

Article

Flexible Power-Normal Models with Applications

Guillermo Martínez-Flórez ¹, Diego I. Gallardo ², Osvaldo Venegas ^{3,*}, Heleno Bolfarine ⁴
and Héctor W. Gómez ⁵

¹ Departamento de Matemáticas y Estadística, Facultad de Ciencias, Universidad de Córdoba, Córdoba 2300, Colombia; guillermomartinez@correo.unicordoba.edu.co

² Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Copiapó 1530000, Chile; diego.gallardo@uda.cl

³ Departamento de Ciencias Matemáticas y Físicas, Facultad de Ingeniería, Universidad Católica de Temuco, Temuco 4780000, Chile

⁴ Departamento de Estatística, IME, Universidade de São Paulo, São Paulo 05508-090, Brazil; hbolfar@ime.usp.br

⁵ Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile; hector.gomez@uantof.cl

* Correspondence: ovenegas@uct.cl; Tel.: +56-45-2205351

Abstract: The main object of this paper is to propose a new asymmetric model more flexible than the generalized Gaussian model. The probability density function of the new model can assume bimodal or unimodal shapes, and one of the parameters controls the skewness of the model. Three simulation studies are reported and two real data applications illustrate the flexibility of the model compared with traditional proposals in the literature.

Keywords: power normal distribution; bimodal asymmetric distribution; moments; maximum likelihood estimators; fisher information matrix



Citation: Martínez-Flórez, G.; Gallardo, D.I.; Venegas, O.; Bolfarine, H.; Gómez, H.W. Flexible Power-Normal Models with Applications. *Mathematics* **2021**, *9*, 3183. <https://doi.org/10.3390/math9243183>

Academic Editor: Christopher Goodrich

Received: 18 November 2021
Accepted: 6 December 2021
Published: 10 December 2021

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



Copyright: © 2021 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

Azzalini [1] introduced the skew-normal (SN) model and Durrans [2] introduced the generalized Gaussian model, also known as the power-normal (PN) model. A limitation of those models is that they only produce uni-modal asymmetric densities, and are therefore not appropriate for fitting bimodal data. An approach typically used for fitting multimodal data is the mixture of normal or asymmetric-normal models, which present difficulties such as identifiability problems (see McLachlan and Peel [3]; Marin et al. [4]) and complicated numerical implementation. Bimodal distributions generated from skew-symmetric distributions can be found in Azzalini and Capitanio [5], Ma and Genton [6], Arellano-Valle et al. [7], Kim [8], Lin et al. [9,10], Elal-Olivero et al. [11], Arnold et al. [12,13], Gómez et al. [14], Braga et al. [15], Venegas et al. [16] and Gómez-Déniz et al. [17], among others.

It is therefore of interest to study asymmetric models that are adequate for fitting possible bimodal data. Hence, the main object of this paper is to propose asymmetric distributions that are more flexible than the SN and PN models; and are also useful for modeling bimodal data, which are very common in practical situations. The proposed model is not a mixture type model, to avoid the identifiability problems mentioned above. Our proposal considers a family of distributions which can adopt both a unimodal shape (competing with the SN and PN models) and a bimodal shape (competing with the mixture of normals).

The following Lemma is introduced in Gómez et al. [14], and is crucial for the construction of the new family of distributions.

Lemma 1. Let f be a symmetric around zero probability density function (pdf), F its respective cumulative distribution function (cdf) and G an absolutely continuous cdf such that G' is a symmetric pdf around zero. Then,

$$g(z; \lambda, \delta) = c_\delta f(|z| + \delta) G(\lambda z), \quad z, \lambda, \delta \in \mathbb{R} \quad (1)$$

is the pdf of a random variable, where $c_\delta^{-1} = 1 - F(\delta)$.

Gómez et al. [14] studied the skew-flexible-normal (SFN) distribution with pdf given by:

$$f(z; \lambda, \delta) = c_\delta \phi(|z| + \delta) \Phi(\lambda z), \quad z, \lambda, \delta \in \mathbb{R}. \quad (2)$$

The authors show that, for $\delta < 0$, the model in (2) is bimodal. We denote this model as SFN(λ, δ).

The “Lehmann’s alternatives” model proposed in Lehmann [18] is based on the distribution of the maximum of a sample. It is a good alternative for extending models because it incorporates a higher asymmetry and/or kurtosis than the basal model. The cdf of this model is given by:

$$\mathcal{F}_F(z; \alpha) = \{F(z)\}^\alpha, \quad z \in \mathbb{R}, \alpha > 0, \quad (3)$$

where F is a cdf. For $\alpha \in \mathbb{Z}$, the cdf corresponds to the random variable defined as (Z_1, \dots, Z_α) , where the Z_i s are independent and identically distributed with common cdf F . Durrans [2] extends this interpretation to fractional order statistics for $\alpha \in \mathbb{R}$. We denote a random variable with cdf defined in (3) as $Z \sim P_F(\alpha)$. The case, $F = \Phi$, corresponds to the PN model with pdf given by

$$\varphi(z; \alpha) = \alpha \phi(z) \{\Phi(z)\}^{\alpha-1}, \quad z \in \mathbb{R}, \alpha > 0, \quad (4)$$

which we denote by $Z \sim PN(\alpha)$. This model has been studied in more detail in Gupta and Gupta [19]. Pewsey et al. [20] derived its Fisher information matrix, showing that it is non-singular in a neighborhood of symmetry point ($\alpha = 1$), which is not satisfied for the SN distribution under the symmetry hypothesis, that is, at $\lambda = 0$, because its Fisher information matrix is singular.

The paper is organized as follows. In Section 2, the flexible power-normal (FPN) model is introduced and some of its main properties are discussed. An inference of the model via maximum likelihood (ML) estimation is considered in Section 3. Three procedures to draw values from the FPN distribution are discussed in Section 4. Three simulation studies are carried out in Section 5 in order to assess different aspects of the model. Two real data applications are reported in Section 6, illustrating the usefulness of the model compared with other common distributions in the literature. Finally, Section 7 presents a discussion of the main results of the paper.

2. Flexible Power-Normal Distribution

In this section, we introduce a more flexible model than the power-normal model introduced in (4) by incorporating a new shape parameter δ . One of the interesting features of this model is that, for certain values of shape parameter δ , a bimodal pdf can be obtained.

Definition 1. A random variable Z follows a FPN distribution with parameters α and δ , if its pdf is given by:

$$\varphi(z; \alpha, \delta) = k(\alpha, \delta) \phi(|z| + \delta) \{\Phi(z)\}^{\alpha-1}, \quad z, \delta \in \mathbb{R}, \alpha \in \mathbb{R}^+, \quad (5)$$

where

$$k(\alpha, \delta) = k_{\alpha, \delta} = \left[\sqrt{2\pi} \phi(\delta) \int_0^1 u^{\alpha-1} \exp\{-\delta |\Phi^{-1}(u)|\} du \right]^{-1},$$

is the normalizing constant. We use the notation $Z \sim FPN(0, 1, \alpha, \delta)$.

2.1. Some Particular Cases of the FPN Model

For the FPN distribution, we have the following particular cases.

Proposition 1. If $Z \sim FPN(0, 1, \alpha, \delta)$, then

1. $\varphi(z; 0, 1, \alpha = 1, \delta = 0) = \phi(z)$.
2. $\varphi(z; 0, 1, \alpha, \delta = 0) = \alpha\phi(z)\{\Phi(z)\}^{\alpha-1}$.
3. $\varphi(z; 0, 1, \alpha = 2, \delta = 0) = 2\phi(z)\Phi(z)$.
4. $\varphi(z; 0, 1, \alpha = 2, \delta) = c_\delta\phi(|z| + \delta)\Phi(z)$.

Remark 1. Note that points 1 and 2 show that N and PN distributions are particular cases of the FPN model. Points 3 and 4 show that the model contains particular cases of the skew-normal and skew-flexible-normal models (SN(1) and SFN(1,δ), respectively).

2.2. Properties of the pdf for the FPN Model

The following propositions illustrate a point where the pdf of the FPN model is not differentiable.

Proposition 2. Let $Z \sim FPN(0, 1, \alpha, \delta)$. If $\delta \neq 0$, then the pdf of the FPN is not differentiable at $z = 0$.

Proof. Note that

$$\lim_{z \rightarrow 0^+} \varphi(z; \alpha, \delta) = (1/2)^{\alpha-2} k_{\alpha, \delta} \phi(\delta) \left(\frac{\alpha - 1}{\sqrt{2\pi}} - \frac{\delta}{2} \right), \quad \text{and}$$

$$\lim_{z \rightarrow 0^-} \varphi(z; \alpha, \delta) = (1/2)^{\alpha-2} k_{\alpha, \delta} \phi(\delta) \left(\frac{\alpha - 1}{\sqrt{2\pi}} + \frac{\delta}{2} \right),$$

from which we conclude that the pdf of the FPN model is not differentiable at $z = 0$ as long as $\delta \neq 0$. □

Proposition 3. Let $Z \sim FPN(0, 1, \alpha, \delta)$. If $\delta < 0$, the distribution of the random variable Z is bimodal.

Proof. Equating to zero the first derivative for the pdf of the random variable Z with FPN distribution, we have the following cases:

1. If $z < 0$, then it follows that the solution is given by $z_1 = (\alpha - 1)\phi(z_1)/\Phi(z_1) + \delta$.
2. If $z \geq 0$, we have the solution $z_2 = (\alpha - 1)\phi(z_2)/\Phi(z_2) - \delta$.

Moreover, if $\alpha \rightarrow 1^-$, $z_1 \in \mathbb{R}^-$ and if $\alpha \rightarrow 1^+$, $z_2 \in \mathbb{R}^+$. Hence, z_1 and z_2 are distinct values. Therefore, Z is a random variable with a bimodal distribution. □

Corollary 1. Let $Z \sim FPN(0, 1, \alpha, \delta)$. If $\alpha = 1$ and $\delta < 0$, then random variable Z is symmetric and bimodal.

Proof. If $\alpha = 1$ then $\varphi(z; \alpha = 1, \delta) = (2c_\delta)^{-1}\phi(|z| + \delta)$, $z, \delta \in \mathbb{R}$. As ϕ is a symmetric function, we conclude that Z is a symmetric random variable. □

Figure 1 shows plots of the pdf for the $FPN(0, 1, \alpha, \delta)$ for different values of α and δ . As mentioned previously, for $\delta < 0$ a bimodal pdf is obtained, but one of the modes “disappears” when α moves away from one ($\alpha \rightarrow 0^-$ or $\alpha \rightarrow +\infty$), whereas for $\delta \geq 0$ a unimodal pdf is obtained with different degrees of kurtosis (leptokurtic, mesokurtic or platykurtic) depending on the value of α . On the other hand, $\alpha < 1$ and $\alpha > 1$ produce a positive and negative skew effect in the pdf, respectively, whereas $\alpha = 1$ corresponds to a symmetrical pdf.

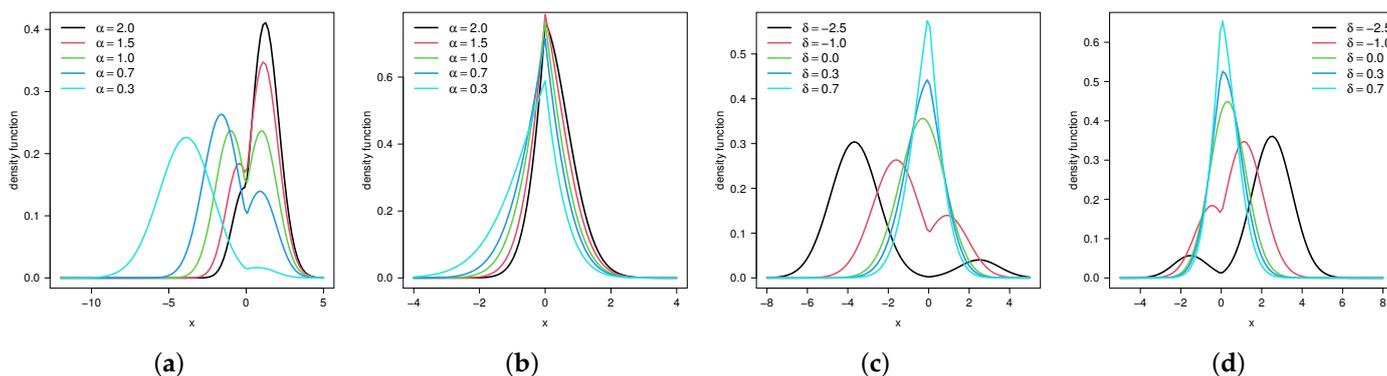


Figure 1. Pdf for $FPN(0, 1, \alpha, \delta)$ with different combinations of α and δ : (a) $FPN(0, 1, \alpha, \delta = -1)$ and varying α ; (b) $FPN(0, 1, \alpha, \delta = 1)$ and varying α ; (c) $FPN(0, 1, \alpha = 0.7, \delta)$ and varying δ ; and (d) $FPN(0, 1, \alpha = 1.5, \delta)$ and varying δ .

2.3. Moments

Definition 2. Let $Z \sim FPN(0, 1, \alpha, \delta)$. For $r = 1, 2, \dots$, we define the lower and upper r -th incomplete moments of Z as

$$\mu_r^-(Z) = \int_{-\infty}^z y^r \varphi(y; \alpha, \delta) dy \quad \text{and} \quad \mu_r^+(Z) = \int_z^{\infty} y^r \varphi(y; \alpha, \delta) dy,$$

respectively. With those definitions, the r -th moment of Z is given by

$$\mu_r = E(Z^r) = \mu_r^-(Z) + \mu_r^+(Z).$$

The next result presents a recursive expression for computing the moments of the FPN random variable.

Proposition 4. Let $Z \sim FPN(0, 1, \alpha, \delta)$. The r -th moment of the random variable Z is given by:

$$\begin{aligned} \mu_r &= (r - 1)\mu_{r-2}(0) + \delta k_{\alpha, \delta} [\mu_{r-1}^-(0) - \mu_{r-1}^+(0)] + \\ &+ (\alpha - 1)k_{\alpha, \delta} \int_0^1 [\Phi^{-1}(v)]^{r-1} \phi[|\Phi^{-1}(v)| + \delta] v^{\alpha-2} dv. \end{aligned} \tag{6}$$

The r -th central moment, say $\mu'_r = E[(Z - E(Z))^r]$, for $r = 2, 3, 4$, can be computed using the expressions:

$$\mu'_2 = \mu_2 - \mu_1^2, \quad \mu'_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 \quad \text{and} \quad \mu'_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4.$$

With those definitions, the variance and the asymmetry and kurtosis coefficients are given by:

$$\sigma^2 = \text{Var}(Z) = \mu'_2, \quad \sqrt{\beta_1} = \mu'_3 / [\mu'_2]^{3/2} \quad \text{and} \quad \beta_2 = \mu'_4 / [\mu'_2]^2,$$

respectively. Figure 2 shows the asymmetry and kurtosis coefficients for the $FPN(0, 1, \alpha, \delta)$ model with different values of α and δ .

2.4. The Location-Scale Extension

We consider now the location-scale extension for the model. The FPN distribution with location parameter $\xi \in \mathbb{R}$ and scale parameter $\eta \in \mathbb{R}^+$, is defined as the distribution of the random variable $X = \xi + \eta Z$, for which the pdf is given by:

$$\varphi(x; \xi, \eta, \alpha, \delta) = \frac{k_{\alpha, \delta}}{\eta} \phi\left(\frac{|x - \xi|}{\eta} + \delta\right) \left\{ \Phi\left(\frac{x - \xi}{\eta}\right) \right\}^{\alpha-1}, \quad x, \xi, \delta \in \mathbb{R}, \eta, \alpha \in \mathbb{R}^+. \tag{7}$$

If X has a density function as in the Equation (7) model, we use the notation $X \sim \text{FPN}(\xi, \eta, \alpha, \delta)$.

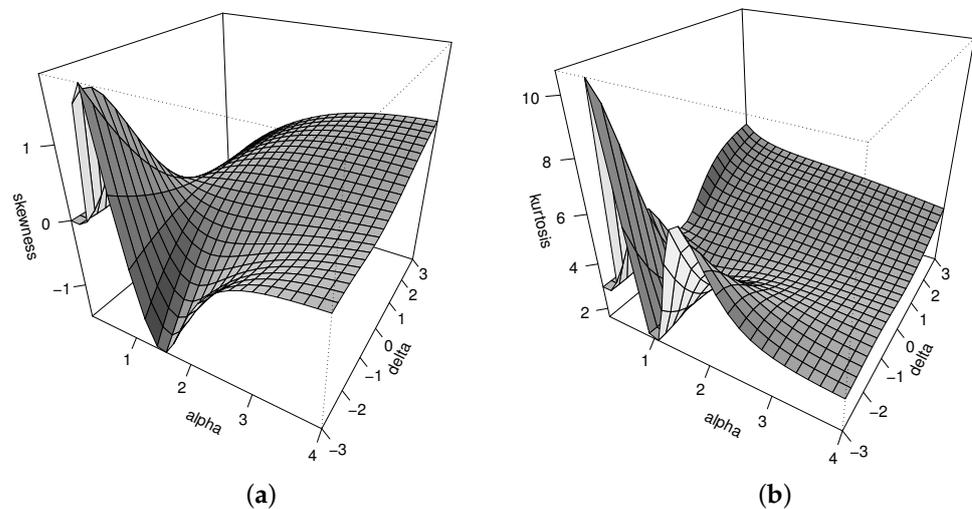


Figure 2. (a) Skewness and (b) kurtosis coefficients for the $\text{FPN}(0, 1, \alpha, \delta)$ model.

3. Inference for the FPN Model

In this section, we discuss the ML estimation of the parameters of the FPN model. Observed and expected (Fisher) information matrices are derived for the standard and location-scale models. For the location-scale situation, we have a four-dimension parameter vector, $\theta = (\xi, \eta, \alpha, \delta)^\top$.

3.1. Standard Case

Considering that Z_1, \dots, Z_n are independent and identically distributed with standard FPN distribution, the log-likelihood function for (α, δ) based on the sample $\mathbf{Z} = (z_1, \dots, z_n)$ is given by:

$$\ell(\alpha, \delta; \mathbf{Z}) = n \log\left(\frac{k_{\alpha, \delta}}{\sqrt{2\pi}}\right) - \frac{1}{2} \sum_{i=1}^n (|z_i| + \delta)^2 + (\alpha - 1) \sum_{i=1}^n \log\{\Phi(z_i)\}, \quad (8)$$

from which the elements of the score function are given by:

$$U(\alpha) = n \left\{ -\sqrt{2\pi}h(\alpha, \delta) + \bar{u} \right\} \quad \text{and} \quad U(\delta) = n \left\{ \sqrt{2\pi}g(\alpha, \delta) - \overline{|z|} - \delta \right\}, \quad (9)$$

where

$$h(\alpha, \delta) = \phi(\delta)k_{\alpha, \delta} \int_0^1 z^{\alpha-1} \log z \exp\{-\delta|\Phi^{-1}(z)|\} dz, \quad \bar{u} = \sum_{i=1}^n u_i/n,$$

$$g(\alpha, \delta) = \phi(\delta)k_{\alpha, \delta} \int_0^1 (|\Phi^{-1}(z)| + \delta) z^{\alpha-1} \exp\{-\delta|\Phi^{-1}(z)|\} dz,$$

and $\overline{|z|} = \sum_{i=1}^n |z_i|/n$.

Equating the above derivatives to zero, we obtain the ML equations, which are given by:

$$\bar{u} = \sqrt{2\pi}h(\alpha, \delta) \quad \text{and} \quad \delta = \sqrt{2\pi}g(\alpha, \delta) - \overline{|z|}.$$

To solve the above equations, a numerical procedure is required since analytical solutions are not available. To derive the Fisher information matrix, we consider the second derivatives of the log-likelihood function for parameters (α, δ) , which we denote by (j_{xy}) ; these are given by:

$$j_{\alpha\alpha} = n[-\Delta_{1\alpha}(\alpha, \delta) + \Delta_{2\alpha}(\alpha, \delta)], \quad j_{\delta\delta} = n[-\sqrt{2\pi}\Delta_{1\delta}(\alpha, \delta) + \Delta_{2\delta}(\alpha, \delta)], \quad \text{and}$$

$$j_{\delta\delta} = n[-\Delta_{1\delta}(\alpha, \delta) + \Delta_{2\delta}(\alpha, \delta) + 1],$$

where

$$\begin{aligned} \Delta_{1\alpha}(\alpha, \delta) &= \left(\frac{\partial \log k_{\alpha,\delta}}{\partial \alpha}\right)^2, & \Delta_{2\alpha}(\alpha, \delta) &= k_{\alpha,\delta} \frac{\partial^2 k_{\alpha,\delta}^{-1}}{\partial \alpha^2}, & \Delta_{1\alpha\delta}(\alpha, \delta) &= g(\alpha, \delta) \left(\frac{\partial \log(k_{\alpha,\delta})}{\partial \alpha}\right), \\ \Delta_{1\delta}(\alpha, \delta) &= \left(\frac{\partial \log k_{\alpha,\delta}}{\partial \delta}\right)^2, & \Delta_{2\delta}(\alpha, \delta) &= k_{\alpha,\delta} \frac{\partial^2 k_{\alpha,\delta}^{-1}}{\partial \delta^2}, & \Delta_{2\alpha\delta}(\alpha, \delta) &= k_{\alpha,\delta} \frac{\partial^2 k_{\alpha,\delta}^{-1}}{\partial \alpha \partial \delta}. \end{aligned}$$

The elements of the Fisher information matrix are obtained by computing the expected values of the corresponding elements of the observed information matrix.

3.2. The Location-Scale Model

For a random sample of size n , X_1, \dots, X_n , from the $FPN(\zeta, \eta, \alpha, \delta)$ distribution, the likelihood function for $\theta = (\zeta, \eta, \alpha, \delta)^T$ based on the sample $\mathbf{X} = (x_1, \dots, x_n)$ can be written as:

$$\begin{aligned} \ell(\theta; \mathbf{X}) &= n\{\log(k_{\alpha,\delta}) - \log(\eta) - \log(\sqrt{2\pi})\} - \frac{1}{2} \sum_{i=1}^n z_i^2 - \frac{n}{2} \delta^2 - \delta \sum_{i=1}^n |z_i| + \\ &+ (\alpha - 1) \sum_{i=1}^n \log\{\Phi(z_i)\}, \end{aligned} \tag{10}$$

where $z_i = \frac{x_i - \zeta}{\eta}$. The elements of the score function are given by:

$$\begin{aligned} U(\zeta) &= \frac{1}{\eta} \sum_{i=1}^n z_i + \frac{\delta}{\eta} \sum_{i=1}^n \text{sign}(z_i) - \frac{\alpha - 1}{\eta} \sum_{i=1}^n w_i, \\ U(\eta) &= -\frac{n}{\eta} + \frac{1}{\eta} \sum_{i=1}^n z_i^2 + \frac{\delta}{\eta} \sum_{i=1}^n z_i \text{sign}(z_i) - \frac{\alpha - 1}{\eta} \sum_{i=1}^n z_i w_i, \\ U(\alpha) &= -\sqrt{2\pi} n h(\alpha, \delta) + \sum_{i=1}^n \log\{\Phi(z_i)\} = n\{-\sqrt{2\pi} h(\alpha, \delta) + \bar{u}\}, \text{ and} \\ U(\delta) &= \sqrt{2\pi} n g(\alpha, \delta) - \sum_{i=1}^n |z_i| - n\delta = n\{\sqrt{2\pi} g(\alpha, \delta) - \bar{|z|} - \delta\}, \end{aligned}$$

where $u_i = \log \Phi(z_i)$, $w_i = \phi(z_i) / \Phi(z_i)$ and sign is the sign function. The solutions for the score equations are given by:

$$\begin{aligned} \delta &= \sqrt{2\pi} g(\alpha, \delta) - \bar{|z|}, & (\alpha - 1)(\bar{z}\bar{w}) + 1 &= \bar{z}^2 + \delta \bar{z} \text{sign}(\bar{z}), \\ \bar{u} &= \sqrt{2\pi} h(\alpha, \delta) & \text{and} & \bar{z} + \delta \text{sign}(\bar{z}) = (\alpha - 1)\bar{w}, \end{aligned}$$

where

$$\bar{w} = \sum_{i=1}^n w_i / n, \bar{z} = \sum_{i=1}^n z_i / n, \bar{|z|} = \sum_{i=1}^n |z_i| / n, \bar{z}^2 = \sum_{i=1}^n z_i^2 / n \text{ and } \bar{z}\bar{w} = \sum_{i=1}^n z_i w_i / n.$$

3.3. Expected Information Matrix

For an observation ($z_i = z$) we have that:

$$\ell_i(\theta; z) = \log(k(\alpha, \delta)) - \log(\eta) - \log \sqrt{2\pi} - \frac{1}{2}(|z| + \delta)^2 + (\alpha - 1) \log\{\Phi(z)\}, \tag{11}$$

where $z = (x - \zeta) / \eta$. The second derivatives of the log-likelihood function (11) with respect to the model parameters are given by:

$$\begin{aligned}
 j_{\xi\xi}^i &= \frac{1}{\eta^2} [1 + (\alpha - 1)(zw + w^2)], & j_{\alpha\xi}^i &= \frac{w}{\eta}, & j_{\delta\xi}^i &= -\frac{1}{\eta} \text{sign}(z), \\
 j_{\eta\xi}^i &= \frac{1}{\eta^2} \left\{ \text{sign}(z)(|z| + \delta) + z + (\alpha - 1) [z^2w + zw^2 - w] \right\}, \\
 j_{\eta\eta}^i &= \frac{1}{\eta^2} [-1 + 2\text{sign}(z)(|z| + \delta)z + z^2 - (\alpha - 1) (2zw - z^2w^2 - z^3w)], \\
 j_{\alpha\eta}^i &= \frac{zw}{\eta}, & j_{\delta\eta}^i &= -\frac{z}{\eta} \text{sign}(z), & j_{\alpha\alpha}^i &= -\Delta_{1\alpha}(\alpha, \delta) + \Delta_{2\alpha}(\alpha, \delta), \\
 j_{\delta\alpha}^i &= [-\sqrt{2\pi}\Delta_{1\alpha\delta}(\alpha, \delta) + \Delta_{2\alpha\delta}(\alpha, \delta)], & \text{and} & & j_{\delta\delta}^i &= -\Delta_{1\delta}(\alpha, \delta) + \Delta_{2\delta}(\alpha, \delta) + 1,
 \end{aligned}$$

where $\Delta_{1\alpha}$, $\Delta_{2\alpha}$, $\Delta_{1\delta}$, $\Delta_{2\delta}$, $\Delta_{1\alpha\delta}$ and $\Delta_{2\alpha\delta}$ are defined in Section 3.1. Note that $z\text{sign}(z) = |z|$, $|z|\text{sign}(z) = z$, $E(\text{sign}(Z)) = \mu_0^+(\xi) - \mu_0^-(\xi)$ and $E(|Z|) = \mu_1^+(\xi) - \mu_1^-(\xi)$. Denoting $i_{\xi\xi}, i_{\eta\xi}, \dots, i_{\delta\delta}$ as the expected values of the elements of the observed information matrix and defining $a_{kj} = E(Z^k W^j)$ for $k = 0, 1, 2, 3$, and $j = 0, 1, 2$ we obtain:

$$\begin{aligned}
 i_{\xi\xi} &= \frac{1}{\eta^2} [1 + (\alpha - 1)(a_{11} + a_{02})], & i_{\alpha\xi} &= \frac{a_{01}}{\eta}, & i_{\delta\xi} &= -\frac{\mu_0^+(\xi) - \mu_0^-(\xi)}{\eta}, \\
 i_{\eta\xi} &= \frac{1}{\eta^2} \left\{ \delta(\mu_0^+(\xi) - \mu_0^-(\xi)) + 2a_{10} + (\alpha - 1)(a_{21} + a_{12} - a_{01}) \right\}, \\
 i_{\eta\eta} &= \frac{1}{\eta^2} [-1 + 2\delta(\mu_1^+(\xi) - \mu_1^-(\xi)) + 3a_{20} - (\alpha - 1)(2a_{11} - a_{22} - a_{31})], \\
 i_{\alpha\eta} &= \frac{a_{11}}{\eta}, & i_{\delta\eta} &= -\frac{\mu_1^+(\xi) - \mu_1^-(\xi)}{\eta}, & i_{\alpha\alpha} &= -E(\Delta_{1\alpha}(\alpha, \delta)) + E(\Delta_{2\alpha}(\alpha, \delta)), \\
 i_{\delta\alpha} &= [-\sqrt{2\pi}E(\Delta_{1\alpha\delta}(\alpha, \delta)) + E(\Delta_{2\alpha\delta}(\alpha, \delta))], & \text{and} & & & \\
 i_{\delta\delta} &= -E(\Delta_{1\delta}(\alpha, \delta)) + E(\Delta_{2\delta}(\alpha, \delta)) + 1.
 \end{aligned}$$

a_{kj} must be computed numerically since they have no closed forms.

4. Simulating Values from the FPN Model

In this section, we discuss three methods for drawing values from the FPN distribution. This model has no stochastic representation, nor does it have a tractable cdf, making data simulation by conventional methods difficult. On the other hand, as the FPN is a location-scale model, we only discuss the case $\xi = 0$ and $\eta = 1$.

4.1. Acceptance-Rejection Method: Way 1

This method can be used for $\delta \geq 0$. Note that the pdf of the FPN(0, 1, α , δ) model can be written as

$$\varphi(z; \alpha, \delta) = \underbrace{\alpha^{-1} k_{\alpha, \delta} \exp\left(-\frac{\delta^2}{2}\right)}_c \times \underbrace{\exp(-\delta|z|)}_{g_1(z; \delta)} \times \underbrace{\alpha \phi(z) \{\Phi(z)\}^{\alpha-1}}_{h_1(z; \alpha)},$$

where $g_1(z; \delta) \in [0, 1]$, $\forall z \in \mathbb{R}, \forall \delta \geq 0$ and $h_1(z; \alpha)$ is the pdf of the PN(α). For this reason, to draw values from the standard FPN model, we can use the following algorithm:

1. Simulate $U \sim U(0, 1)$.
2. Simulate $Y \sim PN(\alpha)$.
 - Simulate $U_1 \sim U(0, 1)$.
 - Do $Y = \Phi^{-1}(U_1^{1/\alpha})$.
3. If $U \leq \exp(-\delta|Y|)$, accept Y . Otherwise, back to step 1.

The probability of a value being accepted is $1/C = \alpha k_{\alpha,\delta}^{-1} \exp(\delta^2/2)$.

4.2. Acceptance-Rejection Method: Way 2

This method can be used for $\alpha \geq 1$. Note that:

$$\varphi(z; \alpha, \delta) = \underbrace{2k_{\alpha,\delta}c_\delta^{-1}}_C \times \underbrace{\Phi^{\alpha-1}(z)}_{g_1(z;\alpha)} \times \underbrace{0.5c_\delta\phi(|z| + \delta)}_{h_1(z;\alpha)},$$

where $g_1(z; \alpha) \in [0, 1], \forall z \in \mathbb{R}, \forall \alpha \geq 1$ and $h_1(z; \delta)$ is the pdf of the $SFN(0, \delta)$. For this reason, to draw values from the standard FPN model, we can use the following algorithm:

1. Simulate $U \sim U(0, 1)$.
2. Simulate $Y \sim SFN(0, \delta)$.
 - Simulate $U_1, U_2 \sim U(0, 1)$ independently.
 - Do $Y_1 = \Phi^{-1}(\Phi(\delta) + U_1[1 - \Phi(\delta)]) - \delta$.
 - If $U_2 \leq 1/2$, do $S = 1$. Otherwise, do $S = -1$.
 - Do $Y = Y_1 \times S$.
3. If $U \leq \{\Phi(Y)\}^{\alpha-1}$, accept Y . Otherwise, back to step 1.

The probability of a value being accepted is $1/C = 0.5k_{\alpha,\delta}^{-1}c_\delta$. Figure 3 shows the constant C for the two ways discussed of drawing values from the FPN model. Results suggest that neither method is appropriate for drawing values from the FPN model for large values of δ .

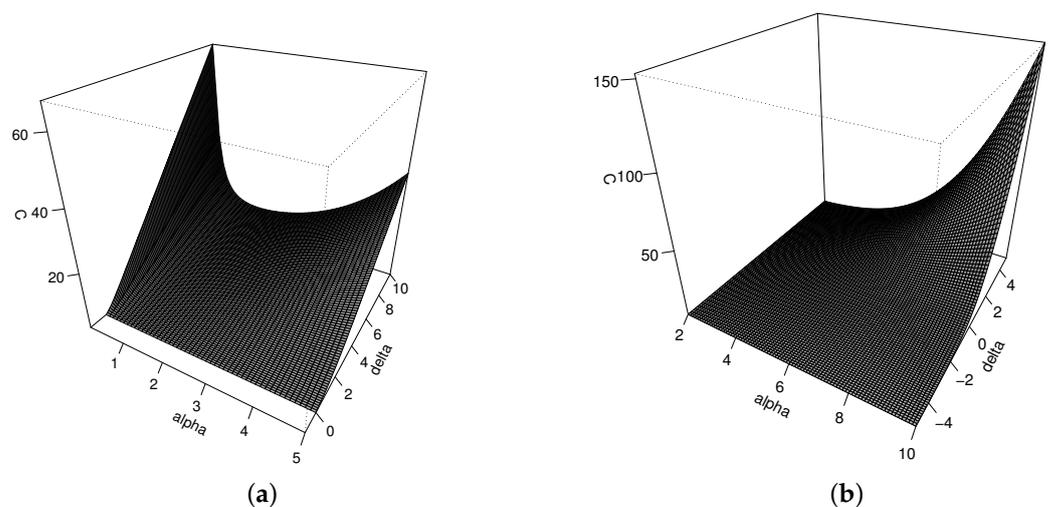


Figure 3. Constant C for the acceptance-rejection method to draw values from FPN model: (a) way 1 and (b) way 2.

4.3. A Metropolis-Hastings Algorithm

The Metropolis–Hastings algorithm is used extensively in a Bayesian context, because it allows values to be drawn from a given pdf, as long as we know its kernel (and not necessarily a possible normalizing constant). For the FPN distribution, that kernel is given by:

$$\varphi^*(z; \alpha, \delta) \propto \phi(|z| + \delta) \{\Phi(z)\}^{\alpha-1}, \quad z, \delta \in \mathbb{R}, \alpha \in \mathbb{R}^+.$$

The method requires a proposal distribution, for which we considered the $N(0, \sigma^2)$ model. The detail of the algorithm is as follows:

- Define an initial value z_0 .
- For $i = 1, \dots, n$, draw $v_i \sim N(0, \sigma^2)$ and $U_i \sim U(0, 1)$. Do $z_i^* = z_{i-1} + v_i$.

- Define $p_i = \varphi^*(z_i^*; \alpha, \delta) / \varphi^*(z_{i-1}; \alpha, \delta)$.
- If $U_i \leq p_i$, do $z_i = z_i^*$. Otherwise, $z_i = z_{i-1}$.

z_1, \dots, z_n represent a sample of size n from the FPN distribution. In order to decrease the correlation of the values drawn, we draw $b + t \times n$ values. The first b values are discarded (the burn-in) and for t values we only consider 1 (the thin). In particular, we consider $b = 1000, t = 10$ and $\sigma = 1$, showing a satisfactory performance of the method for all the combinations of α and δ considered in this work.

5. Simulation Studies

In this Section, we present three simulation studies. The first study illustrates the performance of the three simulation procedures in drawing values from the model discussed in Section 4. The second study is devoted to assessing the performance of the ML estimators in finite samples. The third study shows a model selection problem to select among the N, PN and FPN models when the data are generated under different scenarios.

5.1. Assessing the Simulation Procedures for the FPN Model

In this Section, we study the three methods discussed in Section 4 to simulate values from the FPN distribution. Figure 4 shows the histogram versus the theoretical curve for 10,000 values drawn from three combinations of parameters for the FPN model, using the three methods discussed previously. Note that, except for restrictions in the acceptance-rejection method given in the two ways (i.e., $\delta \geq 0$ and $\alpha \geq 1$ for way 1 and way 2, respectively) the three methods seem suitable for drawing values from the model.

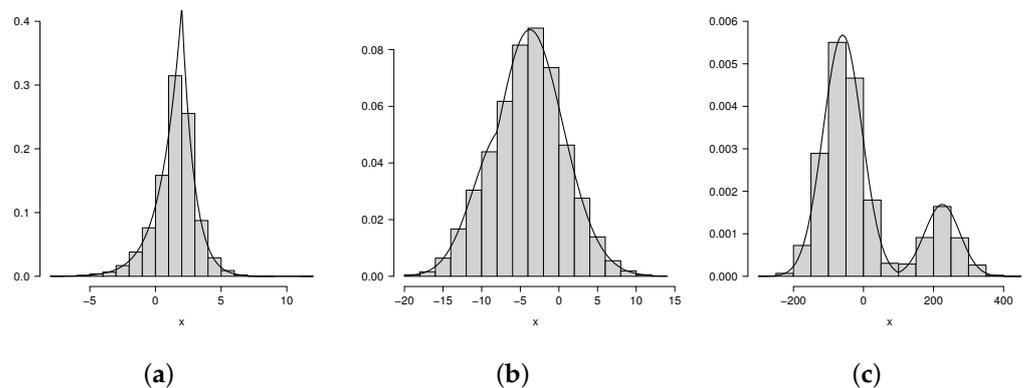


Figure 4. Ten thousand values simulated from the $FPN(\zeta, \eta, \alpha, \delta)$ model under different scenarios: (a) $\zeta = 2, \eta = 3, \alpha = 0.2, \delta = 2.5$ and acceptance-rejection method: way 1; (b) $\zeta = -8, \eta = 5, \alpha = 2, \delta = -0.5$ and acceptance-rejection method: way 2; (c) $\zeta = 100, \eta = 50, \alpha = 0.8, \delta = -2.5$ and Metropolis-Hastings method.

On the other hand, we evaluate the execution times of the three ways of drawing values (when the comparison is possible). For this, we draw 100 samples for the FPN model of size 1000; 2000 and 5000 and compute the average of these execution times. We consider $\zeta = 0, \eta = 1, \alpha$ ranging in $\{0.1, 0.2, \dots, 5.0\}$ and δ ranging in $\{-5.0, -4.9, \dots, 5.0\}$. The results for the three methods were very similar (with differences lower than 0.01 s). For instance, considering all the combinations for α and δ and $n = 5000$ (the case that took the longest), the maximum average of the execution times was 0.9134 s, suggesting that is relatively quick to simulate values from the model in all cases. As the Metropolis-Hastings algorithm can be applied for the FPN model in all situations, this method was used for data generation hereinafter.

5.2. Performance of ML Estimators in Finite Samples

In order to evaluate the properties of ML estimators in finite samples, the samples were generated based on the Metropolis-Hastings algorithm discussed in the previous section.

In all cases, we consider $\xi = 0$ and $\eta = 1$ as true values. The sample sizes considered were $n = 200$ and $n = 500$ with three values for α (0.5, 1.0 and 1.5) and δ (−0.5, 0.0 and 0.5). The three values for δ represent the bimodal asymmetric, bimodal symmetric and unimodal cases respectively, totaling 18 different cases. For each, we drew 1000 replicates and we computed the ML estimators using the *optim* function in R. We present the mean of the estimates and the mean of the estimated standard errors for the 1000 replicates. Results are presented in Table 1, showing that the mean of the estimates converges to the true value, and the mean of the estimated standard errors is reduced for all the parameters when n is increased, suggesting that the ML estimators for the FPN model are consistent.

Table 1. Mean of the estimates and mean of the estimated standard errors (in parentheses) based on 1000 replicates for the FPN model. In all cases $\xi = 0, \eta = 1$ are maintained and different sample sizes and values for α and δ are considered.

Parameter	$\delta = -0.5$		$\delta = 0$		$\delta = 0.5$		
	$n = 200$	$n = 500$	$n = 200$	$n = 500$	$n = 200$	$n = 500$	
$\alpha = 0.1$	ξ	0.4861 (4.4259)	−0.2412 (2.0642)	−0.1071 (1.5502)	−0.0133 (1.2664)	−0.0065 (1.0174)	−0.0029 (0.8505)
	η	0.9068 (0.5161)	0.9280 (0.4248)	0.9340 (0.3629)	0.9258 (0.3173)	0.9382 (0.2802)	0.9469 (0.2468)
	α	0.1195 (0.0990)	0.1109 (0.0774)	0.1136 (0.0719)	0.1101 (0.0637)	0.1093 (0.0640)	0.1081 (0.0601)
	δ	−0.8799 (1.3637)	−0.6915 (0.6524)	−0.4955 (0.4797)	−0.3877 (0.3809)	−0.1621 (0.3485)	−0.0117 (0.3129)
$\alpha = 0.8$	ξ	−0.022 (0.134)	0.000 (0.085)	0.054 (0.150)	0.011 (0.124)	0.000 (0.064)	−0.007 (0.031)
	η	0.971 (0.100)	0.989 (0.065)	0.923 (0.132)	0.953 (0.090)	0.977 (0.189)	1.009 (0.118)
	α	0.835 (0.096)	0.809 (0.060)	0.846 (0.150)	0.870 (0.123)	0.895 (0.155)	0.833 (0.079)
	δ	−0.587 (0.225)	−0.532 (0.145)	−0.163 (0.281)	−0.109 (0.181)	0.450 (0.439)	0.515 (0.277)
$\alpha = 1.0$	ξ	0.006 (0.126)	−0.007 (0.081)	0.018 (0.129)	0.017 (0.105)	−0.003 (0.069)	−0.013 (0.038)
	η	0.965 (0.108)	0.984 (0.070)	0.928 (0.135)	0.947 (0.088)	0.991 (0.206)	1.008 (0.126)
	α	1.011 (0.112)	1.009 (0.072)	1.065 (0.178)	1.058 (0.138)	1.184 (0.352)	1.060 (0.104)
	δ	−0.578 (0.243)	−0.540 (0.156)	−0.167 (0.309)	−0.122 (0.197)	0.493 (0.516)	0.510 (0.298)
$\alpha = 1.5$	ξ	−0.003 (0.122)	0.008 (0.079)	0.050 (0.112)	−0.016 (0.096)	−0.016 (0.096)	−0.018 (0.03)
	η	0.970 (0.115)	0.983 (0.074)	0.921 (0.136)	0.964 (0.092)	0.963 (0.092)	1.011 (0.126)
	α	1.515 (0.190)	1.498 (0.115)	1.479 (0.241)	1.504 (0.207)	1.604 (0.207)	1.501 (0.154)
	δ	−0.584 (0.273)	−0.541 (0.173)	−0.175 (0.343)	−0.104 (0.228)	0.545 (0.527)	0.524 (0.330)

5.3. A Model Selection Study

In this Section, we consider a simulation study in order to assess two model selection criteria in different contexts based on the FPN model: the AIC (see Akaike [21]), namely

$$AIC = -2\ell(\hat{\theta}; \mathbf{X}) + 2k,$$

where $\hat{\theta}$ is the ML estimator of θ and k is the number of parameters for the fitted model; and the Bayesian criterion (see Schwarz [22]), defined as

$$BIC = -2\ell(\hat{\theta}; \mathbf{X}) + \log(n)k.$$

Models with lower AIC and BIC are preferred. The data were drawn from the FPN($\xi = 0, \eta = 1, \alpha, \delta$) model, with α varying in (0.5, 0.8, 1.0, 1.5, 3.0) and δ varying in (−1.5, −0.5, 0.0, 0.5, 2.0). Note that the FPN(0, 1, $\alpha = 1, \delta = 0$) corresponds to the N distribution and FPN(0, 1, $\alpha \neq 1, \delta = 0$) corresponds to the PN model. Again, we consider two sample sizes: 200 and 500. For each simulated sample, we fitted the N, PN and FPN models and selected a model based on AIC and BIC criteria. Table 2 summarizes the percentage of times where each model is chosen based on the two criteria. Note that for $\alpha = 1$ and $\delta = 0$ (where the data are drawn from the N model), both criteria effectively chose the N model more frequently than the other models. However, for $\alpha \neq 1$ and $\delta = 0$ (where the data are drawn from the PN model), both criteria continue to choose the N model more frequently (except for the case $\alpha = 3, \delta = 0$ and $n = 500$), suggesting that AIC and BIC criteria choose the N model over the other models in scenarios where the asymmetry and the sample size are greater. For all other cases, we note that the BIC criterion tends to

choose the N model—incorrectly—much more frequently than the AIC criterion, and the advantages of the FPN model are evidenced in cases where $|\delta|$ is increased.

Table 2. Percentage of times where AIC and BIC choose the N, PN and FPN models based on 1000 replicates for different scenarios for the FPN model and sample size.

			δ										
			−1.5		−0.5		0.0		0.5		2.0		
n	α	Fitted Model	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	
200	0.5	N	0.0	0.2	42.0	85.4	65.2	93.6	35.0	77.9	2.6	19.5	
		PN	2.2	6.6	8.9	3.7	25.0	6.1	16.0	9.4	0.9	1.3	
		FPN	97.8	93.2	49.1	10.9	9.8	0.3	49.0	12.7	96.5	79.2	
	0.8	N	0.0	0.0	41.4	90.2	76.0	97.5	52.0	91.8	5.3	35.1	
		PN	0.1	0.1	9.3	1.8	12.7	2.1	9.0	2.1	1.6	1.3	
		FPN	99.9	99.9	49.3	8.0	11.3	0.4	39.0	6.1	93.1	63.6	
	1.0	N	0.0	0.0	48.2	92.4	76.3	97.7	53.1	92.0	6.0	38.4	
		PN	0.0	0.0	10.9	3.1	12.9	1.9	10.2	2.9	2.3	2.6	
		FPN	100.0	100.0	40.9	4.5	10.8	0.4	36.7	5.1	91.7	59.0	
	1.5	N	0.0	1.4	63.1	96.8	71.6	94.8	43.2	82.3	5.7	35.5	
		PN	4.1	9.9	11.5	2.1	16.7	4.8	25.7	12.3	6.1	9.3	
		FPN	95.9	88.7	25.4	1.1	11.7	0.4	31.1	5.4	88.2	55.2	
	3.0	N	50.1	88.0	71.3	94.0	49.6	83.3	26.6	63.6	1.7	13.2	
		PN	19.3	11.1	20.5	6.0	37.7	15.6	48.2	33.8	25.7	48.7	
		FPN	30.6	0.9	8.2	0.0	12.7	1.1	25.2	2.6	72.6	38.1	
	500	0.5	N	0.0	0.0	11.8	69.2	47.9	87.2	10.7	52.8	0.1	0.4
			PN	0.1	0.4	4.2	1.8	39.0	12.1	9.8	11.2	0.0	0.0
			FPN	99.9	99.6	84.0	29.0	13.1	0.7	79.5	36.0	99.9	99.6
0.8		N	0.0	0.0	13.6	76.3	72.2	97.6	25.6	82.6	0.0	1.4	
		PN	0.0	0.0	3.9	1.1	16.1	2.3	3.5	1.8	0.0	0.0	
		FPN	100.0	100.0	82.5	22.6	11.7	0.1	70.9	15.6	100.0	98.6	
1.0		N	0.0	0.0	17.4	80.8	77.4	98.4	25.4	84.0	0.0	2.2	
		PN	0.0	0.0	5.7	1.7	13.7	1.6	4.6	1.9	0.0	0.0	
		FPN	100.0	100.0	76.9	17.5	8.9	0.0	70.0	14.1	100.0	97.8	
1.5		N	0.0	0.0	41.3	95.2	60.6	94.1	15.7	68.5	0.0	2.1	
		PN	0.0	0.3	10.8	2.3	29.0	5.6	26.0	19.3	0.9	2.0	
		FPN	100.0	99.7	47.9	2.5	10.4	0.3	58.3	12.2	99.1	95.9	
3.0		N	27.6	91.3	53.1	93.5	18.3	61.1	3.1	25.5	0.0	0.3	
		PN	4.4	1.0	27.5	6.4	66.6	38.5	57.0	70.9	8.7	25.8	
		FPN	68.0	7.7	19.4	0.1	15.1	0.4	39.9	3.6	91.3	73.9	

6. Numerical Illustrations

In this Section, we illustrate with two real datasets the better performance of the FPN distribution compared with other models in the literature. Codes were developed in Core Team R [23] and are available as supplementary material for this manuscript.

6.1. Illustration 1

The first dataset corresponds to 481 observations of the variable “Crack” included in the dataset verb “Pollen5.Dat” available at <http://lib.stat.cmu.edu/datasets/pollen.data>, accessed on 15 October 2021. A descriptive summary of the data can be found in Table 3. Quantities $\sqrt{b_1}$ and b_2 denote, respectively, the sample asymmetry and kurtosis coefficients.

Table 3. Summary statistics for precipitation data.

n	Mean	Variance	$\sqrt{b_1}$	b_2
481	−0.0483	26.9980	0.2329	2.5944

Clearly, the values taken by the asymmetry and kurtosis coefficients indicate that an asymmetric model may be more adequate than a symmetric model for fitting the pollen data. Moreover, the data histogram presented in Figure 5a shows that a model with the ability to fit bimodal data may present better results than just the PN model. To quantify the findings we test the hypothesis

$$H_0 : \delta = 0 \text{ against } H_1 : \delta \neq 0,$$

by using the likelihood ratio statistic

$$\Lambda = \frac{\ell_{PN}(\hat{\xi})}{\ell_{FPN}(\hat{\xi})},$$

leading to

$$-2 \log(\Lambda) = 11.80,$$

which is greater than the 5% critical value for the chi-square distribution with one degree of freedom, given by $\chi_1^2 = 3.8414$. Hence, there is a clear indication that the FPN model can be quite useful in fitting bimodal data such as the pollen data described above. Table 4 presents parameter estimates (standard errors in parentheses) for both the FPN and ordinary models, including the mixture of normals (MN). Finally, plots for the fitted PN, MN and FPN models are presented in Figure 5b. Likewise, we present the qq-plot that follows by using parameter estimates for the FPN model, suggesting a good fit of the model for this dataset.

To compare model fit, we consider the AIC and BIC criteria. Hence, we have that the FPN model presents the best fit for the pollen dataset, of all the models considered.

Table 4. Parameter estimates (standard errors in parentheses) for N, SN, PN, MN and FPN models in the pollen dataset.

Parameter	N	SN	PN	FPN	Parameter	MN
ξ	-0.0489 (0.2367)	-4.9482 (0.8164)	-11.8216 (7.2910)	-2.0300 (0.2797)	ξ_1	-3.5847 (0.8948)
η	5.1902 (0.1673)	7.1379 (0.6058)	8.4627 (1.8070)	3.5192 (0.1982)	η_1	3.3009 (0.3662)
α	-	1.6568 (0.4856)	7.5132 (7.9838)	0.7316 (0.0464)	ξ_2	3.8531 (1.5545)
δ	-	-	-	-0.6946 (0.1316)	η_2	3.9516 (0.6768)
-	-	-	-	-	p	0.5245 (0.1593)
Log-likelihood	-1474.64	-1472.08	-1472.24	-1466.33		-1466.30
AIC	2953.28	2950.16	2950.47	2940.66		2942.60
BIC	2961.63	2962.68	2963.00	2957.36		2963.48

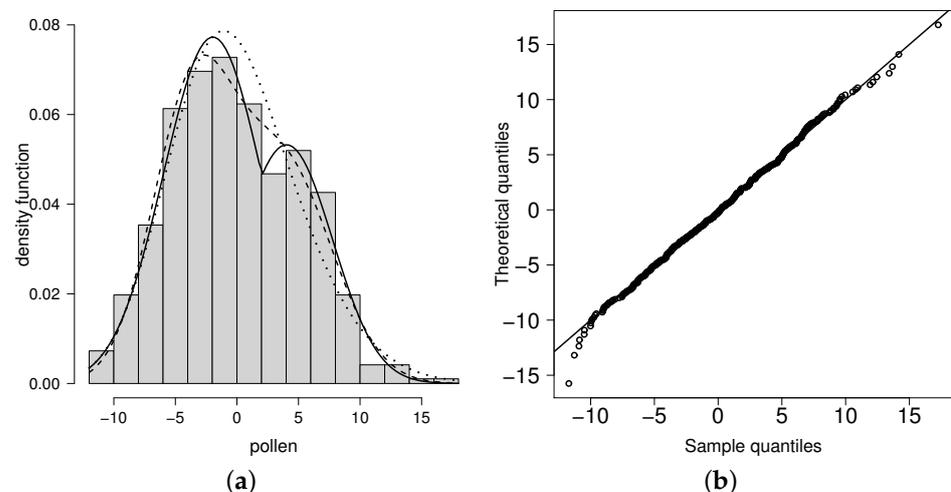


Figure 5. (a) Fitted distributions: SN (dotted line), MN (dashed line) and FPN (solid line) models. (b) Simulated QQ-plot for the fitted FPN model and the variable pollen.

6.2. Illustration 2

The dataset consists of 1150 heights measured at 1 micron intervals along the drum of a roller (i.e., parallel to the axis of the roller). This was part of an extensive study of surface roughness of the rollers. The units of height are not given, because the data are automatically rescaled as they are recorded, and the scaling factor is imperfectly known. The zero reference height is arbitrary. Data are available for downloading at <http://lib.stat.cmu.edu/jasadata/laslett>, accessed on 15 October 2021. More details about the data can be found in Laslett [24]. Table 5 shows the summary statistics for the data. Note that the data have a large, negative sample asymmetry and sample kurtosis greater than three, indicating that the ordinary normal model may present a poor fit. Table 6 presents the fit for the N, SN, PN and FPN models. Note that the minimum AIC and BIC criteria are attached by the FPN distribution. Finally, Figure 6 presents the histogram for the data and the fitted pdf for the models, showing the better performance of the FPN distribution over the other models considered.

Table 5. Summary statistics for roller data.

n	Mean	Variance	$\sqrt{b_1}$	b_2
1150	3.535	0.422	−0.986	4.855

Table 6. Parameter estimates (standard errors in parentheses) for N, SN, PN and FPN models in roller dataset.

Parameter	N	SN	PN	FPN
ζ	3.5347 (0.0192)	4.2475 (0.0284)	4.5494 (0.0570)	3.8529 (0.0106)
η	0.6497 (0.0135)	0.9644 (0.0304)	0.1983 (0.0279)	0.7589 (0.0634)
α	-	−2.7578 (0.2529)	0.0479 (0.0155)	0.2082 (0.0568)
δ	-	-	-	1.3045 (0.2064)
Log-likelihood	−1135.87	−1071.35	−1085.24	−1065.92
AIC	2275.73	2148.69	2176.84	2139.84
BIC	2285.83	2163.84	2191.98	2160.03

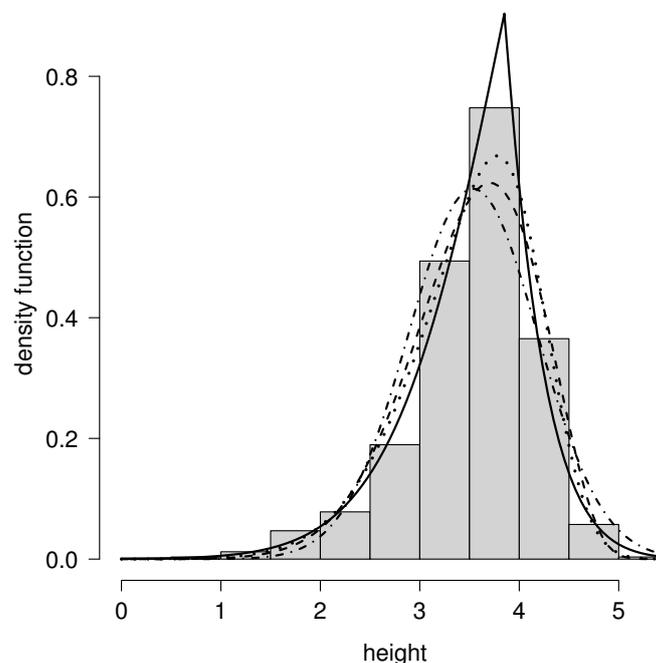


Figure 6. Histogram of heights dataset and fitted models N (dotted-dashed line), SN (dotted line), PN (dashed line) and FPN (solid line).

7. Conclusions

The main object of this paper is to propose a new asymmetric model that is more flexible than the generalized Gaussian model studied in Durrans [2]. This new model includes a parameter that makes it more flexible, because it is useful for unimodal or bimodal data and/or cases with skewness and kurtosis larger than those expected for the N and PN models. Given the flexibility of the model, further work could study the FPN distribution in different contexts: regression, measurement with error-in-variable models, random effect models, and so forth.

Author Contributions: Conceptualization, G.M.-F. and H.W.G.; methodology, H.W.G.; software, D.I.G. and G.M.-F.; validation, G.M.-F., H.W.G. and H.B.; formal analysis, G.M.-F., O.V. and H.W.G.; investigation, H.W.G., O.V. and G.M.-F.; writing—original draft preparation, G.M.-F. and D.I.G.; writing—review and editing, O.V. All authors have read and agreed to the published version of the manuscript.

Funding: Héctor W. Gómez's research has been partially supported by SEMILLERO UA-2021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available for downloading at <http://lib.stat.cmu.edu/datasets/pollen.data> (accessed on 15 October 2021) and <http://lib.stat.cmu.edu/jasadata/laslett> (accessed on 15 October 2021).

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

References

- Azzalini, A. A class of distributions which includes the normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
- Durrans, S.R. Distributions of fractional order statistics in hydrology. *Water Resour. Res.* **1992**, *28*, 1649–1655. [[CrossRef](#)]
- McLachlan, G.; Peel, D. *Finite Mixture Models*; John Wiley & Sons, Inc.: New York, NY, USA, 2000.
- Marin, J.; Mengersen, K.; Robert, C.P. Bayesian Modeling and Inference for mixtures of distributions. In *Handbook of Statistics*; Dey, D., Rao, C., Eds.; Elsevier: Amsterdam, The Netherlands, 2005; pp. 454–508.
- Azzalini, A.; Capitanio, A. Distributions generate by perturbation of symmetry with emphasis on a multivariate skew-t distribution. *J. R. Stat. Soc. Ser. B* **2003**, *65*, 367–389. [[CrossRef](#)]
- Ma, Y.; Genton, M.G. Flexible class of skew-symmetric distributions. *Scand. J. Stat.* **2004**, *31*, 459–468. [[CrossRef](#)]
- Arellano-Valle, R.B.; Gómez, H.W.; Quintana, F.A. Statistical inference for a general class of asymmetric distributions. *J. Stat. Plan. Inference* **2005**, *128*, 427–443. [[CrossRef](#)]
- Kim, H.J. On a class of two-piece skew-normal distributions. *Statistics* **2005**, *39*, 537–553. [[CrossRef](#)]
- Lin, T.I.; Lee, J.C.; Hsieh, W.J. Robust mixture models using the skew-t distribution. *Stat. Comput.* **2007**, *17*, 81–92. [[CrossRef](#)]
- Lin, T.I.; Lee, J.C.; Yen, S.Y. Finite mixture modeling using the skew-normal distribution. *Stat. Sin.* **2007**, *17*, 909–927.
- Elal-Olivero, D.; Gómez, H.W.; Quintana, F.A. Bayesian Modeling using a class of Bimodal skew-Elliptical distributions. *J. Stat. Plan. Inference* **2009**, *139*, 1484–1492. [[CrossRef](#)]
- Arnold, B.C.; Gómez, H.W.; Salinas, H.S. On multiple constraint skewed models. *Statistics* **2009**, *43*, 279–293. [[CrossRef](#)]
- Arnold, B.C.; Gómez, H.W.; Salinas, H.S. A doubly skewed normal distribution. *Statistics* **2015**, *49*, 842–858. [[CrossRef](#)]
- 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](#)]
- Braga, A.D.S.; Cordeiro, G.M.; Ortega, E.M.M. A new skew-bimodal distribution with applications. *Commun. Stat.-Theory Methods* **2018**, *47*, 2950–2968. [[CrossRef](#)]
- Venegas, O.; Salinas, H.S.; Gallardo, D.I.; Bolfarine, B.; Gómez, H.W. Bimodality based on the generalized skew-normal distribution. *J. Stat. Comput. Simul.* **2018**, *88*, 156–181. [[CrossRef](#)]
- Gómez-Déniz, E.; Pérez-Rodríguez, J.V.; Reyes, J.; Gómez, H.W. A Bimodal Discrete Shifted Poisson Distribution. A Case Study of Tourists' Length of Stay. *Symmetry* **2020**, *12*, 442. [[CrossRef](#)]
- Lehmann, E.L. The power of rank tests. *Ann. Math. Stat.* **1953**, *24*, 23–43. [[CrossRef](#)]
- Gupta, D.; Gupta, R.C. Analyzing skewed data by power normal model. *Test* **2008**, *17*, 197–210. [[CrossRef](#)]
- Pewsey, A.; Gómez, H.W.; Bolfarine, H. Likelihood-based inference for power distributions. *Test* **2012**, *21*, 775–789. [[CrossRef](#)]
- Akaike, H. A new look at statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [[CrossRef](#)]
- Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *1*, 461–464. [[CrossRef](#)]

-
23. Core Team R. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.
 24. Laslett, G.M. Kriging and Splines: An Empirical Comparison of Their Predictive Performance in Some Applications: Rejoinder. *J. Am. Stat. Assoc.* **1994**, *89*, 406–409. [[CrossRef](#)]