THE WEIGHTED SABUR DISTRIBUTION WITH APPLICATIONS OF LIFE TIME DATA

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

In this paper, we propose a weighted version of Sabur distribution. The Stability of distribution are studied with structural properties, moments generating functions, likelihood ratio test, entropy measures, order statistics and Fisher's information matrix. The new model provides flexibility to analyse complex real data. Application of model on real data sets shows that the weighted Sabur distribution is quite effective. In this paper we utilize Monte Carlo simulation to evaluate the effectiveness of estimators. We used our weighted Sabur distribution on two real data set, Anderson-Darling and Cramer-von Mises class of quadratic EDF statistics utilize to test whether a given sample of data is drawn from a weighted Sabur distribution.

Key Words: Weighted distribution, Sabur distribution, Entropy, Order statistics.

1.Introduction

The concept of weighted distribution was first utilized by Fisher 1934 [8] in study of effect on form of distribution of recorded observations because of methods of ascertainment. The same concept was demonstrated and formulated by Rao 1965 [18] on modelling statistical data. The weighted distribution reduces to length biased distribution when the weight function considers only the length of units. The concept of length biased sampling was introduced by Cox [7] and Zelen [21]. Many newly introduced distributions along with their weighted versions exist in literature whose statistical behaviour is extensively studied during decades.

In recent years, researchers have made significant advancements in the study of the Lindley distribution and have proposed various one and two-parameter distributions to model complex datasets effectively. A notable contribution was made by Ghitney et al. [12], who conducted an extensive study on the Lindley distribution. They demonstrated that the Lindley distribution outperforms the exponential distribution when applied to modelling waiting times before bank customer service. Additionally, they highlighted that the contours of the hazard rate function for the Lindley distribution show an increasing trend, while the mean residual life function is a decreasing function of the random variable. Many authors modify the Lindley distribution by introducing new parameters and evaluating performance of these extended distribution with various dataset.

In this paper, we introduce a new distribution with three parameter, namely as weighted Sabur distribution with the hope that it provides more flexibility in various applications of Reliability, Survival Analysis, Biology etc.

2. Weighted Sabur Distribution

2.1 Density and Cumulative Density functions

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

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

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

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

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

$$E(w(x))=\int w(x)f(x)\,dx$$

.In this paper, we will consider the weight function was w(x) = xc, and using the definition of weighted distribution, the pdf of the weighted Sabur distribution is given as

$$f_w(x)=\frac{x^c f(x)}{E(x^c)}\ ,c>0 \quad (2)$$

Expected value is defined as

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

$$E(x^c)=\frac{\beta^2}{\alpha\beta+\beta^2+1}\left[\frac{\alpha+\beta}{\beta^{c+1}}\Gamma c+1+\frac{\Gamma c+3}{2\beta^{c+2}}\right] \quad (3)$$

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

$$f_w(x,\alpha,\beta)=\frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)(\Gamma c+1)+(\Gamma c+3)} \quad (4)$$

and the cumulative density function (cdf) of weighted Sabur distribution is obtained by

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

$$F_w(x)=\int_0^x \frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)(\Gamma c+1)+(\Gamma c+3)}\ dx \quad (5)$$

After simplification, the cdf of the weighted Sabur distribution is given by

$$F_w(x)=\frac{2\beta(\alpha+\beta)\gamma(c+1,\beta x)+\gamma(c+3,\beta x)}{2\beta(\alpha+\beta)(\Gamma c+1)+(\Gamma c+3)} \quad (6)$$

Fig.1 and Fig. 2 visually illustrates the pdf and cdf of Weighted Sabur Distribution.

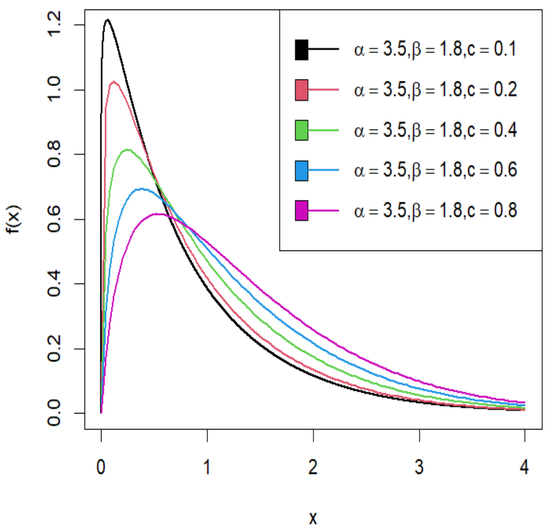

Fig. 1: pdf plot of Weighted Sabur distribution

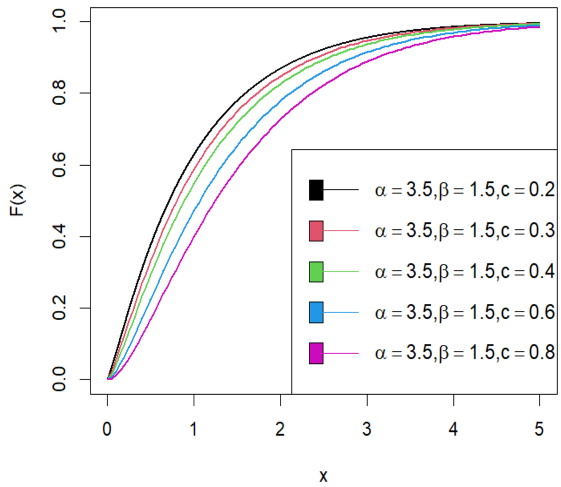

Fig. 2: pdf plot of Weighted Sabur distribution

2.2 Survival, Hazard and Reversed Hazard Functions

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

$$S(x) = 1 - F_w(x)$$

$$S(x) = 1 - \left(\frac{2\beta(\alpha+\beta)\gamma(c+1,\beta x)+\gamma(c+3,\beta x)}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}\right) \tag{7}$$

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

$$h(x) = \frac{f_w(x)}{s(x)}$$

$$h(x) = \frac{\dfrac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}}{1-\left(\dfrac{2\beta(\alpha+\beta)\gamma(c+1,\beta x)+\gamma(c+3,\beta x)}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}\right)} \tag{8}$$

$$h(x) = \frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{(2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)) - (2\beta(\alpha+\beta)\gamma(c+1,\beta x)+\gamma(c+3,\beta x))} \tag{9}$$

The reverse hazard function of the weighted Sabur distribution is given by

$$h_r(x) = \frac{f_w(x)}{F_w(x)}$$

$$h_r(x) = \frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)\gamma(c+1,\beta x)+\gamma(c+3,\beta x)} \tag{10}$$

Fig. 3 and Fig. 4 depicts the graphical survival function and Hazard function plot of Weighted Sabur distribution.

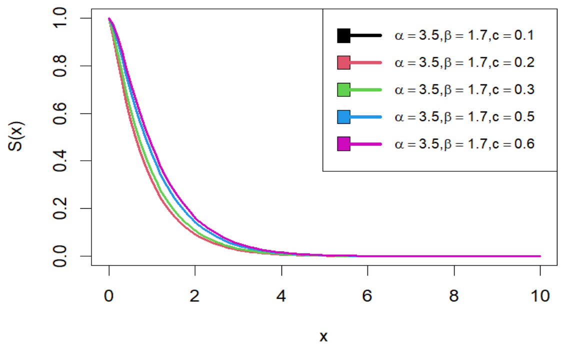

Fig. 3: Plots of Survival function

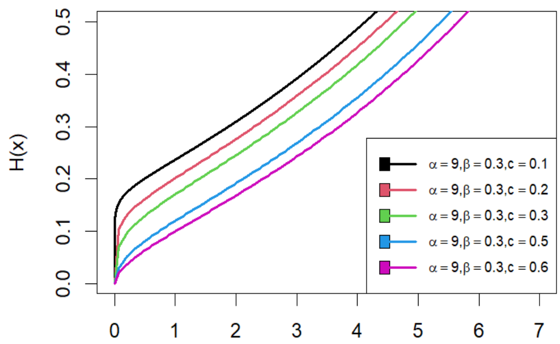

Fig. 4: Plots of hazard rate function

3. Structural properties

In this section we investigate various structural properties of the weighted Sabur distribution. Let X denote the random variable of weighted Sabur distribution with parameters α, β and c, then its r th order moment about origin is given by

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

$$E(x^r) = \int_0^\infty \mathrm{x}^{\mathrm{r}} \frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}\,dx \tag{11}$$

After simplifying the expression, we get

$$E(x^r) = \frac{[2\beta(\alpha+\beta)(\Gamma r+c+1)+(\Gamma r+c+3)]}{\beta^r[2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)]} \tag{12}$$

Putting r = 1, we get the expected value of weighted Sabur distribution as follows

$$E(x) = \frac{[2\beta(\alpha+\beta)(\Gamma c+2)+(\Gamma c+4)]}{\beta[2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)]} \tag{13}$$

Put r = 2, we obtained second moment as

$$E(x^2) = \frac{[2\beta(\alpha+\beta)(\Gamma c+3)+(\Gamma c+5)]}{\beta^2[2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)]} \tag{14}$$

The variance of Weighted Sabur distribution is calculated as

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

$$V(x) = \frac{2[\beta(\alpha+\beta)(\Gamma c+3)+(\Gamma c+5)]}{\beta^2[2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)]} - \left[\frac{[2\beta(\alpha+\beta)(\Gamma c+2)+(\Gamma c+4)]}{\beta[2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)]}\right]^2 \tag{15}$$

3.1 Harmonic mean

The harmonic mean of the weighted Sabur distribution of random variable x can be written as

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

$$H = \int_0^\infty \frac{1}{x} \frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)} dx \tag{16}$$

After simplification we get

$$\mathrm{H} = \frac{\beta[2\,\beta(\alpha+\beta)\Gamma c+\Gamma c+2]}{2\,\beta(\alpha+\beta)\Gamma c+1+\Gamma c+3} \tag{17}$$

3.2 Moment generating function and characteristic function

Let X have a weighted Sabur distribution, 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

$$M_X(t) = E(e^{tx}) = \int_0^\infty \left(1 + tx + \frac{(tx)^2}{2!} + ..\right) f_w(x)dx \tag{18}$$

$$M_X(t) = \int_0^\infty \sum_{i=0}^\infty \frac{t^i}{i!} x^i f_w(x)dx \tag{19}$$

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

$$M_X(t) = \sum_{j=0}^\infty \frac{t^j}{j!} \frac{2[\beta(\alpha+\beta)(\Gamma j+c+1)+(\Gamma j+c+3)]}{\beta^j[2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)]} \tag{20}$$

Similarly, the characteristic function of weighted Sabur distribution of random variable X can obtain as

$$\Phi_X(t) = M_X(it) = \sum_{j=0}^\infty \frac{(it)^j}{j!} \frac{2[\beta(\alpha+\beta)(\Gamma j+c+1)+(\Gamma j+c+3)]}{\beta^j[2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)]} \tag{21}$$

4. Likelihood Ratio Test

Let X_1, X_2, X_3….be a random sample from the weighted Sabur distribution, we use the hypothesis

$$H_0 : f(x) = f(x:\alpha,\beta) \;\; against \;\; H_1: f(x) = f_w(x:\alpha,\beta,c)$$

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

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^n \frac{f_w(x;\alpha,\beta,c)}{f(x;\alpha,\beta)} \tag{22}$$

$$\Delta = \prod_{i=1}^n x_i{}^c \frac{2\beta^c(\alpha\beta+\beta^2+1)}{2\beta(\alpha+\beta)\Gamma c+1 + \Gamma c+3} \tag{23}$$

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

$$A = \frac{2\beta^c(\alpha\beta+\beta^2+1)}{2\beta(\alpha+\beta)\Gamma c+1 + \Gamma c+3} \tag{24}$$

We reject the null hypothesis, if

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

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 $\prod_{i=1}^n x_i{}^c$ is the observed value of the statistics Δ^*.

5. 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.

5.1 Renyi Entropy

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

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

$$e(\xi) = \frac{1}{1-\xi}\log\left(\int_0^\infty \left(\frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}\right)^{\xi} dx\right) \tag{25}$$

After simplifying the equation, we get

$$e(\xi) = \frac{1}{1-\xi}\log\left(\left(\frac{2\beta^{c+2}}{2\beta(\alpha+\beta)\Gamma c+1 + \Gamma c+3}\right)^{\xi}\sum_{i=0}^{\infty}\binom{\xi}{i}(\alpha+\beta)^{\xi-i}\left(\frac{\beta}{2}\right)^{i}\frac{\Gamma(c\xi+2i+1)}{\beta\xi^{c\xi+2i+1}}\right) \tag{26}$$

5.2 Tsallis Entropy

A generalization of Boltzmann-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

$$S_\lambda = \frac{1}{\lambda-1}\left(1 - \int_0^\infty f^{\lambda}(x)\,dx\right) \tag{27}$$

$$S_\lambda = \frac{1}{\lambda-1}\left(1 - \int_0^\infty \left(\frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}\right)^{\lambda} dx\right) \tag{28}$$

After simplifying the expression, we get

$$S_\lambda = \frac{1}{\lambda-1}\left[\left(1 - \left(\frac{2\beta^{c+2}}{2\beta(\alpha+\beta)\Gamma c+1 + \Gamma c+3}\right)^{\lambda}\right)\sum_{i=0}^{\infty}\binom{\lambda}{i}(\alpha+\beta)^{\lambda-i}\left(\frac{\beta}{2}\right)^{i}\frac{\Gamma(c\lambda+2i+1)}{\beta\lambda^{c\lambda+2i+1}}\right] \tag{29}$$

6. Order Statistics

Let $X_{(1)}$, $X_{(2)}$, $X_{(3)}$…...$X_{(n)}$ be the order statistics of a random sample X_1,X_2,X_3…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

$$f_{x(r)}(x) = \frac{n!}{(r-1)!(n-r)!} f_x(x)[F_x(x)]^{r-1}[1-F_x(x)]^{n-r} \tag{30}$$

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

$$f_{x(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!}\left(\frac{2\beta^{c+2}x^c\left(\alpha+\beta+\frac{\beta}{2}x^2\right)e^{-\beta x}}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}\right)$$
$$\times\left(\frac{2\beta(\alpha+\beta)\gamma(c+1,\beta x)+\gamma(c+3,\beta x)}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}\right)^{r-1}$$
$$\times\left(1 - \left(\frac{2\beta(\alpha+\beta)\gamma(c+1,\beta x)+\gamma(c+3,\beta x)}{2\beta(\alpha+\beta)(\Gamma c+1) + (\Gamma c+3)}\right)\right)^{n-r} \tag{31}$$

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

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

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

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

7. 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

$$\mu'_1 = \frac{[2\beta(\alpha+\beta)\Gamma c+2+\Gamma c+4]}{\beta[2\beta(\alpha+\beta)\Gamma c+1+\Gamma c+3]} \quad (34)$$

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

$$B(p) = \frac{2\beta(\alpha+\beta)\gamma(c+2,Bq)+\gamma(c+4,Bq)}{p\left(2\beta(\alpha+\beta)\Gamma(c+2)+\Gamma(c+4)\right)} \quad (35)$$

$$L(p) = \frac{2\beta(\alpha+\beta)\gamma(c+2,Bq)+\gamma(c+4,Bq)}{(2\beta(\alpha+\beta)\Gamma(c+2)+\Gamma(c+4))} \quad (36)$$

8. Estimation

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

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

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

The loglikelihood function is obtained as

$$Log\ L = nlog2 + n(c+2)log\beta - nlog(2\beta(\alpha+\beta)\Gamma c+1+\Gamma c+3) + clog\sum x_i + \sum log(\alpha+\beta\frac{\beta}{2}\sum x_i{}^2) - \beta\sum x_i \quad (39)$$

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

$$\frac{\partial log\ L}{\partial\beta} = \frac{n(c+2)}{\beta} - \frac{n}{\beta} - \frac{2n}{\beta} + \frac{1+\frac{1}{2}\sum x_i{}^2}{\alpha+\beta+\frac{\beta}{2}\sum x_i{}^2} - \sum x_i = 0 \quad (40)$$

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

$$\frac{\partial log\ L}{\partial c} = nlog\beta - nlog\Psi(2\beta(\alpha+\beta)\Gamma c+1+\Gamma c+3) + log\sum x_i = 0 \quad (42)$$

Where $\Psi(.)$ is the digamma function. Because of the complicated form of the above likelihood equations, algebraically it is very difficult to solve the system of nonlinear equations. Therefore, we use R and Wolfram Mathematica for estimating the required parameters. 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

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

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

$$I(\lambda) = -\frac{1}{n}\begin{pmatrix} E\left(\frac{\partial^2 \log l}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 \log l}{\partial \alpha \partial \beta}\right) & E\left(\frac{\partial^2 \log l}{\partial \alpha \partial c}\right) \\ E\left(\frac{\partial^2 \log l}{\partial \beta \partial \alpha}\right) & E\left(\frac{\partial^2 \log l}{\partial \beta^2}\right) & E\left(\frac{\partial^2 \log l}{\partial \beta \partial c}\right) \\ E\left(\frac{\partial^2 \log l}{\partial c \partial \alpha}\right) & E\left(\frac{\partial^2 \log l}{\partial c \partial \beta}\right) & \left(\frac{\partial^2 \log l}{\partial c^2}\right) \end{pmatrix} \tag{43}$$

Here we define

$$\frac{\partial^2 \log L}{\partial \beta^2} = -\frac{(cn-n)}{\beta^2} - \frac{\left(1+\frac{1}{2}\sum {x_i}^2\right)^2}{(\alpha+\beta+\frac{\beta}{2}\sum {x_i}^2)^2} \tag{44}$$

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

$$\frac{\partial^2 \log L}{\partial c^2} = -n\Psi'\left((2\beta(\alpha+\beta)\Gamma c + 1 + \Gamma c + 3)\right) \tag{46}$$

$$\frac{\partial^2 \log L}{\partial c \partial \beta} = \frac{n}{\beta}$$

$$\frac{\partial^2 \log L}{\partial c \partial \alpha} = -n\Psi'\left((2\beta(\alpha+\beta)\Gamma c + 1 + \Gamma c + 3)\right) \tag{47}$$

$$\frac{\partial^2 \log L}{\partial \beta \partial \alpha} = -\frac{\left(1+\frac{\sum {x_i}^2}{2}\right)^2}{(\alpha+\beta+\frac{\beta}{2}\sum {x_i}^2)^2} \tag{48}$$

where $\Psi'(.)$ is the first order derivative of digamma function. Since λ being unknown, we estimate $I^{-1}(\lambda)$ by $I^{-1}(\hat{\lambda})$ and this can be used to obtain asymptotic confidence intervals for α, β, c.

9. Simulation

Simulations offer a comprehensive and flexible approach to comprehending the behaviour of maximum likelihood estimators across diverse sample sizes. This understanding serves as a valuable guide for enhanced decision-making, risk mitigation, and the enhancement of reliability and efficiency in statistical analysis within various domains, including finance, healthcare, and engineering. By utilizing simulations, we gain the ability to anticipate the behaviour of maximum likelihood estimators across a broad spectrum of sample sizes, even those challenging to attain in practical scenarios. This predictive capability aids in grasping how the bias, variance, and efficiency of the estimator evolve with fluctuations in sample size. Simulations play a crucial role in identifying the optimal sample size for the application of maximum likelihood estimators. Our investigation has delved into the performance of ML estimators across different sample sizes, namely n=25, 50, 75, 100, 200, and 300.

The inverse cumulative distribution function (cdf) technique was utilized for data simulation using the R-software, and this process was iterated 700 times to compute bias, variance, and mean squared error (MSE). Analysis of Table 1 reveals a consistent trend across various parameter values and sample sizes of the Weighted Sabur distribution, indicating a decrease in variance, bias, and MSE as the sample size increases. The diminishing bias suggests that Maximum Likelihood (ML) estimation approaches the true parameter values with an expanding sample size. Simultaneously, the declining variance indicates that the estimators exhibit increased precision and stability with larger sample sizes, displaying reduced variability across repeated simulations.

Table 1: *Estimation of Bias, Variance and MSE for different sample sizes*

n	$\beta=1$			$\alpha=1.5$			C=2		
	Bias	Variance	MSE	Bias	Var.	MSE	Bias	Var.	MSE
20	1.288	0.54226	2.20300	-1.4594	0.01567	2.1456	4.01555	11.02994	27.15457
30	1.452	0.36401	2.47277	-1.4999	0	2.2470	5.12954	7.216834	33.52903
50	1.520	0.21295	2.52622	-1.4999	0	2.2470	5.48211	5.147884	35.20144
75	1.400	0.08699	2.04928	-1.4999	0	2.2470	4.84093	0.868977	24.30359
100	1.626	0.19394	2.83889	-1.4999	0	2.2470	5.64059	4.11263	35.92895
200	1.458	0.11823	2.24448	-1.499	0	2.2470	5.08705	2.425165	28.30325
300	1.436	0.02933	2.09309	-1.499	0	2.2470	5.18147	0.397811	27.24549
	$\beta=1$			$\alpha=2$			C=1.2		
	Bias	Variance	MSE	Bias	Var.	MSE	Bias	Var.	MSE
20	1.775	1.155957	3.91724	-1.999	0	3.9960	3.91724	7.448396	22.79317
30	1.041	0.081519	1.16448	-1.9999	0	3.9603	3.67425	6.102347	19.60251
50	1.216	0.132173	1.61203	-1.9999	0	3.9960	2.96087	1.464504	10.23129
75	1.271	0.16421	1.78079	-1.999	0	3.9960	2.65069	1.099146	8.12533
100	1.229	0.102218	1.61407	-1.999	0	3.9960	2.68728	0.816670	8.038174
200	1.079	0.073882	1.23887	-1.9999	0	3.9960	2.24187	0.459616	5.485583
300	1.046	0.018731	1.11406	-1.9999	0	3.9960	2.30047	0.149156	5.441353
	$\beta=1$			$\alpha=1.5$			C=0.9		
	Bias	Variance	MSE	Bias	Var.	MSE	Bias	Var.	MSE
20	1.279	0.360584	1.99845	-1.999	0	3.9960	2.6694	3.71722	10.84287
30	1.161	0.211514	1.56024	-1.9999	0	3.9603	2.1697	1.086474	5.794367
50	1.199	0.202216	1.64052	-1.9999	0	3.9960	2.3040	1.318587	6.627044
80	1.139	0.170998	1.47002	-1.999	0	3.9960	2.0790	1.098943	5.421277
100	1.014	0.157921	1.18775	-1.999	0	3.9960	1.6903	0.799668	3.656934
200	0.976	0.094923	1.04934	-1.9999	0	3.9960	1.6897	0.579610	3.434744
300	0.918	0.025004	0.8694	-1.9999	0	3.9960	1.5269	0.128867	2.460371
	β= 1.5			α= 0.5			C=0.8		
	Bias	Variance	MSE	Bias	Var.	MSE	Bias	Variance	MSE
20	1.631	0.73142	3.3928	-0.499	0	0.2490	2.1509	2.752424	7.379041
30	1.413	0.61655	2.6138	-0.499	0	0.2490	1.7945	1.282839	4.503185
50	1.173	0.19589	1.5725	-0.499	0	0.2490	1.5009	0.681053	2.933925
75	1.516	0.26232	2.5634	-0.499	0	0.2490	1.9730	0.482302	4.375146
100	1.354	0.14775	1.9834	-0.499	0	0.2490	1.5451	0.351610	2.738881
200	1.210	0.05982	1.5258	-0.499	0	0.2490	1.4653	0.141682	2.288914
300	1.177	0.03884	1.4216	-0.499	0	0.2490	1.4216	0.101695	2.122741

9. Application

In this section we consider survival period data of 45 patients treated with chemotherapy only were made by Bekker et al.4 and Fulment et al.10. The data set are:

0.047, 0.115, 0.121, 0.132, 0.164, 0.197, 0.203, 0.260, 0.282, 0.296, 0.334, 0.395, 0.458, 0.466, 0.501, 0.507, 0.529, 0.534, 0.540, 0.641, 0.644, 0.696, 0.841, 0.863, 1.099, 1.219, 1.271, 1.326, 1.447, 1.485, 1.553, 1.581, 1.589, 2.178, 2.343, 2.416, 2.444, 2.825, 2.830, 3.578, 3.658, 3.743, 3.978, 4.003, 4.033.
The second data set represents the failure time of 50 items [13]
0.12, 0.43, 0.92, 1.14, 1.24, 1.61, 1.93, 2.38, 4.51, 5.09, 6.79, 7.64, 8.45, 11.9, 11.94, 13.01, 13.25, 14.32, 17.47, 18.1, 18.66, 19.23, 24.39, 25.01, 26.41, 26.8, 27.75, 29.69, 29.84, 31.65, 32.64, 35, 40.7, 42.34, 43.05, 43.4, 44.36, 45.4, 48.14, 49.1, 49.44, 51.17, 58.62, 60.29, 72.13, 72.22, 72.25, 72.29, 85.2, 89.52.

In order to compare the weighted Sabur distribution with the Erlang Truncated Exponential distribution, Exponential distribution, Power Lindley distribution, we consider the criteria like Bayesian information criterion (BIC), Akaike Information criterion (AIC), Akaike Information Criterion Corrected (AICC) and -2logL. The distribution having lower values of BIC, AIC, AICC and -2log L can be consider better. Along with this we calculate goodness of (GoF) metrics statistic Schwarz Information (SIC), Hannan-Quinn Information (HQIC) criteria, we also assess the Anderson–Darling (A*), Cramer-Von Mises (W*), Kolmogorov–Smirnov (K–S) statistic and associated P-value (PV). Table 2 and Table 3 represents parameter estimation of data set 1 and set 2 with GoF metrics. Figure 5 and Figure 6 shows the diagrammatic representation of density curve of data set 1 and data set 2. Figure 7 and Figure 8 shows the QQ plot of weighted Sabur distribution of data set 1 and data set 2.

$$AIC = 2k - 2logL, \qquad BIC = klogn - 2logL, AICC = AIC + \frac{2k(k+1)}{(n-k-1)}$$

Table 2 : *Parameter estimation and goodness of fit test statistics for survival data set 1*

Distributions and estimations	Weighted Sabur distribution	Erlang Truncated Exp distribution	Power Lindley Distribution	Exponential Distribution
MLE	$\hat{\alpha}$=0.00100 $\hat{\beta}$=2.09088 C=0.66942	$\hat{\beta}$= 1.140789 $\hat{\theta}$= 1.059767	$\hat{\beta}$= 0.9465414 $\hat{\theta}$=1.135077	$\hat{\theta}$= 0.7454655
SE	$\hat{\alpha}$=NaN $\hat{\beta}$=0.278 C=0.276	$\hat{\beta}$=70.36128 $\hat{\theta}$= 116.3132	$\hat{\beta}$= 0.107625 $\hat{\theta}$=0.146510	$\hat{\theta}$=0.111127
-2log L	96.2613	116.437	116.805	116.437
AIC	107.681	120.4372	120.8056	118.4372
BIC	107.681	124.0506	124.4189	120.2439
AICC	108.26666	120.72291	118.72291	120.89862
SIC	107.681	124.7229	124.4189	124.0506
HQIC	104.282	121.7842	122.1526	121.7842

A*	2.507	0.44535	0.56555	0.44535
W*	0.255	0.05897	0.08453	0.05897
K-S	0.255	0.09083 (0.819)	0.11044 (0.603)	0.16968 (0.1332)

Table 3: *Parameter estimation and goodness of fit test statistics for failure data set 2*

Distributions and estimations	Weighted Sabur distribution	Erlang Truncated Exp distribution	Power Lindley Distribution	Exponential Distribution
MLE	$\hat{\alpha}$=6.997702e+04 $\hat{\beta}$=.00332 C=.001	$\hat{\beta}$= 0.1437169 $\hat{\theta}$= 0.2621020	$\hat{\beta}$= 0.72807201 $\hat{\theta}$=0.07252873	$\hat{\theta}$= 0.033142
SE	$\hat{\alpha}$=1.186329e+04 $\hat{\beta}$=.0074 C=.0176	$\hat{\beta}$= 1.9452922 $\hat{\theta}$= 4.0561881	$\hat{\theta}$=0.16807702 $\hat{\theta}$=0.04478477	$\hat{\theta}$=0.004683
-2log L	440.7604	440.7133	442.1001	440.7133
AIC	446.7604	444.7133	446.1001	442.7133
BIC	452.4965	448.5374	449.9242	444.6253
AICC	447.28214	444.96861	446.35542	446.35541
SIC	452.4965	448.5374	449.9242	448.5374
HQIC	448.9448	446.1695	447.5563	446.1695
A*	0.8861	0.8852	0.9367	169
W*	0.11426	0.13171	0.13959	8.571
K-S	0.11426 (0.495)	0.11434 (0.4947)	0.13959 (0.4631)	0.6958 (8.882e-16)

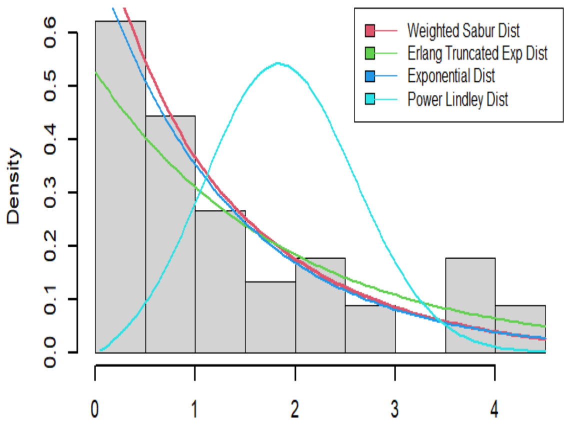

Fig. 5: Fitting density curves of data set 1

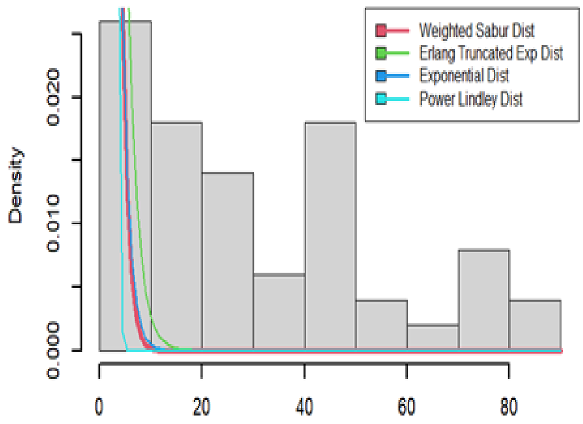

Fig. 6: Fitting density curves of data set 2

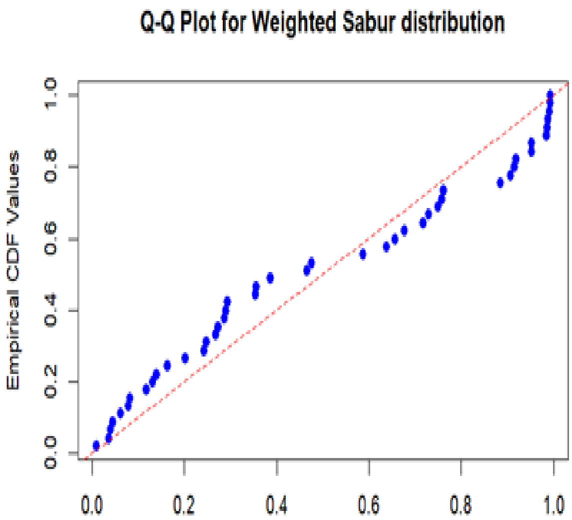

Fig. 7: Custom cdf values for data set 1

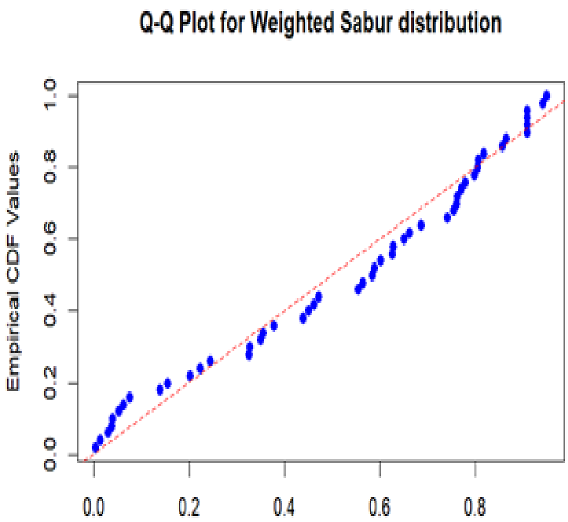

Fig.8 : Custom cdf values for data set 2

10. Conclusion

This paper introduces the weighted Sabur distribution with three parameters, a novel extension of the Sabur distribution, and explores its comprehensive statistical properties. The model parameters are estimated using maximum likelihood estimation, incorporating a weighted approach to enhance precision. The analysis encompasses various mathematical aspects and reliability measures, including the hazard rate function, to evaluate the distribution's performance as a lifetime model. Additionally, we benchmark the weighted Sabur distribution against other established distributions such as the exponential, power Lindley, and Erlang truncated exponential, using two sets of real-world data for validation. This comparative analysis confirms the potential of the weighted Sabur distribution as a robust and versatile model for lifetime data analysis.

References

[1] Ahmad Aijaz, Rather A. A., Gemeay A. M., Nagy M., Sapkota L. P., Mansi A. H. (2024), Novel Sin-G Class of Distributions with an Illustration of Lomax Distribution: Properties and Data Analysis, *AIP Advances*, Vol. 14, No. 035132, pp 1-17.

[2] Aijaz, A., Ahmad, A, Tripathi, R. (2020), Transmuted inverse Lindley distribution: Statistical properties Applications, *Science, technology and development*, 9(7), 1-10.

[3] Aijaz, A., Ahmad, Afaq Ahmad, Aafaq A. Rather, The Sabur Distribution: Properties and application related to engineering data, *RT&A*, No 4 (76), pp 882-892.

[4] Aijaz. A, Jallal. M, Qurat Ul Ain.S, Tripathi . R (2020). The Hamza distribution with Statistical properties and Applications. *Asian Journal of Probability and Statistics*, 8(1):28-42.

[5] Bekker, A., Roux, J. J. J. & Mostert, P. J. A generalization of the compound Rayleigh distribution: Using a Bayesian method on cancer survival times. *Commun. Stat. Theory Methods* 29(7), 1419–1433 (2000).

[6] Bhaumik, D. K., Kapur, K. and Gibbons, R. D. (2009). Testing parameters of a gamma distribution for small samples, *Technimetrics*, 51(3), 326-334.

[7] Cox D. R. (1969). Some sampling problems in technology, In New Development in Survey Sampling, Johnson, N. L. and Smith, H., Jr .(eds.) New York, *Wiley- Inte sciences*, 506-527

[8] Fisher, R.A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics*, 6, 13- 25.

[9] Flaih A., Elsalloukh, H., Mendi, E., Milanova, M., 2012. The exponentiated inverted Weibull distribution. *Appl. Math. Inf. Sci* 6 (2), 167–171.

[10] Fuller E.J., Friema S., Quinn, J., Quinn, G., and Carter, W. (1994): Fracture mechanics approach to the design of glass aircraft windows: A case study, *SPIE* Proc 2286, 419-430.

[11] Fulment A. K., Gadde, S. R. & Peter, J. K. Te odd log-logistic generalized exponential distribution: Application on survival times of chemotherapy patients' data. F1000research 11, 1444 (2023). 30. Aarset, M. V. How to identify a bathtub hazard rate. *IEEE* Trans. R

[12] Ghitany, M.E., Atieh B., Nadarajah S., (2008). Lindley distribution and its applications, Mathematics Computing and Simulation, 78, 493-506

[13] Gross, A. J., Clark, V. A., (1975). Survival Distributions: Reliability Applications in the Biometrical Sciences, *John Wiley*, New York

[14] Lai, C.D.; Murthy, D.; Xie, M. Weibull Distributions and Their Applications. In Springer Handbook of Engineering Statistics; Pham, H., Ed.; *Springe*r: London, UK, 2006.

[15] Lindley, D, V. (1958). Fiducial distributions and Bayes' theorem, *Journal of Royal Statistical Society,* Series B, 20, 102-107.

[16] Merovci, F., (2013) Transmuted Lindley distribution, *International Journal of Open problems in Computer Science and Mathematics*, 6, 63-72.

[17] Nadarajah, S., BAKOUCH, H. S., TAHMASBI, R., (2011). A Generalized Lindley distribution, *Sankhya Series* B, 73, 331-359.

[18] Rao, C. R. (1965), ON Discrete Distributions Arising out of Methods of Ascertainment. *Sankhya*; The Ind. J. Statist, Series A, 27, 311-324.

[19] 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

[20] Rather, A. A. and Subramanian, C. (2020), A New Exponentiated Distribution with *Engineering Science Applications, J. Stat. Appl*. Pro. 9, No. 1, pp 127-137.

[21] Zelen, M. (1974). Problems in cell kinetic and the early detection of disease, in Reliability and Biometry, F. Proschan & R. J. Sering, eds, *SIAM*, Philadelphia, 701-706