A New Extension of Shanker distribution with Real Life Data

C. Subramanian
Department of Statistics
Annamalai University, Annamalai nagar, Tamil Nadu, India
E-mail: manistat@yahoo.co.in

Aafaq A. Rather
Department of Statistics
Annamalai University, Annamalai nagar, Tamil Nadu, India
E-mail: aafaq7741@gmail.com

Abstract- The length biased distribution is a special case of weighted distribution. In this paper, we proposed a new probability model called as the length biased weighted Shanker (LBWS) distribution and discussed its various statistical properties. To estimate the unknown parameters, we use maximum likelihood method. For illustration and importance of the new distribution, we use two real lifetime data sets.

Keywords: Shanker distribution, Weighted technique, Structural properties, Maximum likelihood Estimation

1. INTRODUCTION

The concept of weighted distribution was given by Fisher (1934). Later it was modified by Rao (1965) in a unified manner, where by weighted distributions many situations can be solved. If the weight function considers the length only in units, then the weighted distribution reduces to length biased weighted distribution. Generally, The size-biased distribution is when the sampling mechanism selects the units with probability which is proportional to some measure of the unit size. The length biased version of weighted distributions are applied in various research areas related to biomedicine, reliability, ecology and branching processes. Further, Van Deusen (1986) fitting data related to diameter at breast height (DBH) arising from horizontal point sampling (HPS) in a size biased distribution. Also Lappi and Bailey (1987), use size biased distributions for analysing HPS diameter. Now recent times, Shanker & Shukla (2018) discussed a new generalized size-biased, Poisson-Lindley distribution with its applications to model size distribution. Rather and Subramanian (2018), discussed on characterization and estimation of length biased weighted generalized uniform distribution. Recently, Rather and Subramanian (2019), obtained the length biased erlang truncated exponential distribution with properties and its applications, which shows more flexibility than the classical distribution.

In this paper we consider the length biased weighted Shanker distribution. Shanker distribution was introduced by Rama Shanker (2015) is a newly proposed one parametric lifetime model for various engineering and medical science applications. The usefulness and importance of proposed distribution in lifetime data was better as compared to lindley and exponential distribution. The Shanker distribution has been modified and generalized by so many researchers like, Borah and Hazarika (2017), discussed discrete Shankar distribution and its derived distributions for modelling lifetime data. Later Shanker and Shukla in (2017) discussed weighted Shanker distribution and its applications to model lifetime data. In this paper, our motive is to prove that the length biased weighted version of Shanker distribution is more flexible and fits better in real lifetime data as compared to Shanker, Lindley and Exponential distributions.

The probability density function (pdf) of one parametric Shankar distribution is

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

and its corresponding cdf is

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

and its mean and variance is given by

$$\mu_1^{'}=\frac{\theta^2+2}{\theta(\theta^2+1)} \quad , \tag{2}$$

2. LENGTH-BIASED WEIGHTED SHANKER (LBWS) DISTRIBUTION

A non-negative random variable X is said to have weighted distribution, if the pdf of weighted random variable X_w is given by

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

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

For different weighted models, we have different choice of the weight function $w(x)$. When $w_x(x)=x$, the resulting distribution is termed as length-biased and its pdf is given by:

$$f_l(x)=\frac{xf(x)}{E(x)} \quad ; \qquad x>0 \tag{3}$$

By applying the weights x^c , where c = 1 to the weighted Shanker distribution we will get the Length biased weighted Shanker distribution.

Substitute the values of (1) and (2) in equation (3), we will get the pdf of Length biased weighted Shanker distribution.

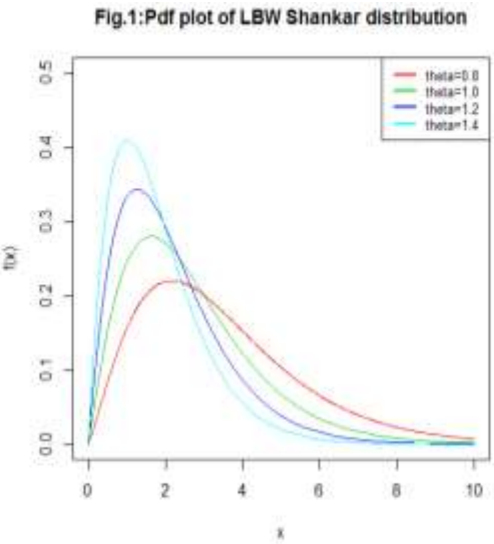
$$f_l(x)=\frac{x\theta^3(\theta+x)e^{-\theta x}}{\theta^2+2} \quad ; \qquad x>0, \theta>0 \tag{4}$$

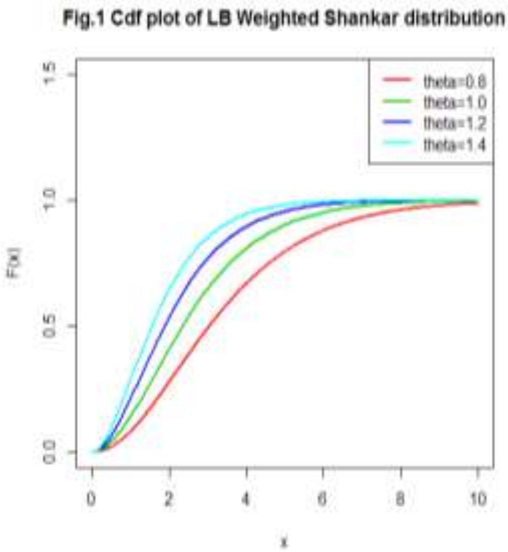
and the corresponding cdf is given by

$$F_l(x)=1-\frac{e^{-\theta x}}{(\theta^2+2)}\left\{(1+\theta x)(\theta^2+2)+\theta^2x^2\right\} \quad ; \qquad x>0, \theta>0 \tag{5}$$

where θ is the positive parameter.

Following below figures shows the graphical curves of pdf and cdf of length biased weighted shanker distribution on different values of parameter.





3. RELIABILITY ANALYSIS

In this section, we will discuss the reliability function, hazard rate and reverse hazard rate function for the proposed Length biased weighted Shanker distribution.

3.1 Reliability function

The reliability function is also known as the survival or survivor function and is defined as the probability that a system survives beyond a specified time. The reliability function R(x), is given by

$$R(x) = 1 - F(x)$$

The reliability function of Length biased weighted Shanker distribution is given by

$$R(x) = \frac{e^{-\theta x}}{(\theta^2 + 2)} \left\{ (1 + \theta x)(\theta^2 + 2) + \theta^2 x^2 \right\}$$

3.2 Hazard function

The hazard function h(x) is also known as the force of mortality, hazard rate or failure rate and is defined by

$$h(x) = \frac{f(x)}{1 - F(x)}$$

Therefore, the hazard rate function for Length biased weighted Shanker distribution is given by

$$h(x) = \frac{x\theta^3(\theta + x)}{(1 + \theta x)(\theta^2 + 2) + \theta^2 x^2}$$

Fig.3: Showing the Survival function curves of LBWSD

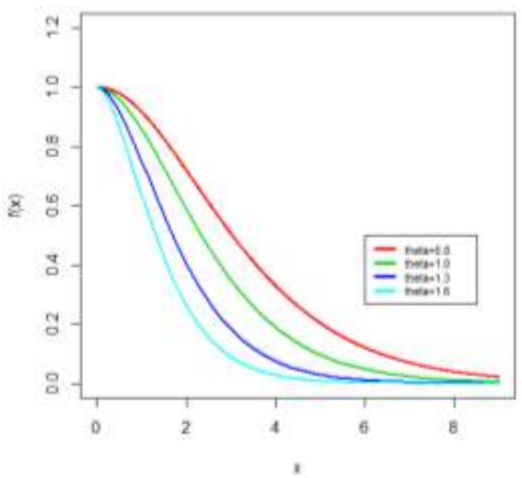
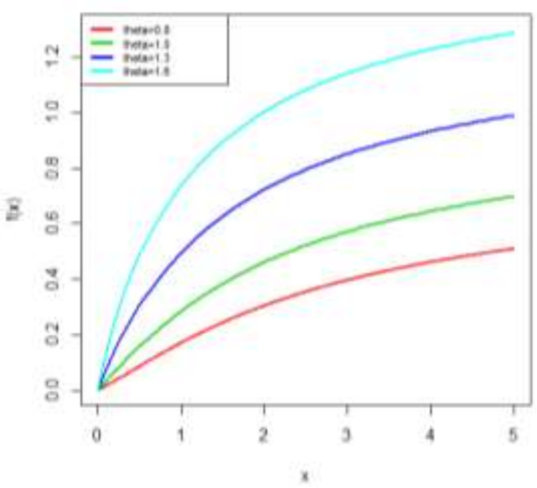


Fig.4: Showing the failure rate curves of LBWSD



3.3 Reverse Hazard rate

The reverse hazard function for Length biased weighted Shanker distribution is given by

$$h_r(x) = \frac{f(x)}{F(x)}$$

$$h_r(x) = \frac{x\theta^3(\theta + x)e^{-\theta x}}{(\theta^2 + 2) - e^{-\theta x}\{(1+\theta x)(\theta^2 + 2) + \theta^2 x^2\}}$$

4. STATISTICAL PROPERTIES

In this section, we will discuss the properties of Length biased weighted Shanker distribution like moments, moment generating function and characteristics function.

4.1 Moments

Let *X* be the random variable of the Length biased weighted Shankar distribution with parameter θ, then the r^{th} moment of LBWSD is given by

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

$$\mu_r' = \frac{\theta^3}{(\theta^2 + 2)} \int_0^\infty x^{r+1}(\theta + x)e^{-\theta x} dx$$

$$\mu_r' = \frac{\theta^3}{(\theta^2 + 2)} \left[\theta \int_0^\infty e^{-\theta x} x^{(r+2)-1} dx + \int_0^\infty e^{-\theta x} x^{(r+3)-1} dx \right]$$

$$\mu_r' = \frac{\theta^3}{(\theta^2 + 2)} \left[\theta \frac{\Gamma(r+2)}{\theta^{r+2}} + \frac{\Gamma(r+3)}{\theta^{r+3}} \right]$$

$$\mu_r' = \frac{\theta^2 \Gamma(r+2) + \Gamma(r+3)}{\theta^r(\theta^2 + 2)} \tag{6}$$

Put r = 1,2,3 and 4, in equation (6), we will get the first four moments

Issn No : 1006-7930

$$Mean = \mu_1' = \frac{2(\theta^2+3)}{\theta(\theta^2+2)}$$

$$\mu_2' = \frac{6(\theta^2+4)}{\theta^2(\theta^2+2)}$$

$$\mu_3' = \frac{24(\theta^2+5)}{\theta^3(\theta^2+2)}$$

$$\mu_4' = \frac{120(\theta^2+6)}{\theta^4(\theta^2+2)}$$

$$Variance = \mu_2 = \frac{2\{3(\theta^2+2)(\theta^2+4)-2(\theta^2+3)^2\}}{\theta^2(\theta^2+2)^2}$$

The Standard deviation (**σ**), Coefficient of variation (**C.V**.) and Index of dispersion (**γ**) are obtained as

$$\sigma = \frac{\sqrt{6(\theta^2+2)(\theta^2+4)-4(\theta^2+3)^2}}{\theta(\theta^2+2)}$$

$$C.V. = \frac{\sigma}{\mu_1'} = \frac{\sqrt{6(\theta^2+2)(\theta^2+4)-4(\theta^2+3)^2}}{2(\theta^2+3)}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{3(\theta^2+2)(\theta^2+4)-2(\theta^2+3)^2}{\theta(\theta^2+2)(\theta^2+3)}$$

4.2 Moment generating function and characteristic function of LBWS distribution

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} f_1(x;\theta)dx$$

Using Taylor Series

$$M_x(t) = \int_0^\infty \left(1 + tx + \frac{(tx)^2}{2!} + ...\right) f_1(x;\theta)dx$$

$$M_x(t) = \int_0^\infty \sum_{j=0}^\infty \frac{t^j}{j!} x^j f_1(x;\theta)dx$$

$$M_x(t) = \sum_{j=0}^\infty \frac{t^j}{j!} \mu_j'$$

$$M_x(t) = \sum_{j=0}^\infty \frac{t^j}{j!} \left(\frac{\theta^2\Gamma(j+2)+\Gamma(j+3)}{\theta^j(\theta^2+2)}\right)$$

$$M_x(t) = \frac{1}{(\theta^2+2)} \sum_{j=0}^\infty \frac{t^j}{j!} \left(\frac{\theta^2\Gamma(j+2)+\Gamma(j+3)}{\theta^j}\right)$$

Similarly, the characteristic function of LBWS distribution is given by

$$\phi_x(t) = M_x(it)$$

$$\phi_x(t) = \frac{1}{(\theta^2+2)} \sum_{j=0}^{\infty} \frac{(it)^j}{j!} \left(\frac{\theta^2 \Gamma(j+2) + \Gamma(j+3)}{\theta^j} \right)$$

5. ORDER STATISTICS

In this section, we will discuss the distributions of the order statistics of the length biased weighted Shanker distribution. Let $X_1, X_2, \ldots, X_n$ be the independent and identically distributed Length Biased Weighted Shanker random variables with parameter θ. Also, let $X_{(1)} < \ldots < X_{(n)}$ denotes the order statistics obtained from these n variables, then the pdf of r-th order statistics $X_{(r)}$, $1 \le r \le n$, for $x > 0$ is:

$$f_{X(r)}(x,\theta) = \frac{n!}{(n-r)!(r-1)!} f(x)[F(x)]^{r-1}[1-F(x)]^{n-r}. \tag{7}$$

For r =1, 2, ..., n.

Substitute the value of (4) and (5) in equation (7), we will get the pdf of r-th order statistics $X_{(r)}$ for LBWS distribution and is given by

$$f_{X(r)}(x) = \frac{n!}{(n-r)!(r-1)!} \frac{x\theta^3(\theta+x)e^{-\theta x}}{(\theta^2+2)} \left(1 - \frac{e^{-\theta x}}{(\theta^2+2)}\left((1+\theta x)(\theta^2+2)+\theta^2 x^2\right)\right)^{r-1} \times \left(\frac{e^{-\theta x}}{(\theta^2+2)}\left((1+\theta x)(\theta^2+2)+\theta^2 x^2\right)\right)^{n-r} \tag{8}$$

Using equation (8), the pdf of n-th order length biased weighted Shanker statistics $X_{(n)}$ is given by

$$f_{X(n)}(x) = \frac{nx\theta^3(\theta+x)e^{-\theta x}}{(\theta^2+2)} \left(1 - \frac{e^{-\theta x}}{(\theta^2+2)}\left((1+\theta x)(\theta^2+2)+\theta^2 x^2\right)\right)^{n-1}$$

Similarly, the pdf of 1st order statistics $X_{(1)}$ is given by

$$f_{X(1)}(x) = \frac{nx\theta^3(\theta+x)e^{-\theta x}}{(\theta^2+2)} \left(\frac{e^{-\theta x}}{(\theta^2+2)}\left((1+\theta x)(\theta^2+2)+\theta^2 x^2\right)\right)^{n-1}$$

6. PARAMETER ESTIMATION

In this section, we will discuss the maximum likelihood estimator which has been most widely used for estimating the parameter of the length biased weighted Shanker distribution. Let $X_1, X_2, X_3, \ldots, X_n$ be the random sample of size n from the length biased weighted Shanker distribution, then the likelihood function is given by

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

The log likelihood function becomes

$$\log L = 3n\log\theta + \sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\log(\theta + x_i) - \theta\sum_{i=1}^{n} x_i - n\log(\theta^2+2) \tag{9}$$

Differentiate (9) with respect to θ , which must satisfy the normal equation

$$\frac{\partial \log L}{\partial \theta} = \frac{3n}{\theta} - \frac{2n\theta}{(\theta^2 + 2)} + \frac{1}{\sum_{i=1}^{n}(\theta + x_i)} - \sum_{i=1}^{n} x_i = 0 \tag{10}$$

By solving equation (10), we will get the maximum likelihood estimator of the parameter of length biased weighted Shanker distribution. It is also more convenient to use Quasi-Newton or Newton-Raphson algorithms to maximize the log likelihood function (10).

7. TEST FOR LENGTH-BIASEDNESS OF LENGTH BIASED WEIGHTED SHANKER DISTRIBUTION

Let X_1, X_2, ... , X_n be a random sample from the length biased weighted Shankar distribution. To test the hypothesis

$$H_o : f(x) = f(x;\theta) \quad \text{against} \quad H_1 : f(x) = f(x;\theta)$$

For testing whether the random sample of size n comes from the Shanker distribution or Length biased weighted Shanker distribution, the following test statistic statistics is used

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{f_1(x;\theta)}{f_0(x;\theta)}$$

$$\Delta = \prod_{i=1}^{n} \left[\frac{x\theta^3(\theta + x)e^{-\theta x}}{(\theta^2 + 2)} \; \frac{(\theta^2 + 1)}{\theta^2(\theta + x)e^{-\theta x}} \right]$$

$$\Delta = \prod_{i=1}^{n} \left[\frac{\theta(\theta^2 + 1)}{(\theta^2 + 2)} x_i \right]$$

$$\Delta = \frac{L_1}{L_0} = \left[\frac{\theta(\theta^2 + 1)}{\theta^2 + 2} \right]^n \prod_{i=1}^{n} x_i$$

We reject the null hypothesis if

$$\left[\frac{\theta(\theta^2 + 1)}{\theta^2 + 2} \right]^n \prod_{i=1}^{n} x_i > k$$

Equivalently we reject the null hypothesis where

$$\Delta^* = \prod_{i=1}^{n} x_i > k^* , \quad \text{where } k^* = k\left(\frac{(\theta^2 + 2)}{\theta(\theta^2 + 1)} \right)^n > 0$$

For large sample size n, 2 logΔ is distributed as chi-square distribution with one degree of freedom and also p-value is obtained from the chi-square distribution. Thus we reject the null hypothesis, when the probability value is given by

$$p(\Delta^* > \alpha^*), \text{where } \alpha^* = \prod_{i=1}^{n} x_i \text{ , is the observed value of statistic } \Delta^* .$$

8. APPLICATIONS

In this section, we have fitted the two real lifetime data sets in length biased weighted Shanker distribution and both the data sets shows fit better than Shankar , Lindely and Exponential distribution.

Data set 1: Gross and Clark (1975) reported a set of data relating relief in minutes receiving analgesic of 20 patients. The data is given below:

1.1	1.4	1.3	1.7	1.9	1.8	1.6	2.2	1.7	2.7
4.1	1.8	1.5	1.2	1.4	3.0	1.7	2.3	1.6	2.0

Data set 2: The second data set is reported by Fuller et al (1994) which is related with strength data of window glass of the aircraft of 31 windows. The data are

18.83	20.80	21.657	23.03	23.23	24.05	24.321	25.50
25.52	25.80	26.69	26.77	26.78	27.05	27.67	29.90
31.11	33.20	33.73	33.76	33.89	34.76	35.75	35.91
36.98	37.08	37.09	39.58	44.045	45.29	45.381	

We will compare Length biased weighted Shanker distribution with Shanker, Lindley and Exponential by fitting these two data sets. By comparing these, we use the criteria like, -2 logL, AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and AICC (Akaike Information Criterion Corrected) for these two data sets mentioned in table 1.

Table 1: Fitted distributions of the two data sets and criteria for comparison

Data sets	Distribution	Parameter	-2 logL	AIC	AICC	BIC
1	**LB Shanker**	1.3142	50.2	52.2	52.4	53.2
	Shanker	0.8039	59.7	61.8	62.0	62.8
	Lindley	0.8161	60.5	62.5	62.7	63.5
	Exponential	0.5263	65.7	67.7	67.9	68.7
2	**LB Shanker**	0.0971	240.5	242.5	242.6	243.9
	Shanker	0.0647	252.3	254.3	254.5	255.8
	Lindley	0.0630	254.0	256.0	256.1	257.4
	Exponential	0.0325	274.5	276.7	276.7	277.9

Where, $AIC = 2k - 2\log L,\quad BIC = k\log n - 2\log L$ and $AICC = AIC + \frac{2k(k+1)}{(n-k-1)}$

k = number of parameters, n = sample size and 2logL is the maximized value of log-likelihood function.

It can be seen easily from table 1, that the Length biased weighted Shanker distribution gives the better fit as compared to Shanker, Lindley and Exponential distributions in both the two data sets.

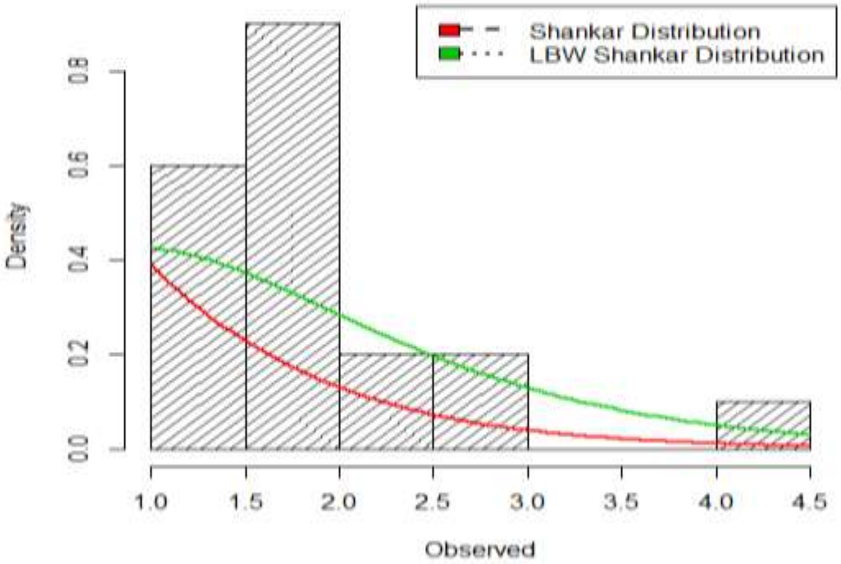


Fig.5, Shows the graphical curves of Shanker distribution and Length biased Shanker distribution of lifetime data set 1.

9. CONCLUSION

In this article we have studied a new distribution called as the length biased weighted Shanker distribution. The distribution has one parameter and the Shanker, Lindley and Exponential distributions are the particular cases of it. By certain special formulas, moments,failure rate, survival function, order statistics and the parameters have been estimated. The criterion for AIC, BIC and AICC has been studied and then compared with Shanker, Lindley and Exponential distributions. The real lifetime of two data sets has been fitted and the fit has been found better.

REFERENCES

[1] Borah, M. and Hazarika, J., (2017), Discrete Shanker distribution and its derived distributions. *Biom Biostat Int J* 5(4): 00140. DOI: 10.15406/bbij.2017.05.00140.
[2] Fisher, R. A., (1934), The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6, 13-25.
[3] Fuller, E. J., Frieman, S., Quinn, J., Quinn, G., Carter, W., (1994), Fracture mechanics approach to the design of glass aircraft windows: A case study. *SPIE Proc* 2286: 419-430.
[4] Gross, A. J. and Clark, V.A., (1975), Survival distributions: *Reliability Applications in the Biometrical Sciences*, John Wiley, New York.
[5] Lappi, J., Bailey, R.L.,(1987), Estimation of diameter increment function or other tree relations using angle-count samples, *Forest science* 33: 725-739.
[6] 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.
[7] Rather, A. A., and Subramanian, C., (2018), characterization and estimation of length biased weighted generalized uniform distribution, *International Journal of Scientific Research in Mathematical and Statistical Sciences*, Vol 5(5), pp. 72-76.
[8] Rather, A. A., and Subramanian, C., (2019), The length-biased erlang-truncated exponential distribution with life time data, *Journal of Information and Computational Science*, Vol 9(8), pp. 340-355.
[9] Shankar, R., (2015), Shanker Distribution and its Applications, *International Journal of Statistic and Applications,* 5(6):338-348.
[10] Shanker, R. and Shukla, K. K., Weighted Shanker distribution and its applications to model lifetime data, *Journal of Applied Quantitative Methods*, 12(2), (2017).
[11] Shanker, R., & Shukla, K. K., (2018), A Generalized Size-Biased Poisson-Lindley Distribution and Its Applications to Model Size Distribution of Freely-Forming Small Group, *Journal of Scientific Research*, *10*(2), 145.
[12] Van Deusen P.C., (1986), Fitting assumed distributions to horizontal point sample diameters. *For Sci* 32(1): 146-148.