Cubo
A Mathematical Journal

CUBO, A Mathematical Journal
Vol. 23, no. 01, pp. 01–20, April 2021
DOI: 10.4067/S0719-06462021000100001

Tan-G class of trigonometric distributions and its applications

Luciano Souza[1]
Wilson Rosa de O. Júnior[2]
Cícero Carlos R. de Brito[3]
Christophe Chesneau[4]
Renan L. Fernandes[5]
Tiago A. E. Ferreira[6]

[1] *UFAPE, Federal University of Agreste of Pernambuco, Garanhuns / PE, Brazil.*
`lcnsza@gmail.com`

[2,6] *PPGBEA, Federal Rural University of Pernambuco, Recife / PE, Brazil.*
`wilson.rosa@gmail.com`;
`taef.first@gmail.com`,

[3] *Federal Institute of Pernambuco, Pernambuco / PE, Brazil.*
`cicerocarlosbrito@yahoo.com.br`

[4] *LMNO, University of Caen-Normandie, Caen, 14032, France.*
`christophe.chesneau@unicaen.fr`

[5] *Centro de Informática, Universidade Federal de Pernambuco, Recife/PE, Brazil.*
`leandrorenanf@gmail.com`

ABSTRACT

In this paper, we introduce a new general class of trigonometric distributions based on the tangent function, called the Tan-G class. A mathematical procedure of the Tan-G class is carried out, including expansions for the probability density function, moments, central moments and Rényi entropy. The estimates are acquired in a non-closed form by the maximum likelihood estimation method. Then, an emphasis is put on a particular member of this class defined with the Burr XII distribution as baseline, called the Tan-BXII distribution. The inferential properties of the Tan-BXII model are investigated. Finally, the Tan-BXII model is applied to a practical data set, illustrating the interest of the Tan-G class for the practitioner.

RESUMEN

En este artículo, introducimos una nueva clase general de distribuciones trigonométricas basada en la función tangente, llamada la clase Tan-G. Se lleva a cabo un procedimiento matemático para la clase Tan-G, incluyendo expansiones para la función de densidad de probabilidad, momentos, momentos centrales y entropía de Rényi. Las estimaciones se obtienen en forma no-cerrada para el método de estimación de máxima verosimilitud. Luego, se pone énfasis en un miembro particular de esta clase definido con la distribución Burr XII como línea de base, llamada la distribución Tan-BXII. Se investigan las propiedades inferenciales del modelo Tan-BXII. Finalmente, el modelo Tan-BXII es aplicado para un conjunto de datos prácticos, ilustrando el interés de la clase Tan-G para el practicante.

Keywords and Phrases: Trigonometric class of distributions, Tangent function, Burr XII distribution, Maximum likelihood estimation, Entropy.

2020 AMS Mathematics Subject Classification: 60E05, 62E15, 62F10.

Accepted: 23 December, 2020
Received: 12 February, 2020

1 Introduction

The recent years of research on probabilistic distributions have been marked by the rise of general classes of trigonometric distributions, more or less sophisticated. Modern statistical developments can be found in, e.g., [10], [16], [18], [19], [11], [4] and [8]. In particular, among the most fundamental of them, [18] introduced the Sin-G class defined by the cumulative distribution function (cdf) given by

$$H_G^{(1)}(x) = \sin\left(\frac{\pi}{2}G(x)\right), \quad x \in \mathbb{R},$$

where $G(x)$ denotes a baseline cdf of a continuous distribution and [19] proposed the Cos-G class defined by the cdf given by

$$H_G^{(2)}(x) = 1 - \cos\left(\frac{\pi}{2}G(x)\right), \quad x \in \mathbb{R}.$$

One can notice that the eventual parameter(s) of these classes is (are) (the one) (those) of $G(x)$ only, and that the following elementary equation hold: $[H_G^{(1)}(x)]^2+[1-H_G^{(2)}(x)]^2 = 1$, i.e., $H_G^{(2)}(x) = 1-\sqrt{1-[H_G^{(1)}(x)]^2}$ (showing that $H_G^{(2)}(x)$ belongs to the so-called Kum-G class with the parameters 1/2 and 2 and the baseline cdf $H_G^{(1)}(x)$, see [5]). In addition to their simplicity, both of these two trigonometric classes benefit from the smooth periodic oscillations of the involved trigonometric functions to attain new levels of flexibility in statistical modeling. In [18] and [19], this fact is illustrated by means of several practical data sets, with winning results in comparison to useful model competitors. In this study, following the spirit of [18] and [19], we introduce a new and simple general class of trigonometric distributions having the feature to be centered around the tangent function. For the purpose of this paper, we call it the Tan-G class. It is defined by the following cdf:

$$H_G(x) = \tan\left(\frac{\pi}{4}G(x)\right), \quad x \in \mathbb{R}. \tag{1.1}$$

Several existing constructions give this cdf, beginning by the integral techniques developed by [2]; we have $H_G(x) = \int_0^{(\pi/4)G(x)} \sec^2(t)dt$, where $\sec(t) = 1/\cos(t)$. After some algebra, one can also notice that $H_G(x)$ can be expressed in terms of the cdfs $H_G^{(1)}(x)$ and $H_G^{(2)}(x)$ as

$$H_G(x) = \frac{\sqrt{1-[1-H_G^{(2)}(x)]^2}}{2-H_G^{(2)}(x)}, \quad H_G(x) = \frac{H_G^{(1)}(x)}{1+\sqrt{1-[H_G^{(1)}(x)]^2}}.$$

From these expressions, we immediately get the following stochastic ordering: $H_G(x) \leq H_G^{(1)}(x)$, attesting that $H_G(x)$ can provide different statistical models to those of $H_G^{(1)}(x)$. In full generality, the main qualities of the Tan-G class are to be simple: there is no additional parameter and the related functions are very tractable, and its ability to create flexible statistical models, well-adapted to fit with precision several kinds of data sets, beyond those related to the Sin-G or Cos-G class.

All these aspects are developed in this paper according to the following plan. In Section 2, the main theoretical features of the Tan-G class are presented. Section 3 is devoted to a special member of the class defined with the Burr XII distribution as baseline. Concluding remarks are given in Section 4.

2 Main theoretical features of the Tan-G class

A theoretical treatment of the Tan-G class is performed in this section, investigating the related distributional functions, asymptotic and critical points, useful expansion, moments and central moments, expansion for the general coefficient, entropy and the mathematics of the maximum likelihood estimation.

2.1 Distributional functions

We recall that the Tan-G class of distributions is defined by the cdf given by (1.1). Upon differentiation, the corresponding pdf is given by

$$h_G(x) = \frac{\pi}{4} g(x) \sec^2\left(\frac{\pi}{4} G(x)\right), \quad x \in \mathbb{R}, \tag{2.1}$$

where $g(x)$ denotes the pdf corresponding to $G(x)$. The hazard function (hf) of the Tan-G class is given by

$$R_G(x) = \frac{\frac{\pi}{4} g(x) \sec^2\left(\frac{\pi}{4} G(x)\right)}{1 - \tan\left(\frac{\pi}{4} G(x)\right)}, \quad x \in \mathbb{R}. \tag{2.2}$$

The curvatures properties of $h_G(x)$ and $R_G(x)$ are crucial to define an appropriate statistical model for a given data set. Further elements on these curvature properties will be presented in the subsection below. Another important function is the quantile function (qf) given by

$$Q(u) = H_G^{-1}(u) = G^{-1}\left[\frac{4}{\pi} \arctan(u)\right], \quad u \in (0, 1).$$

That is, the median of the Tan-G class is given by

$$M = Q(0.5) \approx G^{-1}(0.5903345).$$

Other properties of the Tan-G class can be studied through this qf. For instance, the main steps to generate random numbers from the Tan-G class via the qf are described in Table 1.

Table 1: Generated numbers from the Tan-G class by the use of the qf

Algorithm
1. Generate n values from $u \sim U(0,1)$
2. Specify $G^{-1}(x)$
3. Obtain an outcome of X with cdf (1.1) by $X = Q(u)$

2.2 Asymptotic and critical points

Let us now investigate the asymptotic and critical points for $h_G(x)$ and $R_G(x)$. Owing to (2.1) and (2.2), when $G(x) \to 0$, we have

$$H_G(x) \sim \frac{\pi}{4} G(x), \quad h_G(x) \sim \frac{\pi}{4} g(x), \quad R_G(x) \sim \frac{\pi}{4} g(x).$$

Also, when $G(x) \to 1$, we have

$$H_G(x) \sim 1 - \frac{\pi}{2}(1 - G(x)), \quad h_G(x) \sim \frac{\pi}{2} g(x), \quad R_G(x) \sim \frac{g(x)}{1 - G(x)}.$$

If x_* denotes a critical point for $h_G(x)$, then it satisfies the following equation: $\{\ln[h_G(x)]\}' \mid_{x=x_*} = 0$, i.e.,

$$g(x)' \mid_{x=x_*} + \frac{\pi}{2} g(x_*)^2 \tan\left(\frac{\pi}{4} G(x_*)\right) = 0.$$

With similar arguments, if x_{**} denotes a critical point for $R_G(x)$, then it satisfies the following equation: $\{\ln[R_G(x)]\}' \mid_{x=x_{**}} = 0$, i.e.,

$$\left[g(x)' \mid_{x=x_{**}} + \frac{\pi}{2} g(x_{**})^2 \tan\left(\frac{\pi}{4} G(x_{**})\right)\right] \left[1 - \tan\left(\frac{\pi}{4} G(x_{**})\right)\right] + \frac{\pi}{4} g(x_{**})^2 \sec^2\left(\frac{\pi}{4} G(x_{**})\right) = 0.$$

None of these non-linear equations has solution(s) with closed form. That is, for a specific $G(x)$, we can determine x_* and x_{**} numerically by the use of any scientific software as R, Matlab, Mathematica...

2.3 Useful expansion

The following result presents an useful expansion of the pdf of the Tan-G class involving functions of the exponentiated-G class (see [7]).

Theorem 2.1. *The pdf of the Tan-G class given by* (2.1) *can be expressed as a linear combination of pdfs of the exponentiated-G class as*

$$h_G(x) = \sum_{k=1}^{+\infty} \omega_k g_{(2k-1)}(x),$$

where

$$\omega_k = \left(\frac{\pi}{4}\right)^{2k-1} \frac{B_{2k}(-4)^k(1-4^k)}{(2k)!}, \tag{2.3}$$

B_{2k} is the so-called 2kth Bernoulli number and $g_{(2k-1)}(x) = (2k-1)g(x)G^{2k-2}(x)$ is the pdf of the exponentiated-G class with parameter $2k-1$.

Proof. Using the Taylor series for the tangent function, since $(\pi/4)G(x) \in (0, \pi/2)$, we have

$$\tan\left(\frac{\pi}{4}G(x)\right) = \sum_{k=1}^{+\infty} \frac{B_{2k}(-4)^k(1-4^k)}{(2k)!} \left(\frac{\pi}{4}G(x)\right)^{2k-1}.$$

Thus, we obtain the following expansion for $H_G(x)$:

$$H_G(x) = \sum_{k=1}^{+\infty} \left(\frac{\pi}{4}\right)^{2k-1} \frac{B_{2k}(-4)^k(1-4^k)}{(2k)!} G^{2k-1}(x)\cdot$$

The desired expansion for $h_G(x)$ is deduced by differentiation. This ends the proof of Theorem 2.1. □

2.4 Moments and central moments

An expansion for the moment of order m of the Tan-G class is studied in the following result.

Theorem 2.2. *Let μ_m be the moment of order m of the Tan-G class and $\mu_m^{(2k-1)}$ be the moment of order m of the exponentiated-G class with parameter $2k-1$. Then, we have*

$$\mu_m = \sum_{k=1}^{+\infty} \omega_k \mu_m^{(2k-1)},$$

where ω_k is given by (2.3).

Proof. The moment of order m of the Tan-G class is defined by

$$\mu_m = \int_{-\infty}^{+\infty} x^m dH_G(x).$$

It follows from Theorem 2.1 that

$$\mu_m = \int_{-\infty}^{+\infty} x^m \sum_{k=1}^{+\infty} \omega_k g_{(2k-1)}(x)dx = \sum_{k=1}^{+\infty} \omega_k \int_{-\infty}^{+\infty} x^m g_{(2k-1)}(x)dx = \sum_{k=1}^{+\infty} \omega_k \mu_m^{(2k-1)}.$$

This ends the proof of Theorem 2.2. □

The mean is given by $\mu = \mu_1$.

Remark 2.3. *By applying the change of variable $u = G(x)$, we can express $\mu_m^{(2k-1)}$ as*

$$\mu_m^{(2k-1)} = (2k-1)\int_{-\infty}^{+\infty} x^m g(x)G^{2k-2}(x)dx = (2k-1)\int_0^1 \left[G^{-1}(u)\right]^m u^{2k-2}du.$$

Similarly, we can obtain an expansion of the central moments of order m by using Theorem 2.2.

Corollary 2.4. *Let μ'_m be the central moment of order m of the Tan-G class and $\mu_m^{(2k-1)}$ be the moment of order m of the exponentiated-G class with parameter $2k-1$. Then, we have*

$$\mu'_m = \sum_{k=1}^{+\infty}\sum_{r=0}^{m} \gamma_{k,m,r}\mu_{m-r}^{(2k-1)},$$

where

$$\gamma_{k,m,r} = \omega_k \binom{m}{r}(-1)^r \mu^r$$

and ω_k is defined by (2.3).

Proof. The central moment of order m of the Tan-G class is defined by

$$\mu'_m = \int_{-\infty}^{+\infty} (x-\mu)^m dH_G(x).$$

By using the binomial theorem and Theorem 2.2, we have

$$\begin{aligned}\mu'_m &= \sum_{r=0}^{m}\binom{m}{r}(-1)^r\mu^r \int_{-\infty}^{+\infty} x^{m-r} dH_G(x) = \sum_{r=0}^{m}\binom{m}{r}(-1)^r\mu^r \mu_{m-r} \\ &= \sum_{r=0}^{m}\binom{m}{r}(-1)^r\mu^r \sum_{k=1}^{+\infty}\omega_k \mu_{m-r}^{(2k-1)} = \sum_{k=1}^{+\infty}\sum_{r=0}^{m}\gamma_{k,m,r}\mu_{m-r}^{(2k-1)}.\end{aligned}$$

The proof of Corollary 2.4 is ended. □

By considering $m = 2$, the variance is given by

$$\sigma^2 = \mu'_2 = \sum_{k=1}^{+\infty}\sum_{r=0}^{2}\gamma_{k,2,r}\mu_{2-r}^{(2k-1)}.$$

By using similar summation techniques, one can set expansions of the incomplete moments, the moment generating function and the characteristic function, among others.

2.5 Expansion to the general coefficient

The general coefficient of the Tan-G class is defined by

$$C_m = \frac{\mu'_m}{\sigma^m}.$$

By applying Corollary 2.4, it can be written as

$$C_m = \frac{\sum_{k=1}^{+\infty}\sum_{r=0}^{m}\gamma_{k,m,r}\mu_{m-r}^{(2k-1)}}{\left[\sum_{k=1}^{+\infty}\sum_{r=0}^{2}\gamma_{k,2,r}\mu_{2-r}^{(2k-1)}\right]^{\frac{m}{2}}}.$$

So, the asymmetry and kurtosis of the Tan-G class can be respectively expressed by

$$C_3 = \frac{\sum_{k=1}^{+\infty}\sum_{r=0}^{3}\gamma_{k,3,r}\mu_{3-r}^{(2k-1)}}{\left[\sum_{k=1}^{+\infty}\sum_{r=0}^{2}\gamma_{k,2,r}\mu_{2-r}^{(2k-1)}\right]^{\frac{3}{2}}}, \quad C_4 = \frac{\sum_{k=1}^{+\infty}\sum_{r=0}^{4}\gamma_{k,4,r}\mu_{4-r}^{(2k-1)}}{\left[\sum_{k=1}^{+\infty}\sum_{r=0}^{2}\gamma_{k,2,r}\mu_{2-r}^{(2k-1)}\right]^{2}}.$$

2.6 Entropy

Entropy measures the uncertainty; the greater the entropy, the higher the disorder and the less likely it will be to observe a phenomenon; the lower the entropy, the lower its disorder and the higher the probability of observing a particular event. Among the most useful entropy, there is the Rényi entropy introduced by [13]. In the context of the Tan-G class, it is defined by

$$\mathfrak{L}_G(\gamma) = \frac{1}{1-\gamma}\ln\left[\int_{-\infty}^{+\infty} h_G^{\gamma}(x)\right]dx,$$

where $\gamma > 0$ with $\gamma \neq 1$ and

$$h_G^{\gamma}(x) = \left(\frac{\pi}{4}\right)^{\gamma} g^{\gamma}(x)\sec^{2\gamma}\left(\frac{\pi}{4}G(x)\right).$$

Let us now consider the function $W(s) = \sec^{2\gamma}[(\pi/4)s]$, $s \in (0,1)$. By applying the Taylor series formula to $W(s)$ at a fixed point $s_0 \in (0,1)$ (say $s_0 = 0.5$), we get

$$\sec^{2\gamma}\left[\frac{\pi}{4}s\right] = \sum_{k=0}^{+\infty} a_k(s-s_0)^k = \sum_{k=0}^{+\infty}\sum_{r=0}^{k}\binom{k}{r}a_k s^r(-1)^{k-r}s_0^{k-r},$$

where $a_k = W^{(k)}(s)\mid_{s=s_0}/k!$. We are now able to derive an expansion of the Rényi entropy of the Tan-G class. After some algebra, we obtain

$$\mathfrak{L}_G(\gamma) = \frac{1}{1-\gamma}\left\{\gamma\ln\left(\frac{\pi}{4}\right) + \ln\left[\sum_{k=0}^{+\infty}\sum_{r=0}^{k} a_k s^r(-1)^{k-r}s_0^{k-r}I_r\right]\right\}, \tag{2.4}$$

where

$$I_r = \int_{-\infty}^{+\infty} G^r(x)g^{\gamma}(x)dx.$$

Even if it has no closed form, the integral I_r can be computed numerically. The Shannon entropy, pioneered by [15], is given by

$$\mathfrak{S}_G = -\int_{-\infty}^{+\infty}\ln[h_G(x)]h_G(x)dx.$$

It can deduced from $\mathfrak{L}_G(\gamma)$ via the relation $\lim_{\gamma\to 1}\mathfrak{L}_G(\gamma) = \mathfrak{S}_G$.

2.7 Maximum likelihood estimation and scores

Here, we consider the estimation of the parameters of the Tan-G class by the method of maximum likelihood. Let $\tilde{x} = (x_1, \ldots, x_n)^\top$ be a random sample observations from the Tan-G class with vector parameter $\tilde{\theta} = (\theta_1, \ldots, \theta_p)$ (thus, p is the number of parameters of the distribution). Then, the log-likelihood (LL) function for the Tan-G class is given by

$$\ell(\tilde{\theta}) = n \ln\left(\frac{\pi}{4}\right) + \sum_{i=1}^{n} \ln\left(g(x_i|\tilde{\theta})\right) + 2\sum_{i=1}^{n} \ln\left[\sec\left(\frac{\pi}{4} G(x_i|\tilde{\theta})\right)\right].$$

The maximum likelihood estimators (MLEs) are obtained by maximizing this function according to $\tilde{\theta}$. In this regards, if $G(x|\tilde{\theta})$ is differentiable according to $\tilde{\theta}$, one can consider the jth score given by

$$U(\theta_j) = \frac{\partial \ell(\tilde{\theta})}{\partial \theta_j} = \sum_{i=1}^{n} \frac{1}{g(x_i|\tilde{\theta})} \frac{\partial g(x_i|\tilde{\theta})}{\partial \theta_j} + \frac{\pi}{2} \sum_{i=1}^{n} \tan\left(\frac{\pi}{4} G(x_i|\tilde{\theta})\right) \frac{\partial G(x_i|\tilde{\theta})}{\partial \theta_j}$$

and consider the following equations: $U(\theta_1) = 0, \ldots, U(\theta_p) = 0$. Thus, the MLEs are defined as the simultaneous solutions of these equations.

3 The Tan-BXII distribution

We now focus on a special distribution of the Tan-G class, called the Tan-BXII distribution.

3.1 Definition

Tan-BXII distribution is defined by the cdf given by (1.1) with the cdf $G(x)$ of the Burr XII distribution, i.e., $G(x) = 1 - \left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa}$, $x, s, c, \kappa > 0$. Hence, the cdf of the Tan-BXII distribution is given by

$$H_G(x) = \tan\left\{\frac{\pi}{4}\left(1 - \left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa}\right)\right\}, \quad x > 0.$$

The corresponding pdf is given by

$$h_G(x) = \frac{\pi}{4}\left\{x^{c-1} c\kappa s^{-c}\left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa-1}\right\} \sec^2\left\{\frac{\pi}{4}\left(1 - \left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa}\right)\right\}, \quad x > 0.$$

Finally, the corresponding hf is given by

$$R_G(x) = \frac{\frac{\pi}{4}\left\{x^{c-1} c\kappa s^{-c}\left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa-1}\right\} \sec^2\left\{\frac{\pi}{4}\left(1 - \left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa}\right)\right\}}{1 - \tan\left\{\frac{\pi}{4}\left(1 - \left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa}\right)\right\}}, \quad x > 0.$$

It is expected that the hf is unimodal or decreasing, as it can be seen in Figures 3 and 4, respectively, but an analytic verification of this fact using all three parameters is an unnecessarily complicated computation. One can check for given parameters that it is indeed the case using computing software.

3.2 Shape characteristics of probability density and hazard functions

The asymptotic and critical points for $h_G(x)$ and $R_G(x)$ can be obtained in non-closed form by applying Subsection 2.2. Also, some possible shapes of $h_G(x)$ for some parameter values are displayed in Figure 1. Some plots of $H_G(x)$ are given in Figure 2.

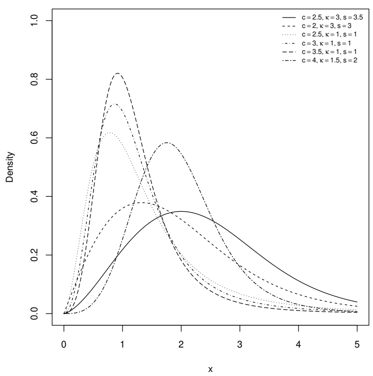


Figure 1: Plots of the pdf of the Tan-BXII distribution

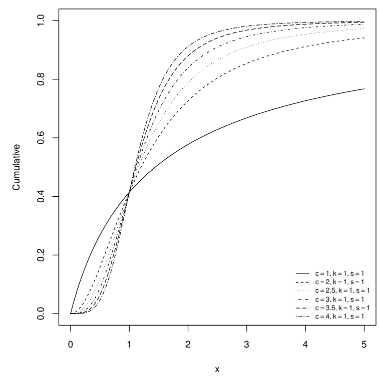


Figure 2: Plots of the cdf of the Tan-BXII distribution

Figures 3 and 4 present plots of $R_G(x)$ for some parameter values. We observe that the hf can be unimodal or only be decreasing.

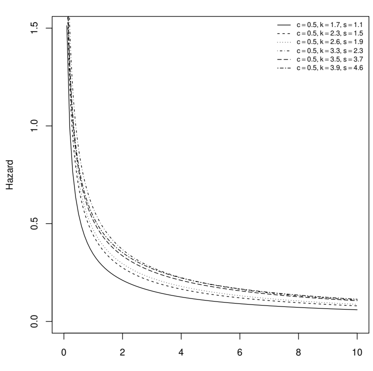


Figure 3: Plots of decreasing hf of the Tan-BXII distribution.

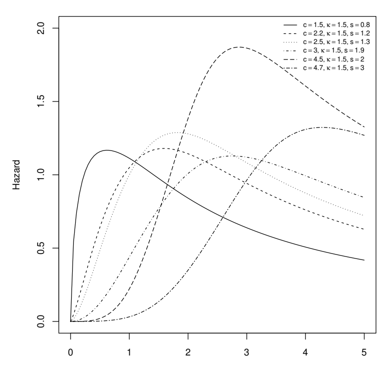


Figure 4: Plots of unimodal hf of the Tan-BXII distribution.

3.3 Expansion of the probability density function

Here, we use the general results proved for the Tan-G class of distributions to reveal properties for the Tan-BXII distribution. An useful expansion of the pdf is presented below.

Theorem 3.1. *The pdf of the Tan-G class can be expanded as a mixture of pdfs of the Burr XII distribution, i.e.,*

$$h_G(x) = \sum_{k=1}^{+\infty} \sum_{j=0}^{2k-2} \omega_{j,k} g_{BurrXII}(x; s, c, \kappa(j+1)),$$

where

$$\omega_{j,k} = \omega_k (2k-1) \binom{2k-2}{j} (-1)^j \frac{1}{j+1}, \tag{3.1}$$

ω_k *is given by* (2.3) *and* $g_{BurrXII}(x; s, c, \kappa(j+1))$ *is the pdf of the Burr XII distribution with parameters* s, c *and* $\kappa(j+1)$, *i.e.,* $g_{BurrXII}(x; s, c, \kappa(j+1)) = x^{c-1} c\kappa(j+1) s^{-c} \left[1 + (x/s)^c\right]^{-\kappa(j+1)-1}$, $x > 0$.

Proof. Owing to Theorem 2.1, we can write

$$h_G(x) = \sum_{k=1}^{+\infty} \omega_k g_{(2k-1)}(x),$$

where ω_k is given by (2.3) and

$$\begin{aligned} g_{(2k-1)}(x) &= (2k-1) g(x) G^{2k-2}(x) \\ &= (2k-1) x^{c-1} c\kappa s^{-c} \left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa-1} \left\{1 - \left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa}\right\}^{2k-2}. \end{aligned}$$

The standard binomial theorem gives

$$\begin{aligned} g_{(2k-1)}(x) &= (2k-1) x^{c-1} c\kappa s^{-c} \sum_{j=0}^{2k-2} \binom{2k-2}{j} (-1)^j \left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa(j+1)-1} \\ &= (2k-1) \sum_{j=0}^{2k-2} \binom{2k-2}{j} (-1)^j \frac{1}{j+1} g_{BurrXII}(x; s, c, \kappa(j+1)). \end{aligned}$$

The proof ends by putting the above equalities together. □

3.4 Moments and central moments

By using identical manipulations to those used in Theorem 2.2, we introduce the moment expansion of the Tan-BXII distribution in the following result.

Theorem 3.2. *First of all, the moment of order m of the Tan-BXII distribution exists if and only if $c\kappa > m$. In this case, the moment of order m of the Tan-BXII distribution is given by*

$$\mu_m = \sum_{k=1}^{+\infty}\sum_{j=0}^{2k-2} \omega_{j,k} s^m \kappa(j+1) B\left(\kappa(j+1) - mc^{-1}, 1 + mc^{-1}\right),$$

where $\omega_{j,k}$ is given by (3.1) *and $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$, $a, b > 0$ (the standard beta function).*

Proof. It follows from Theorem 3.1 that

$$\mu_m = \sum_{k=1}^{+\infty}\sum_{j=0}^{2k-2} \omega_{j,k} J_{j,k,m},$$

where

$$J_{j,k,m} = \int_0^{+\infty} x^m g_{BurrXII}(x; s, c, \kappa(j+1))dx = \int_0^{+\infty} x^m x^{c-1} c\kappa(j+1)s^{-c}\left[1 + \left(\frac{x}{s}\right)^c\right]^{-\kappa(j+1)-1} dx.$$

By applying the changes of variables $u = \left(\frac{x}{s}\right)^c$ and $\nu = (1+u)^{-1}$, in turn, we get

$$\begin{aligned} J_{j,k,m} &= s^m \kappa(j+1) \int_0^{+\infty} u^{\frac{m}{c}} (1+u)^{-\kappa(j+1)-1} du \\ &= s^m \kappa(j+1) \int_0^1 \nu^{\kappa(j+1)-\frac{m}{c}-1}(1-\nu)^{\frac{m}{c}} d\nu \\ &= s^m \kappa(j+1) B\left(\kappa(j+1) - mc^{-1}, 1 + mc^{-1}\right). \end{aligned}$$

By combining the above equalities together, we end the proof of Theorem 3.2. □

The mean is given by $\mu = \mu_1$.

Remark 3.3. *By adopting the notations introduced in Section 2, following the lines of the proof of Theorem 3.2, one can show that*

$$\mu_m^{(2k-1)} = (2k-1)s^m \kappa \sum_{j=0}^{2k-2} \binom{2k-2}{j} (-1)^j B\left(\kappa(j+1) - mc^{-1}, 1 + mc^{-1}\right).$$

Similarly to Corollary 2.4, the central moment of order m of the Tan-BXII distribution is given

$$\mu'_m = \sum_{r=0}^{m} \binom{m}{r} (-1)^r \mu^r \mu_{m-r} = \sum_{k=1}^{+\infty}\sum_{j=0}^{2k-2}\sum_{r=0}^{m} \rho_{j,k,m,r} B\left(\kappa(j+1) - (m-r)c^{-1}, 1 + (m-r)c^{-1}\right),$$

where

$$\rho_{j,k,m,r} = \omega_{j,k} s^{m-r} \kappa(j+1) \binom{m}{r} (-1)^r \mu^r.$$

By considering $m = 2$, we get the following expansion for variance of the distribution:

$$\sigma^2 = \mu'_2 = \sum_{k=1}^{+\infty}\sum_{j=0}^{2k-2}\sum_{r=0}^{2} \rho_{j,k,2,r} B\left(\kappa(j+1) - (2-r)c^{-1}, 1 + (2-r)c^{-1}\right).$$

3.5 Expansion to the general coefficient

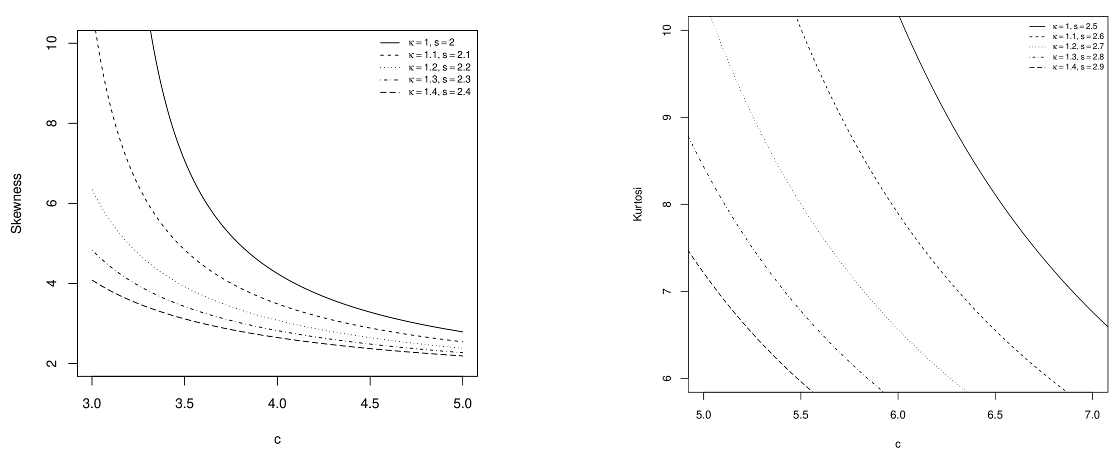
The general coefficient of the Tan-BXII distribution can be expressed as

$$C_m = \frac{\mu'_m}{\sigma^m} = \frac{\sum_{k=1}^{+\infty}\sum_{j=0}^{2k-2}\sum_{r=0}^{m}\rho_{j,k,m,r}B\left(\kappa(j+1)-(m-r)c^{-1},1+(m-r)c^{-1}\right)}{\left\{\sum_{k=1}^{+\infty}\sum_{j=0}^{2k-2}\sum_{r=0}^{2}\rho_{j,k,2,r}B\left(\kappa(j+1)-(2-r)c^{-1},1+(2-r)c^{-1}\right)\right\}^{m/2}}.$$

Thus, the asymmetry and kurtosis can be expressed by taking $m = 3$ and $m = 4$, respectively, which is the object of the next part.

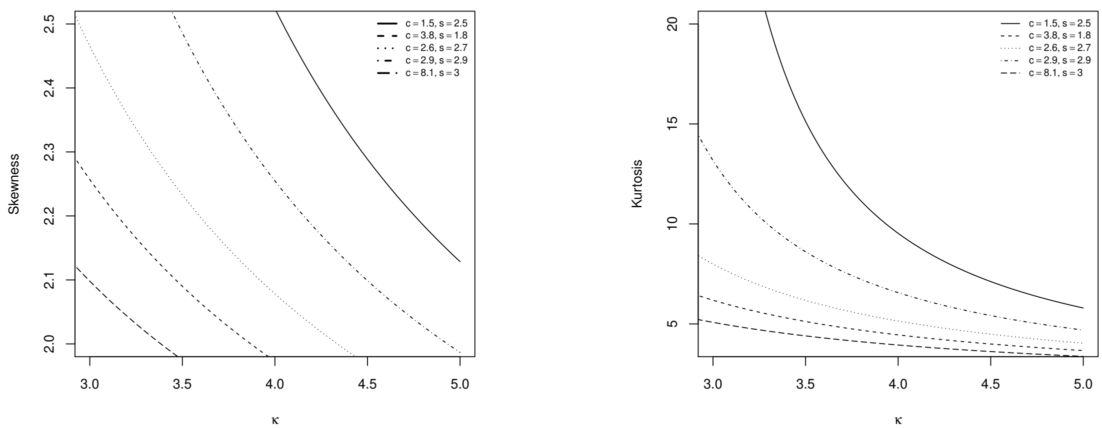
3.6 Figures of asymmetry and kurtosis

In Figures 5, 6 and 7, we present the asymmetry and kurtosis graphs for the Tan-BXII distribution. It is possible to observe that this new distribution has a great flexibility on these aspects, showing varying values, small and large.



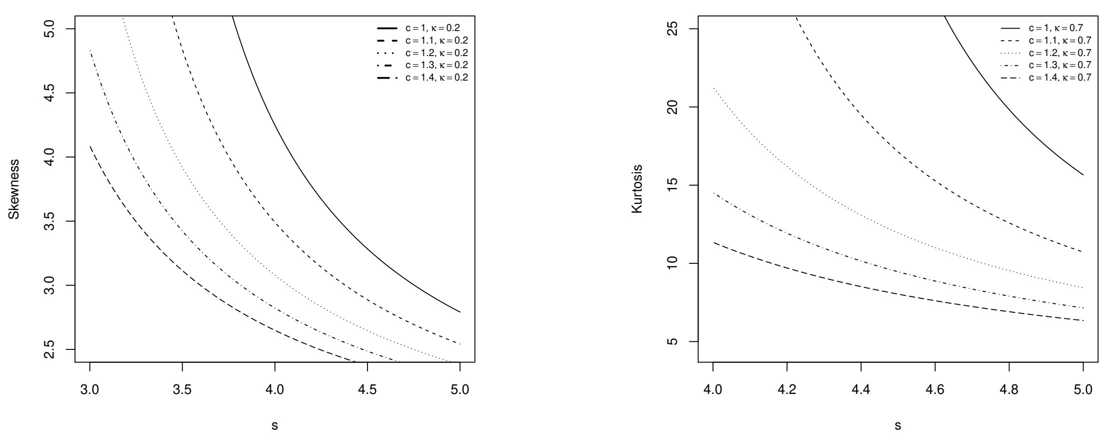
(a) (b)

Figure 5: Plots of the skewness and kurtosis coefficients of the Tan-BXII distribution as a function of c for selected values of κ and s



(a) (b)

Figure 6: Plots of the skewness and kurtosis coefficients of the Tan-BXII distribution as a function of κ for selected values of c and s



(a) (b)

Figure 7: Plots of the skewness and kurtosis coefficients of the Tan-BXII distribution as a function of s for selected values of c and κ

3.7 Entropy

By applying (2.4), the Rényi entropy is given by

$$\mathfrak{L}_G(\gamma) = \frac{1}{1-\gamma}\left\{\gamma \ln\left(\frac{\pi}{4}\right) + \ln\left[\sum_{k=0}^{+\infty}\sum_{r=0}^{k} a_k s^r (-1)^{k-r} s_0^{k-r} I_r\right]\right\},$$

where $\gamma > 0$ with $\gamma \neq 1$ and, after some algebra,

$$\begin{aligned} I_r &= \int_{-\infty}^{+\infty} G^r(x) g^\gamma(x) dx \\ &= \sum_{j=0}^{r}\binom{r}{j}(-1)^j \kappa^\gamma s^{-(\gamma-1)} c^{\gamma-1} B(\kappa(j+\gamma)+(\gamma-1)c^{-1}, (\gamma-1)(c-1)c^{-1}+1), \end{aligned}$$

assuming that $\kappa\gamma + (\gamma-1)c^{-1} > 0$ and $(\gamma-1)(c-1)c^{-1}+1 > 0$.

Figure 8 displays this Rényi entropy for some values of the parameters.

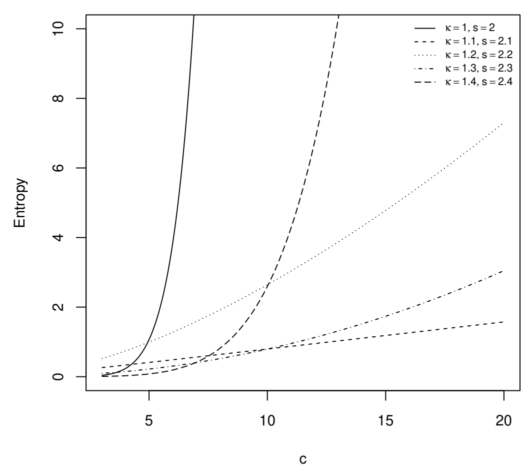


Figure 8: Plots of the Rényi entropy of the Tan-BXII distribution as a function of c for selected values of κ and s

3.8 Maximum likelihood estimation

Here, we provide the mathematical background related to the MLEs of the Tan-BXII model parameters, i.e., c, κ and s. Let $\mathbf{x} = \{x_1, \ldots, x_n\}^\top$ be n independent random variables from the Tan-BXII distribution. Then, the log-likelihood function is given by

$$\begin{aligned} L = {}& n\ln\left(\frac{\pi}{4}\right) + n\ln(c) + n\ln(\kappa) - nc\ln(s) + (c-1)\sum_{i=1}^{n}\ln(x_i) \\ & - (\kappa+1)\sum_{i=1}^{n}\ln\left[1+\left(\frac{x_i}{s}\right)^c\right] + 2\sum_{i=1}^{n}\ln\left[\sec\left\{\frac{\pi}{4}\left(1-\left[1+\left(\frac{x_i}{s}\right)^c\right]^{-\kappa}\right)\right\}\right]. \end{aligned}$$

The scores are presented below:

$$U_c = \frac{n}{c} - n\ln(s) + \sum_{i=1}^{n} \ln(x_i) - (\kappa+1)\sum_{i=1}^{n} \frac{x_i^c \ln\left(\frac{x_i}{s}\right)}{s^c + x_i^c}$$
$$+ \frac{\pi}{2}\kappa \sum_{i=1}^{n} \left(\frac{x_i}{s}\right)^c \ln\left(\frac{x_i}{s}\right) \left[1+\left(\frac{x_i}{s}\right)^c\right]^{-\kappa-1} \tan\left\{\frac{\pi}{4}\left(1-\left[1+\left(\frac{x_i}{s}\right)^c\right]^{-\kappa}\right)\right\},$$

$$U_\kappa = \frac{n}{\kappa} - \sum_{i=1}^{n} \ln\left[1+\left(\frac{x_i}{s}\right)^c\right]$$
$$+ \frac{\pi}{2} \sum_{i=1}^{n} \left[1+\left(\frac{x_i}{s}\right)^c\right]^{-\kappa} \ln\left[1+\left(\frac{x_i}{s}\right)^c\right] \tan\left\{\frac{\pi}{4}\left(1-\left[1+\left(\frac{x_i}{s}\right)^c\right]^{-\kappa}\right)\right\}$$

and

$$U_s = -\frac{nc}{s} + c(\kappa+1)s^{-1} \sum_{i=1}^{n} \frac{x_i^c}{s^c + x_i^c}$$
$$- \frac{\pi}{2} c\kappa s^{-(c+1)} \sum_{i=1}^{n} x_i^c \left[1+\left(\frac{x_i}{s}\right)^c\right]^{-\kappa-1} \tan\left\{\frac{\pi}{4}\left(1-\left[1+\left(\frac{x_i}{s}\right)^c\right]^{-\kappa}\right)\right\}.$$

The MLEs of c, κ and s are defined by the simultaneous solutions of the following non-linear equations: $U_c = 0$, $U_\kappa = 0$ and $U_s = 0$ according to c, κ and s. Under some standard regularity conditions, the well-known theory on MLE can be applied, ensuring nice asymptotic properties (see [3]).

3.9 Simulation

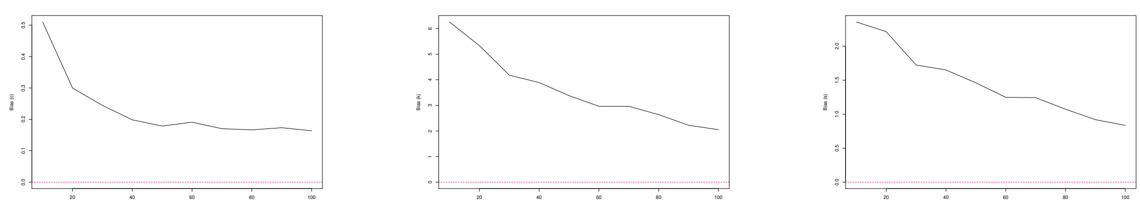
Using the `TanB` R package [17], we perform a simulation study using several random samples of the Tan-BXII distribution. For each sample, we calculate the MLEs using native R language's `optim` implementation. Biases, and Mean Square Errors (MSEs) are also calculated using the MLEs obtained.

For this simulation, we use samples with sizes $10, 20, 30, \ldots, 100$ and 1000 replicas for the parameter's configuration: $c = 1$, $\kappa = 1.4$ and $s = 0.15$. Figures 9a, 9b and 9c show the bias for c, κ and s, respectively, in this simulation and we can see it decreasing over the sample sizes. Figures 10a, 10b and 10c show the MSE for the same parameters and also decreases over the sample sizes.

Table 2 summarizes the simulation, given the means of MLEs, biases and MSEs of the samples with sizes of 10, 20, 30, 50 and 100. We can see in the table that all the parameters are overestimated by the maximum likelihood method. The biases and MSEs decrease over the sample sizes as we see in Figures 9a, 9b, 9c, 10a, 10b and 10c.

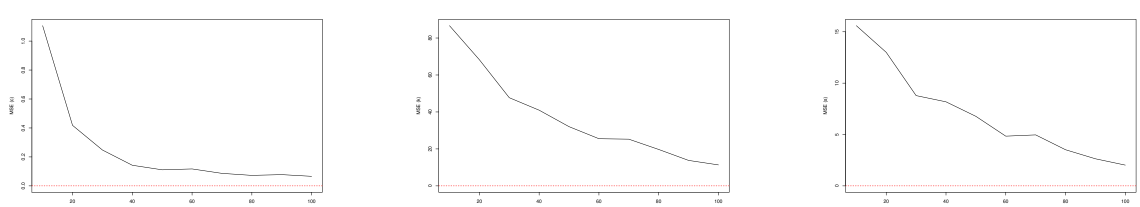
Table 2: MLEs, Biases and MSEs for $c = 1$, $\kappa = 1.4$, $s = 0.15$ using 1000 replicas

Sample size(n)	Parameters	MLEs	Biases	MSEs
10	c	1.5102	0.5102	1.1065
	κ	7.6587	6.2587	86.6797
	s	2.5062	2.3562	15.5951
20	c	1.2998	0.2998	0.4181
	κ	6.7327	5.3327	68.2502
	s	2.3631	2.2131	12.9993
30	c	1.2444	0.2444	0.2478
	κ	5.5806	4.1806	47.7063
	s	1.8732	1.7232	8.7874
50	c	1.1787	0.1787	0.111
	κ	4.7807	3.3807	32.0412
	s	1.6109	1.4609	6.7689
100	c	1.1636	0.1636	0.066
	κ	3.4506	2.0506	11.3414
	s	0.9844	0.8344	2.0205



(a) Plots of bias(c) (b) Plots of bias(κ) (c) Plots of bias(s)

Figure 9: Plots of the biases for the simulated experiment related to the Tan-BurXII model parameters



(a) Plots of MSE(c) (b) Plots of MSE(κ) (c) Plots of MSE(s)

Figure 10: Plots of the MSEs for the simulated experiment related to the Tan-BurXII model parameters

3.10 Application

Now, we apply the Tan-BXII model to fit a practical data set and compare it with three other models, namely Kum-BXII, BurrXII and Kum-W models. These data are on the Aircraft windshield failures (thousands of hours) reported in Murthy [12] (see Table 3). A brief statistical description of these data can be found in Table 4. Table 5 shows the MLEs of the parameters of the Tan-BXII, Kum-BXII, BurrXII and Kum-W models with error in parentheses, as well as the related Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC), Cramér-von Mises (W^*) and Anderson-Darling (A^*) statistics. We refer to [1], [6] and the book of [9] for precise definitions and use of these fundamental statistical tools.

Table 3: Data on aircraft windshield failures (thousands of hours)

0.040	1.866	2.385	3.443	0.301	1.876	2.481	3.467	0.309	1.899	2.610
3.478	0.557	1.911	2.625	3.578	0.943	1.912	2.632	3.595	1.070	1.914
2.646	3.699	1.124	1.981	2.661	3.779	1.248	2.010	2.688	3.924	1.281
2.038	2.823	4.035	1.281	2.085	2.890	4.121	1.303	2.089	2.902	4.167
1.432	2.097	2.934	4.240	1.480	2.135	2.962	4.255	1.505	2.154	2.964
4.278	1.506	2.190	3.000	4.305	1.568	2.194	3.103	4.376	1.615	2.223
3.114	4.449	1.619	2.224	3.117	4.485	1.652	2.229	3.166	4.570	1.652
2.300	3.344	4.602	1.757	2.324	3.376	4.663				

Table 4: Descriptive statistics of the considered data

Min.	Q_1	Median	Mean	Q_3	Max.	Var.
0.040	1.839	2.354	2.557	3.393	4.663	1.252

Table 5: MLEs of the parameters of the Tan-BXII, Kum-BXII, Kum-W and BurrXII models, with errors in parentheses, and AIC, BIC, CAIC, W^* and A^* statistics

Models	Estimates					AIC	BIC	CAIC	W^*	A^*
Tan-BXII(c, κ, s)	2.27	186.02	26.00	—	—	267.76	275.09	268.06	0.06	0.58
	(0.20)	(659.52)	(41.42)	—	—					
Kum-BXII(a, b, c, d, k)	0.28	1.96	7.17	4.54	5.82	267.95	280.17	268.71	0.08	0.64
	(0.11)	(1.36)	(2.38)	(5.07)	(1.46)					
Kum-W(a, b, c, β)	0.38	8.53	5.78	0.13	—	268.82	278.59	269.32	0.06	0.56
	(0.04)	(6.89)	(0.06)	(0.04)	—					
BXII(a, c, k)	2.48	11.31	7.47	—	—	270.24	277.57	270.54	0.06	0.63
	(0.23)	(8.05)	(2.57)	—	—					

It follows from Table 5 that, when compared to other ones, the Tan-BXII model is the best. We illustrate this claim by showing the fits of the estimated pdfs and cdfs in Figures 11 and 12, respectively. Thus, we conclude that the Tan-BXII distribution is quite flexible in the modeling of the proposed data.

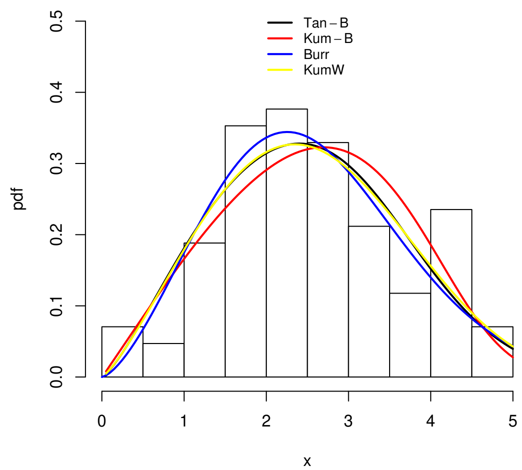


Figure 11: Some fitted pdfs of the data

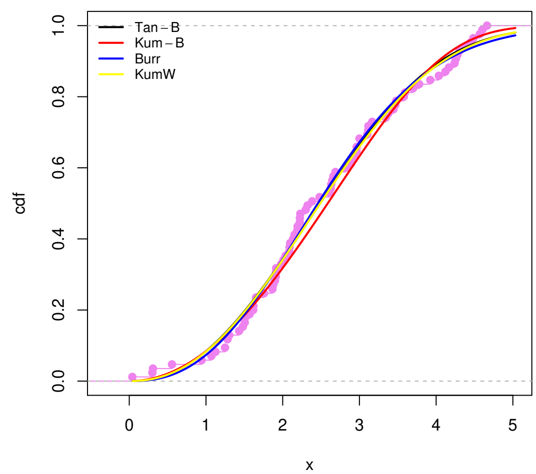


Figure 12: Some fitted cdfs of the data

4 Concluding remarks

In this paper, we introduced and discussed a new class of trigonometric distributions, called the Tan-G class, with a focus on a new lifetime trigonometric distribution of the class, called the Tan-BXII distribution. We obtain probability density function, cumulative distribution function, hazard function and various moments. The entropy is also calculated. A complete part is devoted to the estimation of the model parameters via the maximum likelihood method. We put the light on the applicability of the new related models by considering a practical data set. Even though our class of distributions does not optimally fit the data presented, it still proves to be a powerful tool for statistical analysis. We will apply this distribution to other data sets to show its full power and it will be reported elsewhere.

Acknowledgments

We would like to thank the reviewer and the associated editor for constructive comments on the article, improving it on several important aspects.

References

[1] T. W. Anderson and D. A. Darling, "A Test of Goodness-of-Fit", Journal of the American Statistical Association, vol. 49, pp. 765–769, 1954.

[2] C. C. R. Brito, "Método Gerador de Distribuicoes e Classes de Distribuicoes Probabilisticas", Tese de doutorado (Doutorado em Biometria e Estatistica Aplicada), Universidade Federal Rural de Pernambuco, Recife, 2014.

[3] G. Casella, and R. L. Berger, *Statistical Inference*, Brooks/Cole Publishing Company, California, 1990.

[4] C. Chesneau, H. S. Bakouch, and T. Hussain, "A new class of probability distributions via cosine and sine functions with applications", Communications in Statistics - Simulation and Computation, vol. 48, no. 8, pp. 2287–2300, 2019.

[5] G. M. Cordeiro, and M. de Castro, "A new family of generalized distributions", Journal of Statistical Computation and Simulation, vol. 81, no. 7, pp. 883–893, 2011.

[6] A. Darling, "The Kolmogorov-Smirnov, Cramer-von Mises tests", Annals of Mathematical Statistics, vol. 28, no 4, pp. 823–838, 1957.

[7] R. D. Gupta, and D. Kundu, "Exponentiated exponential family: an alternative to gamma and Weibull distributions", Biometrical Journal, vol. 43, no. 1, pp. 117–130, 2001.

[8] F. Jamal, and C. Chesneau, "A new family of polyno-expo-trigonometric distributions with applications", Infinite Dimensional Analysis, Quantum Probability and Related Topics, vol. 22, no. 04, 1950027, pp. 1–15, 2019.

[9] S. Konishi, and G. Kitagawa, *Information Criteria and Statistical Modeling.* Springer, New York, 2007.

[10] D. Kumar, U. Singh, and S. K. Singh, "A new distribution using sine function: its application to bladder cancer patients data", Journal of Statistics Applications and Probability, vol. 4, no. 3, pp. 417–427, 2015.

[11] Z. Mahmood, C. Chesneau, and M. H. Tahir, "A new sine-G family of distributions: properties and applications", Bulletin of Computational Applied Mathematics, vol. 7, no. 1, pp. 53–81, 2019.

[12] D. N. P. Murthy, M. Xie, and R. Jiag, *Weibull Models*, John Wiley and Sons, Inc. Hoboken, New Jersey, 2004.

[13] A. Rényi, "On measures of entropy and information", In: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, vol. 1, pp. 547–561, 1961.

[14] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, 2012.

[15] C. E. Shannon, "Prediction and entropy of printed English", The Bell System Technical Journal, vol. 30, no. 1, pp. 50–64, 1951.

[16] L. Souza, "New trigonometric classes of probabilistic distributions", Thesis, Universidade Federal Rural de Pernambuco, 2015.

[17] L. Souza, L. Gallindo, and L. Serafim-de-Souza, (2016). *TanB*: The TanB Distribution. R package version 0.2. Available at `https://cran.r-project.org/web/packages/TanB/index.html` or by running `install.packages("TanB");library("TanB");help("rtanb")` `inside R([14])`.

[18] 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, vol. 7, no. 3, pp. 357–379, 2019.

[19] 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, vol. 2, no. 2, pp. 63–79, 2019.