Eur. J. Stat. 4 (2024) 5
doi: 10.28924/ada/stat.4.5

New Sine Inverted Exponential Distribution: Properties, Simulation and Application

I. J. David[1], S. Mathew[2*], J. Y. Falgore[3]

[1]*Department of Mathematics and Statistics, Federal University Wukari, 200, Katsina-Ala Road, Wukari, Wukari, Nigeria*

davidij@fuwukari.edu.ng

[2]*Department of Mathematics, Statistics and Computer Science, Kwararafa University Wukari, Wukari Nigeria*

matsteve231@gmail.com

[3]*Department of Statistics, Ahmadu Bello University, Zaria, Nigeria*

jamiluyf@gmail.com

**Correspondence: matsteve231@gmail.com*

ABSTRACT: The New Sine Inverted Exponential Distribution, a new distribution model with just one parameter, is suggested in this study. The suggested model has several statistical qualities and reliability properties that have been constructed and explored. The MLE estimations of the parameters were determined using R's adequacy model package. To calculate the bias of the model parameter and the root mean square error, a simulation study was done. The simulation study revealed that the proposed model is well-behaved. The findings also showed that the suggested model outperforms the current listed models on two real datasets when performance was compared.

INTRODUCTION

Probability models stand as indispensable tools in the realm of statistical research, enabling the representation of random processes. Recent decades have witnessed significant advancements in the development of flexible models, discussed and presented extensively in the literature. This progression has significantly empowered researchers to delve deeper into the analysis of real-world phenomena. The growing diversity of data generation processes across various fields, including medicine, genetics, agronomy, hydrology, engineering, economics, and more, has spurred statisticians to explore more adaptable probability distributions [1]. The aim is to enhance the modeling of diverse data types for practical applications and to bolster the precision of predictions.

Received: 6 Nov 2023.

Key words and phrases. Raw moments; Parsimony; Maximum likelihood estimation; Inverted exponential distribution.

In the pursuit of increased flexibility in statistical models, a majority of recent proposals in literature involve the introduction of one or more parameters to an existing probability model. For instance, [2] introduced the Exponentiated Exponential Distribution, while [3] proposed the Transmuted-G family of distribution. [4] proposed the Weibull-G family of distribution, and [5] contributed research on the Logistic-X family of distribution, among others. The addition of parameters may seem at odds with the principle of parsimony. However, maintaining parsimony remains essential when modifying probability distributions, as it is preferable to construct models with a minimal number of parameters while preserving high flexibility in data modeling [6].

In line with this perspective, recent scholars, including [7] have introduced a way of modifying a density function without adding more parameter(s) In the context of enhancing reliability analysis and modeling, this study endeavors to introduce a novel probability model, the New Sine Inverted Exponential distribution by implementing the New Sine G modification, as recommended by [7]. The cumulative density function (CDF), which is based on this method is represented by the relationship

$$F_v^*(v) = \sin\left(\pi 4^{-1} G(v)\left[G(v)+1\right]\right) \tag{1}$$

and upon differentiation, produces the new distribution's probability density function (PDF).

As presented by [8], a random variable v is said to follow an inverted exponential distribution if its probability density function (PDF) and cumulative density function (CDF) is in the form presented (2) and (3).

$$g(v) = \frac{1}{\mho v^2}\exp\left(-\frac{1}{\mho v}\right) \tag{2}$$

$$G(v) = \exp\left(-\frac{1}{\mho v}\right) \tag{3}$$

For $v > 0$ and $\mho \in \mathbb{R}$

Therefore, the New Sine G family of distribution has CDF in the form presented below;

$$F_v^*(v) = \sin\left(\pi 4^{-1} G(v)\left[G(v)+1\right]\right) \tag{4}$$

where $G(v)$ is the cumulative density of the distribution to be transformed.

For the New Sine Inverted Exponential Distribution (*NSIvED*), the CDF is obtained by substituting (3) into (4) as presented below;

$$F_v^*(v) = \sin\left(\pi 4^{-1}\left(\exp\left(-\frac{1}{\mho v}\right)\right)\left[\exp\left(-\frac{1}{\mho v}\right)+1\right]\right) \tag{5}$$

To obtain the density function, the above equation (5) is differentiated with respect to *v* as follows;

$$f_v^*(v) = \frac{d}{dv}\left(F_v^*(v)\right) = \frac{d}{dv}\left\langle \sin\left(\pi 4^{-1} \exp\left(-\frac{1}{\mho v}\right)\left[\exp\left(-\frac{1}{\mho v}\right)+1\right]\right)\right\rangle$$

$$f_v^*(v) = \pi 4^{-1}\left(\frac{1}{\mho^2 v^4}\exp\left(-\frac{1}{\mho v}\right)\right)\cos\left(\pi 4^{-1}\left[\exp\left(-\frac{1}{\mho v}\right)\left[\exp\left(-\frac{1}{\mho v}\right)+1\right]\right]\right) \quad (6)$$

For $v > 0$ and $\mho \in \mathbb{R}$

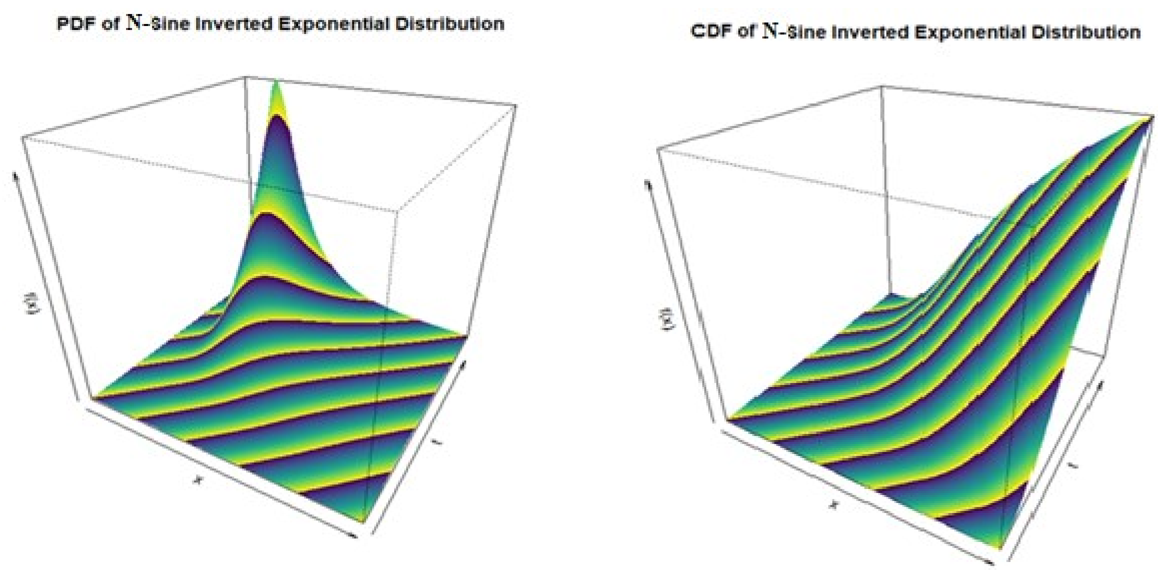

Figure I: PDF and CDF of *NSIvED*

MIXTURE REPRESENTATION

To establish some of the statistical properties, we have derived a mixture representation for the Probability Density Function (PDF) of the *NSIvED* in this section (for a comprehensive explanation, please refer to [7]). A random variable *v* (v~*NSIvED*) which is a member of the family of New Sine family of probability models, a general representation for the PDF is given as;

$$f_v^*(v) = \sum_{y=0}^{\infty}\sum_{x=0}^{x+2y}(-1)^y(1+2y+x)\left[(2y+1)!\right]^{-1}\left(4\pi^{-1}\right)^{2y+1}\left(G(v)\right)^{2y+x} g(v) \quad (7)$$

Let

$$A_\theta = \sum_{y=0}^{\infty}\sum_{x=0}^{x+2y}(-1)^y\left[(2y+1)!\right]^{-1}\left(4\pi^{-1}\right)^{2y+1} \quad (8)$$

then,

$$f_v^*\left(v\right)=A_\theta\left(1+2y+x\right)\left(G\left(v\right)\right)^{2y+x}g\left(v\right) \tag{9}$$

Substituting (2) and (3) for $G(v)$ and $g(v)$ in (9), yields the following,

$$f_v^*\left(v\right)=A_\theta\left(1+2y+x\right)\left(\exp\left(-\frac{1}{\mho v}\right)\right)^{2y+x}\left(\frac{1}{\mho v^2}\exp\left(-\frac{1}{\mho v}\right)\right)$$

$$f_v^*\left(v\right)=A_\theta\left(1+2y+x\right)\left(\mho v^2\right)^{-1}\left(\exp\left(-\frac{1}{\mho v}\left(2y+x+1\right)\right)\right) \tag{10}$$

DERIVED PROPERTIES OF *NSIvED*

Reliability Properties

i. Survival function

The survival function only shows the likelihood of that an event will not have happened by time *t*. This can determine its value by subtracting the CDF from one as follows.

$$S\left(v\right)=1-F_v^*\left(v\right) \tag{11}$$

Substituting for $F_v^*\left(v\right)$ in the above, we get

$$S\left(v\right)=1-\sin\left(\pi 4^{-1}\exp\left(-\frac{1}{\mho v}\right)\left[\exp\left(-\frac{1}{\mho v}\right)+1\right]\right) \tag{12}$$

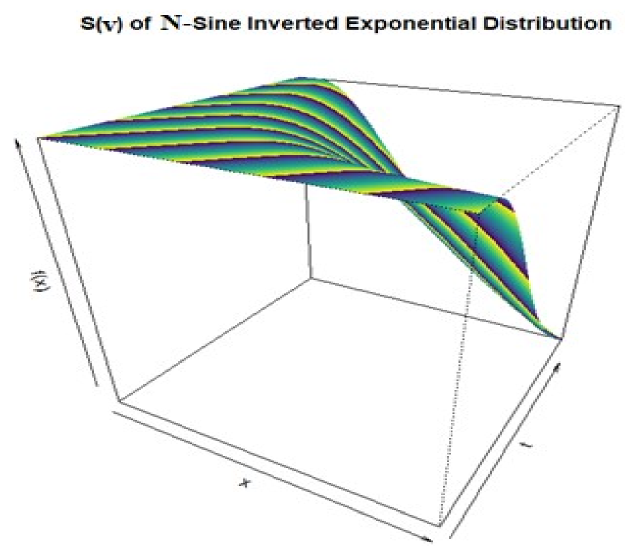

Figure II: S(*v*) of *NSIvED*

ii. Hazard function

In order to model the failure rate in a survival study, the hazard function is utilized. The immediate risk that the desired occurrence will occur in a very short period of time. The hazard function is, more specifically, the ratio of the PDF to the random variable's survival function as shown below;

$$H(v)=\frac{f_v^*(v)}{S(v)} \tag{13}$$

Substituting for $f_v^*(v)$ and $S(v)$, we obtain the hazard function as

$$H(v)=\frac{\pi 4^{-1}\left(\frac{1}{\mho^2 v^4}\exp\left(-\frac{1}{\mho v}\right)\right)\cos\left(\pi 4^{-1}\left[\exp\left(-\frac{1}{\mho v}\right)\left[\exp\left(-\frac{1}{\mho v}\right)+1\right]\right]\right)}{1-\sin\left(\pi 4^{-1}\exp\left(-\frac{1}{\mho v}\right)\left[\exp\left(-\frac{1}{\mho v}\right)+1\right]\right)} \tag{14}$$

iii. Quantile Function

The quantile function of a random variable is obtained by inverting the CDF of the distribution of the random variable. This can be illustrated as follows. Let $v \sim NSIvED(\mho)$, then the quantile function of v is obtained by solving the equation;

$$\Phi_v^* = F_v^{*(-1)}(v) \tag{15}$$

For the family of N-Sine Distribution, the quantile function is given by the expression below;

$$\Phi_v^*(v)=\Phi_G\left(\left(4\pi^{-1}\arcsin(v)+4^{-2}\right)^{1/2}-2^{-1}\right) \tag{16}$$

where $\Phi_G=-\left[\mho\log(q)\right]^{-1}$ is the quantile function of the distribution being transformed (for a comprehensive explanation, please refer to [7]).

Therefore, the quantile function of the $NSIvED$ is thus given as,

$$\Phi_v^*(v)=-\left[\mho\log(q)\right]^{-1}\left(\left(4\pi^{-1}\arcsin(v)+4^{-2}\right)^{1/2}-2^{-1}\right) \tag{17}$$

STATISTICAL PROPERTIES

i. Raw moments

The raw moments of $v \sim NSIvED(\mho)$ is derived as;

$$\Xi^r = \mathrm{E}(v^r)=\int_0^\infty v^r f_v^*(v) dv f_v^*(v)=\int_0^\infty v^r A_\theta (1+2y+x)(\mho v^2)^{-1}\left(\exp\left(-\frac{1}{\mho v}(2y+x+1)\right)\right)dv \tag{18}$$

Let

$$k=(1+2y+x)(\mho v)^{-1}\Rightarrow v=\frac{1+2y+x}{\mho k}\Rightarrow\frac{dk}{dv}=-\frac{(1+2y+x)}{\mho v^{2}} \tag{19}$$

Substituting and simplifying (19) into (18), we have that:

$$=-A_{\theta}\left(\frac{1+2y+x}{\mho}\right)\int_{0}^{\infty}\left(\frac{1+2y+x}{\mho k}\right)^{r-2}\left(\exp(-k)\right).\frac{-\mho\left(\frac{1+2y+x}{\mho k}\right)^{2}}{(1+2y+x)}dk$$

$$=-A_{\theta}(1+2y+x)\left(\frac{1+2y+x}{\mho}\right)^{r-2}\int_{0}^{\infty}(k)^{-(r-2)}\left(\exp(-k)\right).\frac{(1+2y+x)}{(\mho k)^{2}}dk$$

$$=-A_{\theta}\left(\frac{1+2y+x}{\mho}\right)^{2}\left(\frac{1+2y+x}{\mho}\right)^{r-2}\int_{0}^{\infty}(k)^{-(r-2)-2}\left(\exp(-k)\right)dk$$

$$=A_{\theta}(-1)^{r}\left(\frac{1+2y+x}{\mho}\right)^{r}\int_{0}^{\infty}(k)^{-r}\left(\exp(-k)\right)dk$$

$$\Xi^{r}=A_{\theta}(-1)^{r}\left(\frac{2y+x+1}{\mho}\right)^{r}\Gamma(1-r) \tag{20}$$

ii. Moment Generating Function

For $v\sim NSIvED(\mho)$, the MGF is derived as follows;

$M_{v}^{*}(t)=\mathrm{E}\left[\exp(tv)\right]$

Using Taylor's series, it is known that the expansion of $\exp(tv)=\sum_{\rho=0}^{\infty}\frac{(tv)^{\rho}}{\rho!}$

Therefore,

$$M_{v}^{*}(t)=\mathrm{E}\left[\sum_{\rho=0}^{\infty}\frac{(tv)^{\rho}}{\rho!}\right]=\sum_{\rho=0}^{\infty}\frac{t^{\rho}}{\rho!}\mathrm{E}\left[v^{\rho}\right]$$

Since $\mathrm{E}\left[v^{\rho}\right]=\Xi^{r}$ then it implies that

$$M_{v}^{*}(t)=\sum_{\rho=0}^{\infty}\frac{t^{\rho}}{\rho!}A_{\theta}(-1)^{r}\left(\frac{2y+x+1}{\mho}\right)^{r}\Gamma(1-r) \tag{21}$$

iii. Characteristic function

The characteristic function $\left(Q_{V}(t)\right)$, also known as a characteristic generating function, as a concept in probability theory and statistics is used to describe the probability distribution of a random variable. The function is particularly useful in the analysis of sums of random variables as they provide a compact and convenient way to work with probability distributions. For a random variable with *NSIvED*, the $Q_V(t)$ is given as;

$$Q_V(t) = E\left[e^{itv}\right] = \int_0^\infty e^{itv} f_v^*(v)\,dv \tag{22}$$

Using the Taylors series, it can be shown that $e^{itv} = \sum_{j=0}^{\infty} \frac{(it)^n}{n!} v^n ; \Rightarrow E\left[e^{itv}\right] = \sum_{j=0}^{\infty} \frac{(it)^n}{n!} E\left[v^n\right]$

Where $E\left[v^n\right]$is the raw moment of the random variable *v*. Therefore;

$$Q_V(t) = \sum_{j=0}^{\infty} \frac{(it)^n}{n!} E\left[v^n\right] \tag{23}$$

On making the substitution of (20) for $E\left[v^n\right]$ in (23)(22), we obtain the following;

$$Q_V(t) = \sum_{j=0}^{\infty} \frac{(it)^n}{n!} \mathrm{M}_v^*(t) \tag{24}$$

Where $\mathrm{M}_v^*(t) = \sum_{\rho=0}^{\infty} \frac{t^\rho}{\rho!} A_\theta (-1)^r \left(\frac{2y+x+1}{\mho}\right)^r \Gamma(1-r)$

iv. PDF of the n^{th} order statistics of ETAD

Suppose $v_1, v_2, \cdots, v_n$ from the *NSIvED* are ordered in the form $v_{(1)} \le v_{(2)} \le \cdots \le v_{(n)}$ of random samples from the density, then the PDF of the n^{th} order statistics can be defined as follows:

$$f_{(k,n)}(v) = \frac{n!}{(k-1)!(n-k)!} f_v^*(v) F_v^*(v)^{k-1} \left[1 - F_v^*(v)\right]^{n-k}$$

where, $F(v)$ and $f(v)$ are the CDF and PDF of the *NSIvED*. For simplicity, $\left[1 - F_v^*(v)\right]^{n-k}$ can be represented using the binomial series as; $\left[1 - F_v^*(v)\right]^{n-k} = \sum_{c=1}^{\infty} \binom{n-k}{c} (-1)^c \left[F_v^*(v)\right]^c$. Thus,

$$f_{(k,n)}(v) = \frac{n!}{(k-1)!(n-k)!} f_v^*(v) F_v^*(v)^{k-1} \sum_{c=1}^{\infty} \binom{n-k}{c} (-1)^c \left[F_v^*(v)\right]^c \tag{25}$$

Using the above, the n^{th} order statistic for a random variable with *NSIvED* is thus obtained by substituting (6) and (5) into (25) as follows:

$$f_{(k,n)}(v)=\frac{n!}{(k-1)!(n-k)!}\left\{\pi 4^{-1}\left(\frac{1}{\mho^2 v^4}e^{\left(-\frac{1}{\mho v}\right)}\right)\cos\left(\pi 4^{-1}\left[e^{\left(-\frac{1}{\mho v}\right)}\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right]\right)\right\}.$$

$$\left[\sin\left(\pi 4^{-1}\left(e^{\left(-\frac{1}{\mho v}\right)}\right)\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right)\right]^{k-1}\sum_{c=1}^{\infty}\binom{n-k}{c}(-1)^c\left[\sin\left(\pi 4^{-1}\left(e^{\left(-\frac{1}{\mho v}\right)}\right)\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right)\right]^c$$

$$f_{(k,n)}(v)=\frac{n!}{(k-1)!(n-k)!}\left\{\frac{\pi e^{\left(-\frac{1}{\mho v}\right)}}{4\mho^2 v^4}\cos\left(\pi\frac{e^{\left(-\frac{1}{\mho v}\right)}}{4}\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right)\right\}\beta_\theta\left[\sin\left(\pi\frac{e^{\left(-\frac{1}{\mho v}\right)}}{4}\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right)\right]^{k-1+c} \quad (26)$$

Where $\beta_\theta=\sum_{c=1}^{\infty}\binom{n-k}{c}(-1)^c$.

The expression in equation (26) presents the n^{th} order statistic for a random variable with *NSIvED*. To obtain the smallest and n^{th} order statistic, 1 and *n* will be substituted into (26) and simplified.

MEASURES OF INFORMATION

Entropy

Entropy is a metric indicating how much information or uncertainty there is in a random observation of a population's real composition. For a continuous random variable *v* with *NSIvED,* the Renyi entropy is can be derived as follows:

$$R_\tau=\frac{1}{1-\tau}\log\left(\int_0^\infty\left(f_v^*(v)\right)^\tau dv\right) \quad (27)$$

Substituting for (10) for $f_v^*(v)$ in (27), yields the following;

$$R_\tau=\frac{1}{1-\tau}\log\left(\frac{A_{R\theta}}{\mho}\int_0^\infty\left(\frac{e^{\left(-\frac{1}{\mho v}(2y+x+1)\right)}}{v^2}\right)^\tau dv\right) \quad (28)$$

Where $A_{R\theta}=A_\theta(1+2y+x)$

By transformation, let

$$c=\frac{2y+x+1}{\mho v}\Rightarrow v=\frac{2y+x+1}{c\mho};\frac{dc}{dv}=-\frac{2y+x+1}{\mho v^2}\Rightarrow dv=-\frac{\mho v^2}{2y+x+1}dc \quad (29)$$

On substituting (29) into (28), the following is obtained.

$$R_{\tau}=\frac{1}{1-\tau}\log\left(\frac{A_{R\theta}}{\mho}\int_{0}^{\infty}\left(\frac{e^{-c}}{\left(\frac{2y+x+1}{c\mho}\right)^{2}}\right)^{\tau}.-\frac{\mho\left(\frac{2y+x+1}{c\mho}\right)^{2}}{2y+x+1}dc\right)$$

$$R_{\tau}=\frac{1}{1-\tau}\log\left(-\frac{A_{R\theta}\mho^{2\tau}}{\mho}\left(2y+x+1\right)^{1-2\tau}\int_{0}^{\infty}c^{2\tau-1}e^{-c\tau}\,dc\right)$$

$$R_{\tau}=\frac{1}{1-\tau}\log\left(-\frac{A_{R\theta}\mho^{2\tau}}{\mho}\left(2y+x+1\right)^{1-2\tau}\Gamma\left(2\tau\right)\right) \tag{30}$$

PARAMETER ESTIMATION

Using the method of maximum likelihood, the parameter of the *NSIvED* will be estimated as follows. The likelihood function $L(\theta)$ is given by (31).

$$L(\theta)=\prod_{j=1}^{n}f_{v}^{*}\left(v\right) \tag{31}$$

For a random variable with *NSIvED*, $L(\theta)$is obtained by making the substitution of (10) for $f_{v}^{*}(v)$ in (31). Thus; (27)

$$L(\theta)=\left(\frac{\pi}{4\mho^{2}}\right)^{n}\prod_{j=1}^{n}\left(\frac{1}{v^{4}}e^{\left(-\frac{1}{\mho v}\right)}\right)\cos\left(\pi 4^{-1}\left[e^{\left(-\frac{1}{\mho v}\right)}\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right]\right) \tag{32}$$

$$L(\theta)=\left(\frac{\pi}{4\mho^{2}}\right)^{n}\sum_{j=1}^{n}\left\{\left(\frac{1}{v^{4}}e^{\left(-\frac{1}{\mho v}\right)}\right)\cos\left(\pi 4^{-1}\left[e^{\left(-\frac{1}{\mho v}\right)}\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right]\right)\right\}$$

Taking the log of $L(\theta)$, we have

$$l\left[L(\theta)\right]=n\log\left(\frac{\pi}{4\mho^{2}}\right)+\sum_{j=1}^{n}\log\left\{\left(\frac{1}{v^{4}}e^{\left(-\frac{1}{\mho v}\right)}\right)\cos\left(\pi 4^{-1}\left[e^{\left(-\frac{1}{\mho v}\right)}\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right]\right)\right\} \tag{33}$$

Differentiating (33) with respect to the unknown parameter $\mho$, it can be observed (in (34)) that the parameter estimate does not exist in a closed form. Thus, a numerical estimation will be used to obtain the parameter estimate.

$$\frac{dl\left[L(\theta)\right]}{d\mho}=\frac{d}{d\mho}n\log\left(\frac{\pi}{4\mho^{2}}\right)+\frac{d}{d\mho}\sum_{j=1}^{n}\log\left\{\left(\frac{1}{v^{4}}e^{\left(-\frac{1}{\mho v}\right)}\right)\cos\left(\pi 4^{-1}\left[e^{\left(-\frac{1}{\mho v}\right)}\left[e^{\left(-\frac{1}{\mho v}\right)}+1\right]\right]\right)\right\}=0 \tag{34}$$

SIMULATION STUDY

In this part, we use simulations to evaluate the relative bias and relative mean square error for the estimates. The "Adequacy models" package of the R program is used to create the sample parameter estimations. For various sample sizes n = [20, 40, 60, 80, 100, 150, 300, 600 and 1000] and various values of the parameter [0.3, 0.5, 0.7 and 0.9], the sampling distributions are obtained.

$$Bias = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta_i\right) \tag{35}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta_i\right)^2} \tag{36}$$

Table 1 contains a summary of the evaluation of the bias and RMSE features of the MLE parameter estimates. Nine (9) different sample sizes were used to determine the sampling distributions.

Table 1: Showing the Behavior of Parameters of *NSIvED*

Sample size	True Parameters	MLE	BIAS	RMSE
20	0.3	0.2651	0.0349	0.0001
	0.5	0.3830	0.1170	0.0007
	0.7	0.4761	0.2239	0.0025
	0.9	0.5522	0.3478	0.0060
40	0.3	0.2190	0.0810	0.0002
	0.5	0.3177	0.1823	0.0008
	0.7	0.3966	0.3034	0.0023
	0.9	0.4625	0.4375	0.0048
60	0.3	0.2175	0.0825	0.0001
	0.5	0.3173	0.1827	0.0006
	0.7	0.3955	0.3045	0.0015
	0.9	0.4605	0.4395	0.0032
80	0.3	0.2399	0.0601	0.0000
	0.5	0.3434	0.1566	0.0003
	0.7	0.4180	0.2820	0.0010
	0.9	0.4903	0.4097	0.0021
100	0.3	0.2163	0.0837	0.0001
	0.5	0.3109	0.1891	0.0004
	0.7	0.3862	0.3138	0.0010
	0.9	0.4490	0.4510	0.0020

150	0.3	0.2212	0.0788	0.0000
	0.5	0.3164	0.1836	0.0002
	0.7	0.3920	0.3080	0.0006
	0.9	0.4560	0.4440	0.0013
300	0.3	0.2174	0.0826	0.0000
	0.5	0.3106	0.1894	0.0001
	0.7	0.3847	0.3153	0.0003
	0.9	0.4463	0.4537	0.0007
600	0.3	0.2100	0.0900	0.0000
	0.5	0.3025	0.1975	0.0001
	0.7	0.3766	0.3234	0.0002
	0.9	0.4386	0.4614	0.0004
1000	0.3	0.2093	0.0907	0.0000
	0.5	0.3020	0.1980	0.0000
	0.7	0.3764	0.3236	0.0001
	0.9	0.4387	0.4613	0.0002

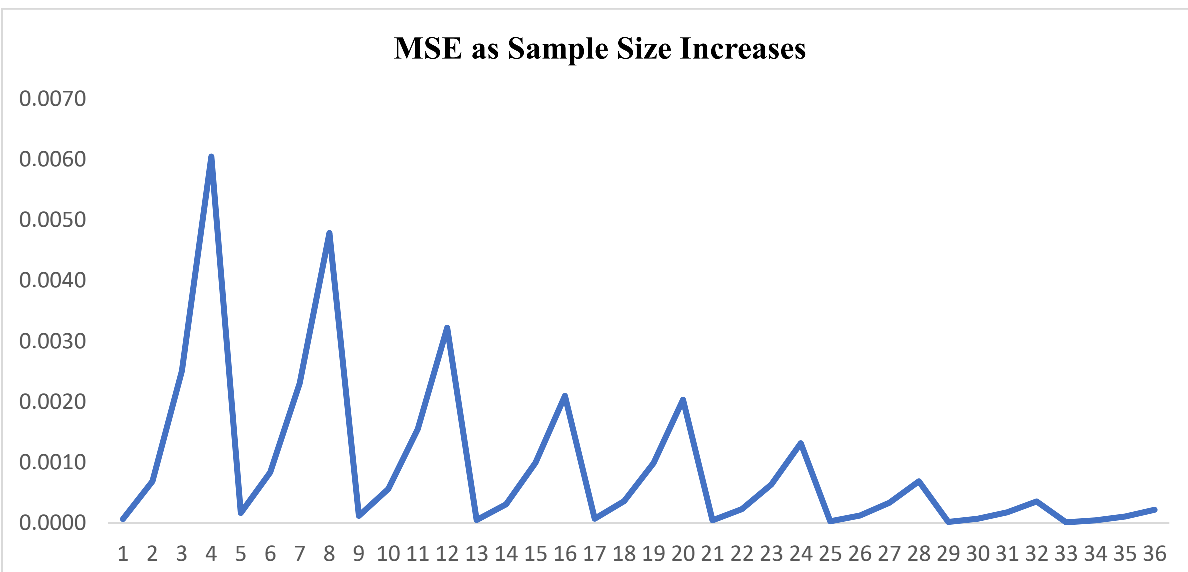

Figure III: Showing the Behavior of the *NSIvED* MSE as the Sample Size Increases

As inferred by the central limit theorem, as the sample size increases, it is expected that the parameter estimates tend to the true population parameter. This implies a reduction in the root mean square error of the estimates. Figure III shows that as the number of the sample increases from 20-1000, the magnitude of the RMSE reduces and tends to zero.

APPLICATION AND GOODNESS OF FIT USING SOME REAL DATASETS

Several distributions have been constructed in statistical literature that models experimental data. The New Sine Inverse Exponential Distribution (NSIED) presented by [9], the Prakaamy Distribution (PkD) by [10], and the Ram Awadh Distribution (R-AD) produced by [11] were the distributions that were compared with the *NSIvED* in the current study.

Using the generated distribution and certain preexisting probability distributions, two actual datasets were modeled. The first data set includes the total of 202 athletes' skin folds, which were measured at the Australian Institute of Sports and were published by [12]. The second dataset included 72 guinea pig survival periods in days that were voluntarily exposed to various dosages of tubercle bacilli in Bjerkedal, (1960) (as cited in [13]).

First data set

"148.9, 149.0, 156.0, 156.9, 157.9, 158.9, 162.0, 162.0, 162.5, 163.0, 163.9, 165.0, 166.1, 166.7, 167.3, 167.9, 168.0, 168.6, 169.1, 169.8, 169.9, 170.0, 170.0, 170.3, 170.8, 171.1, 171.4, 171.4, 171.6, 171.7, 172.0, 172.2, 172.3, 172.5, 172.6, 172.7, 173.0, 173.3, 173.3, 173.5, 173.6, 173.7, 173.8, 174.0, 174.0, 174.0, 174.1, 174.1, 174.4, 175.0, 175.0, 175.0, 175.3, 175.6, 176.0, 176.0, 176.0, 176.0, 176.8, 177.0, 177.3, 177.3, 177.5, 177.5, 177.8, 177.9, 178.0, 178.2, 178.7, 178.9, 179.3, 179.5, 179.6, 179.6, 179.7, 179.7, 179.8, 179.9, 180.2, 180.2, 180.5, 180.5, 180.9, 181.0, 181.3, 182.1, 182.7, 183.0, 183.3, 183.3, 184.6, 184.7, 185.0, 185.2, 186.2, 186.3, 188.7, 189.7, 193.4, 195.9".

Second data set

"12, 15, 22, 24, 24, 32, 32, 33, 34, 38, 38, 43, 44, 48 , 52, 53, 54, 54, 55, 56, 57, 58, 58, 59, 60, 60, 60, 60, 61, 62, 63, 65, 65, 67, 68, 70, 70, 72, 73, 75, 76, 76, 81, 83, 84, 85, 87, 91, 95, 96, 98, 99, 109, 110, 121, 127, 129, 131, 143, 146, 146, 175, 175, 211, 233, 258, 258, 263, 297, 341, 341, 376".

Table 1: Application to First Dataset

Models	AIC	BIC	CAIC	HQIC	MLE	Rank
NSIvED	**5598.165**	**5601.473**	**5598.19**	**5599.50**	**0.0143**	**1**
NSIED	66894.560	66897.87	66894.58	66895.90	1.1952	2
PkD	11415.070	11418.38	11415.09	11416.41	0.5942	3
R-AD	31705.48	31708.79	31705.50	31706.82	1.3952	4

Key: New Sine Inverted Exponential Distribution = *NSIvED*

New Sine Inverse Exponential Distribution = NSIED

Prakaamy Distribution = PkD

Ram Awadh Distribution = R-AD

Table 2: Application to Second Dataset

Models	AIC	BIC	CAIC	HQIC	MLE	Rank
NSIvED	**720.5444**	**722.8751**	**720.5984**	**721.4758**	**0.3409**	**1**
NSIED	2029.2640	2031.5950	2029.3180	2030.1950	1.8104	4
PkD	856.7214	858.9981	856.7786	857.6278	0.0603	3
R-AD	856.7165	858.9932	856.7737	857.6229	0.0601	2

Key: New Sine Inverted Exponential Distribution = NSIvED

New Sine Inverse Exponential Distribution = NSIED

Prakaamy Distribution = PkD

Ram Awadh Distribution = R-AD

Based on the analysis of the two datasets that have been fitted with distinct probability models, a notable trend emerges: the *NSIvED* consistently produced the smallest value among all the information criterion measures used for evaluation. This observation suggests that, in terms of the results obtained, the *NSIvED* model outperforms the others and is, therefore, the most favorable choice. This finding underscores the superiority of the *NSIvED* model in capturing and explaining the underlying data, as it exhibits a superior fit when compared to the alternative models.

CONCLUSION

An Inverted Exponential Distribution modification known as the N Sine Inverted Exponential Distribution is discussed in this study. Through simulation studies, the behavior of the distributions was found to be in line with the central limit theorem as the value of the RMSE was decreasing with increase in the sample size. Using two actual data sets, the relevance of this distribution for modeling lifetime data was demonstrated. The NSIED, PkD, and R-AD were among one-parameter probability distributions with which the performance of the NSIvED was compared. The Kolmogorov-Smirnov test, the Akaike Information Criterion, the Consistent Akaikes Information Criterion, the Bayesian Information Criterion, and the Hannan-Quinn Information Criterion were used to compare distributions. The study showed that when modeling lifetime data, the NSIvED outperforms the NSIED, PkD, and the R-AD as it has the smallest information criterions.

References

[1] M. Muhammad, H.M. Alshanbari, A.R.A. Alanzi, L. Liu, W. Sami, C. Chesneau, F. Jamal, A new generator of probability models: The exponentiated sine-G family for lifetime studies, Entropy 23 (2021) 1394. https://doi.org/10.3390/e23111394.

[2] R.D. Gupta, D. Kundu, Exponentiated exponential family: an alternative to gamma and Weibull distributions, Biom. J. 43 (2001) 117–130. https://doi.org/10.1002/1521-4036(200102)43:1<117::aid-bimj117>3.0.co;2-r.

[3] W.T. Shaw, I.R. Buckley, The alchemy of probability distributions: beyond Gram-Charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map, Department of Mathematics, King's College London, UK, 2007.

[4] M. Bourguignon, R. Silva, The Weibull–G family of probability distributions, J. Data Sci. 12 (2014), 53–68.

[5] H.M. Tahir, M.G. Cordeiro, A. Alzaatreh, M. Mansoor, M. Zubair, The logistic-X family of distributions and its applications, Comm. Stat. - Theory Meth. 45 (2016), 7326–7349. https://doi.org/10.1080/03610926.2014.980516.

[6] I. Malakar, Theorizing probability distribution in applied statistics, J. Popul. Develop. 1 (2020), 79-95.

[7] Z.C. Mahmood, Chesneau, M.H. Tahir, A new sine-G family of distributions: Properties and applications, Bull. Comp. Appl. Math. 7 (2019), 53–81.

[8] S. Dey, Inverted exponential distribution as a life distribution model from a Bayesian viewpoint, Data Sci. J. 6 (2007), 107-113. https://doi.org/10.2481/dsj.6.107.

[9] M. Zafar, C. Christophe, A new sine-G family of distributions: Properties and applications, hal-02079224, pp. 1-20, 2019. https://hal.science/hal-03580871.

[10] K.K. Shukla, Prakaamy distribution with properties and applications, J. Appl. Quant. Meth. 13 (2018), 30-38.

[11] K.K. Shukla, Ram Awadh distribution and its applications, Biom. Biostat. Int. J. 7 (2018), 515–523.

[12] R. Al-Aqtash , C. Lee, F. Famoye, Gumbel - Weibull distribution: properties and application, J. Mod. Appl. Stat. Meth. 13 (2014), 201–225. https://doi.org/10.22237/jmasm/1414815000.

[13] L. Tomy, C. Chesneau, Applications of the sine modified Lindley distribution to biomedical data, Math. Comp. Appl. 27 (2022), 43. https://doi.org/10.3390/mca27030043.

[14] K.K. Shukla, Ram Awadh distribution and its applications, Biom. Biostat. Int. J. 7 (2018), 515–523.