symmetry

MDPI

Article

A Bimodal Model Based on Truncation Positive Normal with Application to Height Data

Héctor J. Gómez [1], Wilson E. Caimanque [2], Yolanda M. Gómez [2], Tiago M. Magalhães [3], Miguel Concha [2] and Diego I. Gallardo [2,*]

[1] Departamento de Ciencias Matemáticas y Físicas, Facultad de Ingeniería, Universidad Católica de Temuco, Temuco 4780000, Chile; hgomez@uct.cl
[2] Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Copiapó 1530000, Chile; wilson.caimanque.21@alumnos.uda.cl (W.E.C.); yolanda.gomez@uda.cl (Y.M.G.); miguel.concha.21@alumnos.uda.cl (M.C.)
[3] Department of Statistics, Institute of Exact Sciences, Federal University of Juiz de Fora, Juiz de Fora 36036-900, Brazil; tiago.magalhaes@ufjf.edu.br
* Correspondence: diego.gallardo@uda.cl; Tel.: +56-52-2255565

Abstract: In this work, we propose a new bimodal distribution with support in the real line. We obtain some properties of the model, such as moments, quantiles, and mode, among others. The computational implementation of the model is presented in the tpn package of the software R. We perform a simulation study in order to assess the properties of the maximum likelihood estimators in finite samples. Finally, we present an application to a bimodal data set, where our proposal is compared with other models in the literature.

Keywords: bimodal distribution; unimodal distribution; tpn package; maximum likelihood estimation; asymmetric

check for updates

Citation: Gómez, H.J.; Caimanque, W.E.; Gómez, Y.M.; Magalhães, T.M.; Concha, M.; Gallardo, D.I. A Bimodal Model Based on Truncation Positive Normal with Application to Height Data. *Symmetry* **2022**, *14*, 665. https://doi.org/10.3390/sym14040665

Academic Editor: Jinyu Li

Received: 13 February 2022
Accepted: 17 March 2022
Published: 24 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Describing a phenomenon by a probability distribution is very useful because of the properties associated with it: expectation, shape, range, etc. However, this description can be difficult when a phenomenon (in practice, an observed dataset) is bimodal, which occurs commonly in areas like astrophysics, ecology and genetics; see [1–3], respectively. The first approach to fit a bimodal data is using a mixture of two unimodal distributions, for instance, a mixture of gaussian distributions; see [4]. The main disadvantage of this procedure is the non-identifiability of the proposed mixture model. The second and the most workable practical approach is to use distributions which already have bimodal properties. Because of these properties, there is an increasing interest to derive bimodal distributions in the literature: refs. [5,6] presented extensions of the skew-normal, ref. [7] proposed a generalization of the Burr type X distribution and [8] derived an extension of the sinh Cauchy distribution. In this paper, we will discuss an extension of the half normal distribution proposed by [9], the truncated positive normal (tpn) model. The probability density function (pdf) for the tpn model is given by

$$f(x;\sigma,\lambda) = \frac{1}{\sigma\Phi(\lambda)}\phi\left(\frac{x}{\sigma}-\lambda\right), \quad x,\sigma \in \mathbb{R}_+, \lambda \in \mathbb{R},$$

where σ and λ are the scale and shape parameters, respectively, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cumulative distribution function (cdf), respectively, of the standard normal distribution. The corresponding cdf of the tpn model is

$$F_X(x;\sigma,\lambda) = \frac{\Phi(\frac{x}{\sigma}-\lambda)+\Phi(\lambda)-1}{\Phi(\lambda)}.$$

Note that the cdf above has a closed-form expression, which is useful for generating random data besides defining quantiles. For more properties of the tpn model, see [9]. The restriction to positive values is a limitation of the tpn model. To overcome this limitation, the chief goal of this paper is to derive an extension of the tpn model which has support in the real line. We describe in detail the model, studying its main properties and related functions. Moreover, we show analytically the regions in which the model is unimodal and bimodal, and such regions depend only on one parameter.

The paper is organized as follows. In Section 2, we derive an extension of the tpn with support in the real line and study some properties of the distribution. The inference for parameter estimation in the proposed model and computational aspects are presented in Section 3. In Section 4, we perform a simulation study to evaluate the parameter estimation in finite samples. An application to real data is discussed in Section 5. Finally, conclusions are given in Section 6.

2. A Bimodal Truncation Positive Normal Distribution

In this section, we present the stochastic representation for the bimodal truncation positive normal (btpn) distribution and some properties, such as its pdf and its cdf. We also discuss some particular cases of the model.

2.1. Stochastic Representation, pdf and cdf

Let T be a discrete random variable such as

$$T = \begin{cases} -(1+\epsilon) & \text{, with probability } (1+\epsilon)/2 \\ 1-\epsilon & \text{, with probability } (1-\epsilon)/2 \end{cases},$$

where $\epsilon \in (-1,1)$. If $Z \sim \text{tpn}(\sigma, \lambda)$, independent from T, then we define a new random variable given by $X = ZT$. We say that X follows a btpn distribution.

Proposition 1. *The pdf for the btpn distribution is given by*

$$f(x;\sigma,\lambda,\epsilon) = \begin{cases} \dfrac{\phi\left(\frac{x}{\sigma(1+\epsilon)}+\lambda\right)}{2\sigma\Phi(\lambda)} & \text{, } if\ x < 0 \\ \dfrac{\phi\left(\frac{x}{\sigma(1-\epsilon)}-\lambda\right)}{2\sigma\Phi(\lambda)} & \text{, } if\ x \geq 0 \end{cases},$$

where $\sigma > 0, \lambda \in \mathbb{R}$ *and* $\epsilon \in (-1,1)$.

Proof. If $x < 0$, then the cdf for X is

$$\begin{aligned} F_X(x) &= P(X \leq x) = \frac{(1+\epsilon)}{2} P\left(z \geq \frac{-x}{1+\epsilon}\right) = \frac{(1+\epsilon)}{2}\left[1 - P\left(z \leq \frac{-x}{1+\epsilon}\right)\right] \\ &= \frac{(1+\epsilon)}{2}\left[1 - F_Z\left(\frac{-x}{1+\epsilon}\right)\right]. \end{aligned}$$

Deriving the last expression in relation to x, we have

$$f_X(x) = \frac{(1+\epsilon)}{2}\left[-f_Z\left(\frac{-x}{1+\epsilon}\right)\frac{-1}{(1+\epsilon)}\right] = \frac{1}{2}\left[\frac{1}{\sigma\Phi(\lambda)}\phi\left(\frac{\frac{-x}{(1+\epsilon)}}{\sigma} - \lambda\right)\right] = \frac{\phi\left(\frac{x}{\sigma(1+\epsilon)} + \lambda\right)}{2\sigma\Phi(\lambda)}.$$

A similar routine calculation shows that for $x \geq 0$, we have that

$$f_X(x) = \frac{\phi\left(\frac{x}{\sigma(1-\epsilon)} - \lambda\right)}{2\sigma\Phi(\lambda)},$$

completing the proof. $\square$

Figure 1 shows the pdf function for the btpn model with different combination of parameters. Note that the model can assume different shapes, including unimodal, bimodal, symmetric and asymmetric.

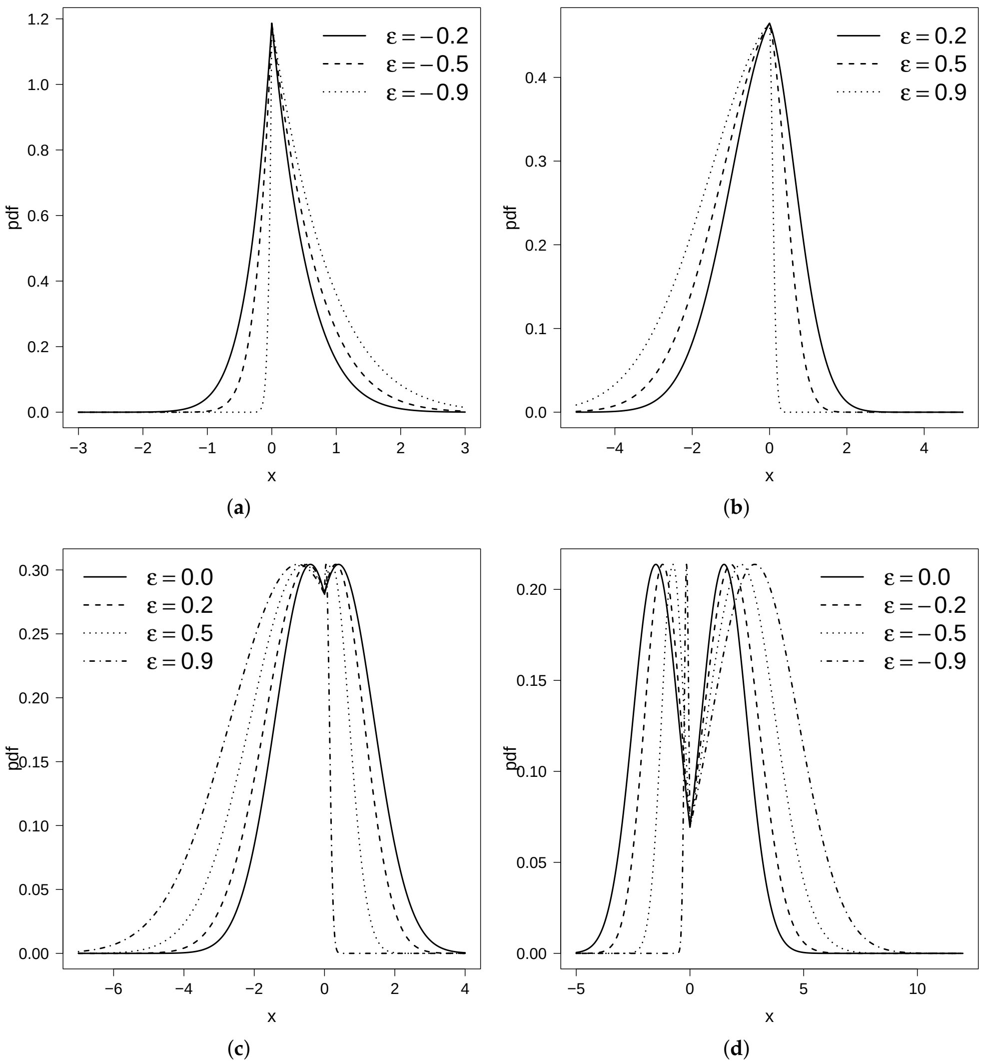


Figure 1. Pdf for btpn($\sigma = 1, \lambda, \epsilon$) with different fixed values for λ and varying ϵ: (**a**) $\lambda = -2$; (**b**) $\lambda = -0.2$; (**c**) $\lambda = 0.4$ and; (**d**) $\lambda = 1.5$.

Proposition 2. *The cdf of* $X \sim btpn(\sigma, \lambda, \epsilon)$ *is given by*

$$F_X(x) = \begin{cases} \frac{(1+\epsilon)}{2\Phi(\lambda)} \Phi\left(\frac{x}{\sigma(1+\epsilon)} + \lambda\right) & , \textit{if } x \leq 0 \\ \frac{(1-\epsilon)}{2\Phi(\lambda)} \left[\Phi\left(\frac{x}{\sigma(1-\epsilon)} - \lambda\right) + \Phi(\lambda) - 1\right] & , \textit{if } x \geq 0 \end{cases}$$

Proof. It is immediate from the last proof. □

Proposition 3. *Let there be* $X \sim btpn(\sigma, \lambda, \epsilon)$. *Its quantile function is given by*

$$Q_X(p) = F_X^{-1}(p) = \begin{cases} \sigma(1+\epsilon)\left[\Phi^{-1}\left(\frac{2p\Phi(\lambda)}{1+\epsilon}\right) - \lambda\right] & , \textit{if } 0 < p \leq \frac{1+\epsilon}{2} \\ \sigma(1-\epsilon)\left[\Phi^{-1}\left(\frac{2p\Phi(\lambda)}{1-\epsilon} - \Phi(\lambda) + 1\right) + \lambda\right] & , \textit{if } \frac{1+\epsilon}{2} < p < 1 \end{cases}$$

Proof. It is immediate from inverting the cdf for the btpn distribution given in Proposition 2. □

Corollary 1. *The median for $X \sim btpn(\sigma, \lambda, \epsilon)$ is given by*

$$Me(X) = \begin{cases} \sigma(1+\epsilon)\left[\Phi^{-1}\left(\frac{\Phi(\lambda)}{1+\epsilon}\right) - \lambda\right] & , \text{if } \epsilon \geq 0 \\ \sigma(1-\epsilon)\left[\Phi^{-1}\left(\frac{\epsilon\Phi(\lambda)}{(1-\epsilon)} + 1\right) + \lambda\right] & , \text{if } \epsilon < 0. \end{cases}$$

Corollary 2. *The median for $X \sim btpn(\sigma, \lambda, \epsilon)$ is < 0, $= 0$ and > 0 if ϵ is > 0, $= 0$ and < 0, respectively.*

2.2. Moments and Moment-Generating Function

The following proposition presents the central moments of the btpn distribution.

Proposition 4. *Let $X \sim btpn(\sigma, \lambda, \epsilon)$. The r-th central moment of X is given by*

$$E(X^r) = \frac{\sigma^r}{2\sqrt{2\pi}\Phi(\lambda)}\left[(-1)^r(1+\epsilon)^{r+1} + (1-\epsilon)^{r+1}\right]\sum_{k=0}^{r}\binom{r}{k}\lambda^{r-k}(2)^{(k-1)/2}\Gamma\left((k+1)/2, \lambda^2/2\right),$$

where $\Gamma(a,b) = \int_b^{+\infty} t^{a-1}e^{-t}dt$ is the upper incomplete gamma function.

Proof. Note that $E(X^r) = E_1(X^r) + E_2(X^r)$, where $E_1(X^r) = \int_{-\infty}^{0} x^r f(x)dx$ and $E_2(X^r) = \int_0^{+\infty} x^r f(x)dx$. For the first term, we perform the change of variable $u = \frac{-x}{\sigma(1+\epsilon)} - \lambda$. With this,

$$\begin{aligned} E_1(X^r) &= \int_{-\infty}^{0} \frac{x^r}{2\sigma(1+\epsilon)}\phi\left(\frac{-x}{\sigma(1+\epsilon)} - \lambda\right)dx \\ &= \frac{(-\sigma)^r(1+\epsilon)^{(r+1)}}{2\Phi(\lambda)}\int_{-\lambda}^{\infty}(u+\lambda)^r\phi(u)du. \end{aligned}$$

Using the binomial theorem and the change of variable $t = u^2/2$ in the last expression, we obtain

$$\begin{aligned} E_1(X^r) &= \frac{(-\sigma)^r(1+\epsilon)^{(r+1)}}{2\sqrt{2\pi}\Phi(\lambda)}\sum_{k=0}^{r}\binom{r}{k}\lambda^{r-k}\int_{-\lambda}^{\infty}u^k e^{-u^2/2}du, \\ &= \frac{(-\sigma)^r(1+\epsilon)^{(r+1)}}{2\sqrt{2\pi}\Phi(\lambda)}\sum_{k=0}^{r}\binom{r}{k}\lambda^{r-k}2^{(k-1)/2}\int_{\lambda^2/2}^{\infty}t^{(k-1)/2}e^{-t}dt, \end{aligned}$$

Note that the last integral corresponds to $\Gamma((k+1)/2, \lambda^2/2)$.

On the other hand, for $E_2(X^r)$, we perform the change of variable $u = \frac{x}{\sigma(1-\epsilon)} - \lambda$, obtaining

$$\begin{aligned} E_2(X^r) &= \int_0^{\infty} \frac{x^r}{2\sigma(1-\epsilon)}\phi\left(\frac{x}{\sigma(1-\epsilon)} - \lambda\right)dx \\ &= \frac{\sigma^r(1-\epsilon)^{(r+1)}}{2\Phi(\lambda)}\int_{-\lambda}^{\infty}(u+\lambda)^r\phi(u)du. \end{aligned}$$

Using the same routine calculation, we obtain

$$E_2(X^r) = \frac{\sigma^r(1-\epsilon)^{(r+1)}}{2\sqrt{2\pi}\Phi(\lambda)}\sum_{k=0}^{r}\binom{r}{k}\lambda^{r-k}2^{(k-1)/2}\int_{\lambda^2/2}^{\infty}t^{(k-1)/2}e^{-t}dt,$$

where again the last integral corresponds to $\Gamma((k+1)/2, \lambda^2/2)$. The final result is obtained by summing $E_1(X^r)$ and $E_2(X^r)$. $\square$

Proposition 5. *Let $X \sim btpn(\sigma, \lambda, \epsilon)$. The moment-generating function (mgf) for X is given by*

$$M_x(t) = \frac{(1+\epsilon)\exp\left\{\frac{(1+\epsilon)^2}{2}t^2\sigma^2 - t\sigma\lambda(1+\epsilon)\right\}\Phi(\lambda - t\sigma(1+\epsilon))}{2\Phi(\lambda)}, + \frac{(1-\epsilon)\exp\left\{\frac{(1-\epsilon)^2}{2}t^2\sigma^2 + t\sigma\lambda(1-\epsilon)\right\}\Phi(\lambda + t\sigma(1-\epsilon))}{2\Phi(\lambda)}.$$

Proof. Note that $M_X(t) = M_{X_1}(t) + M_{X_2}(t)$, where $M_{X_1}(t) = \int_{-\infty}^{0} e^{tx} f(x)dx$ and $M_{X_2}(t) = \int_0^{+\infty} e^{tx} f(x)dx$. For the first integral and using the change of variable $u = x/[\sigma(1+\epsilon)] + \lambda$, we obtain

$$M_{X_1}(t) = \frac{(1+\epsilon)}{2\Phi(\lambda)} \int_{-\infty}^{\lambda} e^{t\sigma(1+\epsilon)(\mu-\lambda)} \phi(u)du.$$

Completing the square of a binomial in the last term of the exponential and using the change of variable $z = \mu - t\sigma(1+\epsilon)$, we have

$$\begin{aligned} M_{X_1}(t) &= \frac{(1+\epsilon)\exp\left\{\frac{t^2\sigma^2(1+\epsilon)^2}{2} - t\sigma\lambda(1+\epsilon)\right\}}{2\Phi(\lambda)} \int_{-\infty}^{\lambda - t\sigma(1+\epsilon)} \frac{e^{-z^2/2}}{\sqrt{2\pi}}dz, \\ &= \frac{(1+\epsilon)\exp\left\{\frac{t^2\sigma^2(1+\epsilon)^2}{2} - t\sigma\lambda(1+\epsilon)\right\}}{2\Phi(\lambda)}\Phi(\lambda - t\sigma(1+\epsilon)). \end{aligned}$$

For $X \geq 0$, and similarly to the previous development, we use the change of variable $u = x/[\sigma(1-\epsilon)] - \lambda$, obtaining that

$$M_{X_2}(t) = \frac{(1-\epsilon)}{2\Phi(\lambda)} \int_{-\lambda}^{\infty} e^{t\sigma(1-\epsilon)(\mu+\lambda)} \phi(u)du.$$

Again, completing the square and using $z = \mu - t\sigma(1-\epsilon)$, we obtain

$$\begin{aligned} M_{X_2}(t) &= \frac{(1-\epsilon)\exp\left\{\frac{t^2\sigma^2(1-\epsilon)^2}{2} + t\sigma\lambda(1-\epsilon)\right\}}{2\Phi(\lambda)} \int_{-\lambda - t\sigma(1-\epsilon)}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}}dz, \\ &= \frac{(1-\epsilon)\exp\left\{\frac{t^2\sigma^2(1-\epsilon)^2}{2} + t\sigma\lambda(1-\epsilon)\right\}}{2\Phi(\lambda)}\Phi(\lambda + t\sigma(1-\epsilon)). \end{aligned}$$

Finally, the result is obtained by summing $M_{X_1}(t)$ and $M_{X_2}(t)$. $\square$

Corollary 3. *Using properties of the mgf, the first four moments of $X \sim btpn(\sigma, \lambda, \epsilon)$ can be obtained from the expression $\mu_r = (\partial^r M_X(t)/\partial t^r)\big|_{t=0}$.*

- $\mu_1 = E[X^1] = -2\epsilon\sigma[\lambda + \Omega(\lambda)]$
- $\mu_2 = E[X^2] = \sigma^2(1+3\epsilon^2)[\lambda^2 + \lambda\Omega(\lambda) + 1]$
- $\mu_3 = E[X^3] = -4\epsilon\sigma^3(1+\epsilon^2)[\lambda^3 + \lambda^2\Omega(\lambda) + \lambda + 2\Omega(\lambda)]$
- $\mu_4 = E[X^4] = \sigma^4[1 + 5\epsilon^2(2+\epsilon)][\lambda^4 + \lambda^3\Omega(\lambda) + 6\lambda^2 + 5\lambda\Omega(\lambda) + 3]$

where $\Omega(\lambda) = \phi(\lambda)/\Phi(\lambda)$ is the reciprocal of the Mill's ratio for the standard normal distribution.

Corollary 4. *The variance, coefficients of skewness and kurtosis for $X \sim btpn(\sigma, \lambda, \epsilon)$ are given by*

$$Var(X) = \sigma^2\left[\lambda^2(1-\epsilon^2) + \lambda\Omega(\lambda)(1-5\epsilon^2) + \epsilon^2(3-4\Omega(\lambda)) + 1\right],$$

$$\sqrt{b_1} = \frac{-4\epsilon(1+\epsilon^2)[\lambda^3 + \lambda^2\Omega(\lambda) + 2\Omega(\lambda)]}{\left[(1+3\epsilon)(\lambda^2 + \lambda\Omega(\lambda) + 1)\right]^{3/2}}, \qquad \textit{and}$$

$$b_2 = \frac{[1+5\epsilon^2(2+\epsilon)][\lambda^4 + \lambda^3\Omega(\lambda) + 6\lambda^2 + 5\lambda\Omega(\lambda) + 3]}{[(1+3\epsilon^2)(\lambda^2 + \lambda\Omega(\lambda) + 1)]^2},$$

respectively.

Figure 2 shows the plots for asymmetry and kurtosis coefficients. Note that a more right-skewed distribution is obtained when $\epsilon \to -1$ and $\lambda \to -\infty$, whereas a more left-skewed model is obtained when $\epsilon \to 1$ and $\lambda \to -\infty$. On the other hand, a greater kurtosis is obtained when $\epsilon \to -1$ and $\lambda \to -\infty$, whereas a lower kurtosis is obtained when $|\epsilon| \to 1$ and $\lambda \to \infty$. Note that this pattern is consistent with the pdf for different parameters presented in Figure 1.

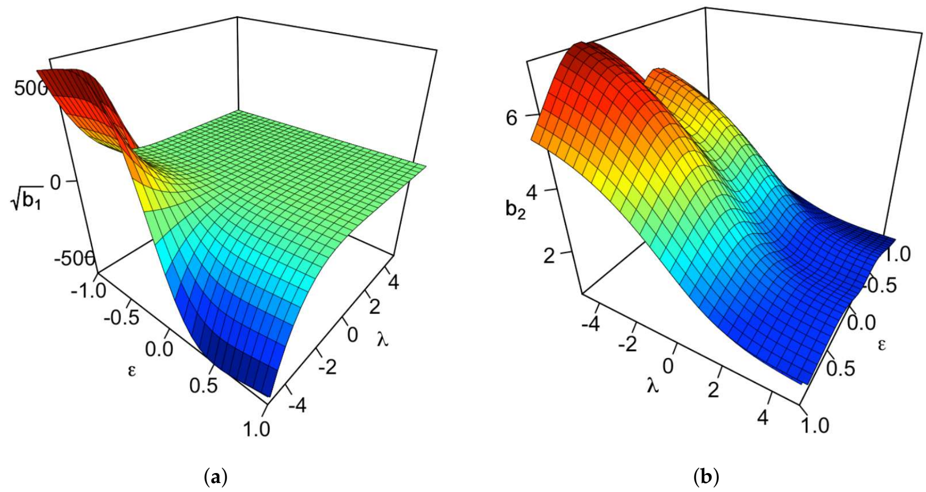


Figure 2. (**a**) Asymmetry coefficient and (**b**) kurtosis coefficient for btpn$(\sigma = 1, \lambda, \epsilon)$ distribution.

2.3. Mode and Unimodality and Bimodality Regions

The next proposition presents the unimodality and bimodality property of the btpn distribution.

Proposition 6. *Let $X \sim btpn(\sigma, \lambda, \epsilon)$. For $\lambda \leq 0$, the model is unimodal, and for $\lambda > 0$, the model is bimodal. Moreover, for the unimodal case, the mode of the model is 0, and for the bimodal case, the two modes are $-\sigma\lambda(1+\epsilon)$ and $\sigma\lambda(1-\epsilon)$, respectively.*

Proof. By definition, the mode is the value that maximizes the pdf or, equivalently, the logarithm of the pdf. For $X \sim btpn(\sigma, \lambda, \epsilon)$, it is straighforward to show that

$$\frac{\partial \log f(x;\sigma,\lambda,\epsilon)}{\partial x} = \begin{cases} -\frac{1}{\sigma(1+\epsilon)}\left(\frac{x}{\sigma(1+\epsilon)} + \lambda\right) & \text{, if } x < 0 \\ -\frac{1}{\sigma(1-\epsilon)}\left(\frac{x}{\sigma(1-\epsilon)} - \lambda\right) & \text{, if } x \geq 0 \end{cases} \quad \text{and}$$

$$\frac{\partial^2 \log f(x;\sigma,\lambda,\epsilon)}{\partial x^2} = \begin{cases} -\frac{1}{\sigma^2(1+\epsilon)^2} & \text{, if } x < 0 \\ -\frac{1}{\sigma^2(1-\epsilon)^2} & \text{, if } x \geq 0 \end{cases}.$$

Therefore, solving the equation $\partial \log f(x;\sigma,\lambda,\epsilon)/\partial x = 0$, we obtain $x_1 = -\sigma\lambda(1+\epsilon)$ and $x_2 = \sigma\lambda(1-\epsilon)$ as the potential mode for each branch, because the second derivative is negative for each respective case. However, this is valid if and only if $x_1 < 0$ and $x_2 > 0$, respectively. In other words, if $-\sigma\lambda(1+\epsilon) < 0$, then x_1 is a mode and if $\sigma\lambda(1-\epsilon) > 0$, then x_2 also is a mode. This is equivalent to

$$\{\lambda > 0 \wedge 1+\epsilon > 0\} \vee \{\lambda < 0 \wedge 1+\epsilon < 0\} \Rightarrow x_1 \text{ is a mode and}$$
$$\{\lambda > 0 \wedge 1-\epsilon > 0\} \vee \{\lambda < 0 \wedge 1-\epsilon < 0\} \Rightarrow x_2 \text{ is a mode.}$$

where $1+\epsilon > 0$ and $1-\epsilon > 0$, $\forall\epsilon(-1,1)$. On the other hand, $1+\epsilon \not< 0$ and $1-\epsilon \not< 0$. For this reason, it is immediate that for $\lambda > 0$, the btpn distribution have two modes, and such modes are x_1 and x_2. Finally, for $\lambda \leq 0$, it is immediate that $\partial \log f(x;\sigma,\lambda,\epsilon)/\partial x > 0$, for $x < 0$ and $\partial \log f(x;\sigma,\lambda,\epsilon)/\partial x \leq 0$ for $x \geq 0$. In other words, the pdf for the btpn distribution is an increasing function in $(-\infty, 0)$ and a decreasing function in $(0,\infty)$, where we can deduce that the model is unimodal and the respective mode is attached in zero. □

Figure 3 shows the regions of unimodality and bimodality for the btpn depending on the parameters λ and ϵ.

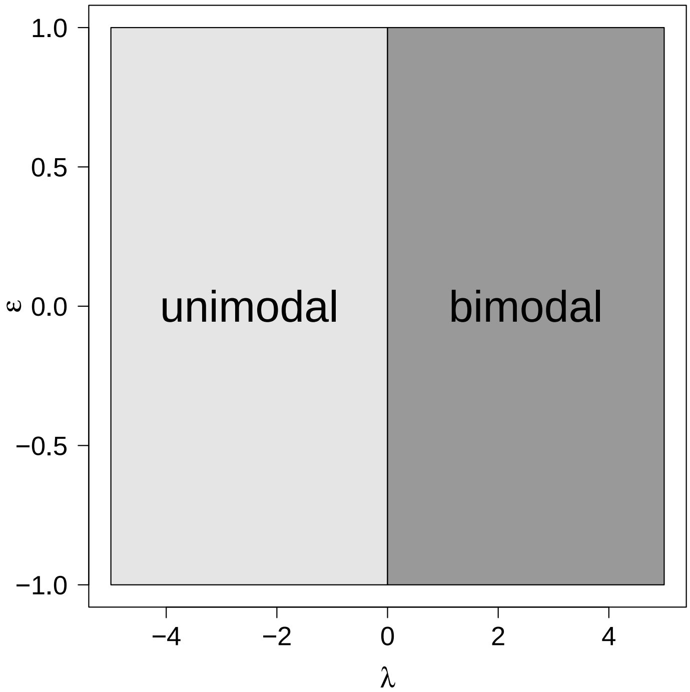


Figure 3. Regions of unimodality and bimodality for the btpn model in terms of λ and η.

2.4. Particular Cases

By construction, the following models are particular cases for the btpn distribution:

- $\text{btpn}(\sigma = \sigma/2, \lambda, \epsilon = -1) \equiv \text{tpn}(\sigma, \lambda)$;
- $\text{btpn}(\sigma, \lambda = 0, \epsilon = 0) \equiv \text{N}(0, \sigma^2)$, i.e., the normal distribution with mean 0 and variance σ^2;
- $\text{btpn}(\sigma = 1, \lambda = 0, \epsilon) \equiv \text{esn}(\sigma, \epsilon)$, i.e., the epsilon skew-normal distribution (Mudholkar and Hutson [10]).

Figure 4 summarizes the relationships among the btpn and its particular cases.

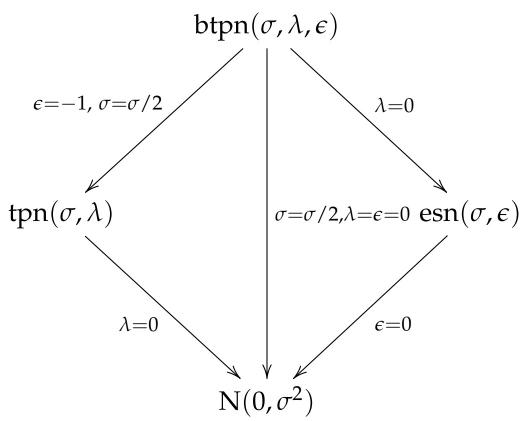


Figure 4. Particular cases for the btpn distribution.

3. Inference

In this section, we discuss the maximum likelihood (ML) method for parameter estimation for the btpn model. We also provide details about the computational aspects.

3.1. *Maximum Likelihood Function*

Hereafter, and to simplify the estimation procedure, we consider the reparameterization $\eta = \frac{\epsilon}{\sqrt{1-\epsilon^2}} \in \mathbb{R}$. Therefore, henceforth, we denote $X \sim \text{btpn}(\sigma, \lambda, \eta)$, with σ, λ and η scale, shape and asymmetry parameters, respectively, if its pdf is given by

$$f(x;\sigma,\lambda,\eta) = \begin{cases} \frac{1}{2\sigma\Phi(\lambda)}\phi\left(\frac{-x\sqrt{1+\eta^2}}{\sigma(\sqrt{1+\eta^2}+\eta)} - \lambda\right) & \text{, if } x < 0 \\ \frac{1}{2\sigma\Phi(\lambda)}\phi\left(\frac{x\sqrt{1+\eta^2}}{\sigma(\sqrt{1+\eta^2}-\eta)} - \lambda\right) & \text{, if } x \geq 0 \end{cases}$$

Given $z_1, \ldots, z_n$, a random sample from the $\text{btpn}(\sigma, \lambda, \eta)$ distribution, the log-likelihood function for $\boldsymbol{\theta} = (\sigma, \lambda, \eta)$ is given by

$$\ell(\boldsymbol{\theta}) = -n\left[\log(2\sigma\Phi(\lambda)) + \frac{1}{2}\log(2\pi) + \frac{\lambda^2}{2}\right] + \ell_1(\boldsymbol{\theta}) + \ell_2(\boldsymbol{\theta}), \tag{1}$$

where

$$\ell_1(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i:z_i\leq 0}\left\{\left(\frac{z_i\sqrt{1+\eta^2}}{\sigma(\sqrt{1+\eta^2}-\eta)}\right)^2 - \frac{2z_i\lambda\sqrt{1+\eta^2}}{\sigma(\sqrt{1+\eta^2}-\eta)}\right\}, \quad \text{and}$$

$$\ell_2(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i:z_i> 0}\left\{\left(\frac{z_i\sqrt{1+\eta^2}}{\sigma(\sqrt{1+\eta^2}+\eta)}\right)^2 + \frac{2z_i\lambda\sqrt{1+\eta^2}}{\sigma(\sqrt{1+\eta^2}+\eta)}\right\}.$$

To find the ML estimator of $\boldsymbol{\theta}$, say $\widehat{\boldsymbol{\theta}}$, we need to maximize $\ell(\boldsymbol{\theta})$ in (1) in relation to $\boldsymbol{\theta}$. However, no closed-form expressions for the ML estimates are possible. Therefore, we must use an iterative method for nonlinear optimization. For instance, we solve this problem using the Broyden-–Fletcher—Goldfarb—Shanno (BFGS) quasi-Newton method; see [11] (p. 199).

3.2. *Computational Aspects*

The ML estimators for the btpn model and the obtaining of their standard errors are included in the tpn package [12] from the R [13] software. The following function can be used to obtain these results:

```
est.btpn(y)
```

where y is the sample. The function returns a list with the estimates, the iterations used for the maximization algorithm, the log-likelihood function evaluated in the parameter estimations and the corresponding Akaike information criterion (AIC, see [14]) and the Bayesian information criterion (BIC, see [15]). Models with lower AIC and/or BIC are preferable. The package also includes the functions to drawn values to evaluate the pdf and the cdf for the btpn model named `rbtpn`, `dbptn` and `pbtpn`, respectively.

4. Simulation Study

In this section, we present a simulation study in order to evaluate the behaviour of the ML estimators in finite samples. The study was conducted using the `tpn` package [12]. Specifically, random samples were generated using the `rbtpn` function, and the estimation was performed using the `est.btpn` function. We considered 5000 Monte Carlo replicates for 3 sample sizes: 50, 100 and 200. We also considered 2 combinations for the scale parameter σ: 2 and 10; 3 values for λ: $-0.75, 1$ and 3; and 2 values for η: -0.5 and 0.75. This setting provides 36 combinations of the parameters σ, λ and η and the sample size. Tables 1 and 2 summarize the empirical bias, the standard errors of the MLE (SE), the root-mean-squared error (RMSE) and the 95% coverage probability (CP) based on the asymptotic distribution of the MLE. In general terms, the bias and RMSE terms are reduced when the sample size is increased, suggesting the consistency of the MLE. Note also that the SE and RMSE terms are closer when the sample size is increased, suggesting that the standard errors of the estimators are also well estimated. Additionally, the CP terms converge reasonably to the nominal value used to their construction (95%), suggesting that the normality is reasonable as an asymptotic distribution to the ML estimators in the btpn model, even in reasonable sample sizes.

Table 1. Empirical bias, SE, RMSE and 95% CP for the ML estimators of σ, λ and η in the btpn distribution with different combinations of parameters (case true $\sigma = 2$).

True Value			$n = 50$				$n = 100$				$n = 200$			
λ	η	par.	bias	SE	RMSE	CP	bias	SE	RMSE	CP	bias	SE	RMSE	CP
-0.75	-0.5	σ	0.894	3.271	8.403	0.814	0.170	1.089	3.165	0.856	0.035	0.514	0.584	0.898
		λ	-0.831	3.405	8.561	0.862	-0.132	1.237	3.050	0.891	-0.028	0.638	0.691	0.917
		η	-0.018	0.112	0.121	0.940	-0.008	0.078	0.081	0.945	-0.003	0.055	0.057	0.947
	0.75	σ	0.736	2.490	7.381	0.802	0.233	1.218	3.497	0.862	0.041	0.538	0.630	0.901
		λ	-0.663	2.663	7.325	0.855	-0.205	1.368	3.453	0.897	-0.029	0.660	0.728	0.924
		η	0.030	0.142	0.167	0.936	0.014	0.099	0.106	0.945	0.009	0.070	0.073	0.947
1	-0.5	σ	-0.057	0.350	0.360	0.887	-0.026	0.248	0.250	0.914	-0.012	0.175	0.176	0.936
		λ	0.073	0.406	0.416	0.938	0.035	0.285	0.289	0.946	0.014	0.201	0.201	0.949
		η	-0.014	0.086	0.095	0.932	-0.008	0.061	0.064	0.940	-0.003	0.043	0.045	0.944
	0.75	σ	-0.055	0.350	0.355	0.884	-0.024	0.248	0.251	0.919	-0.015	0.174	0.176	0.931
		λ	0.072	0.406	0.412	0.935	0.035	0.285	0.288	0.943	0.020	0.200	0.203	0.942
		η	0.028	0.109	0.127	0.940	0.012	0.076	0.082	0.942	0.006	0.054	0.055	0.943
3	-0.5	σ	-0.049	0.203	0.214	0.919	-0.022	0.145	0.152	0.930	-0.013	0.103	0.106	0.936
		λ	0.103	0.356	0.383	0.948	0.047	0.248	0.262	0.947	0.028	0.174	0.180	0.948
		η	-0.007	0.051	0.054	0.937	-0.003	0.036	0.036	0.949	-0.001	0.025	0.026	0.950
	0.75	σ	-0.048	0.203	0.214	0.917	-0.023	0.145	0.148	0.932	-0.012	0.103	0.103	0.946
		λ	0.105	0.356	0.385	0.946	0.046	0.248	0.253	0.952	0.024	0.174	0.175	0.953
		η	0.011	0.064	0.072	0.933	0.005	0.045	0.048	0.939	0.002	0.032	0.033	0.942

Table 2. Empirical bias, SE, RMSE and 95% CP for the ML estimators of σ, λ and η in the btpn distribution with different combinations of parameters (case true $\sigma = 10$).

True Value λ	η	par.	$n=50$ bias	SE	RMSE	CP	$n=100$ bias	SE	RMSE	CP	$n=200$ bias	SE	RMSE	CP
-0.75	-0.5	σ	2.279	10.597	19.571	0.802	0.961	5.254	11.695	0.864	0.221	2.538	3.411	0.900
		λ	-0.400	2.336	4.023	0.857	-0.161	1.209	2.328	0.896	-0.034	0.632	0.780	0.924
		η	-0.020	0.112	0.126	0.938	-0.009	0.079	0.083	0.949	-0.004	0.055	0.056	0.950
	0.75	σ	1.748	9.147	17.694	0.797	0.608	5.063	7.312	0.859	0.157	2.496	3.225	0.891
		λ	-0.272	2.046	3.498	0.856	-0.093	1.172	1.548	0.895	-0.021	0.623	0.734	0.920
		η	0.038	0.143	0.177	0.933	0.013	0.099	0.103	0.951	0.007	0.069	0.072	0.945
1	-0.5	σ	-0.225	1.772	1.829	0.887	-0.126	1.237	1.260	0.916	-0.057	0.876	0.891	0.932
		λ	0.065	0.408	0.415	0.938	0.037	0.285	0.289	0.940	0.015	0.201	0.205	0.943
		η	-0.014	0.086	0.095	0.936	-0.005	0.060	0.063	0.943	-0.002	0.043	0.043	0.952
	0.75	σ	-0.284	1.744	1.756	0.886	-0.136	1.236	1.255	0.913	-0.059	0.875	0.869	0.933
		λ	0.073	0.405	0.405	0.941	0.037	0.285	0.288	0.943	0.016	0.201	0.201	0.946
		η	0.028	0.109	0.131	0.926	0.012	0.076	0.081	0.945	0.008	0.054	0.056	0.946
3	-0.5	σ	1.985	2.784	30.165	0.913	1.699	1.824	26.982	0.928	1.113	0.799	22.887	0.932
		λ	0.023	0.420	1.320	0.941	-0.032	0.287	1.178	0.947	-0.025	0.184	0.955	0.941
		η	-0.004	0.051	0.054	0.934	-0.002	0.036	0.037	0.944	-0.001	0.026	0.026	0.948
	0.75	σ	0.386	4.618	13.399	0.918	0.596	1.816	14.121	0.924	0.355	0.929	10.561	0.943
		λ	0.071	0.474	0.692	0.954	0.016	0.285	0.680	0.941	0.003	0.188	0.503	0.952
		η	0.019	0.065	0.076	0.925	0.015	0.046	0.051	0.935	0.008	0.033	0.034	0.941

5. Application

In this section, we present an application to a real data set in order to illustrate the btpn model. We consider the height data set, which consists of the height of 126 students from the University of Pennsylvania (Cruz-Medina [16]). We compare our proposal with other bimodal proposals, such as the epsilon skew inverted gamma (esig, see Abdulah et al. [17]) and the alpha skew-normal (asn, Elal-Olivero [18]). The pdf for the esig model is given by:

$$f(x;\sigma,\lambda,\eta) = \frac{\lambda^{\sigma}}{2\Gamma(\sigma)} \begin{cases} \left(\frac{x}{1-\eta}\right)^{-(\sigma+1)} e^{-\frac{\lambda(1-\eta)}{x}} & x \geq 0 \\ \left(\frac{-x}{1+\eta}\right)^{-(\sigma+1)} e^{-\frac{\lambda(1+\eta)}{-x}} & x < 0 \end{cases},$$

where $\lambda > 0, \sigma > 0$ and $|\eta| < 1$ are the scale, shape and skewness parameters, respectively.

The pdf for the asn model is given by:

$$f(x;\eta,\lambda,\sigma) = \left(\frac{[1-\eta(\frac{x-\lambda}{\sigma})]^2 + 1}{\sigma(2+\eta^2)}\right)\phi\left(\frac{x-\lambda}{\sigma}\right),$$

where $\eta, \lambda \in \mathbb{R}$ and $\sigma > 0$ are the shape, location and scale parameters, respectively.

Table 3 summarizes some descriptive statistics for the sample, where we highlight the symmetrical behaviour of the data ($\sqrt{b_1} = -0.05$).

Table 3. Descriptive statistics for the *height* data set.

Data Set	n	$\overline{X}$	S^2	$\sqrt{b_1}$	b_2
weight measured	126	0	1	-0.05	3.053

Table 4 presents the estimatives, standard errors, AIC and BIC criteria for the mentioned models. Note that, based on both criteria, btpn presents a better fit than the rest of the distributions. Figure 5 shows the histogram for the height data and the pdf for the three considered distributions, where the better performance for the btpn in this data set is demonstrated. Moreover, as discussed in Proposition 6, $\widehat{\lambda} = 0.496 > 0$ implies a bimodal model, and such modes are equal to $x_1 = -0.402$ and $x_2 = 0.404$. In addition, the distribution of height is very close to symmetry.

We also compute the randomized quantile residuals [19] for the three fitted models. If the model was correctly specified, these residuals should be a random sample from the standard normal distribution. Figure 6 shows the qqplot for such residuals, also suggesting that the btpn is a more appropriated model for this data set.

Table 4. Estimated parameters and their standard errors (in parentheses) for the btpn, esig and asn models for the *height* data set. The AIC and BIC criteria are also presented.

Estimated	btpn	esig	asn
σ	0.813 (0.113)	1.304 (0.148)	0.996 (0.063)
λ	0.496 (0.316)	0.527 (0.073)	0.014 (3.422)
η	−0.002 (0.048)	0.095 (0.059)	0.014 (3.409)
AIC	360.76	415.79	362.67
BIC	369.27	424.30	375.91

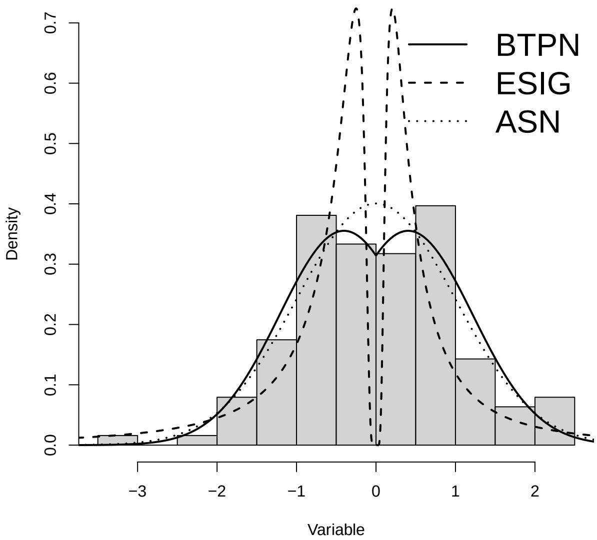
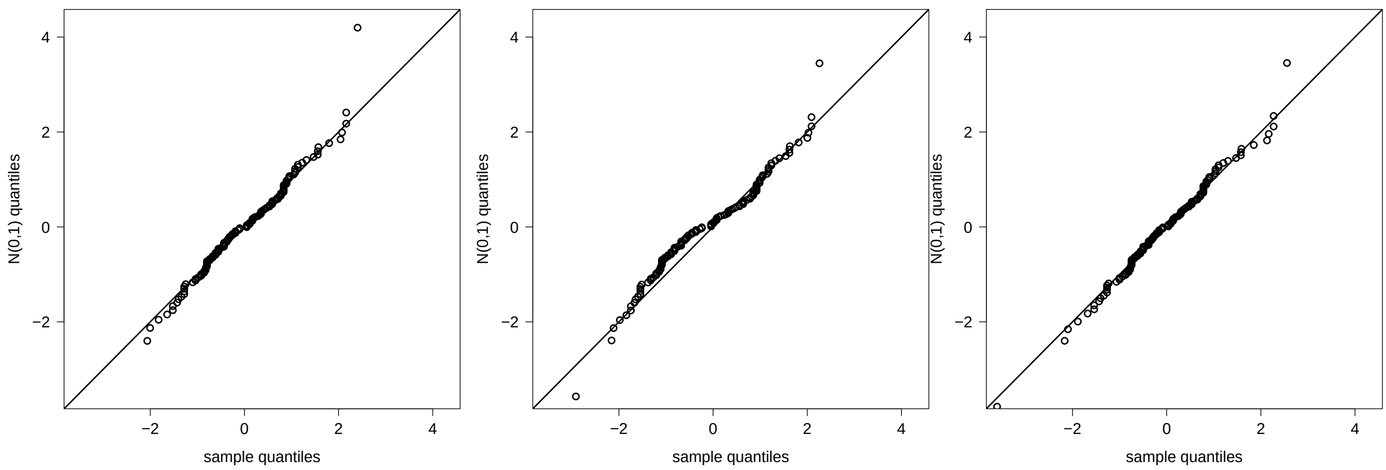


Figure 5. Histogram for the *height* data set and the estimated pdf for the btpn, esig and asn models.



(**a**) (**b**) (**c**)

Figure 6. Quantile residuals for fitted models: (**a**) asn, (**b**) esig, and (**c**) btpn.

6. Conclusions

The importance of fitting an observable dataset by a probability distribution is well-known, since it will be covered with convenient properties. Difficulty arises when the data is bimodal, because there are not traditional distributions with this property. This gap is being filled by an increasing movement in the statistical literature to develop probability distributions which already have a bimodality feature. In this paper, we made our contribution with the bimodal positive truncation normal distribution. The btpn distribution has the following advantages: support in the real line, closed-form cdf and moments, and the ability to generalize the standard normal and treatable maximum likelihood estimators. The ML procedure works very reasonably, i.e, as the sample size increases, the bias and the SE decrease. Since there are models for which the estimation procedure does not work even for large samples, the btpn distribution also has this strength. We ended the advantages of our proposed distribution with an application where btpn was the best choice of fitting. As suggestions for future work, we can mention two possibilities: the first entails the improvement of the asymptotic properties of the ML estimation through bias and variance corrections (see [20,21], respectively), and the second involves the addition of a regression structure. A closed-form cdf allows even a quantile regression structure, see [22,23], for instance, as [24] did for the gamma–sinh Cauchy distribution. For the applicability and possibilities of future works, we think the bimodal positive truncation normal distribution is useful for practitioners and researchers of many different areas.

Author Contributions: Conceptualization, H.J.G., Y.M.G. and D.I.G.; Data curation, M.C.; Formal analysis, W.E.C., Y.M.G. and M.C.; Investigation, W.E.C. and T.M.M.; Methodology, H.J.G., T.M.M. and D.I.G.; Software, H.J.G., Y.M.G. and D.I.G.; Supervision, Y.M.G. and D.I.G.; Validation, M.C.; Visualization, W.E.C.; Writing—review editing, T.M.M. and D.I.G. All authors have read and agreed to the published version of the manuscript.

Funding: The research of H.J.G. was supported by the Proyecto de Investigación de Facultad de Ingeniería. Universidad Católica de Temuco. UCT-FDI032020. W.E.C., Y.M.G., M.C. and D.I.G. also acknowledge the support of Proyecto Gidi: "La estadística como respuesta a problemas de otras áreas" supported by the University of Atacama.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ashman, K.M.; Bird, C.M.; Zepf, S.E. Detecting Bimodality in Astronomical Datasets. *Astron. J.* **1994**, *108*, 2348–2361. [CrossRef]
2. Michele, C.; Accatino, F. Tree cover bimodality in savannas and forests emerging from the switching between two fire dynamics. *PLoS ONE* **2014**, *9*, e91195. [CrossRef] [PubMed]
3. Wang, J.; Wen, S.; Symmans, W.F.; Pusztai, L.; Coombes, K.R. The bimodality index: A criterion for discovering and ranking bimodal signatures from cancer gene expression profiling data. *Cancer Inform.* **2009**, *7*, 199–216. [CrossRef] [PubMed]
4. Robertson, C.A.; Fryer, J.G. Some descriptive properties of normal mixtures. *Scand. Actuar. J.* **1969**, *3–4*, 137–146. [CrossRef]
5. Gómez, H.W.; Elal-Olivero, D.; Salinas, H.S.; Bolfarine, H. Bimodal extension based on the skew-normal distribution with application to pollen data. *Environmetrics* **2011**, *22*, 50–62. [CrossRef]
6. Venegas, O.; Salinas, H.S.; Gallardo, D.I.; Bolfarine, H.; Gómez, H.W. Bimodality based on the generalized skew-normal distribution. *J. Stat. Comput. Simul.* **2018**, *88*, 156–181. [CrossRef]
7. Butt, N.S.; Khalil, M.G. A New Bimodal Distribution for Modeling Asymmetric Bimodal Heavy-Tail Real Lifetime Data. *Symmetry* **2020**, *12*, 2058 [CrossRef]
8. Gómez, Y.M.; Gómez-Déniz, E.; Venegas, O.; Gallardo, D.I.; Gómez, H.W. An Asymmetric Bimodal Distribution with Application to Quantile Regression. *Symmetry* **2019**, *11*, 899. [CrossRef]
9. Gómez, H.J.; Olmos, N.M.; Varela, H.; Bolfarine, H. Inference for a truncated positive normal distribution. *Appl. Math. J. Chin. Univ.* **2018**, *33*, 163–176. [CrossRef]
10. Mudholkar, G.S.; Hutson, A.D. The epsilon–skew–normal distribution for analyzing near-normal data. *J. Stat. Plan. Inference* **2000**, *83*, 291–309. [CrossRef]

11. Mittelhammer, R.C.; Judge, G.G.; Miller, D.J. *Econometric Foundations*; Cambridge University Press: New York, NY, USA, 2000.
12. Gallardo, D.I.; Gómez, H.J.; Gómez, Y.M. tpn: Truncated Positive Normal Model and Extensions. R Package Version 1.1. 2021. Available online: https://cran.r-project.org/web/packages/tpn/index.html (accessed on 25 January 2022).
13. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022. Available online: https://www.R-project.org/ (accessed on 25 January 2022).
14. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
15. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
16. Cruz-Medina I.R.; Olmos, N.M. Almost Nonparametric and Nonparametric Estimation in Mixture Model. Ph.D. Thesis, Pennsylvania State University, State College, PA, USA, 2001.
17. Abdulah, E.K.; Elsalloukh, H. Bimodal Class based on the Inverted Symmetrized Gamma Distribution with Applications. *J. Stat. Appl. Probab.* **2014**, *3*, 1–7. [CrossRef]
18. Elal-Olivero, D. Alpha-skew-normal distribution. *Proyecciones* **2010**, *29*, 224–240. [CrossRef]
19. Dunn, P.K.; Smyth, G.K. Randomized quantile residuals. *J. Comput. Graph. Stat.* **1996**, *5*, 236–244.
20. Magalhães, T.M.; Gómez, Y.M.; Gallardo, D.I.; Venegas, O. Bias reduction for the Marshall-Olkin extended family of distributions with application to an airplane's air conditioning system and precipitation data. *Symmetry* **2020**, *12*, 851. [CrossRef]
21. Magalhães, T.M.; Botter, D.A.; Sandoval, M.C. A general expression for second-order covariance matrices—An application to dispersion models. *Braz. J. Probab. Stat.* **2021**, *35*, 37–49. [CrossRef]
22. Cade, B.S.; Noon, B.R. A gentle introduction to quantile regression for ecologists. *Front. Ecol. Environ.* **2003**, *1*, 412–420. [CrossRef]
23. Alencar, A.P.; Santos, B.R. Association of pollution with quantiles and expectations of the hospitalization rate of elderly people by respiratory diseases in the city of São Paulo, Brazil. *Environmetrics* **2014**, *25*, 165–171. [CrossRef]
24. Gómez, Y.M.; Gallardo, D.I.; Venegas, O.; Magalhães, T.M. An asymmetric bimodal double regression model. *Symmetry* **2021**, *13*, 2279. [CrossRef]