A SINE-ISHITA DISTRIBUTION: PROPERTIES AND APPLICATIONS TO SURVIVAL DATA.

Abstract:
This paper presents a novel extension of the Ishita distribution called "sine-Ishita distribution". The Sine-G family was used to derive the proposed sine-Ishita distribution. The research investigates some properties of the sine-Ishita distribution such as the moments, moments generating function, characteristics function, survival and hazard functions as well as distribution of order statistics. The research also estimates the parameter of the proposed distribution using the method of maximum likelihood. The research illustrated the usefulness of the proposed distribution by applications to four real life datasets and the results show that the proposed distribution fits the data better compared to the other five distributions considered in the study. This performance is an indication that the sine-Ishita distribution can be used for modelling real life event arising from many areas of life.

Keywords: Ishita distribution, Sine-G family, Properties, Maximum Likelihood Estimation, Applications.

1 Introduction

Many distributions have been used over the years for modelling data, and research has shown that extending these probability distributions lead to compound distributions which are more flexible for modelling data in several areas such as the medical/biological sector, engineering, economics, genetics, agronomy (Rao and Aslam, 2021; Strzelecki, 2021; Tung *et al*., 2021; Reynolds *et al*., 2021; Prataviera, 2022).

With the aim of developing new compound distributions, many families of distributions have been defined in literature and some of the recent families include the sine-G family by Kumar et al. (2015), the quadratic rank transmutation map by Shaw and Buckley (2007), Exponentiated T-X by Alzaghal *et al*. (2013), Weibull-X by Alzaatreh *et al*. (2013), Weibull-G by Bourguignon *et al*. (2014), a Lomax-G family by Cordeiro *et al.* (2014), a new Weibull-G family by Tahir *et al.* (2016), a Lindley-G family by Cakmakyapan and Ozel (2016), a Gompertz-G family by Alizadeh *et al.* (2017) and Odd Lindley-G family by Gomes-Silva *et al*. (2017), odd Lomax-G family by Cordeiro *et al.* (2019), an odd Chen-G family of distributions by Anzagra *et al*. (2022), a new sine family of generalized distributions by Benchiha *et al*. (2023), the Topp-Leone type II exponentiated half logistic-G family of distributions by Gabanakgosi and Oluyede (2023), a novel bivariate Lomax-G family of distributions Fayomi *et al*. (2023), X-exponential-G Family of Distributions by Mohammad (2024) and the new Frechet-G family of continuous probability distributions by Ieren *et al*. (2024).

These families have been used to study compound distribution such as exponential-Lindley distribution by Ieren and Balogun (2021), transmuted Kumaraswamy distribution by Khan *et al*. (2016), a Lomax-exponential distribution by Ieren and Kuhe (2018), a sine Lomax-exponential distribution by Joel *et al*. (2024), odd Lindley inverse exponential distribution Ieren and Abdullahi (2020), a Lomax-inverse exponential distribution by Abdulkadir *et al*. (2020), Power Lindley distribution by Ghitany *et al.* (2013), bivariate generalized Rayleigh distribution by Abdel-Hady (2013), Lomax-Frechet distribution by Gupta *et al*. (2015), a

UNDER PEER REVIEW

transmuted Weibull-exponential distribution by Yahaya and Ieren (2017), a transmuted odd Lindley-Rayleigh distribution by Umar *et al*. (2021), Weibull-Frechet by Afify *et al*. (2016), transmuted odd generalized exponential-exponential distribution Abdullahi *et al*. (2018), Sine-exponential distribution by Isa *et al*., (2022a), sine-Lomax distribution by Mustapha et al. (2023), the sine-modified Lindley distribution by Tomy *et al*., (2021) and sine Burr XII distribution by Isa *et al*., (2022b).

The Ishita distribution was proposed by Shanker and Shukla (2017). This distribution is a two-component mixture of exponential distribution with parameter h and a gamma distribution with parameters $(3,\eta)$ using mixing proportion, $\dfrac{\eta^3}{\eta^3+2}$.

The probability density function (pdf) of the Ishita distribution (ISD) according to Shanker and Shukla (2017) is defined by

$$g(x)=\frac{\eta^3}{\eta^3+2}\left(\eta+x^2\right)\mathrm{e}^{-\eta x} \tag{1}$$

The corresponding cumulative distribution function (CDF) of ISD is given by

$$G(x)=1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]\mathrm{e}^{-\eta x} \tag{2}$$

where, $x>0$ and $\eta>0$ is a scale parameter of the ISD. According to Shanker and Shukla (2017), the ISD performed better than the Akash distribution, Lindley distribution and exponential distribution considering two applications to real life data.

The aim of this paper is to introduce a new continuous distribution called the Sine-Ishita distribution (ISD) by using the Sine-G family.

This paper is organized in different sections as follows: definition of the new distribution with its plots is provided in section 2. Section 3 derived some properties of the proposed distribution. The estimation of parameters using maximum likelihood estimation (MLE) and simulation study is presented in section 4. An application of the new model with other existing distributions to real life data is done in section 5 and a conclusion is given in section 6.

2. The Sine-Ishita distribution (SISD)

According to Kumar et al. (2015), the cumulative distribution function (CDF) of the Sine-G family is expressed by:

$$F(x,\xi)=\sin\left[\frac{\pi}{2}G(x,\xi)\right] \tag{3}$$

And the corresponding probability density function (pdf) (for $x>0$) is defined by

UNDER PEER REVIEW

$$f(x,\xi)=\frac{\pi}{2}g(x,\xi)\cos\left[\frac{\pi}{2}G(x,\xi)\right] \tag{4}$$

respectively, where $g(x)$ and $G(x)$ represent the *pdf* and the *cdf* of the continuous distribution to be modified respectively.
Putting equations (1) and (2) into equations (3) and (4) and simplifying, we obtain the cdf and pdf of the SISD given in equations (5) and (6) respectively as follows:

$$F(x,\eta)=\sin\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right] \tag{5}$$

and

$$f(x;\eta)=\frac{\pi}{2}\frac{\eta^3}{\eta^3+2}\left(\eta+x^2\right)e^{-\eta x}\cos\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right] \tag{6}$$

where, $x>0$; and $\eta>0$ is a scale parameter of the sine-Ishita distribution.

Plots of the pdf and cdf of the SISD using some parameter values are presented in Figure 1 as follows.

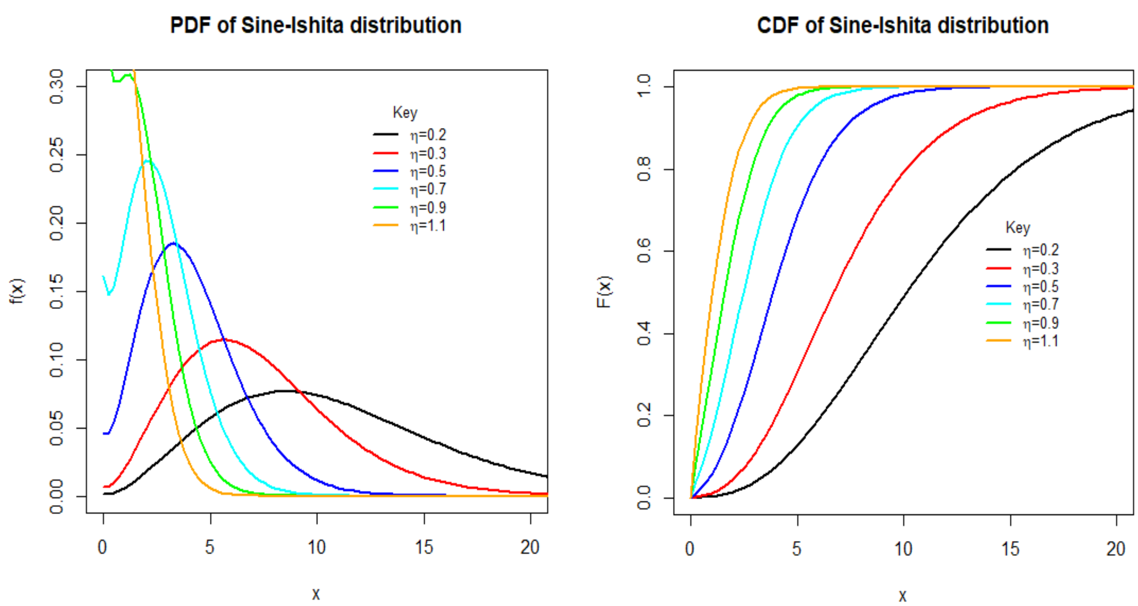

Figure 1: PDF and CDF of the SISD for different values of the parameters.

From the figure above, it can be seen that the pdf SISD is positively skewed and takes various shapes depending on the parameter values. Also, from the above plot of the cdf, it is clear that the cdf equals one when *X* approaches infinity and equals zero when *X* tends to zero as normally expected.

3. Some Properties of SISD

In this section, some properties of the SISD distribution are derived and discussed as follows:

3.1 Moments

Let X denote a continuous random variable the n^{th} moment of X is given by;

$$\mu_n^{'} = E\left(X^n\right) = \int_0^\infty x^n f\left(x\right) dx \tag{7}$$

where f(x), the *pdf* of the SISD is as given in equation (6) as:

$$f\left(x;\eta\right) = \frac{\pi}{2}\frac{\eta^3}{\eta^3+2}\left(\eta + x^2\right)\mathrm{e}^{-\eta x}\cos\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x\left(\eta x+2\right)}{\eta^3+2}\right]\mathrm{e}^{-\eta x}\right)\right] \tag{8}$$

Before substituting (8) in (7), the expansion, simplification and linear representation of the pdf of the SISD is done as follows:

First, by using power series expansion on the last term in (8), we obtain:

$$\cos\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x\left(\eta x+2\right)}{\eta^3+2}\right]\mathrm{e}^{-\eta x}\right)\right] = \sum_{k=0}^{\infty}\frac{\left(-1\right)^k\left(\pi\right)^{2k}}{\left(2k\right)!\left(2\right)^{2k}}\left(1-\left[1+\frac{\eta x\left(\eta x+2\right)}{\eta^3+2}\right]\mathrm{e}^{-\eta x}\right)^{2k} \tag{9}$$

Making use of the result in (9) above and simplifying, equation (8) becomes

$$f\left(x;\eta\right) = \sum_{k=0}^{\infty}\frac{\left(-1\right)^k\left(\pi\right)^{2k+1}}{\left(2k\right)!\left(2\right)^{2k+1}}\frac{\eta^3}{\eta^3+2}\left(\eta + x^2\right)\mathrm{e}^{-\eta x}\left(1-\left[1+\frac{\eta x\left(\eta x+2\right)}{\eta^3+2}\right]\mathrm{e}^{-\eta x}\right)^{2k} \tag{10}$$

Using binomial expansion on the last term in (10) gives:

$$\left(1-\left[1+\frac{\eta x\left(\eta x+2\right)}{\eta^3+2}\right]\mathrm{e}^{-\eta x}\right)^{2k} = \sum_{l=0}^{\infty}\left(-1\right)^l\binom{2k}{l}\left(\left[1+\frac{\eta x\left(\eta x+2\right)}{\eta^3+2}\right]\mathrm{e}^{-\eta x}\right)^l \tag{11}$$

Making use of the result in (11) above in equation (10) and simplifying, we obtain:

$$f\left(x;\eta\right) = \sum_{k,l=0}^{\infty}\frac{\left(-1\right)^{k+l}\left(\pi\right)^{2k+1}}{\left(2k\right)!\left(2\right)^{2k+1}}\binom{2k}{l}\frac{\eta^3}{\eta^3+2}\left(\eta + x^2\right)\mathrm{e}^{-\eta\left(l+1\right)x}\left(1+\frac{\eta x\left(\eta x+2\right)}{\eta^3+2}\right)^l \tag{12}$$

Again using binomial expansion on the last term in (16) gives:

UNDER PEER REVIEW

$$\left(1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right)^l=\sum_{m=0}^{\infty}\binom{l}{m}\left(\frac{\eta x(\eta x+2)}{\eta^3+2}\right)^m \tag{13}$$

Making use of the result (13) in equation (12) and simplifying, we obtain:

$$f(x;\eta)=\sum_{k,l,m=0}^{\infty}\frac{(-1)^{k+l}(\pi)^{2k+1}\eta^{m+3}}{(2k)!(2)^{2k+1}(\eta^3+2)^{m+1}}\binom{2k}{l}\binom{l}{m}x^m(\eta x+2)^m(\eta+x^2)\mathrm{e}^{-\eta(l+1)x} \tag{14}$$

Also, using binomial expansion on the last term in (14) gives:

$$(\eta x+2)^m=\sum_{r=0}^{\infty}\binom{m}{r}2^{m-r}(\eta x)^r \tag{15}$$

Making use of the result (15) in equation (14) and simplifying, we obtain:

$$f(x;\eta)=\sum_{k,l,m,r=0}^{\infty}\frac{(-1)^{k+l}(\pi)^{2k+1}\eta^{m+r+3}2^{m-r}}{(2k)!(2)^{2k+1}(\eta^3+2)^{m+1}}\binom{2k}{l}\binom{l}{m}\binom{m}{r}x^{m+r}(\eta+x^2)\mathrm{e}^{-\eta(l+1)x} \tag{16}$$

Now, let $\psi_{klmr}=\sum_{k,l,m,r=0}^{\infty}\frac{(-1)^{k+l}(\pi)^{2k+1}\eta^{m+r+3}2^{m-r}}{(2k)!(2)^{2k+1}(\eta^3+2)^{m+1}}\binom{2k}{l}\binom{l}{m}\binom{m}{r}$ be a constant, which implies that the pdf in equation (16) can also be written in its simple and linear form as:

$$f(x;\eta)=\psi_{klmr}x^{m+r}(\eta+x^2)\mathrm{e}^{-\eta(l+1)x} \tag{17}$$

Now, using the linear representation of the pdf of the SISD in equation (17), the n^{th} ordinary moment of the SISD is derived as follows:

$$\mu_n^{'}=E(X^n)=\int_0^{\infty}x^n f(x)\,dx=\psi_{klmr}\int_0^{\infty}x^{m+n+r}(\eta+x^2)\mathrm{e}^{-\eta(l+1)x}dx \tag{18}$$

Making use of integration by substitution method in equation (18), we perform the following operations:

Let $u=\eta(l+1)x\Rightarrow x=u(\eta(l+1))^{-1}$ such that $\frac{du}{dx}=\eta(l+1)\Rightarrow dx=\frac{du}{\eta(l+1)}$

Substituting for x, u and dx in equation (18) and simplifying; we have:

$$\mu_n^{'}=E(X^n)=\psi_{klmr}\int_0^{\infty}\left(\frac{u}{\eta(l+1)}\right)^{m+n+r}\left(\eta+\left(\frac{u}{\eta(l+1)}\right)^2\right)\mathrm{e}^{-u}\,\frac{du}{\eta(l+1)}$$

UNDER PEER REVIEW

$$\mu_n^{'} = E\left(X^n\right) = \psi_{klmr}\left[\eta\int_0^{\infty}\frac{u^{m+n+r}}{\left[\eta(l+1)\right]^{m+n+r+1}}\mathrm{e}^{-u}du + \int_0^{\infty}\frac{u^{m+n+r+2}}{\left[\eta(l+1)\right]^{m+n+r+3}}\mathrm{e}^{-u}du\right] \quad (19)$$

Now, recall that $\int_0^{\infty} t^{k-1}e^{-t}dt = \Gamma(k)$ and that $\int_0^{\infty} t^{k}e^{-t}dt = \int_0^{\infty} t^{k+1-1}e^{-t}dt = \Gamma(k+1)$

Using the statement above, the n^{th} ordinary moment of X for the SISD is obtained as:

$$\mu_n^{'} = E\left(X^n\right) = \psi_{klmr}\left[\frac{\eta\Gamma(m+n+r+1)}{\left[\eta(l+1)\right]^{m+n+r+1}} + \frac{\Gamma(m+n+r+3)}{\left[\eta(l+1)\right]^{m+n+r+3}}\right] \quad (20)$$

where $\psi_{klmr} = \sum_{k,l,m,r=0}^{\infty}\frac{(-1)^{k+l}(\pi)^{2k+1}\eta^{m+r+3}2^{m-r}}{(2k)!(2)^{2k+1}\left(\eta^3+2\right)^{m+1}}\binom{2k}{l}\binom{l}{m}\binom{m}{r}$ is a constant.

3.2 Moment Generating Function

The moment generating function of a continuous random variable X can be obtained as

$$M_x(t) = E\left[e^{tx}\right] = \int_{-\infty}^{\infty} e^{tx} f(x)dx \quad (21)$$

Using the nth ordinary moment of X in equation (20), the moment generating function of a random variable X for the SISD can be derived based on power series expansion as follows:

$$M_X(t) = E\left[\mathrm{e}^{tx}\right] = E\left[\sum_{n=0}^{\infty}\frac{(tx)^n}{n!}\right] = \sum_{n=0}^{\infty}\frac{t^n}{n!}\int_0^{\infty}x^n f(x)dx = \sum_{n=0}^{\infty}\frac{t^n}{n!}E\left(X^n\right) \quad (22)$$

Substituting the result from equation (20) into equation (22) and simplifying, the moment generating function of the SISD is obtained as:

$$M_X(t) = E\left[\mathrm{e}^{tx}\right] = \sum_{n=0}^{\infty}\frac{t^n}{n!}\psi_{klmr}\left[\frac{\eta\Gamma(m+n+r+1)}{\left[\eta(l+1)\right]^{m+n+r+1}} + \frac{\Gamma(m+n+r+3)}{\left[\eta(l+1)\right]^{m+n+r+3}}\right] \quad (23)$$

where $\psi_{klmr} = \sum_{k,l,m,r=0}^{\infty}\frac{(-1)^{k+l}(\pi)^{2k+1}\eta^{m+r+3}2^{m-r}}{(2k)!(2)^{2k+1}\left(\eta^3+2\right)^{m+1}}\binom{2k}{l}\binom{l}{m}\binom{m}{r}$ is a constant.

3.3 Characteristics Function

A representation for the characteristics function is given by

$$\phi_x(t) = E\left(e^{itx}\right) = \int_0^{\infty} e^{itx} f(x)dx \quad (24)$$

UNDER PEER REVIEW

The characteristics function of the SISD can be obtained based on the nth ordinary moment using power series expansion as follows:

$$\phi_X(t) = E\left[e^{itx}\right] = E\left[\sum_{n=0}^{\infty} \frac{(itx)^n}{n!}\right] = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \int_0^{\infty} x^n f(x)\,dx = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} E\left(X^n\right) \quad (25)$$

Again, substituting for $E\left(X^n\right)$ in equation (25) and simplifying, the characteristic function of the SISD is determined as:

$$\phi_X(t) = E\left[e^{itx}\right] = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \psi_{klmr} \left[\frac{\eta\Gamma(m+n+r+1)}{\left[\eta(l+1)\right]^{m+n+r+1}} + \frac{\Gamma(m+n+r+3)}{\left[\eta(l+1)\right]^{m+n+r+3}}\right] \quad (26)$$

where $\psi_{klmr} = \sum_{k,l,m,r=0}^{\infty} \frac{(-1)^{k+l}(\pi)^{2k+1}\eta^{m+r+3}2^{m-r}}{(2k)!(2)^{2k+1}\left(\eta^3+2\right)^{m+1}} \binom{2k}{l}\binom{l}{m}\binom{m}{r}$ is a constant.

3.4 Reliability analysis of the SISD.

A derivation and study of the survival function and the hazard rate function is presented in this section.

The Survival function describes the likelihood that a system or an individual will not fail after a given time. Mathematically, the survival function is given by:

$$S(x) = 1 - F(x) \quad (27)$$

Applying the cdf of the SISD in (27), the survival function for the SISD is obtained as:

$$S(x) = 1 - \sin\left[\frac{\pi}{2}\left(1 - \left[1 + \frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right] \quad (28)$$

A hazard function is the probability that a component will fail or die for an interval of time. The hazard function is defined as;

$$h(x) = \frac{f(x)}{1-F(x)} = \frac{f(x)}{S(x)} \quad (29)$$

Meanwhile, the expression for the hazard rate of the SISD is given by:

$$h(x) = \frac{\pi\eta^3\left(\eta + x^2\right)e^{-\eta x}\cos\left[\frac{\pi}{2}\left(1 - \left[1 + \frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right]}{2\left(\eta^3+2\right)\left(1 - \sin\left[\frac{\pi}{2}\left(1 - \left[1 + \frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right]\right)} \quad (30)$$

The plot of the survival and hazard functions for some chosen parameter values are presented in Figure **2** as follows below:

UNDER PEER REVIEW

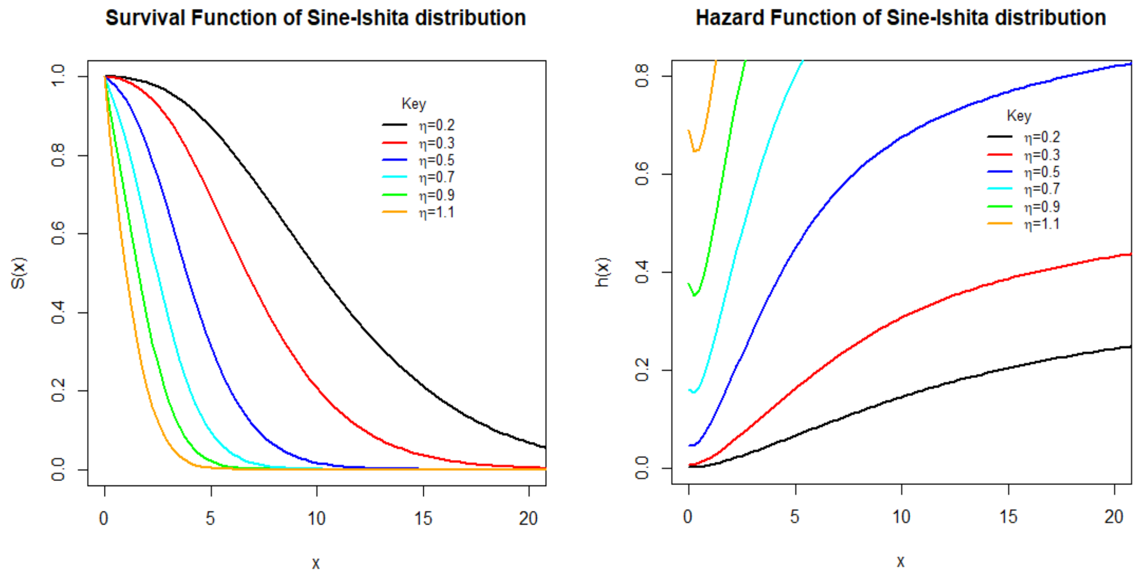

Figure 2: The Survival Function and Hazard Function of SISD.

The plot in Figure **2** shows that the probability of survival is always sure at an initial time or early age and it decreases as time increases up to zero (0) at infinity.
The figure above revealed that the SISD has an increasing failure rate which implies that the probability of failure for any random variable following a SISD increases as time increases, that is, the probability of failure or death increases with age.

3.5 Order Statistics

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from the SISD and let $X_{1:n}, X_{2:n}, \ldots, X_{i:n}$ denote the corresponding order statistic obtained from this same sample. The pdf, $f_{i:n}(x)$ of the i^{th} order statistic can be obtained by;

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} \sum_{k=0}^{\infty} (-1)^k \binom{n-i}{k} f(x) \left[F(x)\right]^{k+i-1} \qquad (31)$$

Using (5) and (6), the pdf of the i^{th} order statistics $X_{i:n}$, can be expressed from (31) as;

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} \sum_{k=0}^{\infty} (-1)^k \binom{n-i}{k} \frac{\pi\eta^3\left(\eta + x^2\right)}{2\left(\eta^3+2\right)e^{\eta x}} \cos\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right] \left[\sin\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right]\right]^{k+i-1} \qquad (32)$$

UNDER PEER REVIEW

Hence, the *pdf* of the minimum order statistic $X_{(1)}$ and maximum order statistic $X_{(n)}$ of the SISD are respectively given by;

$$f_{1:n}(x)=\frac{n!}{(n-1)!}\sum_{k=0}^{\infty}(-1)^k\binom{n-1}{k}\frac{\pi\eta^3\left(\eta+x^2\right)}{2\left(\eta^3+2\right)e^{\eta x}}\cos\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right]$$
$$\left[\sin\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right]\right]^k \qquad (33)$$

and

$$f_{n:n}(x)=n\sum_{k=0}^{\infty}(-1)^k\binom{n-i}{k}\frac{\pi\eta^3\left(\eta+x^2\right)}{2\left(\eta^3+2\right)e^{\eta x}}\cos\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right]$$
$$\left[\sin\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x(\eta x+2)}{\eta^3+2}\right]e^{-\eta x}\right)\right]\right]^{k+n-1} \qquad (34)$$

4. Maximum Likelihood Estimation of the Parameters of the SISD

Let $X_1, X_2, ..., X_n$ be a sample of size 'n' independently and identically distributed random variables from the SISD with unknown parameter, η defined previously.

The likelihood function of the SISD using the pdf in equation (6) is given by:

$$L\left(\underline{X}\mid\eta\right)=\left(\frac{\pi\eta^3}{2\left(\eta^3+2\right)}\right)^n e^{-\eta\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}\left(\eta+x_i^2\right)\prod_{i=1}^{n}\left(\cos\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x_i\left(\eta x_i+2\right)}{\eta^3+2}\right]e^{-\eta x_i}\right)\right]\right) \qquad (35)$$

Let the natural logarithm of the likelihood function be, $l=\log L\left(\underline{X}\mid\eta\right)$, therefore, taking the natural logarithm of the function equation (35) above gives:

$$l=n\log\left(\frac{\pi}{2}\right)+3n\log\eta-n\log\left(\eta^3+2\right)+\sum_{i=1}^{n}\log\left[\eta+x_i^2\right]-\eta\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}\log\left\{\cos\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x_i\left(\eta x_i+2\right)}{\eta^3+2}\right]e^{-\eta x_i}\right)\right]\right\}$$

(36)

Differentiating l partially with respect to η gives the following result:

$$\frac{\partial l}{\partial\eta}=\frac{3n}{\eta}-\frac{2n\eta^2}{\left(\eta^3+2\right)}+\sum_{i=1}^{n}\left(\eta+x_i^2\right)^{-1}-\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}\left\{\frac{\pi\eta^2 x_i e^{-\eta x_i}\tan\left[\frac{\pi}{2}\left(1-\left[1+\frac{\eta x_i\left(\eta x_i+2\right)}{\eta^3+2}\right]e^{-\eta x_i}\right)\right]}{\left(\eta^3+2\right)^2\left(\eta^4+\eta^3x_i^2+3\eta x_i+8\eta+2x_i^2\right)^{-1}}\right\}$$

(37)

UNDER PEER REVIEW

Equating (37) to zero (0) and solving for the solution of the non-linear system of equation produce the maximum likelihood estimate of parameter η. It is however difficult to solve this equation analytically and hence more appropriate to employ Newton-Raphson's iteration methods using computing software such as R in this case.

5 Applications

This section presents four applications of the proposed distribution to real life data. The MLEs of the model parameters are determined and some goodness-of-fit statistics for this distribution are compared with other competitive models. For all the datasets, the fits of the sine-Ishita distribution (SISD) are compared to sine Lomax-exponential distribution (SLED) sine Lomax distribution (SLD), Ishita distribution (ISD), Akash distribution (AKD) and the sine exponential distribution (SED).

The selection of the best model was done based on the value of the log-likelihood function evaluated at the MLEs (ℓ), Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC) and Hannan Quin Information Criterion (HQIC). Meanwhile, the smaller these statistics are, the better the fit of the distribution considered to be. The required computations are carried out using the R package "AdequacyModel" which is freely available from http://cran.r-project.org/web/packages/AdequacyModel/AdequacyModel.pdf.

Data set I: This dataset reflects the body fat percentage of 202 Australian athletes, it was extracted from Oguntunde *et al.*, (2018) and has also been used by Al-Noor and Hadi (2021) and Ieren *et al*. (2024).

Table 1: Summary Statistics for data set I

parameters	**n**	**Mini**	Q_1	**Median**	Q_3	**Mean**	**Maximum**	**Variance**	**Skewness**	**Kurtosis**
Values	202	5.630	8.545	11.650	18.080	13.507	35.520	38.31395	0.75955	-0.1733

Table 2: Maximum Likelihood Parameter Estimates for dataset I

Distribution	Parameter Estimates		
SISD	$\hat{\eta} = 0.16354061$	-	-
SLED	$\hat{\lambda} = 0.11571342$	$\hat{\alpha} = 3.732268$	$\hat{\beta} = 7.983072$
SLD	$\hat{\alpha} = 0.56743184$	$\hat{\beta} = 8.483704$	-
ISD	$\hat{\eta} = 0.22106492$	-	-
AKD	$\hat{\eta} = 0.21864352$	-	-
SED	$\hat{\lambda} = 0.04240049$	-	-

Table 3: The statistics ℓ, AIC, CAIC, BIC and HQIC for dataset I

UNDER PEER REVIEW

Distribution	$\hat{\ell}$	*AIC*	*CAIC*	*BIC*	*HQIC*
SISD	641.4218	1284.844	1284.864	1288.152	1286.182
SLED	734.1326	1474.265	1474.386	1484.190	1478.281
SLD	790.5107	1585.021	1585.082	1591.638	1587.699
ISD	647.8899	1297.780	1297.800	1301.088	1299.118
AKD	650.1304	1302.261	1302.281	1305.569	1303.599
SED	718.2667	1438.533	1438.553	1441.842	1439.872

The following figure presents a plot of estimated PDFs (densities) and CDFs of the fitted models to dataset I.

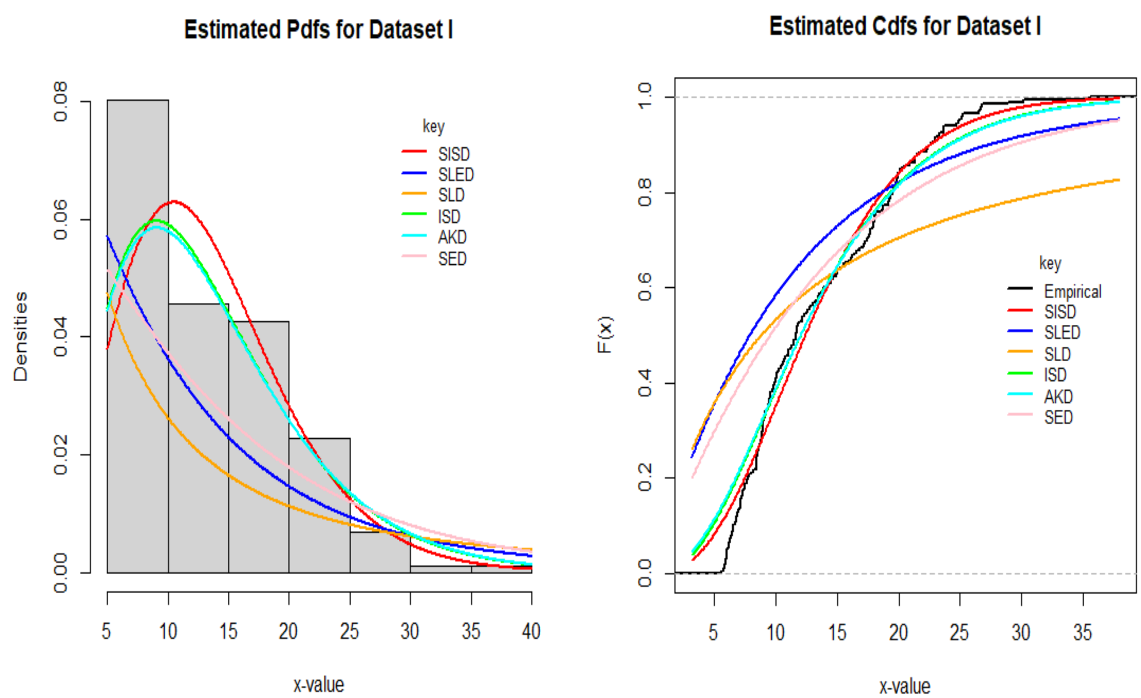

Figure 3: Plots of the estimated densities and CDFs of the fitted distributions to dataset I.

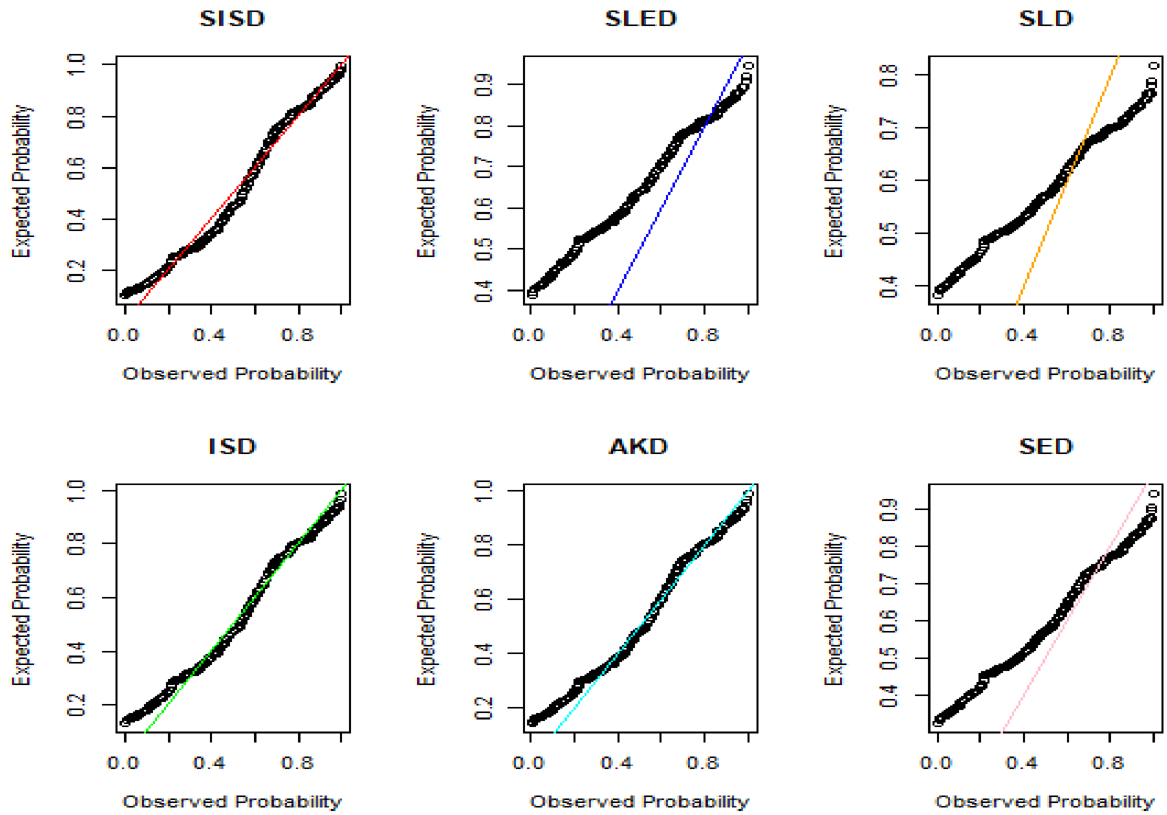

Figure 4: Probability plots for the fitted distributions to dataset I.

Data set II: This second dataset is on the sum of skin folds in 202 athletes collected at the Australian Institute of Sports. The data set was analyzed previously by Hosseini *et al*. (2018) and Reis et al. (2022). The summary of the data set is as follows:

Table 4: Summary Statistics for data set II

parameters	n	Mini	Q_1	Median	Q_3	Mean	Maximum	Variance	Skewness	Kurtosis
Values	202	28.00	43.85	58.6	90.35	69.02	200.8	1060.51	1.17467	1.36514

Table 5: Maximum Likelihood Parameter Estimates for dataset II

Distribution	Parameter Estimates		
SISD	$\hat{\eta} = 0.033578649$	-	-
SLED	$\hat{\lambda} = 0.019797070$	$\hat{\alpha} = 2.541371$	$\hat{\beta} = 5.955829$
SLD	$\hat{\alpha} = 0.232531300$	$\hat{\beta} = 7.160846$	-
ISD	$\hat{\eta} = 0.044254857$	-	-
AKD	$\hat{\eta} = 0.043657840$	-	-
SED	$\hat{\lambda} = 0.008408739$	-	-

UNDER PEER REVIEW

Table 6: The statistics ℓ, AIC, CAIC, BIC and HQIC for dataset II

Distribution	$\hat{\ell}$	*AIC*	*CAIC*	*BIC*	*HQIC*
SISD	971.0088	1944.018	1944.038	1947.326	1945.356
SLED	1064.5279	2135.056	2135.177	2144.981	2139.071
SLD	1215.7041	2435.408	2435.469	2442.025	2438.085
ISD	976.1181	1954.236	1954.256	1957.545	1955.575
AKD	976.1390	1954.278	1954.298	1957.586	1955.617
SED	1047.7759	2097.552	2097.572	2100.860	2098.890

The following figure presents a plot of estimated PDFs (densities) and CDFs of the fitted models to dataset II.

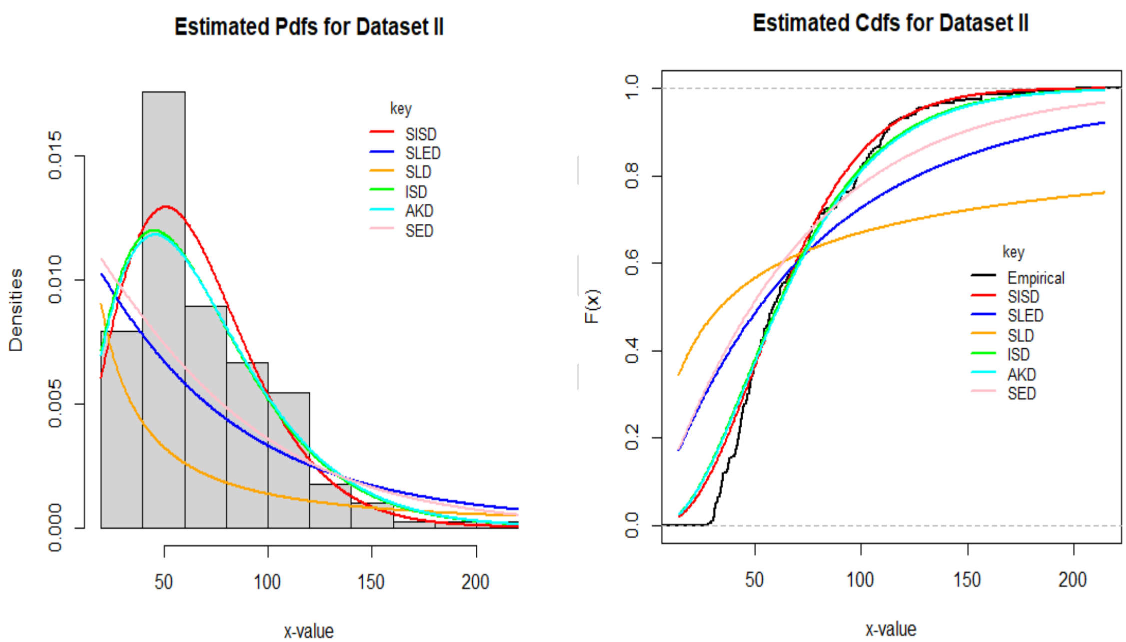

Figure 5: Plots of the estimated densities and CDFs of the fitted distributions to dataset II.

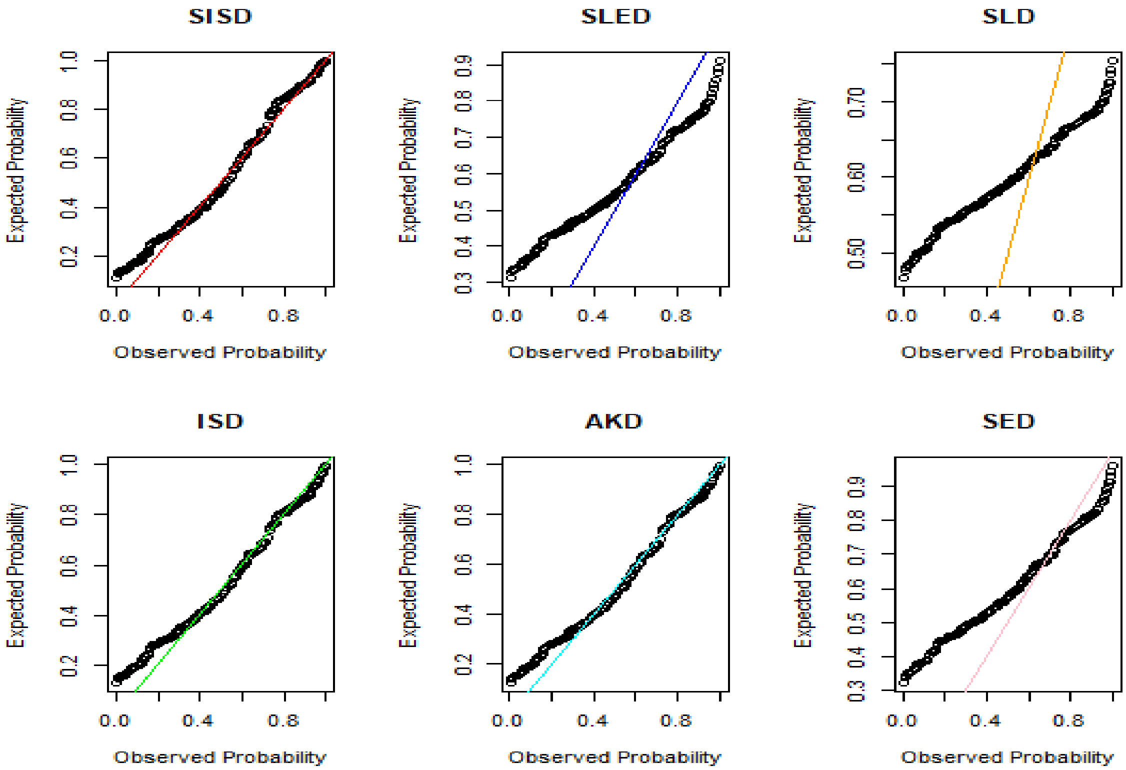

Figure 6: Probability plots for the fitted distributions to dataset II.

Data set III: The fourth data set has been used by Xu *et al*. (2003), Cordeiro *et al*. (2019) and Reis *et al*. (2022). It consists of the time-to-failure of a particular type of engine (turbocharger data). The summary statistics is given as follows:

Table 7: Summary Statistics for data set III

parameters	n	Mini	Q_1	Median	Q_3	Mean	Maximum	Variance	Skewness	Kurtosis
Values	40	1.60	5.075	6.5	7.825	6.255	9.00	3.83587	-0.65791	-0.36311

Table 8: Maximum Likelihood Parameter Estimates for dataset III

Distribution	Parameter Estimates		
SISD	$\hat{\eta}=0.35047881$	-	-
SLED	$\hat{\lambda}=0.07642464$	$\hat{\alpha}=4.869435$	$\hat{\beta}=3.996847$
SLD	$\hat{\alpha}=1.08964058$	$\hat{\beta}=9.546098$	-
ISD	$\hat{\eta}=0.46050280$	-	-
AKD	$\hat{\eta}=0.45007719$	-	-
SED	$\hat{\lambda}=0.09208243$	-	-

Table 9: The statistics ℓ, AIC, CAIC, BIC and HQIC for dataset III

Distribution	$\hat{\ell}$	*AIC*	*CAIC*	*BIC*	*HQIC*
SISD	91.86304	185.7261	185.8313	187.4150	186.3367
SLED	113.15717	232.3143	232.9810	237.3810	234.1463
SLD	119.77697	243.5539	243.8783	246.9317	244.7752
ISD	95.70210	193.4042	193.5095	195.0931	194.0149
AKD	96.88118	195.7624	195.8676	197.4512	196.3730
SED	111.11235	224.2247	224.3300	225.9136	224.8353

The following figure presents a plot of estimated PDFs (densities) and CDFs of the fitted models to dataset III.

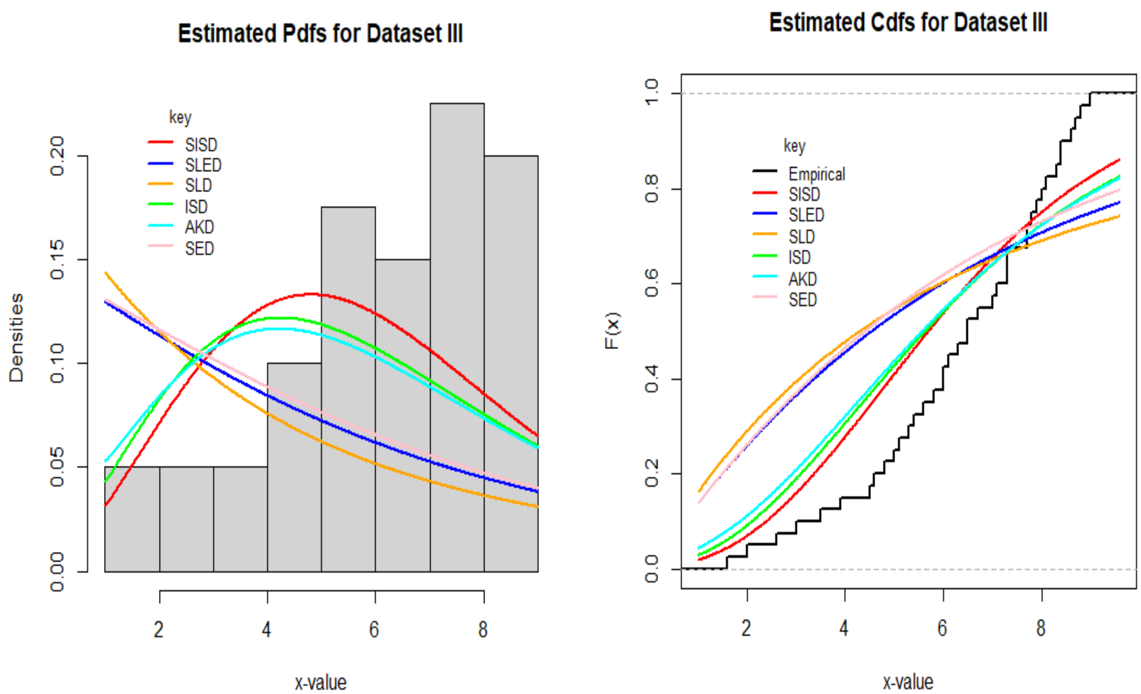

Figure 7: Plots of the estimated densities and CDFs of the fitted distributions to dataset III.

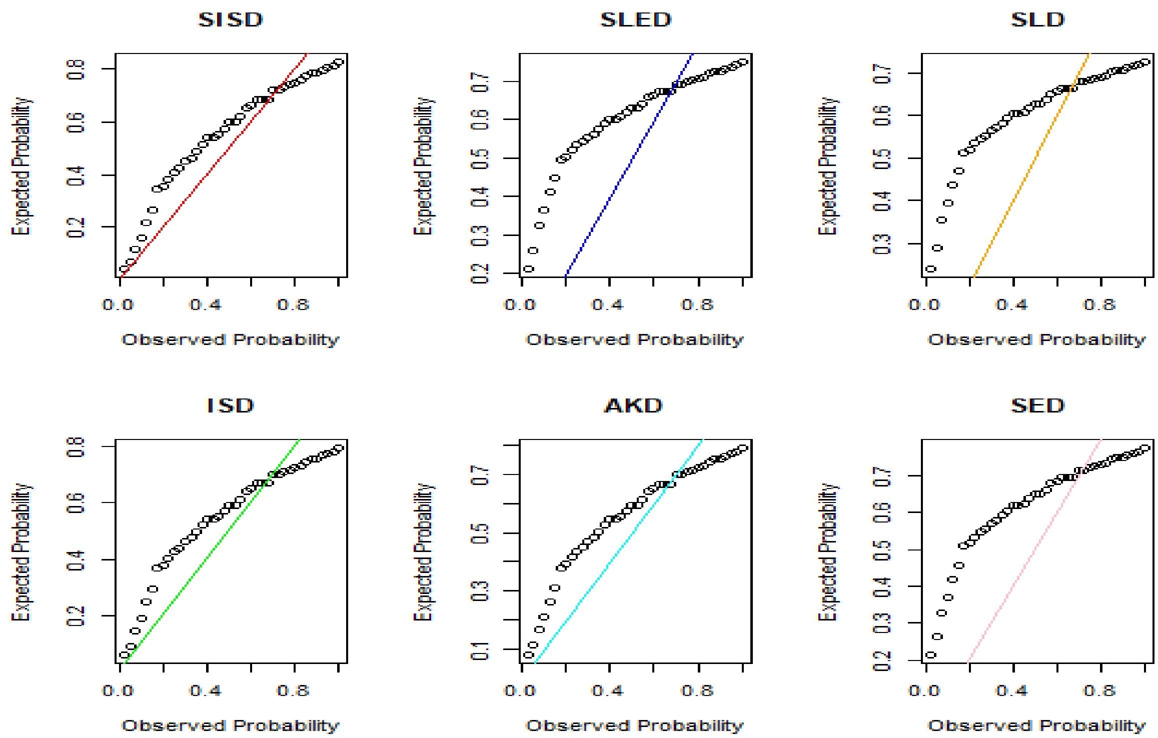

Figure 8: Probability plots for the fitted distributions to dataset III.

Data set IV: The third set of data on failure times for a particular model of windshield are given in Murthy *et al*. (2004) and was recently studied by Ramos *et al*. (2013). The summary statistics for the data is given as follows:

Table 10: Summary Statistics for data set IV

parameters	n	Mini	Q_1	Median	Q_3	Mean	Maximum	Variance	Skewness	Kurtosis
Values	84	0.040	1.839	2.3545	3.393	2.55745	4.663	1.25177	0.09949	-0.65232

Table 11: Maximum Likelihood Parameter Estimates for dataset IV

Distribution	Parameter Estimates		
SISD	$\hat{\eta}=0.7251658$	-	-
SLED	$\hat{\lambda}=0.1653491$	$\hat{\alpha}=6.133407$	$\hat{\beta}=4.330152$
SLD	$\hat{\alpha}=2.3991189$	$\hat{\beta}=9.848023$	-
ISD	$\hat{\eta}=0.9194357$	-	-
AKD	$\hat{\eta}=0.9351114$	-	-
SED	$\hat{\lambda}=0.2230054$	-	-

Table 12: The statistics ℓ, AIC, CAIC, BIC and HQIC for dataset IV

Distribution	$\hat{\ell}$	*AIC*	*CAIC*	*BIC*	*HQIC*
SISD	136.7063	275.4125	275.4613	277.8433	276.3897
SLED	161.9884	329.9767	330.2767	337.2692	332.9082
SLD	166.5233	337.0466	337.1947	341.9082	339.0009
ISD	145.0972	292.1943	292.2431	294.6251	293.1715
AKD	145.7071	293.4142	293.4629	295.8450	294.3913
SED	158.9431	319.8863	319.9351	322.3171	320.8634

The following figure presents a plot of estimated PDFs (densities) and CDFs of the fitted models to dataset IV.

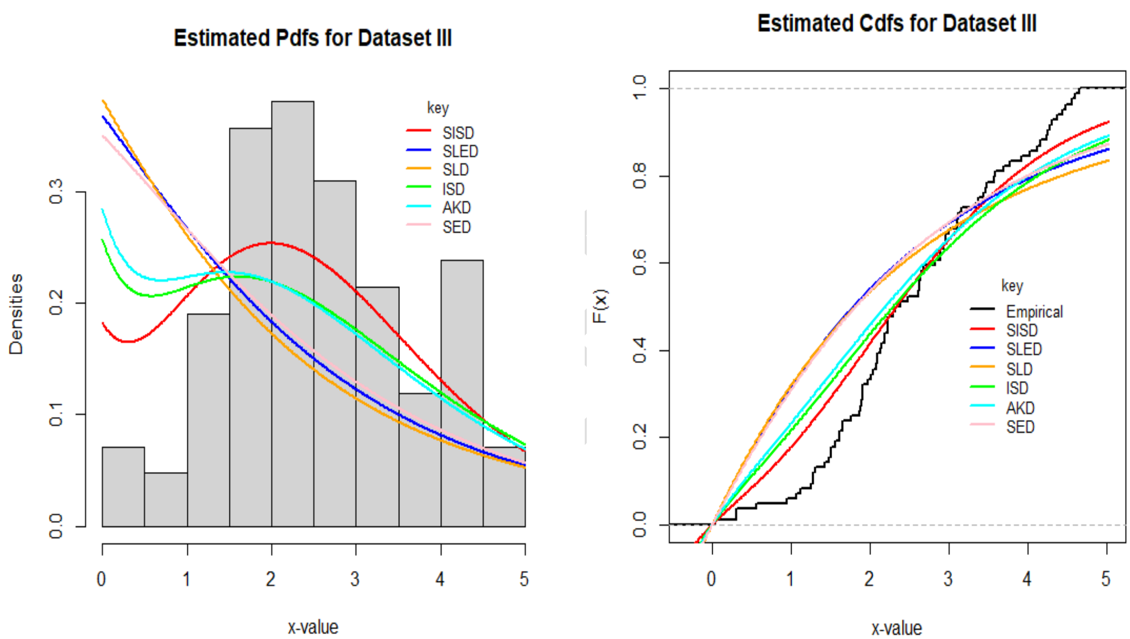

Figure 9: Plots of the estimated densities and CDFs of the fitted distributions to dataset IV.

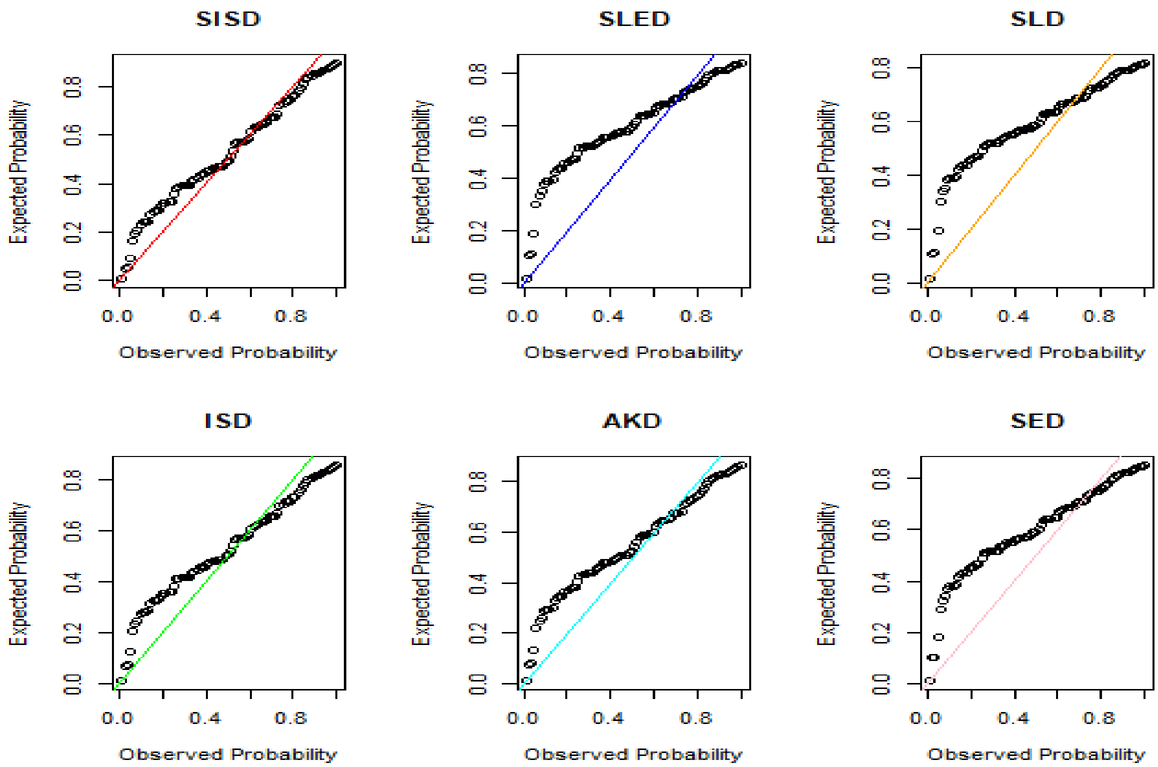

Figure 10: Probability plots for the fitted distributions to dataset IV.

Tables 2, 5, 8 and 11 present the maximum likelihood parameter estimates of the fitted SISD and the other five fitted distributions (SLED, SLD, ISD, AKD and SED) based on dataset I, dataset II, dataset III and dataset IV respectively and Tables 3, 6, 9 and 12 list the values of $\hat{\ell}$, AIC, CAIC, BIC and HQIC for the fitted SISD and the other five distributions based on dataset I, dataset II, dataset III and dataset IV respectively. The plots of the fitted densities and CDFs of the SISD and the other five distributions for dataset I, dataset II, dataset III and dataset IV are also shown in Figures 3, 5, 7 and 9 respectively and the probability plots of the fitted SISD and the other five distributions for dataset I, dataset II, dataset III and dataset IV are shown in figures 4, 6, 8 and 10 respectively.

The values of $\hat{\ell}$, AIC, CAIC, BIC and HQIC in Tables 3, 6, 9 and 12 for dataset I, dataset II, dataset III and dataset IV respectively are lower for the proposed SISD compared to the other fitted five distributions. This result indicates that the SISD fits the four datasets better than the other five fitted distributions. This result which is based on the values of $\hat{\ell}$, AIC, CAIC, BIC and HQIC also agrees with the plots of the estimated densities and CDFs of the fitted distributions shown in Figures 3, 5, 7 and 9 as well as the probability plots in Figures 4, 6, 8 and 10 for dataset I, dataset II, dataset III and dataset IV respectively.

This overall performance of the SISD is an indication that the Sine-G family of distributions is effective for modifying continuous distributions as previously reported by other researchers (Joel et al., 2024; Isa *et al*., 2022a; Mustapha et al., 2023; Tomy *et al*., 2021; Isa *et al*., 2022b).

6. Summary and Conclusion

This article proposed a new extension of the Ishita distribution called "Sine Ishita distribution". The article discussed some properties of the proposed distribution. The maximum likelihood method was used to estimate the model parameter. Some plots of the pdf of the distribution generated with arbitrary parameter values show that it is skewed and flexible and that its shape depends on the values of the parameters. Also, the plots of the survival function show that it is monotone decreasing, where the probability of survival decreases over time. Also, the plot of the hazard rate of the distribution shows that it is increasing for all parameter values and this shape is proof that the proposed model would be appropriate for analyzing events whose failure rate increases with time. The proposed model was fitted to four real life datasets to illustrate its capability over other existing distributions. The results of the analysis show that the proposed model, Sine-Ishita distribution has a better fit to the datasets than the other five considered models for all the datasets used. This implies that the proposed SISD and its properties will be useful in several areas of statistical theory and applications due to its better performance.

UNDER PEER REVIEW

REFERENCES

1. Abdel-Hady, D. H. (2013). Bivariate Generalized Rayleigh Distribution. *Journal of Applied Sciences Research, 9* (9): 5403-5411
2. Abdulkadir, S. S., Joel, J. & Ieren, T. G. (2020). Statistical Properties of Lomax-Inverse Exponential Distribution and Applications to Real Life Data. FUDMA Journal of Sciences, 4(2): 680-694.
3. Abdullahi, J., Abdullahi, U. K., Ieren, T. G., Kuhe, D. A. and Umar, A. A. (2018). On the properties and applications of transmuted odd generalized exponential-exponential distribution. *Asian Journal of Probability and Statistics*, 1(4):1-14. DOI: 10.9734/AJPAS/2018/44073
4. Afify, M. Z., Yousof, H. M., Cordeiro, G. M., Ortega, E. M. M. and Nofal, Z. M. (2016).The Weibull-Frechet Distribution and Its Applications. *Journal ofApplied Statistics*.1-22.
5. Alizadeh, M., Cordeiro, G. M., Pinho, L. G. B. and Ghosh, I. (2017). The Gompertz-G family of distributions. *Journal of Statistical Theory and Practice*, 11(1), 179–207, https://doi.org/10.1080/15598608.2016.1267668
6. Al-Noor, N. H. and Hadi, H. H. (2021). Properties and Applications of Truncated Exponential Marshall Olkin Weibull Distribution, *Journal of Physics: Conference Series*, 1879 032024, doi:10.1088/1742-6596/1879/3/032024
7. Alzaatreh, A., Famoye, F. and Lee, C. (2013). A new method for generating families of continuous distributions. *Metron*, 71, 63–79. https://doi.org/10.1007/s40300-013-0007-y
8. Alzaghal, A., Famoye, F. and Lee, C. (2013). Exponentiated T-X family of distributions with some applications. *International Journal of Probability and Statistics*, 2, 31–49. https://doi.org/10.5539/ijsp.v2n3p31
9. Anzagra, L, Sarpong, S. and Nasiru, S. (2022). Odd Chen-G family of distributions, *Ann. Data Sci.* 9, 369–391.

10. Benchiha, S, Sapkota, L. P., Al Mutairi, A., Kumar, V., Khashab, R. H., Gemeay, A. M., Elgarhy, M. and Nassr, S. G. (2023). A new sine family of generalized distributions: Statistical inference with applications. *Math. Comput. Appl*. 28, 83.
11. Bourguignon, M., Silva, R. B. and Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. *Journal of Data Science*, 12: 53-68.
12. Cakmakyapan, S. and Ozel, G. (2016). The Lindley Family of Distributions: Properties and Applications. *Hacettepe Journal of Mathematics and Statistics*, 46, 1-27.
13. Cordeiro, G. M., Afify, A. Z., Ortega, E. M. M., Suzuki, A. K. and Mead, M. E. (2019). The odd Lomax generator of distributions: Properties, estimation and applications. Journal of Computational and Applied Mathematics, 347, 222–237. https://doi.org/10.1016/j.cam.2018.08.008
14. Cordeiro, G. M., Ortega, E. M. M., Popovic, B. V and Pescim, R. R. (2014). The Lomax generator of distributions: Properties, minification process and regression model. *Applied Mathematics and Computation*, 247:465-486
15. Fayomi, A., Almetwally, E. M. and Qura, M. E. (2023). A novel bivariate Lomax-G family of distributions: Properties, inference, and applications to environmental, medical, and computer science data. *AIMS Math*. 8, 17539–17584.
16. Gabanakgosi, M. and Oluyede, B. (2023). The Topp-Leone type II exponentiated half logistic-G family of distributions with applications," Int. J. Math. Oper. Res. 25, 85–117.
17. Ghitany, M., Al-Mutairi, D., Balakrishnan, N. and Al-Enezi, L. (2013). Power Lindley distribution and associated inference. Comput. Stat. Data Anal. 64, 20–33.
18. Gomes-Silva, F., Percontini, A., De Brito, E., Ramos, M. W., Venancio, R. and Cordeiro, G. M. (2017). The Odd Lindley-G Family of Distributions. *Austrian Journal of Statistics*, 46, 65-87. https://doi.org/10.17713/ajs.v46i1.222
19. Gupta, V., Bhatt, M. and Gupta, J. (2015). The Lomax-Frechet distribution. *Journal of Rajasthan Academy of Physical Sciences*, 14(1): 25-43
20. Hosseini, B., Afshari, M. and Alizadeh, M. (2018). The generalized odd gamma-G family of distributions: Properties and applications. *Austrian Journal of Statistics*, 47, 69–89.
21. Ieren, T. G. & Abdullahi, J. (2020). Properties and Applications of a Two-Parameter Inverse Exponential Distribution with a Decreasing Failure Rate. Pakistan Journal of Statistics, 36(3): 183-206.
22. Ieren, T. G. and Balogun, O, S. (2021). Exponential-Lindley Distribution: Theory and Application to Bladder Cancer Data. *Journal of Applied Probability and Statistics*, 16(2): 129-146.
23. Ieren, T. G. and Kuhe, A. D. (2018). On the Properties and Applications of Lomax-Exponential Distribution. *Asian Journal of Probability and Statistics*, 1(4), 1-13. DOI: 10.9734/AJPAS/2018/42546
24. Ieren, T. G., Abdulkadir, S. S., Okolo, A. & Jibasen, D. (2024). A New Fréchet-G Family of Continuous Probability Distributions: Special Models, Properties, Simulation and Applications. *Journal of the Royal Statistical Society Nigeria Group (JRSS-NIG Group)*, 1(1), 46-71.

UNDER PEER REVIEW

25. Isa A. M., Ali B. A. and Zannah U. (2022a). Sine Burr XII Distribution: Properties and Application to Real Data Sets. Arid Zone Journal of Basic and Applied Sciences, 1(3): 48-58.
26. Isa, A. M., Sule O. B., Buhari A. A., Akeem A. A. and Ibrahim I. I. (2022b). Sine-Exponential Distribution: Its Mathematical Properties and Application to Real Dataset. UMYU Scientifica, 1(1): 127-131.
27. Joel, J., Yakura, B. S., Aniah-Betiang, E. I., Iseyemi, S. O. and Ieren, T. G. (2024). A Sine Lomax-Exponential Distribution: Its Properties, Simulation and Applications to Survival Data. *African Journal of Mathematics and Statistics Studies*, 7(4), 296-319. DOI:10.52589/AJMSS-IHSYZU29
28. Khan, M. S., King, R. and Hudson, I. L. (2016). Transmuted kumaraswamy distribution. *Statistics in Transition*, 17(2): 183-210
29. 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.
30. Mohammad, S. (2024). X-exponential-G Family of Distributions With Applications, International Journal of Statistics and Probability; Vol. 13(1), 40-54. https://doi.org/10.5539/ijsp.v13n1p40
31. Murthy, D. P., Xie, M. and Jiang, R. (2004). Weibull Models; John Wiley & Sons: Hoboken, NJ, USA.
32. Mustapha, B. A., Isa, A, M., Bashiru, S. O. and Itopa, I. I. (2023). Sine-Lomax Distribution: Properties and Applications to Real Data Sets. FUDMA Journal of Sciences (FJS), 7(4): 60-66.
33. Oguntunde, P. E., Khaleel, M. A., Adejumo, A. O., Okagbue, H. I., Opanuga, A. A. and Owolabi, F. O. (2018). The Gompertz inverse exponential distribution with applications *Cogent Mathematics and Statistics*, 5(1), 1-11. https://doi.org/10.1080/25742558.2019.1568662.
34. Prataviera F. (2022). Reparameterized flexible Weibull distribution with some applications. American Journal of Mathematical and Management Sciences, 41(3), 259–277. https://doi.org/10.1080/01966324.2021.1957731
35. Ramos, M. W. A., Marinho, P. R. D., da Silva, R. V. and Cordeiro, G. M. (2013). The exponentiated Lomax Poisson distribution with an application to lifetime data. Adv. Appl. Stat., 34, 107.
36. Rao G. S., & Aslam M. (2021). Inspection plan for COVID-19 patients for Weibull distribution using repetitive sampling under indeterminacy. BMC Medical Research Methodology, 21(1), 1–15. https://doi.org/10.1186/s12874-021-01387-7 PMID: 34696713
37. Reis, L. D. R., Cordeiro, G. M., Lima, M. d.C. S. (2022). The Stacy-G Class: A New Family of Distributions with Regression Modeling and Applications to Survival Real Data. Stats, 5, 215–257. https://doi.org/10.3390/stats5010015
38. Reynolds W. D., Drury C. F., Phillips L. A., Yang X., & Agomoh I. V. (2021). An adapted Weibull function for agricultural applications. Canadian Journal of Soil Science, 101(4), 680–702. https://doi.org/10.1139/cjss-2021-0046
39. Shanker, R. and Shukla, K. (2017). Ishita distribution and its Applications. Biometrics & Biostatistics International Journal, 5(2), 1-9.

UNDER PEER REVIEW

40. Shaw, W. and Buckley, I. (2007). The alchemy of probability distributions: beyond gram-charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map. Research Report. https://doi.org/10.48550/arXiv.0901.0434
41. Strzelecki P. (2021). Determination of fatigue life for low probability of failure for different stress levels using 3-parameter Weibull distribution. International Journal of Fatigue, 145, 106080. https://doi.org/ 10.1016/j.ijfatigue.2020.106080
42. Tahir, M. H., Zubair, M., Mansoor, M., Cordeiro, G. M. and Alizadeh, M. (2016). A New Weibull-G family of distributions. *Hacettepe Journal of Mathematics and Statistics*, 45(2), 629-647. https://doi.org/10.1186/s40488-014-0024-2
43. Tomy, L., Veena, G., Chesneau, C. (2021). The Sine Modified Lindley Distribution. Math. Comput. Appl. 2021, 26, 81. https://doi.org/10.3390/mca26040081
44. Tung Y. L., Ahmad Z., & Mahmoudi E. (2021). The Arcsine-X Family of Distributions with Applications to Financial Sciences. Computer System Science and Engineering, 39(3), 351–363. https://doi.org/10.32604/csse.2021.014270
45. Umar, S. A., Bukar, A. B., Makama, M. S. and Ieren, T. G. (2021). Some Results on the Transmuted Odd Lindley-Rayleigh Distribution. *Benin Journal of Statistics*, 4:135– 153.
46. Xu, K., Xie, M., Tang, L. C. and Ho, S. L. (2003). Application of neural networks in forecasting engine systems reliability. Applied Software Computation., 2, 255–268.
47. Yahaya, A. and Ieren, T. G. (2017). On Transmuted Weibull-Exponential Distribution: Its Properties and Applications. *Nigerian Journal of Scientific Research, 16(3), 289-297.*

UNDER PEER REVIEW