RT&A, No 3 (79)
Volume 19, September 2024

STATISTICAL PROPERTIES AND APPLICATIONS OF TRANSMUTED SKEW STUDENT *t* DISTRIBUTION

DAVID Ikwuoche John, MATHEW Stephen
•

Department of Mathematics and Statistics, Federal University Wukari, Nigeria
davidij@fuwukari.edu.ng, matsteve231@gmail.com
https://orcid.org/0000-0002-7100-5357

Abstract

In this study, a modified 2-parameter skew t distribution called the transmuted skew student t distribution (TSStD) was presented. Some statistical and reliability properties of TSStD such as the quantile function, the raw moments, and the moment generating function (among others), were derived. Through the method of maximum likelihood, the two parameters of the model were estimated. The stability of the model was studied via Montecarlo simulations utilizing bias, mean square error, and root mean square error as metrics. The results from the stability study revealed that the TSStD was well-behaved. Four datasets were modeled with the transmuted skewed student t distribution and four other probability density models. On the basis of information criteria, the results revealed that the transmuted skew student t distribution provides a better fit for all the datasets compared to the other competing models.

Keywords: Transmuted, Skew, Raw moments, Quantile, Reliability function, Hazard function

I. Introduction

Empirical probability models, or probability distributions, are an essential aspect of parametric statistical investigation. Classical distributions are more susceptible to an anomaly when characterizing various data and data generating processes, according to a prevalent reality across a variety of sectors, including finance, environmental science, biological sciences, engineering, and others. This emphasizes the need to hybridize classical probability models in order to meet this complex task [1]. Owing to the applicability of distribution theory and the availability of diverse data in today's world, the thirst for improved statistical distributions that might be used to describe and model these events have increased spontaneously [2], [3], [4], [5], and [6]. Many scholars have over the years help modified simple statistical methodologies in relation to distribution theory and these methods have been found immensely useful in statistical modeling. Several methods of modifying probability distributions have been proposed over the years by scholars to improve statistical methodology of distribution theory. Among the methods are Exponentiated Exponential Distribution [7], the Sine-G family [8], the New Sine-G Family [9], the G-families using the transformed-transformer [10], Transmuted G family by [11], and amongst others. Several distributions have been modified using the Transmuted G family of distributions over the years and some of them include Transmuted Lomax and Transmuted Exponentiated Lomax distributions [12], Transmuted Frechet distribution [13], Transmuted Exponentiated Gamma distribution [14], Transmuted additive Weibull distribution [15], Transmuted generalized Gompertz distribution [16], Kumaraswamy Transmuted Exponentiated modified Weibull distribution [17], Transmuted Exponential power distribution [18], etc. Using the method of transmutation by [19], this paper proposes a new probability distribution called the Transmuted Skew Student *t* Distribution (TSS*t*D).

II. Methods

I. The Transmuted Family of Distribution

This section presents the TSS*t*D, some of its statistical properties, simulation study as well as application to real life data. The cumulative density function (CDF) and probability density function (PDF) of the Transmuted family of distribution generator is given by;

$$F(v) = (1+\alpha)G(v) - \alpha G(v)^2 \tag{1}$$

$$f(v) = g(v)\{(1+\alpha) - 2\alpha G(v)\} \tag{2}$$

where, α is the transmuted parameter (shape), $F(v)$ is the CDF and $f(v)$ is the PDF of the baseline distribution.

II. The Skew Student t Distribution

Using the simplified version of the Skew Student *t* Distribution (SS*t*D) with 2 degrees of freedom introduced by Jonson *et al*., [20] whose CDF and PDF are expressed as;

$$\left.\begin{aligned} G(v) &= \frac{1}{2}\left(1+\frac{v}{\sqrt{\Lambda+v^2}}\right) \\ g(v) &= \frac{\Lambda}{2\left(\Lambda+v^2\right)^{3/2}} \end{aligned}\right\}; -\infty < v < \infty \tag{3}$$

where, Λ is the shape parameter.

III. Transmuted Skew Student *t* Distribution (TSS*t*D)

On substituting *G(v)* and *g(v)* in equation (3) into (1) and (2) the PDF and CDF of TSS*t*D are obtained as;

$$f(v;\Lambda,\alpha) = \frac{\Lambda}{2\left(\Lambda+v^2\right)^{\frac{3}{2}}}\left[(1+\alpha) - 2\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right] \tag{4}$$

and

$$\mathrm{F}(\mathrm{v};\Lambda,\alpha) = (1+\alpha)\left(\frac{1}{2}+\frac{\mathrm{v}}{2\sqrt{\Lambda+\mathrm{v}^2}}\right) - \alpha\left(\frac{1}{2}+\frac{\mathrm{v}}{2\sqrt{\Lambda+\mathrm{v}^2}}\right)^2 \tag{5}$$

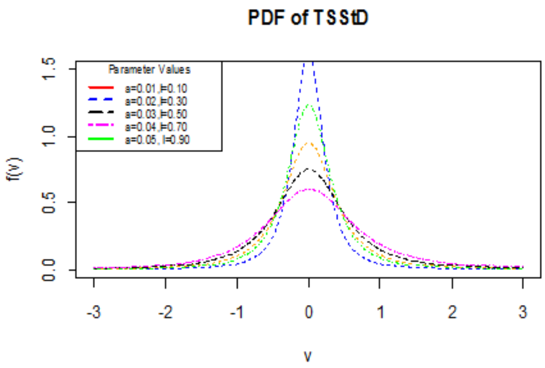

Figure 1: The *PDF Plot of TSStD*

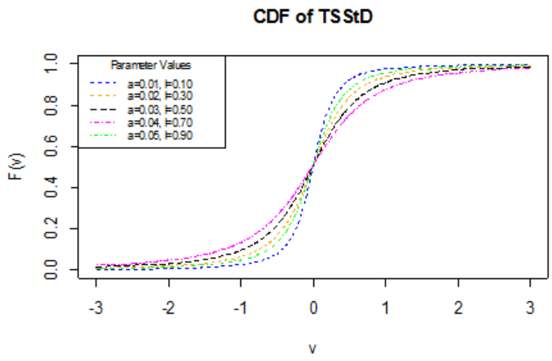

Figure 2: The *CDF Plot of TSStD*

IV. Properties of Transmuted Skew Student *t* Distribution

The survival [$S(x)$], hazard [$h(x)$], odd [$O(x)$], and quantile [$Q(v)$] functions, skewness, and kurtosis statistics are presented as well as the $S(x)$ plot as follows

$$S(x)=1-F(V\leq v)=1-\left((1+\alpha)\left(\frac{1}{2}+\frac{v}{2\sqrt{\Lambda+v^2}}\right)-\alpha\left(\frac{1}{2}+\frac{v}{2\sqrt{\Lambda+v^2}}\right)^2\right) \tag{6}$$

$$h(x)=\frac{f(x)}{S(x)}=\frac{\dfrac{\Lambda}{2\left(\Lambda+v^2\right)^{\frac{3}{2}}}\left[(1+\alpha)-2\alpha\left(\dfrac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]}{1-\left((1+\alpha)\left(\dfrac{1}{2}+\dfrac{v}{2\sqrt{\Lambda+v^2}}\right)-\alpha\left(\dfrac{1}{2}+\dfrac{v}{2\sqrt{\Lambda+v^2}}\right)^2\right)} \tag{7}$$

$$O(x)=\frac{F(x)}{S(x)}=\frac{(1+\alpha)\left(\dfrac{1}{2}+\dfrac{v}{2\sqrt{\Lambda+v^2}}\right)-\alpha\left(\dfrac{1}{2}+\dfrac{v}{2\sqrt{\Lambda+v^2}}\right)^2}{1-\left((1+\alpha)\left(\dfrac{1}{2}+\dfrac{v}{2\sqrt{\Lambda+v^2}}\right)-\alpha\left(\dfrac{1}{2}+\dfrac{v}{2\sqrt{\Lambda+v^2}}\right)^2\right)} \tag{8}$$

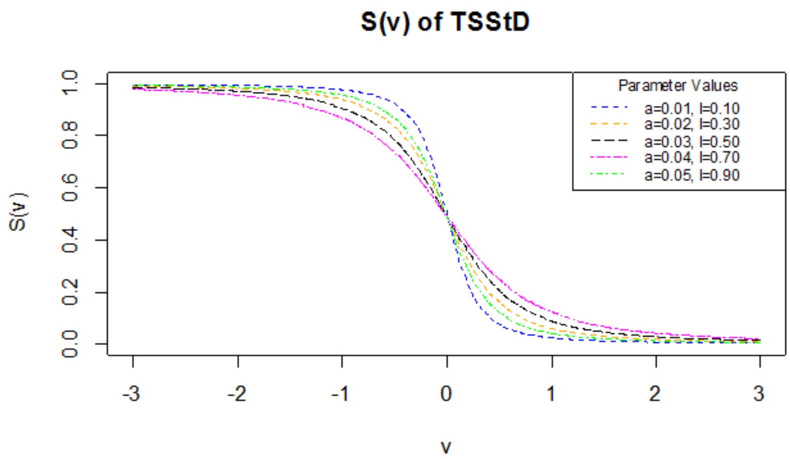

Figure 3: The S(v) *Plot of TSStD for varied parameter value*

Quantile Function: the inverse of the CDF of *TSStD* gives the quantile function of *TSStD* and after numerous algebraic simplifications, this is expressed as;

$$Q(v)=\pm\left(-\frac{\left(\alpha^2\Lambda^2-4\alpha\Lambda^2u^*+2\alpha\Lambda^2+\Lambda^2\right)^{\frac{1}{2}}}{8\left(u^{*2}-u^*\right)}+\frac{\alpha\Lambda u^*}{4\left(u^{*2}-u^*\right)}-\frac{\alpha\Lambda}{8\left(u^{*2}-u^*\right)}-\frac{\Lambda u^{*2}}{u^{*2}-u^*}+\frac{\Lambda u^*}{u^{*2}-u^*}-\frac{\Lambda}{8\left(u^{*2}-u^*\right)}\right)^{\frac{1}{2}} \quad (9)$$

Skewness: the Galton measure of skewness [GSK] which measures the presence and lack of symmetry of a probability distribution is presented for *TSStD* using *Q*(*v*) as follows;

$$GSK=\frac{Q\left(\frac{1}{8}\right)+Q\left(\frac{3}{4}\right)-Q\left(\frac{1}{4}\right)-Q\left(\frac{1}{2}\right)}{Q\left(\frac{3}{4}\right)-Q\left(\frac{1}{4}\right)} \quad (10)$$

Kurtosis: the kurtosis measure whether or not the probability distribution is heavy-tailed. For TSS*t*D the Moor's kurtosis measure is derived using *Q*(*v*) as follows;

$$MKT=\frac{Q\left(\frac{1}{8}\right)+Q\left(\frac{3}{8}\right)-Q\left(\frac{5}{8}\right)-Q\left(\frac{1}{4}\right)}{Q\left(\frac{3}{4}\right)-Q\left(\frac{1}{4}\right)} \quad (11)$$

IV. Moments, Moment Generating Function, and Characteristic Function

Moments: A random variable v with TSS*t*D has its moment in the form;

$$u_r=\frac{\lambda}{2}\left[G\beta\left[\frac{r^2-r+2}{2},\frac{r^2}{2}\right]-H\beta\left[\frac{r+2}{2},\frac{r}{2}\right]\right] \quad (12)$$

Proof

The r^{th} moment of a random variable *v* with a valid probability distribution is given by;

$$u_r=E\left[v^r\right]=\int_{-\infty}^{\infty}v^r f\left(v\right)dv \quad (13)$$

Let *v* be a random variable following the TSS*t*D. Then the moment is derived as follows;

$$u_r=E\left[v^r\right]=\int_{-\infty}^{\infty}v^r\frac{\Lambda}{2\left(\Lambda+v^2\right)^{\frac{3}{2}}}\left[1-\frac{\alpha v}{\left(\Lambda+v^2\right)^{\frac{1}{2}}}\right]dv \quad (14)$$

$$u_r=\int_{-\infty}^{\infty}\frac{\Lambda v^r}{2}\left[\frac{1}{\left(\Lambda+v^2\right)^{\frac{3}{2}}}-\frac{\alpha v}{\left(\Lambda+v^2\right)^{\frac{1}{2}+\frac{3}{2}}}\right]dv=\frac{2\Lambda}{2}\int_{0}^{\infty}\left[\frac{v^r}{\left(\Lambda+v^2\right)^{\frac{3}{2}}}-\frac{\alpha v v^r}{\left(\Lambda+v^2\right)^{2}}\right]dv$$

$$u_r=\Lambda\int_{0}^{\infty}\left[\frac{v^r}{\left(\Lambda+v^2\right)^{\frac{3}{2}}}-\frac{\alpha v^{r+1}}{\left(\Lambda+v^2\right)^{2}}\right]dv \quad (15)$$

By transformation, let

$u=v^r$; $v=u^{\frac{1}{r}}$; $\frac{du}{dv}=rv^{r-1}$; $dv=\frac{1}{rv^{r-1}}du$. Substituting the transformations into (15) to obtain;

$$u_r = \Lambda \int_0^\infty \left[\frac{u^{r\left(\frac{1}{r}\right)}}{\left(\Lambda + u^{2\left(\frac{1}{r}\right)}\right)^{\frac{3}{2}}} - \frac{\alpha u^{\frac{1}{r}(r+1)}}{\left(\Lambda + u^{2\left(\frac{1}{r}\right)}\right)^2} \right] \frac{1}{r u^{\left(\frac{1}{r}\right)(r-1)}} du \tag{16}$$

$$u_r = \frac{\Lambda}{r} \left[\int_0^\infty \frac{u^{1-\frac{r-1}{r}}}{\left(\Lambda + u^{\frac{2}{r}}\right)^{\frac{3}{2}}} du - \int_0^\infty \frac{\alpha u^{\frac{r+1}{r} - \left(\frac{r-1}{r}\right)}}{\left(\Lambda + u^{\frac{2}{r}}\right)^2} du \right] = \frac{\lambda}{r} \left[\int_0^\infty \frac{u^{\frac{r-r+1}{r}}}{\left(\Lambda + u^{\frac{2}{r}}\right)^{\frac{3}{2}}} du - \int_0^\infty \frac{\alpha u^{\frac{r+1-r+1}{r}}}{\left(\Lambda + u^{\frac{2}{r}}\right)^2} du \right]$$

$$u_r = \frac{\Lambda}{r} \left[\Lambda^{-\frac{3}{2}} \int_0^\infty \frac{u^{\frac{1}{r}}}{\left(1 + \frac{u^{\frac{2}{r}}}{\Lambda}\right)^{\frac{3}{2}}} du - \frac{\alpha}{\Lambda^2} \int_0^\infty \frac{u^{\frac{2}{r}}}{\left(1 + \frac{u^{\frac{2}{r}}}{\Lambda}\right)^2} du \right] \tag{17}$$

Also, letting $k = \frac{u^{\frac{2}{r}}}{\Lambda}, \Rightarrow u = (\Lambda k)^{\frac{r}{2}}, \Rightarrow \frac{dk}{du} = \frac{2}{\Lambda r} u^{\frac{2-r}{r}}, \Rightarrow du = \frac{\Lambda r}{2u^{\frac{2-r}{r}}} dk$ then by substituting for u and du in (17), we obtain the following;

$$u_r = \frac{\Lambda}{r} \left[\Lambda^{-\frac{3}{2}} \int_0^\infty \frac{(\Lambda k)^{\frac{r}{2}\left(\frac{1}{r}\right)}}{(1+k)^{\frac{3}{2}}} \frac{\Lambda r}{2(\Lambda k)^{\frac{r}{2}\left(\frac{2-r}{r}\right)}} dk - \frac{\alpha}{\Lambda^2} \int_0^\infty \frac{(\Lambda k)^{\frac{r}{2}\left(\frac{2}{r}\right)}}{(1+k)^2} \frac{\Lambda r}{2(\Lambda k)^{\frac{r}{2}\left(\frac{2-r}{r}\right)}} dk \right]$$

$$u_r = \frac{\Lambda}{2} \left[\frac{\Lambda^{\frac{r^2-r}{2}}}{\Lambda^{\frac{1}{2}}} \int_0^\infty \frac{k^{\frac{r^2-r}{2}}}{(1+k)^{\frac{3}{2}}} dk - \frac{\alpha \Lambda^{\frac{r}{2}}}{\Lambda} \int_0^\infty \frac{k^{\frac{r}{2}}}{(1+k)^2} dk \right] = \frac{\Lambda}{2} \left[G \int_0^\infty \frac{k^{\frac{r^2-r}{2}}}{(1+k)^{\frac{3}{2}}} dk - H \int_0^\infty \frac{k^{\frac{r}{2}}}{(1+k)^2} dk \right] \tag{18}$$

where, $G = \frac{\Lambda^{\frac{r^2-r}{2}}}{\Lambda^{\frac{1}{2}}}$; $H = \frac{\alpha \Lambda^{\frac{r}{2}}}{\Lambda}$. Therefore, the Beta function is given as;

$$\beta[p,q] = \int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}} dt \tag{19}$$

Therefore, on transforming (18) into the form presented in (19), the following is obtained;

$$u_r = \frac{\Lambda}{2} \left[G \int_0^\infty \frac{k^{\frac{r^2-r}{2}+1-1}}{(1+k)^{\frac{3}{2}+\frac{r^2}{2}-\frac{r^2}{2}}} dk - H \int_0^\infty \frac{k^{\frac{r}{2}+1-1}}{(1+k)^{2+\frac{r^2}{2}-\frac{r^2}{2}}} dk \right] = \frac{\Lambda}{2} \left[G\beta \left[\frac{r^2-r+2}{2}, \frac{r^2}{2} \right] - H\beta \left[\frac{r+2}{2}, \frac{r}{2} \right] \right] \tag{20}$$

which completes the proof.

The *mean* of TSS*t*D, that is, E[v] based on equation (20) by substituting r to be 1 is given as;

$$E[v] = u_1 = \frac{\Lambda}{2} \left[G\beta \left[\frac{1^2-1+2}{2}, \frac{1^2}{2} \right] - H\beta \left[\frac{1+2}{2}, \frac{1}{2} \right] \right] = \frac{\Lambda}{2} \left[G\beta \left[1, \frac{1}{2} \right] - H\beta \left[\frac{3}{2}, \frac{1}{2} \right] \right] \tag{21}$$

Moment Generating Function (MGF): The MGF of the random variable v with TSS*t*D is given by:

$$M_v(t) = \sum_{m=0}^\infty \frac{t^m}{m!} \cdot \frac{\Lambda}{2} \left[G\beta \left[\frac{m^2-m+2}{2}, \frac{m^2}{2} \right] - H\beta \left[\frac{m+2}{2}, \frac{m}{2} \right] \right] \tag{22}$$

Proof

The MGF is given by:

$$M_v(t) = E[e^{tv}] = \int_{-\infty}^{\infty} e^{tv} f(v) d \tag{23}$$

By McLaurin's series expansion, e^{tv} is expressed as $e^{tv} = \sum_{n=0}^{\infty} \frac{(tv)^m}{m!}$ then equation (23) becomes

$$M_v(t) = E[e^{tv}] = E\left[\sum_{m=0}^{\infty} \frac{(tv)^m}{m!}\right] = \sum_{m=0}^{\infty} \frac{t^m}{m!} E\left[v^m\right] \tag{24}$$

As obtained in equation (20), $E\left[v^m\right] = u_m$, therefore,

$$M_v(t) = \sum_{m=0}^{\infty} \frac{t^m}{m!} . \frac{\Lambda}{2}\left[G\beta\left[\frac{m^2 - m + 2}{2}, \frac{m^2}{2}\right] - H\beta\left[\frac{m+2}{2}, \frac{m}{2}\right]\right] \tag{25}$$

Characteristic Function [CF]: for a random variable v with TSS*t*D, the CF is given as;

$$f_v(t) = \sum_{c=0}^{\infty} \frac{(it)^c}{c!} . \frac{\Lambda}{2}\left[G\beta\left[\frac{r^2 - r + 2}{2}, \frac{r^2}{2}\right] - H\beta\left[\frac{r+2}{2}, \frac{r}{2}\right]\right] \tag{26}$$

Proof

The CF of a valid probability distribution is expressed as;

$$f_v(t) = E\left[e^{itv}\right] = \int_{-\infty}^{\infty} e^{itv} f(v) dv \tag{27}$$

By Taylor's series expansion of $e^{itv} = \sum_{i=0}^{\infty} \frac{(itv)^c}{c!}$, therefore,

$$f_v(t) = \sum_{c=0}^{\infty} \frac{(itv)^c}{c!} \frac{\Lambda}{2}\left[G\beta\left[\frac{r^2 - r + 2}{2}, \frac{r^2}{2}\right] - H\beta\left[\frac{r+2}{2}, \frac{r}{2}\right]\right] \tag{28}$$

V. Order Statistics for TSS*t*D

Sample values such as the smallest, largest, or middle observation from a random sample provide important information. Order Statistics could be used to determine the distribution of the smallest

(minimum) order statistic and the largest (maximum) order statistic of a given distribution. Let V_1, V_2, …, V_n denote n-independent random sample from a distribution function $F(v)$ and probability density function, $f(v)$, then $v_1, v_2, \ldots, v_n$ represent the order sample arrangement and the pdf of $v_{(n)}$ is given by:

$$f_{i:n}(v) = \frac{n!}{(i-1)!(n-i)!} f(v) F(v)^{i-1}\left[1 - F(v)\right]^{n-i} \text{ ; for } i = 1, 2, \ldots, n. \tag{29}$$

For simplicity, $[1 - F(v)]^{n-i}$ in (29) can be expressed using the sum of a binomial series as $\sum_{l=0}^{\infty} \binom{n-i}{m} (-1)^m \left[F(v)\right]^m$. By substitution the following is obtained:

$$f_{i:n}(v) = \frac{n!}{(i-1)!(n-i)!} f(v) \sum_{m=0}^{\infty} \binom{n-i}{m} (-1)^m F(v)^{m+i-1} \tag{30}$$

Now, making the substitution of $f(v)$ and $F(v)$ into (30) will yield the following;

$$\left.\begin{aligned} f_{i:n}(v) &= \frac{n!}{(i-1)!(n-i)!}\left[\frac{\Lambda}{2\left(\Lambda+v^2\right)^{\frac{3}{2}}}\left[1-\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]\right] \\ &\times\sum_{m=0}^{\infty}\binom{n-i}{m}(-1)^m\left[(1+\alpha)\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)-\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)^2\right]^{m+i-1}\end{aligned}\right\} \tag{31}$$

Applying binomial expansion to the term $\left[(1+\alpha)\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)-\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)^2\right]^{m+i-1}$ to get

$\sum_{q=0}^{m+i-1}\binom{m+i-1}{q}\left[(1+\alpha)\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]^q\left[-\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)^2\right]^{m+i-1-q}$ and when substituted into (31), the $f_{i:n}(v)$ becomes;

$$\left.\begin{aligned} f_{i:n}(v) &= \frac{n!}{(i-1)!(n-i)!}\left[\frac{\Lambda}{2\left(\Lambda+v^2\right)^{\frac{3}{2}}}\left[1-\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]\right] \\ &\times\sum_{m=0}^{\infty}\binom{n-i}{m}(-1)^m\sum_{q=0}^{m+i-1}\binom{m+i-1}{q}\left[(1+\alpha)\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]^q\left[-\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)^2\right]^{m+i-1-q}\end{aligned}\right\} \tag{32}$$

Expanding the term $\left[(1+\alpha)\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]^q$ binomially yields $\sum_{r=0}^{q}\binom{q}{r}\alpha^{q-r}\left[\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]^q$ and on substitution into (32), $f_{i:n}(v)$ becomes,

$$\left.\begin{aligned} f_{i:n}(v) &= \frac{n!}{(i-1)!(n-i)!}\left[\frac{\Lambda}{2\left(\Lambda+v^2\right)^{\frac{3}{2}}}\left[1-\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]\right] \\ &\times\sum_{m=0}^{\infty}\binom{n-i}{m}(-1)^m\sum_{q=0}^{m+i-1}\binom{m+i-1}{q}\sum_{r=0}^{q}\binom{q}{r}(-\alpha)^{m+i-1-q+q-r}\left[\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)^2\right]^{m+i-1-q+q}\end{aligned}\right\} \tag{33}$$

$$f_{i:n}(v) = \frac{n!}{(i-1)!(n-i)!}\left[\frac{\Lambda}{2\left(\Lambda+v^2\right)^{\frac{3}{2}}}\left[1-\alpha\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)\right]\right]H_{i:n}(-\alpha)^{m+i-1-r}\left[\left(\frac{\sqrt{\Lambda+v^2}+v}{2\sqrt{\Lambda+v^2}}\right)^2\right]^{m+i-1} \tag{34}$$

where, $H_{i:n} = \sum_{m=0}^{\infty}\binom{n-i}{m}(-1)^m\sum_{q=0}^{m+i-1}\binom{m+i-1}{q}\sum_{r=0}^{q}\binom{q}{r}$.

VI. Parameter Estimation Using Maximum Likelihood Estimation Technique

Let $l(\theta)$ be a parameter vector for the transmuted family of distributions. Consider a random variable $v \sim TSStD(\Lambda, \alpha)$, then by definition, the likelihood function of v with PDF $f(v)$ is given as;

$$l(\theta) = l\left(f(v;\Lambda)\right) = \prod_{i=1}^{n}\left[\Lambda\left[\frac{\sqrt{\left(\Lambda+v^2\right)}-\alpha v}{\sqrt{\left(\Lambda+v^2\right)}}\right]2\left(\Lambda+v^2\right)^{-\frac{3}{2}}\right] \tag{35}$$

$$= \Lambda^n 2^{-n} \left[1 - \frac{\alpha \sum_{i=1}^{n} v}{\sum_{i=1}^{n} \sqrt{\left(\Lambda + v^2\right)}} \right] \sum_{i=1}^{n} \left(\Lambda + v^2\right)^{-\frac{3}{2}} \tag{36}$$

The log of $l(\theta)$ is as follows;

$$\log\left(l(\theta)\right) = \log\left(\Lambda^n \left[\frac{\sum_{i=1}^{n} \sqrt{\left(\Lambda + v^2\right)} - \alpha^n \sum_{i=1}^{n} v}{\sum_{i=1}^{n} \sqrt{\left(\Lambda + v^2\right)}} \right] \right) - \log\left(2^n \sum_{i=1}^{n} \left(\Lambda + v^2\right)^{\frac{3}{2}} \right) \tag{37}$$

Differentiating equation (37) with respect to the parameters Λ and α will yield the estimates of the parameters. The differentiation w.r.t. Λ is obtained as

$$\frac{d\left[\log\left(l(\theta)\right)\right]}{d\Lambda} = \frac{1 + n\alpha^n \sum_{i=1}^{n} v \left(\sum_{i=1}^{n} \left(\Lambda + v^2\right)^{-\frac{3}{2}} \right)}{n} \left[\frac{\sum_{i=1}^{n} \sqrt{\left(\Lambda + v^2\right)} - \alpha^n \sum_{i=1}^{n} v}{\sum_{i=1}^{n} \sqrt{\left(\Lambda + v^2\right)}} \right]^{-1} - \left(3 \times 2^{n-1} \sum_{i=1}^{n} \left(\Lambda + v^2\right)^{\frac{3}{2}} \right)^{-1} \sum_{i=1}^{n} \left(\Lambda + v^2\right)^{\frac{5}{2}} = 0 \cdot$$

The result obtained shows that the parameter Λ does not exist in a closed form. A numerical estimate for the parameter will be obtained using R-software. Differentiation w.r.t. α gives

$$\frac{d\left[\log\left(l(\theta)\right)\right]}{d\alpha} = -\frac{n\alpha^{n-1} \sum_{i=1}^{n} v}{\sum_{i=1}^{n} \sqrt{\left(\Lambda + v^2\right)}} \left[\frac{\sum_{i=1}^{n} \sqrt{\left(\Lambda + v^2\right)} - \Omega^n \sum_{i=1}^{n} v}{\sum_{i=1}^{n} \sqrt{\left(\Lambda + v^2\right)}} \right]^{-1} = 0$$

$$\hat{\alpha} = \sqrt[n]{\frac{\sum_{i=1}^{n} \left(\Lambda + v^2\right)^{1/2}}{\sum_{i=1}^{n} v}} \tag{38}$$

VII. Measures of Goodness-of-Fit Adopted

This section presents the measures used in model selection. They include the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Corrected AIC (CAIC), and Hannan-Quinn Information Criterion (HQIC).

$$AIC = 2k - 2(ll) \tag{39}$$

$$BIC = k\left(\log_{10}(n)\right) - (2ll) \tag{40}$$

$$CAIC = \frac{AIC(n-k-1) + 2k(k+1)}{(n-k-1)} \tag{41}$$

$$HQIC = 2k\log_{10}\left(\ln(n)\right) - 2(ll) \tag{42}$$

where, *lL* is the log-likelihood value, *n* is the in-sample size, and *k* is the parameter.

III. Results

I. Stability Study via Monte Carlo Simulations

In this section, we examined the stability of the probability model with an increase in sample size using simulation. The parameters were fixed at (0.6, 0.5) and (0.3, 0.7). Utilizing the quantile function presented in Equation (8), a random sample of sizes—30, 75, 300, 500, and 1000—was generated. The measures used for assessing the models were the Average Absolute Bias (AAB), Mean Square Error (MSE), and Root Mean Square Error (RMSE), respectively. These measures were calculated using the expressions given below:

$$AAB = \sum_{q=1}^{Q} \left|\hat{u}_q - u\right| \Big/ q \tag{43}$$

$$MSE = \sum_{q=1}^{Q} \left(\hat{u}_q - u\right)^2 \Big/ q \tag{44}$$

$$RMSE = \sqrt{MSE} \tag{45}$$

In accordance with the central limit theorem, an increase in sample sizes is expected to result in the reduction of estimation errors, approaching zero. Analyzing the results of the simulation as presented in Table 1, it becomes evident that this holds true for the new model. An increase in sample sizes corresponds to a decrease in bias and mean square error, as demonstrated.

Table 1: The *AAB, MSE and RMSE*

Size (n)	Parameter	Parameter Value	MLE	AAB	MSE	RMSE
30	λ	0.21	-2.929E-06	0.21000293	0.04410123	0.21000293
30	α	0.72	-0.3688692	1.08886921	1.18563616	1.08886921
100	λ	0.21	0.00062679	0.20937321	0.04383714	0.20937321
100	α	0.72	-0.378557	1.09855699	1.20682746	1.09855699
300	λ	0.21	0.08369549	0.12630451	0.01595283	0.12630451
300	α	0.72	0.74825993	0.02825993	0.00079862	0.02825993
500	λ	0.21	0.2257876	0.0157876	0.00024925	0.0157876
500	α	0.72	0.57707861	0.14292139	0.02042652	0.14292139

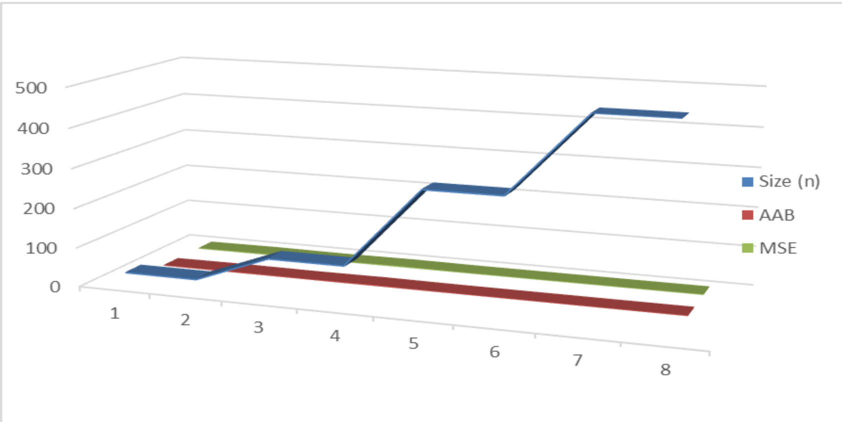

Figure 4: *The AAB and MSE*

As illustrated in the plot of the AAB and MSE presented in Figure 4, it can be observed that the probability model is well-behaved. This was due to the decay in the value of the AAB and MSE. The larger the sizes of the sample, the better the estimates are, the smaller the error and the more consistent the parameters are.

II. Application to Real-Life Data

This section provides an application of the model to real data. Other competing models were fitted to the data, and the goodness of fit of these models was assessed using various statistical information criteria. This illustration involves four data sets. The first and second datasets were sourced from [21]. The first dataset pertains to the duration of symptom decrease or disappearance in patients with bladder cancer, measured in months for one hundred and twenty-eight patients. The second dataset focuses on the response time of patients to treatments, measured in minutes from the moment the treatment was administered.

First data set: 0.08, 2.09, 2.73, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.22, 3.52, 4.98, 6.99, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 15.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.93, 8.65, 12.63, 22.69

Second data set: 1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0

The third and fourth dataset were both sourced from [22]. The third dataset consisted of three hundred and forty-six measures of nicotine taken as obtained from different cigarette product category. The fourth data was on the windshield of an aircraft. The data comprises of one hundred and fifty-three measurements, of which eighty-eight were categorized as failed windshields and the remaining sixty-five were service times of windshields that were in good condition at the time of the observation. 1000h is the measuring unit.

Third data set: 1.3, 1.0, 1.2, 0.9, 1.1, 0.8, 0.5, 1.0, 0.7, 0.5, 1.7, 1.1, 0.8, 0.5, 1.2, 0.8, 1.1, 0.9, 1.2, 0.9, 0.8, 0.6, 0.3, 0.8, 0.6, 0.4, 1.1, 1.1, 0.2, 0.8, 0.5, 1.1, 0.1, 0.8, 1.7, 1.0, 0.8, 1.0, 0.8, 1.0, 0.2, 0.8, 0.4, 1.0, 0.2, 0.8, 1.4, 0.8, 0.5, 1.1, 0.9, 1.3, 0.9, 0.4, 1.4, 0.9, 0.5, 1.7, 0.9, 0.8, 0.8, 1.2, 0.9, 0.8, 0.5, 1.0, 0.6, 0.1, 0.2, 0.5, 0.1, 0.1, 0.9, 0.6, 0.9, 0.6, 1.2, 1.5, 1.1, 1.4, 1.2, 1.7, 1.4, 1.0, 0.7, 0.4, 0.9, 0.7, 0.8, 0.7, 0.4, 0.9, 0.6, 0.4, 1.2, 2.0, 0.7, 0.5, 0.9, 0.5, 0.9, 0.7, 0.9, 0.7, 0.4, 1.0, 0.7, 0.9, 0.7, 0.5, 1.3, 0.9, 0.8, 1.0, 0.7, 0.7, 0.6, 0.8, 1.1, 0.9, 0.9, 0.8, 0.8, 0.7, 0.7, 0.4, 0.5, 0.4, 0.9, 0.9, 0.7, 1.0, 1.0, 0.7, 1.3, 1.0, 1.1, 1.1, 0.9, 1.1, 0.8, 1.0, 0.7, 1.6, 0.8, 0.6, 0.8, 0.6, 1.2, 0.9, 0.6, 0.8, 1.0, 0.5, 0.8, 1.0, 1.1, 0.8, 0.8, 0.5, 1.1, 0.8, 0.9, 1.1, 0.8, 1.2, 1.1, 1.2, 1.1, 1.2, 0.2, 0.5, 0.7, 0.2, 0.5, 0.6, 0.1, 0.4, 0.6, 0.2, 0.5, 1.1, 0.8, 0.6, 1.1, 0.9, 0.6, 0.3, 0.9, 0.8, 0.8, 0.6, 0.4, 1.2, 1.3, 1.0, 0.6, 1.2, 0.9, 1.2, 0.9, 0.5, 0.8, 1.0, 0.7, 0.9, 1.0, 0.1, 0.2, 0.1, 0.1, 1.1, 1.0, 1.1, 0.7, 1.1, 0.7, 1.8, 1.2, 0.9, 1.7, 1.2, 1.3, 1.2, 0.9, 0.7, 0.7, 1.2, 1.0, 0.9, 1.6, 0.8, 0.8, 1.1, 1.1, 0.8, 0.6, 1.0, 0.8, 1.1, 0.8, 0.5, 1.5, 1.1, 0.8, 0.6, 1.1, 0.8, 1.1, 0.8, 1.5, 1.1, 0.8, 0.4, 1.0, 0.8, 1.4, 0.9, 0.9, 1.0, 0.9, 1.3, 0.8, 1.0, 0.5, 1.0, 0.7, 0.5, 1.4, 1.2, 0.9, 1.1, 0.9, 1.1, 1.0, 0.9, 1.2, 0.9, 1.2, 0.9, 0.5, 0.9, 0.7, 0.3, 1.0, 0.6, 1.0, 0.9, 1.0, 1.1, 0.8, 0.5, 1.1, 0.8, 1.2, 0.8, 0.5, 1.5, 1.5, 1.0, 0.8, 1.0, 0.5, 1.7, 0.3, 0.6, 0.6, 0.4, 0.5, 0.5, 0.7, 0.4, 0.5, 0.8, 0.5, 1.3, 0.9, 1.3, 0.9, 0.5, 1.2, 0.9, 1.1, 0.9, 0.5, 0.7, 0.5, 1.1, 1.1, 0.5, 0.8, 0.6, 1.2, 0.8, 0.4, 1.3, 0.8, 0.5, 1.2, 0.7, 0.5, 0.9, 1.3, 0.8, 1.2, 0.9

Fourth data set: 0.040, 1.866, 2.385, 3.443, 0.301, 1.876, 2.481, 3.467, 0.309, 1.899, 2.610, 3.478, 0.557, 1.911, 2.625, 3.578, 0.943, 1.912, 2.632, 3.595, 1.070, 1.914, 2.646, 3.699, 1.124, 1.981, 2.661, 3.779,1.248, 2.010, 2.688, 3.924, 1.281, 2.038, 2.82,3, 4.035, 1.281, 2.085, 2.890, 4.121, 1.303, 2.089, 2.902, 4.167, 1.432, 2.097, 2.934, 4.240, 1.480, 2.135, 2.962, 4.255, 1.505, 2.154, 2.964, 4.278, 1.506, 2.190, 3.000, 4.305, 1.568, 2.194, 3.103, 4.376, 1.615, 2.223, 3.114, 4.449, 1.619, 2.224, 3.117, 4.485, 1.652, 2.229, 3.166, 4.570, 1.652, 2.300, 3.344, 4.602, 1.757, 2.324, 3.376, 4.663.

Five models were fitted to the above four datasets using the *Adequacy Model* package in R [23], these models include; Exponentiated Ailamujia distribution (EAD), Exponentiated Exponential distribution (EED), Exponentiated Weibull (EWD), the Logistic distribution (LD) and the transmuted skew student t distribution **(**TSS*t*D**)** respectively. The resulting fitted models selected on the basis of Akaike Information Criterion, (AIC), Bayesian Information Criterion (BIC), Consistent Akaikes Information Criterion (AIC) and Hannan-Quinn information criterion, (HQIC).

Table 2: The *AIC, CAIC, BIC, HQIC and MLE of the First data*

Models	AIC	BIC	CAIC	HQIC	MLE	Rank
TSS*t*D	-642.6124	-636.9719	-634.9719	-627.3313	0.0006 1.1348	**1**
EED	799.7295	805.3700	799.8286	802.0208	1.2682 0.1041	3
EWD	275.1219	282.4499	275.4182	278.0694	0.4493 1.6494 1.7404	2
EAD	7285.8160	7291.4570	7285.9150	7288.1080	1.4819 1.8290	5
LD	1080.3360	1085.9770	1080.4350	1082.6270	1.9978 1.8904	4

Table 2 to Table 5 presents the model estimates for each of the datasets. The results revealed that the model with the smallest measure of the entire information criterion was the TSS*t*D. The ranks **for** the performance of the models were based on the information criterions of each of the models. From the results obtained, for the five models estimated, the TSS*t*D was the models with the best fit.

Table 3: The *AIC, CAIC, BIC, and HQIC of the second data*

Models	AIC	BIC	CAIC	HQIC	MLE	Rank
TSS*t*D	-642.6124	-640.621	-638.621	-634.6295	9.3568e-03 -7.6898e-16	**1**
EED	36.3450	38.3365	37.0509	36.7338	54.366966 2.172273	2
EWD	46.9561	49.9433	48.4561	47.5392	0.6234 1.7261 1.7733	3
EAD	48.3887	50.3802	49.0946	48.7775	1.8667 0.7558	4
LD	1204.3680	1210.0080	1204.4670	1206.6590	1.7387 1.7264	5

Table 4: The *AIC, CAIC, BIC, and HQIC of the* Third data

Models	AIC	BIC	CAIC	HQIC	MLE	Ranks
TSS*t*D	-642.6124	-634.9196	-632.9196	-623.2267	-0.1381 1.4114	**1**
EED	325.3694	333.0623	325.4044	328.4327	6.0022 2.3477	4
EWD	261.0285	272.5679	261.0987	265.6235	1.3488 1.7973 1.5733	2
EAD	300.5637	308.2566	300.5987	303.6271	1.9331 1.5668	3
LD	293.6612	301.3540	293.6961	296.7245	0.9789 0.2148	5

Table 5: The *AIC, CAIC, BIC, and HQIC of the* fourth data

Models	AIC	BIC	CAIC	HQIC	MLE	Ranks
TSS*t*D	-347.3415	-342.4562	-340.4562	-333.5709	-0.1613 1.1483	**1**
EED	292.3541	297.2394	292.5004	294.3191	3.6582 0.6240	4
EWD	275.2268	282.5548	275.5231	278.1743	0.4305 1.7537 1.3413	2
EAD	283.6236	288.5089	283.7699	285.5886	1.7293 0.5065	3
LD	412.6227	417.5080	412.7690	414.5877	0.8074 1.3006	5

IV. Conclusion

This research paper presented a novel two-parameter distribution known as the Transmuted Skew Student *t* distribution. Some of the statistical and reliability properties for the TSS*t*D were derived and they included the survival function, the r^{th} moment, the hazard function, the mean, the quantile function, the moment generating function, the characteristic function and the order statistics. Before application to real dataset, a Monte-Carlo simulation study was conducted to assess the stability of the model with more sample sizes. The results revealed that the model was consistent with increase in the number of samples. The new PDF was applied to four different real datasets. Using information criterions, it was found that TSS*t*D performs better than other competing models.

References

[1] David, I. J., Mathew, S., & Falgore, J. Y. (2024). Reliability Analysis with New Sine Inverted Exponential Distribution: Properties, Simulation and Application. *European Journal of Statistics, 4*, 1-14.

[2] David, I. J., Mathew, S., & Eghwerido, J. T. (2023). Reliability Analysis with New Sine Inverse Rayleigh Distribution. *Journal of Reliability and Statistical Studies, 16*(2), 255-268.

[3] David, I. J., Asiribo, O. E., Dikko, H. G., & Ikwuoche, P. O. (2023). Johnson-Schumacher Split-Plot Design Modelling of Rice Yield. *Biometrical Letters, 60*(1), 37-52.

[4] David, I. J., Asiribo, O. E., & Dikko, H. G. (2023). Nonlinear Split-Plot Design Modeling and Analysis of Rice Varieties Yield. *Scientific African, 19*(9):e01444.

[5] David, I. J., Asiribo, O. E., & Dikko, H. G. (2022). A Weibull Split-Plot Design Model and Analysis. *Thailand Statistician, 20*(2), 420-434.

[6] David, I. J., Asiribo, O. E., & Dikko, H. G. (2022). A Bertalanffy-Richards Split-Plot Design Model and Analysis. *Journal of Statistical Modeling and Analysis, 4*(1), 56-71.

[7] Gupta, R. D., & Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, *43*(1), 117-130.

[8] Kumar, D., Singh, U., & Singh, S. K. (2015). A new distribution using sine function-its application to bladder cancer patients data. *Journal of Statistics Applications & Probability*, *4*(3), 417.

[9] Mahmood, Z., Chesneau, C., & Tahir, M. H. (2019). A new sine-G family of distributions: properties and applications. *Bull. Comput. Appl. Math.*, *7*(1), 53-81.

[10] Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, *71*(1), 63-79.

[11] Ashour S. K. & Eltehiwy M. A., (2013). Transmuted exponentiated Lomax distribution. Australian. *Journal of Basic and Applied Sciences, 7*(7), 658-667.

[12] Ashour, S. K., & Eltehiwy, M. A. (2013). Transmuted lomax distribution. *American Journal of Applied Mathematics and Statistics*, *1*(6), 121-127.

[13] Mahmoud, M. R., & Mandouh, R. M. (2013). On the transmuted Fréchet distribution. *Journal of Applied*

Sciences Research, *9*(10), 5553-5561.

[14] Hussian, M. A. (2014). Transmuted exponentiated gamma distribution: A generalization of the exponentiated gamma probability distribution. *Applied Mathematical Sciences*, *8*(27), 1297-1310.

[15] Elbatal, I., & Aryal, G. (2013). On the transmuted additiveweibull distribution. *Austrian Journal of statistics*, *42*(2), 117-132.

[16] Khan, M. S., King, R., & Hudson, I. L. (2017). Transmuted Generalized Gompertz distribution with application. *Journal of Statistical Theory and Applications*, *16*(1), 65-80.

[17] Al-Babtain, A., Fattah, A. A., Ahmed, A. H. N., & Merovci, F. (2017). The Kumaraswamy-transmuted exponentiated modified Weibull distribution. *Communications in Statistics-Simulation and Computation*, *46*(5), 3812-3832.

[18] Saraçoğlu, B., & Taniş, C. (2021). A new lifetime distribution: transmuted exponential power distribution. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, *70*(1), 1-14.

[19] Shaw, W. T., & Buckley, I. R. (2009). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint arXiv:0901.0434*.

[20] Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Volume 2, Second Edition. New York: Wiley.

[21] Jones M. C., (2001). A skew t distribution. In Probability and Statistical Models with Applications (eds C. A. Charalambides, M. V. Koutras and N. Balakrishnan), pp. 269–277. London: Chapman and Hall.

[22] Jones M. C. & Faddy M., (2003). A skew extension of the t-distribution, with applications.," *Journal of the Royal Statistical Society Series B: Statistical Methodology,* 65(1)159-174.

[23] Rather, A. A., Subramanian, C., Al-Omari, A. I., & Alanzi, A. R. (2022). Exponentiated Ailamujia distribution with statistical inference and applications of medical data. *Journal of Statistics and Management Systems*, *25*(4), 907-925.