

Harris Extended Burr XII Distribution and its Applications

Remya Sivadas^a and K.K.Jose^b

^a Department of Statistics, M.S.M College, Kayamkulam ^bHon. Director, School of Maths, Statistics & Data Analytics, Mahatma Gandhi University, Kottayam.

ARTICLE HISTORY

Compiled August 18, 2022

Received 05 November 2021; Accepted 19 June 2022

ABSTRACT

Aly and Benkherouf (2011) introduced a new method for generating a new class of distributions based on the probability generating function. Here we consider the Harris extended Burr XII distribution and explore their properties. Moments, quantiles, etc are derived. The model parameters are estimated by maximum likelihood method. Simulation is done to see the behavior of maximum likelihood estimates. Application is given to a real life data sets to compare our model with other computational models.

KEYWORDS

Sections; Harris distribution; Harris Extended Burr XII distribution; Hazard rate function; Marshall-Olkin family of distribution; Order statistics; Quantile function; Survival function.

1. Introduction

Aly and Benkherouf (2011) introduced a new method for generating a new class of distributions based on the probability generating function. By adding two new parameters to a baseline distribution and called the new family as the Harris extended family of distributions. Harris extended family contains the baseline distribution as a basic exemplar and gives more flexibility to distributions. Harris (1948) introduced a probability generating function

$$\phi(s, \theta, k) = \left[\frac{\theta s^k}{1 - \theta s^k} \right]^{\frac{1}{k}}$$

where $k > 0$ and $0 < \theta < 1$. When k is a positive integer then it can be put forwarded as a result of examining a simple discrete branching processes where a particle either splits into $(k+1)$ identical branches or remains the same for a short interval. When $k = 1$, ϕ reduces to the probability generating function of positive Geometric distribution.

Using survival function \bar{F} , the Harris family of survival functions is given by

$$\bar{G}(x, \theta, k) = \left[\frac{\theta(\bar{F}(x))^k}{1 - \bar{\theta}(\bar{F}(x))^k} \right]^{\frac{1}{k}}; 0 < \theta < \infty, k > 0.$$

This can be considered as a generalization of the Marshall-Olkin family of distributions introduced by Marshall and Olkin(1997). Sandhya et al.(2008) proposed Harris family of discrete distributions. Alice et al.(2003, 2005) discussed Marshall-Olkin Pareto Process and Marshall-Olkin semi-Weibull minification processes. Ghitany (2005) discussed Marshall-Olkin extended Pareto distribution and its application, El-Bassiouny et al.(2010) proposed Reliability properties of seven parameter Burr XII distribution, Jayakumar and Thomas (2008) discussed a generalization of Marshall-Olkin scheme and its applications to Burr type XII distribution. Nadarajah et al.(2012) introduced a new family of life time models. Surles et al.(2001) discussed inference for reliability and stress-strength for a scaled Burr X distribution. Surles et al.(2005) discussed some properties of a scaled Burr Type X distribution. Jose and Paul(2018) discussed Reliability test plans for percentiles based on the harris generalized linear exponential distribution. Gupta and Kundu (2001) discussed exponentiated exponential family as an alternative to gamma and Weibull distribution. Jose and Paul (2020) discussed Harris Generalized Rayleigh Distribution and its Applications in Industrial Reliability Test Plan.

This paper is organized as follows. Section 1.2 deals with the Harris extended family of distribution. Section 1.3 deals with the Harris Extended Burr XII Distribution. Moments are considered in section 1.4. In section 1.5 we discuss the estimation of parameters using maximum likelihood method. In section 1.6 we discuss simulation study. In section 1.7, we carry a data analysis and apply this model to a real data set. In section 1.8 offers the concluding remarks.

1.1. Harris Extended family

Let $F(x) = F(x; \eta)$ be a baseline cumulative distribution function and $\bar{F}(x) = 1 - F(x; \eta)$ be the corresponding survival function of a lifetime random variable X , where $\eta = (\eta_1, \eta_2, \dots, \eta_q)'$ is a parameter vector of dimension q . Let $f(x) = f(x, \eta)$ be the probability density function of X . The survival function of Harris extended family of distribution is given by

$$\bar{G}(x) = \left[\frac{\theta(\bar{F}(x))^k}{1 - \bar{\theta}(\bar{F}(x))^k} \right]^{\frac{1}{k}}; x > 0. \quad (1)$$

where $\bar{\theta} = 1 - \theta$, $\theta > 0$ and $k > 0$. The new parameters $\theta > 0$ and $k > 0$ are additional shape parameters to those in η and gives more flexibility.

The Harris extended probability density function is

$$g(x) = \frac{\theta^{\frac{1}{k}} f(x)}{[1 - \bar{\theta}(\bar{F}(x))^k]^{\frac{k+1}{k}}}; x > 0. \quad (2)$$

and the Harris extended failure rate function is

$$h(x) = \frac{rF(x)}{[1 - \bar{\theta}(\bar{F}(x))^k]} \quad (3)$$

where $rF(x)$ denotes the failure rate function of the baseline distribution.

When $k = 1$, the above equations reduces to Marshall-Olkin (1997) family of distributions. Therefore Harris extended family of distributions is a generalization of Marshall-Olkin family of distributions.

The Harris Extended probability density function can be expressed as an infinite linear combination of exponentiated base line survival function. Barreto-Souza et al. (2013) discussed general mathematical properties of Marshall-Olkin family of distributions. Similarly using the algebraic equations,

$$g(x) = f(x) \sum_{i=0}^{\infty} w_i (\bar{F}(x))^{ki} \quad (4)$$

for $\theta \in (0, 1)$, where $w_i = w_i(\theta, k) = \theta^{\frac{1}{k}} \bar{\theta}^i \frac{\Gamma(k^{-1}+i+1)}{\Gamma(k^{-1}+1)i!}$.

For $\theta > 1$

$$g(x) = f(x) \sum_{i=0}^{\infty} v_i (\bar{F}(x))^{ki} \quad (5)$$

where $v_i = v_i(\theta, k) = (-1)^i \theta^{-1} \sum_{j=i}^{\infty} \binom{j}{i} \left(\frac{\theta-1}{\theta}\right)^j \frac{\Gamma(k^{-1}+j+1)}{\Gamma(k^{-1}+1)j!}$ i.e,

$$g(x) = f(x) \begin{cases} \sum_{i=0}^{\infty} w_i (\bar{F}(x))^{ki}; & 0 < \theta < 1 \\ \sum_{i=0}^{\infty} v_i (\bar{F}(x))^{ki}; & \theta > 1 \end{cases}$$

(4) and (5) have the same representation except for the coefficients. The HE density function for any $\theta > 0$ can be expressed as the baseline density $f(x)$ multiplied by an infinite power series $\bar{F}(x)$. For more details see Batsidis and Lemonte(2014).

1.2. Harris Extended Burr XII distribution

Consider the baseline survival function of Burr XII distribution is

$\bar{F}(x) = (1 + x^c)^{-b}$, $x > 0, c > 0, b > 0$. Then consider the Harris Extended Burr XII distribution denoted as $HEBXII(\theta, k, c, b)$. The survival function given by

$$\bar{G}(x) = \left[\frac{\theta(1 + x^c)^{-kb}}{1 - \theta(1 + x^c)^{-kb}} \right]^{\frac{1}{k}} \quad (6)$$

The probability density function of HEBXII distribution is given by

$$g(x) = \frac{\theta^{\frac{1}{k}} b c x^{c-1} (1+x^c)^{-(b+1)}}{[1 - \bar{\theta}(1+x^c)^{-kb}]^{\frac{k+1}{k}}}; x > 0, c > 0, b > 0, \theta > 0, k > 0. \tag{7}$$

and the corresponding hazard rate function is given by

$$h(x) = \frac{bcx^{c-1}(1+x^c)^{-1}}{1 - \bar{\theta}(1+x^c)^{-kb}} \tag{8}$$

When $k=1$, (7) becomes Marshall-Olkin Burr XII distribution. When $\theta = 1, k=1$ it becomes Burr XII distribution.

Figure 1.1 and figure 1.2 shows the probability density function and hazard rate function of HEBXII distribution with different combinations of parameter values.

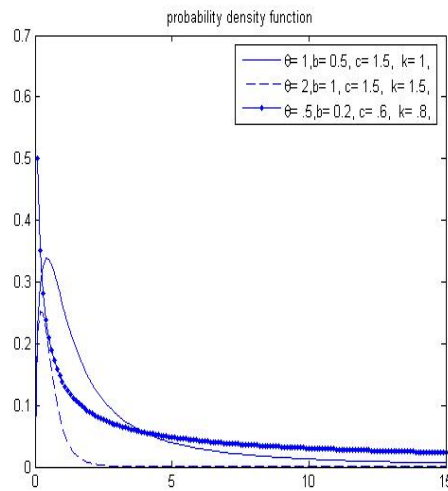


Figure 1.1

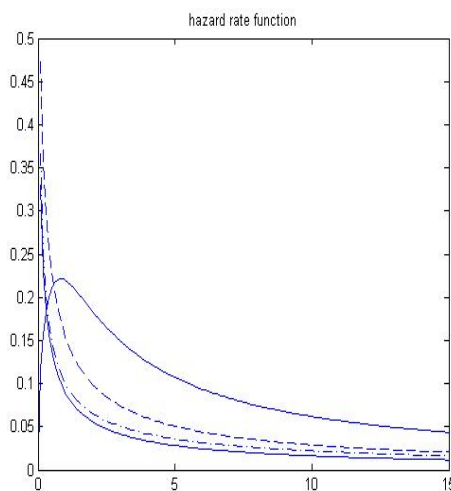


Figure 1.2

The quantile function of the HEBXII distribution is given by

$$x_p = \left\{ [\bar{\theta} + \theta(1-p)^{-k}]^{\frac{1}{kb}} - 1 \right\}^{\frac{1}{c}}$$

1.3. Moments

Consider the general expression for the moments of Harris Extended distribution in terms of the probability weighted moments of the base line distribution. The probability weighted moments introduced by Greenwood et al.(1979) are expectations of certain functions of a random variable whose mean exists. It can be defined as

$$\tau_{p,r} = \int_{-\infty}^{\infty} x^p (\bar{F}(x))^r f(x) dx.$$

From (4) and for $0 < \theta < 1$, the s^{th} moment of the Harris extended distribution can be expressed as

$$\mu'_s = \sum_{j=0}^{\infty} w_j \tau_{s,jk} \quad (9)$$

For $\theta > 1$, (9) holds replacing w_j with v_j .

Now we consider the moments of HEBXII distribution. Let $\tau_{s,jk}$ can be expressed as

$$\begin{aligned} \tau_{s,jk} &= \int_0^{\infty} x^s [\bar{F}(x)]^{jk} f(x) \\ &= bc \int_0^{\infty} x^{s+c-1} (1+x^c)^{-b(jk+1)-1} dx \\ &= bB\left(b(jk+1) - \frac{s}{c}, \frac{s}{c} + 1\right) \end{aligned}$$

Substituting in (9), we get

$$\mu'_s = \sum_{j=0}^{\infty} w_j bB\left(b(jk+1) - \frac{s}{c}, \frac{s}{c} + 1\right)$$

Now

$$\mu'_1 = b \sum_{j=0}^{\infty} w_j B\left(b(jk+1) - \frac{1}{c}, \frac{1}{c} + 1\right)$$

$$\mu'_2 = b \sum_{j=0}^{\infty} w_j B \left(b(jk+1) - \frac{2}{c}, \frac{2}{c} + 1 \right)$$

$$\mu_2 = \left[b \sum_{j=0}^{\infty} w_j B \left(b(jk+1) - \frac{2}{c}, \frac{2}{c} + 1 \right) \right] - \left[b \sum_{j=0}^{\infty} w_j B \left(b(jk+1) - \frac{1}{c}, \frac{1}{c} + 1 \right) \right]^2$$

1.4. Estimation of Parameters

Let x_1, x_2, \dots, x_n be a random sample of size n from HEBXII distribution with parameters (θ, k, η) . where $\eta = (\eta_1, \eta_2, \dots, \eta_q)'$ is a parameter vector of dimension q . Here $\eta = (\eta_1, \eta_2)' = (\eta_1(b), \eta_2(c))'$. Let $\xi = (\theta, k, \eta)'$ be the parameter vector. The log likelihood function for ξ based on a given random sample is

$$\begin{aligned} \log L(\xi) &= \frac{n}{k} \log \theta + \sum_{i=1}^n \log [bc(1+x_i^c)^{-(b+1)} x_i^{c-1}] - \\ &\quad \frac{k+1}{k} \sum_{i=1}^n \log [1 - \bar{\theta}(1+x_i^c)^{-bk}] \end{aligned}$$

Taking partial derivatives with respect to model parameters are

$$\frac{\partial \log L(\xi)}{\partial \theta} = \frac{n}{k\theta} - \frac{k+1}{k} \sum_{i=1}^n \frac{(1+x_i^c)^{-bk}}{[1 - \bar{\theta}(1+x_i^c)^{-bk}]}$$

$$\begin{aligned} \frac{\partial \log L(\xi)}{\partial k} &= \frac{-n \log \theta}{k^2} + \frac{1}{k^2} \sum_{i=1}^n \log [1 - \bar{\theta}(1+x_i^c)^{-bk}] + \\ &\quad \frac{\bar{\theta}(k+1)}{k} \sum_{i=1}^n \frac{(1+x_i^c)^{-bk} \log(1+x_i^c)^{-b}}{[1 - \bar{\theta}(1+x_i^c)^{-bk}]} \end{aligned}$$

$$\frac{\partial \log L(\xi)}{\partial \eta} = \sum_{i=1}^n \frac{\partial \log [f_0(x_i; \eta)]}{\partial \eta} + \bar{\theta}(k+1) \sum_{i=1}^n \frac{\bar{F}_0(x_i; \eta)^{k-1}}{1 - \bar{\theta} \bar{F}_0(x_i; \eta)^k} \frac{\partial \bar{F}_0(x_i; \eta)}{\partial \eta}$$

Therefore,

$$\frac{\partial \log L(\xi)}{\partial \eta_1(b)} = \sum_{i=1}^n \left[\frac{1}{b} - \log(1 + x_i^c) \right] + \bar{\theta}(k+1) \sum_{i=1}^n (1 + x_i^c)^{-b} \log(1 + x_i^c) \frac{(1 + x_i^c)^{-b(k-1)}}{[1 - \bar{\theta}(1 + x_i^c)^{-bk}]}$$

$$\frac{\partial \log L(\xi)}{\partial \eta_2(c)} = \sum_{i=1}^n \frac{1}{c} - (b+1) \frac{x_i^c \log x_i}{1 + x_i^c} + \log x_i - \bar{\theta}b(k+1) \sum_{i=1}^n \frac{(1 + x_i^c)^{-b(k-1)-b-1} x_i^c \log x_i}{1 - \bar{\theta}(1 + x_i^c)^{-bk}}$$

The maximum likelihood estimator $\hat{\xi} = (\hat{\theta}, \hat{k}, \hat{\eta}')'$ of $\xi = (\theta, k, \eta)'$ can be obtained by solving the equations $\frac{\partial \log L(\xi)}{\partial \theta} = 0$, $\frac{\partial \log L(\xi)}{\partial k} = 0$ and $\frac{\partial \log L(\xi)}{\partial \eta} = 0$. For that we use nlm software in R.

1.5. Simulation Study

A simulation study is conducted to test the performance of the mle's for estimating the parameters of HEBXII distribution. For this, we consider, $\theta = 2$, $k = 2.5$, $b = 0.2$ and $c = 1.5$. We simulate data from HEBXII model for different sample sizes $n = 200, n = 300, n = 400$ and $n = 500$ and calculate the mle by maximizing the likelihood function. We repeat the process 1000 times and we can see that as the sample size increases mean squared error (MSE) decreases. Thus we can conclude that the mle method performs good for estimating the parameters of HEBXII distribution. The results are listed in Table 1.

Table 1. Simulation Results: Average Estimates (AE), Biases and MSE's

n	Parameters	AE	Bias	MSE
200	t	-0.0831	-2.0832	4.3881
	k	3.1313	0.6313	0.8554
	b	1.2334	1.0338	2.4214
	c	2.6153	1.1153	1.5180
300	t	-0.3504	-2.0350	4.1548
	k	3.1937	0.6938	0.8263
	b	0.9759	0.7759	2.2966
	c	2.6712	1.1712	1.2632
400	t	-0.0119	-2.0119	4.0600
	k	3.2781	0.7781	0.7928
	b	0.7111	0.51111	1.3932
	c	2.6567	1.1567	1.2393
500	t	-0.0009	-2.0009	4.0525
	k	3.3430	0.8430	0.7200
	b	0.5128	0.3128	1.5351
	c	2.7873	1.2873	1.2233

1.6. Data Analysis

Here we consider the application of a data set. The data set is used to compare Harris Extended Burr type XII with Weibull Burr XII (WBXII) (Arslan Nasiret al., 2017) and Kumaraswamy BXII (KwBXII) (Paraniaba et al., 2013). To compare the goodness of fit test using information criteria, we use $AIC = -2 \log L + 2k$, $BIC = -2 \log L + k \log n$ and the Kolmogorov-Smirnov statistic, where k is the number of parameters and n is the sample size. The results of these applications are listed in Tables 2. These results show that the HEBXII distribution has the lowest AIC, BIC and K-S values and largest p-value of the K-S statistics among the fitted models. Hence we conclude that HEBXII distribution is a better model for this data set.

The data is taken from Lee and Wang (2003) and data represents the remission times (in months) of a random sample of 128 bladder cancer patients

0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

Table 2. Estimates, SE, AIC, BIC, K-S statistic and p value for the data set

Model	Parameter	Estimates	SE	AIC	BIC	K-S statistic	p value
HEBXII	θ	0.0315	0.1392	818.8546	827.9613	0.0162	0.9453
	k	0.6144					
	b	0.0999					
	c	0.1000					
WBXII	α	0.020	0.285	829.588	840.996	0.047	0.937
	β	3.4075	0.835				
	a	0.612	0.226				
	b	2.011	8.403				
KwBXII	α	12.022	6.837	829.962	841.370	0.0503	0.0503
	β	41.101	6.848				
	c	0.284	0.088				
	k	1.261	0.632				

1.7. Conclusion

In this paper we study in detail the Harris extended Burr XII distribution and its applications. Various properties of these distributions such as probability density function, hazard function, moments etc are obtained. Maximum likelihood estimates are derived and applied to a real data set on remission times (in months) of a random sample of 128 bladder cancer patients.

Acknowledgement

The authors are grateful to the Reviewer for the valuable comments which helped in improving the paper.

References

- [1] Mathai, A.M. and Davis, T.A. (1974). Constructing the sunflower head, *Mathematical Biosciences*, 20(1-2), 117-133. DOI: 10.1016/0025-5564(74)90072-8.
- [2] Mathai, A.M. and Pederzoli, G. (1977). *Characterizations of the Normal Probability Law*, John Wiley & Sons, New York.
- [3] Ahsanullah, M. (1995). *Record Statistics*. Nova Science Publishers. New York.
- [4] Ahsanullah, M. (1997). A Note on record values of uniform distribution, *Pakistan Journal of Statistics*, 13(1), 1-8.
- [5] Alice, T. and Jose, K.K. (2003). Marshall-Olkin Pareto Process, *Far East Journal of Theoretical Statistics*, 9, 117-132.
- [6] Alice, T. and Jose, K.K. (2005). Marshall-Olkin semi-Weibull minification processes, *Recent Advances in Statistical Theory and Applications*, 1, 6-17.
- [7] Aly, E.A.A. and Benkherouf, L.(2011). A new family of distributions based on probability generating functions, *Sankhya B*, 73, 70-80.
- [8] Balakrishnan, N. and Ahsanullah, M.(1994). Recurrence relations for single and product moments of record values from generalized Pareto distribution, *Communications in Statistics - Theory and Methods*, 23 (10), 2841- 2852.
- [8] Barreto-Souza, W., Lemonte, A.J. and Cordeiro, G.M.(2013). General results for the Marshall-Olkin's family of distributions. *Ann Braz Acad Science*, 85, 3-21.
- [9] Batsidis, A. and Lemonte, A.J. (2014). On the Harris extended family of distributions, *Statistics*, <http://dx.doi.org/10.1080/02331888.2014.969732>
- [10] Chandler, K. N. (1952). The distribution and frequency of record values, *J. Roy. Stat. Soci Ser. B*, 14, 220 -228.
- [11] Choulakian, V. and Stephens, M.A. (2001). Goodness of fit for the generalized Pareto distribution, *Technometrics*, 43, 478-484.
- [12] El-Bassiouny, A.H., Abdo, N.F. (2010). Reliability properties of seven parameters Burr XII distribution. *Comput Methods Sci Technol*, 16, 127–133.
- [13] Ghitany, M.E.(2005). Marshall-Olkin extended Pareto distribution and its application, *Int J Appl Math.*, 18, 17–31.
- [14] Greenwood, J.A., Landwehr, J.M., Matalas, N.C and Wallis, J.R.(1979). Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form, *Water Resour Res*, 1049-1054.
- [15] Gupta, R.D. and Kundu, D.(2001). Exponentiated exponential family: an alternative to Gamma and Weibull, *Biometrical Journal*, 43, 117-130.
- [16] Harris, T.E.(1948). Branching processes. *Annals of Mathematical Statistics*, 19, 474-494.
- [17] Jayakumar, K. and Thomas, M.(2008). On a generalization of Marshall-Olkin scheme and its application to Burr type XII distribution, *Statistical Papers*, 49, 421-439.
- [18] Jose, K.K., Krishna, E. and Ristic, M.M.(2014). On Record Values and Reliability Properties of Marshall-Olkin Extended Exponential Distribution, *Journal of Statistical Theory and Applications*, 13(3), 247-262.
- [19] Jose, K.K., Paul, A.(2018). Reliability test plans for percentiles based on the harris generalized linear exponential distribution, *Stochastics and Quality Control*, 33(1), 61–70.
- [20] Jose, K.K., Paul, A.(2020). Harris Generalized Rayleigh Distribution and its Applications in Industrial Reliability Test Plan, *International Journal of Statistics and Reliability Engineering* ISSN(P) 2350-0174, ISSN(O) 2456-2378, 3.
- [21] Marshall, A.W. and Olkin, I.(1997). A new method for adding a parameter to a family

- of distributions with application to the exponential and Weibull families, *Biometrika*, 84, 641-652.
- [22] Nadarajah, S., Jayakumar, K. and Ristić, M.M. (2012). A new family of lifetime models. *Journal of Statistical Computation and Simulation*, 1-16.
- [23] Nasir, A.M, Korkmaz, C.M, Jamal, F. and Yousof, M.H. (2018). On A New Weibull Burr XII Distribution for Lifetime Data, *Sohag J. Math*, 5(2), 1-10
- [24] Paranaíba, P.F., Ortega, E.M., Cordeiro, G.M., and Pascoa, M.A.D. (2013). The Kumaraswamy Burr XII distribution: theory and practice. *Journal of Statistical Computation and Simulation*, 83(11), 2117–2143.
- [25] Raqab, M.Z. (2001). Some results on the moments of record values from linear exponential distribution, *Math. Comput. Mode.*, 34, 1-8.
- [26] Sandhya, E., Sherly, S. and Raju, N (2008). Harris family of discrete distributions, in *Some Recent Innovations in Statistics*, Published by the Department of Statistics, University of Kerala, 57-72.
- [27] Saran, J. and Singh, S. K. (2008). Recurrence relations for single and product moments of the K -th record values from Linear Exponential Distribution and a characterization, *Asian Journal of Mathematical Statistics*, 1(3), 159-164.
- [28] Surles, J.G. and Padgett, W.J. (2001). Inference for reliability and stress-strength for a scaled Burr X distribution, *Lifetime Data Analysis*, 7, 187-200.
- [29] Surles, J.G. and Padgett, W.J. (2005). Some properties of a scaled Burr Type X distribution. *Journal of Statistical Planning and Inference*, 128, 271-280.