

THE TRANSMUTED HALF-NORMAL DISTRIBUTION WITH APPLICATION TO PRECIPITATION DATA

André Menezes^{1*}, Josmar Mazucheli²,
Jean Cardoso³ and Subrata Chakraborty⁴

Received November 23, 2018 / Accepted March 28, 2020

ABSTRACT. The Half-Normal distribution has been intensively extended in the recent years. A review of the literature showed that at least 10 extensions of the Half-Normal distribution were introduced between 2008 and 2016. These extensions generalized the behavior of the density and hazard functions, which are restricted to monotonous decreasing and monotonically increasing, respectively. In this paper we propose a new extension called the transmuted Half-Normal distribution using the quadratic rank transmutation map, introduced by Shaw & Buckley (2009). A comprehensive account of mathematical properties of the new distribution is presented. We provide explicit expressions for the moments, moment-generating function, Shannon's entropy, mean deviations, Bonferroni and Lorenz curves, order statistics, and reliability. The estimation of the parameters is implemented by the maximum likelihood method. The bias and accuracy of the estimators are assayed by the Monte Carlo simulations. This proposed distribution allows us to incorporate covariates directly in the mean and consequently to quantify their influences on the average of the response variable. Experiment with two real data sets show usefulness and its value as a good alternative to several extensions of the Half-Normal distribution in data modeling with and without covariates.

Keywords: Half-Normal distribution, moments, parametric estimation, precipitation data, transmutation.

1 INTRODUCTION

Underlying any parametric inference procedure a probability distribution is used to describe the behavior of a random variable in the population. Over the last years, an uncountable number of probability distributions, mostly having support of positive real numbers, have been proposed in

*Corresponding author

¹Department of Statistics, Universidade Estadual de Maringá, DEs, PR, Brazil – E-mail: andrefelipemaringa@gmail.com – <https://orcid.org/0000-0002-3320-9834>

²Department of Statistics, Universidade Estadual de Maringá, DEs, PR, Brazil – E-mail: jmazucheli@gmail.com – <https://orcid.org/0000-0001-6740-0445>

³Department of Statistics, Universidade Estadual de Maringá, DEs, PR, Brazil – E-mail: jeancarlos.card@gmail.com – <https://orcid.org/0000-0002-1539-1564>

⁴Department of Statistics, Dibrugarh University, India – E-mail: subrata_stats@dibru.ac.in – <https://orcid.org/0000-0002-6405-1486>

the literature. Many strategies can be used to generate or extend a probability distribution. Most of these strategies adds one or more parameters to some basic distribution (Normal, Gumbel, Exponential, Weibull, Range, Log-Normal, among many others). In general, introducing one or more parameters brings greater flexibility in the behavior of the density and hazard functions of the distributions. A recent survey of the main methods used to extend a baseline distribution is available, for example, in Nadarajah & Rocha (2016), Tahir & Nadarajah (2015), de Brito et al. (2015), Aljarrah et al. (2014), Lee et al. (2013), Lai (2011) and Gupta & Kundu (2009).

It is important to emphasize that the transformation of a random variable Z into another X , of the form $X = g(Z)$, is the simplest way to generate or extend a base probability distribution. An example of a distribution obtained by transformation is the one parameter Half-Normal (HN) distribution, resulting from the transformation $X = \theta |Z|$, where Z has standard Normal distribution and $\theta > 0$ is a scale parameter.

It should be mentioned here that the Half-Normal distribution is a special case of Nakagami- m distribution introduced by Nakagami (1960). In the recent years several extensions of the Half-Normal distribution were proposed. These extensions include: the general Half-Normal distribution (Pewsey, 2002, 2004), the generalized Half-Normal distribution (Cooray & Ananda, 2008), the Beta (log-Beta) generalized Half-Normal distribution (Pescim et al., 2010; Cordeiro et al., 2013; Pescim et al., 2013), the Kumaraswamy generalized Half-Normal distribution (Cordeiro et al., 2012), the Beta generalized Half-Normal geometric distribution (Ramires et al., 2013), the extension generalized Half-Normal distribution (Olmos et al., 2014), the Power Half-Normal distribution (Gómez & Bolfarine, 2015), the generalized Half-Normal extended (Sanchez et al., 2016), the Odd Log-Logistic generalized Half-Normal (Cordeiro et al., 2017) and the transmuted generalized Gamma distribution (Saboor et al., 2019).

This article introduces the transmuted Half-Normal distribution (THN), derived from the HN distribution. The extension of a baseline distribution via transmutation was proposed by Shaw & Buckley (2009). Tahir & Cordeiro (2016) have enumerated 51 distributions extended by transmutation. In Section 2 we present the HN distribution used as the baseline distribution to be extended. Its transmuted version is presented in Section 3. Various statistical and reliability properties of the THN are explored and discussed in Section 4. A few characterizations are considered in Section 5. Estimation by the maximum likelihood method is presented in Section 6. In Section 7, Monte Carlo simulations are conducted to investigate the bias and accuracy the maximum likelihood estimators. Two applications considering the proposed distribution are presented in Section 8. Section 9, with some concluding remarks, closes the paper.

2 THE HALF-NORMAL DISTRIBUTION

If a nonnegative random variable X has HN distribution with scale parameter $\theta > 0$, then the probability density function (p.d.f.) and the cumulative distribution function (c.d.f.) are written, respectively, as:

$$g(x | \theta) = \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \text{ and } G(x | \theta) = 2\Phi\left(\frac{x}{\theta}\right) - 1 \quad (1)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote, respectively, the p.d.f. and c.d.f. of distribution of a random variable with standard Normal distribution. The corresponding hazard rate function (h.r.f.) is written as:

$$h(x | \theta) = \frac{1}{\theta} \phi\left(\frac{x}{\theta}\right) \Phi\left(-\frac{x}{\theta}\right)^{-1}$$

which is monotonically increasing for every θ . It is worth remembering that p.d.f. is monotonically decreasing for all θ .

Formally, if a random variable Z is normally distributed with mean zero and variance one, then the HN distribution is the distribution of $X = \theta |Z|$. For Z normally distributed with mean μ and variance one, the transformation $X = \theta |Z|$ leads to Leone et al. (1961); Tsagris et al. (2014). The folded-Normal distribution is a special case of the Normal distribution whenever the sign of the measured variable is unknown, lost or not relevant. In other words, it is used when one is interested in the size of the measured variable and not in the direction or sign (Chou & Liu, 1998). It is clear that the HN distribution is a particular case of the folded-Normal distribution when $\mu = 0$, which also coincides with the normal distribution truncated at zero (Nadarajah & Kotz, 2006b). In addition, the expressions in (1) can be obtained with the particular case of the generalized Rayleigh distribution (Vodā, 1976a,b), of the generalized range distribution Stacy (1962) and the square root of a chi-square random variable with one degree of freedom (Johnson et al., 1994).

For $X \sim HN(\theta)$ the k -th moment of X is $E(X^k) = \frac{1}{\sqrt{\pi}} 2^{\frac{k}{2}} \theta^k \Gamma\left(\frac{k+1}{2}\right)$ such that for $k = 1$ we have $E(X) = \sqrt{\frac{2}{\pi}} \theta$ and for $k = 2$, $E(X^2) = \theta^2$. The third and fourth standardized moments (asymmetry and kurtosis, respectively) do not depend on θ and are written as:

$$\sqrt{\beta_1} = \frac{(4-\pi)\sqrt{2}}{(\pi-2)^{\frac{3}{2}}} = 0.9952 \quad \text{e} \quad \beta_2 = \frac{3\pi^2-4\pi-12}{(\pi-2)^2} = 3.8691.$$

The quantile function can be written in the form:

$$Q(p | \theta) = G^{-1}(p | \theta) = \theta \Phi^{-1}\left(\frac{p+1}{2}\right) \tag{2}$$

where $0 < p < 1$ and $\Phi^{-1}(\cdot)$ is the quantile function of a standard Normal distribution. Note that from (2) we can generate pseudo-random values of X . An alternative strategy can be found in Singh (1994) as a consequence of the application of the Box-Muller transformation Singh (1995).

3 THE TRANSMUTED HALF-NORMAL DISTRIBUTION

Motivated by the need for more versatile density and hazard functions, Shaw & Buckley (2009) proposed a strategy that has proved useful in extending a baseline distribution. The distribution obtained by this process is called the transmuted-baseline distribution, for example, transmuted Weibull distribution. In this section, from the transmutation procedure, we will introduce the transmuted HN distribution (THN).

From the strategy proposed by Shaw & Buckley (2009), a random variable X is said to have a transmuted distribution if its cdf is written in the form:

$$F(x | \theta, \lambda) = (1 + \lambda)G(x | \theta) - \lambda [G(x | \theta)]^2, \quad |\lambda| \leq 1.$$

Consequently:

$$f(x | \theta, \lambda) = g(x | \theta)[(1 + \lambda) - 2\lambda G(x | \theta)]$$

where $G(x | \theta)$ and $g(x | \theta)$ are respectively the c.d.f. and p.d.f. of the baseline distribution, indexed by a parameter vector θ . For $\lambda = 0$, we have the distribution of the base random variable as a particular case.

Therefore, a non-negative random variable X has THN distribution, with scale parameter $\theta > 0$ and $-1 \leq \lambda \leq 1$ if its respective p.d.f. and c.d.f. are given, respectively, by

$$f(x | \theta, \lambda) = \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \tag{3}$$

and

$$F(x | \theta, \lambda) = (1 + \lambda) \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right) - \lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)^2. \tag{4}$$

Figure 1 illustrates how the parameter λ influences the behavior of (3). Since θ is the scale parameter it is set to $\theta = 1$.

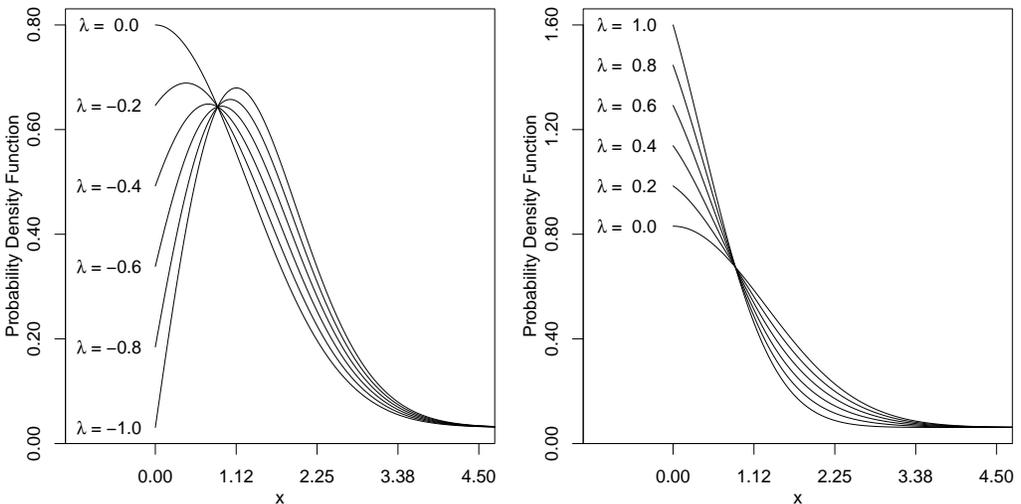


Figure 1 – Behavior of the p.d.f. of the THN distribution.

It can be checked that the cdf $F(x | \theta, \lambda)$ of $\text{THN}(\theta, \lambda)$ is a convex combination (finite mixture) of the cdf of maximum and minimum of two i.i.d $\text{HN}(\theta)$ random variables by writing $F(x | \theta, \lambda)$ as

$$F(x | \theta, \lambda) = \frac{1 + \lambda}{2} \left[2 \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right) - \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)^2\right] + \frac{1 - \lambda}{2} \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)^2.$$

Consequently the following result follows

$$\left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)^2 \leq F(x | \theta, \lambda) \leq 2\left(2\Phi\left(\frac{x}{\theta}\right) - 1\right) - \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)^2.$$

Therefore a THN(θ, λ) distribution is stochastically larger (smaller) than the distribution of the minimum (maximum) of two i.i.d. HN(θ) random variables.

4 STATISTICAL AND RELIABILITY PROPERTIES

In this section, we provide some important statistical and reliability properties of the THN distribution.

4.1 Survival, hazard rate and residual life functions

The survival function $S(x | \theta, \lambda)$ describes the probability of an item or individual surviving the time x . For the THN distribution we have:

$$S(x | \theta, \lambda) = 1 - (1 + \lambda) \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right) + \lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)^2. \tag{5}$$

The hazard rate function (h.r.f.) specifies the instantaneous rate of death or failure at time x , given that the individual has survived up to x . Mathematically we have $h(x | \theta, \lambda) = \frac{f(x|\theta,\lambda)}{S(x|\theta,\lambda)}$, consequently for THN

$$h(x | \theta, \lambda) = \frac{\frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) [1 + \lambda - 2\lambda (2\Phi\left(\frac{x}{\theta}\right) - 1)]}{1 - (1 + \lambda) (2\Phi\left(\frac{x}{\theta}\right) - 1) + \lambda (2\Phi\left(\frac{x}{\theta}\right) - 1)^2}. \tag{6}$$

Figure 2 illustrates how the parameter λ influences the behavior of (6). Since θ is the scale parameter it is set to $\theta = 1$.

The h.r.f. of THN(θ, λ) can be written in terms of that of HN(θ) as

$$\begin{aligned} h(x | \theta, \lambda) &= \frac{[1 + \lambda - 2\lambda (2\Phi\left(\frac{x}{\theta}\right) - 1)]}{1 - \lambda (2\Phi\left(\frac{x}{\theta}\right) - 1)} \frac{\frac{2}{\theta} \phi\left(\frac{x}{\theta}\right)}{2 - 2\Phi\left(\frac{x}{\theta}\right)} \\ &= \frac{[1 + \lambda - 2\lambda (2\Phi\left(\frac{x}{\theta}\right) - 1)]}{1 - \lambda (2\Phi\left(\frac{x}{\theta}\right) - 1)} h_{HN}(x | \theta), \end{aligned}$$

where $h_{HN}(x | \theta)$ is the hazard function of the HN(θ) distribution.

Using the above relation we can see that for all $0 \leq \lambda \leq 1$

$$h_{HN}(x | \theta) \leq h(x | \theta, \lambda) \leq (1 + \lambda) h_{HN}(x | \theta)$$

and for all $-1 \leq \lambda \leq 0$

$$(1 - \lambda) h_{HN}(x | \theta) \leq h(x | \theta, \lambda) \leq h_{HN}(x | \theta)$$

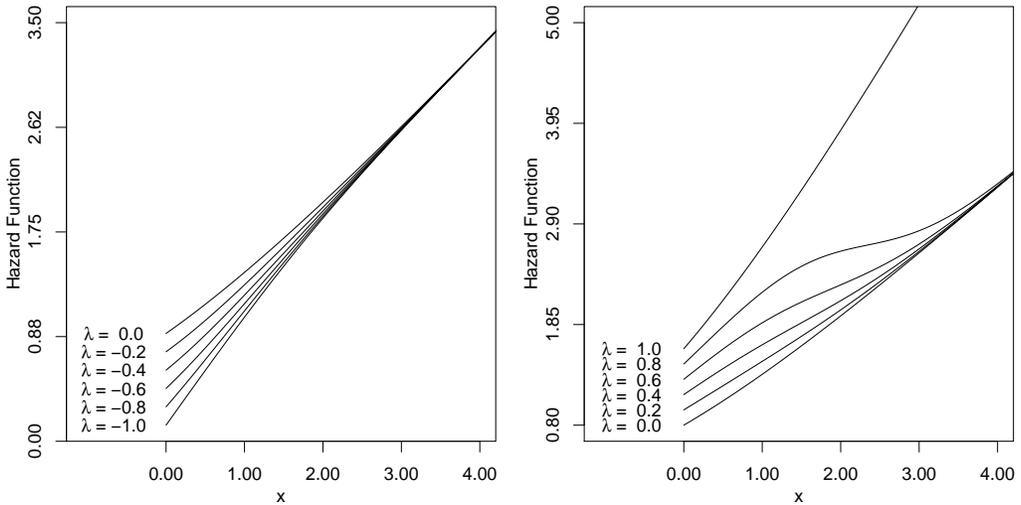


Figure 2 – Behavior of the h.r.f. of the THN distribution.

That is $HN \geq_{\text{hazard order}} THN$ for all $0 \leq \lambda \leq 1$ and $HN \leq_{\text{hazard order}} THN$ for all $-1 \leq \lambda \leq 0$.

In fact it can be shown that $THN \leq_{\text{likelihood order}} HN$ for all $0 \leq \lambda \leq 1$ and $THN \geq_{\text{likelihood order}} HN$ for all $-1 \leq \lambda \leq 0$.

Moreover if $X \sim THN(\theta, \lambda_1)$ and $Y \sim THN(\theta, \lambda_2)$ then $X \leq_{\text{likelihood order}} Y$ if $\lambda_2 < \lambda_1$.

The mean residual life for individuals at age t is the average remaining life time and corresponds to the ratio of the area under the survival curve to the right of t and $S(x | \theta, \lambda)$. From (5) we have:

$$m(t | \theta, \lambda) = \mathbb{E}(X - t | X > t) = \frac{1}{S(t | \theta, \lambda)} \int_t^\infty S(x | \theta, \lambda) dx. \tag{7}$$

Theorem 4.1. *The mean residual life of the THN distribution is given by:*

$$m(t | \theta, \lambda) = \frac{-\pi^{-\frac{1}{2}}}{S(t | \theta, \lambda)} \left\{ 2^{\frac{3}{2}} \theta \lambda e^{-\frac{t^2}{2\theta^2}} \gamma(t) - (\lambda + 1) \theta \sqrt{2} e^{-\frac{t^2}{2\theta^2}} - 2t \sqrt{\pi} \left[\Phi\left(\frac{t}{\theta}\right) - 1 \right] + \right. \\ \left. - 4\theta \lambda \left[\Phi\left(\frac{t\sqrt{2}}{\theta}\right) - 1 \right] + \lambda t \sqrt{\pi} \gamma(t) (\gamma(t) - 1) \right\}$$

where $\gamma(t) = 2\Phi\left(\frac{t}{\theta}\right) - 1$.

Proof. In fact, by replacing the equation (5) in (7), we obtain

$$m(t | \theta, \lambda) = \frac{1}{S(t | \theta, \lambda)} \int_t^\infty 1 - (1 + \lambda) \left[2\Phi\left(\frac{x}{\theta}\right) - 1 \right] + \lambda \left[2\Phi\left(\frac{x}{\theta}\right) - 1 \right]^2 dx.$$

Solving the integral, we obtain:

$$m(t | \theta, \lambda) = \frac{-\pi^{-\frac{1}{2}}}{S(t | \theta, \lambda)} \left\{ 2^{\frac{3}{2}} \theta \lambda e^{-\frac{t^2}{2\theta^2}} \left[2\Phi\left(\frac{t}{\theta}\right) - 1 \right] - 2\theta\lambda \left[2\Phi\left(\frac{t\sqrt{2}}{\theta}\right) - 1 \right] + \right. \\ \left. + \lambda t\sqrt{\pi} \left[2\Phi\left(\frac{t}{\theta}\right) - 1 \right]^2 + t\sqrt{\pi} - t\sqrt{\pi} \left[2\Phi\left(\frac{t}{\theta}\right) - 1 \right] + \right. \\ \left. - \lambda t\sqrt{\pi} \left[2\Phi\left(\frac{t}{\theta}\right) - 1 \right] - \lambda\theta\sqrt{2}e^{-\frac{t^2}{2\theta^2}} - \theta\sqrt{2}e^{-\frac{t^2}{2\theta^2}} + 2\lambda\theta \right\}.$$

From algebraic manipulations we arrive at the following expression:

$$m(t | \theta, \lambda) = \frac{-\pi^{-\frac{1}{2}}}{S(t | \theta, \lambda)} \left\{ 2^{\frac{3}{2}} \theta \lambda e^{-\frac{t^2}{2\theta^2}} \left[2\Phi\left(\frac{t}{\theta}\right) - 1 \right] - (\lambda + 1)\theta\sqrt{2}e^{-\frac{t^2}{2\theta^2}} - 2t\sqrt{\pi} \left[\Phi\left(\frac{t}{\theta}\right) - 1 \right] + \right. \\ \left. - 4\theta\lambda \left[\Phi\left(\frac{t\sqrt{2}}{\theta}\right) - 1 \right] + \lambda t\sqrt{\pi} \left[2\Phi\left(\frac{t}{\theta}\right) - 1 \right] \left[\left(2\Phi\left(\frac{t}{\theta}\right) - 1 \right) - 1 \right] \right\}.$$

Replacing $2\Phi\left(\frac{t}{\theta}\right) - 1 = \gamma(t)$, we get the result. □

4.2 Asymptotic behavior of the tails

The behavior of $\text{THN}(\theta, \lambda)$ in the tails for $x \rightarrow 0$ and for $x \rightarrow \infty$ can be stated respectively as follows:

$$f(x | \theta, \lambda) \sim (1 + \lambda) \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right),$$

$$F(x | \theta, \lambda) \sim (1 + \lambda) \left\{ 2\Phi\left(\frac{x}{\theta}\right) - 1 \right\},$$

$$h(x | \theta, \lambda) \sim (1 + \lambda) h_{HN}(x | \theta),$$

and

$$f(x | \theta, \lambda) \sim (1 - \lambda) \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right),$$

$$F(x | \theta, \lambda) \sim (1 + \lambda) \left\{ 2\Phi\left(\frac{x}{\theta}\right) - 1 \right\} - \lambda,$$

$$h(x | \theta, \lambda) \sim h_{HN}(x | \theta).$$

4.3 Quantile function

The p^{th} quantile of $\text{THN}(\theta, \lambda)$ is given by

$$Q(p | \theta, \lambda) = \theta \Phi^{-1} \left[\frac{1 + 3\lambda - \sqrt{(1 + \lambda)^2 - 4\lambda p}}{4\lambda} \right].$$

4.4 Moments and associated measures

Moments are measures capable of characterizing a probability distribution, with the first four moments indicative of central tendency, dispersion, asymmetry and kurtosis in that order. Unlike many extensions of the HN distribution, the k -th moment of the THN distribution can be obtained analytically.

Theorem 4.2. *If X has THN distribution then the k -th moment can be written as follows:*

$$E(X^k) = \frac{\theta^k 2^{\frac{k}{2}}}{\pi^{3/2}} \left\{ (\lambda + 1) \pi \Gamma\left(\frac{k+1}{2}\right) - 2\lambda k \sqrt{\pi} \text{Hgeo}\left[\frac{1}{2}, \frac{k}{2} + 1; \frac{3}{2}; -1\right] \Gamma\left(\frac{k}{2}\right) \right\}$$

where *Hgeo* is the Hypergeometric distribution Feller (1968).

Proof. From the moment of the order k of a continuous random variable definition:

$$\begin{aligned} E(X^k) &= \int_0^\infty x^k \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] dx \\ &= (1 + \lambda) \int_0^\infty \frac{2}{\theta} x^k \phi\left(\frac{x}{\theta}\right) dx - 2\lambda \int_0^\infty \frac{2}{\theta} x^k \phi\left(\frac{x}{\theta}\right) \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right) dx \end{aligned}$$

which can be written as:

$$E(X^k) = \frac{\theta^k 2^{\frac{k}{2}}}{\pi^{3/2}} \left\{ (\lambda + 1) \pi \Gamma\left(\frac{k+1}{2}\right) - 2\lambda k \sqrt{\pi} \text{Hgeo}\left[\frac{1}{2}, \frac{k}{2} + 1; \frac{3}{2}; -1\right] \Gamma\left(\frac{k}{2}\right) \right\}. \quad (8)$$

□

From (8) we have:

$$E(X) = \frac{\theta}{\sqrt{\pi}} \left[(1 + \lambda) \sqrt{2} - 2\lambda \right], \quad (9)$$

and

$$\text{Var}(X) = \frac{\theta^2}{\sqrt{\pi}} \left[(\pi - 2) + (4\sqrt{2} - 6)(\lambda - \lambda^2) \right].$$

The expressions of the coefficient of variation, skewness and kurtosis are obtained from the following relations:

$$\text{CV}(X) = \frac{\sqrt{\text{Var}(X)}}{E(X)} = \frac{\left[(\pi - 2) + (4\sqrt{2} - 6)(\lambda - \lambda^2) \right]^{1/2}}{\sqrt{2}(\lambda - 1) - 2\lambda},$$

$$\begin{aligned} \text{Skw}(X) &= \frac{E(X - E(X))^3}{\text{Var}(X)^{3/2}} \\ &= \left[\sqrt{2}(4 - \pi) + \lambda \left(\sqrt{2}(18 - \pi) + \pi - 24 \right) + \lambda^2 (42\sqrt{2} - 60) \right. \\ &\quad \left. + \lambda^3 (28\sqrt{2} - 40) \right] \times \left[(\pi - 2) + (4\sqrt{2} - 6)(\lambda - \lambda^2) \right]^{-3/2} \end{aligned}$$

and

$$\begin{aligned}
 Kur(X) &= \frac{E(X - E(X))^4}{Var(X)^2}, \\
 &= \left[(3\pi^2 - 4\pi - 12) + \lambda (4\pi (3\sqrt{2} - 4) - 48\sqrt{2} - 72) \right. \\
 &\quad + \lambda^2 (4\pi (3\sqrt{2} - 5) + 192\sqrt{2} - 264) + \lambda^3 (288\sqrt{2} - 408) \\
 &\quad \left. + \lambda^4 (144\sqrt{2} - 204) \right] \times \left[(\pi - 2) + (4\sqrt{2} - 6) (\lambda - \lambda^2) \right]^{-2}.
 \end{aligned}$$

4.5 Moment generating function

Theorem 4.3. *If X has THN distribution then the moment generating function of X , $M_X(t)$, is given by:*

$$\begin{aligned}
 M_X(t) &= -2\lambda \sum_{n=0}^{\infty} \frac{t^n}{n!} H_{geo} \left[\frac{1}{2}, \frac{n}{2} + 1; \frac{3}{2}; -1 \right] \Gamma\left(\frac{n}{2} + 1\right) \frac{\theta^n 2^{n/2+1}}{\pi} + \\
 &\quad + (1 + \lambda) \sum_{n=0}^{\infty} \frac{t^n}{n!} \Gamma\left(\frac{n+1}{2}\right) \frac{\theta^n 2^{n/2}}{\sqrt{\pi}}.
 \end{aligned}$$

Proof. Let X be a random variable with THN distribution, then the moment generating function of X is given by:

$$M_X(t) = \int_0^{\infty} e^{xt} \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1 \right) \right] dx. \tag{10}$$

From the Taylor series expansion of the function e^{tx} we have that (10) can be written in the form:

$$\begin{aligned}
 M_X(t) &= (1 + \lambda) \sum_{n=0}^{\infty} \frac{t^n}{n!} \int_0^{\infty} \frac{2}{\theta} x^n \phi\left(\frac{x}{\theta}\right) dx + \\
 &\quad - 2\lambda \sum_{n=0}^{\infty} \frac{t^n}{n!} \int_0^{\infty} \frac{2}{\theta} x^n \phi\left(\frac{x}{\theta}\right) \left(2\Phi\left(\frac{x}{\theta}\right) - 1 \right) dx.
 \end{aligned}$$

Solving integrals, we get:

$$\begin{aligned}
 M_X(t) &= -2\lambda \sum_{n=0}^{\infty} \frac{t^n}{n!} H_{geo} \left[\frac{1}{2}, \frac{n}{2} + 1; \frac{3}{2}; -1 \right] \Gamma\left(\frac{n}{2} + 1\right) \frac{\theta^n 2^{n/2+1}}{\pi} + \\
 &\quad + (1 + \lambda) \sum_{n=0}^{\infty} \frac{t^n}{n!} \Gamma\left(\frac{n+1}{2}\right) \frac{\theta^n 2^{n/2}}{\sqrt{\pi}}.
 \end{aligned} \tag{11}$$

□

4.6 Differential entropy

Here we investigate differential entropy of a continuous random variable. This is a measure of the uncertainty variation, and a large entropy value indicates a greater uncertainty in the data. One of the most popular measures is the Shannon entropy Shannon (1951).

The concept of Shannon's entropy refers to the uncertainty of a probability distribution. An important fact is that the entropy \mathcal{H}_{Sh} is not a function of the random variable X , but rather of the probability distribution of that variable.

Definition 4.1. *The differential entropy \mathcal{H}_{Sh} of a continuous random variable X with a probability density function $f(x)$ is defined as*

$$\mathcal{H}_{Sh}(X) = - \int_S f(x) \log f(x) dx,$$

where S is the support set of the random variable.

Theorem 4.4. *The Shannon's entropy for a continuous random variable X with THN distribution is given by:*

$$\begin{aligned} \mathcal{H}_{Sh}(X) = & -\log\left(\frac{2}{\theta}\right) + \frac{\pi(1 - \log 2\pi) - 2\lambda}{2\pi} + \\ & + 2\lambda^2 \{1 - \log[(1 + \lambda)(1 - \lambda)]\} - \lambda(\lambda^2 + 1) \log\left(\frac{1 + \lambda}{1 - \lambda}\right). \end{aligned}$$

Proof. In fact, it follows directly from (3) that

$$\begin{aligned} \mathcal{H}_{Sh}(X) = & - \int_0^\infty \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \times \\ & \times \log\left\{\frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right]\right\} dx. \end{aligned}$$

Using the distributive properties of the logarithmic function we have:

$$\begin{aligned} \mathcal{H}_{Sh}(X) = & - \int_0^\infty \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \times \\ & \times \left\{ \log\left(\frac{2}{\theta}\right) + \log\left(\phi\left(\frac{x}{\theta}\right)\right) + \log\left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \right\} dx \\ = & - \int_0^\infty \log\left(\frac{2}{\theta}\right) \left\{ \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \right\} + \\ & + \log\left(\phi\left(\frac{x}{\theta}\right)\right) \left\{ \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \right\} + \\ & + \log\left(1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right) \left\{ \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \right\} dx. \end{aligned}$$

Since the functions in question are all integrable, we have

$$\begin{aligned} \mathcal{H}_{Sh}(X) = & - \int_0^\infty \log\left(\frac{2}{\theta}\right) \left\{ \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \right\} dx + \\ & - \int_0^\infty \log\left(\phi\left(\frac{x}{\theta}\right)\right) \left\{ \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \right\} dx + \\ & - \int_0^\infty \log\left(1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right) \left\{ \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)\right] \right\} dx. \end{aligned}$$

Solving the integrals, we have that the Shannon’s entropy for the THN distribution is given by:

$$\begin{aligned} \mathcal{H}_{Sh}(X) = & - \log\left(\frac{2}{\theta}\right) + \frac{\pi(1 - \log 2\pi) - 2\lambda}{2\pi} + \\ & + 2\lambda^2 \{1 - \log[(1 + \lambda)(1 - \lambda)]\} - \lambda(\lambda^2 + 1) \log\left(\frac{1 + \lambda}{1 - \lambda}\right). \end{aligned}$$

□

4.7 Mean deviations

The amount of dispersion in a population can be measured by deviations around the mean and median, defined by:

$$\delta_1 = \int_0^\infty |x - \mu| f(x) dx \quad \text{and} \quad \delta_2 = \int_0^\infty |x - M| f(x) dx,$$

where μ and M denote the mean and median, respectively.

To calculate these measures we can use the following relations presented in Nadarajah & Kotz (2006a):

$$\delta_1 = 2\mu F(\mu) - 2 \int_0^\mu x f(x) dx \quad \text{and} \quad \delta_2 = \mu + 2MF(M) - M - 2 \int_0^M x f(x) dx,$$

where $F(\mu)$ and $F(M)$ can be calculated according to the equation (4). Taking the $f(x)$ as the p.d.f. of the THN distribution, we have:

$$\begin{aligned} \int_0^\mu x f(x | \theta, \lambda) dx = & \frac{\theta}{\sqrt{\pi}} \left\{ \sqrt{2}(1 + \lambda) + 2\lambda \left[2\Phi\left(\frac{\mu\sqrt{2}}{\theta}\right) - 1 \right] \right\} + \\ & - \theta^2 \left\{ \frac{2}{\theta} \phi\left(\frac{\mu}{\theta}\right) \left[1 + \lambda - 2\lambda \left(2\Phi\left(\frac{\mu}{\theta}\right) - 1 \right) \right] \right\}, \end{aligned}$$

In an analogous way we obtain $\int_0^M x f(x | \theta, \lambda) dx$.

4.8 Bonferroni and Lorenz curves

The Bonferroni and Lorenz curves proposed by Bonferroni (1930), commonly used in areas such as reliability, demography, economics, medicine and insurance, are applications of the mean

deviations and are considered by economists as a measure of social inequality, since they relate the accumulated percentages of income and population . The Bonferroni and Lorenz curves are defined as:

Definition 4.2. Suppose that X is a nonnegative random variable with probability density function $f(x)$ and cumulative distribution function $F(x)$. The Bonferroni and Lorenz curves denoted by $B(p)$ and $L(p)$, respectively, are defined as:

$$B(p) = \frac{1}{p\mu} \int_0^q x f(x) dx \quad e \quad L(p) = \frac{1}{\mu} \int_0^q x f(x) dx,$$

on what $\mu = \mathbb{E}(X)$, $q = F^{-1}(p)$ and $p \in [0, 1]$.

In particular, for the THN distribution we have:

Theorem 4.5. Let X be a nonnegative continuous random variable that has THN distribution. The Bonferroni and Lorenz curves are given by:

$$B(p) = \frac{-\theta^2}{p\mu} \left\{ \frac{2}{\theta} \phi \left(\frac{q}{\theta} \right) \left[1 + \lambda - 2\lambda \left(2\Phi \left(\frac{q}{\theta} \right) - 1 \right) \right] + \right. \\ \left. - \frac{1}{\theta\sqrt{\pi}} \left[\sqrt{2}(1 + \lambda) + 2\lambda \left(2\Phi \left(\frac{q\sqrt{2}}{\theta} \right) - 1 \right) \right] \right\}.$$

and

$$L(p) = \frac{-\theta^2}{\mu} \left\{ \frac{2}{\theta} \phi \left(\frac{q}{\theta} \right) \left[1 + \lambda - 2\lambda \left(2\Phi \left(\frac{q}{\theta} \right) - 1 \right) \right] + \right. \\ \left. - \frac{1}{\theta\sqrt{\pi}} \left[\sqrt{2}(1 + \lambda) + 2\lambda \left(2\Phi \left(\frac{q\sqrt{2}}{\theta} \right) - 1 \right) \right] \right\}.$$

Proof. In fact, by applying the THN distribution we have

$$B(p) = \frac{1}{p\mu} \int_0^q x \frac{2}{\theta} \phi \left(\frac{x}{\theta} \right) \left[1 + \lambda - 2\lambda \left(2\Phi \left(\frac{x}{\theta} \right) - 1 \right) \right] \\ = \frac{-\theta^2}{p\mu} \left[\frac{2}{\theta} \phi \left(\frac{q}{\theta} \right) 2\lambda \left(2\Phi \left(\frac{q}{\theta} \right) - 1 \right) - \lambda \frac{2}{\theta} \phi \left(\frac{q}{\theta} \right) + \right. \\ \left. - \frac{2}{\theta} \phi \left(\frac{q}{\theta} \right) - \frac{2\lambda}{\theta\sqrt{\pi}} \left(2\Phi \left(\frac{q\sqrt{2}}{\theta} \right) - 1 \right) + \frac{\sqrt{2}\lambda + \sqrt{2}}{\theta\sqrt{\pi}} \right] \\ = \frac{-\theta^2}{p\mu} \left\{ \frac{2}{\theta} \phi \left(\frac{q}{\theta} \right) \left[1 + \lambda - 2\lambda \left(2\Phi \left(\frac{q}{\theta} \right) - 1 \right) \right] + \right. \\ \left. - \frac{1}{\theta\sqrt{\pi}} \left[\sqrt{2}(1 + \lambda) + 2\lambda \left(2\Phi \left(\frac{q\sqrt{2}}{\theta} \right) - 1 \right) \right] \right\}.$$

□

In an analogous way we obtain the Lorenz curve.

4.9 Order statistics

Order statistics, as well as sample moments, play an important role in statistical inference (David & Nagaraja, 2003). Order statistics moments play an important role in quality control and reliability testing to predict the failure of future items based on the times of some initial failures.

Let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of a random sample X_1, \dots, X_n obtained from a population following a distribution $f(x)$, then the p.d.f. of the j -th order statistic is given by:

$$f_j(x) = \frac{n!}{(j-1)!(n-j)!} f(x) [F(x)]^{j-1} [1 - F(x)]^{n-j}$$

for $j = 1, \dots, n$.

Applying the equations (3) and (4) we obtain the density of the j -th order statistic of the THN.

$$\begin{aligned} f_j(x | \theta, \lambda) &= \frac{n!}{(j-1)!(n-j)!} \left\{ \frac{2}{\theta} \phi \left(\frac{x}{\theta} \right) \left[1 + \lambda - 2\lambda \left(2\Phi \left(\frac{x}{\theta} \right) - 1 \right) \right] \right\} \times \\ &\quad \times \left\{ (1 + \lambda) \left(2\Phi \left(\frac{x}{\theta} \right) - 1 \right) - \lambda \left(2\Phi \left(\frac{x}{\theta} \right) - 1 \right)^2 \right\}^{j-1} \times \\ &\quad \times \left\{ 1 - (1 + \lambda) \left(2\Phi \left(\frac{x}{\theta} \right) - 1 \right) + \lambda \left(2\Phi \left(\frac{x}{\theta} \right) - 1 \right)^2 \right\}^{n-j} \end{aligned}$$

Considering the substitution $\gamma(x) = 2\Phi \left(\frac{x}{\theta} \right) - 1$ we have:

$$\begin{aligned} f_j(x | \theta, \lambda) &= \frac{n!}{(j-1)!(n-j)!} \left[\frac{2}{\theta} \phi \left(\frac{x}{\theta} \right) (1 + \lambda - 2\lambda \gamma(x)) \right] \times \\ &\quad \times \left[(1 + \lambda) \gamma(x) - \lambda \gamma(x)^2 \right]^{j-1} \left[1 - (1 + \lambda) \gamma(x) + \lambda \gamma(x)^2 \right]^{n-j}. \end{aligned}$$

Reordering the terms we obtain the general formula for density of the j -th order statistic THN:

$$\begin{aligned} f_j(x | \theta, \lambda) &= \frac{n!}{(j-1)!(n-j)!} \frac{2}{\theta} \phi \left(\frac{x}{\theta} \right) \gamma(x)^{j-1} (1 + \lambda - 2\lambda \gamma(x)) \times \\ &\quad \times (1 + \lambda - \lambda \gamma(x))^{j-1} (\gamma(x)\lambda - 1)^{n-j} (\gamma(x) - 1)^{n-j}. \end{aligned}$$

We can use the THN distribution to model maximum or minimum events, so in what follows we obtain the densities of the maximum and the minimum order statistic. For the order statistic n , we have to replace $j = n$ in the general formula (12). Therefore, the n -th order statistic representing the distribution of the maximum of the THN distribution has the p.d.f. is given by:

$$f_n(x) = n \frac{2}{\theta} \phi \left(\frac{x}{\theta} \right) \gamma(x)^{n-1} (1 + \lambda - 2\lambda \gamma(x)) (1 + \lambda - \lambda \gamma(x))^{n-1}. \tag{12}$$

For first order statistic, we replace in the general formula $j = 1$. The result generates the p.d.f. of the minimum of the THN distribution given by:

$$f_1(x) = n \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right) (1 + \lambda - 2\lambda\gamma(x)) (\gamma(x)\lambda - 1)^{n-1} (\gamma(x) - 1)^{n-1}.$$

4.10 Extreme Values

The $\text{HN}(\theta)$ distribution belongs to the max domain of attraction of the Gumbel extreme value distribution. Hence, there must exist a strictly positive function, say $h(t)$, such that

$$\lim_{t \rightarrow \infty} \frac{1 - G(t + xh(t))}{1 - G(t)} = e^{-x},$$

for every $x \in (0, \infty)$ (Leadbetter et al., 1987).

Then it can be shown that

$$\lim_{t \rightarrow \infty} \frac{1 - F(t + xh(t))}{1 - F(t)} = e^{-ax},$$

for every $x \in (0, \infty)$. Therefore $\text{THN}(\theta, \lambda)$ belongs to the max domain of attraction of the Gumbel extreme value distribution with

$$\lim_{n \rightarrow \infty} P[a_n(X_{n:n} - b_n \leq x)] = \exp[-\exp(x)]$$

for some suitable norming constants $a_n > 0$ and $b_n \in \mathbb{R}$, where $X_{n:n}$ is the maximum order statistic.

4.11 Stress strength reliability

Stress strength reliability estimation is of great interest in engineering, being used in stress-force models or as a measure of performance in electrical and electronic systems. However, it can also be applied in other areas, as it allows a general measure of the differences between two populations Asgharzadeh et al. (2011).

Theorem 4.6. Let $X_1 \sim THN(\theta_1, \lambda_1)$ and $X_2 \sim THN(\theta_2, \lambda_2)$, are independent then

$$\begin{aligned}
 R &= P(X_2 < X_1) = \int_0^\infty f_1(x | \theta, \lambda) F_2(x | \theta, \lambda) dx \\
 &= (1 + \lambda_1)(1 + \lambda_2) \frac{4}{\theta_1 \pi} \sum_{j=1}^\infty \frac{(-1)^{j+1} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+1} \theta^{2j+2} 2^{-\frac{1}{2}+j} \Gamma(j+1)}{(2j-1)(j-1)!} + \\
 &\quad - (1 + \lambda_1) \lambda_2 \frac{8}{\theta_1 \pi^{\frac{3}{2}}} \sum_{j=1}^\infty \sum_{l=1}^\infty \frac{(-1)^{j+l+2} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+2l+2} \theta^{2j+2l+3} 2^{j+l} \Gamma(j+l+\frac{3}{2})}{(2j-1)(j-1)!(2l-1)(l-1)!} + \\
 &\quad - \lambda_1 (1 + \lambda_2) \frac{16}{\theta_1 \pi^{\frac{3}{2}}} \sum_{k=1}^\infty \sum_{j=1}^\infty \frac{(-1)^{k+j+2} \left(\frac{\sqrt{2}}{2\theta_1}\right)^{2k+1} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+1} \theta^{2k+2l+3} 2^{k+l} \Gamma(k+l+\frac{3}{2})}{(2k-1)(k-1)!(2j-1)(j-1)!} + \\
 &\quad + \lambda_1 \lambda_2 \frac{32}{\theta_1 \pi^2} \sum_{k=1}^\infty \sum_{j=1}^\infty \sum_{l=1}^\infty \frac{(-1)^{k+j+l+3} \left(\frac{\sqrt{2}}{2\theta_1}\right)^{2k+1} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+2l+2} \theta^{2k+2l+2j+4} 2^{\frac{1}{2}+k+l+j}}{(2k-1)(k-1)!(2j-1)(j-1)!(2l-1)(l-1)!} \times \\
 &\quad \times \Gamma(k+l+j+2).
 \end{aligned}$$

Proof. Using the p.d.f. and c.d.f. defined in (3) and (4) we have:

$$\begin{aligned}
 R = P(X_2 < X_1) &= \int_0^\infty \frac{2}{\theta_1} \phi\left(\frac{x}{\theta_1}\right) \left[1 + \lambda_1 - 2\lambda_1 \left(2\Phi\left(\frac{x}{\theta_1}\right) - 1\right)\right] \times \\
 &\quad \times \left[(1 + \lambda_2) \left(2\Phi\left(\frac{x}{\theta_2}\right) - 1\right) - \lambda_2 \left(2\Phi\left(\frac{x}{\theta_2}\right) - 1\right)^2 \right] dx.
 \end{aligned}$$

Replacing $2\Phi\left(\frac{x}{\theta}\right) - 1$ by $\frac{2}{\sqrt{\pi}} \sum_{k=1}^\infty \frac{(-1)^{k+1} \left(\frac{\sqrt{2}}{2\theta}\right)^{2k+1} x^{2k+1}}{(2k-1)(k-1)!}$, we have to

$$\begin{aligned}
 R &= \int_0^\infty \frac{2}{\theta_1} \phi\left(\frac{x}{\theta_1}\right) \left\{ 1 + \lambda_1 - 2\lambda_1 \left[\frac{2}{\sqrt{\pi}} \sum_{k=1}^\infty \frac{(-1)^{k+1} \left(\frac{\sqrt{2}}{2\theta_1}\right)^{2k+1} x^{2k+1}}{(2k-1)(k-1)!} \right] \right\} \times \\
 &\quad \times \left\{ (1 + \lambda_2) \left(2\Phi\left(\frac{x}{\theta_2}\right) - 1\right) - \lambda_2 \left[\frac{2}{\sqrt{\pi}} \sum_{j=1}^\infty \frac{(-1)^{j+1} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+1} x^{2j+1}}{(2j-1)(j-1)!} \right]^2 \right\} dx.
 \end{aligned}$$

Applying the distributive law, using properties of integration and isolating terms independent of x , we get:

$$\begin{aligned}
 R &= (1 + \lambda_1)(1 + \lambda_2) \frac{4}{\theta_1 \sqrt{\pi}} \sum_{j=1}^{\infty} \frac{(-1)^{j+1} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+1} \int_0^{\infty} \phi\left(\frac{x}{\theta_1}\right) x^{2j+1} dx}{(2j-1)(j-1)!} + \\
 &- (1 + \lambda_1)\lambda_2 \frac{8}{\theta_1 \pi} \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} \frac{(-1)^{j+l+2} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+2l+2} \int_0^{\infty} \phi\left(\frac{x}{\theta_1}\right) x^{2j+2l+2} dx}{(2j-1)(j-1)!(2l-1)(l-1)!} + \\
 &- \lambda_1(1 + \lambda_2) \frac{16}{\theta_1 \pi} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \frac{(-1)^{k+j+2} \left(\frac{\sqrt{2}}{2\theta_1}\right)^{2k+1} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+1} \int_0^{\infty} \phi\left(\frac{x}{\theta_1}\right) x^{2k+2j+2} dx}{(2k-1)(k-1)!(2j-1)(j-1)!} + \\
 &+ \lambda_1 \lambda_2 \frac{32}{\theta_1 \pi^{\frac{3}{2}}} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} \frac{(-1)^{k+j+l+3} \left(\frac{\sqrt{2}}{2\theta_1}\right)^{2k+1} \left(\frac{\sqrt{2}}{2\theta_2}\right)^{2j+2l+2} \int_0^{\infty} \phi\left(\frac{x}{\theta_1}\right) x^{2k+2j+2l+3} dx}{(2k-1)(k-1)!(2j-1)(j-1)!(2l-1)(l-1)!}
 \end{aligned}$$

On solving the integrals we arrive at the desired result. □

Note that for the case where $\theta_1 = \theta_2$, it immediately follows that

$$R = \frac{1}{6}(\lambda_2 - \lambda_1) + \frac{1}{2}, \quad 1/3 < R < 2/3.$$

5 CHARACTERIZATION

Here we discuss characterizations of $THN(\theta, \lambda)$ distribution based on: (i) simple relationship between two truncated moments and (ii) maximum and minimum order statistics.

5.1 Characterizations by two truncated moments

Here we present a characterization of $THN(\theta, \lambda)$ distribution based on a simple relationship between two truncated moments using the following theorem due to Glänzel (1987) and used in Hamedani et al. (2017) and Yousof et al. (2017). This theorem also holds if the interval I is not closed and also when the *cdf* F is not in a compact form. This characterization is stable in the sense of weak convergence.

Theorem 5.1. *Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $I = [l, u]$ be an interval for some $l < u$ ($l = -\infty, u = \infty$ might as well be allowed). Let $X : \Omega \rightarrow I$ be a continuous random variable with the *cdf* F and let w_1 and w_2 be two real valued functions defined on I such that*

$$\mathbf{E}[w_2(X) \mid X \geq x] = \mathbf{E}[w_1(X) \mid X \geq x] \xi(x), \quad x \in I,$$

is defined for some real function ξ . Assume that $w_1, w_2 \in C^1(I)$, $\xi \in C^2(H)$ and F are twice continuously differentiable and strictly monotone function on H . More over, assume that the

equation $\xi w_1 = w_2$ has no real solution in the interior of I . Then F is uniquely determined by the functions w_1, w_2 and ξ , particularly

$$F(x) = \int_I^x K \left| \frac{\xi'(u)}{\xi(u) w_1(u) - w_2(u)} \right| \exp(-s(u)) du,$$

where the function s is a solution of the differential equation $s'(x) = \frac{\xi'(x) w_1(x)}{\xi(x) w_1(x) - w_2(x)}$ and K is the normalization constant, such that $\int_I dF = 1$.

Proposition 5.1. *Let $X : \Omega \rightarrow (0, \infty)$ be a continuous random variable and let $w_1(x) = 1 + \lambda - 2\lambda (2\Phi(\frac{x}{\theta}) - 1)$ and $w_2(x) = w_1(x) \{2\Phi(\frac{x}{\theta}) - 1\}$ for $x \in (0, \infty)$. Then the random variable X follows $THN(\theta, \lambda)$ with p.d.f. in (3) if and only if the function ξ defined in Theorem 5.1 has the form*

$$\xi(x) = \Phi\left(\frac{x}{\theta}\right), \quad x \geq 0.$$

Proof. Let X be a random variable with pdf (3), then

$$(1 - F(x)) E[w_1(X) | X \geq x] = \left[1 - \left\{ 2\Phi\left(\frac{x}{\theta}\right) - 1 \right\} \right], \quad x \geq 0,$$

and

$$(1 - F(x)) E[w_2(X) | X \geq x] = \left[\frac{1}{2} - \frac{1}{2} \left\{ 2\Phi\left(\frac{x}{\theta}\right) - 1 \right\}^2 \right], \quad x \geq 0,$$

hence

$$\xi(x) = \frac{E[w_2(X) | X \geq x]}{E[w_1(X) | X \geq x]} = \Phi\left(\frac{x}{\theta}\right), \quad x \geq 0.$$

Conversely, if $\xi = \Phi\left(\frac{x}{\theta}\right)$, then

$$s'(x) = \frac{\frac{1}{\theta} \phi\left(\frac{x}{\theta}\right) w_1(x)}{\Phi\left(\frac{x}{\theta}\right) w_1(x) - w_2(x) \{2\Phi\left(\frac{x}{\theta}\right) - 1\}} = \frac{d}{dx} \left\{ -\log \left(1 - \Phi\left(\frac{x}{\theta}\right) \right) \right\}, \quad x \geq 0,$$

and hence

$$s(x) = -\log \left(1 - \Phi\left(\frac{x}{\theta}\right) \right), \quad x \geq 0.$$

Now, in view of Theorem 5.1, X has density (2). □

Corollary 5.1. *Let $X : \Omega \rightarrow (0, \infty)$ be a continuous random variable and let $w_1(x)$ be as in Proposition 5.1. Then pdf of X follows $THN(\theta, \lambda)$ with p.d.f. in (3) if and only if there exist functions w_2 and ξ defined in Theorem 5.1 satisfying the differential equation*

$$\frac{\xi'(x) w_1(x)}{\xi(x) w_1(x) - w_2(x)} = \frac{\frac{1}{\theta} \phi\left(\frac{x}{\theta}\right)}{1 - \Phi\left(\frac{x}{\theta}\right)} \quad x \geq 0.$$

Remark: The general solution of the differential equation in Corollary 5.1 is

$$\xi(x) = \left[1 - \Phi\left(\frac{x}{\theta}\right)\right]^{-1} \left[-\int \frac{\frac{1}{\theta}\phi\left(\frac{x}{\theta}\right)}{\{1 - \Phi\left(\frac{x}{\theta}\right)\}} dx + C\right],$$

where C is a constant. The set of functions given in Proposition 5.1 satisfies this differential equation when $C = 0$. It need to be noted that there are other triplets (w_1, w_2, ξ) satisfying the conditions of Theorem 5.1.

5.2 Characterizations by order statistics

Here we use two results from Hamedani et al. (2017) stated in the theorem below to characterize THN(θ, λ) distribution by maximum $(X_{n:n})$ and minimum $(X_{1:n})$ order statistics from a sample of size n from THN(θ, λ) distribution.

Theorem 5.2. Let $X : \Omega \rightarrow (0, \infty)$ be a continuous random variable with cdf F and let $\xi(x)$ and $w(x)$ be two differentiable functions in $(0, \infty)$ such that

(i) $\lim_{x \rightarrow 0} \xi(x) F(x)^n = 0,$

$$\int_0^\infty \frac{w'(t)}{\xi(t) - w(t)} dt = \infty,$$

then

$$\mathbf{E}[\xi(X_{n:n}) \mid X_{n:n} < t] = w(t), \quad t > 0,$$

implies

$$F(x) = \exp\left[\int_x^\infty \frac{w'(t)}{n[\xi(t) - w(t)]} dt\right], \quad x \geq 0$$

(ii) $\lim_{x \rightarrow \infty} \xi(x) [1 - F(x)]^n = 0,$

$$\int_0^\infty \frac{w'(t)}{w(t) - \xi(t)} dt = \infty,$$

then

$$\mathbf{E}[\xi(X_{1:n}) \mid X_{1:n} > t] = w(t), \quad t > 0,$$

implies

$$F(x) = 1 - \exp\left[\int_0^x \frac{w'(t)}{n[w(t) - \xi(t)]} dt\right], \quad x \geq 0.$$

Proposition 5.2. *If we consider $\xi(x) = \left\{ (1 + \lambda) (2\Phi(\frac{x}{\theta}) - 1) - \lambda (2\Phi(\frac{x}{\theta}) - 1)^2 \right\}^n$ and $w(x) = \frac{1}{2}\xi(x)$ in Theorem 5.2(i). Then the random variable X follows THN(θ, λ) with c.d.f. in (4).*

Proposition 5.3. *If we consider $\xi(x) = \left\{ 1 - (1 + \lambda) (2\Phi(\frac{x}{\theta}) - 1) - \lambda (2\Phi(\frac{x}{\theta}) - 1)^2 \right\}^n$ and $w(x) = \frac{1}{2}\xi(x)$ in Theorem 5.2(ii). Then the random variable X follows THN(θ, λ) with c.d.f. in (4).*

6 MAXIMUM LIKELIHOOD ESTIMATOR

Let $x = (x_1, \dots, x_n)^\top$ be a random sample from the THN distribution with p.d.f. expressed by (3). Then, the log-likelihood function, a part from constant terms, can be written as:

$$\ell(\theta, \lambda | x) \propto -n \log(\theta) + \sum_{i=1}^n \log \left[\phi \left(\frac{x_i}{\theta} \right) \right] + \sum_{i=1}^n \log \left\{ 1 + \lambda - 2\lambda \left[2\Phi \left(\frac{x_i}{\theta} \right) \right] \right\}. \quad (13)$$

The MLEs $\hat{\theta}$ and $\hat{\lambda}$ of θ and λ , respectively may be obtained by maximization of (13), or solving the following likelihood equations:

$$\begin{aligned} \frac{\partial \ell(\theta, \lambda | x)}{\partial \lambda} &= \sum_{i=1}^n \frac{-4\Phi(\frac{x_i}{\theta}) + 3}{-4\lambda\Phi(\frac{x_i}{\theta}) + 3\lambda + 1} \\ \frac{\partial \ell(\theta, \lambda | x)}{\partial \theta} &= -\frac{n}{\theta} - \frac{1}{\theta^2} \sum_{i=1}^n \frac{\phi'(\frac{x_i}{\theta})}{\phi(\frac{x_i}{\theta})} x_i + \frac{4\lambda}{\theta^2} \sum_{i=1}^n \frac{\phi(\frac{x_i}{\theta})}{-4\lambda\Phi(\frac{x_i}{\theta}) + 3\lambda + 1} x_i \end{aligned}$$

where $\phi'(\frac{x_i}{\theta}) = \frac{d}{d\theta} \phi(\frac{x_i}{\theta})$.

7 SIMULATION STUDY

In this section we present the results of a Monte Carlo simulation used to evaluate the bias and mean square error of the estimates obtained by the maximum likelihood method. Samples of size $n = 20, 50, \dots, 170, 200$ with $\lambda = -0.9, -0.7, \dots, 0.7, 0.9$ and $\theta = 1.0$ were generated. For each of the combinations of n, θ and λ the inverse transformation method was used to generate, $N = 10.000$ pseudo-random samples of the THN. The results are reported in the Figures 3 and 4.

In analyzing the bias of $\hat{\theta}$, we note that for values of $\lambda = -0.9, -0.7, -0.5, -0.3, 0, 0.3, 0.5$, $\hat{\theta}$ presented an excellent estimate converging to zero even for small samples. Although for $\lambda = 0.7$ and 0.9 the convergence of $\hat{\theta}$ is a little more time consuming, we have observed that with the increase in the sample size the estimated bias is very low, around -0.04 and -0.08 respectively.

We also note that for the values of $\lambda = -0.9, -0.7, -0.5, -0.3, 0, 0.3, 0.5$, $\hat{\lambda}$ is estimated accurately. Like in the case of θ , $\lambda = 0.7, 0.9$ also has a slower convergence, with the estimated bias close to -0.11 and -0.02 , respectively.

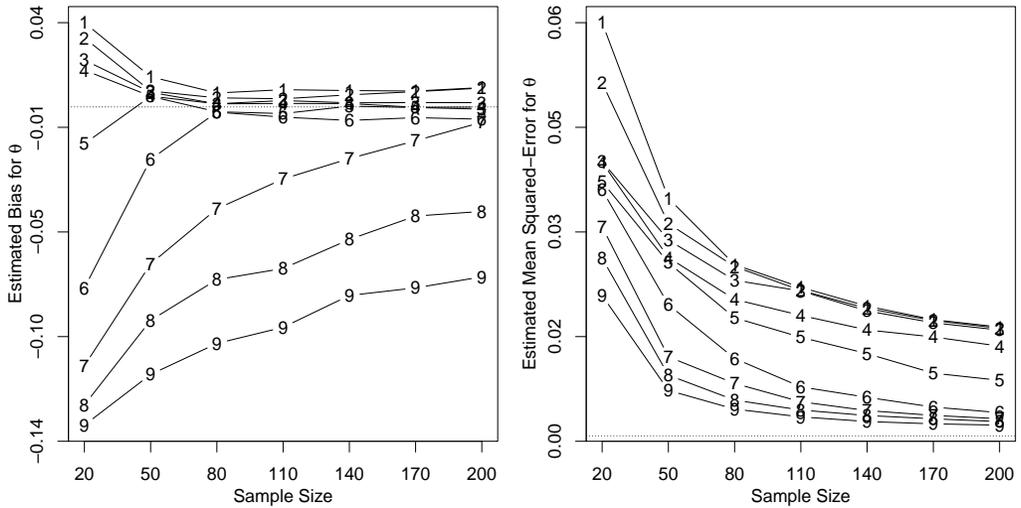


Figure 3 – Estimated bias and mean square error of θ ($\theta = 1, 1 : \lambda = -0.9, 2 : \lambda = -0.7, 3 : \lambda = -0.5, 4 : \lambda = -0.3, 5 : \lambda = 0, 6 : \lambda = 0.3, 7 : \lambda = 0.5, 8 : \lambda = 0.7$ and $9 : \lambda = 0.9$).

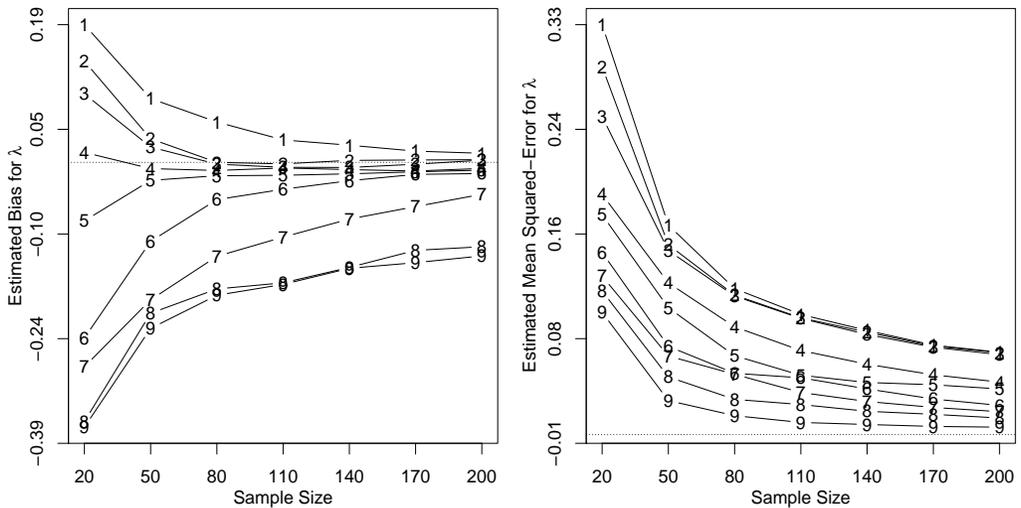


Figure 4 – Estimated bias and mean square error of λ ($\theta = 1, 1 : \lambda = -0.9, 2 : \lambda = -0.7, 3 : \lambda = -0.5, 4 : \lambda = -0.3, 5 : \lambda = 0, 6 : \lambda = 0.3, 7 : \lambda = 0.5, 8 : \lambda = 0.7$ and $9 : \lambda = 0.9$).

Though the bias are very low for both the estimated parameters, it is comparatively higher for $\hat{\lambda}$. Moreover higher the amplitude of λ higher is the bias in the estimates of both the parameters showing that the transmutation parameter λ , exerts influence in the estimation of the scale parameter θ .

The mean square errors of $\hat{\theta}$ are extremely low, with the positive values of λ the ones that present a more precise convergence. Although the positive λ parameters do not converge directly to zero,

we can note from the graph that the errors in the estimate are less than 0.03. Thus we can conclude that the errors in the estimates are practically insignificant.

Finally, when we observe the mean square error of the parameter $\hat{\lambda}$, we realize that θ exerts influence on its estimation. Analogously to the previous one, we can see that, for all scenarios, the errors tend to zero, the furthest being close to 0.1.

In general, we can conclude that the estimators have the property of asymptotic unbiasedness, since the bias tends to zero as n increases, while trend in the mean squared error show the consistency, because when the value of n increases the errors tend to zero.

8 REAL DATA ANALYSIS

In this section, we illustrate the applicability of THN distribution using two real data sets that were not been analyzed before in the literature. Our objective is to evaluate its adjustment in relation to other distributions already presented in the literature.

In both applications the data used were compiled from the daily series of daily precipitation obtained at the portal of the National Institute of Meteorology (<http://www.inmet.gov.br>). In the adjustment of distributions, the cumulative total for the month was considered as a response. The cumulative total monthly, quarterly, and so forth is widely used in the calculation of the Standardized Precipitation Index (SPI).

Table 1 – Descriptive statistics.

N. obs	30
Mean	206.93 mm
Std.	128.55 mm
Min.	8.20 mm
Q25	107.18 mm
Median	191.80 mm
Q75	270.45 mm
Max.	500.30 mm

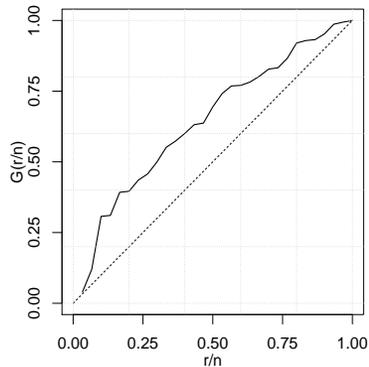


Figure 5 – TTT Plot.

8.1 Monthly Precipitation

For the first application, we consider the data for the month of February of each year, from 1974 to 2016, of the city of Chapecó localized in the state of Santa Catarina. It is important to note that due to some fault it does not record measurement in the month of February of some years in the considered period. Table 1 shows the summary measures of the data used. Looking at Figure 5, it is possible to note that the risk is increasing, an indication that the THN distribution may be an appropriate model for adjustment.

For this data set, we fit the following models:

1. Half-Normal (HN) due to Daniel (1959).

$$f_1(x) = \frac{2}{\theta} \phi\left(\frac{x}{\theta}\right), \quad \theta > 0.$$

2. General Half-Normal (GHN) due to Pewsey (2002, 2004).

$$f_2(x) = \frac{2}{\theta} \phi\left(\frac{x-\alpha}{\theta}\right), \quad \alpha = \min(x), \theta > 0.$$

3. Power Half-Normal (PHN) due to Gómez & Bolfarine (2015).

$$f_3(x) = \frac{2\alpha}{\theta} \phi\left(\frac{x}{\theta}\right) \left(2\Phi\left(\frac{x}{\theta}\right) - 1\right)^{\alpha-1}, \quad \alpha, \beta > 0.$$

4. Generalized Half-Normal (GHN II) due to Cooray & Ananda (2008).

$$f_4(x) = \frac{2\alpha}{\theta} \left(\frac{x}{\theta}\right)^{\alpha-1} \phi\left[\left(\frac{x}{\theta}\right)^\alpha\right], \quad \theta, \alpha > 0.$$

5. Gamma Half-Normal (GMHN) due to Alzaatreh & Knight (2013).

$$f_5(x) = \frac{2}{\Gamma(\alpha)b^\alpha} \phi\left(\frac{x}{\theta}\right) \left\{-\log\left[2\Phi\left(\frac{-x}{\theta}\right)\right]\right\}^{\alpha-1} \left[2\Phi\left(\frac{-x}{\theta}\right)\right]^{\frac{1}{b}-1}, \quad \theta, \alpha, b > 0.$$

6. Odd Log-Logistic Generalized Half-Normal (OLLGHN) due to Cordeiro et al. (2017).

$$f_6(x) = \frac{a\sqrt{\frac{2}{\pi}} \left(\frac{\alpha}{x}\right) \left(\frac{x}{\theta}\right)^\alpha e^{-\frac{1}{2}\left(\frac{x}{\theta}\right)^{2\alpha}} \left\{\left[2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) - 1\right] \left[2 - 2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right)\right]\right\}^{a-1}}{\left\{\left[2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) - 1\right]^a + \left[2 - 2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right)\right]^a\right\}^2}, \quad \theta, \alpha, a > 0.$$

7. Transmuted Generalized Half-Normal (TGHN).

$$f_7(x) = \frac{2\alpha}{\theta} \left(\frac{x}{\theta}\right)^{\alpha-1} \phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) \left[(1+\lambda) - 2\lambda \left(2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) - 1\right)\right], \quad \theta, \alpha > 0, |\lambda| \leq 1$$

8. Beta Generalized Half-Normal (BGHN) due to Pescim et al. (2010).

$$f_8(x) = \frac{2^b \alpha}{a\theta B(a,b)} \left(\frac{x}{\theta}\right)^{\alpha-1} \phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) \left[2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) - 1\right]^{a-1} \left[\Phi\left(-\left(\frac{x}{\theta}\right)^\alpha\right)\right]^{b-1}, \quad \theta, \alpha, a, b > 0$$

9. Extended Generalized Half-Normal (EGHN) due to Sanchez et al. (2016).

$$f_9(x) = ab\sqrt{\frac{2}{\pi}} \left(\frac{\alpha}{x}\right) \left(\frac{x}{\theta}\right)^\alpha \exp\left(-\frac{1}{2}\left(\frac{x}{\theta}\right)^{2\alpha}\right) \left[2 - 2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right)\right]^{a-1} \\ \times \left\{1 - \left[2 - 2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right)\right]^a\right\}^{b-1}, \quad \theta, \alpha, a, b > 0.$$

10. Kumaraswamy Generalized Half-Normal (KGHN) due to Cordeiro et al. (2012).

$$f_{10}(x) = ab\sqrt{\frac{2}{\pi}} \left(\frac{\alpha}{x}\right) \left(\frac{x}{\theta}\right)^\alpha \exp\left(-\frac{1}{2}\left(\frac{x}{\theta}\right)^{2\alpha}\right) \left[2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) - 1\right]^{a-1} \\ \times \left\{1 - \left[2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) - 1\right]^a\right\}^{b-1}, \quad \theta, \alpha, a, b > 0.$$

11. Beta Generalized Half-Normal Geometric (BGHNG) due to Ramires et al. (2013).

$$f_{11}(x) = (1-p) \frac{2^{b-\frac{1}{2}} \left(\frac{x}{\theta}\right)^\alpha \exp\left(-\frac{1}{2} \left(\frac{x}{\theta}\right)^{2\alpha}\right)}{\sqrt{\pi} B(a,b)} \left[2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right) - 1\right]^{a-1}$$

$$\times \left[1 - \Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right)\right]^{b-1} \left[1 - p \left(1 - I_{2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right)^{-1}}(a,b)\right)\right]^{-2}, \theta, \alpha, a, b > 0, 0 \leq p < 1.$$

where $I_{2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right)^{-1}}(a,b) = B_{2\Phi\left(\left(\frac{x}{\theta}\right)^\alpha\right)^{-1}}(a,b)/B(a,b).$

Table 2 shows the maximum likelihood estimates (standard errors) of the fitted distribution. All estimates were obtained in SAS/NLMIXED procedure SAS (2010), applying the Newton-Raphson optimization technique. Although, for all distributions, the resulting variance-covariance matrices were positive definite and $\max\left[\frac{\partial}{\partial \Theta} \ell(\Theta | x) \Big|_{\Theta = \hat{\Theta}}\right] < 0.000001$, we observed atypical standard errors for some parameters in the GMHN, BGHN, EGHN and KGHN distributions. Our guess is that they converge to a local minimum or, most likely, that the parameters are linear functions of each other (or almost collinear) on the data in question. In fact, from the correlation matrices we obtain $\text{corr}(\hat{\theta}, \hat{\alpha}) = 0.9594$, $\text{corr}(\hat{b}, \hat{\alpha}) = -0.9707$ and $\text{corr}(\hat{\theta}, \hat{b}) = -0.9984$ for the GMHN distribution, $\text{corr}(\hat{\alpha}, \hat{b}) = 0.7179$, $\text{corr}(\hat{\alpha}, \hat{a}) = -0.8255$, $\text{corr}(\hat{\theta}, \hat{\alpha}) = 0.8587$ and $\text{corr}(\hat{\theta}, \hat{a}) = 0.9711$ for the BGHN distribution, $\text{corr}(\hat{\theta}, \hat{b}) = 0.9817$, $\text{corr}(\hat{a}, \hat{b}) = 0.9896$ and $\text{corr}(\hat{\theta}, \hat{a}) = 0.9988$ for the EGHN distribution, $\text{corr}(\hat{\alpha}, \hat{a}) = -0.8899$, $\text{corr}(\hat{\theta}, \hat{a}) = 0.8977$, $\text{corr}(\hat{\alpha}, \hat{b}) = -0.9535$, $\text{corr}(\hat{a}, \hat{b}) = 0.9568$, $\text{corr}(\hat{\theta}, \hat{b}) = 0.9726$ and $\text{corr}(\hat{\theta}, \hat{\alpha}) = -0.9807$ for the KGHN distribution. It was not possible to estimate the standard error of all the parameters of the BGHNG distribution and its variance-covariance matrix was also not completely filled, but we can observe that $\text{corr}(\hat{\alpha}, \hat{p}) = 0.7391$, $\text{corr}(\hat{\alpha}, \hat{a}) = -0.7480$, $\text{corr}(\hat{\alpha}, \hat{b}) = -1.0000$ and $\text{corr}(\hat{b}, \hat{p}) = -1.0000$.

To compare the distributions, we consider the statistics based on the likelihood $-2 \times \text{Log-Like}$, AIC, AICc and BIC and the measures of good adjustment KS, AD and CvM. The best model is one that provides the minimum values of these criteria. Table 3 shows such values and the index indicates the classification obtained for each distribution. We also have in the last column a total rating (sum of ratings) for each distribution. Since there are large uncertainties in the estimates of the GMHN, BGHN, EGHN, KGHN and BGHNG distributions, they will not be considered. When looking at the table, we can see that the THN distribution ranked first, followed by the TGHN and PHN distribution. Both distributions, THN and PHN, have two parameters. The TGHN distribution has three parameters, and the additional parameter (λ) has zero in its respective confidence interval. In view of this, we can infer that the models that obtained the best fit were THN (17) and PHN (24).

8.2 Ten-Day Precipitation

The data consists of the ten-day accumulated precipitation between April and June in the station WMO 83498 localized in the state of Bahia, Brazil. We have considered the historical series from

Table 2 – Maximum likelihood estimates (standard errors).

Distribution	θ	α	λ	a	b	p
HN	242.4746 (31.3031)
GHN	235.5162 (30.4049)
PHN	211.7649 (28.9254)	1.4462 (0.3307)
THN	202.6784 (24.8288)	.	-0.6767 (0.3097)	.	.	.
GHN II	260.1236 (28.3470)	1.2904 (0.1952)
TGHN	201.8721 (59.6638)	0.9952 (0.3247)	-0.6852 (0.6450)	.	.	.
OLLGHN	262.3615 (33.1280)	1.2245 (0.4707)	.	1.0701 (0.4822)	.	.
GMHN	120.00 (324.49)	1.2035 (0.9856)	.	.	2.4547 (12.5272)	.
BGHN	94.6287 (189.44)	1.0069 (0.4544)	.	1.1730 (0.8355)	0.2371 (0.6290)	.
EGHN	49.9935 (989.30)	0.9418 (0.4977)	.	0.0947 (3.3346)	1.0646 (4.4612)	.
KGHN	640.00 (298.26)	8.0006 (0.9535)	.	0.1852 (0.0772)	3.9998 (5.5428)	.
BGHNG	86.8732 ($\hat{}$ 0.001)	1.1999 ($\hat{}$ 0.001)	.	0.9728 (0.2289)	0.0806 ($\hat{}$ 0.001)	0.5568 ($\hat{}$ 0.001)

Table 3 – Comparison criterion.

Distribution	$-2 \times \text{Log-Like}$	AIC	AICc	BIC	KS	AD	CvM	Total
HN	373.0016 ⁷	375.0016 ⁵	375.1444 ⁵	376.4028 ²	0.8602 ⁷	0.7372 ⁶	0.1312 ⁷	39 ⁷
GHN	371.2544 ⁶	373.2544 ¹	373.3972 ¹	374.6556 ¹	0.7476 ⁶	23.9874 ⁷	0.0954 ⁶	28 ⁴
PHN	370.5842 ⁵	374.5842 ⁴	375.0287 ⁴	377.3866 ⁵	0.3984 ³	0.1498 ²	0.0167 ¹	24 ³
THN	370.4147 ²	374.4147 ²	374.8591 ²	377.2170 ³	0.3802 ²	0.1500 ³	0.0192 ³	17 ¹
GHN II	370.4468 ⁴	374.4468 ³	374.8913 ³	377.2492 ⁴	0.4327 ⁵	0.1721 ⁵	0.0224 ⁵	29 ⁵
TGHN	370.4144 ¹	376.4144 ⁶	377.3375 ⁶	380.6180 ⁶	0.3777 ¹	0.1488 ¹	0.0189 ²	23 ²
OLLGHN	370.4232 ³	376.4232 ⁷	377.3463 ⁷	380.6268 ⁷	0.4035 ⁴	0.1589 ⁴	0.0196 ⁴	36 ⁶

1961 to 2017. In this application we adopted a regression structure for the scale parameter, that is,

$$\log(\theta_i) = \beta_0 + \beta_1 x_i \tag{14}$$

where x_i denote the i th observation associate the i th ten-day period. The periods of ten days were considered for the months of April, May, June and July. We have considered the distributions HN, GHN II and THN since only they have closed analytical expressions for the mean.

In Table 4 are reported the parameter estimates and standard errors. Since the $\hat{\beta}_1$ has a negative sign for all distributions, decrease in the accumulated precipitation is indicated as the time passes. The empirical and estimated means (95% confidence intervals) are presented in Table 5 and Figure 6. Also, in Table 6 are present several criteria to discriminate between the HN, GHN II and THN distributions. From these results we can conclude that GHN II and THN provide better fit than the HN distribution. It is also observed that THN has the lowest values of those criteria.

Table 4 – Parameter estimates and 95% confidence intervals.

	HN		GHN II		THN	
	Estimate	C.I.	Estimate	C.I.	Estimate	C.I.
β_0	5.0942	(4.7007, 5.4876)	5.0119	(4.5392, 5.4846)	5.4141	(4.9208, 5.9075)
β_1	-0.0785	(-0.1064, -0.0507)	-0.0795	(-0.1129, -0.0462)	-0.0792	(-0.1123, -0.0462)
α	—	—	0.8348	(0.7700, 0.8996)	—	—
λ	—	—	—	—	0.7822	(0.5896, 0.9747)

Table 5 – Empirical and estimated means for each ten-day period.

Ten-Day	Empirical	HN	GHN II	THN
10	51.6089	59.3420	55.1367	54.8366
11	48.1578	54.8614	50.9212	50.6590
12	49.1844	50.7191	47.0280	46.7996
13	41.0744	46.8896	43.4325	43.2342
14	41.1256	43.3492	40.1118	39.9404
15	37.8395	40.0761	37.0451	36.8976
16	32.3732	37.0502	34.2128	34.0866
17	31.9098	34.2528	31.5970	31.4897
18	26.2780	31.6665	29.1813	29.0907

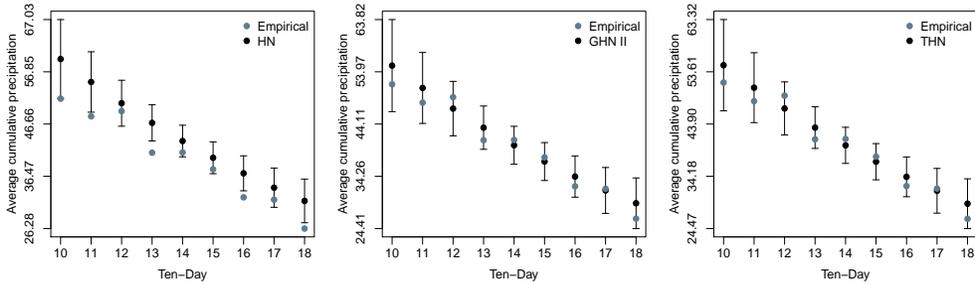


Figure 6 – The empirical and estimated means.

Table 6 – Model selection measures.

Criteria	HN	GHN II	THN
$-2 \times \text{Log-Like}$	3631.9093	3608.9037	3597.1897
AIC	3635.9093	3614.9037	3603.1897
BIC	3643.8106	3626.7556	3615.0416
SSR	207.2568	43.8623	40.3417

9 CONCLUSION

In this paper, we present the THN distribution formulated from the quadratic transmutation proposed by Shaw & Buckley (2009). Some characteristics and mathematical properties of the proposed distribution are studied. It is important to note that the moment generating function, moment of order k , mean, variance, asymmetry and kurtosis have explicit analytic expressions, which depend only of the parameters of the THN distribution. Due to the simplicity of the distribution, it was possible to calculate the uncertainty measures as a Shannon entropy and the mean deviations. As the Bonferroni and Lorenz curves as well as the reliability characteristic are presented, giving opportunities to areas such as engineering to extract benefits from the use of the proposed distribution. Distribution of order statistic was calculated for the distribution of THN, as well as their respective expressions for the densities of the maximum and minimum distribution. A Monte Carlo simulation study showed that the parameters are efficiently estimated by maximum likelihood method and show a low bias and a low accuracy even for a small sample sizes, which indicates the potential that the new distribution provides for modeling. In the first application using daily precipitation data and 11 models proposed in the literature (derived from the HN distribution) were placed in competition with the distribution proposed in this work. Five models were withdrawn from the analysis because of their inconsistent estimates and by looking at the statistics based on likelihood $-2 \times \text{Log-Like}$, AIC, AICc and BIC, as well as KS, AD and CvM, the THN distribution showed a better fit in comparison with the models used. Since a distribution presents an explicit and simple expression for the mean, it was possible to use it for an application using regression structure. When observing the $-2 \times \text{Log-Like}$, AIC, BIC and SSR criteria, we noticed that the THN distribution presented a good of fit, reinforcing its supremacy when compared to the models used here.

It is important to mention that during the peer review process the THN distribution was considered by Balaswamy (2018). Although both work propose the same distribution, we emphasize that in our paper a more comprehensive account of mathematical properties of the new distribution was presented (survival, hazard rate and residual life functions and their properties; asymptotic behavior of the tails; moments, associated measures and moment generating function; differential entropy; mean deviations; Bonferroni and Lorenz curves; order statistics; extreme values; stress strength reliability; characterizations by two truncated moments; characterizations by order statistics). In addition, we studied the bias and accuracy of the parameters estimated by the maximum likelihood method and illustrated the applicability of THN distribution using two real data sets that were not been analyzed before in the literature.

References

- [1] ALJARRAH MA, LEE C & FAMOYE F. 2014. On generating T-X family of distributions using quantile functions. *Journal of Statistical Distributions and Applications*, **1**(1): 1–17.
- [2] ALZAATREH A & KNIGHT K. 2013. On The Gamma-Half Normal Distribution and Its Applications, **12**(1): 103–119.
- [3] ASGHARZADEH A, VALIOLLAHI R & RAQAB MZ. 2011. Stress-strength reliability of Weibull distribution based on progressively censored samples. *SORT-Statistics and Operations Research Transactions*, **35**(2): 103–124.
- [4] BALASWAMY S. 2018. Transmuted Half Normal Distribution. *International Journal of Scientific Research in Mathematical and Statistical Sciences*, **5**(4): 163–170.
- [5] BONFERRONI CE. 1930. *Elementi di statistica generale*. Libreria Seber Firenze.
- [6] CHOU CY & LIU HR. 1998. Properties of the Half-Normal distribution and its application to quality control. *Journal of Industrial Technology*, **14**(3): 4–7.
- [7] COORAY K & ANANDA MMA. 2008. A generalization of the Half-Normal distribution with applications to lifetime data. *Communications in Statistics – Theory and Methods*, **37**(9): 1323–1337.
- [8] CORDEIRO GM, ALIZADEH M, PESCIM RR & ORTEGA EMM. 2017. The odd log-logistic generalized Half-Normal lifetime distribution: Properties and applications. *Communications in Statistics - Theory and Methods*, **46**(9): 4195–4214.
- [9] CORDEIRO GM, PESCIM RR & ORTEGA EMM. 2012. The Kumaraswamy Generalized Half-Normal Distribution for Skewed Positive Data. *Journal of Data Science*, **10**: 195–224.
- [10] CORDEIRO GM, PESCIM RR, ORTEGA EMM & DEMÉTRIO CGB. 2013. The Beta generalized Half-Normal distribution: New properties. *Journal of Probability and Statistics*, **2013**: 1–18.

- [11] DANIEL C. 1959. Use of Half-Normal plots in interpreting factorial two-level experiments. *Technometrics*, **1**(4): 311–341.
- [12] DAVID HA & NAGARAJA HN. 2003. *Order statistics*. 3rd ed.. Wiley Series in Probability and Statistics. John Wiley & Sons.
- [13] DE BRITO CCR, RÊGO LC & DE OLIVEIRA WR. 2015. Method for Generating Distributions and Classes of Probability Distributions: The Univariate Case. *arXiv:1504.01062*, pp. 1–50.
- [14] FELLER W. 1968. *An introduction to probability theory and its applications*. vol. 1. 3rd ed.. John Wiley & Sons.
- [15] GLÄNZEL W. 1987. *A Characterization Theorem Based on Truncated Moments and its Application to Some Distribution Families*. Dordrecht: Springer Netherlands.
- [16] GÓMEZ YM & BOLFARINE H. 2015. Likelihood-based inference for the power Half-Normal distribution. *Journal of Statistical Theory and Applications*, **14**(4): 383–398.
- [17] GUPTA RD & KUNDU D. 2009. Introduction of Shape/Skewness Parameter(s) in a Probability Distribution. *Journal of Applied Statistical Science*, **7**(2): 153–171.
- [18] HAMEDANI GG, CORDEIRO GM, LIMA MCS & NASCIMENTO ADC. 2017. Some Extended Classes of Distributions: Characterizations and Properties. *Pakistan Journal of Statistics and Operation Research*, **13**(4): 893–908.
- [19] JOHNSON NL, KOTZ S & BALAKRISHNAN N. 1994. *Continuous univariate distributions*. Vol. 1. Second ed.. New York: John Wiley & Sons Inc.
- [20] LAI DC. 2011. Constructions and applications of lifetime distributions. *Applied Stochastic Models in Business and Industry*, **29**(2): 127–140.
- [21] LEADBETTER MR, LINDGREN G & ROOTZÉN H. 1987. *Extremes and related properties of random sequences and processes*. 1st ed.. Springer Science & Business Media.
- [22] LEE C, FAMOYE F & ALZAATREH AY. 2013. Methods for generating families of univariate continuous distributions in the recent decades. *Wiley Interdisciplinary Reviews: Computational Statistics*, **5**(3): 219–238.
- [23] LEONE FC, NELSON LS & NOTTINGHAM RB. 1961. The Folded Normal Distribution. *Technometrics*, **3**(4): 543–550.
- [24] NADARAJAH S & KOTZ S. 2006a. The beta exponential distribution. *Reliability Engineering & System Safety*, **91**(6): 689–697.
- [25] NADARAJAH S & KOTZ S. 2006b. R Programs for Truncated Distributions. *Journal of Statistical Software*, **16**(1): 1–8.

- [26] NADARAJAH S & ROCHA R. 2016. Newdists: An R Package for New Families of Distributions. *Journal of Statistical Software*, **69**(1): 1–32.
- [27] NAKAGAMI N. 1960. *The m-distribution a general formulation of intensity distribution of rapid fading*. Proc. Symp. Statist. Methods Radio Wave Propag.
- [28] OLMOS NM, VARELA H, BOLFARINE H & GÓMEZ HW. 2014. An extension of the generalized Half-Normal distribution. *Statistical Papers*, **55**(4): 967–981.
- [29] PESCIM RR, DEMÉTRIO CGB, CORDEIRO GM, ORTEGA EMM & URBANO MR. 2010. The Beta generalized Half-Normal distribution. *Computational Statistics & Data Analysis*, **54**(4): 945–957.
- [30] PESCIM RR, ORTEGA EMM, CORDEIRO GM & DEMÉTRIO CGB. 2013. The Log-Beta Generalized Half-Normal Regression Model. *Journal of Statistical Theory and Applications*, **12**(4): 330–347.
- [31] PEWSEY A. 2002. Large-sample inference for the general Half-Normal distribution. *Communications in Statistics - Theory and Methods*, **31**(7): 1045–1054.
- [32] PEWSEY A. 2004. Improved likelihood based inference for the general Half-Normal distribution. *Communications in Statistics - Theory and Methods*, **33**(2): 197–204.
- [33] RAMIRES TG, ORTEGA EMM, CORDEIRO GM & HAMEDANI GG. 2013. The Beta generalized Half-Normal geometric distribution. *Studia Scientiarum Mathematicarum Hungarica*, **50**(4): 523–554.
- [34] SABOOR A, KHAN MN, CORDEIRO GM, PASCOA MA, RAMOS PL & KAMAL M. 2019. Some new results for the transmuted generalized gamma distribution. *Journal of Computational and Applied Mathematics*, **352**: 165–180.
- [35] SANCHEZ JJD, FREITAS WWL & CORDEIRO GM. 2016. The extended generalized Half-Normal distribution. *Brazilian Journal of Probability and Statistics*, **30**: 366–384.
- [36] SAS. 2010. *The NLMIXED Procedure, SAS/STAT® User's Guide, Version 9.22*. Cary, NC: SAS Institute Inc.
- [37] SHANNON CE. 1951. Prediction and entropy of printed English. *Bell Labs Technical Journal*, **30**(1): 50–64.
- [38] SHAW W & BUCKLEY IRC. 2009. The alchemy of probability distributions: Beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv:0901.0434v1 [q-fin.ST]*, pp. 1–8.
- [39] SINGH R. 1994. Simulation of Observations for the Half-Normal Distribution. *Sankhyā: The Indian Journal of Statistics, Series B*, **56**(2): 137–139.

- [40] SINGH R. 1995. Editorial Note on the Paper “Simulation of observations for the Half-Normal distribution”. *Sankhyā: The Indian Journal of Statistics, Series B*, **57**(3): 461–461.
- [41] STACY EW. 1962. A generalization of the Gamma distribution. *The Annals of Mathematical Statistics*, **33**(3): 1187–1192.
- [42] TAHIR MH & CORDEIRO GM. 2016. Compounding of distributions: A survey and new generalized classes. *Journal of Statistical Distributions and Applications*, **3**(1): 3–13.
- [43] TAHIR MH & NADARAJAH S. 2015. Parameter induction in continuous univariate distributions: Well-established G families. *Journal of Probability and Statistics*, **87**(2): 539–568.
- [44] TSAGRIS M, BENEKI C & HASSANI H. 2014. On the Folded Normal Distribution. *Mathematics*, **2**(1): 12–28.
- [45] VODĀ VG. 1976a. Inferential procedures on a generalized Rayleigh variate I. *Aplikace Matematiky*, **21**(6): 395–412.
- [46] VODĀ VG. 1976b. Inferential procedures on a generalized Rayleigh variate II. *Aplikace Matematiky*, **21**(6): 413–419.
- [47] YOUSOF HM, ALIZADEH M, JAHANSHAH SMA, RAMIRES TG, GHOSH I & HAMEDANI G. 2017. The Transmuted Topp-Leone G Family of Distributions: Theory, Characterizations and Applications. *Journal of Data Science*, **15**(4): 723–740.

How to cite

MENEZES A, MAZUCHELI J, CARDOSO J & CHAKRABORTY S. 2020. The transmuted Half-Normal distribution with application to precipitation data. *Pesquisa Operacional*, **40**: e216792. doi: 10.1590/0101-7438.2020.040.00216792.