A New Sine-G Family of Distributions: Properties and Applications

Zafar Mahmood, Christophe Chesneau, Muhammad Hussain Tahir

CompAMa
www.compama.co.usb.ve

Bull. Comput. Appl. Math., Vol.7, 2019
ISSN 2244-8659

A New Sine-G Family of Distributions: Properties and Applications

Zafar Mahmood [1], *Christophe Chesneau* [2,3], *Muhammad Hussain Tahir* [4]

CompAMa Vol.7, pp.53-81, 2019 - Received April 1, 2019 - Accepted October 30, 2019

Abstract

This paper is devoted to the study of a new family of distributions based on a sine transformation. In some situations, we show that the new family provides a suitable alternative to the so-called sine-G family of distributions, with the same number of parameters. Among others, some of its significant mathematical properties are derived, including shapes of probability density and hazard rate functions, asymptotic, quantile function, useful expansions, moments and moment generating function. Then, a special member with two parameters, using the inverse Weibull distribution as baseline, is introduced and investigated in detail. By considering this new distribution as a statistical model, the parameters are estimated via the maximum likelihood method. A simulation study is carried out to assess the performance of the obtained estimators. The applications on two real data sets are explored, showing the ability of the proposed model to fit various type of data sets.

Keywords: Trigonometric distributions; Moments; Inverse Weibull distribution; Real life data sets.

Resumen

[1]Department of Statistics, Government Degree College, Khairpur Tamewali, Bahawalpur, Pakistan; and Department of Statistics, The Islamia University of Bahawalpur, Pakistan. (zafii68@gmail.com).

[2]Université de Caen, LMNO, Campus II, Science 3, 14032, Caen, France. (christophe.chesneau@gmail.com).

[3]Corresponding author.

[4]Department of Statistics, The Islamia University of Bahawalpur, Pakistan. (mtahir.stat@gmail.com).

Este artículo esta dedicado al estudio de una nueva familia de distribuciones basadas en transformaciones de la función seno. En algunas situaciones, mostramos que la nueva familia ofrece una alternativa adecuada a la llamada familia de distribuciones sine-G, con el mismo número de parámetros. Entre otros, se derivan algunas de sus propiedades matematicas significativas, que incluyen formas densidad de probabilidad y funciones de tasa de riesgo, función asintotica, funciones de cuantil, expansiones útiles, momentos y función generadora de momentos. Luego, se introduce e investiga en detalle un miembro especial con dos parámetros, que utiliza la distribución inversa Weibull como línea base. Al considerar esta nueva distribución como un modelo estadístico, los parámetros se estiman vía los métodos de máxima verosimilitud. Se lleva a cabo un estudio de simulación para evaluar la eficiencia de los estimadores obtenidos. Se exploran las aplicaciones en dos conjunto de datos reales, mostrando la abilidad del modelo propuesto para ajustar varios tipos de conjuntos de datos.

Palabras claves: Distribuciones trigonométricas; distribución inversa Weibull; conjuntos de datos de la vida real.

1 Introduction

A challenging work for the statistician is to construct flexible models for modeling various types of data. Generally, this allows to reveal new features of real life phenomena and provide advised predictions. In this regards, numerous families of distributions have been created via various techniques (differential equations, induction of location, scale, shape parameters, compounding, weighting ...), each giving flexible models, with specific properties. Among the most useful families of distributions, there are the Marshall-Olkin-G family introduced by [1], the exp-G family introduced by [2], the beta-G family introduced by [3], the gamma-G family developed by [4], the RB-G family introduced by [5], the TX-G family introduced by [6], the Weibull-G family developed by [7], the sine-G family developed by [8] and [9], the cos-G family developed by [10] and the generalized odd Gamma-G family introduced by [11].

This study proposes a new family of distributions following the spirit of the sine-G family by [8] and [9]. A brief description of the sine-G family is presented below. For a given cumulative distribution function (cdf) $G(x)$, the sine-G family is defined by the cdf given by

$$F(x) = \sin\left(\frac{\pi}{2}G(x)\right), \qquad x \in \mathbb{R}.$$

This family has multiple merits including the following ones. (i) It is simple (ii) $F(x)$ and $G(x)$ have the same number of parameters; there is no additional parameter, avoiding any problem of over parametrization (iii) Thanks to the trigonometric function, $F(x)$ has the ability to increase the flexibility of $G(x)$, providing new flexible models. Thus, it enriches the literature of new trigonometric distributions and models, which is welcome in view of the statistical impact of the few existing ones (as the sine distribution introduced by [12], the cosine distribution introduced by [13], the circular Cauchy distribution introduced by [14], the beta trigonometric distribution developed by [15], the sine square distribution introduced by [16] or the new trigonometric exponential distribution introduced by [17]). All these aspects are described in details in [8], [18], [9] and [10], with a special focus on the exponential cdf for $G(x)$ in [8] and a special focus on the inverse Weibull cdf for $G(x)$ in [18]. In these cases, complete data analyzes show that the sine-G model considerably increases the flexibility properties of the former model (corresponding to $G(x)$), showing better fits in comparison to some serious competitors.

These nice features are the motor of this study. Indeed, we introduce a new family of distributions characterized by a cdf based on the sine function, called the new sine-G family of distributions. We show that, in some situations, the new sine-G models provide an interesting alternative to the sine-G models, with possible different targets in terms of modeling. In a first part, we define the new family, with comments, discussions and comparisons with the former sine-G family. Then, we give a comprehensive account of its general mathematical properties, such as shapes of probability density and hazard rate functions, asymptotic, quantile function, useful expansions, moments and moment generating function. As in [9] for the sine-G family, we focus our attention on a special member based on the inverse Weibull cdf for $G(x)$, providing a new two-parameter distribution with heavy right skewed tail. For the corresponding model, we investigate the estimation of the parameters by the method of maximum likelihood, with a simulation study illustrating their convergence. Then, two practical data sets are analyzed, showing that the corresponding model has a better fit to the standard sine-G model, and other useful competitors. Several numerical and graphical reference tools are considered (AIC, BIC, A^*, W^*, K-S, P-P Plots...), all of them are favorable to the proposed model, attesting its interest for the statistical society.

The new sine-G family is presented in Section 2, along with some of its general mathematical properties. The special member using the inverse

Weibull distribution as baseline is investigated in Section 3. Then, it is considered as a statistical model in Section 4, with estimation of the parameters, simulation and applications. Some concluding remarks end the study in Section 5.

2 New sine-G family of distributions

2.1 Definition and motivations

Let us now define the new sine-G family of distributions (N-sine-G for short), with discussion. The N-sine-G family is characterized by the cdf given by

$$F(x) = \sin\left(\frac{\pi}{4}G(x)(G(x)+1)\right), \qquad x \in \mathbb{R}. \tag{1}$$

The idea behind the N-sine-G family of distributions is to put into the sine-G family of distributions a balanced compromise between two cdfs: the cdf $G(x)$ and the squaring cdf $[G(x)]^2$. Indeed, we can write $F(x)$ as

$$F(x) = \sin\left(\frac{\pi}{2}H(x)\right),$$

where $H(x)$ denotes a cdf defined as the uniform mixture of $G(x)$ and $[G(x)]^2$, i.e.,

$$H(x) = \frac{1}{2}G(x) + \frac{1}{2}[G(x)]^2.$$

The idea of doing a compromise between $G(x)$ and $[G(x)]^2$ also belongs to the so-called transmuted-G family of distributions introduced by [19]. Moreover, one can observe that $H(x)$ is a central member of the transmuted-G family; we can express $H(x)$ as: $H(x) = G(x) + \lambda G(x)(1-G(x))$ with $\lambda = -1/2$. Also, one can note that the N-sine-G corresponds to the sine-H family.

On the other side, thanks to the inequality $[G(x)]^2 \leq H(x) \leq G(x)$ and the fact that the sine function is increasing on $(0, \pi/2)$, we have an immediate stochastic ordering; the N-sine-G cdf can be bounded by two cdfs: one of the sine-G family and the other of the sine-G^2 family, as

$$\sin\left(\frac{\pi}{2}[G(x)]^2\right) \leq F(x) \leq \sin\left(\frac{\pi}{2}G(x)\right).$$

In this sense, the N-sine-G family provides a simple alternative to the sine-G family, with different target in terms of modelling. This practical aspect will

be developed in Section 4. Also, observe that no new parameter has been added, respecting the prime idea of the sine-G family.

Table 1 lists some members of the N-sine-G family of (continuous) distributions, selected for their potential usefulness. At this stage, let us mention that the one using the inverse Weibull cdf will be the object of all the attentions in Section 4.

Some secondary remarks are now formulated below. Using standard trigonometric formulas, we can express $F(x)$ as

$$F(x) = \sin\left(\frac{\pi}{4}G(x)\right)\cos\left(\frac{\pi}{4}[G(x)]^2\right) + \sin\left(\frac{\pi}{4}[G(x)]^2\right)\cos\left(\frac{\pi}{4}G(x)\right).$$

Using well-known inequalities involving the sine function, we have the following bounds for $F(x)$ in terms of power cdf of $G(x)$:

$$\frac{\pi}{2}[G(x)]^2 - \frac{\pi^3}{48}[G(x)]^6 \leq F(x) \leq \frac{\pi}{2}G(x).$$

2.2 General mathematical properties

The general properties of the N-sine-G family of distributions are described in this subsection.

2.2.1 Central functions

First of all, suppose that $G(x)$ is the cdf of a given univariate continuous distribution. Let $g(x)$ be a pdf corresponding to $G(x)$. Then, by differentiation (almost surely), the pdf corresponding to $F(x)$ is given by

$$f(x) = \frac{\pi}{4}g(x)\left[2G(x)+1\right]\cos\left(\frac{\pi}{4}G(x)(G(x)+1)\right), \quad x \in \mathbb{R}. \tag{2}$$

The hrf corresponding to $F(x)$ is given by

$$h(x) = \frac{f(x)}{1-F(x)} = \frac{(\pi/4)g(x)\left[2G(x)+1\right]\cos\left((\pi/4)G(x)(G(x)+1)\right)}{1-\sin\left((\pi/4)G(x)(G(x)+1)\right)}, \quad x \in \mathbb{R}. \tag{3}$$

Note that, using standard trigonometric formulas, we can express $f(x)$ as

$$\begin{aligned} f(x) =& \frac{\pi}{4}g(x)\left[2G(x)+1\right]\left[\cos\left(\frac{\pi}{4}G(x)\right)\cos\left(\frac{\pi}{4}[G(x)]^2\right)\right. \\ & \left. - \sin\left(\frac{\pi}{4}[G(x)]^2\right)\sin\left(\frac{\pi}{4}G(x)\right)\right]. \end{aligned}$$

2.2.2 Critical points and asymptotic

Some features on the variations of the functions $F(x)$, $f(x)$ and $h(x)$ are described below. The critical points x_0 of $f(x)$ satisfy $f'(x_0) = 0$, i.e.,

$$\cos\left(\frac{\pi}{4}G(x_0)(G(x_0)+1)\right)\left[\frac{\pi}{2}G(x_0)g'(x_0)+\frac{\pi}{4}g'(x_0)+\frac{\pi}{2}[g(x_0)]^2\right] - \left[\frac{\pi}{2}G(x_0)g(x_0)+\frac{\pi}{4}g(x_0)\right]^2\sin\left(\frac{\pi}{4}G(x_0)(G(x_0)+1)\right)=0. \quad (4)$$

The critical points x_* of $h(x)$ satisfy $h'(x_*) = 0$, i.e.,

$$4[2G(x_*)+1]g'(x_*)\cos\left(\frac{\pi}{4}G(x_*)(G(x_*)+1)\right)+[g(x_*)]^2\Bigg[4\pi[G(x_*)]^2 + 4\pi G(x_*)+8\cos\left(\frac{\pi}{4}G(x_*)(G(x_*)+1)\right)+\pi\Bigg]=0. \quad (5)$$

We can determine the nature of the critical point by determining the sign of the second derivative of the function taken at this point.

The asymptotes for $F(x)$, $f(x)$ and $h(x)$ are given below. When $G(x) \to 0$, using $\sin(y) \sim y$ when $y \to 0$, we have

$$F(x) \sim \frac{\pi}{4}G(x), \quad f(x) \sim \frac{\pi}{4}g(x), \quad h(x) \sim \frac{\pi}{4}g(x).$$

When $G(x) \to 1$, using $\sin(y) = \cos(\pi/2 - y) \sim 1 - (\pi/2 - y)^2/2$ and $\cos(y) = \sin(\pi/2 - y) \sim \pi/2 - y$, when $y \to \pi/2$, we have

$$F(x) \sim 1 - \frac{\pi^2}{8}\left(1-\frac{1}{2}G(x)(G(x)+1)\right)^2,$$

$$f(x) \sim 3\frac{\pi^2}{8}g(x)\left(1-\frac{1}{2}G(x)(G(x)+1)\right)$$

and

$$h(x) \sim \frac{3g(x)}{1-(1/2)G(x)(G(x)+1)}.$$

2.2.3 Quantile function

Let $Q_G(x)$ be the quantile function (qf) corresponding to $G(x)$, i.e., satisfying $G(Q_G(y)) = y$ for $y \in (0, 1)$. Then, the qf corresponding to $F(x)$ is given by

$$Q_F(y) = Q_G\left(\sqrt{\frac{4}{\pi}\arcsin(y)+\frac{1}{4}}-\frac{1}{2}\right), \quad y \in (0,1). \quad (6)$$

In particular, the median given by $Med_F = Q_F(0.5) = Q_G(x_*)$ with $x_* \approx 0.4574271$.

Some important practical applications of $Q_F(y)$ are the following. Let U be a random variable following the uniform distribution over $(0,1)$. Then, the random variable $X = Q_F(U)$ has the cdf $F(x)$ given by (1). For a given $G(x)$, this can be used to simulate different values distributed following the corresponding N-sine-G distribution. Moreover, the quantile density function of X can be obtained by differentiating $Q_F(y)$ with respect to y.

On the other side, the analysis of the variability of the skewness and kurtosis can be investigated based on quantile measures as the Bowley skewness (see [20]) and the Moors kurtosis (see [21]), respectively. The Bowley skewness based on quartiles is given by

$$B = \frac{Q_F(3/4) + Q_F(1/4) - 2Q_F(2/4)}{Q_F(3/4) - Q_F(1/4)}.$$

The Moors kurtosis based on octiles is given by

$$M = \frac{Q_F(3/8) - Q_F(1/8) + Q_F(7/8) - Q_F(5/8)}{Q_F(6/8) - Q_F(2/8)}.$$

2.2.4 Useful expansions

Proposition 1. *The cdf $F(x)$ given by (1) can be expressed as sums of power cdfs, i.e., of the form $[G(x)]^\theta$, where θ is an integer.*

Proof. Using the series expansions for the sine function, we have

$$\begin{aligned} F(x) &= \sin\left(\frac{\pi}{4}G(x)(G(x)+1)\right) \\ &= \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)!}\left(\frac{\pi}{4}\right)^{2k+1} [G(x)]^{2k+1}(G(x)+1)^{2k+1}. \end{aligned}$$

Then, the binomial formula gives

$$F(x) = \sum_{k=0}^{+\infty}\sum_{\ell=0}^{2k+1} a_{k,\ell}[G(x)]^{\ell+2k+1},$$

where

$$a_{k,\ell} = \frac{(-1)^k}{(2k+1)!}\left(\frac{\pi}{4}\right)^{2k+1} \frac{(2k+1)!}{\ell!(2k+1-\ell)!}. \tag{7}$$

This completes the proof of Proposition 1. □

It follows from Proposition 1 that, by differentiation under the sums, we can express the pdf as

$$f(x) = \sum_{k=0}^{+\infty} \sum_{\ell=0}^{2k+1} b_{k,\ell} [G(x)]^{\ell+2k} g(x), \tag{8}$$

where $b_{k,\ell} = (\ell + 2k + 1)a_{k,\ell}$. Moreover, the properties of the power cdfs of the form $[G(x)]^\theta$ can be used to determine transformation involving $f(x)$ as the integrals, and, a fortiori, statistical properties on X. The next subsection applies this result to express moments of various kind.

2.2.5 Moments

Let X be a random variable having the cdf $F(x)$. Then the rth moment of X is given by $\mu'_r = \mathbb{E}(X^r) = \int_{-\infty}^{+\infty} x^r f(x)dx$. Using the series expansions (8), assuming that the sum and integral terms exist, we obtain

$$\mu'_r = \sum_{k=0}^{+\infty} \sum_{\ell=0}^{2k+1} b_{k,\ell} \int_{-\infty}^{+\infty} x^r [G(x)]^{\ell+2k} g(x)dx. \tag{9}$$

Note that $\int_{-\infty}^{+\infty} x^r [G(x)]^{\ell+2k} g(x)dx = \int_0^1 x^{\ell+2k} [Q_G(x)]^r dx$. This integral has not necessarily a closed form. We can at least compute numerically by using a standard software (R, Matlab, Mathematica...).

From μ'_r, we can deduce the mean of X given by $\mathbb{E}(X) = \mu'_1$, the variance of X given by $\mathbb{V}(X) = \mu'_2 - (\mu'_1)^2$, the standard deviation of X given by $\sigma(X) = \sqrt{\mu'_2 - (\mu'_1)^2}$, the coefficient of variation given by $CV = \sigma(X)/\mu'_1$, the rth central moment of X given by

$$\mu_r = \mathbb{E}\left[(X - \mu'_1)^r\right] = \sum_{k=0}^{r} \frac{r!}{k!(r-k)!} (-1)^k (\mu'_1)^k \mu'_{r-k},$$

the coefficient of skewness given by $CS = \mu_3/\mu_2^{3/2}$, the coefficient of kurtosis given by $CK = \mu_4/\mu_2^2$, and the rth descending factorial moment of X is given by

$$\mu_{(r)} = \mathbb{E}\left[X(X-1)\dots(X-r+1)\right] = \sum_{k=0}^{r} s_{r,k} \mu'_k,$$

where $s_{r,k}$ denotes the Stirling number of the first kind.

The moment generating function of X is given by $M(t) = \mathbb{E}(e^{tX})$ with $t \leq 0$ (to ensure its existence, this domain of definition can be refined according to the definition $G(x)$). By assuming that the sum and integral terms exist, we obtain

$$M(t) = \sum_{k=0}^{+\infty} \sum_{\ell=0}^{2k+1} b_{k,\ell} \int_{-\infty}^{+\infty} e^{tx} [G(x)]^{\ell+2k} g(x) dx.$$

Note that $\int_{-\infty}^{+\infty} e^{tx} [G(x)]^{\ell+2k} g(x) dx = \int_0^1 x^{\ell+2k} e^{tQ_G(x)} dx$. Again, we can determine it numerically for a given $G(x)$. As usual, we have the following relation between the rth moments and the moment generating function: $\mu'_r = M^{(r)}(t) \mid_{t=0}$ for any integer r.

Other probabilistic can be express in a similar manner, as the characteristic function, the conditional moments and the mean deviations. See, for instance, the methodology of [11].

3 The N-sine-IW distribution

3.1 Presentation

We now present a special member of N-sine-G family of distributions with support on $(0, +\infty)$ using the cdf $G(x)$ of the inverse Weibull distribution given by $G(x) = e^{-(\beta/x)^\alpha}$, $\alpha, \beta, x > 0$ (see [22]). We thus aim to construct a new heavy right skew model for real life data, by increasing the flexibility of the former inverse Weibull distribution. Hereafter, this special member will be called the N-sine-IW(α, β) distribution. By using the cdf given by (1), the N-sine-IW(α, β) distribution is characterized by the cdf given by

$$F(x) = \sin\left(\frac{\pi}{4} e^{-(\beta/x)^\alpha} (e^{-(\beta/x)^\alpha} + 1)\right), \quad x > 0. \tag{10}$$

By using (2) and (3) with $g(x) = \alpha\beta^\alpha x^{-\alpha-1} e^{-(\beta/x)^\alpha}$, the corresponding pdf is given by

$$f(x) = \frac{\pi}{4} \alpha\beta^\alpha x^{-\alpha-1} e^{-(\beta/x)^\alpha} \left[2e^{-(\beta/x)^\alpha} + 1\right] \cos\left(\frac{\pi}{4} e^{-(\beta/x)^\alpha} (e^{-(\beta/x)^\alpha} + 1)\right), \quad x > 0, \tag{11}$$

and the corresponding hrf is given by

$$h(x) = \frac{(\pi/4)\alpha\beta^{\alpha}x^{-\alpha-1}e^{-(\beta/x)^{\alpha}}\left[2e^{-(\beta/x)^{\alpha}}+1\right]\cos\left((\pi/4)e^{-(\beta/x)^{\alpha}}(e^{-(\beta/x)^{\alpha}}+1)\right)}{1-\sin\left((\pi/4)e^{-(\beta/x)^{\alpha}}(e^{-(\beta/x)^{\alpha}}+1)\right)},$$

$x > 0$.

The critical points for $f(x)$ and $h(x)$ can be obtained by solving the equations (4) and (5), respectively. The asymptotes for $F(x)$, $f(x)$ and $h(x)$ are given below. When $x \to 0$, we have

$$F(x) \sim \frac{\pi}{4}e^{-(\beta/x)^{\alpha}} \to 0, \qquad f(x) \sim \frac{\pi}{4}\alpha\beta^{\alpha}x^{-\alpha-1}e^{-(\beta/x)^{\alpha}} \to 0$$

and

$$h(x) \sim \frac{\pi}{4}\alpha\beta^{\alpha}x^{-\alpha-1}e^{-(\beta/x)^{\alpha}} \to 0.$$

When $x \to +\infty$, we have

$$F(x) \sim 1 - 9\frac{\pi^2}{32}\beta^{2\alpha}x^{-2\alpha} \to 1, \quad f(x) \sim 9\frac{\pi^2}{16}\alpha\beta^{2\alpha}x^{-2\alpha-1} \to 0$$

and

$$h(x) \sim 2\alpha x^{-1} \to 0.$$

We can remark that $f(x)$ as a polynomial decay when $x \to +\infty$. Moreover, when $x \to +\infty$, the asymptote of $h(x)$ depends only on the parameter α.

Figure 1 shows the plots for $f(x)$ and $h(x)$ respectively, for selected values for α and β. Various forms of right skew tail curvatures are observed, attesting the ability of the N-sine-IW(α, β) to model a wide variety of life-time data sets having such form of histogram.

3.2 Quantile function

Let us remark that the qf corresponding to $G(x)$ is given by $Q_G(y) = \beta\left[-\log(y)\right]^{-1/\alpha}$, $y \in (0,1)$. Then, by virtue of (6), the qf of the N-sine-IW(α, β) distribution is given by

$$Q_F(y) = \beta\left[-\log\left(\sqrt{\frac{4}{\pi}\arcsin(y)+\frac{1}{4}}-\frac{1}{2}\right)\right]^{-1/\alpha}, \quad y \in (0,1).$$

Let U be a random variable following the uniform distribution over $(0,1)$. Then $X = Q_F(U)$ follows the N-sine-IW(α, β) distribution. As an immediate

consequence, data distributed following the N-sine-IW(α, β) distribution can be simulated. The median of the N-sine-IW(α, β) distribution is given by

$$Med_F = Q_F(0.5) = \beta \left[-\log \left(\sqrt{\frac{4}{\pi} \arcsin(0.5) + \frac{1}{4}} - \frac{1}{2} \right) \right]^{-1/\alpha}.$$

In a similar way, we can express the corresponding Bowley skewness and Moors kurtosis. Table 2 indicates the values of the first quartile, median, third quartile, Bowley skewness and Moors kurtosis of the N-sine-IW(α, β) distribution for selected values for α and β. We observe that the Bowley skewness and Moors kurtosis do not depend on the parameter β.

3.3 Moments

Let X be a random variable following the N-sine-IW(α, β) distribution. Then X has a rth moment if and only if $r \in (0, 2\alpha)$. Indeed, there is no problem for $x \to 0$ and for $x \to +\infty$, we have

$$x^r f(x) \sim 9\frac{\pi^2}{16}\alpha\beta^{2\alpha}x^{r-2\alpha-1},$$

and $\int_1^{+\infty} x^{r-2\alpha-1}dx$ exists as a Riemann integral if and only if $r \in (0, 2\alpha)$. For given values for r, α and β, the integral expression of μ'_r can be evaluated numerically. On the other hand, for $r \in (0, \alpha)$, the rth moment of X can be obtained by the formula given by (9), i.e.,

$$\mu'_r = \sum_{k=0}^{+\infty} \sum_{\ell=0}^{2k+1} b_{k,\ell} \int_{-\infty}^{+\infty} x^r [G(x)]^{\ell+2k} g(x)dx.$$

The integral terms can be expressed via gamma functions, as developed below. Let us consider the gamma function $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$ with $x > 0$. By the change of variable $y = (\ell + 2k + 1)(\beta/x)^\alpha$, we have

$$\begin{aligned}\int_{-\infty}^{+\infty} x^r [G(x)]^{\ell+2k} g(x)dx &= \alpha\beta^\alpha \int_0^{+\infty} x^{r-\alpha-1} e^{-(\ell+2k+1)(\beta/x)^\alpha} dx \\ &= \beta^r (\ell + 2k + 1)^{r/\alpha-1} \Gamma\left(1 - \frac{r}{\alpha}\right).\end{aligned}$$

By noticing that $(\ell+2k+1)^{r/\alpha-1}b_{k,\ell} = (\ell+2k+1)^{r/\alpha}a_{k,\ell}$, where $a_{k,\ell}$ is given by (7), we have

$$\mu'_r = \beta^r \Gamma\left(1-\frac{r}{\alpha}\right) \sum_{k=0}^{+\infty} \sum_{\ell=0}^{2k+1} a_{k,\ell}(\ell+2k+1)^{r/\alpha}.$$

In particular, the variance $\mathbb{V}(X)$ and the standard deviation $\sigma(X)$ can be determined, as well as the coefficients CV, CS and CK. Table 3 shows numerical values for for some moments of X with selected values for α and β. In particular, we see that the moment increases as β increase. Moreover, we see that the CV, CS and CK do not depend on the parameter β.

4 The N-sine-IW model

This section is devoted to the consideration of the N-sine-IW(α,β) distribution as statistical model. In particular, we show that the N-sine-IW model (NSIW for short) can outperform, in some senses, some existing models in the literature.

4.1 Maximum likelihood estimation

We propose to estimate the parameters α and β of the NSIW model by the maximum likelihood method. Let $x_1,\ldots,x_n$ be a sample of a random variable following the N-sine-IW(α,β) distribution. By using the pdf $f(x)$ given by (11), the likelihood function is defined by

$$L(\alpha,\beta) = \prod_{i=1}^{n} f(x_i) = \left(\frac{\pi}{4}\right)^n \alpha^n \beta^{n\alpha} \left(\prod_{i=1}^{n} x_i\right)^{-\alpha-1} e^{-\sum_{i=1}^{n}(\beta/x_i)^\alpha} \times$$
$$\prod_{i=1}^{n}\left[2e^{-(\beta/x_i)^\alpha}+1\right] \prod_{i=1}^{n} \cos\left(\frac{\pi}{4}e^{-(\beta/x_i)^\alpha}(e^{-(\beta/x_i)^\alpha}+1)\right).$$

The log-likelihood function is given by

$$\begin{aligned}
&\ell(\alpha,\beta) = \log L(\alpha,\beta) \\
&= n\log\left(\frac{\pi}{4}\right) + n\log(\alpha) + n\alpha\log(\beta) - (\alpha+1)\sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}\left(\frac{\beta}{x_i}\right)^{\alpha} \\
&+ \sum_{i=1}^{n}\log\left[2e^{-(\beta/x_i)^{\alpha}}+1\right] + \sum_{i=1}^{n}\log\left[\cos\left(\frac{\pi}{4}e^{-(\beta/x_i)^{\alpha}}(e^{-(\beta/x_i)^{\alpha}}+1)\right)\right].
\end{aligned}$$

The maximum likelihood estimates (MLEs) are given by the simultaneous solutions of the nonlinear equations: $\partial\ell(\alpha,\beta)/\partial\alpha = 0$ and $\partial\ell(\alpha,\beta)/\partial\beta = 0$ according to α and β, with

$$\begin{aligned}
&\frac{\partial}{\partial\alpha}\ell(\alpha,\beta) = n\left(\frac{1}{\alpha}+\log(\beta)\right) - \sum_{i=1}^{n}\log\left(x_i\right) - \sum_{i=1}^{n}\left(\frac{\beta}{x_i}\right)^{\alpha}\log\left(\frac{\beta}{x_i}\right) \\
&- 2\sum_{i=1}^{n}\frac{e^{-(\beta/x_i)^{\alpha}}(\beta/x_i)^{\alpha}\log\left(\beta/x_i\right)}{2e^{-(\beta/x_i)^{\alpha}}+1} \\
&+ \frac{\pi}{4}\sum_{i=1}^{n}e^{-(\beta/x_i)^{\alpha}}\left[2e^{-(\beta/x_i)^{\alpha}}+1\right]\left(\frac{\beta}{x_i}\right)^{\alpha}\log\left(\frac{\beta}{x_i}\right)\tan\left(\frac{\pi}{4}e^{-(\beta/x_i)^{\alpha}}(e^{-(\beta/x_i)^{\alpha}}+1)\right)
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\partial}{\partial\beta}\ell(\alpha,\beta) = \frac{\alpha n}{\beta} - \alpha\sum_{i=1}^{n}\frac{1}{x_i}\left(\frac{\beta}{x_i}\right)^{\alpha-1} - 2\alpha\sum_{i=1}^{n}\frac{e^{-(\beta/x_i)^{\alpha}}(\beta/x_i)^{\alpha-1}}{x_i\left(2e^{-(\beta/x_i)^{\alpha}}+1\right)} \\
&+ \frac{\pi}{4}\frac{\alpha}{\beta}\sum_{i=1}^{n}e^{-(\beta/x_i)^{\alpha}}\left[2e^{-(\beta/x_i)^{\alpha}}+1\right]\left(\frac{\beta}{x_i}\right)^{\alpha}\tan\left(\frac{\pi}{4}e^{-(\beta/x_i)^{\alpha}}(e^{-(\beta/x_i)^{\alpha}}+1)\right).
\end{aligned}$$

These equations can not be solved analytically. Numerical solutions exist by the use of iterative methods such as Newton-Raphson type algorithms. Under standard regularity conditions, it is well established that the MLEs are asymptotically unbiased and normal. This last property allows us to construct approximate confidence intervals (CI) and Likelihood ratio tests for the parameters. In particular, the CIS of α and β are of the form [L.bound, U.bound], where L.bound denotes the lower bound of the interval and U.bound the upper bound, both depending on the fixed level of the CI and the components of the estimated Fisher information matrix. See, for instance, [23].

4.2 Simulations

In this subsection, a Monte Carlo simulations study is performed to attest the convergence of the MLEs. This study is repeated 1000 times each with sample sizes $n = 50, 100, 200, 300, 500, 1000$ and parameter combinations: I: $\alpha = 0.5$ and $\beta = 1.25$, II: $\alpha = 1.0$ and $\beta = 1.5$, III: $\alpha = 3.0$ and $\beta = 5.5$ and IV: $\alpha = 5.5$ and $\beta = 7.25$. The R-software is used. The empirical Biases, mean squared errors (MSEs), coverage probabilities (CP) of the 95% two-sided CIs with L.bound and U.bound for the model NSIW parameters can be found in Table 4. In particular, with respect to the theory, we observe that the biases and MSEs decrease with increasing sample size. Also, in coherence with the theory, the CPs of the confidence intervals are quite close to the 95% nominal levels. Therefore, the MLEs and their asymptotic results can be used for estimating and constructing confidence intervals for the model parameters.

4.3 Applications

In this section, we presented the analysis of two practical data sets via different models, with a focus on the NSIW model.

4.3.1 Data sets

Data set 1. The first data set contains 72 survival times in days of guinea pigs, voluntary contaminated with different doses of tubercle bacilli. The source of this data set is [24]. The 72 values of this data set are listed below: 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.
Figure 2 (a) presents the histogram of data set 1, showing a heavy right tail, also indicated by the TTT plot in Figure 2 (b). The boxplot of data set 1 can be seen in Figure 3 (a) and the Q-Q plot in Figure 3 (b). From these graphics, we clearly see that the normal distribution is misappropriated, motivating the use of a model with heavy right tail, as the NSIW model for instance.

Data set 2. The second data set contains 23 numbers of million of revolutions before failure of a ball bearing. The source of this data sets is [25].

The 23 values of this data set are listed below: 17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 51.84, 51.96, 54.12, 55.56, 67.80, 68.44, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04, 173.40

The histogram of data set 2 is given in Figure 4 (a). We clearly see a heavy right tail, confirmed by the TTT plot in Figure 4 (b). The boxplot of data set 2 is presented in Figure 5 (a) and the Q-Q plot in Figure 5 (b). The normal model is clearly not the best for this data set; a model with heavy right tail is required, motivating the use of the NSIW model.

4.3.2 Analysis

In order to analyze data sets 1 and 2 and compare the fitted models, we compute some well-known measures of goodness-of-fit statistics: the log-likelihood function evaluated at the MLEs ($\hat{\ell}$), Akaike information criterion (AIC), Anderson-Darling (A^*), Cramér–von Mises (W^*), Kolmogorov-Smirnov (K-S) and P-values. The statistics A^* and W^* are described in details in [26]. The required computations are carried out using the R-software. The lower values of AIC, A^*, W^* , K-S statistics, and high P-values indicate the better fit. The considered models are the proposed NSIW model, sine inverse Weibull model (SIW) (see [8] and [9]), inverse Weibull model (IW) (see [22]), inverse Nadarajah-Haghighi model (INH) (see [27]), inverse exponential model (IED) (see [28]) and inverse Rayleigh model (IRD) (see [29]). Their cdfs are respectively given by

$$
\begin{aligned}
F_{NSIW}(x) &= \sin\left(\frac{\pi}{4}e^{-(\beta/x)^{\alpha}}(e^{-(\beta/x)^{\alpha}}+1)\right), \quad x,\alpha,\beta>0,\\
F_{SIW}(x) &= \sin\left(\frac{\pi}{2}e^{-(\beta/x)^{\alpha}}\right), \quad x,\alpha,\beta>0,\\
F_{IW}(x) &= e^{-(\beta/x)^{\alpha}}, \quad x,\alpha,\beta>0,\\
F_{INH}(x) &= e^{1-(1+\beta/x)^{\alpha}}, \quad x,\alpha,\beta>0,\\
F_{IED}(x) &= \mathrm{e}^{-\alpha/x}, \quad x,\alpha>0,\\
F_{IRD}(x) &= e^{-\alpha/x^2}, \quad x,\alpha>0.
\end{aligned}
$$

Table 5 lists the MLEs and their corresponding standard errors (in parentheses) of the model parameters for the data sets 1 and 2.

All the results from Table 6 indicate that the NSIW model provides the better fit as compared to other models.

Figures 6 and 7 represent all the estimated pdfs over the histogram of data sets 1 and 2, respectively. Figures 8 and 9 represent all the estimated cdfs with the empirical cdf of data sets 1 and 2, respectively. Figures 10 and 11 show the P-Plot for the estimated models for data sets 1 and 2, respectively. In all these figures, we observe a nice fit for the NSIW model. In particular, the NSIW model reveals to be visually favorable to the SIW model, which remains the best of the considered competitors. Last but not least, the NSIW model can present better goodness-of-fits to more sophisticated model, with three parameters or more. For instance, this is the case for the three parameter gamma inverse Weibull (GIW) model introduced by [30] satisfying AIC = 786.5 and BIC = 793.3 for data set 1 (see [30, Table 1]), and AIC = 232.5 and BIC = 235.9 for data set 2 (see [30, Table 2]).

5 Concluding remarks

In this paper, we introduce the N-sine-G family of distributions as an alternative to the sine-G family proposed by [8] and [9], including a two-parameter NSIW distribution with decreasing and upside-down bathtub hazard rates. We investigate several of its structural properties such as asymptotes and shapes, quantile function, linear representation in terms of exponentiated distributions, moments and moment generating function. The model parameters are estimated by the maximum likelihood method. A Monte Carlo simulation study is presented to verify the adequacy of the estimates. Then, two practical data sets are considered, as well as strong competitor models. According to several goodness-of-fit statistics, the proposed model provides consistently better fits than the others. We hope that the proposed family and its generated models will attract wider applications in several areas such as reliability engineering, insurance, hydrology, economics and survival analysis.

References

[1] A.N. Marshall and I. Olkin. A new method for adding a parameter to a family of distributions with applications to the exponential and weibull families. *Biometrica*, 84(3):641–652, 1997. https://www.jstor.org/stable/2337585.

[2] R.C. Gupta, P.I. Gupta, and R.D. Gupta. Modeling failure time data by lehmann alternatives. *Communications in statistics-Theory and Methods*, 27(4):887–904, 1998. doi: 10.1080/03610929808832134.

[3] N. Eugene, C. Lee, and F. Famoye. Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*, 31(4):497–512, 2002. doi: 10.1081/STA-120003130.

[4] K. Zografos and N. Balakrishnan. On the families of beta-and gamma-generated generalized distribution and associated inference. *Statistical Methodology*, 6(4):344–362, 2009. doi: 10.1016/j.stamet.2008.12.003.

[5] M.M. Ristić and N. Balakrishnan. The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, 82(8):1191–1206, 2012. doi: 10.1080/00949655.2011.574633.

[6] A. Alzaatreh, C. Lee, and F. Famoye. A new method for generating families of continuous distributions. *Metron*, 71(1):63–79, 2013. doi: 10.1007/s40300-013-0007-y.

[7] M. Bourguignon, R.B. Silva, and G.M. Cordeiro. The weibull-g family of probability distributions. *Journal of Data Science*, 12:53–68, 2014. doi: 10.6339/JDS.2014.12(1).1210.

[8] D. Kumar, U. Singh, and S. K. Singh. A new distribution using sine function- its application to bladder cancer patients data. *J. Stat. Appl. Pro.*, 4(3):417–427, 2015. doi: 10.12785/jsap/040309, http://www.naturalspublishing.com/files/published/j9wsil53h390x8.pdf.

[9] L. Souza, W.R.O. Junior, C.C.R. de Brito, C. Chesneau, T.A.E. Ferreira, and L. Soares. On the sin-g class of distributions: theory, model and application. *Journal of Mathematical Modeling*, 7(3):357–379, 2019. doi: 10.22124/jmm.2019.13502.1278.

[10] L. Souza, W.R.O. Junior, C.C.R. de Brito, C. Chesneau, T.A.E. Ferreira, and L. Soares. General properties for the cos-g class of distributions with applications. *Eurasian Bulletin of Mathematics*, 2(2):63–79, 2019. http://www.ebmmath.com/index.php/EBM/article/view/42.

[11] B. Hosseini, M. Afshari, and M. Alizadeh. The generalized odd gamma-g family of distributions: Properties and applications. *Austrian Journal of Statistics*, 47(2):69–89, 2018. doi: 10.17713/ajs.v47i2.580.

[12] G.K. Gilbert. The moon's face; a study of the origin of its features. *Bulletin of the Philosophical Society of Washington*, 12:241–292, 1895. https://catalog.hathitrust.org/Record/100359666.

[13] D.H. Raab and E.H. Green. A cosine approximation to the normal distribution. *Psychometrika*, 26(4):447–450, 1961. doi: 10.1007/BF02289774.

[14] J.T. Kent and D.E. Tyler. Maximum likelihood estimation for the wrapped cauchy distribution. *J. Appl. Statist.*, 15(2):247–254, 1988. doi: 10.1080/02664768800000029.

[15] S. Nadarajah and S. Kotz. Beta trigonometric distribution. *Portuguese Economic Journal*, 5(3):207–224, 2006. doi: 10.1007/s10258-006-0013-6.

[16] R.Q. Al-Faris and S. Khan. Sine square distribution: A new statistical model based on the sine function. *Journal of Applied Probability and Statistics*, 3(1):163–173, 2008. http://japs.isoss.net/isoss16a.pdf.

[17] H.S. Bakouch, C. Chesneau, and J. Leao. A new lifetime model with a periodic hazard rate and an application. *Journal of Statistical Computation and Simulation*, 88(11):2048–2065, 2018. doi: 10.1080/00949655.2018.1448983.

[18] L. Souza. *New trigonometric classes of probabilistic distributions*. PhD thesis, Universidade Federal Rural de Pernambuco, Recife, Brasil, 7 2009. http://www.tede2.ufrpe.br:8080/tede2/handle/tede2/5127.

[19] W.T. Shaw and I.R.C. Buckley. The alchemy of probability distributions: beyond gram-charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. arXiv preprint arXiv:0901.0434, 1 2009. An optional note.

[20] J.F. Kenney and E.S. Keeping. *Mathematics of Statistics, 3rd ed.*, volume 76. Chapman and Hall Ltd, New Jersey, USA, 1962.

[21] J.J. Moors. A quantile alternative for kurtosis. *Journal of the Royal Statistical Society. Series D*, 37:25–32, 1988. https://www.jstor.org/stable/2348376.

[22] A.Z. Keller and A.R. Kamath. Alternative reliability models for mechanical systems. In Centre national d'études spatiales and European Space Agency, editors, *Proceedings of the Third International Conference on Reliability and Maintainability*, volume 179 of *Economie et statistiques avancées*, pages 411–415, Toulouse, France, 1982. https://books.google.co.ve/books?id=NeMIjwEACAAJ.

[23] R.B. Millar. *Maximum likelihood estimation and inference: with examples in R, SAS, and ADMB. Statistics in practice.* Wiley, Chichester, West Sussex, 2011. ISBN 978-0-470-09482-2.

[24] T. Bjerkedal. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Journal of Statistical Computation and Simulation*, 72(1):130–148, 1960. doi: 10.1093/oxfordjournals.aje.a120129.

[25] J. Lawless. *Statistical Models and Methods for Lifetime Data, 2nd ed.* John Wiley and Sons, New York, USA, 1982. ISBN 0-471-37215-3.

[26] G. Chen and N. Balakrishnan. A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, 27(2):154–161, 1995. doi: 10.1080/00224065.1995.11979578.

[27] M.H. Tahir, G.M. Cordeiro, S. Ali, S. Dey, and A. Manzoor. The inverted nadarajah-haghighi distribution: estimation methods and applications. *Journal of Statistical Computation and Simulation*, 88(14):2775–2798, 2018. doi: 10.1080/00949655.2018.1487441.

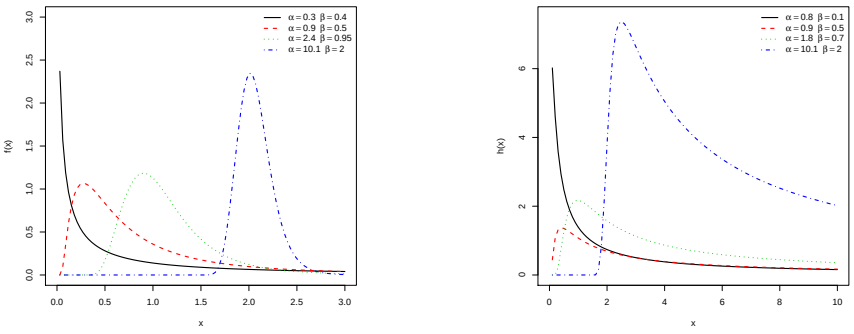
[28] A.Z. Keller and A.R. Kamath. Reliability analysis of cnc machine tools. *Reliability Engineering*, 3(6):449–473, 1982. doi: 10.1016/0143-8174(82)90036-1.

[29] V.G. Voda. On the inverse rayleigh random variable. *Rep. Statist. Appl. Res. Un. Japan. Sci. Engrs.*, 19(1):13–21, 1972.

[30] M. Pararai, G. Warahena-Liyanage, and B.O. Oluyede. A new class of generalized inverse weibull distribution with applications. *Journal of Applied Mathematics and Bioinformatics*, 4(2):17–35, 2014. http://www.scienpress.com/Upload/JAMB/Vol

Tab. 1: Some members of the N-sine-G family of distributions described by their cdfs.

cdf $G(x)$	Support	cdf $F(x)$	Parameters
Uniform	$(0,\theta)$	$\sin\left(\frac{\pi}{4}\frac{x}{\theta}(\frac{x}{\theta}+1)\right)$	(θ)
Arcsin	$(0,1)$	$\sin\left(\frac{1}{2}\arcsin(\sqrt{x})(\frac{2}{\pi}\arcsin(\sqrt{x})+1)\right)$	-
Exponential	$(0,+\infty)$	$\sin\left(\frac{\pi}{4}(1-e^{-\lambda x})(2-e^{-\lambda x})\right)$	(λ)
Weibull	$(0,+\infty)$	$\sin\left(\frac{\pi}{4}(1-e^{-(\beta x)^{\alpha}})(2-e^{-(\beta x)^{\alpha}})\right)$	(α,β)
Inverse Weibull	$(0,+\infty)$	$\sin\left(\frac{\pi}{4}e^{-(\beta/x)^{\alpha}}(1+e^{-(\beta/x)^{\alpha}})\right)$	(α,β)
Burr XII	$(0,+\infty)$	$\sin\left(\frac{\pi}{4}\{1-[1+(x/s)^{c}]^{-k}\}\{2-[1+(x/s)^{c}]^{-k}\}\right)$	(c,k,s)
Logistic	$\mathbb{R}$	$\sin\left(\frac{\pi}{4}[1+e^{-(x-\mu)/s}]^{-1}\left([1+e^{-(x-\mu)/s}]^{-1}+1\right)\right)$	(μ,s)
Gumbel	$\mathbb{R}$	$\sin\left(\frac{\pi}{4}\exp(-e^{-(x-\mu)/\sigma})\left\{\exp(-e^{-(x-\mu)/\sigma})+1\right\}\right)$	(μ,σ)
Normal	$\mathbb{R}$	$\sin\left(\frac{\pi}{4}\Phi((x-\mu)/\sigma)\left\{\Phi((x-\mu)/\sigma)+1\right\}\right)$	(μ,σ)



(a) (b)

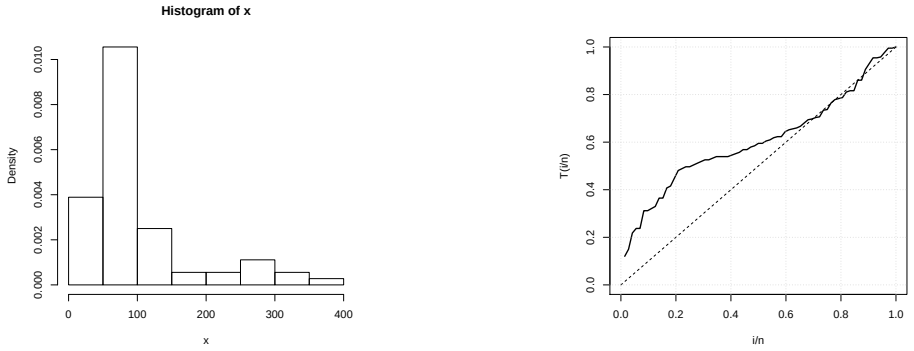
Fig. 1: (a) Curves for the pdf $f(x)$ (b) Curves for the hrf $h(x)$ for selected values of the parameters.

Tab. 2: First quartile, median, third quartile, Bowley skewness and Moors kurtosis of the N-sine-IW(α, β) distribution for the following selected parameters values in order (α, β): (i): $(2.5, 0.3)$, (ii): $(2.5, 1)$, (iii): $(2.5, 2)$, (iv): $(3, 5)$ and (v): $(6, 1)$.

	$Q_F(1/4)$	Med_F	$Q_F(3/4)$	B	M
(i)	0.265	0.331	0.422	0.161	1.332
(ii)	0.884	1.103	1.407	0.161	1.332
(iii)	1.767	2.207	2.814	0.161	1.332
(iv)	4.511	5.427	6.645	0.142	1.314
(v)	0.950	1.042	1.153	0.094	1.279

Tab. 3: Some moments of the N-sine-IW(α, β) distribution for the following selected parameters values in order (α, β): (i): $(2.5, 0.3)$, (ii): $(2.5, 1)$, (iii): $(2.5, 2)$, (iv): $(3, 5)$ and (v): $(6, 1)$.

	$\mathbb{E}(X)$	$\mathbb{E}(X^2)$	$\mathbb{E}(X^3)$	$\mathbb{E}(X^4)$	$\mathbb{V}(X)$	$\sigma(X)$	CV	CS	CK
(i)	0.364	0.156	0.084	0.069	0.023	0.153	0.420	2.861	32.967
(ii)	1.213	1.731	3.105	8.496	0.259	0.509	0.420	2.861	32.967
(iii)	2.426	6.923	24.842	135.939	1.036	1.018	0.420	2.861	32.967
(iv)	5.814	37.634	279.681	2543.236	3.828	1.957	0.337	2.183	16.624
(v)	1.065	1.163	1.303	1.505	0.028	0.167	0.156	1.164	6.219

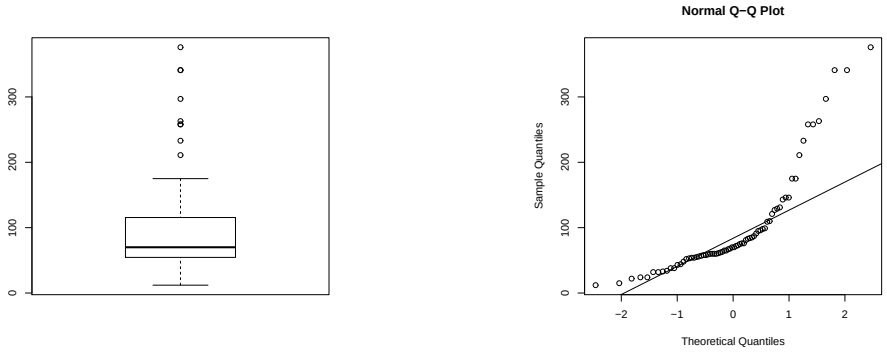


(a) (b)

Fig. 2: (a) Histogram (b) TTT plot for data set 1.

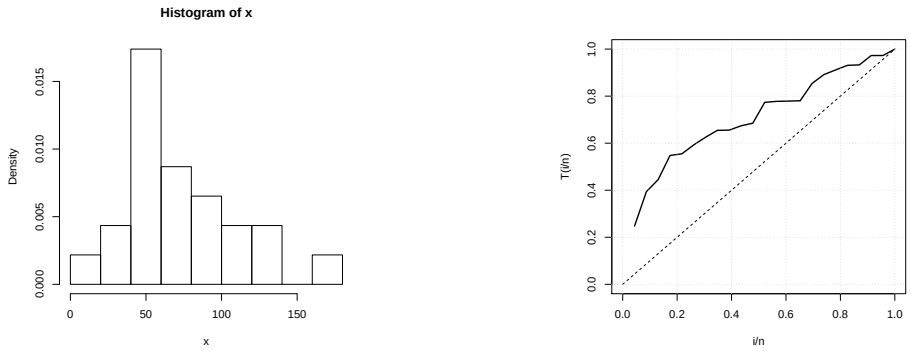
Tab. 4: The Bias, MSE, CP, L.bound and U.bound obtained from simulation of the NSIW model.

n		Initial	Bias	MSE	CP	L.bound	U.bound	Initial	Bias	MSE	CP	L.bound	U.bound
50	α	0.5	0.015	0.003	0.95	0.410	0.619	1	0.028	0.013	0.94	0.819	1.236
	β	1.25	0.045	0.095	0.95	0.696	1.894	1.5	0.017	0.033	0.94	1.167	1.867
100	α	0.5	0.008	0.002	0.94	0.435	0.580	1	0.015	0.006	0.95	0.869	1.161
	β	1.25	0.039	0.050	0.95	0.865	1.712	1.5	0.009	0.016	0.94	1.261	1.757
200	α	0.5	0.003	0.001	0.95	0.452	0.554	1	0.009	0.003	0.96	0.907	1.111
	β	1.25	0.016	0.023	0.95	0.970	1.562	1.5	0.000	0.008	0.95	1.325	1.674
300	a	0.5	0.004	0.000	0.95	0.463	0.546	1	0.003	0.002	0.95	0.920	1.086
	β	1.25	0.000	0.014	0.95	1.013	1.488	1.5	0.002	0.006	0.94	1.359	1.646
500	α	0.5	0.001	0.000	0.96	0.469	0.534	1	0.004	0.001	0.97	0.94	1.068
	β	1.25	0.009	0.009	0.95	1.073	1.445	1.5	0.001	0.003	0.95	1.39	1.612
1000	α	0.5	0.001	0.000	0.95	0.478	0.524	1	0.001	0.001	0.95	0.956	1.046
	β	1.25	0.000	0.004	0.96	1.119	1.381	1.5	0.001	0.002	0.95	1.422	1.579
n		Initial	Bias	MSE	CP	L.bound	U.bound	Initial	Bias	MSE	CP	L.bound	U.bound
50	α	3	0.085	0.109	0.95	2.458	3.712	5.5	0.162	0.385	0.96	4.512	6.813
	β	5.5	0.020	0.051	0.94	5.096	5.944	7.25	0.017	0.026	0.94	6.963	7.571
100	α	3	0.036	0.051	0.94	2.601	3.472	5.5	0.074	0.181	0.95	4.775	6.374
	β	5.5	-0.001	0.023	0.95	5.198	5.801	7.25	0.010	0.012	0.95	7.043	7.477
200	α	3	0.030	0.027	0.94	2.723	3.337	5.5	0.046	0.085	0.95	4.985	6.108
	β	5.5	0.004	0.013	0.94	5.290	5.717	7.25	0.001	0.006	0.95	7.098	7.405
300	a	3	0.009	0.017	0.94	2.760	3.258	5.5	0.018	0.054	0.95	5.062	5.974
	β	5.5	0.002	0.008	0.94	5.327	5.678	7.25	0.000	0.004	0.95	7.124	7.376
500	α	3	0.005	0.010	0.95	2.813	3.197	5.5	0.017	0.033	0.94	5.164	5.871
	β	5.5	0.005	0.005	0.95	5.369	5.641	7.25	0.001	0.002	0.95	7.153	7.348
1000	α	3	0.003	0.005	0.95	2.867	3.139	5.5	0.002	0.018	0.94	5.253	5.751
	β	5.5	0.000	0.002	0.94	5.404	5.596	7.25	0.000	0.001	0.96	7.181	7.319



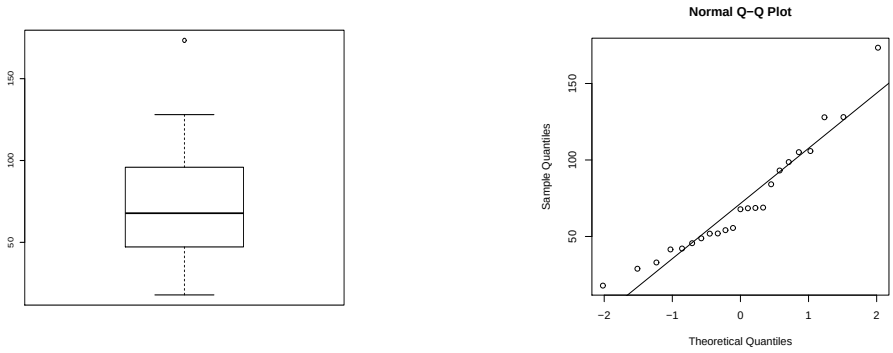
(a) (b)

Fig. 3: (a) Boxplot (b) Normal Q-Q plot for data set 1.



(a) (b)

Fig. 4: (a) Histogram (b) TTT plot for data set 2.



(a) (b)

Fig. 5: (a) Boxplot (b) Normal Q-Q plot for data set 2.

Tab. 5: MLEs and their standard errors (in parentheses) for the data sets 1 and 2.

Distribution	α	β
	Data set 1	
NSIW	1.187246 (0.09749533)	59.277943 (4.89936841)
SIW	1.087296 (0.09010003)	78.679016 (7.12214888)
IW	1.415099 (0.1172978)	54.154478 (4.7819604)
INH	1.837227 (0.6195688)	25.777213 (11.9194201)
IED	60.09902 (7.082733)	- -
IRD	2124.003 (250.2112)	- -
	Data set 2	
NSIW	1.555214 (0.221727)	52.354276 (5.828250)
SIW	1.418507 (0.2050217)	64.874725 (7.9449506)
IW	1.834041 (0.2691721)	48.612750 (5.8731302)
INH	12.265650 (46.28737)	2.913329 (11.59465)
IED	54.98987 (11.46618)	- -
IRD	2235.817 (466.2137)	- -

Tab. 6: The statistics $\hat{\ell}$, AIC, BIC, A^*, W^*, K-S and P-value for the data sets 1 and 2.

Distribution	$\hat{\ell}$	AIC	BIC	A^*	W^*	K-S	P-value
			Data set 1				
NSIW	**391.114**	**786.228**	**790.7813**	**0.7421414**	**0.128227**	**0.12022**	**0.2491**
SIW	391.8296	787.6592	792.2125	0.815797	0.1390462	0.12695	0.1962
IW	395.6491	795.4722	799.8516	1.283477	0.2148432	0.15231	0.07082
INH	400.4679	804.9357	809.4891	1.427511	0.2398301	0.14121	0.1132
IED	402.6718	807.3437	809.6203	0.9225515	0.1561167	0.18466	0.01474
IRD	406.7674	815.5347	817.8114	1.994931	0.3371843	0.26145	0.0001062
			Data set 2				
NSIW	**113.9501**	**231.9001**	**234.1711**	**0.2910172**	**0.03847493**	**0.09927**	**0.9604**
SIW	114.3037	232.6073	234.8783	0.340215	0.04477491	0.10618	0.9335
IW	115.7833	235.5666	237.8376	0.5546877	0.07517144	0.13261	0.7654
INH	119.027	242.054	244.325	0.76066	0.1063972	0.22877	0.1536
IED	121.7296	245.4592	246.5947	0.3238192	0.04286789	0.30567	0.02099
IRD	115.967	233.934	235.0695	0.6101612	0.0835341	0.14293	0.6831

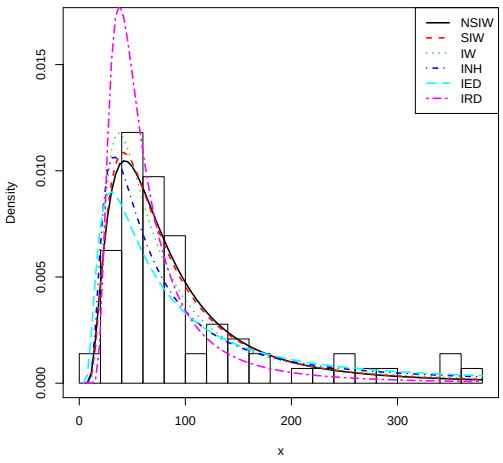


Fig. 6: Plots of the estimated pdfs for data set 1.

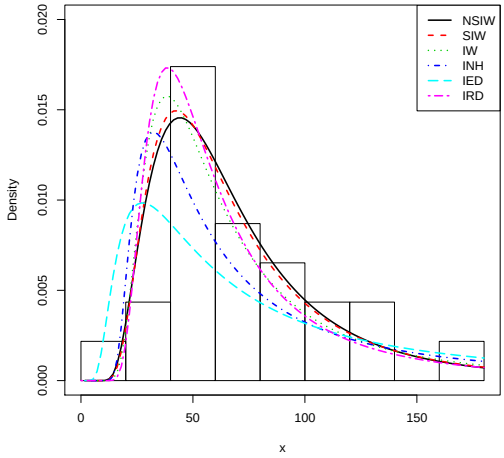


Fig. 7: Plots of the estimated pdfs for data set 2.

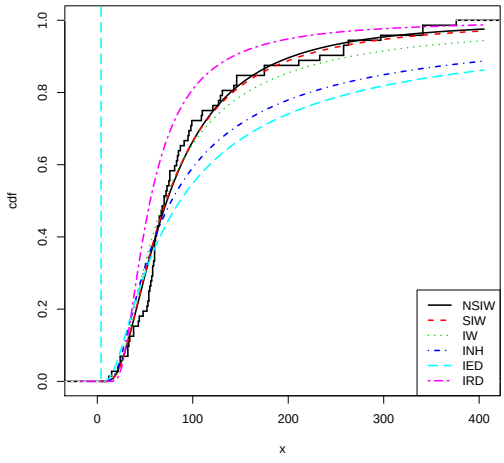


Fig. 8: Plots of the estimated cdfs for data set 1.

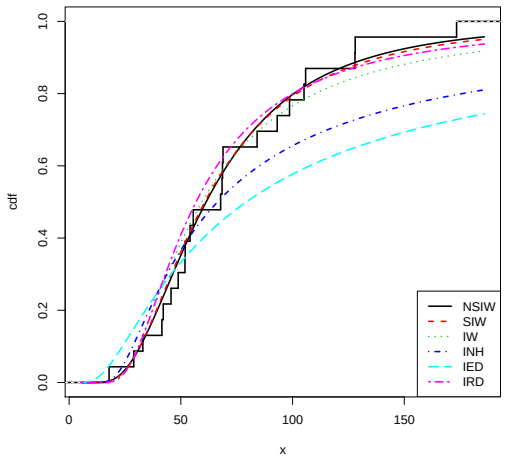


Fig. 9: Plots of the estimated cdfs for data set 2.

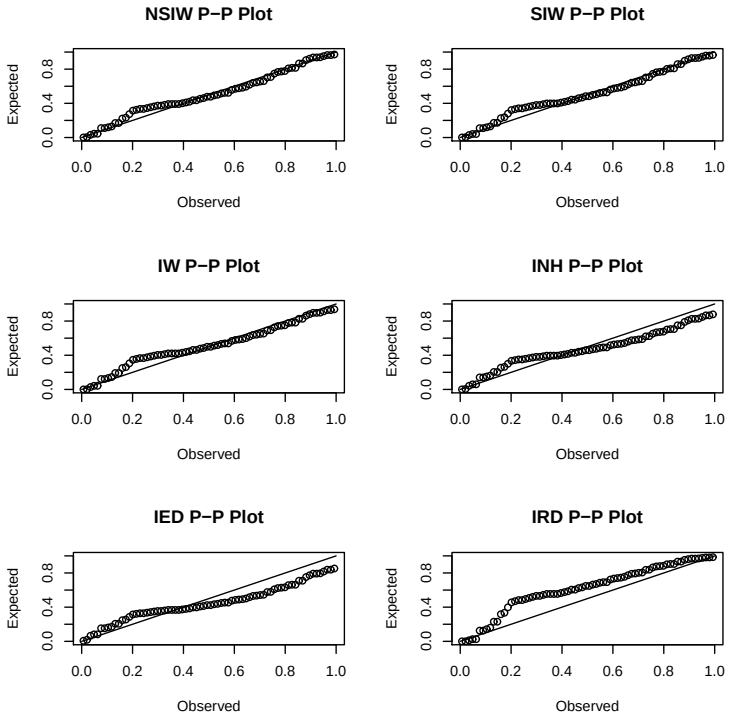


Fig. 10: P-P Plots of the estimated models for data set 1.

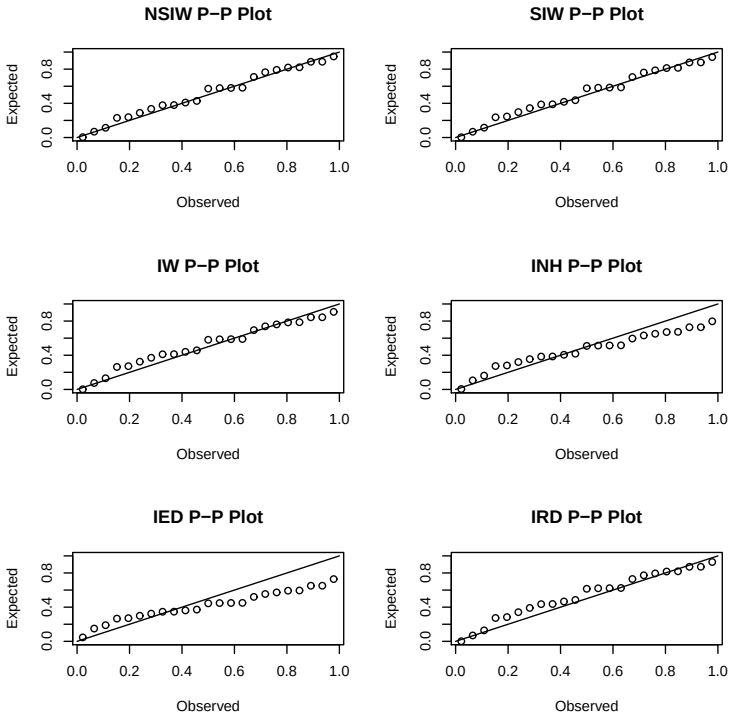


Fig. 11: P-P Plots of the estimated models for data set 2.