Contents list available http://www.kinnaird.edu.pk/

Journal of Natural and Applied Sciences Pakistan

Journal homepage: http://jnasp.kinnaird.edu.pk/

AREA-BIASED LOG-LOGISTIC DISTRIBUTION: PROPERTIES AND APPLICATION

Sidra Suhaib[1], Saleha Naghmi Habibullah[1] & Kalsoom Akhtar Chaudhry[1*]
[1]Department of Statistics, Kinnaird College for Women, Lahore, Pakistan.

Article Info
*Corresponding Author
Email: kalsoom.akhtar@kinnaird.edu.pk

Keywords
Area-Biased, Self-Inverse, Moments, Moment-Ratios, Kolmogrov Smirnov

Abstract

The development of new probability models is always a matter of interest for researchers and many techniques have been developed for the obtaining new probability distributions. The size-biasedness technique encompasses the development of length-biased and area-biased distributions. The objective of this research is to develop a new density function by applying area-biasing to the log-logistic distribution which is an important member of the class of self-inverse distributions. We derive a number of properties of the newly developed distribution including moments and moment ratios, hazard rate and survival function. In the last section of this paper, the area-biased log-logistic distribution is fitted to a real-life dataset for which an earlier research-work determined that the log-logistic distribution is a pretty well-fitting model. The Kolmogorov Smirnov test of goodness of fit indicates that the area-biased log-logistic distribution fits this data-set better than the original log-logistic distribution which points to the advantageousness of the newly derived density functions for purposes of modelling.

1. Introduction

The main intention of researchers in the area of distribution theory is to develop such distributions which yield better-fitting models for real-world data-sets. For this purpose, they endeavor to develop different approaches and techniques. A substantial amount of work has been carried out related to classes of distributions of non-negative continuous random variables. Size-biased distributions, also known as weighted distributions, are defined as follows:

$$f_{w}(x)=\frac{w(x)f(x)}{\mathrm{E}\left(w\left(X\right)\right)} \qquad (1.1)$$

Letting w(x) = x^2,we obtain area-biased distribution, which has been developed in the past. Bashir and Rasul (2016, 2018) found that the estimated parameters of Poisson area-biased Lindley distribution are biased and consistent asymptotically, and the distribution has its application in biological sciences. Parveen, Ahmed and Ahmad (2016) presented a new area-biased weighted Weibull distribution and compared the results with size-biased Rayleigh and Maxwell distribution and concluded that their proposed distribution fits better to data related to numerous fields of theoretical and applied sciences. Oluwafemi and Olalekan (2017) developed the length and area-biased exponentiated Weibull distribution based on forestinventories. Al-Omari, Al-Naseer and Ciavolino (2019) proposed size-biased Ishita distribution, derived various structural statistical properties, compared the model with other distributions and concluded that the newly developed distribution is the most appropriate model for ball bearings data. Shukla and Shankar (2021) introduced a new size-biased Poisson Ishita distribution relating to thunderstorm events and estimated its parameters by the method of moments and the maximum likelihood estimator approach. The proposed distribution was a better fit for thunderstorm events compared to other distributions. A new size-biased Zeghoudi distribution was proposed by Chouia, Zeghdoudi, Raman and Beghriche (2021), a number of properties were developed and comparisons were made with other distributions using real-life data and through simulation.

An interesting class is that which has been given the nomenclature ‘log-symmetric’ by Jones (2008) and the nomenclature ‘self-inverse’ by Habibullah and Saunders (2011). In this class, every distribution of a non-negative continuous random variable X is such that the pdf of X/A is identical to the pdf of A/X, where A is the median of the distribution. A special case is the one when the median is equal to unity so that the pdf of X is identical to the pdf of its reciprocal. Seshadri (1965) developed a functional equation for this type of distributions and presented a few examples of such distributions. Saunders (1974) presented a generalization of Seshadri (1965)’s reciprocal property for the normal family of distributions. Also, he proposed a new method for generating distributions possessing the reciprocal property. Habibullah (2012) presented the abbreviation ‘SIU’ for distributions self-inverse at unity.

Habibullah and Fatima (2015) adopted the abbreviation 'SIA' for distributions self-inverse at A. Habibullah (2017) proposed that the two nomenclatures be merged, and such distributions be called 'SIA log-symmetric distributions'. Habibullah and Xavier (2018) fitted the well-known Log-Logistic distribution, which is an SIA log-symmetric distribution, on bladder cancer data and, using the method of SIA-estimation, obtained a fit that turned out to be better than a number of distributions used by previous researchers for modelling that particular dataset. Nadeem and Habibullah (2018) obtained the life distribution of a parallel system consisting of identical components possessing Log-logistic life.

In this paper, we obtain a new density function by applying area-biasing to the log-logistic distribution. Moreover, a number of properties of the newly derived distribution are developed and the area-biased log-logistic distribution is fitted to a dataset related to relief time of patients.

2. Newly Derived Area-Biased Log-Logistic Distribution

The Log-Logistic distribution is:

$$f(x)=\frac{\left(\frac{\beta}{A}\right)\left(\frac{x}{A}\right)^{\beta-1}}{\left[1+\left(\frac{x}{A}\right)^{\beta}\right]^{2}}\quad 0<x<\infty,\ A>0,\ \beta>0 \qquad (1.2)$$

Now, using expression (1.1), newly derived area-biased log-logistic distribution is developed and is as follows

$$f(x;A,\beta)=\frac{\beta^{2}x^{\beta+1}\sin\left(\frac{2\pi}{\beta}\right)}{2A^{\beta+2}\pi\left[1+\left(\frac{x}{A}\right)^{\beta}\right]^{2}}\quad 0<x<\infty,\ A>0,\ \beta>0 \qquad (1.3)$$

The area-biased log-logistic distribution has a positively skewed shape that shows its applicability to the positively skewed datasets.

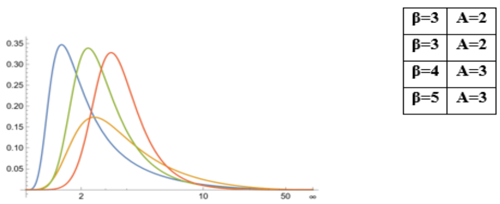


Figure 1.1: Density function of area-biased log-logistic distribution

The expression of the cumulative distribution function of the area-biased log-logistic distribution is given as:

$$F(x)=\frac{\left(\beta\sin\left(\frac{2\pi}{\beta}\right)(x)^{\beta+2}\right)\left(A^{\beta}(2+\beta)-2\left(A^{\beta}+x^{\beta}\right){}_2F_1\left(1,\frac{2+\beta}{\beta};2+\frac{2}{\beta};-A^{-\beta}x^{\beta}\right)\right)}{2\pi A^{\beta+2}(2+\beta)\left(A^{\beta}+x^{\beta}\right)} \quad , \qquad (1.4)$$

Where ${}_2F_1$ is the confluent hyper-geometric function. The graph of the CDF for different values of its parameters is given as follows

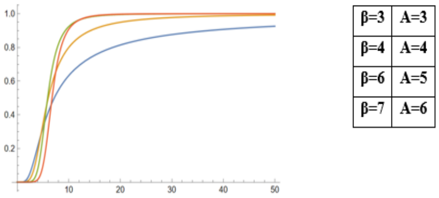


Figure 1.2: Cumulative function of area-biased log-logistic distribution

3. Some Properties of the Newly Derived Distribution

In this section, some properties of the newly derived area-biased log-logistic distribution are explored.

The expression of mean is

$$E(X)=\int_0^{\infty} x\frac{\beta^2 x^{\beta+1}\sin\left(\frac{2\pi}{\beta}\right)}{2A^{\beta+2}\pi\left[1+\left(\frac{x}{A}\right)^{\beta}\right]^2}dx$$

After substitution and simplification, we get the expression of mean

$$E(X)=\frac{3A\sin\left(\frac{2\pi}{\beta}\right)}{2\sin\left(\frac{3\pi}{\beta}\right)} \qquad (2.1)$$

The expression of the variance is

$$V(X)=A^2\left(\frac{4\sin\left(\frac{2\pi}{\beta}\right)}{\sin\left(\frac{4\pi}{\beta}\right)}-\frac{9\sin^2\left(\frac{2\pi}{\beta}\right)}{4\sin^2\left(\frac{2\pi}{\beta}\right)}\right) \qquad (2.2)$$

The higher order moment for the newly derived area-biased log-logistic distribution is:

$$E\left(X^{r}\right)=\frac{A^{r}(r+2)\sin\left(\frac{2\pi}{\beta}\right)}{2\sin\left(\frac{\pi(r+2)}{\beta}\right)} \tag{2.3}$$

By putting r=1, 2, 3 and 4 in expression (2.3) we get the first four moments about origin which is as follows:

$$\mu_1'=\frac{3A\sin\left(\frac{2\pi}{\beta}\right)}{2\sin\left(\frac{3\pi}{\beta}\right)} \tag{2.4}$$

$$\mu_2'=\frac{2A^{2}\sin\left(\frac{2\pi}{\beta}\right)}{\sin\left(\frac{4\pi}{\beta}\right)} \tag{2.5}$$

$$\mu_3'=\frac{5A^{3}\sin\left(\frac{2\pi}{\beta}\right)}{2\sin\left(\frac{5\pi}{\beta}\right)} \tag{2.6}$$

$$\mu_4'=\frac{3A^{4}\sin\left(\frac{2\pi}{\beta}\right)}{\sin\left(\frac{6\pi}{\beta}\right)} \tag{2.7}$$

The second moment about mean is obtained as

$$\mu_2=A^{2}\left[\frac{2\sin\left(\frac{2\pi}{\beta}\right)}{\sin\left(\frac{4\pi}{\beta}\right)}-\frac{9\sin^{2}\left(\frac{2\pi}{\beta}\right)}{4\sin^{2}\left(\frac{3\pi}{\beta}\right)}\right] \tag{2.8}$$

The third moment about mean is obtained after simplification

$$\mu_3=A^{3}\left[\frac{5\sin\left(\frac{2\pi}{\beta}\right)}{2\sin\left(\frac{5\pi}{\beta}\right)}-\frac{9\sin^{2}\left(\frac{2\pi}{\beta}\right)}{\sin\left(\frac{4\pi}{\beta}\right)\sin\left(\frac{3\pi}{\beta}\right)}+\frac{27\sin^{3}\left(\frac{2\pi}{\beta}\right)}{4\sin^{3}\left(\frac{3\pi}{\beta}\right)}\right] \tag{2.9}$$

The fourth moment about mean is obtained after simplification

$$\mu_4=A^{4}\left(\frac{3\sin\left(\frac{2\pi}{\beta}\right)}{\sin\left(\frac{6\pi}{\beta}\right)}-\frac{15\sin^{2}\left(\frac{2\pi}{\beta}\right)}{\sin\left(\frac{5\pi}{\beta}\right)\sin\left(\frac{3\pi}{\beta}\right)}+\frac{27\sin^{3}\left(\frac{2\pi}{\beta}\right)}{\sin\left(\frac{3\pi}{\beta}\right)\sin\left(\frac{4\pi}{\beta}\right)}-\frac{243\sin^{4}\left(\frac{2\pi}{\beta}\right)}{2\sin^{4}\left(\frac{3\pi}{\beta}\right)}\right) \tag{2.10}$$

Moment generating function of the area-biased log-logistic density is as follows

$$E\left(e^{tx}\right)=\frac{\beta^2\sin\left(\frac{2\pi}{\beta}\right)}{2A\pi}\int_0^{\infty}\left(1+tx+\frac{t^2x^2}{2!}+\ldots\right)\frac{\left(\frac{x}{A}\right)^{\beta+1}}{\left[1+\left(\frac{x}{A}\right)^{\beta}\right]^2}dx$$

After suitable substitution and simplification, the expression of moment generating function becomes

$$E\left(e^{tx}\right)=\frac{\beta\sin\left(\frac{2\pi}{\beta}\right)}{2\pi}\left[B\left(1-\frac{1}{\beta},1+\frac{1}{\beta}\right)+AtB\left(1-\frac{2}{\beta},1+\frac{2}{\beta}\right)+\frac{At^2}{\pi 2!}B\left(1-\frac{3}{\beta},1+\frac{3}{\beta}\right)+\cdots\right]$$

(2.11)

The mode is obtained by taking derivative of expression (1.3) and equating it to zero

$$\frac{\beta^2\sin\left(\frac{2\pi}{\beta}\right)\left(\left(1+\left(\frac{x}{A}\right)^{\beta}\right)^2(\beta+1)x^{\beta}-x^{\beta+1}2\left(1+\left(\frac{x}{A}\right)^{\beta}\right)\beta\left(\frac{x}{A}\right)^{\beta-1}\left(\frac{1}{A}\right)\right)}{2A^{\beta+1}\pi\left(1+\left(\frac{x}{A}\right)^{\beta}\right)^4}=0$$

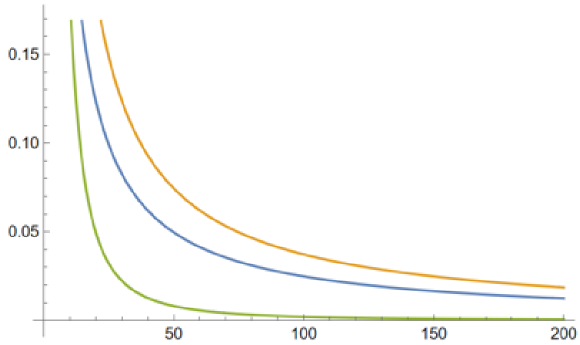
After mathematical manipulation, the expression of mode is as follows

$$\hat{x}=A\left(\frac{\beta+1}{\beta-1}\right)^{\frac{1}{\beta}} \qquad (2.12)$$

Survival function of the area-biased log-logistic distribution is obtained

$$S(x)=1-\frac{\left(\beta\sin\left(\frac{2\pi}{\beta}\right)(x)^{\beta+2}\right)\left(A^{\beta}(2+\beta)-2\left(A^{\beta}+x^{\beta}\right){}_2F_1\left(1,\frac{2+\beta}{\beta};2+\frac{2}{\beta};-A^{-\beta}x^{\beta}\right)\right)}{2\pi A^{\beta+2}(2+\beta)\left(A^{\beta}+x^{\beta}\right)}$$

(2.13)



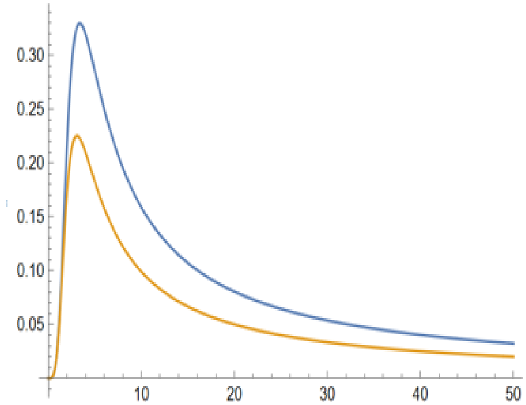
β=3	A=3
β=4	A=4
β=6	A=5

Figure 1.3: Survival function of area-biased log-logistic distribution

It can be observed from the survival graph that survival function is a decreasing function.

The hazard function of density is

$$h(x)=\frac{\dfrac{\beta^{2}x^{\beta+1}\sin\left(\dfrac{2\pi}{\beta}\right)}{2A^{\beta+2}\pi\left[1+\left(\dfrac{x}{A}\right)^{\beta}\right]^{2}}}{1-\dfrac{\left(\beta\sin\left(\dfrac{2\pi}{\beta}\right)(x)^{\beta+2}\right)\left(A^{\beta}(2+\beta)-2\left(A^{\beta}+x^{\beta}\right){}_{2}F_{1}\left(1,\dfrac{2+\beta}{\beta};2+\dfrac{2}{\beta};-A^{-\beta}x^{\beta}\right)\right)}{2\pi A^{\beta+2}(2+\beta)\left(A^{\beta}+x^{\beta}\right)}} \tag{2.14}$$



β=3	A=3
β=4	A=4

Figure 1.4: Hazard function of area-biased log-logistic distribution

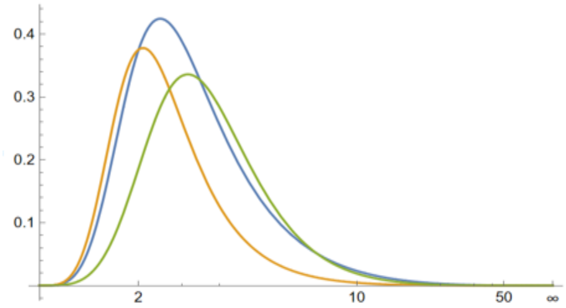
It can be observed from the hazard graph that hazard function has the increasing as well as the decreasing trend. The expression for Renyi Entropy is:

$$RI(\alpha)=\frac{1}{1-\alpha}\left(\alpha\log\left(\frac{\beta^{2}\sin\left(\dfrac{2\pi}{\beta}\right)}{2A\pi}\right)+\frac{A}{\beta}\log\left(B\left(1+\left((\alpha-1)-\frac{1}{\beta}(\alpha+\beta)\right),1+\left(\alpha-1+\frac{1}{\beta}\right)-1\right)\right)\right) \tag{2.15}$$

The distribution of order statistics for the area-biased log-logistic distribution is derived

$$f_{X(1)}(x)=n\left[1-\frac{\left(\beta\sin\left(\frac{2\pi}{\beta}\right)(x)^{\beta+2}\right)\left(A^{\beta}(2+\beta)-2\left(A^{\beta}+x^{\beta}\right){}_2F_1\left(1,\frac{2+\beta}{\beta};2+\frac{2}{\beta};-A^{-\beta}x^{\beta}\right)\right)}{2\pi A^{\beta+2}(2+\beta)\left(A^{\beta}+x^{\beta}\right)}\right]^{n-1}\left[\frac{\beta^{2}x^{\beta+1}\sin\left(\frac{2\pi}{\beta}\right)}{2A^{\beta+2}\pi\left[1+\left(\frac{x}{A}\right)^{\beta}\right]^{2}}\right]$$

(2.16)



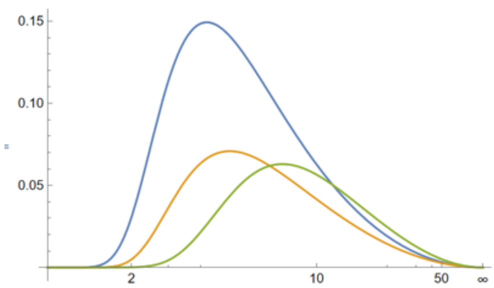
β=3	**A=2**
β=3	**A=2**
β=4	**A=3**

Figure 1.5: Graph of density of first order statistic of area-biased log-logistic distribution

The density of largest order statistic is as follows

$$f_{X(n)}(x)=n\left[\frac{\left(\beta\sin\left(\frac{2\pi}{\beta}\right)(x)^{\beta+2}\right)\left(A^{\beta}(2+\beta)-2\left(A^{\beta}+x^{\beta}\right){}_2F_1\left(1,\frac{2+\beta}{\beta};2+\frac{2}{\beta};-A^{-\beta}x^{\beta}\right)\right)}{2\pi A^{\beta+2}(2+\beta)\left(A^{\beta}+x^{\beta}\right)}\right]^{n-1}\left[\frac{\beta^{2}x^{\beta+1}\sin\left(\frac{2\pi}{\beta}\right)}{2A^{\beta+2}\pi\left[1+\left(\frac{x}{A}\right)^{\beta}\right]^{2}}\right]$$

(2.17)



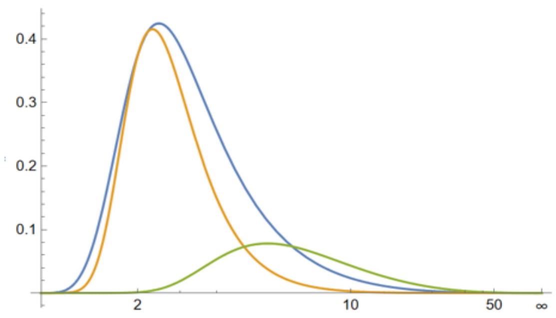
β=3	**A=2**
β=3	**A=2**
β=4	**A=3**

Figure 1.6: Graph of density of largest order statistic of area-biased log-logistic distribution

The density of median order statistic is as follows

$$f_{X(m)}(x)=\frac{n!}{(r-1)!(n-r)!}\cdot\frac{\beta^{2}x^{\beta+1}\sin\left(\frac{2\pi}{\beta}\right)}{2A^{\beta+2}\pi\left[1+\left(\frac{x}{A}\right)^{\beta}\right]^{2}}\times\left[\frac{\left(\beta\sin\left(\frac{2\pi}{\beta}\right)(x)^{\beta+2}\right)\left(A^{\beta}(2+\beta)-2\left(A^{\beta}+x^{\beta}\right){}_{2}F_{1}\left(1,\frac{2+\beta}{\beta};2+\frac{2}{\beta};-A^{-\beta}x^{\beta}\right)\right)}{2\pi A^{\beta+2}(2+\beta)\left(A^{\beta}+x^{\beta}\right)}\right]^{r-1}$$

$$\times\left[1-\frac{\left(\beta\sin\left(\frac{2\pi}{\beta}\right)(x)^{\beta+2}\right)\left(A^{\beta}(2+\beta)-2\left(A^{\beta}+x^{\beta}\right){}_{2}F_{1}\left(1,\frac{2+\beta}{\beta};2+\frac{2}{\beta};-A^{-\beta}x^{\beta}\right)\right)}{2\pi A^{\beta+2}(2+\beta)\left(A^{\beta}+x^{\beta}\right)}\right]^{n-r} \quad (2.18)$$



β=3	A=2
β=3	A=2
β=4	A=3

Figure 1.7: Graph of density of median order statistic of area-biased log-logistic distribution

4. Application to the Real Life Datasets

Area-biased log-logistic distribution is fitted on a real-life dataset and tested for the goodness of fit to explore its usefulness for modelling. This dataset consists of 20 observations which depict the relief times of patients. This dataset was firstly used by Gross and Clark (1975) and later by Chouia *et al*. (2021) for modelling. The dataset is presented 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

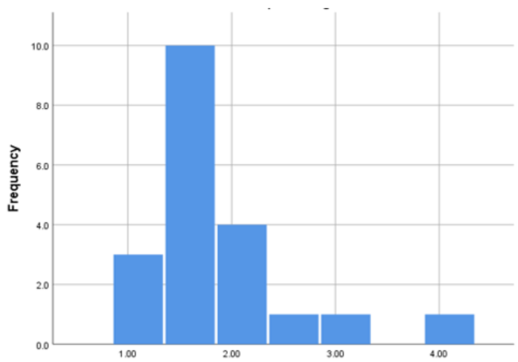


Figure 1.8:

Histogram of area-biased log-logistic distribution

5. Fitting of Log-Logistic Distribution

The mean and mode of the dataset are 1.90 and 1.70 respectively. Then for estimating the parameters of A and B. We have,

$$\frac{A\pi}{\beta\sin\left(\frac{\pi}{\beta}\right)}=1.90 \qquad (2.19)$$

$$A\left(\frac{\beta\text{-}1}{\beta+1}\right)^{\frac{1}{\beta}}=1.70 \qquad (2.20)$$

On simplification,

$$A=(0.6048)\beta\sin\left(\frac{\pi}{\beta}\right) \qquad (2.21)$$

Substituting A in (2.20), we get

$$\beta\sin\left(\frac{\pi}{\beta}\right)\left(\frac{\beta\text{-}1}{\beta+1}\right)^{\frac{1}{\beta}}=2.8108 \qquad (2.22)$$

Now, by inserting different values of β in (2.22), the expression (2.22) is achieved at β=5.75. Using the obtained value of β in (2.21), the value of parameter A comes out to be 1.81. Moreover, using Kolmogorov-Smirnov goodness of fit test, the maximum difference value D is 0.1580 whereas the critical value at 5% level of significance is 0.2940. The mean and mode are used for estimating the parameters of area-biased log-logistic distribution.

$$\frac{3A\sin\left(\frac{2\pi}{\beta}\right)}{2\sin\left(\frac{3\pi}{\beta}\right)}=1.90 \qquad (2.23)$$

$$A\left(\frac{\beta+1}{\beta\text{-}1}\right)^{\frac{1}{\beta}}=1.70 \qquad (2.24)$$

Using (2.24) and simplifying

$$A=(1.267)\frac{\sin\left(\frac{3\pi}{\beta}\right)}{\sin\left(\frac{2\pi}{\beta}\right)} \qquad (2.25)$$

Substituting the value of A in (2.25), we get

$$\frac{\sin\left(\frac{3\pi}{\beta}\right)}{\sin\left(\frac{2\pi}{\beta}\right)}\left(\frac{\beta+1}{\beta-1}\right)^{\frac{1}{\beta}}=1.3421 \qquad (2.26)$$

Now by plugging in different values of β in (2.26), the above expression exactly equal to 1.3421at β=7.84. Moreover, using Kolmogorov-Smirnov goodness of fit test, the maximum difference value D is 0.1072 whereas the critical value at 5% level of significance is 0.1129. Since, the maximum difference value of the area-biased log-logistic distribution is less than Log-Logistic distribution therefore, it is concluded that newly-derived area-biased distribution gives better fit compared to log-logistic distribution for this dataset.

6. Conclusion

In this paper, the area-biased log-logistic distribution is derived. Various properties of this distribution have been obtained including moments about origin, central moments, Renyi Entropy and distributions of order statistics. The shape of the density function is moderately positively skewed which indicates that the distribution should be appropriate for fitting to hump-shaped real-life datasets. As such, the area-biased log-logistic distribution is fitted to the relief time dataset with the result that the area-biased log-logistic distribution fitted the dataset better than the log-logistic distribution. This demonstrates the usefulness of this distribution for modeling purposes.

References

Al-Omari, A.I, Al-Naseer, A.D., & Ciavolino, E., (2019). A size biased Ishita Distribution and application to real data, *Quant*,53, 493-512.

Bashir, S. & Rasul, M. (2016). Some properties of the size-biased Janardan distribution, *American Journal of theoretical and applied statistics*, 5(5), 305-310.

Bashir, S. & Rasul, M. (2018). A Characterization and Recurrence relations of moments of the size-biased power function distribution by record values, *International Journal of statistics and Probability*, 7, No. 5, 1-8.

Chouia, S., Zeghdoudi, H., Raman, V., &Beghriche, A., (2021). A new size biased distribution with application. *Journal of Applied probability and Statistics*, 16(1), 111-125.

Habibullah, S.N. (2017). On the Power Generalization of Log-Symmetric SIA Distributions,*Journal of ISOSS*, 3(1), 45-54.

Habibullah, S.N. and Fatima, S.S. (2015). On a Newly Developed Estimator for More Accurate Modelling with an Application to Civil Engineering,*12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP12).*

Habibullah, S.N. and Saunders, S.C. (2011).A Role for Self-Inversion. Proceedings of *International Conference on Advanced Modeling and Simulation(ICAMS, Nov 28-30, 2011)* published by Department of Mechanical Engineering, College of Electrical and Mechanical Engineering, Islamabad, Pakistan, Copyright 2011, ISBN 978-869-8535-11-7.

Habibullah, S.N. and Xavier, K. (2018).Modeling bladder cancer patients' data through an SIA Log-Symmetric Distribution for achieving better fit.*Journal of Research and Reviews in Social Sciences Pakistan*, 1(2), 205-211.

Jones, M.C. (2008). On Reciprocal Symmetry.*Journal of Statistical Planning and Inference*, 138(10), 3039-3043.

Nadeem, S. and Habibullah, S.N. (2018). On the life distribution of a parallel system consisting of identical components possessing Log-logistic life,*16th international conference on Statistical Sciences*, March, 5-8, 2018.

Oluwafemi, O.S. and Olalekan, D. M. (2017). Length and area biased exponentiated Weibull distribution based on forest inventories, *Biometrics and Biostatistics International Journal*, 6(2), 311-320.

Parveen, Z., Munir,M., Ahmed, Z., & Ahmad, M. (2016). On area-biased weighted Weibull distribution, *Sci.Int.*(Lahore),28(4),3669-3679.

Saunders, S.C. (1974). A Family of Rando Variables Closed Under Reciprocation. J. Amer. Statist. Assoc., 69(346), 533-539.

Shukla, K. K. and Shankar, R. (2021).Size-biased Poison-Ishita distribution and its application to thunderstorm events.*Journal of Statistics Applications and Probability*, 10(1), 113-122.