli J, and Achcar J. A (2011). The Lindley distribution applied to competing risks lifetime data. Compmajtut. Methods Programs Biomed. 104(2):188-192.

Rama S and Mishra A (2013). A quasi Lindley distribution. African Journal of Mathematics and Computer Science Research. 6(4), 64-71.

Sankaran M (1970). The discrete Poisson-Lindley distribution. Biometrics pp. 145-149.
Swain, J , Venkatraman,S and Wilson, J (1988) Least squares estimation of distribution function in Johnson's translation system. Journal of Statistical Computation and Simulation, 29 , 271 - 297.

Zakerzadah H, Dolati A (2010). Generalized Lindley distribution. J. Math. Ext. 3(2):13-25.

Zenga M.,(2007). Inequality curve and inequality index based on the ratios between lower and upper arithmetic means. Statistica e Applicazioni 4: 3–27

Wall, H. S., (1948). Analytic Theory of Continued Fractions. New York: Chelsea.