THE LENGTH BIASED NEW QUASI LINDLEY DISTRIBUTION: STATISTICAL PROPERTIES AND APPLICATION

N. W. Andure (Yawale)[1] and R. B. Ade[2*]

•

[1] Department of Statistics, Government Vidarbha Institute of Science and Humanities, Amravati, Maharashtra, India, neetayawale@gmail.com

[2*]Department of Statistics, Government Vidarbha Institute of Science and Humanities, Amravati, Maharashtra, India, rajeshwarb.sc@gmail.com

Abstract

In this paper, a new distribution namely the length biased new quasi-Lindley distribution is proposed with the different weight function. The different mathematical and statistical properties of the proposed distribution are derived and discussed. The survival function, hazard rate function and mean residual life function for the length biased new quasi Lindley distribution is discussed. Also, concepts like stochastic ordering and entropy for proposed distribution are studied. The parameters of the proposed distribution are estimated by using the method of maximum likelihood estimation. The performance of the newly introduced distribution is studied using a real- life data set.

Keywords: Length Biased Distribution, New Quasi Lindley Distribution, Reliability Analysis, Stochastic ordering, Maximum Likelihood Estimation.

I. Introduction

The concept of weighted distributions was first introduced by [8] to model ascertainment bias, weighted distributions were later formalized in a unifying theory by [16]. Weighted distribution is used in a variety of research fields related to reliability, environment, engineering and biomedicine. The weighted distribution reduces to length-biased distribution when the weight function considers only the length of the units. The concept of length-biased sampling was first introduced by [5] and [24]. [14] studied size-biased sampling and related form-invariant weighted distributions. Refer [15] for a general statistical discussion of weighted distributions. [12] proposed a useful result by giving a relationship between the original random variable X and its length-biased form Y when X is either Inverse Gaussian or Gamma distribution. Several researchers have studied length biased versions of different distributions see, [10], [6], [7], [17], [13] and [19].

Initially Quasi Lindley distribution was proposed by [22]. Later New Quasi Lindley (NQL) distribution was studied by [21] for modelling various data sets with probability density function (p.d.f) as follows

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

The NQL distribution in (1) is a mixture of exponential and gamma distribution [exponential (θ) and gamma (2, θ)].

In the present work, length biased new Quasi Lindley distribution is proposed and discussed in next section.

II. Length Biased New Quasi Lindley Distribution

Suppose X be a non-negative random variable with pdf $f(x)$. Let $w(x)$ be the non-negative weight function, and then the pdf of the weighted random variable X_w is given by:

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))} \quad , \quad x > 0$$

where $w(x)$ be a non - negative weight function and

$$E\big(w(x)\big) = \int w(x)f(x)dx < \infty$$

When $w(x) = x^c$,the resulting distribution is termed as weighted distribution. When $w(x) = x$ the resultant is known as size or length biased distribution. In this paper, the length biased version of new quasi-Lindley distribution is proposed. The weight function used is as follows

$$w(x) = \frac{nx}{\theta} \tag{2}$$

According to [1], let X is a non-negative random variable with pdf $f(x)$. Let $w(x)$ be the non-negative weight function then the pdf $f_l(x)$ for a length biased distribution of X is given by:

$$f_l(x) = \frac{w_*(x)w(x)f(x)}{E(w_*(x)w(x))} \quad , \quad x > 0 \tag{3}$$

Assuming the $E\big(w(x)w_*(x)\big) = \int w(x)w_*(x)f(x)dx < \infty$

Provided that $w_*(x) = x$

Using equation (1) and (2) in (3), the pdf of length biased new quasi-Lindley (LBNQL) distribution is

$$f_l(x;\theta,\alpha) = \frac{\frac{nx}{\theta}\frac{x\theta^2}{\theta^2+\alpha}(\theta+\alpha x)e^{-\theta x}}{E\big(w(x)w_*(x)\big)}$$

where

$$E\big(w(x)w_*(x)\big) = \int_0^\infty \frac{nx}{\theta}\frac{x\theta^2}{\theta^2+\alpha}(\theta+\alpha x)e^{-\theta x}dx = \frac{2n(\theta^2+3\alpha)}{\theta^3(\theta^2+\alpha)}$$

$$f_l(x;\theta,\alpha) = \frac{\theta^4}{2(\theta^2+3\alpha)}x^2(\theta+\alpha x)e^{-\theta x} \quad ;x > 0, \theta > 0, \alpha > 0 \tag{4}$$

and the cumulative distribution function (CDF) of LBNQL distribution is obtained as

$$F(x) = 1 - Pr(X > x) = 1 - \int_x^\infty f_l(t;\theta,\alpha)dt$$

$$F(x) = 1 - \frac{\theta^4}{2(\theta^2+3\alpha)}\int_x^\infty t^2(\theta+\alpha t)e^{-\theta t}dt$$

$$F(x) = 1 - \frac{\theta^4}{2(\theta^2+3\alpha)}\left(\theta\int_x^\infty t^2\, e^{-\theta t}dt + \alpha\int_x^\infty t^3\, e^{-\theta t}dt\right)$$

after the simplification, the CDF of the LBNQL distribution is

$$F(x) = 1 - \left(1 + \frac{\alpha\theta^3x^3 + x(6\alpha\theta+2\theta^3) + (3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x} \tag{5}$$

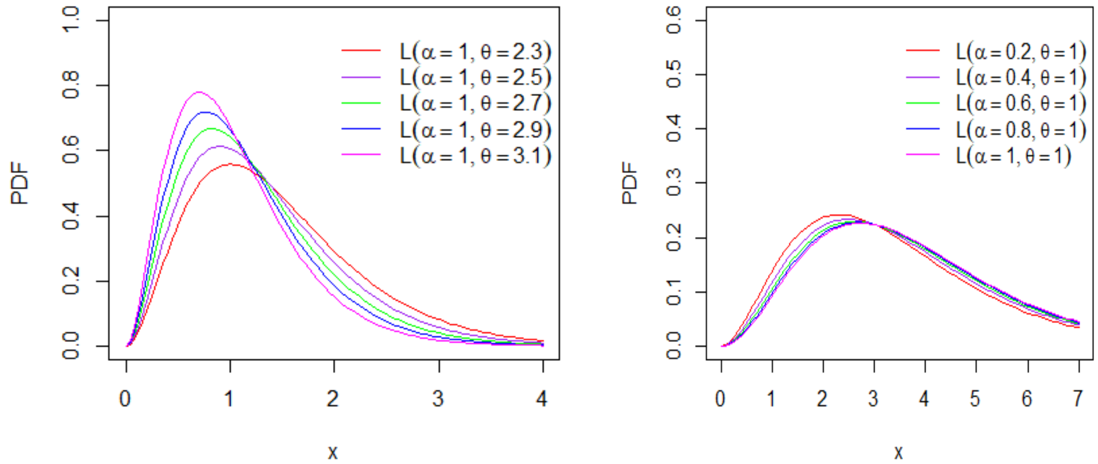

Figure 1: *PDF of LBNQL distribution for different values of θ and α.*

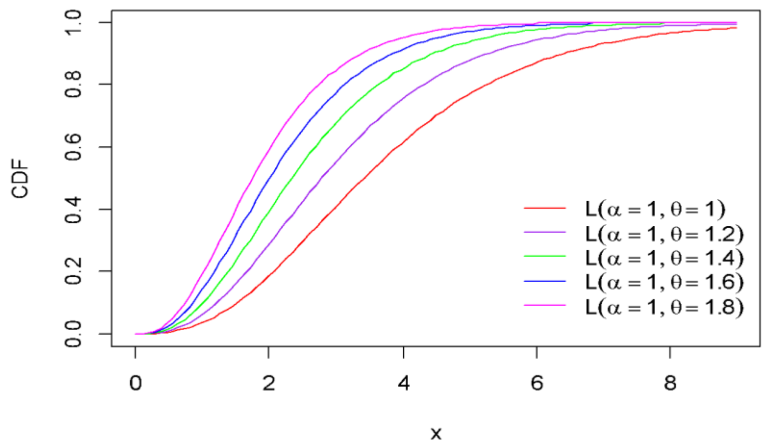

Figure 2: *CDF of LBNQL distribution for different values of θ and α.*

III. Reliability Analysis

In this section, the reliability function or survival function, hazard rate function and mean residual life for the LBNQL distribution is discussed.

I. Survival Function of LBNQL Distribution

The survival function or the reliability function of Length biased new quasi-Lindley distribution (LBNQLD) is defined as

$$S(x) = 1 - F(x) \tag{6}$$

Substituting from equation (5) in (6),

$$S(x) = \left(1 + \frac{\alpha\theta^3 x^3 + x(6\alpha\theta + 2\theta^3) + (3\alpha\theta^2 + \theta^4)x^2}{2(\theta^2 + 3\alpha)}\right) e^{-\theta x} \tag{7}$$

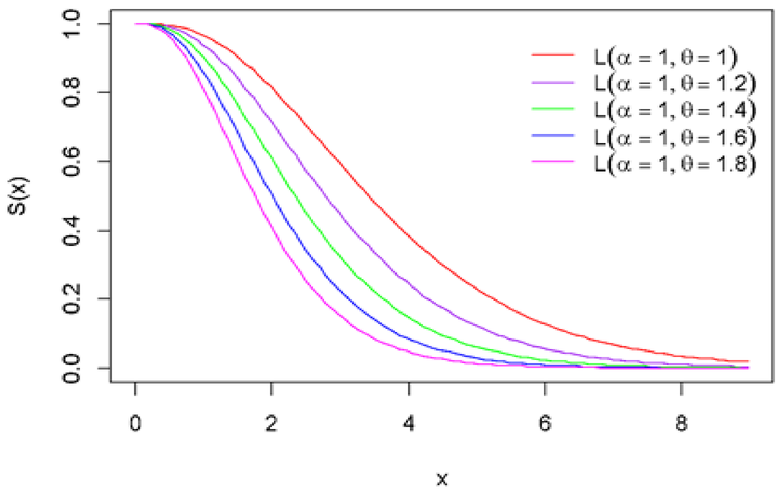

Figure 3: *survival function of LBNQL distribution for different values of θ and α.*

II. Hazard Rate Function of LBNQL Distribution

The basic tool for studying the aging and reliability characteristics of the system is the hazard rate (HR). The hazard function is also known as the hazard rate. Thus, the hazard rate function of the LBNQL distribution is given by

$$h(x) = \frac{f_l(x;\theta,\alpha)}{S(x)} \tag{8}$$

Substitute the value of (4) and (7) in (8),

$$h(x) = \frac{x^2\theta^4(\theta + \alpha x)}{2(\theta^2 + 3\alpha) + \alpha\theta^3 x^3 + (6\alpha\theta + 2\theta^3)x + (3\alpha\theta^2 + \theta^4)x^2}$$

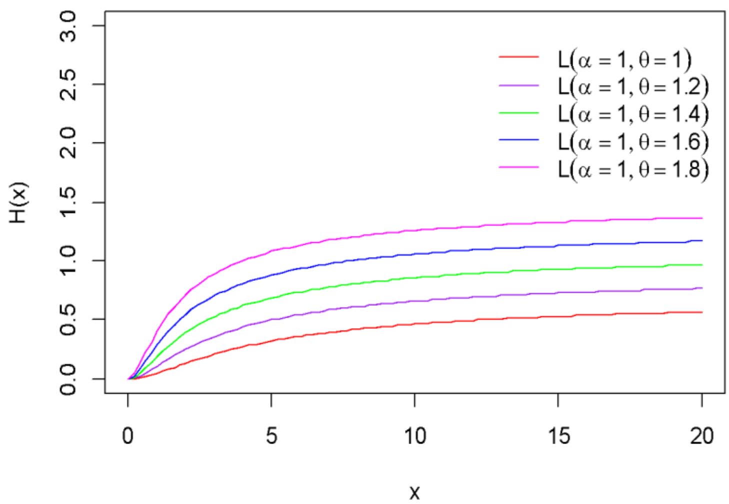

Figure.4: *Hazard rate function of LBNQL distribution for different values of θ and α.*

Figure (4) shows the behavior of hazard function. For different choices of α and θ it shows increasing failure rate.

III. Mean Residual Life Function of LBNQL Distribution

The mean residual life function is defined as

$$m(x) = E[X > x] = \frac{1}{1-F(x)}\int_x^\infty [1 - F(t)]\, dt$$

$$= \frac{1}{\left(1 + \frac{\alpha\theta^3x^3 + x(6\alpha\theta + 2\theta^3) + (3\alpha\theta^2 + \theta^4)x^2}{2(\theta^2 + 3\alpha)}\right)e^{-\theta x}}$$

$$\times \int_x^\infty \left(1 + \frac{\alpha\theta^3t^3 + t(6\alpha\theta + 2\theta^3) + (3\alpha\theta^2 + \theta^4)t^2}{2(\theta^2 + 3\alpha)}\right)e^{-\theta t} dt$$

After simplifying, we get

$$m(x) = \frac{24\alpha + 6\theta^2 + \alpha\theta^3x^3 + x(18\alpha\theta + 4\theta^3) + (6\alpha\theta^2 + \theta^4)x^2}{\theta(6\alpha + 2\theta^2 + \alpha\theta^3x^3 + (3\alpha\theta^2 + \theta^4)x^2 + x(6\alpha\theta + 2\theta^3))}$$

It can be easily verified that $m(0) = \frac{3(\theta^2+4\alpha)}{\theta(\theta^2+3\alpha)} = \mu_1'$

IV. Moments and Associated Measures

Let X denotes the random variable of LBNQL distribution with parameters α and θ then the r^{th} order moment of LBNQL distribution can be defined as

$$E(X^r) = \mu_r' = \int_0^\infty x^r f_l(x;\theta,\alpha)\, dx$$

$$= \int_0^\infty x^{r+2} \frac{\theta^4}{2(\theta^2+3\alpha)}(\theta + \alpha x)e^{-\theta x}dx$$

$$= \frac{\theta^4}{2(\theta^2+3\alpha)}\int_0^\infty x^{r+2}(\theta + \alpha x)e^{-\theta x}dx$$

$$= \frac{\theta^4}{2(\theta^2+3\alpha)}\left(\theta\int_0^\infty x^{r+3-1} e^{-\theta x}dx + \alpha\int_0^\infty x^{r+4-1} e^{-\theta x}dx\right)$$

$$\mu_r' = \frac{\theta^4}{2(\theta^2+3\alpha)}\left(\frac{\Gamma(r+3)}{\theta^{r+2}} + \frac{\alpha\Gamma(r+4)}{\theta^{r+4}}\right) \tag{9}$$

Putting $r = 1$ in (9), the mean of LBNQL distribution is given by

$$\mu_1' = E(X) = \frac{3(\theta^2 + 4\alpha)}{\theta(\theta^2 + 3\alpha)}$$

and putting $r = 2,3,4$ in (9), the second, third and fourth raw moments are

$$\mu_2' = E(X^2) = \frac{12(\theta^2+5\alpha)}{\theta^2(\theta^2+3\alpha)}, \quad \mu_3' = E(X^3) = \frac{60(\theta^2+6\alpha)}{\theta^3(\theta^2+3\alpha)}, \quad \mu_4' = E(X^4) = \frac{360(\theta^2+7\alpha)}{\theta^4(\theta^2+3\alpha)}$$

Therefore,

$$Variance = \sigma^2 = \frac{3(12\alpha^2+8\alpha\theta^2+\theta^4)}{\theta^2(\theta^2+3\alpha)^2}$$

$$Standard\ Deviation = \sigma = \frac{\sqrt{3(12\alpha^2 + 8\alpha\theta^2 + \theta^4)}}{\theta(\theta^2 + 3\alpha)}$$

$$Coefficient\ of\ Variation(c.v) = \frac{\sigma}{\mu} = \frac{\sqrt{3(12\alpha^2 + 8\alpha\theta^2 + \theta^4)}}{3(\theta^2 + 4\alpha)}$$

$$Coefficient\ of\ Dispersion(\gamma) = \frac{\sigma^2}{\mu} = \frac{(12\alpha^2 + 8\alpha\theta^2 + \theta^4)}{\theta(\theta^2 + 3\alpha)(\theta^2 + 4\alpha)}$$

I. Moment Generating Function and Characteristic Function

Let the random variable X follows the LBNQL distribution. By definition of moment generating function of X and using equation (4), we get

$$M_X(t) = E(e^{tx}) = \int_0^\infty e^{tx}\ f_l(x;\theta,\alpha)dx$$

$$= \int_0^\infty (1 + (tx) + \tfrac{(tx)^2}{2!} + \tfrac{(tx)^3}{3!} + \cdots) f_l(x;\theta,\alpha)dx$$

$$= \int_0^\infty \sum_{r=0}^\infty \tfrac{(tx)^r}{r!} f_l(x;\theta,\alpha)dx$$

$$= \sum_{r=0}^\infty \tfrac{(t)^r}{r!} \int_0^\infty x^r\ f_l(x;\theta,\alpha)dx$$

$$= \sum_{r=0}^\infty \tfrac{(t)^r}{r!} E(X^r) \tag{10}$$

Substituting from equation (9) in (10),

$$M_X(t) = \sum_{r=0}^\infty \frac{(t)^r}{r!}\left\{\frac{\theta^4}{2(\theta^2+3\alpha)}\left(\frac{\Gamma(r+3)}{\theta^{r+2}} + \frac{\alpha\Gamma(r+4)}{\theta^{r+4}}\right)\right\}$$

Similarly, the characteristic function of LBNQL distribution is obtained as

$$\phi_X(t) = M_X(it)$$

$$\Rightarrow \phi_x(t) = M_X(it) = \sum_{r=0}^\infty \frac{(it)^r}{r!}\left\{\frac{\theta^4}{2(\theta^2+3\alpha)}\left(\frac{\Gamma(r+3)}{\theta^{r+2}} + \frac{\alpha\Gamma(r+4)}{\theta^{r+4}}\right)\right\}$$

V. Entropy

The concept of entropy is important in a variety of topics such as probability and mathematics, physics, communication theory and economics. The entropy of random variable X is a measure of the variability of uncertainty.

I. R´enyi Entropy

R´enyi entropy [18] is important in nature and mathematics as an indicator of diversity. R´enyi entropy is also important in quantum data, where it can be used as a catch measure. R´enyi entropy is provided by

$$Re(\delta) = \frac{1}{1-\delta} log\left(\int_0^\infty f_l{}^\delta(x;\theta,\alpha)\, dx\right)$$

where $\delta > 0$ and $\delta \neq 1$

$$Re(\delta) = \frac{1}{1-\delta} log\left(\int_0^\infty \left(\frac{\theta^4}{2(\theta^2+3\alpha)} x^2(\theta+\alpha x)e^{-\theta x}\right)^\delta dx\right)$$

$$Re(\delta) = \frac{1}{1-\delta} log\left(\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^\delta \int_0^\infty x^{2\delta}\left(1+\frac{\alpha}{\theta}x\right)^\delta e^{-\theta\delta x}dx\right)$$

putting $\left(1+\frac{\alpha}{\theta}x\right)^\delta = \sum_{j=0}^\infty \binom{\delta}{j}\left(\frac{\alpha}{\theta}x\right)^j$

$$Re(\delta) = \frac{1}{1-\delta} log\left\{\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^\delta \sum_{j=0}^\infty \binom{\delta}{j}\left(\frac{\alpha}{\theta}\right)^j \int_0^\infty (x)^{2\delta+j+1-1}\, e^{-\theta\delta x}dx\right\}$$

$$Re(\delta) = \frac{1}{1-\delta} log\left\{\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^\delta \sum_{j=0}^{\infty}\binom{\delta}{j}\left(\frac{\alpha}{\theta}\right)^j \frac{\Gamma(2\delta+j+1)}{(\theta\delta)^{2\delta+j+1}}\right\}$$

II. Tsallis Entropy

A generalization of Boltzmann-Gibbs (B.G) statistical properties initiated by Tsallis has received great attention. This generalization of BG statistic was proposed firstly by introducing the mathematical expression of Tsallis entropy by [23] for a continuous random variable and is defined as follows

$$S_\lambda = \frac{1}{1-\lambda}\left(1-\int_0^\infty f_l^{\lambda}(x;\theta,\alpha)dx\right)$$

$$S_\lambda = \frac{1}{1-\lambda}\left(1-\int_0^\infty \left(\frac{\theta^4}{2(\theta^2+3\alpha)}x^2(\theta+\alpha x)e^{-\theta x}\right)^\lambda dx\right)$$

$$S_\lambda = \frac{1}{1-\lambda}\left(1-\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^\lambda \int_0^\infty x^{2\lambda}\left(1+\frac{\alpha}{\theta}x\right)^\lambda e^{-\lambda\theta x}dx\right)$$

putting $\left(1+\frac{\alpha}{\theta}x\right)^\lambda = \sum_{j=0}^{\infty}\binom{\lambda}{j}\left(\frac{\alpha}{\theta}x\right)^j$

$$S_\lambda = \frac{1}{1-\lambda}\left\{1-\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^\lambda \sum_{j=0}^{\infty}\binom{\lambda}{j}\left(\frac{\alpha}{\theta}\right)^j \int_0^\infty (x)^{2\lambda+j}\, e^{-\theta\lambda x}dx\right\}$$

$$S_\lambda = \frac{1}{1-\lambda}\left\{1-\left(\frac{\theta^5}{2(\theta^2+3\alpha)}\right)^\delta \sum_{j=0}^{\infty}\binom{\lambda}{j}\left(\frac{\alpha}{\theta}\right)^j \frac{\Gamma(2\lambda+j+1)}{(\theta\lambda)^{2\lambda+j+1}}\right\}$$

VI. Order Statistic of LBNQL Distribution

Order statistic have a central role in statistical theory. Suppose $X_{(1)}, X_{(2)}, \dots\dots\dots\dots, X_{(n)}$ be the continuous ascending order statistic. The probability density function of the j^{th} order statistic $X_{(j)}$ for $1 \le j \le n$ is

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!}[F(x)]^{j-1}[1-F(x)]^{n-j}f_1(x) \tag{11}$$

Substitute the value of pdf and cdf of LBNQL distribution in (11), we get

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!}\left[1-\left(1+\frac{\alpha\theta^3x^3+x(6\alpha\theta+2\theta^3)+(3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x}\right]^{j-1}$$
$$\times\left[\left(1+\frac{\alpha\theta^3x^3+x(6\alpha\theta+2\theta^3)+(3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x}\right]^{n-j} \times \frac{\theta^4}{2(\theta^2+3\alpha)}x^2(\theta+\alpha x)e^{-\theta x} \tag{12}$$

Put $j = 1$ in equation (12), the probability density function of first order statistics of LBNQL distribution.

$$f_{X_{(1)}}(x) = n\left[\left(1+\frac{\alpha\theta^3x^3+x(6\alpha\theta+2\theta^3)+(3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x}\right]^{n-1} \times \frac{\theta^4}{2(\theta^2+3\alpha)}x^2(\theta+\alpha x)e^{-\theta x}$$

Put $j = n$ in equation (12), the probability density function of n^{th} order statistics of LBNQLD.

$$f_{X_{(n)}}(x) = n\left[1-\left(1+\frac{\alpha\theta^3x^3+x(6\alpha\theta+2\theta^3)+(3\alpha\theta^2+\theta^4)x^2}{2(\theta^2+3\alpha)}\right)e^{-\theta x}\right]^{n-1}$$

$$\times \frac{\theta^4}{2(\theta^2+3\alpha)} x^2(\theta+\alpha x)e^{-\theta x}$$

VII. Stochastic ordering

Stochastic ordering of positive continuous random variables is an important tool for judging their comparative behavior. A random variable X is said to be smaller than a random variable Y in the
1) Stochastic order $(X \le_{st} Y)$ if $F_X(x) \ge F_Y(x)$ for all x
2) Hazard rate order $(X \le_{hr} Y)$ if $h_X(x) \ge h_Y(x)$ for all x
3) Mean residual life order $(X \le_{mrl} Y)$ if if $m_X(x) \le m_Y(x)$ for all x
4) Likelihood ratio order $(X \le_{lr} Y)$ if $\frac{f_X(x)}{f_Y(x)}$ decreases in x

The following important interrelations due to [20] are well-known for establishing stochastic ordering of distributions

$$X \le_{lr} Y \Rightarrow X \le_{hr} Y \Rightarrow X \le_{mrl} Y$$
$$\Downarrow$$
$$X \le_{st} Y$$

The LBNQL distribution is ordered with respect to the strongest 'likelihood ratio ordering' as shown in the following theorem:

Theorem 7.1: Let $X \sim LBNQL(\theta_1,\alpha_1)$ and $Y \sim LBNQL(\theta_2,\alpha_2)$.If $\theta_1=\theta_2$ and $\alpha_1 \le \alpha_2$ or ($\alpha_1=\alpha_2$ 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: From the pdf of LBNQL distribution (4), we have

$$\frac{f_X(x)}{f_Y(x)} = \frac{{\theta_1}^4({\theta_2}^2+3\alpha_2)}{{\theta_2}^4({\theta_1}^2+3\alpha_1)}\left(\frac{\theta_1+\alpha_1 x}{\theta_2+\alpha_2 x}\right)e^{-(\theta_1-\theta_2)x} \qquad ;x>0$$

Now

$$\log\frac{f_X(x)}{f_Y(x)} = \log\left[\frac{{\theta_1}^4({\theta_2}^2+3\alpha_2)}{{\theta_2}^4({\theta_1}^2+3\alpha_1)}\right] + \log(\theta_1+\alpha_1 x) - \log(\theta_2+\alpha_2 x) - (\theta_1-\theta_2)x$$

This gives

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

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

Thus, LBNQL distribution follows the strongest likelihood ratio ordering.

VIII. Bonferroni and Lorenz curve

The most important inequality curves are called Bonferroni and Lorenz curve, which have some application in applied science such as economics, reliability, demography and medicine. Bonferroni and Lorenz curves are proposed by [3]. The Bonferroni and Lorentz curves for the LBNQL distribution is obtained as

$$B(p) = \frac{1}{p\mu}\int_0^q x\, f_l(x,\theta,\alpha)dx$$

and

$$L(p) = pB(p) = \frac{1}{\mu}\int_0^q x\, f_l(x;\theta,\alpha)dx$$

where $E(x) = \mu = \frac{3(\theta^2+4\alpha)}{\theta(\theta^2+3\alpha)}$ and $q = F^{-1}(p)$

$$\therefore\ B(p) = \frac{\theta(\theta^2+3\alpha)}{p3(\theta^2+4\alpha)}\int_0^q x^3 \frac{\theta^4}{2(\theta^2+3\alpha)}(\theta+\alpha x)e^{-\theta x}dx$$

$$B(p) = \frac{\theta^5}{p6(\theta^2+4\alpha)}\int_0^q x^3\,(\theta+\alpha x)e^{-\theta x}dx$$

after simplification,

$$B(p) = \frac{\theta^2\gamma(4,\theta q) + \alpha\gamma(5,\theta q)}{p6(\theta^2+4\alpha)}$$

and

$$L(p) = pB(p) = \frac{1}{6(\theta^2+4\alpha)}(\theta^2\gamma(4,\theta q) + \alpha\gamma(5,\theta q))$$

IX. Maximum Likelihood Estimation

The method of maximum likelihood is the most frequently used method of parameter estimation given in [4]. The maximum likelihood method of estimation has been adopted to estimate the unknown parameter α and θ of the LBNQL distribution. Consider the random sample of size n from the LBNQL distribution, the likelihood function is given by

$$L(x;\alpha,\theta) = \left(\frac{\theta^4}{2(\theta^2+3\alpha)}\right)^n \prod_{i=1}^{n} {x_i}^2\,(\theta+\alpha x_i)e^{-\theta\sum_{i=1}^n x_i}$$

The log likelihood function is given by

$$\log l = 4n\log\theta - n\log\big(2(\theta^2+3\alpha)\big) + 2\sum_{i=1}^n \log x_i + \sum_{i=1}^n \log(\theta+\alpha x_i) - \theta\sum_{i=1}^n x_i \qquad (13)$$

Now maximize the above log-likelihood function given in equation (13) to get maximum likelihood estimate of unknown parameters of length biased new quasi-Lindley distribution. For this purpose, take the first derivative of the above log-likelihood equation with respect to parameters α and θ and equate to zero respectively.

$$\Rightarrow \frac{4n}{\theta} - \frac{2n\theta}{(\theta^2+3\alpha)} + \sum_{i=1}^n \frac{1}{\theta+\alpha x_i} - \sum_{i=1}^n x_i = 0 \qquad (14)$$

$$\Rightarrow \frac{-3n}{(\theta^2+3\alpha)} + \sum_{i=1}^n \frac{x_i}{\theta+\alpha x_i} = 0 \qquad (15)$$

Equations (14) and (15) are nonlinear equation. The exact solution of above equation is not possible numerically. Above nonlinear equations are solved with the help of R Software.

Using the large sample property of MLE, $\hat{\lambda}$ can be treated as being approximately normal with mean λ and variance covariance matrix equal to the inverse of the expected information matrix, i.e.,

$$\sqrt{n}(\hat{\lambda} - \lambda) \to N(0, I^{-1}(\lambda))$$

$I(\lambda)$ is the information matrix then its inverse matrix is $I^{-1}(\lambda)$. The $I(\hat{\lambda})$ variance-covariance matrix is essentially equal to the inverse of the expected information matrix $I^{-1}(\hat{\lambda})$, the observed information matrix is given by

$$I(\lambda) = -\frac{1}{n}\begin{bmatrix} E\left(\frac{\partial^2 \log l}{\partial \theta^2}\right) & E\left(\frac{\partial^2 \log l}{\partial \theta \partial \alpha}\right) \\ E\left(\frac{\partial^2 \log l}{\partial \alpha \partial \theta}\right) & E\left(\frac{\partial^2 \log l}{\partial \alpha^2}\right) \end{bmatrix}$$

where $I(\lambda)$ is Fisher's Information Matrix.

$$\frac{\partial^2 \log l}{\partial \theta^2} = \frac{-4n}{\theta^2} + \frac{2n(\theta^2 - 3\alpha)}{(\theta^2 + 3\alpha)^2} - \sum_{i=1}^{n} \frac{1}{(\theta + \alpha x_i)^2}$$

$$\frac{\partial^2 \log l}{\partial \alpha^2} = \frac{9n}{(\theta^2 + 3\alpha)^2} - \sum_{i=1}^{n} \left(\frac{x_i^2}{(\theta + \alpha x_i)^2}\right)$$

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

Since λ being unknown, $I^{-1}(\lambda)$ is estimated by using $I^{-1}(\hat{\lambda})$ and we obtain the asymptotic confidence intervals for α and θ. Hence the approximate $100(1 - \psi)\%$ confidence interval for α and θ are respectively given by

$$\hat{\alpha} \pm z_{\frac{\psi}{2}} \sqrt{I_{\alpha\alpha}{}^{-1}(\hat{\lambda})}, \qquad \hat{\theta} \pm z_{\frac{\psi}{2}} \sqrt{I_{\theta\theta}{}^{-1}(\hat{\lambda})}$$

Where $z_{\frac{\psi}{2}}$ is the ψ^{th} percentile of the standard distribution.

X. Application

In this section, one real life data set is analyzed for the purpose of illustration to show the usefulness and flexibility of the LBNQL distribution. The LBNQL model is compared with other distributions, such as, New Quasi Lindley (NQL) distribution, [21], length biased weighted New Quasi Lindley (LBWNQL) distribution [9]. The ML Estimates of the unknown parameters are determined for the LBNQL distribution and two other models along with goodness of fit test.

Data set I: Following data depicts the fatigue life of some aluminum's coupons cut in specific manner (see,[2]). The dataset (after subtracting 65) is given below:

5, 25, 31, 32, 34, 35, 38, 39, 39, 40, 42, 43, 43, 43, 44, 44, 47, 47, 48, 49, 49, 49 ,51, 54, 55, 55 ,56 ,56, 56, 58, 59, 59, 59, 59, 59, 63, 63, 64 ,64, 55, 65, 65, 65, 66, 66, 66, 66, 67, 67, 67, 68, 69, 69, 69 ,69, 71, 71, 72, 73, 73, 73, 74, 74, 76, 76, 77, 77, 77, 77, 77 ,77, 79, 79, 80, 81, 83, 83, 84, 86, 86, 87, 90, 91, 92, 92, 92 ,92, 93, 93, 94, 97, 98, 98, 99, 101, 101, 103, 105, 109, 139, 147

The data set is modeled by LBNQL distribution and compared with the New Quasi Lindley, length biased weighted New Quasi Lindley distribution. Table 1 describes estimated unknown parameters, -log likelihood (-LL), the values of the AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and K-S Statistics calculated for above data using LBNQL, LBWNQL, and NQL distributions.

Table 1. *Estimate and goodness of fit measures under considered distribution based on data set.*

Model	Estimated Parameter		-2LL	AIC	BIC	K-S	P-Value
	$\hat{\alpha}$	$\hat{\theta}$					
LBNQLD	**74.01668**	**0.05801**	**945.7779**	**949.7779**	**955.0082**	**0.14788**	**0.2403**
LBWNQLD	425.13398	0.04350	962.6170	966.6170	971.8473	0.18094	0.1436
NQL	168.18633	0.02900	992.7213	996.7213	1001.9516	0.23557	0.05719

From table 1 it can be seen that the value of the statistics -2LL, AIC, BIC values of LBNQL distribution are comparatively smaller than the other distributions on a data set. Therefore, the result shows that LBNQL distribution provides a significantly better fit than the other models.

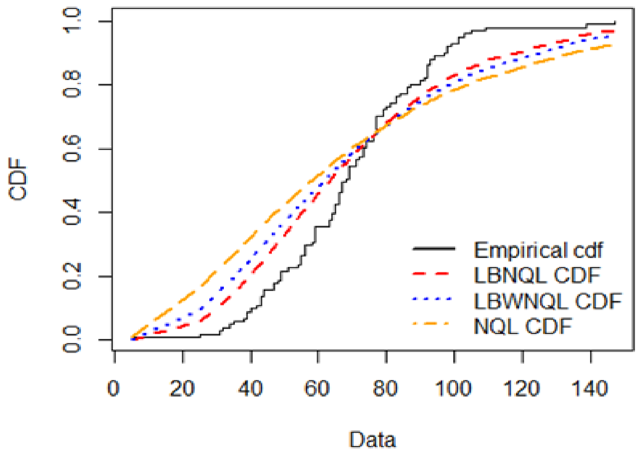

Figure 5: *Empirical CDF and fitted CDF plot of data set*.

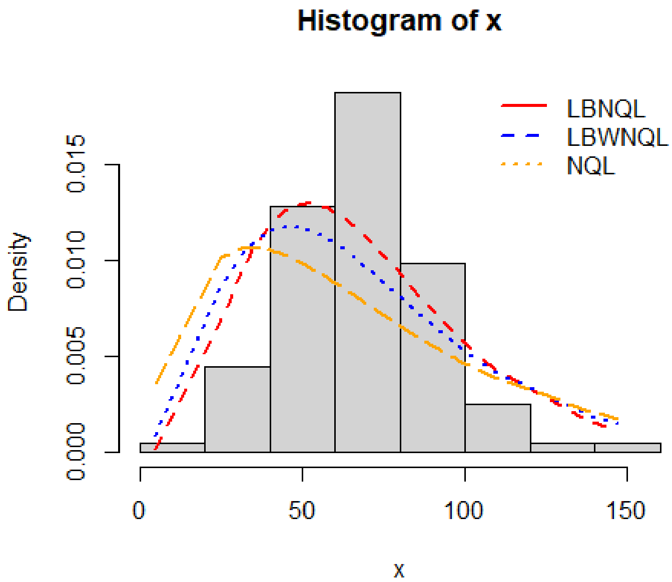

Figure 6: *fitted PDF plot of data set*.

XI. Conclusion

In the present study, a Length biased NQL distribution is proposed. Some statistical properties along with Reneyi entropy, and Tsallis entropy, Bonferroni and Lorenz curves have been discussed. Various reliability properties such as hazard rate function, mean residual life function, stochastic orderings have been obtained. It is proved that LBNQL distribution follows the strongest likelihood ratio ordering. For different choices of the parameters α and θ increasing failure rate is observed. The parameters of the proposed distribution are obtained by using the maximum likelihood estimation technique. Finally, the new proposed distribution is tested by applying to a real-life data set and compared with new quasi Lindley distribution and length biased weighted new quasi Lindley distribution. It is observed from the table 1 that LBNQL distribution gives better fit over both distributions on a data set.

References

[1] Al-kadim, K. A. and Hussain, N. A. (2014). New Proposed Length-Biased Weighted Exponential and Rayleigh Distribution with Application. Mathematical Theory and Modelling.4(7),137-152

[2] Birnbaum, Z. W. and Saunders, S. C. (1969). Estimation for a Family of Life Distributions with Applications to Fatigue. Journal of Applied Probability, 6(2), 328–347.

[3] Bonferroni, CE. Elementi di statisca gnerarale, Seeber, Firenze (1930).

[4] Casela, G. and Berger, R. L. Statistical inference. Pacific Grove, Calif: Brooks/Cole Pub. (1990).

[5] Cox, D. R. Some sampling problems in technology, In New Development in Survey Sampling, Johnson, N. L. and Smith, H., Jr.(eds.) New York, Wiley- Interscience, 506-527, (1969).

[6] Das, K. K. and Roy, T. D. (2011). Applicability of length biased generalized Rayleigh distribution. Advances in Applied Science Research, 2, 320–327

[7] Das, K. K. and Roy, T. D. (2011a). On some length-biased weighted Weibull distribution. Advances in Applied Science Research, 2, 465–475.

[8] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, Ann. Eugenics, 6, 13-25.

[9] Ganaie, R. A. and Rajgopalan, V. (2021). Length biased Weighted New Quasi Lindley Distribution: Statistical Properties and Applications. Pakistan Journal of Statistics and Operation Research. 17(1), 123-136.

[10] Gove, J. (2003). Moment and maximum likelihood estimators for Weibull distributions under length- biased and area-biased sampling. Environmental and Ecological Statistics. 10. 455-467.

[11] Jones, O., Maillardet, R. and Robinson, A. Introduction to scientific Programming and Simulation Using R, New York: Taylor and Francis Group, (2009).

[12] Khattree, R. (1989). Characterization of inverse-Gaussian and gamma distributions through their length-biased distributions. IEEE Transactions on Reliability, 38, 610-611.

[13] Nanuwong, N. and Bodhisuwan, W. (2014). Length Biased Beta-Pareto Distribution and Its Structural Properties with Application. Journal of Mathematics and Statistics, 10(1), 49-57.

[14] Patil, G. P. and Ord, J. K. (1976). On size biased sampling and related form invariant weighted distributions. Sankhya, Series B 38, 48-61

[15] Patil, G. P. and Rao, C. R. (1978). Weighted distributions and size biased sampling with applications to wildlife populations and human families. Biometrics 34, 179-189.

[16] Rao, C. R. (1965). On discrete distributions arising out of method of ascertainment in classical and Contagious Discrete, G.P. Patiled; Pergamum Press and Statistical publishing Society, Calcutta. 320-332.

[17] Ratnaparkhi, M. V. and Naik-Nimbalkar, U. V. (2012). The length biased lognormal distribution and its application in the analysis of data from oil field exploration studies. Modern Applied Statistical Methods, 11, 225–260.

[18] Renyi, A. (1961). On measures of entropy and information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. Contributions to the Theory of Statistics, Berkeley, California: University of California Press, (1): 547–561

[19] Seenoi, P., Supapa, K. T. and Bodhisuwan, W. (2014). The length biased exponentiated inverted Weibull Distribution. International Journal of Pure and Applied Mathematics, 92, 191–206.

[20] Shaked, M. and Shanthikumar, J. G. Stochastic Orders and Their Applications, Academic Press, New York. (1994).

[21] Shanker, R. and Ghebretsadik, A. H. (2013). A New Quasi Lindley distribution. International Journal of Statistics and Systems, 8(2), 143-156.

[22] Shanker, R. and Mishra, A. (2013). A quasi Lindley distribution. African Journal of Mathematics and Computer Science Research. Vol. 6(4), 64-71.

[23] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibb's statistics. *Journal of statistical physics*, 52(1-2), 479-487.

[24] Zelen, M. (1974). Problems in Cell Kinetics and Early Detection of Disease. Reliability and Biometry, SIAM, Philadelphia, 701-726.