Pak. J. Statist.
2024 Vol. 40(2), 173-197

PROPERTIES AND APPLICATIONS OF THE TRANSFORMED SINE DAGUM DISTRIBUTION

Anuwoje Ida L. Abonongo
Department of Statistics and Actuarial Science
School of Mathematical Sciences
C. K. Tedam University of Technology and Applied Sciences
Navrongo, Ghana.
Email: idalogubayom@yahoo.com

ABSTRACT

In this paper, the transformed sine Dagum distribution is proposed using the Dagum distribution as a baseline distribution in the transformed sin-G family of distributions. The purpose of this paper is to propose and explore a modified version of the Dagum distribution, based on the transformed sin-G family of distributions. The density of the transformed sine Dagum distribution shows that it can adequately model datasets that are positively skewed, approximately symmetric and decreasing. The hazard rate plot shows that the proposed distribution can adequately model both monotonic and non-monotonic hazard rate datasets. Statistical properties such quantile function, moments, inverse moments, moments generating functions, mean and median deviation and order statistics are derived. The plots of the skewness and kurtosis show that the skewness is always positive and the kurtosis is increasing. Maximum likelihood estimation is used to estimate the parameters of the transformed sine Dagum distribution. Monte Carlo simulations are performed to ascertain the behavior of the estimators. The results show that the estimators are consistent. The proposed distribution is applied to two real life datasets. The results show that the proposed distribution provide a reasonable parametric fit to the datasets compared with the other competing distribution.

KEYWORDS

Dagum distribution, transformed sin-G, monotonic hazard rate, non-monotonic hazard rate.

1. INTRODUCTION

Probability distributions play a major role in various statistical analysis. Advancement in probability distribution theory in developing new families of distributions have given rise to the sin-G family of distributions which is a trigonometric family of continuous distributions. Trigonometric families are simple alternatives of deriving models without introducing new parameter(s). The unique features of the sin-G family inspired other trigonometric extension such as Cos-G family by Souza et al. (2019), CS-G family by Chesneau et al. (2019), NSin-G family by Mahmood et al. (2019) most of which are based on the sin-G family with no additional parameter(s) or transformation. The transformed sin-G (TS-G) family is an extension of the sin-G family proposed by Jamal et al. (2021). Its cumulative distribution function (CDF) is derived using a simple one parameter polynomial trigonometric transformation. The TS-G family has many desirable characteristics, including being analytically straightforward, the non-transformed case, having continuous CDF properties so that values are within the unit interval, differentiable, and increasing, and having the ability to be either convex, concave, or none of the above for different parameter values. The polynomial trigonometric functions, flexibility, kurtosis, skewness, adaptable distribution tails, and a variety of hazard rate shapes make the TS-G family different from other distributions and provide it with a wide range of fitting abilities. The Transformed sin-G family of distributions by Jamal et al. (2021) has CDF given by

$$F(x;\lambda,\boldsymbol{\Phi}) = \sin\left[\frac{\pi}{2}G(x;\boldsymbol{\Phi})\right] - \lambda\frac{\pi}{2}G(x;\boldsymbol{\Phi})\cos\left[\frac{\pi}{2}G(x;\boldsymbol{\Phi})\right], x \in \mathbb{R}, \quad (1)$$

where $\lambda \in [0,1]$, $G(x;\boldsymbol{\Phi})$ is the CDF of continuous distribution and $\boldsymbol{\Phi}$ is a $k \times 1$ vector of parameter(s).

The Dagum distribution was introduced by Dagum (Dagum, 1977). It is a continuous probability distribution defined over positive real numbers. This distribution arose from several variants of a new model on the size distribution of personal income and is mostly applied in economics. The Dagum distribution can be three-parameter specification (Type I Dagum distribution) or a four-parameter specification (Type II Dagum distribution). It was developed with the aim of getting a distribution that accommodates the heavy tails present in empirical income and wealth data as well as permitting interior mode. An important feature of the Dagum distribution is that its hazard rate function can be monotonically decreasing, bathtub and upside-down bathtub shaped (Domma, 2002). Although the Pareto distribution accommodates heavy tails, it does not permit interior mode. Also, the log-normal permits interior mode, but does not accommodate heavy tails. The Dagum distribution have Pareto, Fisk, log-logistic and Singh-Maddala distributions as special cases.

In recent time, the Dagum distribution has been studied from a reliability point of view and used to analyze survival data (Domma et al., 2011; Domma et al., 2013). Sakthivel and Dhivakar (2021) proposed and studied the transmuted sine-Dagum distribution using the transmuted sine-G family of distributions. Khadim et al. (2021) proposed the log-Dagum Weibull distribution on the basis of the T-X family technique. Domma and Condino (2013) proposed the five parameter beta-Dagum distribution. Oluyede and Ye (2014) presented the class weighted Dagum and related distributions. Also, other generalization of the Dagum distribution include the Log-Dagum distribution by Domma (2004), Transmuted Dagum by Elbatal and Aryal (2015), Gamma-Dagum by Rodrigues and Silva (2015), and Odd Log-Logistic Dagum by Domma et al. (2018).

The goal of this paper is to propose and explore a modified version of the Dagum distribution base on the TS-G family of distributions. Thus, the proposed distribution is known as Transformed sine Dagum distribution (TSDa). This distribution presents desirable features and hence is flexible. To the best of my knowledge the Dagum distribution has not been modified using the transformed sin-G family of distributions. Therefore, this is an attempt in modifying the Dagum distribution via the TS-G family of distributions.

The rest of the paper is organized as follows: Section 2 presents the basics of the Transformed Sine Dagum distribution. The statistical properties of the distribution are presented in Section 3. In Section 4, the parameter estimation of the proposed distribution using maximum likelihood estimation is presented. The behaviour of the estimators of the parameters of the proposed distribution is ascertained in Section 5 using Monte Carlo simulations. In Section 6, the applications of the TSDa distribution are illustrated using two real life datasets and the conclusion is presented in Section 7.

2. BASICS ON THE TSDa DISTRIBUTION

Taking inspiration from Section 1, a modified version of the three-parameter Dagum distribution known as Transformed Sine Dagum (TSDa) distribution is proposed. Consider the three-parameter Dagum as the baseline distribution with CDF and PDF as $G(x;\alpha,\beta,\gamma) = (1+\alpha x^{-\gamma})^{-\beta}$ and $g(x;\alpha,\beta,\gamma) = \alpha\beta\gamma x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1)}$ for $\alpha,\beta,\gamma > 0$ and $x > 0$ respectively. Substituting the CDF and PDF of the Dagum distribution into equation (1), the TSDa distribution is obtained with CDF given by;

$$F(x;\ \lambda,\alpha,\beta,\gamma) = \sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] - \lambda\frac{\pi}{2}\ (1+\alpha x^{-\gamma})^{-\beta} \cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right], x>0, \alpha,\beta,\gamma>0, 0\le\lambda\le 1, \quad (2)$$

where β and γ are shape parameters and α is a scale parameter.

The related PDF is given by

$$f(x;\ \lambda,\alpha,\beta,\gamma) = \frac{\pi\alpha\beta\gamma}{2}x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1)}\left[\lambda\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta} \sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] + (1-\lambda)\cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\right], x>0. \quad (3)$$

Lemma 1: The PDF of the TSDa distribution has a mixture representation of the form

$$f(x;\ \lambda,\alpha,\beta,\gamma) = \alpha\beta\gamma\sum_{n=0}^{\infty}\Psi_n(2n+1)x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1+2n\beta)}, \quad (4)$$

where $\Psi_n = (-1)^n\left(\frac{\pi}{2}\right)^{2n+1}\ \left(\frac{1-\lambda(2n+1)}{(2n+1)!}\right)$.

Proof: Using the Taylor series expansion of the sine and cosine functions given by $\sin(x) = \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}x^{2n+1}$ and $\cos(x) = \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}x^{2n}$ respectively, gives

$$\sin\left[\frac{\pi}{2}G(x;\boldsymbol{\Phi})\right] = \sum_{n=0}^{\infty}\frac{(-1)^n\left(\frac{\pi}{2}\right)^{2n+1}}{(2n+1)!}G(x;\boldsymbol{\Phi})^{2n+1} \quad (5)$$

and

$$\cos\left[\frac{\pi}{2}G(x;\boldsymbol{\Phi})\right] = \sum_{n=0}^{\infty}\frac{(-1)^n\left(\frac{\pi}{2}\right)^{2n}}{(2n)!}G(x;\boldsymbol{\Phi})^{2n}. \quad (6)$$

Substituting equations (5) and (6) into equation (1), one will obtain

$$F(x;\ \lambda, \boldsymbol{\Phi}) = \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{\pi}{2}\right)^{2n+1}}{(2n+1)!} G(x; \boldsymbol{\Phi})^{2n+1} - \lambda \frac{\pi}{2} G(x; \boldsymbol{\Phi}) \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{\pi}{2}\right)^{2n}}{(2n)!} G(x; \boldsymbol{\Phi})^{2n}$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{\pi}{2}\right)^{2n+1}}{(2n+1)!} G(x; \boldsymbol{\Phi})^{2n+1} - \lambda \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{\pi}{2}\right)^{2n+1}}{(2n)!} G(x; \boldsymbol{\Phi})^{2n+1}.$$

The fact that $2n! = (2n+1)!/2n+1$, implies that

$$F(x;\ \lambda, \alpha, \beta, \gamma) = \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{\pi}{2}\right)^{2n+1}}{(2n+1)!} G(x; \boldsymbol{\Phi})^{2n+1} - \lambda \sum_{n=0}^{\infty} \frac{(-1)^n (2n+1) \left(\frac{\pi}{2}\right)^{2n+1}}{(2n+1)!} G(x; \boldsymbol{\Phi})^{2n+1}.$$

Letting $\Psi_n = (-1)^n \left(\frac{\pi}{2}\right)^{2n+1} \left(\frac{1-\lambda(2n+1)}{(2n+1)!}\right)$, implies that

$$F(x;\ \lambda, \alpha, \beta, \gamma) = \sum_{n=0}^{\infty} \Psi_n\, G(x; \boldsymbol{\Phi})^{2n+1}. \tag{7}$$

Differentiating equation (7), one obtains;

$$f(x;\ \lambda, \alpha, \beta, \gamma) = \sum_{n=0}^{\infty} \Psi_n\, (2n+1) g(x; \boldsymbol{\Phi}) G(x; \boldsymbol{\Phi})^{2n}. \tag{8}$$

Therefore, the mixture representation of the PDF of the TSDa is obtained by substituting the CDF and PDF of the Dagum distribution into equation (8). This is given by;

$$f(x;\ \lambda, \alpha, \beta, \gamma) = \alpha\beta\gamma \sum_{n=0}^{\infty} \Psi_n (2n+1) x^{-(\gamma+1)} (1 + \alpha x^{-\gamma})^{-(\beta+1+2n\beta)}. \tag{9}$$

The hazard rate function is given by

$$h(x;\lambda,\alpha,\beta,\gamma)=\frac{\pi\alpha\beta\gamma x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1)}\begin{bmatrix}\lambda\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\\+(1-\lambda)\cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\end{bmatrix}}{2\begin{bmatrix}1-\sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\\-\lambda\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\end{bmatrix}}, x>0. \tag{10}$$

From equation (2), the following remarks are made:

Remark 1: When $\alpha = 1$, the TSDa distribution becomes Transformed Sine Burr III (TSBIII) distribution with CDF given by

$$F(x;\ \lambda,\beta,\gamma)=\sin\left[\frac{\pi}{2}(1+x^{-\gamma})^{-\beta}\right]-\lambda\frac{\pi}{2}(1+x^{-\gamma})^{-\beta}\cos\left[\frac{\pi}{2}(1+x^{-\gamma})^{-\beta}\right],$$
$$x>0,\beta,\gamma>0,0\le\lambda\le 1.$$

Remark 2: When $\gamma = 1$, the TSDa distribution becomes Transformed Sine log-logistic (TSLL) distribution with CDF given by

$$F(x;\ \lambda,\alpha,\beta)=\sin\left[\frac{\pi}{2}(1+\alpha x)^{-\beta}\right]-\lambda\frac{\pi}{2}(1+\alpha x)^{-\beta}\cos\left[\frac{\pi}{2}(1+\alpha x)^{-\beta}\right],$$
$$x>0,\alpha,\beta,>0,0\le\lambda\le 1.$$

Figure 1 displays the density plot of the TSDa distribution. For various parameter combinations, it is seen that its density can be decreasing, right skewed and approximately symmetric. There is evident of heavy-tailed behavior with various kurtosis values. This implies that the TSDa distribution can adequately model data set whose density exhibit any of the characteristics mentioned.

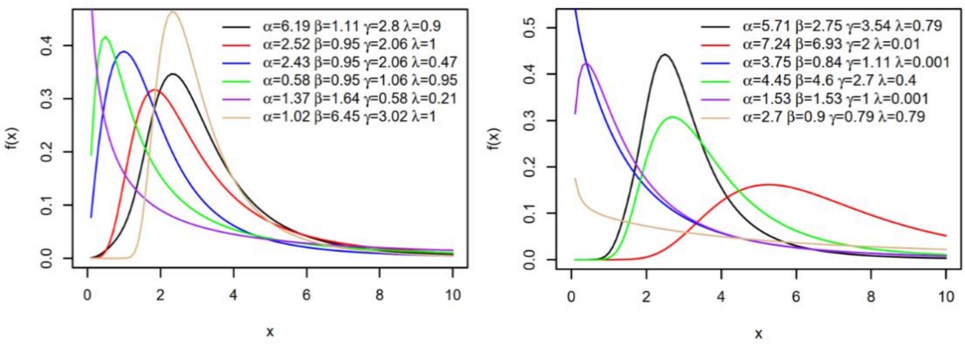


Figure 1: Plots of the Density of the TSDa Distribution

As shown in Figure 2, the hazard rate function of the TSDa distribution can be increasing, decreasing, bathtub, upside-down-bathtub, modified upside-down-bathtub among others. This implies the TSDa distribution can adequately model both monotonic and non-monotonic failure rate data.

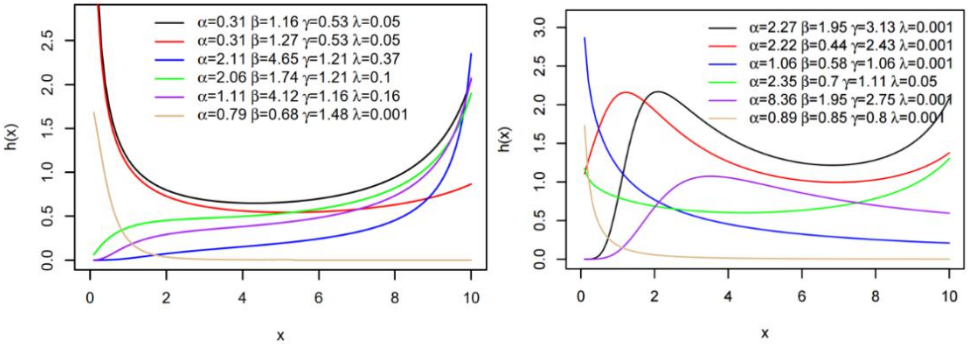


Figure 2: Plots of the Hazard Rate Function of the TSDa Distribution

3. STATISTICAL PROPERTIES

In this section, the statistical properties of the TSDa distribution are derived.

3.1 Quantile Function

Quantile functions are vital in describing the distribution of a random variable. It helps in generating random samples which are useful in simulations. It can also be used to compute measures of shape such as skewness and kurtosis.

The quantile function of the TSDa distribution for $u \in (0, 1)$ is given by

$$\sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] - \lambda\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] - u = 0. \quad (11)$$

The first quartile, the median, and the upper quartile are obtained by substituting u = 0.25, 0.5, and 0.75, respectively into equation (11).

3.2 Moments

The moments of a distribution are important in estimating measures of variation like the variance, standard deviation, coefficient of variation, mean deviation, median deviation and measures of shapes such as kurtosis, skewness amongst others.

Proposition 1: The m^{th} non-central moment of the TSDa distribution is given by

$$\mu'_m = \alpha^{m/\gamma}\beta\sum_{n=0}^{\infty}\Psi_n\,(2n+1)B\left(1+\frac{m}{\gamma}-(\beta-1+2n\beta), 1-\frac{m}{\gamma}\right), \gamma > m. \quad (12)$$

Proof: By definition, the m^{th} non-central moment is given by

$$\mu'_m = \int_0^\infty x^m f(x)dx.$$

This implies that,

$$\mu'_m = \alpha\beta\gamma \sum_{n=0}^{\infty} \Psi_n\,(2n+1) \int_0^{\infty} x^m\, x^{-(\gamma+1)} (1+\alpha x^{-\gamma})^{-(\beta+1+2n\beta)} dx.$$

Let $y = (1+\alpha x^{-\gamma})^{-1}$, then if $x \to 0$, $y \to 0$ and if $x \to \infty$, $y \to 1$. Also, $dx = \frac{dy}{\alpha\gamma(1+\alpha x^{-\gamma})^{-2} x^{-(\gamma+1)}}$ and $x = \alpha^{\frac{1}{\gamma}} y^{\frac{1}{\gamma}} (1-y)^{-1/\gamma}$.

After some algebra, manipulations, and making use of the beta function; $B(a,b) = \int_0^1 y^{a-1}(1-y)^{b-1}\,dy$, one obtains

$$\mu'_m = \alpha^{m/\gamma}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1) B\left(1+\frac{m}{\gamma} - (\beta-1+2n\beta), 1-\frac{m}{\gamma}\right), \gamma > m.$$

Remark 3: By substituting $m = 1$ into equation (12), one obtains the mean of the TSDa distribution given by

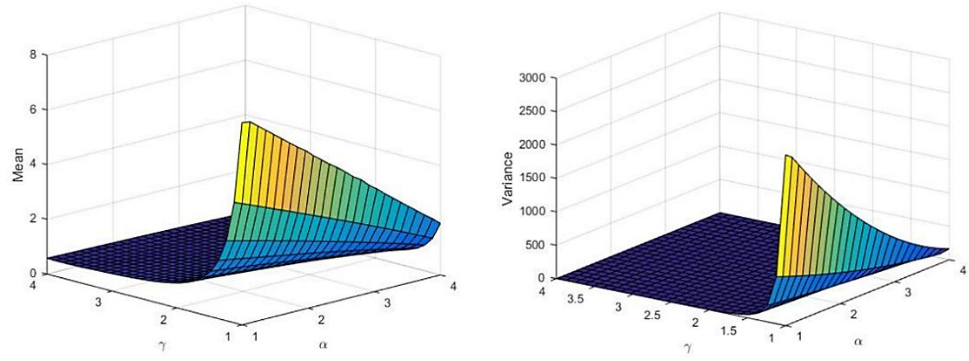
$$\mu'_1 = \alpha^{1/\gamma}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1) B\left(1+\frac{1}{\gamma} - (\beta-1+2n\beta), 1-\frac{1}{\gamma}\right), \gamma > 1. \qquad (13)$$

The values for the first four moments, standard deviation (SD), coefficient of variation (CV), coefficient of skewness (CS), and coefficient of kurtosis (CK) of the TSDa distribution for selected values of the parameters are shown in Table 1. The values of the first four moments are obtained by using numerical integration. The values of SD, CV, CS and CK are computed using the equations; $SD = \sqrt{\mu'_2 - (\mu'_1)^2}$, $CV = \sqrt{\mu'_2 - (\mu'_1)^2}/\mu'_1$, $CS = \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3/[\mu'_2 - (\mu'_1)^2]^{3/2}$, and $CK = \mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4/[\mu'_2 - (\mu'_1)^2]^2$ respectively.

Table 1
First Four Moments, SD, CS, CK and CV of the TSDa Distribution for Some Parameter Values

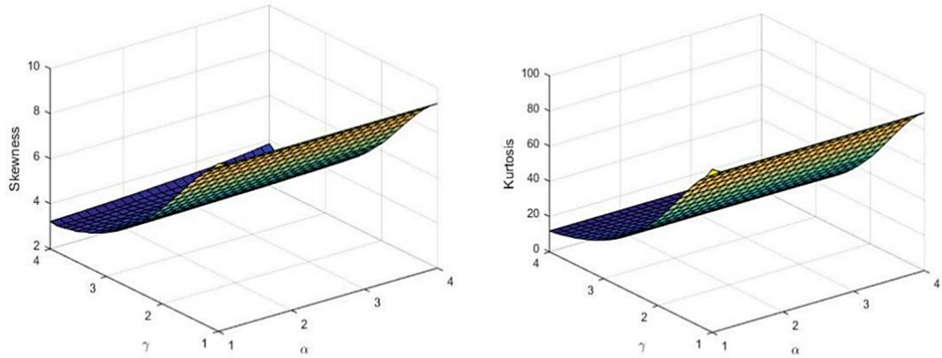
$(\alpha,\beta,\gamma,\lambda)$	μ'_1	μ'_2	μ'_3	μ'_4	**SD**	**CS**	**CK**	**CV**
(0.1,2.3,4.5,0.5)	0.867	0.871	0.111	2.655	0.345	3.627	72.726	0.398
(1.2,0.1,5.1,0.1)	0.251	0.146	0.128	0.172	0.288	21.771	6.571	1.474
(1.4,0.1,5.1,0.82)	0.606	0.558	0.683	1.230	0.437	1.362	10.973	0.721
(2.5,0.2,6.4,0.3)	0.663	0.593	0.646	0.852	0.392	2.249	5.231	0.591
(1.3,0.2,7.2,0.9)	0.877	0.867	0.951	1.167	0.313	6.789	5.974	0.361

Figures 3 and 4 show the mean, variance, skewness and kurtosis plots of the TSDa distribution for $\beta = 0.13$, $\lambda = 0.25$ and a range of values for α and γ respectively. From the plots, the mean and variance are increasing. It can also be seen that the skewness is positive, an indication that the TSDa is right skewed and the kurtosis is increasing; meaning TSDa distribution is leptokurtic.



(a) (b)

Figure 3: Mean (a) and Variance (b) Plots of the TSDa Distribution



(a) (b)

Figure 4: Skewness (a) and Kurtosis (b) Plots of the TSDa Distribution

3.3 Incomplete Moment

The incomplete moment of a distribution is vital in estimating the mean deviation, median deviation, and measures of inequalities like Bonferroni and Lorenz curves.

Proposition 2: The m^{th} incomplete moment of the TSDa distribution is given by

$$M_m(x) = \alpha^{m/\gamma}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1) B\left((1+\alpha x^{-\gamma})^{-1}; 1+\frac{m}{\gamma} - (\beta - 1 + 2n\beta), 1 - \frac{m}{\gamma}\right), m < \gamma, \quad (14)$$

where $B(.;.,.)$is an incomplete beta function and $m = 1, 2, \ldots$

Proof: Using the identity; $B(s; a, b) = \int_0^s y^{a-1}(1-y)^{b-1}\,dy$ and the concept in proving the moment, the incomplete moment of the TSDa distribution is given

$$M_m(x) = \alpha\beta\gamma \sum_{n=0}^{\infty} \Psi_n\,(2n+1) \int_0^{(1+\alpha x^{-\gamma})^{-1}} x^m\, x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1+2n\beta)}dx. \tag{15}$$

On solving equation (15), one obtains

$$M_m(x) = \alpha^{\frac{m}{\gamma}}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1)B$$
$$\left((1+\alpha x^{-\gamma})^{-1}; 1+\frac{m}{\gamma}-(\beta-1+2n\beta), 1-\frac{m}{\gamma}\right), m<\gamma.$$

3.4 Inverted Moments

The inverse moments can be applied in many practical applications. For example, they appear in Stein estimation and Bayesian post-stratification (Wooff, 1985; Pittenger, 1990), evaluating risks of estimators and powers of test statistics (Marciniak and Wesolowski, 1999; Fujioka, 2001), expected relaxation times of complex systems (Jurlewicz and Weron, 2002), insurance and financial mathematics (Ramsay, 1993).

Proposition 3: The m^{th} inverted moment of the TSDa distribution is given by

$$\mu_m^* = \alpha^{-m/\gamma}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1)B\left(1-\frac{m}{\gamma}-(\beta-1+2n\beta), 1-\frac{m}{\gamma}\right), \gamma>m, \tag{16}$$

where $m = 1, 2, \ldots$.

Proof: By definition, the m^{th} inverted moment of a random variable X is given by

$$\mu_m^* = \int_0^{\infty} x^{-m}\, f(x)dx.$$

That is,

$$\mu_m^* = \alpha^{-m/\gamma}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1) \int_0^{\infty} x^{-m}\, x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1+2n\beta)}dx.$$

Using the concept in obtaining the moments and the identity $B(a,b) = \int_0^1 y^{a-1}(1-y)^{b-1}\, dy$,

$$\mu_m^* = \alpha^{-m/\gamma}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1)B\left(1-\frac{m}{\gamma}-(\beta-1+2n\beta), 1-\frac{m}{\gamma}\right), \gamma>m.$$

3.5 Moment Generating Function

The moment generating function (MGF) helps in determining the moments of a random variable.

Proposition 4: The MGF of the TSDa distribution is given by

$$M_X(t) = \alpha^{m/\gamma}\beta \sum_{n=0}^{\infty} \frac{\Psi_n(2n+1)t^m}{m!} B\left(1+\frac{m}{\gamma} - (\beta - 1 + 2n\beta), 1-\frac{m}{\gamma}\right), \gamma > m. \tag{17}$$

Proof: By definition, the MGF is given as;

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_0^{\infty} e^{tx} f(x)dx.$$

Using Taylor series expansion, one obtains

$$M_X(t) = \mathbb{E}\left[\sum_{m=0}^{\infty} \frac{t^m X^m}{m!}\right] = \sum_{m=0}^{\infty} \frac{t^m}{m!}\mu_m'. \tag{18}$$

Substituting equation (12) into equation (18) completes the proof.

3.6 Mean and Median Deviations

The totality of the deviations from the mean and median can be used to estimate the variation in a population with some certainty. If the random variable X follows the TSDa distribution, then the mean and median deviations are given by the following propositions.

Proposition 7: The expected value of the absolute deviation of a random variable X having the TSDa distribution from its mean is given by

$$\kappa_1(x) = 2\mu F(\mu) - 2\alpha^{\frac{1}{\gamma}}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1) B\left((1+\alpha x^{-\gamma})^{-1}; 1+\frac{1}{\gamma} - (\beta - 1 + 2n\beta), 1-\frac{1}{\gamma}\right), \gamma > 1, \tag{19}$$

where $\mu = \mu_1'$ is the mean of X.

Proof: By definition,

$$\begin{aligned}\kappa_1(x) &= \int_0^{\infty} |x-\mu|\, f(x)dx \\ &= 2\mu F(\mu) - 2\int_0^{\mu} x\, f(x)dx\end{aligned}$$

$$\kappa_1(x) = 2\mu F(\mu) - 2\alpha^{1/\gamma}\,\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1) B\left((1+\alpha x^{-\gamma})^{-1}; 1+\frac{1}{\gamma} - (\beta - 1 + 2n\beta), 1-\frac{1}{\gamma}\right), \gamma > 1,$$

where $\int_0^{\mu} x\, f(x)dx$ is simplified using the first incomplete moment.

Proposition 8: The expected value of the absolute deviation of a random variable X following the TSDa distribution from its median is given by

$$\kappa_2(x) = \mu - 2\int_0^{\mu} x\, f(x)dx$$

$$\kappa_2(x) = u - 2\alpha^{\frac{1}{\gamma}}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1)B\left((1+\alpha M^{-\gamma})^{-1}; 1+\frac{1}{\gamma} - (\beta - 1 + 2n\beta), 1-\frac{1}{\gamma}\right), \gamma > 1, \tag{20}$$

where M is the median of X.

Proof: By definition,

$$\kappa_2(x) = \int_0^{\infty} |x - M|\, f(x)dx$$

$$= \mu - 2\int_0^{M} x\, f(x)dx$$

$$\kappa_2(x) = \mu - 2\alpha^{\frac{1}{\gamma}}\beta \sum_{n=0}^{\infty} \Psi_n\,(2n+1)B\left((1+\alpha M^{-\gamma})^{-1}; 1+\frac{1}{\gamma} - (\beta - 1 + 2n\beta), 1-\frac{1}{\gamma}\right), \gamma > 1,$$

where $\int_0^M x\, f(x)dx$ is simplified using the first incomplete moment.

3.7 Order Statistics

Let $X_1, X_2, \dots, X_n$ be a sample of size n and $X_{1:n} \le X_{2:n} \le \cdots X_{n:n}$ denote the order statistics of the sample. The PDF of the i^{th} order statistics $f_{i:n}(x)$ is defined as

$$f_{i:n}(x) = \frac{n!}{(i-1)!\,(n-i)!}[F(x)]^{i-1}[1-F(x)]^{n-i}f(x). \tag{21}$$

Using binomial series expansion,

$$[1-F(x)]^{n-i} = \sum_{r=0}^{n-i}(-1)^r\binom{n-i}{r}[F(x)]^r. \tag{22}$$

That is, equation (21) becomes

$$f_{i:n}(x) = \frac{n!}{(i-1)!\,(n-i)!}f(x)\sum_{r=0}^{n-i}(-1)^r\binom{n-i}{r}[F(x)]^{r+i-1}. \tag{23}$$

Substituting the CDF and PDF of the TSDa distribution into equation (23), the i^{th} order statistics of the TSDa distribution is given by

$$f_{i:n}(x) = \frac{A \times \left[\begin{array}{c} \lambda \frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta} \sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] \\ +(1-\lambda)\cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] \end{array} \right]}{2(i-1)!\,(n-i)!}$$
$$\times \sum_{r=0}^{n-i} (-1)^r \binom{n-i}{r} \left[\begin{array}{l} \sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] \\ -\lambda \frac{\pi}{2}\,(1+\alpha x^{-\gamma})^{-\beta} \cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] \end{array} \right]^{r+i-1} \quad (24)$$

where $A = \pi\alpha\beta\gamma x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1)}$

The PDF of the first-order statistics is defined as

$$f_{1:n}(x) = n[1-F(x)]^{n-1} f(x). \quad (25)$$

Letting $[1-F(x)]^{n-1} = \sum_{k=0}^{n-1}\binom{n-1}{k}[F(x)]^k$, equation (25) can be re-written as

$$f_{1:n}(x) = nf(x) \sum_{k=0}^{n-1} \binom{n-1}{k} [F(x)]^k. \quad (26)$$

From equation (26), the PDF of the first-order statistics of the TSDa distribution is given by

$$f_{1:n}(x) = \frac{n\pi\alpha\beta\gamma}{2} x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1)}$$
$$\left[\lambda \frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta} \sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] + (1-\lambda)\cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\right]$$
$$\left[\sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] - \lambda \frac{\pi}{2}\,(1+\alpha x^{-\gamma})^{-\beta} \cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\right]^k \quad (27)$$

Also, the PDF of the n^{th} order statistics is defined as

$$f_{n:n}(x) = n[F(x)]^{n-1} f(x). \quad (28)$$

Therefore, employing equation (28) the PDF of the n^{th} order statistics as

$$f_{n:n}(x) = n\left[\sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] - \lambda \frac{\pi}{2}\,(1+\alpha x^{-\gamma})^{-\beta} \cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\right]^{n-1}$$
$$\times \frac{\pi\alpha\beta\gamma}{2} x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1)}$$
$$\left[\lambda \frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta} \sin\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right] + (1-\lambda)\cos\left[\frac{\pi}{2}(1+\alpha x^{-\gamma})^{-\beta}\right]\right]. \quad (29)$$

4. PARAMETER ESTIMATION

In this section, the unknown parameters of the TSDa distribution are estimated using the maximum likelihood estimation (MLE) technique.

4.1 Maximum Likelihood Estimation

Let $X_1, X_2, \ldots, X_n$ be n random sample from the TSDa distribution and $\rho = (\alpha, \beta, \gamma, \lambda)'$, then the log-likelihood function, $\ell = \ell(\rho)$, is given by

$$\ell = n\log\left(\frac{\pi\alpha\beta\gamma\lambda}{2}\right) - (\gamma+1)\sum_{i=1}^{n}\log(x_i) - (\beta+1)\sum_{i=1}^{n}\log\left(1+\alpha x_i^{-\gamma}\right) + \sum_{i=1}^{n}\log\left[\lambda\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] + (1-\lambda)\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]. \quad (30)$$

The log-likelihood function in equation (28) is differentiated with respect to each parameter to obtain the score function, $L(\rho) = \left(\frac{\partial\ell}{\partial\alpha}, \frac{\partial\ell}{\partial\beta}, \frac{\partial\ell}{\partial\gamma}, \frac{\partial\ell}{\partial\lambda}\right)^T$. The elements of the score function are as follows:

$$\frac{\partial\ell}{\partial\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n}\frac{1+\beta}{\alpha+x_i^{\gamma}} - \Bigg(2\sum_{i=1}^{n}\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\Bigg(\frac{1}{2}x_i^{-\gamma}\left(1+\alpha x_i^{-\gamma}\right)^{-(\beta+1)}\beta\left(-2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\right)\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] + 2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] - \pi\lambda\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] - \frac{1}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\Bigg(-2x_i^{-\gamma}\left(1+\alpha x_i^{-\gamma}\right)^{\beta-1}\beta\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] + \frac{\pi^2}{2}x_i^{-\gamma}\left(1+\alpha x_i^{-\gamma}\right)^{-(\beta+1)}\beta\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] + 2x_i^{-\gamma}\left(1+\alpha x_i^{-\gamma}\right)^{\beta-1}\beta\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] - \frac{\pi x_i^{-\gamma}\beta\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]}{1+\alpha x_i^{-\gamma}} + \frac{\pi x_i^{-\gamma}\beta\lambda\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]}{1+\alpha x_i^{-\gamma}}\Bigg)\Bigg)\Bigg) / \left(-2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] + 2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] - \pi\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\right), \quad (31)$$

$$
\begin{aligned}
\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} - \sum_{i=1}^{n} \log\left(x_i^{-\gamma}\left(x_i^{\gamma} + \alpha\right)\right) \\
&- \left(2\sum_{i=1}^{n}\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\left(\frac{1}{2}x_i^{-\gamma}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\log\left(1\right.\right.\right. \\
&\left. + \alpha x_i^{-\gamma}\right)\ \left(-2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\right)\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] \\
&+ 2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] - \pi\lambda\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] \\
&- \tfrac{1}{2}\left(1\right. \\
&\left.+ \alpha x_i^{-\gamma}\right)^{-\beta}\left(-2x_i^{-\gamma}\left(1+\alpha x_i^{-\gamma}\right)^{\beta-1}\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\log\left(1\right.\right. \\
&\left.+ \alpha x_i^{-\gamma}\right) + \frac{\pi^2}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\log\left(1+\alpha x_i^{-\gamma}\right) \\
&+ 2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\log\left(1+\alpha x_i^{-\gamma}\right) \\
&- \pi\log\left(1+\alpha x_i^{-\gamma}\right)\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] \\
&\left.\left.\left.+ \pi\lambda\log\left(1+\alpha x_i^{-\gamma}\right)\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\right)\right)\right) \\
&/\left(-2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\right. \\
&\left.+ 2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right] - \pi\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\right),
\end{aligned} \tag{32}
$$

$$\frac{\partial \ell}{\partial \gamma}=\frac{n}{\gamma}-\sum_{i=1}^{n}\log(x_i)-\sum_{i=1}^{n}\frac{x_i^{\gamma}\left[\log(x_i)-x_i^{-\gamma}(x_i^{\gamma}+\alpha\log(x_i)\right]}{\alpha+x_i^{\gamma}}$$
$$-\left(2\sum_{i=1}^{n}\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\left(-\frac{1}{2}\alpha x_i^{-\gamma}\left(1\right.\right.\right.$$
$$+\alpha x_i^{-\gamma}\Big)^{-(\beta+1)}\beta\log(x_i)\left(-2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\right)\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]$$
$$+2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]-\pi\lambda\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]$$
$$-\frac{1}{2}\Big(1$$
$$+\alpha x_i^{-\gamma}\Big)^{-\beta}\left(2\alpha x_i^{-\gamma}\left(1+\alpha x_i^{-\gamma}\right)^{\beta-1}\beta\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\log(x_i)\right.$$
$$-\frac{\pi^2}{2}x_i^{-\gamma}\alpha\left(1+\alpha x_i^{-\gamma}\right)^{-(\beta+1)}\beta\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\log(x_i)$$
$$-2x_i^{-\gamma}\alpha\left(1+\alpha x_i^{-\gamma}\right)^{\beta-1}\beta\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\log(x_i)$$
$$+\frac{\pi x_i^{-\gamma}\alpha\beta\log(x_i)\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]}{1+\alpha x_i^{-\gamma}}$$
$$\left.\left.\left.-\pi x_i^{-\gamma}\alpha\beta\lambda\log(x_i)\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\right)\right)\right)\right)$$
$$/\left(-2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\right.$$
$$+2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]$$
$$\left.-\pi\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]\right) \tag{33}$$

and

$$\frac{\partial \ell}{\partial \lambda}=\sum_{i=1}^{n}\frac{2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]-\pi\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]}{-2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]+2\left(1+\alpha x_i^{-\gamma}\right)^{\beta}\lambda\cos\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]-\pi\sin\left[\frac{\pi}{2}\left(1+\alpha x_i^{-\gamma}\right)^{-\beta}\right]} \tag{34}$$

5. MONTE CARLO SIMULATION

In this section, the simulation results are presented in examining the properties of the maximum likelihood estimators for the parameters of the TSDa distribution. Four different combinations of the parameter values of this distribution are specified and the quantile function is used in generating four different random samples of size, $n = 70$, 100, 200, 300. The simulations are replicated for $N = 1000$ times. The properties of the estimators are investigated by computing average bias (AB) and root mean square error (RMSE) for each of the parameter. The simulation steps are as follows:

i. Specify the values of the parameters and the sample size n.

ii. Generate random samples of size $n =$ 70, 100, 200, 300 from the TSDa distribution using its quantile.

iii. Find the maximum likelihood estimates for parameters.

iv. Repeat steps ii-iii for 1000 times.

v. Calculate the average bias (AB) and root mean square error (RMSE) for the parameters of the distributions.

Table 2 shows the simulation results for the TSDa distribution. It can be observed that as the sample size increase, the AB and RMSE for the estimators of each parameter decreases. This shows that the estimators are consistent. Therefore, the MLEs and their asymptotic results can be adopted in estimating the model parameters.

Table 2
Simulation Results for the Parameters of the TSDa Distribution

N	Parameter value				AB				RMSE			
	α	β	γ	λ	α	β	γ	λ	α	β	γ	λ
70	1.2	0.13	1.52	0.25	0.636	0.037	0.343	0.163	0.473	0.003	0.151	0.049
100	1.2	0.13	1.52	0.25	0.631	0.033	0.330	0.143	0.468	0.002	0.141	0.038
200	1.2	0.13	1.52	0.25	0.573	0.026	0.266	0.117	0.396	0.001	0.099	0.023
300	1.2	0.13	1.52	0.25	0.572	0.022	0.239	0.105	0.396	0.001	0.081	0.019
70	4.81	0.151	1.811	0.28	3.010	0.043	0.338	0.190	9.227	0.004	0.163	0.055
100	4.81	0.151	1.811	0.28	2.967	0.037	0.322	0.175	8.935	0.002	0.141	0.046
200	4.81	0.151	1.811	0.28	2.910	0.028	0.291	0.158	8.543	0.001	0.110	0.035
300	4.81	0.151	1.811	0.28	2.885	0.026	0.283	0.155	8.375	0.001	0.099	0.032
70	1.06	0.17	1.1	0.22	0.655	0.057	0.343	0.181	0.515	0.005	0.187	0.060
100	1.06	0.17	1.1	0.22	0.621	0.048	0.301	0.160	0.474	0.003	0.150	0.047
200	1.06	0.17	1.1	0.22	0.580	0.038	0.231	0.139	0.423	0.002	0.090	0.033
300	1.06	0.17	1.1	0.22	0.554	0.031	0.181	0.129	0.391	0.002	0.056	0.027
70	3.104	0.236	1.824	0.5	1.242	0.063	0.244	0.140	1.637	0.085	0.088	0.029
100	3.104	0.236	1.824	0.5	1.215	0.053	0.221	0.132	1.551	0.006	0.071	0.027
200	3.104	0.236	1.824	0.5	1.165	0.039	0.196	0.111	1.390	0.003	0.051	0.020
300	3.104	0.236	1.824	0.5	1.150	0.033	0.186	0.101	1.342	0.002	0.044	0.016

6. APPLICATIONS

This section illustrates the usefulness and flexibility of the TSDa distribution using real datasets. The performance of the TSDa distribution is compared with other loss distributions. The performance of the distributions about providing proper parametric fit to the dataset is compared using the AIC, BIC, Cram´er-von Misses (W*), Anderson-Darling (A∗) and Kolmogorov Smirnov (K-S) statistics. The distribution with the least of these measures provide a reasonable fit to the dataset. The fit for the TSDa distribution is compared with other distributions, including the 3-parameter Dagum, 2-parameter Weibull, Transmuted sine Dagum (TSD), Burr III (BIII), sine Burr XII (SBXII) and Sine Inverse Lomax Frechet (SILF) distributions. The density functions of the Dagum, TSD, BIII, Weibull, SBXII and SILF are:

$$g(x) = \alpha\beta\gamma x^{-(\gamma+1)}(1+\alpha x^{-\gamma})^{-(\beta+1)}, x>0, \beta>0, \alpha>0, \gamma>0, \tag{35}$$

$$g(x) = \frac{\pi}{2}\sigma\theta\beta x^{-(\theta+1)}\left(1+\sigma x^{-\theta}\right)^{-(\beta+1)}\cos\left[\frac{\pi}{2}\left(1+\sigma x^{-\theta}\right)^{-\beta}\right]$$
$$\left\{(1+\lambda)-2\lambda\cos\left[\frac{\pi}{2}\left(1+\sigma x^{-\theta}\right)^{-\beta}\right]\right\},$$
$$x>0, \beta>0, \sigma>0, \theta>0, -1\le\lambda\le 1, \tag{36}$$

$$g(x) = ckx^{-(c+1)}(1+x^{-c})^{-(k+1)}, x>0, c>0, k>0, \tag{37}$$

$$g(x) = \alpha\beta x^{\alpha-1}e^{-\beta x^{\alpha}}, x\ge 0, \alpha>0, \beta>0, \tag{38}$$

$$g(x) = \frac{\pi ckx^{c-1}}{2(1+x^{c})^{k+1}}\cos\left[\frac{\pi}{2}\left(1-\frac{1}{(1+x^{c})^{k}}\right)\right], x>0, c>0, k>0 \tag{39}$$

and

$$g(x) = \frac{\pi\alpha\delta\mu^{\delta}x^{-(\delta+1)}}{2}e^{-\alpha\left(\frac{\mu}{\sigma}\right)^{\delta}}\cos\left[\frac{\pi}{2}e^{-\alpha\left(\frac{\mu}{\sigma}\right)^{\delta}}\right], x>0, c>0, k>0. \tag{40}$$

6.1 Traffic Data

The first dataset represents the length of intervals between the times at which vehicles pass a point on a road. The data set is given as;

2.50, 2.60, 2.60, 2.70, 2.80, 2.80, 2.90, 3.00, 3.00, 3.10, 3.20, 3.40, 3.70, 3.90, 3.90, 3.90, 4.60, 4.70, 5.00, 5.60, 5.70, 6.00, 6.00, 6.10, 6.60, 6.90, 6.90, 7.30, 7.60, 7.90, 8.00, 8.30, 8.80, 8.80, 9.30, 9.40, 9.50, 10.1, 11.0, 11.3, 11.9, 11.9, 12.3, 12.9, 12.9, 13.0, 13.8, 14.5, 14.9, 15.3, 15.4, 15.9, 16.2, 17.6, 20.1, 20.3, 20.6, 21.4, 22.8, 23.7, 23.7, 24.7 , 29.7, 30.6, 31.0, 34.1, 34.7, 36.8, 40.1, 40.2, 41.3, 42.0, 44.8, 49.8, 51.7, 55.7, 56.5, 58.1, 70.5, 72.6, 87.1, 88.6, 91.7, 119.8. This data was also studied by Lemonte et al. (2013).

The descriptive statistics of the traffic dataset is given in Table 3. The statistics shows that the length of interval data is positively skewed and leptokurtic. The skewness and leptokurtic nature of the data is further confirmed by the box-plot and histogram in Figure 5.

Table 3
Descriptive Statistics of Traffic Dataset

No. of Obs.	Mean	Std. dev.	Skewness	Kurtosis	Min.	Max.
84	21.602	23.97	1.915	6.581	2.500	119.8

The total time on test (TTT) transformed plot of the length of intervals between the times at which vehicles pass a point on a road presented in Figure 6 shows that, the failure rate function of the data set is bathtub shaped.

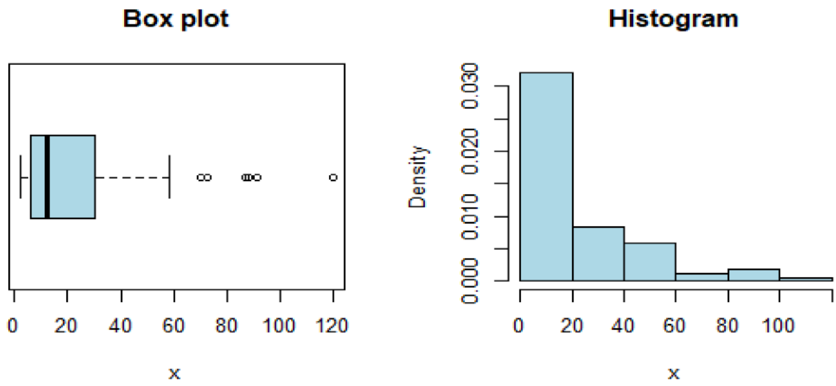


Figure 5: Box Plot and Histogram of the Traffic Dataset

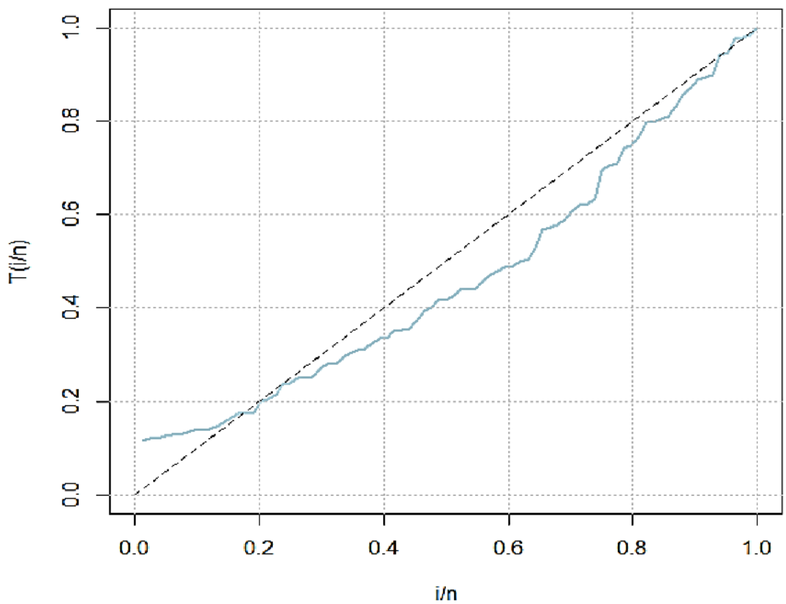


Figure 6: TTT-Transform Plot for Traffic Dataset

Table 4 shows the maximum likelihood estimates for the parameters of the fitted distributions with their corresponding standard errors in brackets and p-values. The parameters of the TSDa, BIII, Weibull and SILF are significant at the 5% level. The parameters of SBXII are significant at 10%. The parameters of TSD are significant at the 10% level with the exception of θ which is significant at 5%. Also, the parameters of the Dagum distribution are significant at 10% with the exception of α which is significant at 5%.

Table 4
Maximum Likelihood Estimates of the Parameters, Standard Errors and p-values of Traffic Dataset

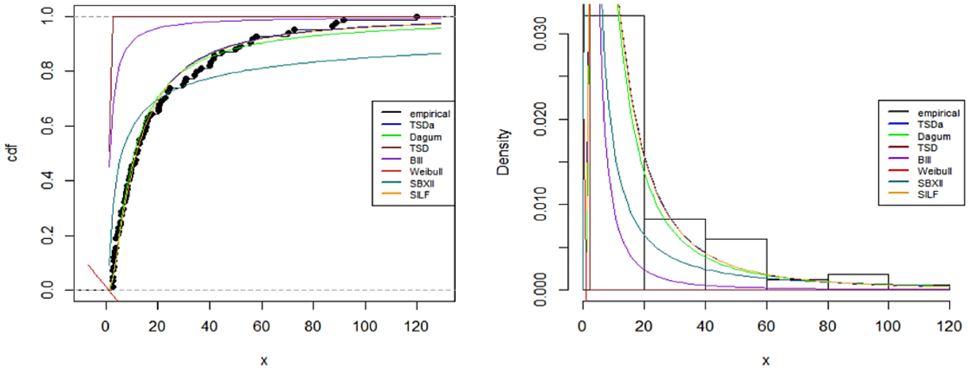
Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{\sigma}$	$\hat{\theta}$	$\hat{c}$	$\hat{k}$
TSDa	0.327 (0.097) 0.002	29.861 (5.711) 0.004	1.217 (0.114) 0.000	0.558 (0.143) 0.000				
Dagum	1.130 (0.101) 0.000	22.202 (33.107) 0.088	0.469 (0.759) 0.064					
TSD		22.019 (14.218) 0.069		-0.029 (0.399) 0.091	0.372 (0.263) 0.084	0.815 (0.074) 0.000		
BIII							1.159 (0.092) 0.000	11.428 (2.046) 0.000
Weibull	0.997 (0.081) 0.000	0.047 (0.014) 0.001						
SBXIII							7.443 (13.718) 0.065	0.030 (0.056) 0.081
SILF	3.486 (0.180) 0.000	$\hat{\mu}$ 2.690 (0.184) 0.000	$\hat{\delta}$ 0.789 (0.064) 0.000					

The model selection and goodness-of-fit statistics for the length of intervals between the times at which vehicles pass a point on a road are presented in Table 5. Among the distributions, the TSDa has the lowest AIC, BIC, W*, A*, K-S statistics values hence provides the best fit to the data set among the candidate distributions.

Table 5
Information Criteria and Goodness-of-Fit of Traffic Dataset

Model	$-2l$	AIC	BIC	W*	A*	K-S	p-value
TSDa	**671.144**	**675.144**	**680.867**	**0.067**	**0.621**	**0.072**	**0.780**
Dagum	672.417	678.417	688.709	0.099	0.826	0.073	0.737
TSD	671.314	679.314	689.038	0.083	0.745	0.296	0.008
BIII	672.552	676.552	681.414	0.095	0.805	0.075	0.756
Weibull	684.230	688.230	693.092	0.241	1.544	0.110	0.261
SBXII	744.587	748.587	753.449	0.076	0.665	0.289	0.006
SILF	675.792	676.792	684.085	0.068	0.621	0.074	0.745

Figure 7 shows the plot of the empirical CDF and the CDFs of the tentative distributions for the length of intervals between the times at which vehicles pass a point on a road. It can be seen that the TSDa provides a better to the data than the competitive distribution.



(a) (b)

Figure 7: CDF (a) and Empirical and fitted Density (b) Plots of Traffic Dataset

6.2 Breaking Stress of Carbon Fibres

The second dataset consists of 100 uncensored data on breaking stress of carbon fibres (in Gba). This data was also studied by Khadim et al. (2021). The data set is given as;

0.39, 0.81, 0.85, 0.98, 1.08, 1.12, 1.17, 1.18, 1.22, 1.25, 1.36, 1.41, 1.47, 1.57, 1.57, 1.59, 1.59, 1.61, 1.61, 1.69, 1.69, 1.71, 1.73, 1.8, 1.84, 1.84, 1.87, 1.89, 1.92, 2, 2.03, 2.03, 2.05, 2.12, 2.17, 2.17, 2.17, 2.35, 2.38, 2.41, 2.43, 2.48, 2.48, 2.5, 2.53, 2.55, 2.55, 2.56, 2.59, 2.67, 2.73, 2.74, 2.76, 2.77, 2.79, 2.81, 2.81, 2.82, 2.83, 2.85, 2.87, 2.88, 2.93, 2.95, 2.96, 2.97, 2.97, 3.09, 3.11, 3.11, 3.15, 3.15, 3.19, 3.19, 3.22, 3.22, 3.27, 3.28, 3.31, 3.31, 3.33, 3.39, 3.39, 3.51, 3.56, 3.6, 3.65, 3.68, 3.68, 3.68, 3.7, 3.75, 4.2, 4.38, 4.42, 4.7, 4.9, 4.91, 5.08, 5.56.

Figure 8 presents the box-plot and histogram of the carbon fibre data set. It can be observed that, the data set is positively skewed and leptokurtic in nature.

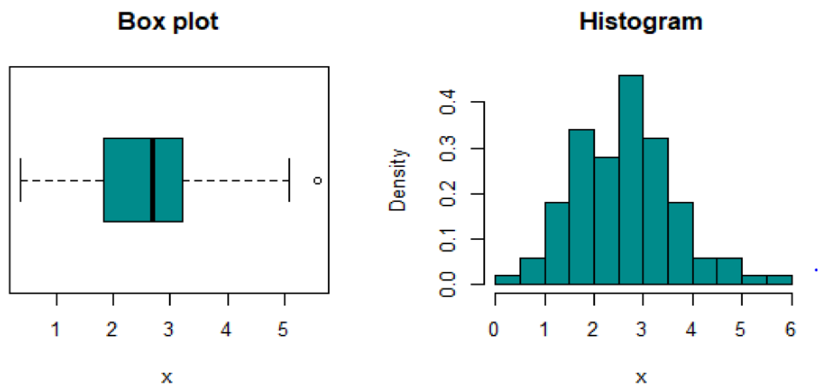


Figure 8: Box Plot and Histogram of the Carbon Fibre Dataset

The positively skewness and leptokurtic nature of the data is further confirmed by descriptive statistics of the breaking stress of carbon fibres dataset given in Table 6 (skewness greater than 0 and kurtosis greater than 3).

Table 6
Descriptive Statistics of Carbon Fibre Dataset

No. of Obs.	Mean	Std. dev.	Skewness	Kurtosis	Min.	Max.
100	2.621	1.014	0.368	3.105	0.390	5.560

The TTT transformed plot of the 100 uncensored data on breaking stress of carbon fibres (in Gba) presented in Figure 9 shows that the hazard function of the data set is increasing.

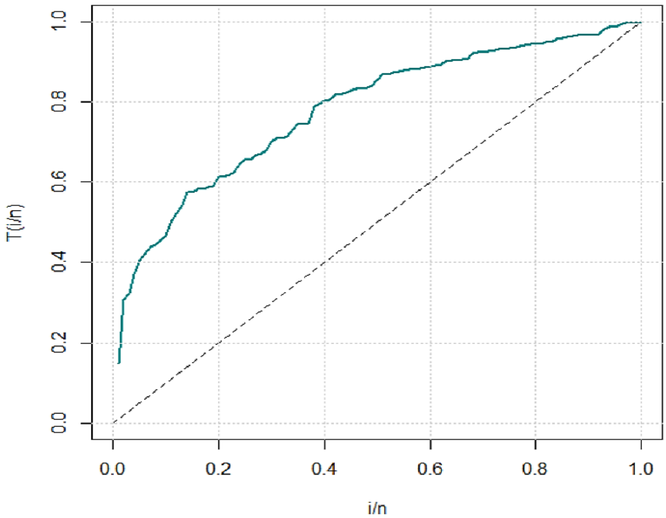


Figure 9: TTT-Transform Plot for Carbon Fibre Dataset

Table 7 shows the maximum likelihood estimates for the parameters of the fitted distributions with their corresponding standard errors in brackets and p-values. The parameters of the BIII, Weibull, TSDa, SBXII and SILF are significant at the 5% level. All the parameters of Dagum distribution are significant at the 5% significant level with the exception of γ which is significant at 10%. The parameters of the TSD distribution are significant at 10% with the exception of θ, which is significant at the 5% level.

Table 7
Maximum Likelihood Estimates of the Parameters, Standard Errors and p-value of Carbon Fibre Dataset

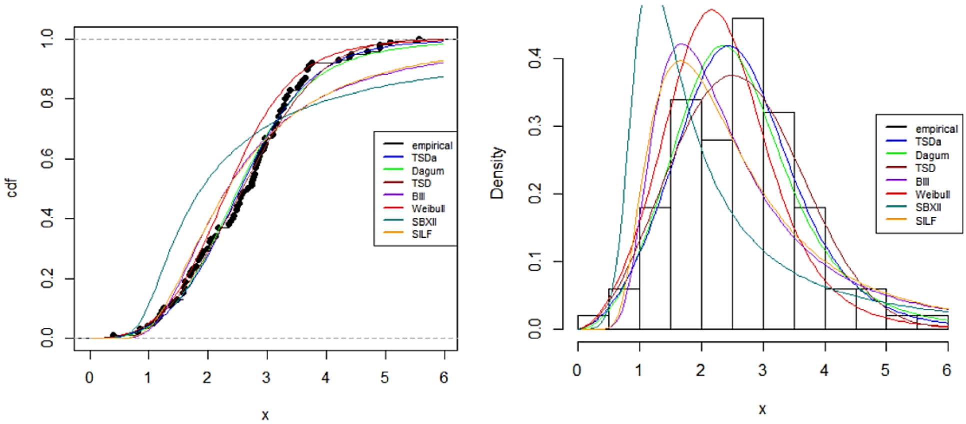
Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{\sigma}$	$\hat{\theta}$	$\hat{c}$	$\hat{k}$
TSDa	80.980 (23.500) 0.003	0.629 (0.154) 0.000	4.616 (0.433) 0.000	0.599 (0.159) 0.000				
Dagum	5.010 (0.514) 0.000	0.609 (0.118) 0.000	214.557 (160.770) 0.073					
TSD		0.739 (0.227) 0.066		-0.376 (0.463) 0.085	78.349 (70.572) 0.072	3.719 (0.515) 0.004		
BIII							2.319 (0.150) 0.000	5.221 (0.598) 0.000
Weibull	2.792 (0.214) 0.000	0.049 (0.014) 0.000						
SBXII							5.822 (1.261) 0.003	0.109 (0.025) 0.001
SILF	3.891 (0.010) 0.000	$\hat{\mu}$ 0.982 (0.060) 0.000	$\hat{\delta}$ 1.463 (0.093) 0.000					

Table 8 presents the information criteria and goodness-of-fit statistics for the 100 uncensored data on breaking stress of carbon fibres (in Gba). Also, the TSDa has the lowest AIC, BIC, W*, A*, K-S statistics values hence provides the best fit to the data among the candidate distributions.

Table 8
Information Criteria and Goodness-of-fit of Carbon Fibre Dataset

Model	$-2l$	AIC	BIC	W*	A*	K-S	p-value
TSDa	**283.059**	**287.059**	**292.269**	**0.062**	**0.416**	**0.060**	**0.858**
Dagum	286.169	292.169	299.984	0.136	0.692	0.087	0.442
TSD	283.593	291.593	302.014	0.537	3.009	0.944	0.002
BIII	321.946	325.946	331.157	0.569	3.173	0.140	0.041
Weibull	287.094	295.094	305.515	0.146	0.745	0.086	0.457
SBXII	369.997	373.997	379.207	0.800	4.497	0.259	0.002
SILF	323.159	329.159	336.975	0.530	2.993	0.136	0.050

The empirical CDF and the CDFs of the tentative distributions for the 100 uncensored data on breaking stress of carbon fibres (in Gba) is presented in Figure 10. It can be seen that, the TSDa provides a better to the data than the competitive distribution.



(a) (b)

Figure 10: CDF (a) and Empirical and Fitted Density (b) Plots of Carbon Fibre Dataset

7. CONCLUSION

Using the Dagum distribution as a baseline distribution in the transformed Sin-G family of distribution, the transformed sine Dagum (TSDa) distribution is proposed and studied. The statistical properties of the proposed distribution are derived and inferences made. The density plots of the TSDa distribution can be decreasing, right skewed and approximately symmetric. The hazard rate function can be increasing, decreasing, bathtub, upside-down-bathtub, modified upside-down-bathtub among others. This implies that the TSDa distribution can adequately model both monotonic and non-monotonic failure rate data. The skewness and kurtosis plots show that the skewness is positive, an indication that the TSDa is right skewed and the kurtosis is increasing; meaning TSDa distribution is leptokurtic. The estimators of the parameters of the TSDa distribution are developed using the maximum likelihood estimation. Monte Carlo simulation is performed on the parameters of the new distribution and the results shows that the parameters are consistent. The new distribution proposed is applied to data on the length of intervals between the times at which vehicles pass a point on a road and 100 uncensored data on breaking stress of carbon fibres (in Gba). Both applications shows that the TSDa distribution provides a better fit than the 3-parameter Dagum, 2-parameter Weibull, Transmuted sine Dagum (TSD), Burr III (BIII), sine Burr XII (SBXII) and Sine Inverse Lomax Frechet distributions. It is hope that the proposed distribution will gain attention in medical, economic, social sciences, engineering and other related field.

ACKNOWLEDGMENTS

The author wants to thank the editor and the reviewers for their constructive comments and suggestions which led to this improved version.

REFERENCES

1. Lemonte, A.J., Barreto-Souza, W. and Cordeiro, G.M. (2013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics*, 27(1), 31-53.
2. Khadim, A., Saghir, A., Hassain, T., Shakil, M. and Ahsanullah, M. (2021). A log-Dagum Weibull distribution: Properties and Characterization. Applied and Computational Mathematics, 10(5), 100-113.
3. Jamal, F., Chesneau, C., Bouali, D.L. and Hassan, M.U. (2021). Beyond the Sin-G family: The transformed Sin-G family. *PLoS ONE*, 16(5), 1-22.
4. Souza, L., Junior, W.R.O., de Brito, C.C.R., Chesneau C., Ferreira T.A.E. and Soares L., (2019). General properties for the Cos-G class of distributions with applications. *Eurasian Bulletin of Mathematics*, 2(2), 63-79.
5. Chesneau, C., Bakouch, H.S. and Hussain, T. (2019). A new class of probability distributions via cosine and sine functions with applications. *Communications in Statistic-Simulation and Computation*, 48(8), 2287-2300. https://doi.org/10.1080/03610918.2018.1440303.
6. Mahmood, Z., Chesneau, C. and Tahir, M.H. (2019). A new sine-G family of distributions: properties and applications. *Bulletin of Computational Applied Mathematics*, 7(1), 53-81.
7. Dagum, C. (1977). A systematic approach to the generation of income distribution models. *Journal of Income Distribution*, 6(1), 1-6.
8. Domma, F. and Condino, F. (2013). The beta-dagum distribution: Definition and properties. *Communications in Statistics-Theory and Methods*, 42(22), 4070-4090.
9. Domma, F., Giordano, S. and Zenga, M.M. (2011). Maximum likelihood estimation in Dagum distribution with censored sample. *Journal of Applied Statistics*, 38(12), 2971-2985.
10. Oluyede, B.O. and Ye, Y. (2014). Weighted Dagum and related distributions. *Afrika Matematika*, 25(4), 1125-1141.
11. Rodrigues, J.A. and Silva, A.P.C.M. (2015). The Gamma-Dagum distribution: Definition, properties and application. *Matematica e Estatistica em Foco*, 3(1), 1-7.
12. Elbatal, I. and Aryal, G. (2015). Transmuted Dagum distribution with applications. *Chilean Journal of Statistics (ChJS)*, 6(2), 31-45.
13. Domma, F. (2004). Kurtosis diagram for the log-Dagum distribution. *Statistica and Applicatzioni*, 2, 3-23.
14. Domma, F., Giordano, S. and Zenga, M. (2013). The Fisher information matrix on a type II doubly censored sample from a Dagum distribution. *Applied Mathematical Sciences*, 7, 3715-3729.
15. Khadim, A., Saghir, A., Hussain, T., Shakil, M. and Ahsanullah, M. (2021). A Log-Dagum Weibull Distribution: Properties and Characterization. *Applied and Computational Mathematics*, 10(5), 100-113.

16. Domma, F. (2002). L'andamento della hazard function nel modello di Dagum a tre parametri. *Quaderni di Statistica*, 4, 1-12.
17. Domma, F., Eftekharian, A., Afify, A., Alizadeh, M. and Ghosh, I. (2018). The odd log-logistic Dagum distribution: properties and applications. *Revista Colombiana de Estadistica*, 41(1), 109-135.
18. Wooff, D.A. (1985). Bounds on reciprocal moments with applications and developments in Stein estimation and post-stratification. *Journal of the Royal Statistical Society*, Series B, 47(2), 362-371.
19. Pittenger, A.O. (1990). Sharp mean-variance bounds for Jensen-type inequalities. *Statistics and Probability Letters*, 10(2), 91-94.
20. Marciniak, E. and Wesolowski, J. (1999). Asymptotic Eulerian expansions for binomial and negative binomial reciprocals. *Proceedings of the American Mathematical Society*, 127(11), 3329-3338.
21. Fujioka, T. (2001). Asymptotic approximations of the inverse moment of the non-central chi-squared variable. *Journal of the Japan Statistical Society*, 31(1), 99-109.
22. Sakthivel, K.M. and Dhivakar, K. (2021). Transmuted Sine-Dagum Distribution and its Properties. *Reliability: Theory and Applications*, 16, 150-166.
23. Jurlewicz, A. and Weron, K. (2002). Relaxation of dynamically correlated clusters, *Journal of Non-Crystalline Solids*, 305(1-3), 112-121.
24. Ramsay, C.M. (1993). A note on random survivorship group benefits. *ASTIN Bulletin*, 23, 149-156.