Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online): 2229-5666
Vol. 12, Issue 1 (2019): 205-223

MULTIVARIATE AREA-BIASED LINDLEY DISTRIBUTION: PROPERTIES AND APPLICATIONS

Shakila Bashir[1*] and Huma Mahmood[2]
[*1]Department of Statistics
Forman Christian College, A Chartered University, Lahore, Pakistan
[2]Kinnaird College for Women, Lahore, Pakistan

E Mail: shakilabashir@fccollege.edu.pk; shakilabashir15@gmail.com

Received July 08, 2018
Modified May 10, 2019
Accepted June 22, 2019

Abstract

In this research paper, we have proposed multivariate area biased Lindley distribution and developed some of its statistical properties like mean, covariance, correlation and joint moments. Before proposing multivariate area biased Lindley distribution, we have first introduced area-biased Lindley distribution and have derived its various statistical properties such as mean, mode, moments, harmonic mean, and coefficient of variation, skewness and kurtosis. Expressions for Lorenz curve and entropy as a measure for uncertainty reduction are derived. The method of moments and maximum likelihood estimates of the parameter of area-biased Lindley distribution are investigated. Some reliability measures such as hazard function, mean residual life, survival function and reverse hazard rate have been studied. Area biased Lindley distribution is compared with one parameter Lindley distribution and size biased Lindley distribution by taking a real life data. We have then proposed that area biased Lindley distribution is expected to serve as a best model as compared to other models available in the statistical literature for modeling positive real data. Finally, we have extended our suggested distribution to multivariate area biased Lindley distribution and developed their expressions for mean, joint moments, covariance and correlation.

Key Words: ALD, SLD, LD, PDF, CDF, Reliability.

1. Introduction

In observational studies of many applied sciences such as medical, engineering, medicine, finance etc. it's often crucial to model and analyze life time data. Various continuous lifetime distributions have been introduced to model such kind of data sets.

In 1958, a British Statistician D.V Lindley proposed a one parameter distribution known as Lindley distribution. The probability distribution function (PDF) of Lindley distribution is given by

$$f(x) = \frac{\theta^2}{(\theta+1)}(1+x)e^{-\theta x} \begin{cases} x > 0 \\ \theta > 0 \end{cases} \quad (1.1)$$

The cumulative distribution function (CDF) of the Lindley distribution has been obtained as

$$F(x) = 1 - \left[\frac{\theta + 1 + \theta x}{\theta + 1}\right] e^{-\theta x} \begin{cases} x > 0 \\ \theta > 0 \end{cases} (1.2)$$

The first four moments about origin of the one parameter Lindley distribution has been obtained as

$$\mu_1^{'} = \frac{\theta + 2}{\theta(\theta + 1)}, \quad \mu_2^{'} = \frac{2(\theta + 3)}{\theta^2(\theta + 1)}, \mu_3^{'} = \frac{6(\theta + 4)}{\theta^3(\theta + 1)},$$

$$\mu_4^{'} = \frac{24(\theta + 5)}{\theta^4(\theta + 1)} \quad (1.3)$$

In last ten years, various works have been done on Lindley distribution. Lindley distribution has been further generalized as a special mixture of gamma distribution, poison, exponential, beta, power and many other distributions. Shanker, Fesshaye and Sharma (2016) introduced two parameter Lindley distribution and its applications to model lifetime Data. Bhati, Malik and Jose (2016) proposed a new class of Lindley generated distributions named as "a new 3-parameter extension of generalized Lindley distribution". Babtain, Ahmed and Merovci (2015) introduced a generalized form of Lindley Distribution having five parameters named FPLD. Dutta and Borah (2014) discussed some further properties and applications of size-biased Poisson-Lindley distribution. Bhati and Malik (2014) introduce a new distribution named as Lindley-exponential distribution generated by Lindley distribution.

In practical and real-life situations, often well-defined sampling frames and research methodologies for selection of sampling units do not exist and are difficult to develop, particularly in observational studies of wildlife, insects, fish, and plants and even for human beings. When a researcher collects data and records an observation by nature according to a certain stochastic model, the recorded sampling units will not have the original distribution unless each unit is given an equal chance of being recorded.

Size-biased and area-biased distributions are special cases of the moment distributions which is a type of weighted distributions. Fisher (1934) introduced weighted distributions to model identified bias in data that was later progressed and formalized in a unifying theory by Rao (1965). Weighted distributions naturally arise when sampling units have an unequal probability of being recorded, such as from probability proportional to size (PPS) designs in sampling.

If a random variable X has PDF $f(x:\theta)$, with unknown parameter θ, then the weighted distribution is of the form

$$f^w(x:\theta) = \frac{w(x)f(x:\theta)}{E[w(x)]} \quad (1.4)$$

Moment distributions arises when,

$$w(x) = x^a$$

And eq. (1.4) can be rewrite as,

$$f_{\alpha}^{*}(x;\theta)=\frac{x^{\alpha}f(x:\theta)}{\mu_{a}^{'}} \quad (1.5)$$

Where, $\mu_{a}^{'}=\int x^{\alpha}\,f(x;\theta)dx$ is the α^{th} raw moment of $f(x;\theta)$. In context of sampling, two of the most common cases of moment distributions occur when we take $\alpha=1$ or $\alpha=2$, these special cases are termed as size-biased/length biased and area biased respectively.

Let $X_1, X_2 \dots, X_p$ be a set of p random variables with probability function $f_i(x_i)$ and distribution function $F_i(x_i)$. Gumbel (1958) provided following method of obtaining joint density of the vector x as

$$f(x_i)=\prod_{i=1}^{p}f_i(x_i)\left[1+\theta\prod_{i=1}^{p}\{2F_i(x_i)-1\}\right] \quad (1.6)$$

Suppose Y is another random variable with density function $f(y)$ and distribution function $F(y)$ then the joint density of Y and X can be obtained by using

$$f(y,x)=f(y)\times\prod_{i=1}^{p}f(x_i)\left[1+\theta\{2F(y)-1\}\times\prod_{i=1}^{p}\{2F(x_i)-1\}\right] \quad (1.7)$$

which can be used to generate multivariate distribution based on any available marginal. Several authors have used (1.7) to propose multivariate distribution.

2. Area Biased Lindley Distribution

Using pdf in (1.1) in (1.5) for α=2, the PDF of the area-biased Lindley distribution (ALD) is

$$f^{*}(x;\theta)=\frac{x^2\theta^4}{2(\theta+3)}(1+x)e^{-\theta x}, \qquad \begin{cases}x>0\\ \theta>0\end{cases} \quad (2.1)$$

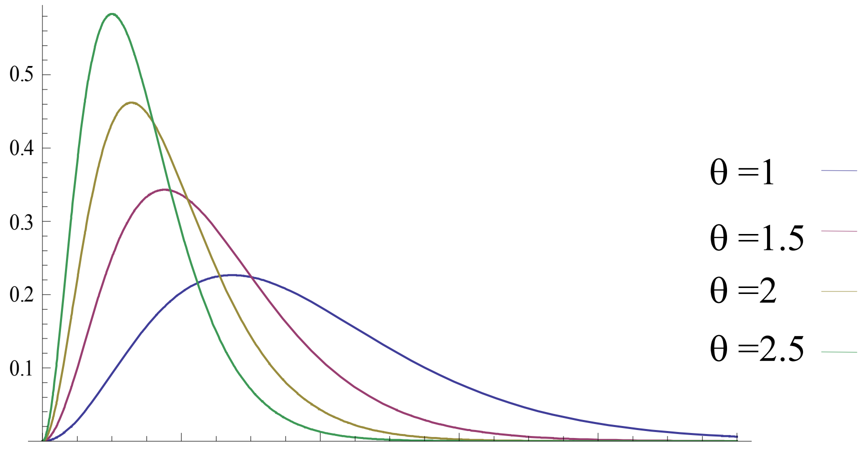

Figure: 1: PDF of Area Biased Lindley Distribution

The cumulative distribution function for area biased Lindley distribution is

$$F(x) = 1 - \left[1 + \theta x + \frac{\theta^2}{2}x^2 + \frac{\theta^3}{2(\theta+3)}x^3\right]e^{-\theta x} \tag{2.2}$$

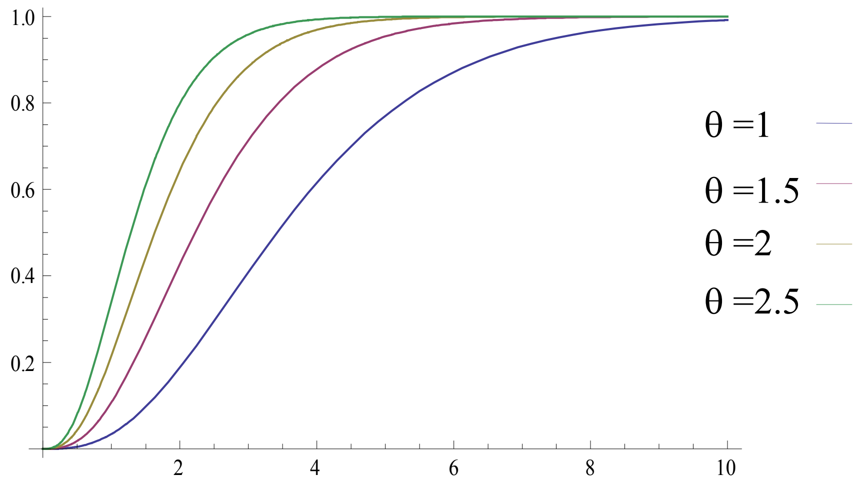

Figure 2: CDF of Area Biased Lindley Distribution

2.1 Properties of ALD

The r^{th} moment about origin of the one parameter area biased Lindley distribution has been obtained as

$$\mu_r' = \frac{(r+2)(r+1)r\Gamma(r)(\theta+r+3)}{2\theta^r(\theta+3)} \tag{2.3}$$

We obtain first four moments about origin by putting r=1, 2, 3, 4 respectively

$$\mu_1' = \frac{3(\theta+4)}{\theta(\theta+3)} \mu_2' = \frac{12(\theta+5)}{\theta^2(\theta+3)} \mu_3' = \frac{60(\theta+6)}{\theta^3(\theta+3)} \mu_4' = \frac{360(\theta+7)}{\theta^4(\theta+3)} \tag{2.4}$$

The central moments of ALD are

$$\mu_2 = \frac{3(\theta^2+8\theta+12)}{\theta^2(\theta+3)^2} \mu_3 = \frac{6(\theta^3+12\theta^2+36\theta+36)}{\theta^3(\theta+3)^3}$$

$$\mu_4 = \frac{3(15\theta^4+240\theta^3+1224\theta^2+2592\theta+1944}{\theta^4(\theta+3)^4}$$

The mean and variance of area biased Lindley distribution are

$$E(x) = \frac{3(\theta+4)}{\theta(\theta+3)} \tag{2.5}$$

$$Var(x) = \frac{3(\theta+2)(\theta+6)]}{\theta^2(\theta+3)^2} \tag{2.6}$$

The harmonic mean of area biased Lindley distribution is derived as

$$HM = \frac{2(\theta+3)}{\theta(\theta+2)} \tag{2.7}$$

The mode of area biased Lindley distribution is

$$x = \frac{-(\theta-3) \pm \sqrt{\theta^2+9-2\theta}}{2\theta} \tag{2.8}$$

The skewness, coefficient of variation and kurtosis of ALD are

$$Skewness = \sqrt{\beta_1} = \frac{2(\theta^3+12\theta^2+36\theta+36)}{3\sqrt{3}(\theta^2+8\theta+12)^{3/2}} \tag{2.9}$$

$$Kurtosis = \beta_2 = \frac{15\theta^4+240\theta^3+1224\theta^2+2592\theta+1944}{3(\theta^2+8\theta+12)^2} \tag{2.10}$$

$$Coefficients\ of\ Variation = \gamma = \frac{\sqrt{3(\theta^2+8\theta+12)}}{3(\theta+4)} \tag{2.11}$$

Some additional properties of the ALD

i. Since

$$\mu - \sigma^2 = \frac{3\theta^3 + 18\theta^2 + 12\theta - 36}{\theta^2(\theta+3)^2}$$

$$\mu = \sigma^2 + \frac{3(\theta^3 + 16\theta^2 + 4\theta - 12)}{\theta^2(\theta+3)^2} \tag{2.12}$$

From equation (2.12), generally it can be seen that ALD is over dispersed but as $\theta \to \infty$ then $\mu = \sigma^2$, and ALD is equi-dispersed. For large θ, ALD is equi-dispersed.

ii. Since

$$\frac{f(x+1;\theta)}{f(x;\theta)} = \left(x + \frac{1}{x}\right)\left(1 + \frac{2}{x}\right)e^{-\theta} \tag{2.13}$$

is a decreasing function in x, therefore ALD is unimodal has an increasing failure rate and decreasing mean residual life.

2.2 Lorenz Curve Property for ALD

For Lorenz Curve

$$\emptyset(x) = \frac{1}{\mu_1'}\int_0^x tf(t;\theta)dt \tag{2.14}$$

Where

$$f^*(t;\theta) = \frac{t^2\theta^4}{2(\theta+3)}(1+t)e^{-\theta t}\begin{cases} t > 0 \\ \theta > 0 \end{cases}$$

$$\mu_1' = \frac{3(\theta+4)}{\theta(\theta+3)}$$

After simplifying

$$\emptyset(x) = 1 - \left[1 + x\theta + \frac{x^2\theta^2}{2} + \frac{x^3\theta^3}{6} + \frac{x^4\theta^4}{6(\theta+4)}\right]e^{-\theta x} \tag{2.15}$$

Theorem 2.2.1: If a random variable follows the area-biased Lindley distribution with cdf F(x) then,

$$\frac{\emptyset'(x)}{F'(x)} = \frac{x}{\mu_1'} \tag{2.16}$$

Proof: Taking first derivative of $\emptyset(x)$

$$\emptyset'(x) = \frac{x^3\theta^5}{6(\theta+4)}(1+x)e^{-\theta x}$$

Now, we need to differentiate the CDF of ALD

$$F'(x) = \left[\frac{x^2\theta^3}{2(\theta+3)}(\theta + x\theta)\right]e^{-\theta x}$$

Now to show

$$\frac{\emptyset'(x)}{F'(x)} = \frac{x}{\mu_1'}$$

$$\Rightarrow \frac{[x^2\theta^4(1+x)e^{-\theta x}]/6(\theta+4)}{[x^2\theta^5(1+x)e^{-\theta x}]/2(\theta+3)} = \frac{x}{3(\theta+4)/\theta(\theta+4)}$$

Simplifying both sides

$$\Rightarrow \frac{x\theta(\theta+3)}{3(\theta+4)} = \frac{x\theta(\theta+3)}{3(\theta+4)}$$

a) $\emptyset(x)$ is an increasing function of x. Its first derivative is always positive and so is its second order derivative (convexity).

That is,

$$\emptyset''(x) = \frac{\theta^5}{6(\theta+4)}\frac{d}{dx}\left[(x^3+x^4)e^{-\theta x}\right]$$

$$\emptyset''(x) = \frac{\theta^5}{6(\theta+4)}\left[-x^3e^{-\theta x} + 3x^2e^{-\theta x} - \theta x^4e^{-\theta x} + 4x^3e^{-\theta x}\right]$$

$$\emptyset''(x) = x^2\theta^5[3 + 4x - \theta x - \theta x^2]e^{-ti\theta} > 0$$

b) It is entirely contained in a square because x is defined over [0,1]
c) The Lorenz curve is not defined if μ is either 0 or ∞ i.e. $\emptyset(0) = 0$
d) If the underlying variable is positive and has a density, the Lorenz curve is a continuous function. It is always below the 45° line or equal to it.

2.3 Shannon Entropy

For entropy, we know that

$$-H(x) = \int_{-\infty}^{\infty} f(x)lnf(x)dx \qquad -\infty < x < \infty$$

$$H(x) = \frac{3(\theta+4)}{\theta+3} - \ln\left(\frac{\theta^4}{2(\theta+3)}\right) - \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\left[\frac{(n+2)(n+1)n\Gamma(n)(\theta+n+3)}{2\theta^n(\theta+3)}\right] - \frac{\theta^4}{(\theta+3)} I_3$$

$$where\ I_3 = \int_0^{\infty} (x^2 + x^3) lnx e^{-\theta x} dx$$

3. Estimation of Parameters

An important part of statistical data analysis is model or distribution fitting. Once an appropriate statistical model has been constructed and its parameters estimated, its information can help predict survival, plan future clinical or laboratory studies, develop optimal treatment regimens and so on. In this section, we have estimated the parameter of area biased Lindley distribution and derived the properties of estimated parameter of ALD.

3.1 Maximum Likelihood Estimator (MLE)

Let $x_1, x_2 \ldots\ldots\ldots x_n$ be the random sample of size n from ALD with PDF (2.1), the maximum likelihood estimate $\tilde{\theta}$ of the parameter $\theta$is obtained as

$$f^*(x_1, x_2 \ldots\ldots\ldots x_n; \theta) = \prod_{i=1}^{n}\left[\frac{x^2\theta^4}{2(\theta+3)}(1+x)e^{-\theta x}\right]$$

$$f^*(x_1, x_2 \ldots\ldots\ldots x_n; \theta) = \left(\frac{\theta^4}{2(\theta+3)}\right)^n e^{-\theta\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} x^2(1+x)$$

Taking natural log on both sides

$$ln\prod_{i=1}^{n} f^*(x_i; \theta) = 4nln\theta - nln2 - nln(\theta+3) - \theta\sum_{i=1}^{n} x_i + ln\sum_{i=1}^{n} {x_i}^2 + ln\sum_{i=1}^{n}(1+x_i)$$

Differentiating with respect to θ and setting to 0, we get

$$\hat{\theta} = \frac{3 - 3\bar{x} + \sqrt{9\bar{x}^2 + 9 + 30\bar{x}}}{2\bar{x}} \tag{3.1}$$

Hence, the sample average is the MLE for $\bar{x}$.

3.2 Methods of Moments (MOM)

Thus, if $x_1, x_2 \ldots\ldots x_n$be the random sample from ALD with PDF f(x) (2.1), the method of moments estimate $\tilde{\theta}$ is given by

$$\tilde{\theta} = \frac{3 - 3\bar{x} + \sqrt{9\bar{x}^2 + 9 + 30\bar{x}}}{2\bar{x}} \tag{3.2}$$

It can be seen that the estimated parameter of θ is same obtained by MLE and MOM.

Theorem 3.1: The MOM/MLE estimator $\tilde{\theta}$ of θ is positively biased

Proof: Let

$$\tilde{\theta} = \psi(x) \tag{3.3}$$

Where

$$\psi(z) = \frac{3 - 3z + \sqrt{9z^2 + 9 + 30z}}{2z} \tag{3.4}$$

If a function is doubly differentiable, then f is convex and only if $d^2f/dx^2 \geq 0$

$$\psi''(z) = \left(\frac{270z^4 + 918z^3 + 810z^2 + 162z + 54z^3\left(\sqrt{9z^2 + 9 + 30z}\right) + 90z^2\left(\sqrt{9z^2 + 9 + 30z}\right) + 54z\sqrt{9z^2 + 9 + 30z})}{2z^4(9z^2 + 9 + 30z)^{3/2}} \right) \tag{3.5}$$

Since

$$\psi\{E(\bar{X})\} = \psi(\mu) = \psi\left[\frac{3(\theta + 4)}{\theta(\theta + 3)}\right] = \theta \tag{3.6}$$

Therefore

$$E(\tilde{\theta}) > \theta \tag{3.7}$$

Theorem 3.2: The MOM/MLE estimator $\tilde{\theta} of \theta$ is consistent and asymptotically normal

$$\sqrt{n}(\tilde{\theta} - \theta) \xrightarrow{d} N[0, v^2(\theta)] \tag{3.8}$$

Where

$$v^2(\theta) = \frac{\theta^2(\theta + 3)}{3(\theta^2 + 8\theta + 12)}$$

Proof:

Asymptotically Normal:

$\psi(\mu)$ is a differentiable function and $\psi'(\mu) \neq 0$ then by using delta method we have

$$\sqrt{n}\left(\psi(\bar{X}) - \psi(\mu)\right) \xrightarrow{d} N[0, [\psi'(\mu)]^2\sigma^2]$$

Finally we have $\psi(\bar{X}) = \tilde{\theta}, \psi(\mu) = \theta and$

$$\psi'(\mu) = \frac{-15\mu - 9 - 3\sqrt{9\mu^2 + 9 + 30\mu}}{2\mu^2\sqrt{9\mu^2 + 9 + 30\mu}} = \frac{\theta^2(\theta + 3)}{3(\theta^2 + 8\theta + 12)} \tag{3.9}$$

The theorem 3.2 follow the asymptotic $100(1 - \alpha)\%$ confidence interval for $\theta$is

$$\tilde{\theta} \pm z\frac{\alpha}{2}\frac{v(\tilde{\theta})}{\sqrt{n}}$$

4. Reliability Measures

Reliability data measure the time to a certain event, such as the development of a given disease, failure, death, relapses, response, parole or divorce. Statistical methods for reliability analysis have an important role in many disciplines of life. In this section, we have derived various reliability measures for ALD.

4.1 Survival Function

Survival function or reliability function of area biased Lindley is given by

$$R(x) = \left[1 + \theta x + \frac{\theta^2}{2}x^2 + \frac{\theta^3}{2(\theta+3)}x^3\right]e^{-\theta x} \tag{4.1}$$

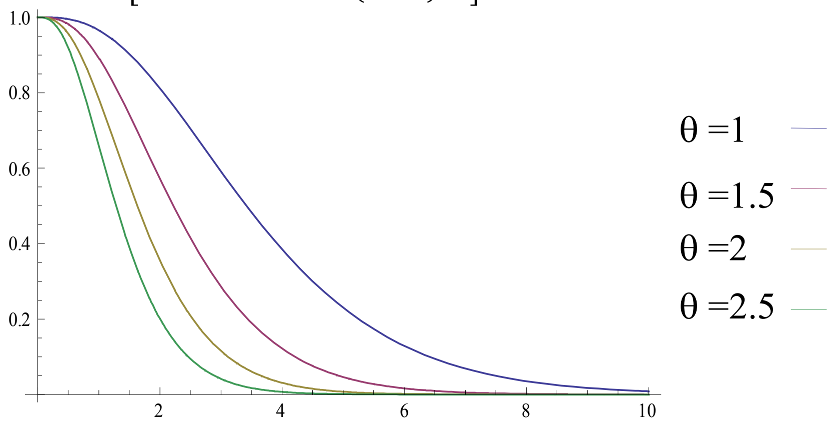

Figure 3: Survival Function of Area Biased Lindley Distribution

4.2 Hazard Rate

The hazard rate of area biased Lindley distribution is derived as

$$H(x) = \frac{x^2\theta^4(1+x)}{2(\theta+3) + 2\theta x(\theta+3) + \theta^2 x^2(\theta+3) + \theta^3 x^3} \tag{4.2}$$

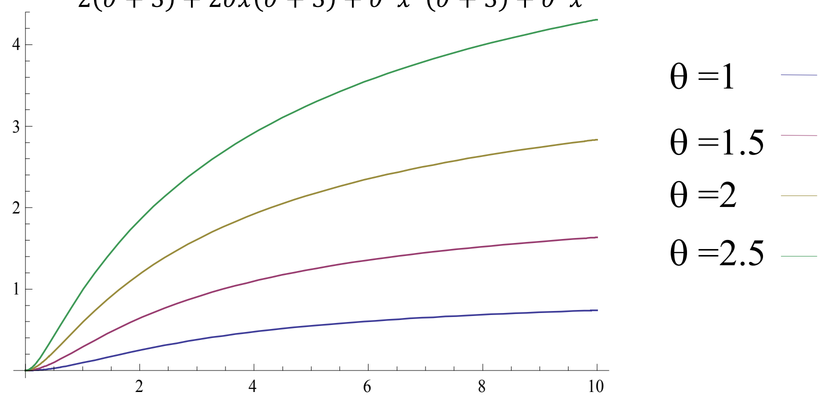

Figure 4: Hazard Rate of Area Biased Lindley Distribution

4.3 Mean Residual Life

The, mean residual life of area biased Lindley distribution can be obtained as

$$m(x) = \frac{6(\theta+4)/\theta + 2x(2\theta+9) + \theta x^2(\theta+6) + x^3\theta^2}{2(\theta+3) + 2\theta x(\theta+3) + \theta^2 x^2(\theta+3) + \theta^3 x^3} \tag{4.3}$$

It can easily be verifying that

$$m(0) = \mu_1^{'}$$

$$m(0) = \frac{6(\theta+4)}{2\theta(\theta+3)}$$

Thus,

$$m(0) = \frac{3(\theta+4)}{\theta(\theta+3)} = \mu_1^{'}$$

4.4 Mean Inactivity Time

The mean activity time of an area biased Lindley distribution can be derived by

$$MIT = \frac{2x(\theta+3) + 6(\theta+4)/\theta - [2x(2\theta+9) + x^2\theta(\theta+6) + x^3\theta^2 + 6(\theta+4)/\theta]e^{-\theta x}}{2(\theta+3) - [2(\theta+3) + 2\theta x(\theta+3) + \theta^2 x^2(\theta+3) + \theta^3 x^3]e^{-\theta x}} \tag{4.4}$$

4.5 Reverse Hazard Rate

$$RHR = \frac{x^2\theta^4(1+x)e^{-\theta x}}{(2(\theta+3) - [2(\theta+3) + 2\theta x(\theta+3) + \theta^2 x^2(\theta+3) + \theta^3 x^3]e^{-\theta x})} \tag{4.5}$$

5. Applications

Data Set 1: 72 Guinea Pigs Data

The first data represent the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by Bjerkedal(1960).

Data set 2: Relief Time of 20 Patients

The second data, reported by Gross and Clark(1975), represents the lifetime data relating to relief time (in minutes) of 20 patients receiving an analgesic. The data set is as follows: 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.2, 2.3, 2.7, 3.0, 4.1

	Model	Parameter Estimate	Test-Statistic	Table-Value	Calculated Value
Data 1	ALD	0.022046	Chi Square	7.815	2.88507
	SLD	0.016519			3.288299
	LD	0.011			7.7712
Data 2	ALD	1.823154	Kolmogorov-Smirnov (K-S Statistic)	0.294	0.17653
	SLD	1.310782			0.18161
	LD	0.7853			0.188931

Table 1: Comparison between model fit

Interpretation

It can easily be seen from the above table that the ALD is better than the LD and SLD for modeling life time data. Thus, the proposed distribution can be used as an alternative model to other models available in the statistical literature.

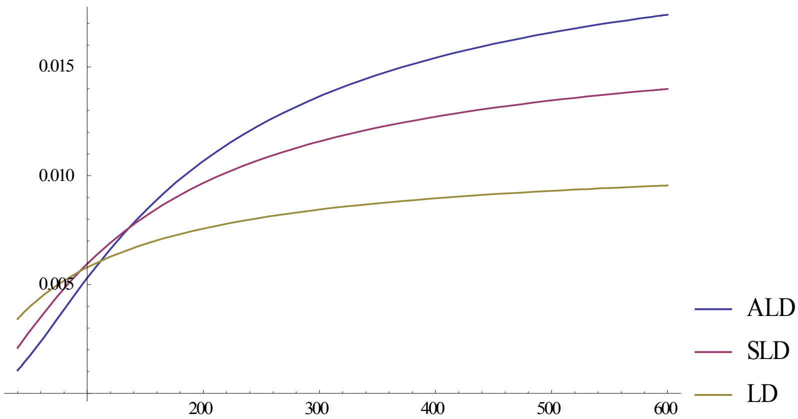

Figure 5: Hazard Rate of 72 guinea pigs

Interpretation

From the above graph, that the hazard rate of 72 guinea pigs infected with virulent tubercle bacilli is monotonically increasing. During an initial period, the risk is low but subsequently increases because of wear-out failures. It may indicate that the 72 guinea pigs infected with virulent tubercle bacilli are not responding to the treatment.

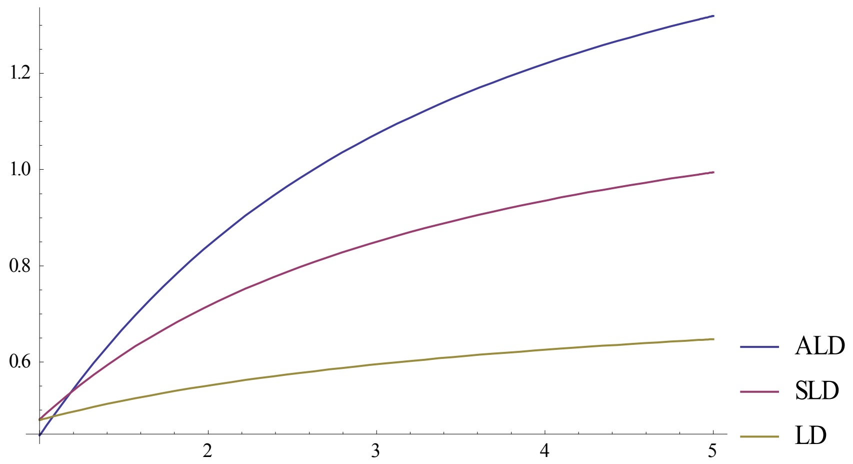

Figure 6: Hazard Rate of 20 Patients

Interpretation

From the above graph, we see that the hazard rate of 20 patients receiving analgesic is monotonically increasing.

6. Multivariate Area Biased Lindley Distribution

In this section, we have developed a bivariate, trivariate and a general expression of multivariate distribution for area biased Lindley distribution. Also, expressions for their respective means, moments, co variances and correlations have been obtained

The bivariate distribution is readily obtained from (1.7) by using p=2 and assuming that density and distribution function if i^{th} random variable is given as

$$f^*(x_i) = \frac{x_i^2\theta^4}{2(\theta_i+3)}(1+x_i)e^{-\theta_i x_i}\begin{cases} x>0 \\ \theta>0 \end{cases}$$

$$F(x_i) = 1 - \left[1+\theta_i x_i + \frac{\theta_i^2}{2}x_i^2 + \frac{\theta_1^3}{2(\theta+3)}x_i^3\right]e^{-\theta_i x_i}\begin{cases} x>0 \\ \theta>0 \end{cases}$$

Such that $F(0) = 0$ and $F(\infty) = 1$

Using (2.1) and (2.2) in (1.7), the PDF of bivariate ALD is given as

$$f(x_i, x_j) = \frac{x_i^2(1+x_i)\theta_i^4 e^{-\theta_i x_i}}{2(\theta_i+3)}\frac{x_j^2(1+x_j)\theta_j^4 e^{-\theta_j x_j}}{2(\theta_j+3)}$$

$$\left[1+\theta\left\{\begin{array}{l}\left(1-2e^{-\theta_i x_i}-2\theta_i x_i e^{-\theta_i x_i}-\theta_i^2 x_i^2 e^{-\theta_i x_i}+\frac{\theta_i^3 x_i^3}{(\theta_i+3)}e^{-\theta_1 x_1}\right)\\ \left(1-2e^{-\theta_j x_j}-2\theta_j x_j e^{-\theta_j x_j}-\theta_j^2 x_j^2 e^{-\theta_j x_j}+\frac{\theta_j^3 x_j^3}{(\theta_j+3)}e^{-\theta_j x_j}\right)\end{array}\right\}\right] \tag{6.1}$$

$$0 < x_i < \infty$$
$$0 < x_j < \infty$$

Some Properties of Bivariate Area Biased Lindley Distribution have been derived. The mean of bivariate area biased Lindley distribution is

$$E(X_i, X_j) = \frac{\theta_i^4\theta_j^4}{4(\theta_i+3)(\theta_j+3)}\left[\frac{36(\theta_i+4)(\theta_j+4)}{\theta_i^5\theta_j^5} + \theta\left(\frac{30\theta_i^2+210\theta_i+315)}{16\theta_i^5(\theta_i+3)}\right)\left(\frac{30\theta_j^2+210\theta_j+315)}{16\theta_j^5(\theta_j+3)}\right)\right] \tag{6.2}$$

The covariance of bivariate area biased Lindley distribution is

$$cov(X_i, X_j) = \theta\frac{(30\theta_i^2+210\theta_i+315)(30\theta_j^2+210\theta_j+315)}{1024\theta_i\theta_j(\theta_i+3)^2(\theta_j+3)^2} \tag{6.3}$$

The correlation of bivariate area biased Lindley distribution is

$$correlation = \rho_{x_i x_j} = \theta\frac{(10\theta_i^2+70\theta_i+105)(10\theta_j^2+70\theta_j+105)}{1024(\theta_i+3)(\theta_j+3)\sqrt{(\theta_i^2+8\theta_i+12)}\sqrt{(\theta_j^2+8\theta_j+12)}} \tag{6.4}$$

The first four moments for bivariate area biased Lindley distribution has been respectively obtained as

$$\mu'_{1,1} = \frac{\theta_1^4\theta_2^4}{4(\theta_1+3)(\theta_2+3)}\left[\frac{36(\theta_1+4)(\theta_2+4)}{\theta_1^5\theta_2^5} + \theta\left(\frac{30\theta_1^2+210\theta_1+315)}{16\theta_1^5(\theta_1+3)}\right)\left(\frac{30\theta_2^2+210\theta_2+315)}{16\theta_2^5(\theta_2+3)}\right)\right] \tag{6.5}$$

$$\mu'_{2,2} = \frac{\theta_1^4\theta_2^4}{4(\theta_1+3)(\theta_2+3)}\left[\frac{576(\theta_1+5)(\theta_2+5)}{\theta_1^6\theta_2^6} + \theta\left(\frac{210\theta_1^2+1680\theta_1+2835)}{16\theta_1^6(\theta_1+3)}\right)\left(\frac{210\theta_2^2+1680\theta_2+2835)}{16\theta_2^6(\theta_2+3)}\right)\right] \tag{6.6}$$

$$\mu'_{3,3} = \frac{\theta_1^4\theta_2^4}{4(\theta_1+3)(\theta_2+3)}\left[\frac{120(\theta_1+6)120(\theta_2+6)}{\theta_1^7\theta_2^7} + \theta\left(\frac{1365\theta_1^2+878885\theta_1+250425)}{16\theta_1^7(\theta_1+3)}\right)\left(\frac{1365\theta_2^2+878885\theta_2+250425}{16\theta_2^7(\theta_2+3)}\right)\right] \quad (6.7)$$

$$\mu'_{4,4} = \frac{\theta_1^4\theta_2^4}{4(\theta_1+3)(\theta_2+3)}\left[\frac{720(\theta_1+7)720(\theta_2+7)}{\theta_1^8\theta_2^8} + \theta\left(\frac{42210\theta_1^2+878885\theta_1+87885}{8\theta_1^8(\theta_1+3)}\right)\left(\frac{42210\theta_2^2+158445\theta_2+87885}{8\theta_2^8(\theta_2+3)}\right)\right] \quad (6.8)$$

The PDF of trivariate area biased Lindley distribution is derived as

$$f(x_1,x_2,x_3) = \frac{x_1^2x_2^2x_3^2(1+x_1)(1+x_2)(1+x_3)\theta_1^4\theta_2^4\theta_3^4e^{-\theta_1x_1}e^{-\theta_2x_2}e^{-\theta_3x_3}}{8(\theta_1+3)(\theta_2+3)(\theta_3+3)}$$

$$\left[1+\theta\left\{\begin{array}{l}\left(1-2e^{-\theta_1x_1}-2\theta_1x_1e^{-\theta_1x_1}-\theta_1^2x_1^2e^{-\theta_1x_1}+\frac{\theta_1^3x_1^3}{(\theta_1+3)}e^{-\theta_1x_1}\right)\\ \left(1-2e^{-\theta_2x_2}-2\theta_2x_2e^{-\theta_2x_2}-\theta_2^2x_2^2e^{-\theta_2x_2}+\frac{\theta_2^3x_2^3}{(\theta_2+3)}e^{-\theta_2x_2}\right)\\ \left(1-2e^{-\theta_3x_3}-2\theta_3x_3e^{-\theta_3x_3}-\theta_3^2x_3^2e^{-\theta_3x_3}+\frac{\theta_3^3x_3^3}{(\theta_3+3)}e^{-\theta_3x_3}\right)\end{array}\right\}\right] \quad (6.9)$$

$$0 < x_1 < \infty$$
$$0 < x_2 < \infty$$
$$0 < x_3 < \infty$$

Some Properties of Trivariate Area Biased Lindley Distribution have been derived. The mean of trivariate area biased Lindley distribution is

$$E(X_1,X_2,X_3) = \frac{\theta_1^4\theta_2^4\theta_3^4}{8(\theta_1+3)(\theta_2+3)(\theta_3+3)}\left[\left(\frac{216(\theta_2+4)(\theta_1+4)(\theta_3+4)}{\theta_1^5\theta_2^5\theta_3^5}\right)+\theta\left(\frac{30\theta_1^2+210\theta_1+315)}{16\theta_1^5(\theta_1+3)}\right)\left(\frac{30\theta_2^2+210\theta_2+315)}{16\theta_2^5(\theta_2+3)}\right)\left(\frac{30\theta_3^2+210\theta_3+315)}{16\theta_3^5(\theta_3+3)}\right)\right] \quad (6.10)$$

The correlation of trivariate area biased Lindley distribution is

$correlation =$

$$\rho_{x_1x_2x_3} = \theta \frac{(10\theta_1^2 + 70\theta_1 + 105)\left(10\theta_2^2 + 70\theta_2 + 105\right)\left(10\theta_3^2 + 70\theta_3 + 105\right)}{32768(\theta_1 + 3)\left(\theta_2 + 3\right)\left(\theta_3 + 3\right)\sqrt{(\theta_1^2 + 8\theta_1 + 12)}\sqrt{\left(\theta_2^2 + 8\theta_2 + 12\right)}\sqrt{\left(\theta_3^2 + 8\theta_3 + 12\right)}}$$

(6.11)

The first four moments for trivariate area biased Lindley distribution has been respectively obtained as

$$\mu'_{1,1,1} = \frac{\theta_1^4\theta_2^4\theta_3^4}{4\left(\theta_1 + 3\right)\left(\theta_2 + 3\right)\left(\theta_3 + 3\right)}\left[\begin{array}{l}\dfrac{216(\theta_1 + 4)\left(\theta_2 + 4\right)\left(\theta_3 + 4\right)}{\theta_1^5\theta_2^5\theta_3^5} + \theta\left(\dfrac{30\theta_1^2 + 210\theta_1 + 315)}{16\theta_1^5(\theta_1 + 3)}\right) \\ \left(\dfrac{30\theta_2^2 + 210\theta_2 + 315)}{16\theta_2^5(\theta_2 + 3)}\right)\left(\dfrac{30\theta_3^2 + 210\theta_3 + 315)}{16\theta_3^5(\theta_3 + 3)}\right)\end{array}\right]$$

(6.12)

$$\mu'_{2,2,2} = \frac{\theta_1^4\theta_2^4\theta_3^4}{4\left(\theta_1 + 3\right)\left(\theta_2 + 3\right)\left(\theta_3 + 3\right)}\left[\begin{array}{l}\dfrac{13824\left(\theta_1 + 5\right)\left(\theta_2 + 5\right)\left(\theta_3 + 5\right)}{\theta_1^6\theta_2^6\theta_3^6} + \theta\left(\dfrac{210\theta_1^2 + 1680\theta_1 + 2835)}{16\theta_1^6\left(\theta_1 + 3\right)}\right) \\ \left(\dfrac{210\theta_2^2 + 1680\theta_2 + 2835)}{16\theta_2^6\left(\theta_2 + 3\right)}\right)\left(\dfrac{210\theta_3^2 + 1680\theta_3 + 2835)}{16\theta_3^6\left(\theta_3 + 3\right)}\right)\end{array}\right]$$

(6.13)

$$\mu'_{3,3,3} = \frac{\theta_1^4\theta_2^4\theta_3^4}{4\left(\theta_1 + 3\right)\left(\theta_2 + 3\right)\left(\theta_3 + 3\right)}\left[\begin{array}{l}\dfrac{120\left(\theta_1 + 6\right)120\left(\theta_2 + 6\right)120\left(\theta_3 + 6\right)}{\theta_1^7\theta_2^7\theta_3^7} + \\ \theta\left(\dfrac{1365\theta_1^2 + 878885\theta_1 + 250425)}{16\theta_1^7\left(\theta_1 + 3\right)}\right)\left(\dfrac{1365\theta_2^2 + 878885\theta_2 + 250425}{16\theta_2^7\left(\theta_2 + 3\right)}\right) \\ \left(\dfrac{1365\theta_3^2 + 878885\theta_3 + 250425}{16\theta_2^7\left(\theta_3 + 3\right)}\right)\end{array}\right]$$

(6.14)

$$\mu'_{4,4,4} = \frac{\theta_1^4\theta_2^4\theta_3^4}{4\left(\theta_1 + 3\right)\left(\theta_2 + 3\right)\left(\theta_3 + 3\right)}\left[\begin{array}{l}\dfrac{720\left(\theta_1 + 7\right)720\left(\theta_2 + 7\right)\left(\theta_3 + 7\right)}{\theta_1^8\theta_2^8\theta_3^8} + \theta\left(\dfrac{42210\theta_1^2 + 878856\theta_1 + 87885}{8\theta_1^8\left(\theta_1 + 3\right)}\right) \\ \left(\dfrac{42210\theta_2^2 + 1584456\theta_2 + 87885}{8\theta_2^8\left(\theta_2 + 3\right)}\right)\left(\dfrac{42210\theta_3^2 + 1584456\theta_3 + 87885}{8\theta_3^8\left(\theta_3 + 3\right)}\right)\end{array}\right]$$

(6.15)

The general form of multivariate ALD has been given by using the method of

Gumbel (1958) assuming that density and distribution function of i^{th} random variable is given as

$$f^*(x_i;\theta_i) = \frac{x_i^2\theta_i^4}{2(\theta_i+3)}(1+x_i)e^{-\theta_i x_i}, \qquad \begin{cases} x_i > 0 \\ \theta_i > 0 \\ i = 1,2,3 \ldots \end{cases}$$

$$F(x_i) = 1 - \left[1 + \theta_i x_i + \frac{\theta_i^2 x_i^2}{2} + \frac{\theta_i^3 x_i^3}{2(\theta_i+3)} x^3\right] e^{-\theta_i x_i} \begin{cases} x > 0 \\ \theta > 0 \\ i = 1,2,3.. \end{cases}$$

such that $F(0) = 0$ and $F(\infty) = \infty$. The parameter θ is a measure of association in defining bivariate and multivariate distributions.

$$f(x_i) = f(x_1)f(x_2)\ldots\ldots f(x_i)[1 + \theta\{2F(x_1) - 1\}\{2F(x_2) - 1\}\ldots..\{2F(x_i) - 1\}]$$

Where

$$\begin{gathered} x_1 \sim ABLD(\theta_1) \\ x_2 \sim ABLD(\theta_2) \\ \vdots \\ x_i \sim ABLD(\theta_i) \end{gathered}$$

Expression for multivariate ALD mean has been obtained as

$$E(X_1 X_2 \cdots X_i) = \iint_0^\infty \cdots \int_0^\infty x_1 x_2 \cdots x_i f(x_1 x_2 \cdots x_k)\, dx_1 dx_2 \cdots dx_i \qquad 0 < x_i < \infty$$

$$E(X_1 X_2 \cdots X_i) = \frac{\theta_2^4\theta_3^4 \cdots \theta_i}{2(\theta_2+3)2(\theta_3+3)\cdots 2(\theta_i+3)}\left[\prod_{i=1}^{p} \frac{6(\theta_i+4)}{\theta_i^5} + \theta \prod_{i=1}^{p}\left\{\left(\frac{30\theta_i^2 + 210\theta_i + 315)}{16\theta_i^5(\theta_i+3)}\right)\right\}\right] i = 1.2.3 \ldots$$

Explicit expression for multivariate ALD correlation has been obtained as

$$correlation = \rho_{x_1 x_2} = \theta \prod_{i=1}^{p}\left\{\frac{1}{2}\left[\frac{(10\theta_i^2 + 70\theta_i + 105)}{16(\theta_i+3)\sqrt{(\theta_i^2 + 8\theta_i + 12)}}\right]\right\} i = 1,2,3 \ldots.$$

r^{th}moment (about origin) expression for multivariate ALD has been derived as

$$\mu'_{r_i} = \iint_{00}^{\infty\infty} \ldots \int_0^\infty x_i^{r_i} f(x_i) dx_i \begin{cases} x_i > 0 \\ i = 1,2,3, \ldots \end{cases}$$

$$\mu'_{r_i} = \frac{\theta_1^4\theta_2^4 \dots \theta_i^4}{2(\theta_1+3)2(\theta_2+3)\dots 2(\theta_i+3)}\left[\prod_{i=1}^{p}\frac{\Gamma(r_i+3)(\theta_i+r_i+3)}{\theta^{r_i+4}} + \theta\prod_{i=1}^{p}\left\{\frac{\Gamma(r_i+3)(2^{r_i+3}-r_i-5)}{(2\theta)^{r_i+3}} + \frac{\Gamma(r_i+4)(2^{r_i+4}-r_i-6)}{(2\theta)^{r_i+4}} + \frac{\theta_i^2\Gamma(r_i+5)}{(2\theta_i)^{r_i+5}}\left(1+\frac{r_i+5}{2\theta_i}\right) + \frac{\theta_i^3\Gamma(r_i+3)}{(\theta_i+3)(2\theta_i)^{r_i+6}}\left(1+\frac{r_i+6}{2\theta_i}\right)\right\}\right]$$

The multivariate area-biased distribution can be applied on at least two of the data sets. The parameters of the multivariate ALD can be estimated through the maximum likelihood estimation (MLE) and non-linear equation of MLE can be solved by simulations.

7. Conclusion

In this research paper, we have proposed a multivariate area biased Lindley distribution and developed some of its statistical properties like mean, covariance, correlation and joint moments. Though area biased Lindley distribution is a special case of two parameter weighted Lindley distribution proposed by Ghitany*et al.* (2010) for (α=3) but we derived additional properties that were not studied by Ghitany et al (2010) for (α=3). So before proposing multivariate ALD distribution, we have presented detailed introduction about ALD and studied some of its statistical properties such as mean, mode, explicit expressions for moments and various other statistical properties. Expressions for Lorenz curve and entropy as a measure for uncertainty reduction are derived. Parameter of the ALD is estimated by the method of moments and maximum likelihood estimation and properties of the parameter has been discussed. Some reliability measures such as hazard function, mean residual life, mean inactivity time, survival function and reverse hazard rate have been studied. We then compared area biased Lindley distribution with one parameter Lindley distribution and size biased Lindley distribution and proposed how ALD provides consistently better fit than the other distributions considered. Thus, ALD distribution is found to be suitable to be used in bio-medical and forestry research. Finally we have extended our distribution to multivariate area biased Lindley, that would be applicable where two or more variables are involved, and derived explicit expressions for mean, joint moments and also for covariance and correlation to determine the relationship and strength among random variables. Hopefully, extended and useful work could be done on multivariate area biased Lindley distribution.

References

1. Abdulhakim, A., Ahmed, M.E., A-Hadi, N.A. and Faton, M.(2015). The five parameter lindley distribution, Pakistan Journal of Statistics, 31(4), p. 363-384.
2. Ghitany, M., Alqallaf, F., Al-Mutairi, D.K. and Husain, H.A. (2010). A two-parameter weighted Lindley distribution and its application to survival data, ELSEVIER, 81(2011), p. 1190-1201.

3. Ibrahim, E. and Faton, M.(2013). A new generalized Lindley distribution, IISTE Mathematical Theory and Modeling, 3(13), p. 2224-5804.
4. Jose, A.R. and Silva, A.P.(2015).The Beta exponentiated Lindley distribution, Journal of Statistical Theory and Applications, 14(1), p. 60-75.
5. Lee, E.T. and Wang, J.W.(2003). Statistical Methods for Survival Data Analysis, Hoboken, New Jersey: A John Wiley & Sons, Inc.
6. Mahmoud, M.M. and Salah M.M.(2015). A new generalized of transmuted Lindley distribution, Applied Mathematical Sciences, 9(55), p. 2729-2748.
7. Mariyam, H., Shahid, K. and Muhammad Q.S.(2014).The multivariate order statistics for exponential and Weibull distributions, Pakistan Journal of Statistics and Operation Research, 10(3), p. 361-368.
8. Mohamed, E.G. and Dhaifalla, K.A.(2008). Size-biased Poisson-Lindley distribution and its application, Metron - International Journal of Statistics, 66, p. 299-311.
9. Porinita, D. and Munindra, B. (2014). Some properties and application of size-biased Poisson-Lindley distribution, International Journal of Mathematical Archive, 5(1), p. 89-96.
10. Rama, S., Hagos, F. and Shambhu, S.(2016). Two parameter Lindley distribution and its applications, Biometrics and Biostatistics International Journal, 3(1), p. 1-7.
11. Rama, S., Shambhu, S. and Ravi, S.(2013). A two-parameter Lindley distribution for modeling waiting and survival times data, Applied Mathematics, 4, p. 363-368.
12. Samir, K.A. and Mahmoud, A.E.(2015). Exponentiated power Lindley distribution, Journal of Advanced Research, 6, p.895-905.
13. Shakila, B. and Mujahid, R.(2015). Some properties of weighted Lindley distribution, ESPRA International Journal of Economic and Business Review, 3(8), p. 11-17.
14. Shakila, B. and Mujahid, R.(2016).Poisson area-biased Lindley distribution and its applications on biological data, Biometrics and Biostatistics International Journal , 3(1), p. 1-10.
15. Tanka, R.A. and Shyyam, S.(2014). Poisson size-biased Lindley distribution, International Journal of Scientific and Research Publications, 4(1), p. 2250-3153.