A NEW GENERALIZATION OF SABUR DISTRIBUTION

Suvarna Ranade [1], Aafaq A. Rather[2,*]

•

[1,2]Symbiosis Statistical Institute, Symbiosis International (Deemed University), Pune-411004, India
[1]Dr. Vishwanath Karad MIT World Peace University, Pune
[1]suvarnamacs@gmail.com, [2,*]Corresponding Author: aafaq7741@gmail.com

Abstract

When the weight function depends on the lengths of the units of interest, the resulting distribution is called length biased. Length biased distribution is thus a special case of the more general form, known as weighted distribution. In this study, we introduce a novel probability distribution named the Length-Biased Sabur distribution (LBSD). This new distribution enhances the traditional Sabur distribution by incorporating a weighted transformation approach. The paper investigates the probability density function (pdf) and the cumulative distribution function (cdf) associated with the LBSD. A thorough examination of the distinctive structural properties of the proposed model is conducted, covering the survival function, conditional survival function, hazard function, cumulative hazard function, mean residual life, moments, moment generating function, characteristic function, likelihood ratio test, ordered statistics, entropy measures, and Bonferroni and Lorenz curve.

Key words: Sabur distribution, Length biased, Weighted transformation, Reliability analysis, Maximum likelihood estimator, Ordered statistics

1. Introduction

Weighted distributions occur when observations from a stochastic process are recorded with unequal probabilities, determined by a specific weighting function. When the weight function depends on the lengths of the units of interest, the resulting distribution is called length biased. Length biased distribution is thus a special case of the more general form, known as weighted distribution. The concept of length-biased distribution finds various applications in biomedical area such as family history and disease survival and intermediate events and latency period of AIDS due to blood transfusion [2]. The study of human families and wildlife populations was the subject of an article developed by Patil and Rao [8]. Patil, et al. [9] presented a list of the most common forms of the weight function useful in scientific and statistical literature as well as some basic theorems for weighted distributions and length-biased as special case. They arrived at the conclusion that the length biased version of some mixture of discrete distributions arises as a mixture of the length biased version of these distributions. Gupta R.D and Kundu D. [3] studied a new class of weighted exponential distribution which has applications in many fields such as: ecology, social and behavioural sciences and species abundance studies. Gupta R.C and Kirmani S. [2], studied the role of weighted distributions in stochastic modelling. Much work was done to characterize relationships between original distributions and their length biased version. A table for some basic distributions and their length biased forms is given by Patil and Rao [8] such as lognormal, Gamma, Pareto, Beta distribution. Khatree [4] presented a useful result by giving a relationship between the original random variable X and its length biased version Y. Recently Mudasir and S.P. Ahmad [5] studied the length biased Nakagami distribution. In subsequent years, Rather and Subramanian [10] explored the length-biased Erlang truncated exponential distribution, highlighting its practical applications, Rather and Ozel [11] introduced a new length-biased power Lindley distribution with applications.

2. Probability density function and cumulative distribution function

The probability density function (pdf) of the Sabur distribution with two parameters α and β is defined as

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

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

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))}, x > 0 \tag{2}$$

Where w(x) is a non-negative weight function and

$$\mathrm{E(w(x))} = \int w(x)f(x)\mathrm{dx} \ and \ w(x) = x^c$$

In this paper , we will consider the weight function was w(x) = x , where c=1 and using the definition of weighted distribution, the pdf of the LBSD on is given as

$$f_w(x) = \frac{x\, f(x)}{E(w(x))} \tag{3}$$

Expected value is defined as

$$E(x) = \int_0^\infty x\ f(x)dx$$

$$E\ (x) = \frac{(\alpha\beta+\beta^2+3)}{\beta(\alpha\beta+\beta^2+1)} \tag{4}$$

Substituting equation (1) and (3) in equation (2) we obtain the density function of LBSD as follows

$$f_w(x,\beta,\alpha) = \frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)} \tag{5}$$

And the cumulative density function (cdf) of LBSD is obtained by

$$F_w(x) = \int_0^x f_w(x)dx$$

$$F_w(x) = \int_0^x \frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)}\, dx \tag{6}$$

After simplification, the cdf of the LBSD is given by

$$F_w(x) = \frac{2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)}{2(\alpha\beta+\beta^2+3)} \tag{7}$$

Fig. 1 and Fig. 2 visually illustrates the pdf and cdf of LBSD.

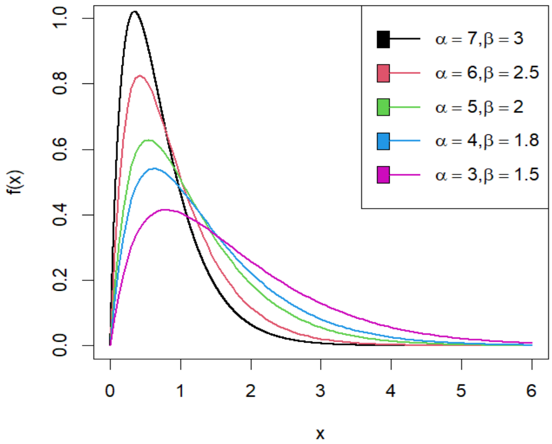

Fig. 1 : *pdf plot of LBSD*

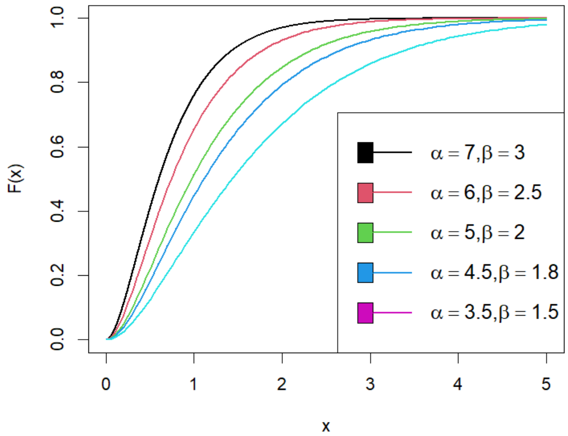

Fig. 2 : *cdf plot of LBSD*

3. Survival, Hazard and Reversed Hazard Functions

In this section we discuss about the survival function, hazard and reverse hazard functions of the LBSD. The survival function or the reliability function of is given by

$$S(x) = 1 - F_w(x) \tag{8}$$

$$S(x) = 1 - \left(\frac{2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)}{2(\alpha\beta+\beta^2+3)}\right) \tag{9}$$

The hazard function is also known as the hazard rate function, instantaneous failure rate or force of mortality and is given for the LBSD as

$$h(x) = \frac{f_W(x)}{s(x)} \tag{10}$$

$$h(x) = \frac{\frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)}}{1-\left(\frac{2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)}{2(\alpha\beta+\beta^2+3)}\right)} \tag{11}$$

$$h(x) = \frac{2x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2(\alpha\beta+\beta^2+3)-2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)} \tag{12}$$

The reverse hazard function of the LBSD is given by

$$h_r(x) = \frac{f_W(x)}{F_W(x)} \tag{13}$$

$$h_r(x) = \frac{2x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)} \tag{14}$$

Fig. 3 and Fig. 4 depicts the graphical survival function and Hazard function plot of LBSD.

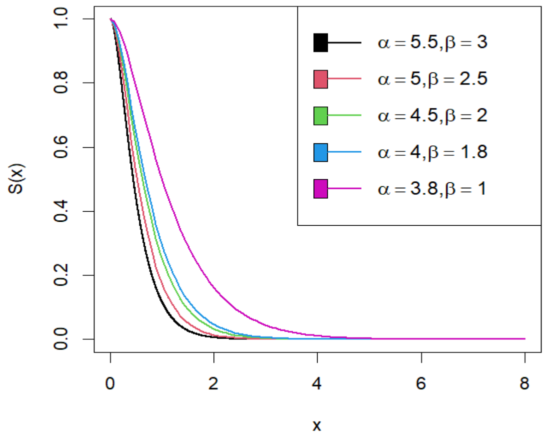

Fig. 3: *Survival function plot of LBSD*

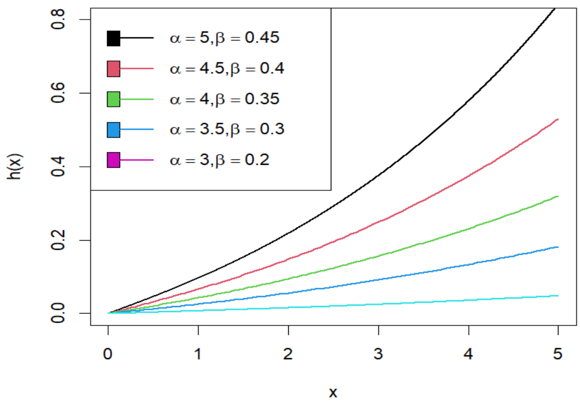

Fig. 4: *Hazard function plot of LBSD*

4. Structural properties

In this section we investigate various structural properties of the LBSD

Let X denote the random variable of LBSD with parameters α , β, then its r th order moment $E(x^r)$ about origin is given by

$$E(x^r) = \mu_r{}' = \int_0^\infty x^r f_w(x)\, dx \tag{15}$$

$$E(x^r) = \int_0^\infty x^r \frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)}\, dx \tag{16}$$

After simplifying the expression, we get

$$E(x^r) = \frac{2\beta(\alpha+\beta)\Gamma(r+2)+\Gamma(r+4)}{2\beta^r(\alpha\beta+\beta^2+3)} \tag{17}$$

Putting r=1, we get the expected value of LBSD as follows

$$E(x) = \frac{2\beta(\alpha+\beta)+12}{\beta(\alpha\beta+\beta^2+3)} \tag{18}$$

Put r=2, we obtained second moment as

$$E(x^2) = \frac{6\beta(\alpha+\beta)+60}{\beta^2(\alpha\beta+\beta^2+3)} \tag{19}$$

The variance of LBSD is calculated as

$$V(x) = E(x^2) - [E(x)]^2$$

$$V(x) = \frac{6\beta(\alpha+\beta)+60}{\beta^2(\alpha\beta+\beta^2+3)} - \left[\frac{2\beta(\alpha+\beta)+12}{\beta(\alpha\beta+\beta^2+3)}\right]^2 \tag{20}$$

4.1 Harmonic mean

The harmonic mean of the LBSD of random variable X can be written as

$$H = E\left(\frac{1}{x}\right) = \int_0^\infty \frac{1}{x} f_w(x)dx$$

$$H = \int_0^\infty \frac{1}{x} \frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)} dx \quad (21)$$

After simplification we get

$$\text{H} = \frac{\beta(\beta^2+\alpha\beta+1)}{(\beta^2+\alpha\beta+3)} \quad (22)$$

4.2 Moment generating function and characteristic function

Let X have a LBSD, then the Moment generating function of X is obtained as

$$M_X(t) = E(e^{tx}) = \int_0^\infty e^{tx} f_w(x)dx$$

Using Tayler's series, we obtain

$$M_X(t) = E(e^{tx}) = \int_0^\infty \left(1 + tx + \frac{(tx)^2}{2!} + ..\right) f_w(x)dx$$

$$M_X(t) = \int_0^\infty \sum_{i=0}^\infty \frac{t^i}{i!} x^i f_w(x)dx$$

$$M_X(t) = \sum_{j=0}^\infty \frac{t^j}{j!} E(x^j)dx$$

$$M_X(t) = \sum_{j=0}^\infty \frac{t^j}{j!} \; \frac{2\beta(\alpha+\beta)\Gamma j+2 + \Gamma j+4}{2\beta^j(\alpha\beta+\beta^2+3)} \quad (23)$$

Similarly, the characteristic function of LBSD of random variable X can obtained as

$$\Phi_X(t) = M_X(it) = \sum_{j=0}^\infty \frac{(it)^j}{j!} \frac{2\beta(\alpha+\beta)(\Gamma j+2\,)+(\,\Gamma j+4)}{2\beta^j(\alpha\beta+\beta^2+3)} \quad (24)$$

5. Likelihood Ratio Test

Let X1, X2, X3….be a random sample from the LBSD, we use the hypothesis

$H_0 : f(x) = f(x:\alpha,\beta)$ *against* $H_1: f(x) = f_w(x:\alpha,\beta,1)$

In order to test whether the random sample of size n comes from the Sabur distribution or weighted Sabur distribution, we will use following statistics

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^n \frac{f_w(x;\alpha,\beta,1)}{f(x;\alpha,\beta)}$$

$$\Delta = \prod_{i=1}^{n} \frac{x\beta(\alpha\beta+\beta^2+1)}{(\alpha\beta+\beta^2+3)} \tag{25}$$

$$\Delta = A^n \prod_{i=1}^{n} x_i \quad \text{where}$$

$$A = \frac{\beta(\alpha\beta+\beta^2+1)}{(\alpha\beta+\beta^2+3)} \tag{26}$$

We reject the null hypothesis, if

$$\Delta = A^n \prod_{i=1}^{n} x_i > \text{k}$$

$$\Delta^* = \prod_{i=1}^{n} x_i > \text{k}\, A^n$$

For large sample size n, 2log Δ 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

$$P(\Delta^* > a^*)$$

Where a^* is less than a specified level of significance and $\prod_{i=1}^{n} x_i$ is the observed value of the statistics Δ^*.

6. Entropy Measures

The concept of entropy is important in different areas such as probability and statistics, physics, communication theory and economics. Entropy measures quantify the diversity, uncertainty or randomness of a system. Entropy of a random variable X is measure of variation of the uncertainty.

6.1 Renyi Entropy

It was proposed by Renyi(1957). The Renyi entropy of order ξ for a random variable X is given by

$$e(\xi) = \frac{1}{1-\xi} \log\left(\int_0^\infty f^{\xi}(x)\, dx\right) \text{ where } \xi > 0 \text{ and } \xi \neq 1$$

$$e(\xi) = \frac{1}{1-\xi} \log\left(\int_0^\infty \left(\frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)}\right)^{\xi} dx\right) \tag{27}$$

After simplifying the equation we get

$$e(\xi) = \frac{1}{1-\xi} \log\left(\left(\frac{\beta^3}{(\alpha\beta+\beta^2+3)}\right)^{\xi} \sum_{i=0}^{\infty} \binom{\xi}{i} (\alpha+\beta)^{\xi-i} \left(\frac{\beta}{2}\right)^{i} \frac{\Gamma(\xi+2i+1)}{\beta\xi^{\xi+2i+1}}\right) \tag{28}$$

6.2 Tsallis Entropy

A generalization of Boltzman-Gibbs(B-G) statistical mechanics initiated by Tsallis has focussed a great deal to attention. This generalization of B-G statistics was proposed firstly by introducing the mathematical expression of Tsallis entropy (Tsallis, 1988) for a continuous random variable. Tsallis entropy of order λ of the weighted Sabur distribution is given by

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

$$S_\lambda = \frac{1}{\lambda-1}\left(1 - \int_0^\infty \left(\frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)}\right)^\lambda dx\right) \tag{29}$$

After simplifying the expression, we get

$$S_\lambda = \frac{1}{\lambda-1}\left[\left(1 - \left(\frac{\beta^3}{(\alpha\beta+\beta^2+3)}\right)^\lambda\right)\sum_{i=0}^\infty \binom{\lambda}{i}(\alpha+\beta)^{\lambda-i}\left(\frac{\beta}{2}\right)^i \frac{\Gamma(\lambda+2i+1)}{\beta\lambda^{(\lambda+2i+1)}}\right] \tag{30}$$

7. Order Statistics

Let $X_{(1)}, X_{(2)}, X_{(3)}\ldots X_{(n)}$ be the order statistics of a random sample $X_1, X_2, X_3 \ldots X_n$ drawn from the continuous population with pdf $f_x(x)$ and cdf $F_x(x)$ then the pdf of r th order statistic X (r) is given by

$$f_{x(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!} f_x(x)[F_x(x)]^{r-1}[1 - F_x(x)]^{n-r}$$

Substituting equation (4) and (5) in equation (6) , the pdf of order statistics X(r) of the weighted Sabur distribution is given by

$$f_{x(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!}\left(\frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)}\right) \times \left(\frac{2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)}{2(\alpha\beta+\beta^2+3)}\right)^{r-1}$$

$$\times\left(1 - \left(\frac{2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)}{2(\alpha\beta+\beta^2+3)}\right)\right)^{n-r} \tag{31}$$

Therefore, the pdf of the higher order statistics X(n) can be obtained as

$$f_{x(n)}(x) = n\left(\frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)}\right) \times \left(\frac{2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)}{2(\alpha\beta+\beta^2+3)}\right)^{n-1} \tag{32}$$

And the pdf of the first order statistics $X_{(1)}$ can be obtained as

$$f_{x(1)}(x) = n\left(\frac{x\beta^3\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(\alpha\beta+\beta^2+3)}\right) \times \left(1 - \left(\frac{2\beta(\alpha+\beta)\gamma(2,\beta x)+\gamma(4,\beta x)}{2(\alpha\beta+\beta^2+3)}\right)\right)^{n-1} \tag{33}$$

8. Income Distribution Curve

The Bonferroni and the Lorenz curves are not only used in economics in order to study the income and poverty, but it is also being used in other fields like reliability, medicine and demography. The Bonferroni and Lorenz curves are given by

$$B(p) = \frac{1}{p\mu_1{}'}\int_0^q x\,f(x)dx \quad \text{and}$$

$$L(p) = PB(p) = \frac{1}{\mu_1{}'}\int_0^q x\,f(x)dx$$

Here, we define the first raw moments as

$$\mu'_1 = \frac{2\beta(\alpha+\beta)+12}{\beta(\alpha\beta+\beta^2+3)} \tag{34}$$

And q = $F^{-1}(p)$, Then we have

$$B(p) = \frac{2\beta(\alpha+\beta)\gamma(3,\beta q)+\gamma(5,\beta q)}{2p(2\beta(\alpha+\beta)+12)} \tag{35}$$

$$L(p) = PB(p) = \frac{2\beta(\alpha+\beta)\gamma(3,\beta q)+\gamma(5,\beta q)}{2(2\beta(\alpha+\beta)+12)} \tag{36}$$

9. Estimation

We will discuss the maximum likelihood estimators (MLEs) of the LBSD. Consider $X_1, X_2, X_3 \ldots, X_n$ be the random sample of size n from the LBSD, then the likelihood function is given by

$$L(x;\ \alpha,\beta) = \prod_{i=1}^{n} x_i \frac{\beta^3\left(\alpha+\beta+\frac{\beta}{2}x_i{}^2\right)e^{-\beta x_i}}{(\alpha\beta+\beta^2+3)} \tag{37}$$

$$L(x;\ \alpha,\beta) = \frac{\beta^{3n}}{(\alpha\beta+\beta^2+3)^n}\prod_{i=1}^{n} x_i\left(\alpha+\beta+\frac{\beta}{2}x_i{}^2\right)e^{-\beta x_i} \tag{38}$$

The loglikelihood function is obtained as

$$Log\ L = 3nlog\beta - nlog(\alpha\beta+\beta^2+3) + log\sum x_i + \sum log(\alpha+\beta+\frac{\beta}{2}\sum x_i{}^2) - \beta\sum x_i \tag{39}$$

The MLEs of α, β can be obtained by differentiating Log L with respect to α, β and must satisfy the normal equation.

$$\frac{\partial log\ L}{\partial\beta} = -\frac{3n}{\beta} - \frac{n(\alpha+2\beta)}{(\alpha\beta+\beta^2+3)} + \sum_{i=1}^{n}\frac{\left(1+\frac{x_i{}^2}{2}\right)}{(\alpha+\beta+\frac{\beta}{2}x_i{}^2)} - \sum x_i = 0 \tag{40}$$

$$\frac{\partial log\ L}{\partial\alpha} = \left[-\frac{n\beta}{(\alpha\beta+\beta^2+3)} + \sum_{i=1}^{n}\frac{1}{\alpha+\beta+\frac{\beta}{2}\sum x_i{}^2}\right] = 0 \tag{41}$$

To obtain confidence interval we use the asymptotic normality results. We have that, if $\hat{\lambda} = (\hat{\alpha},\hat{\beta},\hat{c})$ denotes the MLE of $\lambda = (\alpha,\beta,c)$ we can state the results as follows

$$\left(\hat{\lambda} - \lambda\right) \to N_3\left(0, I^{-1}(\lambda)\right)$$

Where I(λ) is Fisher's Information matrix given by

$$I(\lambda) = -\frac{1}{n}\begin{pmatrix} E(\frac{\partial^2 logl}{\partial\alpha^2}) & E(\frac{\partial logl}{\partial\beta\partial\alpha}) \\ E(\frac{\partial logl}{\partial\beta\partial\alpha}) & E(\frac{\partial^2 logl}{\partial\beta^2}) \end{pmatrix} \tag{42}$$

Here we define

$$\frac{\partial^2 \log L}{\partial \beta^2} = -\frac{3n}{\beta^2} - n\left(\frac{-\alpha^2-2\alpha\beta-2\beta^2+6}{(\alpha\beta+\beta^2+3)^2}\right) - \frac{\left(1+\frac{1}{2}\sum x_i^2\right)^2}{(\alpha+\beta+\frac{\beta}{2}\sum x_i^2)^2} \tag{43}$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{n\beta}{(\alpha\beta+\beta^2+3)^2} - \frac{n}{(\alpha+\beta+\frac{\beta}{2}\sum x_i^2)^2} \tag{44}$$

$$\frac{\partial^2 \log L}{\partial \beta \partial \alpha} = -n\left(\frac{-\beta(\alpha+2\beta)-1}{(\alpha\beta+\beta^2+3)^2}\right) \tag{45}$$

10. Conclusion

In this paper, we introduce a novel extension of the Sabur distribution by incorporating a weighted transformation approach. This extension builds upon the existing two-parameter Sabur distribution, resulting in a three-parameter model known as the Length-Biased Sabur distribution. We conduct a comprehensive analysis of this new distribution, exploring its mathematical formulation and statistical properties in detail. Parameter estimation is performed using maximum likelihood estimation techniques.

References

[1] Alzaatreh, A., Famoye, F. and Lee, C. (2013). Weibull-pareto distribution and its applications, Communications in Statistics-Theory and Methods. 42(9): 1673-1691.

[2] Gupta, R.C. and Kirmani, S. (1990). The role of weighted distributions in stochastic modelling, Communications in Statistics, Theory and Methods. 19, 3147- 3162.

[3] Gupta, R.D. and Kundu, D. (2009). A new class of weighted exponential distribution, Statistics. 43, 621 - 634.

[4] Khatree, R. (1989). Characterization of Inverse-Gaussian and Gamma distributions through their length biased distribution. IEEE Transactions on Reliability. 38 (1) 610-611.

[5] Mudasir, S., & Ahmad, S. P. (2015). Structural Properties of Length-Biased Nakagami Distribution. International Journal of Modern Mathematical Sciences. 13, 217-227.

[6] Maxwell O, Chukwudike NC, and Bright OC (2019). Modelling lifetime data with the odd generalized exponentiated inverse Lomax distribution. Biom Biostat Int J. 8(2), 39–42.

[7] Oguntunde, P. E., Adejumo, A. O., Okagbue, H. I. and Rastogi M. K. (2016). Statistical properties and applications of a new lindley exponential distribution, Gazi University Journal of Science, 29(4): 831-838

[8]Patil, G.P. & Rao, C.R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families, Biometrics. 34, 179– 184.

[9] Patil, G. P., and Ord, J. K. (1976). On size-Biased Sampling and Related Form Invariant Weighted distribution, Sankhya: The Indian Journal of Statistics. 38, Series B, 48-61.

[10] Rather, A. A. & Subramanian, C. (2019), The Length-Biased Erlang–Truncated Exponential Distribution with Life Time Data, Journal of Information and Computational Science, vol-9, Issue 8, pp 340-355.

[11] Rather, A. A. &Ozel G., (2021): A new length-biased power Lindley distribution with properties and its applications, Journal of Statistics and Management Systems, DOI: 10.1080/09720510.2021.1920665.