ISSN No:-2456-2165

On the Application of Two Parameter Shanker Distribution

Olukemi B. Olufemi -Ojo[1], Sidney I. Onyeagu[2], Happiness O. Obiora Ilouno[3]
[1]Department of Statistics Federal Polytechnic Oko, Anambra State.
[2,3]Department of statistics Nnamdi Azikiwe University Awka.

Abstract:- In this study, a two parameter Shanker distribution was proposed. Its statistical Properties including the probability density function, moments, kurtosis, skewness, survival function, hazard function, stochastic ordering and more were discussed. The maximum likelihood method was used to estimate the parameters. The proposed distribution was compared with one parameter shanker, two-parameter Lindley, two-parameter Weibull and two-parameter Gamma distributions using log likelyhood function, the Akaike Information Criterion and Bayesian Information Criterion.

Keywords:- *Shanker Distribution, Goodness of Fit, Maximum Likelyhood, Hazard Function, Moment Generating Function.*

I. INTRODUCTION

Many approaches have been devised by researchers in constructing new distributions with the aim of obtaining more flexible distributions suitable for modelling the complex data being churned out from various areas of human endeavours as a result of the advancement in technology in recent years. Some of the standard probability distributions in the distribution theory literature include the Binomial distribution, Poisson distribution, Geometric distribution, Negative Binomial distribution, Hypergeometric distribution among others. Some commonly used standard continuous probability distributions are Gamma distribution, Exponential distribution, Uniform distribution, Normal distribution, Weibull distribution, Rayleigh distribution, Log-normal distribution, Logistic distribution, Frechet distribution, Gumbel distribution, Beta distribution, Pareto distribution, Triangular distribution, Burr distribution, Cauchy distribution, Chi-square distribution, Student-t distribution, Fisher distribution and so on.These standard distributions have been found to be of immense importance and usage in many fields of study both in theory and practice. In building statistical methods such as regression models, correlation models, analysis of variance models, queueing models, time series models, multivariate statistical techniques, quality control charts, project management techniques, reliability models, survival analysis models and so on, it is often assumed that certain distributional assumptions are satisfied by the research data. For instance, to construct the P-chart, it is assumed that number of defects follow the Poisson distribution. In statistical quality control, the construction of X and R chart rests on the assumption that the population from which the samples are drawn is normally distributed. Also, the sampling distribution of the number of defective is the binomial distribution. The t-statistic for testing the null hypothesis that two population means are equal rests on the assumption that the populations of interests are normally distribution. The normal distribution is also assumed for the error term of a linear regression model. The use of one-way analysis of variance model also rests on the assumption that the sampled populations are normally distributed with equal variance. In time series, it also assumed that the error term is Gaussian. The building block to multivariate statistical methods is the multivariate normal distribution. Even non parametric statistical method for goodness-of-fit, homogeneity and independence rely on the assumption of chi-square distribution, though not stringent.

It is important to note that each of the distributions mentioned in the preceding paragraph is characterized by different shapes and tail behaviours which defines their suitability to different data types. Each of these distributions was characterized by different shapes and tail behaviours which defines their suitability to different data types. While some are right-skewed, others are left-skewed, some are symmetric, others are both skewed to the right, left and symmetric. The tendency of any distribution to exhibit different shape behaviours defines the flexibility of that distribution and adding a new parameter to the existing distribution will create room for additional control on the high degree of skewness and kurtosis.Due to the importance of Shanker distribution, Shanker and Shukla (2017) suggested power Shanker distribution. Borah and Hazarika (2017) offered discrete Shanker distribution. Abushal (2019) considered Bayesian estimation of the reliability characteristic of Shanker distribution.

II. TWO PARAMETER SHANKER DISTRIBUTION

A. The Shanker Distribution

Shanker (2015a) introduced the Shanker distribution with parameter θ whose probability density function (PDF) and cumulative distribution function (CDF) are

$$f(x,\theta)=\frac{\theta^2}{\theta^2+1}[\theta+x]e^{-\theta x};x>0,\theta>0 \quad (1)$$

And

$$F(x,\theta)=1-\left[\frac{(1+\theta^2)+\theta x}{\theta^2+1}\right]e^{-\theta x};x>0,\theta>0 \quad (2)$$

respectively.

The properties of the Shanker distribution such as hazard rate function, survival function, and moments have been studied by Shanker (2015a). The parameters of the Shanker distribution was estimated using the method of maximum likelihood.

B. Probability Density Function of The Two-Parameter Shanker Model

- **Theorem 2.1**: A random variable X is said to have a two-parameter Shanker random distribution with parameters η and τ, denoted by $X \sim TPS(\eta,\tau)$, if its probability density function (pdf) is given by

$$f_{TPSD}(x;\,\eta,\tau)=\frac{\eta^2}{\eta^2+\tau}(\eta+\tau x)e^{-\eta x},x>0,\eta>0,\tau>0 \quad (3)$$

- **Proof:** Consider $X \sim Exp(\eta)$, then

$$f_1(x;\eta)=\eta e^{-\eta x},x>0,\eta>0 \quad (4)$$

and $X \sim Gamma(2,\eta)$, then

$$f_2(x;\eta)=\eta^2 e^{-\eta x},x>0,\eta>0 \quad (5)$$

Motivated by the work of Shanker (2015a), we define the mixing proportion as $\omega_1=\eta^2/(\eta^2+\tau)$. By mixture method of deriving a probability distribution, the two-parameter Shanker distribution is obtained from the following relation:

$$f_{TPSD}(x;\,\eta,\tau)=\omega_1 f_1(x;\eta)+(1-\omega_1)\,f_2(x;\eta),\ x\in\mathbb{R} \quad (6)$$

A substitution of (4) and (5) into (6), leads to:

$$f_{TPSD}(x;\,\eta,\tau)=\frac{\eta^2}{\eta^2+\tau}\cdot\eta e^{-\eta x}+\left(1-\frac{\eta^2}{\eta^2+\tau}\right)\eta^2 e^{-\eta x},\ x>0,\eta>0,\tau>0$$

$$f_{TPSD}(x;\,\eta,\tau)=\frac{\eta^2}{\eta^2+\tau}(\eta+\tau x)e^{-\eta x},\ x>0,\eta>0,\tau>0$$

which completes the proof of Theorem 2.1.

ISSN No:-2456-2165

- **Corollary 2.1:** The pdf of the two-parameter Shanker distribution is a proper density function.
- **Proof:** Corollary 2.1 suffices to show that the pdf in (3) satisfy the conditions:

$$f_{TPSD}(x;\eta,\tau)\geq 0 \tag{7}$$

and $$\int_0^\infty f_{TPSD}(x;\eta,\tau)\,dx=1 \tag{8}$$

Verifying, it is observed from (3.1) that $f_{TPSD}(x;\eta,\tau)\geq 0$ for all $x\geq 0$. Also,

$$\int_0^\infty f_{TPSD}(x;\eta,\tau)\,dx=\int_0^\infty \frac{\eta^2}{\eta^2+\tau}(\eta+\tau x)e^{-\eta x}\,dx$$

$$=\frac{\eta^2}{\eta^2+\tau}\int_0^\infty(\eta+\tau x)e^{-\eta x}\,dx=\frac{\eta^3}{\eta^2+\tau}\int_0^\infty e^{-\eta x}\,dx+\frac{\eta^2\tau}{\eta^2+\tau}\int_0^\infty x e^{-\eta x}\,dx$$

$$=\frac{\eta^3}{\eta^2+\tau}\frac{\Gamma(1)}{\eta}+\frac{\eta^2\tau}{\eta^2+\tau}\frac{\Gamma(2)}{\eta^2}=\frac{\eta^2}{\eta^2+\tau}+\frac{\tau}{\eta^2+\tau}=1$$

which completes the proof of Theorem 2.1.

C. Cumulative Density Function of The Two-Parameter Shanker Model

- **Theorem 2.2:** A random variable X is said to have a two-parameter Shanker random distribution with parameters η and τ, denoted by $X\sim TPS(\eta,\tau)$, if its cumulative distribution function (cdf) is given by

$$F(x;\eta,\tau)=1-\left[1+\frac{\eta\tau x}{\eta^2+\tau}\right]e^{-\eta x},\ x>0,\eta>0,\tau>0 \tag{9}$$

- **Proof:** Let $X\sim Exp(\eta)$, then

$$F_1(x;\eta)=\int_0^x f_1(u;\eta)\,du=\int_0^x \eta e^{-\eta u}du$$

$$=\left[\frac{\eta e^{-\eta u}}{-\eta}\right]_{u=0}^{u=x}=1-e^{-\eta x},\ x>0,\eta>0 \tag{10}$$

and $X\sim Gamma(2,\eta)$, then

$$F_2(x;\eta)=\int_0^x f_2(u;\eta)\,du=\int_0^x \eta^2 u e^{-\eta u}du=\eta^2\left[\left.\frac{ue^{-\eta u}}{-\eta}\right|_{u=0}^{u=x}+\frac{1}{\eta}\int_0^x e^{-\eta u}du\right]$$

$$=\eta^2\left[-\frac{xe^{-\eta x}}{\eta}+\frac{1}{\eta}\left[\frac{e^{-\eta u}}{-\eta}\right]_{u=0}^{u=x}\right]=\eta^2\left[-\frac{xe^{-\eta x}}{\eta}+\frac{1}{\eta^2}\left(1-e^{-\eta x}\right)\right]$$

$$=1-(1-\eta x)e^{-\eta x},x>0,\eta>0 \tag{11}$$

In line with the work of Shanker (2015a), we define the mixing proportion as $\omega_1 = \eta^2 \big/ \left(\eta^2 + \tau\right)$. Thus, by mixture method of deriving a cumulative distribution function(cdf), we obtain the cdf of the two-parameter Shanker distribution from the relation

$$F_{TPSD}\left(x;\,\eta,\tau\right) = \omega_1 F_1\left(x;\eta\right) + \left(1-\omega_1\right) F_2\left(x;\eta\right),\ x \in \mathbb{R} \tag{12}$$

Substituting (10) and (11) into (12), gives

$$F_{TPSD}\left(x;\,\eta,\tau\right) = \frac{\eta^2}{\eta^2+\tau}\cdot\left(1-e^{-\eta x}\right) + \left(1-\frac{\eta^2}{\eta^2+\tau}\right)\left[1-\left(1-\eta x\right)e^{-\eta x}\right],\ x>0,\eta>0,\tau>0$$

$$= \frac{\eta^2}{\eta^2+\tau}\left(1-e^{-\eta x}\right) + \frac{\tau}{\eta^2+\tau}\left[1-\left(1-\eta x\right)e^{-\eta x}\right],\ x>0,\eta>0,\tau>0$$

$$= \frac{\eta^2}{\eta^2+\tau} - \frac{\eta^2}{\eta^2+\tau}e^{-\eta x} + \frac{\tau}{\eta^2+\tau} - \frac{\tau}{\eta^2+\tau}e^{-\eta x} - \frac{\eta\tau}{\eta^2+\tau}xe^{-\eta x},\ x>0,\eta>0,\tau>0$$

$$= 1 - \left[\frac{\eta^2}{\eta^2+\tau} + \frac{\tau}{\eta^2+\tau} + \frac{\eta\tau x}{\eta^2+\tau}\right]e^{-\eta x},\ x>0,\eta>0,\tau>0$$

$$= 1 - \left[1 + \frac{\eta\tau x}{\eta^2+\tau}\right]e^{-\eta x},\ x>0,\eta>0,\tau>0$$

which completes the proof of Theorem 2.2.

- **Corollary 2.2:** The cdf in (9) satisfies the conditions $F_{TPSD}\left(-\infty\right)=0$ and $F_{TPSD}\left(\infty\right)=1$.
- **Proof:** The proof of Corollary 2.2 follows from taking the limit of the cdf in (9) as $x \to -\infty$ and $x \to +\infty$. Thus,

$$F\left(-\infty\right) = \lim_{x\to-\infty} F\left(x;\eta,\tau\right) = \lim_{x\to-\infty}\left(1-\left[1+\frac{\eta\tau x}{\eta^2+\tau}\right]e^{-\eta x}\right) = 0 \tag{13}$$

and $$F\left(+\infty\right) = \lim_{x\to+\infty} F\left(x;\eta,\tau\right) = \lim_{x\to+\infty}\left(1-\left[1+\frac{\eta\tau x}{\eta^2+\tau}\right]e^{-\eta x}\right) = 1 \tag{14}$$

D. Moment Generating Function of The Two-Parameter Shanker Model

The moment generating function of $X \sim TPS\left(\eta,\tau\right)$ is given by

$$M_X\left(t\right) = E\left(e^{tX}\right) = \int_0^\infty e^{tx} f_{TPSD}\left(x;\,\eta,\tau\right)dx = \int_0^\infty e^{tx}\frac{\eta^2}{\eta^2+\tau}\left(\eta+\tau x\right)e^{-\eta x}dx$$

$$= \frac{\eta^2}{\eta^2+\tau}\left[\eta\int_0^\infty e^{-(\eta-t)x}dx + \tau\int_0^\infty x\,e^{-(\eta-t)x}dx\right] = \frac{\eta^2}{\eta^2+\tau}\left[\frac{\eta}{\eta-t} + \frac{\tau}{\left(\eta-t\right)^2}\right]$$

$$= \frac{\eta^2}{\eta^2+\tau}\left[\frac{\eta\left(\eta-t\right)+\tau}{\left(\eta-t\right)^2}\right] = \left(\frac{\eta^2+\tau-\eta t}{\eta^2+\tau}\right)\frac{\eta^2}{\left(\eta-t\right)^2}$$

$$=\left(1-\frac{\eta t}{\eta^2+\tau}\right)\left(1-\frac{t}{\eta}\right)^{-2},\left|\frac{it}{\eta}\right|\le 1 \quad (15)$$

E. Characteristic Function of The Two-Parameter Shanker Model

The characteristic function of $X \sim TPS(\eta,\tau)$ is given by

$$\phi_X(t)=E\left(e^{itX}\right)=\int_0^\infty e^{itx} f_{TPSD}(x;\eta,\tau)\,dx=\int_0^\infty e^{itx}\frac{\eta^2}{\eta^2+\tau}(\eta+\tau x)e^{-\eta x}dx$$

$$=\frac{\eta^2}{\eta^2+\tau}\left[\eta\int_0^\infty e^{-(\eta-it)x}dx+\tau\int_0^\infty x\,e^{-(\eta-it)x}dx\right]=\frac{\eta^2}{\eta^2+\tau}\left[\frac{\eta}{\eta-it}+\frac{\tau}{(\eta-it)^2}\right]$$

$$=\frac{\eta^2}{\eta^2+\tau}\left[\frac{\eta(\eta-it)+\tau}{(\eta-it)^2}\right]=\left(\frac{\eta^2+\tau-\eta it}{\eta^2+\tau}\right)\frac{\eta^2}{(\eta-it)^2}$$

$$=\left(1-\frac{\eta it}{\eta^2+\tau}\right)\left(1-\frac{it}{\eta}\right)^{-2},\ i=\sqrt{-1}\ \text{is the complex unit} \quad (16)$$

F. Survival and Hazard Functions of The Two-Parameter Shanker Model

The survival function of $X \sim TPS(\eta,\tau)$ is given by

$$S(x)=1-F_{TPSD}(x;\eta,\tau)=1-\left(1-\left[\frac{\eta^2+\tau+\eta\tau x}{\eta^2+\tau}\right]e^{-\eta x}\right)=\left(1+\frac{\eta\tau x}{\eta^2+\tau}\right)e^{-\eta x} \quad (17)$$

The hazard rate function of $X \sim TPS(\eta,\tau)$ is given by

$$h(x)=\frac{f_{TPSD}(x;\eta,\tau)}{1-F_{TPSD}(x;\eta,\tau)}=\frac{\frac{\eta^2}{\eta^2+\tau}(\eta+\tau x)e^{-\eta x}}{1-\frac{\eta^2+\tau+\eta\tau x}{\eta^2+\tau}e^{-\eta x}}=\frac{\eta^2(\eta+\tau x)}{\eta^2+\tau+\eta\tau x} \quad (18)$$

Notably,

$$h(0)=\frac{\eta^2(\eta+\tau(0))}{\eta^2+\tau+\eta\tau(0)}=\frac{\eta^3}{\eta^2+\tau}=f_{TPSD}(0) \quad (19)$$

III. METHODOLOGY

A. Maximum Likelihood Estimates of Parameters of Two- Parameter Shanker Model

The likelihood function of a random sample $x_1,x_2,\dots,x_n$ drawn from a two-parameter Shanker distribution with parameters η and τ is given by

$$L(\eta,\tau)=\prod_{i=1}^{m} f_{TPSD}(x_i;\eta,\tau) \quad (20)$$

$$=\prod_{i=1}^{m}\frac{\eta^2}{\eta^2+\tau}(\eta+\tau x_i)e^{-\eta x_i}$$

using (3)

$$=\left(\frac{\eta^2}{\eta^2+\tau}\right)^m\left[\prod_{i=1}^{m}(\eta+\tau x_i)\right]e^{-\eta\sum_{i=1}^{m}x_i} \tag{21}$$

The natural logarithm of $L(\eta,\tau)$, is given by

$$L(\eta,\tau)=\ln\left(\frac{\eta^2}{\eta^2+\tau}\right)^m+\ln\left[\prod_{i=1}^{m}(\eta+\tau x_i)\right]+\ln e^{-\eta\sum_{i=1}^{m}x_i}$$

$$=m\left[2\ln(\eta)-\ln\left(\eta^2+\tau\right)\right]+\sum_{i=1}^{m}\ln(\eta+\tau x_i)-\eta n\bar{x} \tag{22}$$

Taking the partial derivative of $\ln L(\eta,\tau)$ with respect to η and τ, and equating the results to zero, gives

$$\frac{\partial\ln L(\eta,\tau)}{\partial\eta}=f(\eta,\tau)=m\left[\frac{2}{\eta}-\frac{2\eta}{\eta^2+\tau}\right]+\sum_{i=1}^{m}\frac{1}{\eta+\tau x_i}-n\bar{x}$$

$$=\frac{2m\tau}{\eta\left(\eta^2+\tau\right)}+\sum_{i=1}^{m}\frac{1}{\eta+\tau x_i}-n\bar{x}=0 \tag{23}$$

And

$$\frac{\partial\ln L(\eta,\tau)}{\partial\tau}=g(\eta,\tau)=-\frac{m}{\eta^2+\tau}+\sum_{i=1}^{m}\frac{x_i}{\eta+\tau x_i}=0 \tag{24}$$

To solve the two systems of non-linear equations (3.79) and (3.80) for η and τ, the Newton-Raphson method is used. This is given by

$$\boldsymbol{\Theta}_{k+1}=\boldsymbol{\Theta}_k-\boldsymbol{H}_k^{-1}G(\boldsymbol{\Theta}_k),\quad k=0,1,2,\ldots \tag{25}$$

where the vector two equations being solved, $G(\boldsymbol{\Theta})$ and the Hessian matrix, $\boldsymbol{H}$ are given by

$$G(\boldsymbol{\Theta})=\left[f(\eta,\tau),g(\eta,\tau)\right] \tag{26}$$

$$\boldsymbol{H}=-E\begin{pmatrix}\dfrac{\partial^2\ln L(\eta,\tau)}{\partial\eta^2} & \dfrac{\partial^2\ln L(\eta,\tau)}{\partial\eta\partial\tau}\\ \dfrac{\partial^2\ln L(\eta,\tau)}{\partial\tau\partial\eta} & \dfrac{\partial^2\ln L(\eta,\tau)}{\partial\tau^2}\end{pmatrix} \tag{27}$$

The elements of the $\boldsymbol{H}$ are the second-order partial derivative $\ln L(\eta,\tau)$ with respect to η and τ, are respectively given by

ISSN No:-2456-2165

$$\frac{\partial^2 \ln L(\eta,\tau)}{\partial \eta^2} = \frac{\partial}{\partial \eta}\left[\frac{2m\tau}{\eta(\eta^2+\tau)} + \sum_{i=1}^{m}\frac{1}{\eta+\tau x_i} - n\bar{x}\right] = -\frac{2m\tau(3\eta^2+\tau)}{\eta(\eta^2+\tau)^2} - \sum_{i=1}^{m}\frac{1}{(\eta+\tau x_i)^2} \tag{28}$$

$$\frac{\partial^2 \ln L(\eta,\tau)}{\partial \tau^2} = \frac{\partial}{\partial \tau}\left[-\frac{m}{\eta^2+\tau} + \sum_{i=1}^{m}\frac{x_i}{\eta+\tau x_i}\right] = \frac{m}{(\eta^2+\tau)^2} - \sum_{i=1}^{m}\frac{x_i^2}{(\eta+\tau x_i)^2} \tag{29}$$

$$\frac{\partial^2 \ln L(\eta,\tau)}{\partial \eta\,\partial \tau} = \frac{\partial}{\partial \eta}\left[-\frac{m}{\eta^2+\tau} + \sum_{i=1}^{m}\frac{x_i}{\eta+\tau x_i}\right] = \frac{2m\eta}{(\eta^2+\tau)^2} - \sum_{i=1}^{m}\frac{x_i}{(\eta+\tau x_i)^2} \tag{30}$$

The determinant and inverse of $\boldsymbol{H}$ are respectively given by

$$|\boldsymbol{H}| = E\left[\frac{\partial^2 \ln L(\eta,\tau)}{\partial \eta^2}\right]\cdot E\left[\frac{\partial^2 \ln L(\eta,\tau)}{\partial \tau^2}\right] - E\left[\frac{\partial^2 \ln L(\eta,\tau)}{\partial \eta\partial \tau}\right] E\left[\frac{\partial^2 \ln L(\eta,\tau)}{\partial \tau\partial \eta}\right] \tag{31}$$

$$\boldsymbol{H}^{-1} = \frac{1}{|\boldsymbol{H}|} Adj\boldsymbol{H} = \begin{pmatrix} \frac{\partial^2 \ln L(\eta,\tau)}{\partial \tau^2} & -\frac{\partial^2 \ln L(\eta,\tau)}{\partial \eta\partial \tau} \\ -\frac{\partial^2 \ln L(\eta,\tau)}{\partial \tau\partial \eta} & \frac{\partial^2 \ln L(\eta,\tau)}{\partial \eta^2} \end{pmatrix} \tag{32}$$

For ease of computation and accuracy, the R package was used to implement the Newton-Raphson method.

B. *Cumulant Generating Function and Cumulants of The Two-Parameter Shanker Model*

The cumulant generating function of $X \sim TPS(\eta,\tau)$ is given by:

$$K_X(t) = \log \phi_X(t) = \log\left(1 - \frac{\eta i t}{\eta^2+\tau}\right) - 2\log\left(1 - \frac{it}{\eta}\right)$$

Using the expansion $\log(1-z) = -\sum_{r=0}^{\infty}\frac{z^r}{r}$, the above expression becomes

$$K_X(t) = -\sum_{r=0}^{\infty}\frac{\left(\eta i t/(\eta^2+\tau)\right)^r}{r} + 2\sum_{r=0}^{\infty}\frac{(it/\eta)^r}{r}$$

$$= 2\sum_{r=0}^{\infty}\frac{1}{\eta^r}\frac{(it)^r}{r} - \sum_{r=0}^{\infty}\left(\frac{\eta}{\eta^2+\tau}\right)^r\frac{(it)^r}{r}$$

$$= 2\sum_{r=0}^{\infty}\frac{(r-1)!}{\eta^r}\frac{(it)^r}{r!} - \sum_{r=0}^{\infty}(r-1)!\frac{\eta^r}{(\eta^2+\tau)^r}\frac{(it)^r}{r} \tag{33}$$

Thus, the rth cumulants of $X \sim TPS(\eta,\tau)$ is given by

$$\kappa_r = \text{coefficient of } \frac{(it)^r}{r!} \text{ in } K_X(t)$$

ISSN No:-2456-2165

$$=(r-1)!\frac{2}{\eta^r}-(r-1)!\frac{\eta^r}{\left(\eta^2+\tau\right)^r}$$

$$=(r-1)!\left[\frac{2}{\eta^r}-\frac{\eta^r}{\left(\eta^2+\tau\right)^r}\right], r=1,2,\ldots \quad (34)$$

which gives the first four cumulants of $X \sim TPS(\eta,\tau)$ as

$$\kappa_1=(1-1)!\left[\frac{2}{\eta^1}-\frac{\eta^1}{\left(\eta^2+\tau\right)^1}\right]=\frac{\eta^2+2\tau}{\eta\left(\eta^2+\tau\right)}=\mu_1' \quad (35)$$

$$\kappa_2=(2-1)!\left[\frac{2}{\eta^2}-\frac{\eta^2}{\left(\eta^2+\tau\right)^2}\right]=\frac{\eta^4+4\eta^4\tau+2\tau^2}{\eta^2\left(\eta^2+\tau\right)^2}=\mu_2 \quad (36)$$

$$\kappa_3=(3-1)!\left[\frac{2}{\eta^3}-\frac{\eta^3}{\left(\eta^2+\tau\right)^3}\right]==\frac{2\left(\eta^6+6\tau\eta^4+6\tau^2\eta^2+2\tau^3\right)}{\eta^3\left(\eta^2+\tau\right)^3}=\mu_3 \quad (37)$$

$$\kappa_4=\kappa_4+3\kappa_2^2=\frac{3\left(\eta^8+24\tau\eta^6+44\tau^2\eta^4+32\tau^3\eta^2+8\tau^4\right)}{\eta^4\left(\eta^2+\tau\right)^4}=\mu_4 \quad (38)$$

C. Method of Conducting Monte-Carlo Simulation for The Two-Parameter Shanker Model

To assess the performance of the maximum likelihood estimation method of finding model parameters, a simulation study is required. To simulate random samples from the two-parameter Shanker distribution, a sequence of random numbers is first generated from the uniform distribution on the interval (0,1) for specific sample size. A substitution of the uniform random numbers in the quantile function gives the two-parameter Shanker variates. These variates are, therefore, used to compute the maximum likelihood estimates, average bias of the estimates and average of the mean squared errors of parameter estimates, respectively. The steps to be adopted for simulation in this study are:

- **Step 1:** Generate a pseudo-random value drawn from the $TPS(\eta, \tau)$ using the quantile function in Equation (34).
- **Step 2:** Set initial values for the parameters η and τ.
- **Step 3:** Using the values in step 1, calculate $\hat{\mathbf{\Theta}}=(\hat{\eta}, \hat{\tau})$ via MLE method.
- **Step 4:** Repeat steps 1 to 3 N times.
- **Step 5:** Using $\hat{\mathbf{\Theta}}=(\hat{\eta}, \hat{\tau})$ and $\mathbf{\Theta}=(\eta, \tau)$, compute the mean, bias, average bias and average mean squares errors of the N maximum likelihood estimates of each parameter η and τ. The mean estimate of the maximum likelihood estimator $\hat{\mathbf{\Theta}}$ of the parameter $\mathbf{\Theta}$ is given by

$$\bar{\hat{\mathbf{\Theta}}}=\frac{1}{N}\sum_{j=1}^{N}\hat{\Theta}_i \quad (39)$$

The bias of the MLEs $\hat{\mathbf{\Theta}}$ is given by

$$bias\left(\hat{\mathbf{\Theta}}\right)=\hat{\Theta}_i-\Theta \quad (40)$$

The average bias of the MLE of $\hat{\mathbf{\Theta}}$ is given by

ISSN No:-2456-2165

$$Ave.Bias\left(\hat{\boldsymbol{\Theta}}\right)=\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\Theta}_i-\Theta\right) \tag{41}$$

The average bias of the MLE of $\hat{\boldsymbol{\Theta}}$ is given by

$$Ave.MSE\left(\hat{\boldsymbol{\Theta}}\right)=\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\Theta}_i-\Theta\right)^2 \tag{42}$$

- **Step 6:** Repeat steps 1-4 with different samples

D. DATA PRESENTATION AND METHOD OF COMPARISION OF DISTRIBUTIONS

Table 1: Life Time Data on minutes spent at ATM by customers.

13.7,24.4,4.9,3.5,66.9,59.9,0.1,0.1,1.1,1,0.1,0,0.5,0.4,9,6.7,3,2.3,0.1,0.1,0.1,0.1

To demonstrate the flexibility and applicability of the two-parameter Shanker (TPS) distribution with the data in Table 1, the study fitted the TPS distribution using the method of maximum likelihood and compared the results of the TPS distribution with four competing distributions namely,

(1) One-Parameter Shanker (OPS) distribution (Shanker 2015a):

$$f\left(x;\theta\right)=\frac{\theta^2}{\theta^2+1}\left(\theta+x\right)e^{-\theta x},x>0,\theta>0$$

and $F\left(x;\theta\right)=1-\left(1+\frac{\theta x}{\theta^2+1}\right)e^{-\theta x},x>0,\theta>0$

(2) Two-Parameter Lindley (TPL) distribution

$$f\left(x;\theta,\alpha\right)=\frac{\theta^2}{\theta+\alpha}\left(1+\alpha x\right)e^{-\theta x},x>0,\theta>0$$

and $F\left(x;\theta,\alpha\right)=1-\left(1+\frac{\alpha\theta x}{\theta^2+\alpha}\right)e^{-\theta x},x>0,\theta>0$

(3) Two-Parameter Weibull (TPW) distribution:

$$f\left(x;\alpha,\beta\right)=\frac{\beta}{\alpha^\beta}x^{\beta-1}e^{-\left(\frac{x}{\alpha}\right)^\beta},x>0,\alpha>0,\beta>0$$

and $F\left(x;\alpha,\beta\right)=1-e^{-\left(\frac{x}{\alpha}\right)^\beta},x>0,\alpha>0,\beta>0$

(4) Two-parameter Gamma distribution (TPGD):

$$f\left(x;\alpha,\beta\right)=\frac{\beta^\alpha}{\Gamma\left(\alpha\right)}x^{\alpha-1}e^{-\beta x},x>0,\alpha>0,\beta>0$$

and $F\left(x;\alpha,\beta\right)=\frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\gamma\left(\alpha,\beta x\right),x>0,\alpha>0,\beta>0$

where $\gamma\left(\alpha,\beta x\right)$ is the lower incomplete gamma function.

This study used goodness-of-fit test based on the Kolmogorov-Smirnov (KS) test due to Kolmogorov (1933), Smirnoff (1939)and Wolfowitz (1949) with its corresponding p-value to verify that the data in Table 1 actually follow the TPS distribution. the KS statistic is given by

$$KS = \max\left[\frac{i}{n} - \hat{F}\left(x_{(i)}\right), \hat{F}\left(x_{(i)}\right) - \frac{i-1}{n}\right] \quad (43)$$

where $\hat{F}\left(x_{(i)}\right)$ is the estimated distribution function under the ordered data. Since there is more than one distribution to be compared, the distribution with the largest KS p-value will be more appropriate to fit the data in Table 1.Similarly, an alternative goodness-of-fit test to determine whether the data in Table 1 comes from the TPS distribution is the Anderson-Darling (AD) test (Anderson and Darling, 1952) given by

$$AD\left(A^*\right) = -n - \sum_{i=1}^{n} \frac{2i-1}{2n}\left[\ln \hat{F}\left(Y_i\right) + \ln\left(1 - \hat{F}\left(Y_{n+1-i}\right)\right)\right] \quad (44)$$

Alternative to the Anderson-Darling test is the Cramer-von Mises test (Anderson, 1962) given by

$$CVM\left(W^*\right) = \frac{1}{2n} + \sum_{i=1}^{n}\left[\frac{2i-1}{2n} - \hat{F}\left(x_i\right)\right] \quad (45)$$

Once it has been established that the data in Table 1 comes from the TPS distribution, the next task is to determine the determine the best model from amongst the OPS, TPS, TPL and TPW distributions using three discrimination criteria, based on the log-likelihood function evaluated at the maximum likelihood estimates, the Akaike Information Criterion, AIC (Akaike, 1974) and Bayesian Information Criterion, BIC (Schwarz, 1978). To compute the AIC and BIC, the following formulae are used:

$$AIC = -2\ell + 2k \quad (46)$$

$$BIC = k\ln(n) - 2\ell \quad (47)$$

where ℓ is the log-likelihood function evaluated at the maximum likelihood estimates of the TPS distribution, k is the number of parameters in the TPS distribution and n is the sample size of the data in Table 1. In general, for the data in Table 1, any one of the OPS, TPS, TPL and TPW distributions is considered to be best among them if it has the smallest AIC value, the smallest BIC value and the smallest log-likelihood value.

IV. RESULTS AND DISCUSSIONS

The plots of the pdf and hazard functions of the two-parameter Shanker distribution are presented in Figures 1 and 2, respectively. The results obtained using the method of moments and method of maximum likelihood are presented in tables 1 and 2.

Figure 1 shows the plots of the pdf of the TPSD based on several sets of values of the parameters of the distribution. As can be seen in Figure 1, the TPSD has unimodal shape. It is also good for datasets that are positively skewed and heavy tailed. The graph depicted as Figure 2 show that the hazard rate function of TPSD is monotonically increasing.

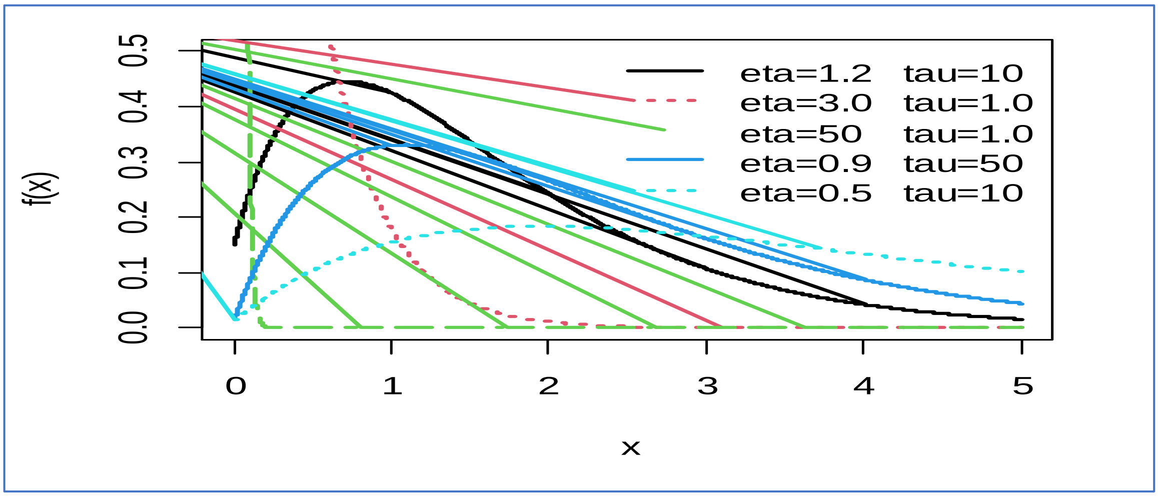

Fig. 1: Plot of pdf of the two-parameter Shanker Distribution

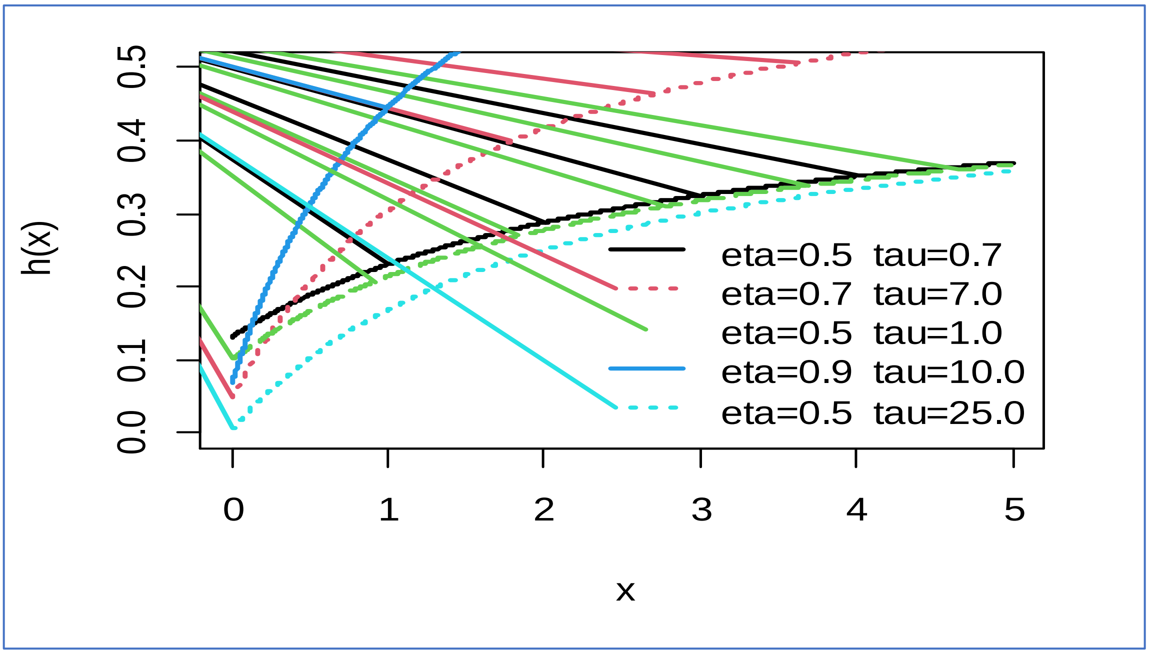

Fig. 2: Plot of hazard function of the two-parameter Shanker Distribution .

Table 2: Method of Maximum Likelihood Results for the two-parameter Shanker Distribution

Model	Estimates	SE	ℓ	AIC	BIC	K-S	P-value
TPSD							
$\hat{\eta}$	0.0390	0.0033	-578.7876	1161.575	1167.167	0.0588	0.7973
$\hat{\tau}$	0.0064	0.0048					
SD							
$\hat{\theta}$	0.0432	0.0028	-582.8922	1167.784	1170.580	0.0824	0.3844
TPLD							
$\hat{\alpha}$	0.1648	0.1289	-578.7876	1161.575	1167.167	0.9610	0.0001
$\hat{\theta}$	0.0390	0.0035					
TPWD							
$\hat{\lambda}$	50.1508	3.6728	-579.0237	1162.047	1167.639	0.0597	0.7815
$\hat{k}$	1.3057	0.0935					
TPGD							
$\hat{\alpha}$	1.4967	0.1749	-579.8047	1163.609	1169.201	0.0762	0.4838
$\hat{\beta}$	0.0323	0.0045					

The maximum likelihood estimates of the parameters of the two-parameter Shanker distribution satisfies the support of the distribution. The value of the log-likelihood function evaluated at these estimates, AIC and AIB values are found to be smaller than those of Shanker distribution, two-parameter Lindley distribution, Weibull distribution and gamma distribution respectively.

The value of the K.S statistics is found to be smaller than those of the competing models with the highest P-values, indicating that the two-parameter Shanker distribution is better than the competing models for modelling the data used in this study.

V. CONCLUSION

Explicit mathematical expressions for some of the basic statistical properties of a new two parameter shanker distribution such as the probability density function, cumulative distribution function, skewness, kurtosis, the survival hazard rate and hazard functions were discussed. Estimation of the model parameters was approached through the method of maximum likelihood estimation and method of moments. A Monte-Carlo simulation was performed to verify the stability of the maximum likelihood estimates of the model parameters. The flexibility and applicability of the new lifetime distribution were illustrated with real data set.Two parameter shanker was recommended for modelling unimodal lifetime data with a non decreasing and heavy-tailed shaped hazard rate function and hope that it would receive significant applications in the future.

REFERENCES

[1]. Abushal, T.A. (2019). Bayesian estimation of the reliability characteristic of Shanker distribution. Journal of the Egyptian mathematical society 27 article number.

[2]. Anderson, T.W. (1962). On the Distribution of the Two-Sample Cramer-von Mises Criterion. *Annals of Mathematical Statistics. Institute of Mathematical Statistics*, 33(3): 1148-1159.

[3]. Anderson, T.W. and Darling, D.A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Mathematical Statistics*, 23(2):193-212.

[4]. Akaike, H. (1974) A new look at the statistical model Identification. IEEE Transactions on Automatic Control AC – 19, 716 – 723.

[5]. Borah, M. and Hazakira, J. (2017). Derived distribution and Biostatistics International Journal, 5(4):146-153.

[6]. Kolmogorov, A.N (1933): foundations of probability theory. Julius Sringer Berlin. Pg 3 – 12. Statistical inference in the non – parametric case. Annals of mathematical statistics, 14, 305 – 332.

[7]. Schwarz, G. (1978) Estimating the Dimension of a Model. Annals of Statistics. 6, 461 – 464.

[8]. Series 1-14.

[9]. Shanker, R. (2015a). Shanker distribution and its application. *International journal of statistics and application*,5(6):338-348.

[10]. Shanker, R. and Shukla, K.K. (2017b). Ishita distribution and its applications. *International Journal of Biometrics and Biostatistics*, 5(2):00126.

[11]. Smirnov, N.V (1939) on the estimation of the discrepancy between empirical curves of distribution for two independent sample Bull. Math. Univ. Moscou, 2, (2).

[12]. Wolfowitz, J. (1949) Non – parametric statistical Inference in proceeding of the Berkeley Symposium on Mathematical statistics and Probability, University of California press, Berkeley, 93 – 113.

APPENDIX

R CODE

```
#Function for Extended Shanker Distribution
#a=eta,b=tau
dESD<-function(x,a,b) a^2/(a^2+b)*(a+b*x)*exp(-a*x)
pESD<-function(q,a,b)  1-(1+(a*b*q)/(a^2+b))*exp(-a*q)
SESD<-function(x,a,b)  1-pPD(x,a,b)
hESD<-function(x,a,b)  dPD(x,a,b)/(1-pPD(x,a,b))
#===================================================================
#Function for Shanker Distribution
#theta=a
dSD<-function(x,a)  (a^2)/(a^2+1)*(a+x)*exp(-a*x)
pSD<-function(q,a) 1-((a^2+1)+a*q)/(a^2+1)*exp(-a*q)
#====================================================================
#Function for two parameter Lindley distribution
#a=alpha;b=theta
dTPLD<-function(x,a,b) b^2/(b+a)*(1+a*x)*exp(-b*x)
pTPLD<-function(q,a,b) 1-(b*a+a*b*q)/(b+a)*exp(-a*q)
#===================================================================
#The Function for Weibull distribution
#a=alpha,k=b
dWD<-function(x,a,b) (b/a)*(x/a)^(b-1)*exp(-(x/a)^b)
pWD<-function(q,a,b)  1-exp(-(q/a)^b)
#===================================================================
#The Function for Rayleigh Distribution
#a=sigma
dRD<-function(x,a)  (x/a^2)*exp(-x^2/(2*a^2))
pRD<-function(q,a)  1-exp(-q^2/(2*a^2))
#===================================================================
#The Function for Gamma Distribution
#a=alpha, b=beta
dGD<-function(x,a,b)  (b^a/gamma(a))*x^(a-1)*exp(-b*x)
pGD<-function(q,a,b)  pgamma(q,a,b)
#====================================================================
#The Function for Exponential Distribution
#a=lambda,
dED<-function(x,a)  a*exp(-a*x)
pED<-function(q,a)  1-exp(-a*q)
#====================================================================
#datasets
x1<-c(13.7,24.4,4.9,3.5,66.9,59.9,0.1,0.1,1.1,1,0.1,0,0.5,0.4,9,6.7,3,2.3,0.1,0.1,0.1,0.1)
#===================================================================
w1<-fitdist(x1,"ESD",start=list(a=0.05,b=0.5))
summary(w1)
gofstat(w1)
ks.test(x1,"pESD",a=w1$estimate[1],b=w1$estimate[2])
#====================================================================
w2<-fitdist(x1,"SD", start=list(a=0.05))
summary(w2)
gofstat(w2)
ks.test(x1,"pSD",a=w2$estimate[1])
#====================================================================
w3<-fitdist(x1,"TPLD", start=list(a=0.5,b=0.05))
summary(w3)
gofstat(w3)
ks.test(x1,"pTPLD",a=w3$estimate[1],b=w3$estimate[2])
#====================================================================
w4<-fitdist(x1,"WD", start=list(a=0.5,b=0.5))
summary(w4)
gofstat(w4)
```

ISSN No:-2456-2165

```
ks.test(x1,"pWD",a=w4$estimate[1],b=w4$estimate[2])
#====================================================================
w5<-fitdist(x1,"RD", start=list(a=15))
summary(w5)
gofstat(w5)
ks.test(x1,"pRD",a=w4$estimate[1])
#====================================================================
w6<-fitdist(x1,"GD", start=list(a=5.0,b=0.5))
summary(w6)
gofstat(w6)
ks.test(x1,"pGD",a=w6$estimate[1],b=w6$estimate[2])
```