Modified Power-Symmetric Distribution

Emilio Gómez-Déniz [1], Yuri A. Iriarte [2], Enrique Calderín-Ojeda [3] and Héctor W. Gómez [2],*

1. Department of Quantitative Methods in Economics and TIDES Institute, University of Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Spain; emilio.gomez-deniz@ulpgc.es
2. Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile; yuri.iriarte@uantof.cl
3. Centre for Actuarial Studies, Department of Economics, The University of Melbourne, Melbourne, VIC 3010, Australia; enrique.calderin@unimelb.edu.au
* Correspondence: hector.gomez@uantof.cl

Abstract: In this paper, a general class of modified power-symmetric distributions is introduced. By choosing as symmetric model the normal distribution, the modified power-normal distribution is obtained. For the latter model, some of its more relevant statistical properties are examined. Parameters estimation is carried out by using the method of moments and maximum likelihood estimation. A simulation analysis is accomplished to study the performance of the maximum likelihood estimators. Finally, we compare the efficiency of the modified power-normal distribution with other existing distributions in the literature by using a real dataset.

1. Introduction

Over the last few years, the search for flexible probabilistic families capable of modeling different levels of bias and kurtosis has been an issue of great interest in the field of distributions theory. This was mainly motivated by the seminal work of Azzalini [1]. In that paper, the probability density function (pdf) of a skew-symmetric distribution was introduced. The expression of this density is given by

$$g(z; \lambda) = 2f(z)G(\lambda z), \quad z, \lambda \in \mathbb{R}, \tag{1}$$

where $f(\cdot)$ is a symmetric pdf about zero; $G(\cdot)$ is an absolutely continuous distribution function, which is also symmetric about zero; and λ is a parameter of asymmetry. For the case where $f(\cdot)$ is the standard normal density (from now on, we reserve the symbol ϕ for this function), and $G(\cdot)$ is the standard normal cumulative distribution function (henceforth, denoted by Φ), the so-called skew-normal ($\mathcal{SN}$) distribution with density

$$\phi_Z(z; \lambda) = 2\phi(z)\Phi(\lambda z), \quad z, \lambda \in \mathbb{R}, \tag{2}$$

is obtained. We use the notation $Z \sim \mathcal{SN}(\lambda)$ to denote the random variable Z with pdf given by Equation (2). A generalization of the $\mathcal{SN}$ distribution is introduced by Arellano-Valle et al. [2] and Arellano-Valle et al. [3]; they study Fisher's information matrix of this generalization. For further details about the $\mathcal{SN}$ distribution, the reader is referred to Azzalini [4]. Martínez-Flórez et al. [5] used generalizations of the $\mathcal{SN}$ distribution to extend the Birnbaum-Saunders model, and Contreras-Reyes and Arellano-Valle [6] utilized the Kullback–Leibler divergence measure to compare the multivariate normal distribution with the skew-multivariate normal.

One of the main limitations of working with the family given by Equation (1) is that the information matrix could be singular for some of its particular models (see Azzalini [1]). This might

lead to some difficulties in the estimation, due to the asymptotic convergence of the maximum likelihood (ML) estimators. To overcome this issue, some authors (see Chiogna [7] or Arellano-Valle and Azzalini [8]) have used a reparametrization of the $\mathcal{SN}$ model to obtain a nonsingular information matrix. However, this methodology cannot be extended to all type of skew-symmetric models which suffers of this convergence problem. On the other hand, the family of power-symmetric ($\mathcal{PS}$) distributions does not have this problem of singularity in the information matrix (see, Pewsey et al. [9]). The pdf of this family of distribution is given by

$$\varphi_F(z;\alpha) = \alpha f(z)\{F(z)\}^{\alpha-1}, \quad z \in \mathbb{R}, \ \alpha \in \mathbb{R}^+. \tag{3}$$

where $F(\cdot)$ is itself a cumulative distribution function (cdf) and α is the shape parameter. For the particular case that $F(\cdot) = \Phi(\cdot)$, the power-normal ($\mathcal{PN}$) distribution is obtained, with density given by

$$f(z;\alpha) = \alpha\phi(z)\{\Phi(z)\}^{\alpha-1}, \quad z \in \mathbb{R}, \ \alpha \in \mathbb{R}^+. \tag{4}$$

For some references where this family is discussed, the reader is referred to Lehmann [10], Durrans [11], Gupta and Gupta [12], and Pewsey et al. [9], among other papers. Other extensions of this model are given in Martínez-Flórez et al. [13], where a multivariate version from the model is introduced; also, Martínez-Flórez et al. [14] carried out applications by using regression models; finally, Martínez-Flórez et al. [15] examined the exponential transformation of the model , and Martínez-Flórez et al. [16] examined a version of the model doubly censored with inflation in a regression context. Truncations of the $\mathcal{PN}$ distribution were considered by Castillo el al. [17].

In this paper, a modification in the pdf of the $\mathcal{PS}$ probabilistic family is implemented to increase the degree of kurtosis. This methodology is later used to explain datasets that include atypical observations. Usually, this methodology is accomplished by increasing the number of parameters in the model.

The paper is organized as follows. In Section 2, first, we introduce the modified power symmetric distribution. Then, the particular case of the modified power normal distribution is derived. Some of the most relevant statistical properties of this model, including moments and kurtosis coefficient, are presented. Next, in Section 3, some methods of estimation are discussed. Later, a simulation study is provided to illustrate the behavior of the shape parameter. A numerical application where the modified power normal distribution is compared to the $\mathcal{SN}$ and $\mathcal{PN}$ distributions is given in Section 4. Finally, Section 5 concludes the paper.

2. Genesis and Properties of Modified Power-Normal Distribution

In this section, we introduce a new family of probability distributions. The idea is to make a transformation to a given probability density, as the skew-symmetric or power-symmetric distributions does. As there exists a certain resemblance between our formula (Equation (6)) and the formula for the power-symmetric distributions (Equation (3)), we agree to name these new distributions as modified power-symmetric ($\mathcal{MPS}$) distributions. From the standard normal distribution, we obtain the so-called Modified Power-Normal ($\mathcal{MPN}$) distribution. The main parameters and properties of this particular distribution will be studied throughout this work.

2.1. Probability Density Function

Definition 1. *Let Z be a continuous and symmetric random variable with cdf $G(z;\boldsymbol{\eta})$ and pdf $g(z;\boldsymbol{\eta})$, where $\boldsymbol{\eta}$ denotes a vector of parameters. We say that, a random variable, X, follows a $\mathcal{MPS}$ distribution, denoted as $X \sim \mathcal{MPS}(\boldsymbol{\eta}, \alpha)$, if its cdf is given by*

$$F(x;\boldsymbol{\eta},\alpha) = \frac{[1 + G(x;\boldsymbol{\eta})]^\alpha - 1}{2^\alpha - 1}, \tag{5}$$

and its pdf is given by

$$f(x; \boldsymbol{\eta}, \alpha) = \frac{\alpha}{2^\alpha - 1} g(x; \boldsymbol{\eta}) \left[1 + G(x; \boldsymbol{\eta}) \right]^{\alpha - 1}. \tag{6}$$

where $x \in \mathbb{R}$ and $\alpha > 0$.

Remark 1. *In the case $\alpha = 1$, the transformation given by Equation (6) is the identity. That is, the $\mathcal{MPS}$ distribution for $\alpha = 1$ always provides the input probability density function.*

Thereforeforth, we proceed to examine the $\mathcal{MPN}$ distribution, whose cdf is provided by

$$F(x; \mu, \sigma, \alpha) = \frac{\left[1 + \Phi\left(\frac{x-\mu}{\sigma}\right) \right]^\alpha - 1}{2^\alpha - 1}, \tag{7}$$

and whose pdf is given by

$$f(x; \mu, \sigma, \alpha) = \frac{\alpha}{(2^\alpha - 1)\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \left[1 + \Phi\left(\frac{x-\mu}{\sigma}\right) \right]^{\alpha - 1}, \tag{8}$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is the location parameter, $\sigma > 0$ is the scale parameter, and $\alpha > 0$ is the shape parameter. Hereafter, this will be denoted as $X \sim \mathcal{MPN}(\mu, \sigma, \alpha)$. Figure 1 depicts some different shapes of the pdf of this model, for selected values of the parameter α with $\mu = -1, 1$ and $\sigma = 1$. The $\mathcal{MPN}$ class of distributions is applicable for the change point problem, due to its favorable properties (see Maciak et al. [18]); moreover, the $\mathcal{MPN}$ model can be utilized in calibration (see Pešta [19]).

Remark 2. *Here, $\mu \in \mathbb{R}$ and $\sigma > 0$ are location and scale parameters of the $\mathcal{MPN}$ distribution, respectively. For the particular case $\alpha = 1$, these are not only location and scale parameters but also the mean and standard deviation of the standard normal distribution.*
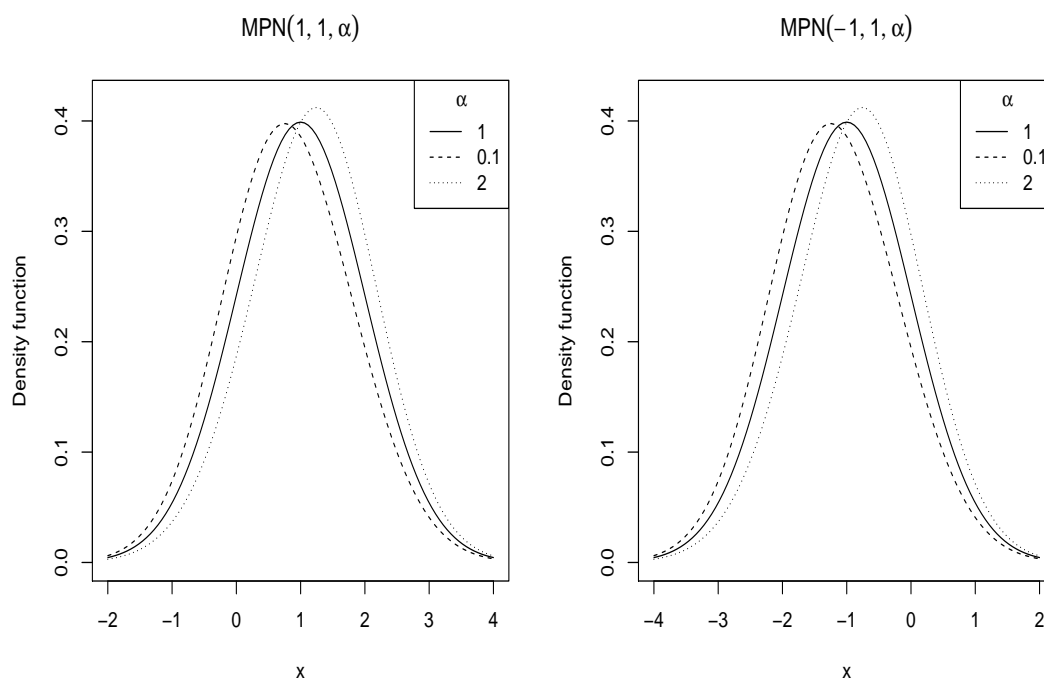
Figure 1. Plot of the pdf of $\mathcal{MPN}$ distribution for selected values of the parameters.

2.2. Statistical Properties

2.2.1. Shape of the Density

The $\mathcal{MPN}$ distribution exhibits a bell-shaped form, which can be symmetric or positively or negatively skewed depending on the value of the parameter α. Now, we derive some analytical expressions that are useful to obtain approximations of modal values and inflection points of this model. In the following, it will be assumed that $\mu = 0$ and $\sigma = 1$.

Proposition 1. *The pdf of $X \sim \mathcal{MPN}(0, 1, \alpha)$ has a local maximum at $(x_1, f(x_1; \alpha))$ and two inflection points at $(x_2, f(x_2; \alpha))$ and $(x_3, f(x_3; \alpha))$, respectively, where x_1 is the root of the equation*

$$x^* = \frac{(\alpha - 1)\phi(x^*)}{1 + \Phi(x^*)}, \tag{9}$$

and x_2 and x_3 are two solutions of the equation

$$1 = \left(-x + \frac{(\alpha - 1)\phi(x)}{1 + \Phi(x)} \right)^2 - \frac{(\alpha - 1)\phi(x)}{1 + \Phi(x)} \left(x + \frac{\phi(x)}{1 + \Phi(x)} \right). \tag{10}$$

Proof. The proof consists of simple derivatives of the function f. From the equation (8), we calculate

$$\frac{\partial}{\partial x} f(x; \alpha) = \frac{\alpha}{2^\alpha - 1} \phi(x)[1 + \Phi(x)]^{\alpha - 1} \left(-x + \frac{(\alpha - 1)\phi(x)}{1 + \Phi(x)} \right).$$

$$\frac{\partial^2}{\partial x^2} f(x; \alpha) = \frac{\alpha}{2^\alpha - 1} \phi(x)[1 + \Phi(x)]^{\alpha - 1} \left\{ \left(-x + \frac{(\alpha - 1)\phi(x)}{1 + \Phi(x)} \right)^2 - \left[1 + \frac{(\alpha - 1)\phi(x)}{1 + \Phi(x)} \right. \right.$$

$$\left. \left. \times \left(x + \frac{\phi(x)}{1 + \Phi(x)} \right) \right] \right\}.$$

By setting Equations (9) and (10) to be equal to zero, the results are obtained after some algebra. Figure 2 displays the graph of the first derivative of $f(\cdot)$, where it is observed that the maximum exists and it is unique. Therefore, the $\mathcal{MPN}$ distribution is unimodal. □
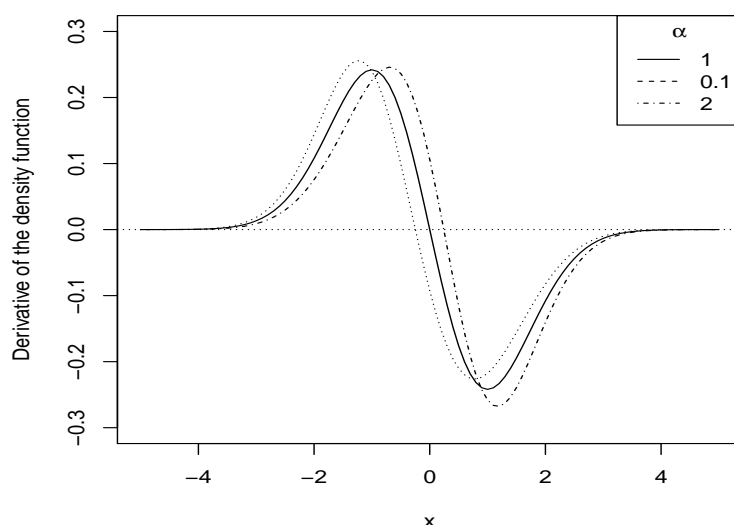
Figure 2. Plot of the first derivative of $\mathcal{MPN}$ distribution for selected values of the parameters.

Remark 3. *The solutions of Equations (9) and (10) can be numerically obtained by using the built-in function "uniroot" in the software package* R. *Table 1 below illustrates some approximations of the roots x_1, x_2, and x_3, and the corresponding figures of the pdf evaluated at these values.*

Table 1. Approximations of the roots of Equations (9) and (10) for some values of α, and the corresponding figures of the pdf of the $\mathcal{MPN}$ evaluated at these roots.

α	x_1	x_2	x_3	$f(x_1;\alpha)$	$f(x_2;\alpha)$	$f(x_3;\alpha)$
0.5	-0.136	-1.135	0.886	0.397	0.239	0.241
1.0	0.000	-1.000	1.000	0.399	0.242	0.242
2.0	0.243	-0.691	1.173	0.412	0.261	0.251
3.0	0.435	-0.414	1.299	0.433	0.282	0.266
4.0	0.586	-0.203	1.396	0.457	0.298	0.284
5.0	0.706	-0.041	1.475	0.481	0.316	0.301

2.2.2. Moments

Proposition 2. *The rth moments of $X \sim \mathcal{MPN}(0,1,\alpha)$ for $r = 1,2,3,\ldots$, are given by*

$$\mathbb{E}(X^r) = \frac{\alpha}{2^\alpha - 1}\, a_r(\alpha), \tag{11}$$

where $a_r(\alpha)$ is defined as

$$a_r(\alpha) = \int_0^1 [\Phi^{-1}(u)]^r (1+u)^{\alpha-1}\, du. \tag{12}$$

Here, $\Phi^{-1}(\cdot)$ is the quantile function of the standard normal distribution.

Proof. By using the change of variable $u = \Phi(x)$, it follows that

$$
\begin{aligned}
\mathbb{E}(X^r) &= \int_{-\infty}^{\infty} x^r \frac{\alpha}{2^\alpha - 1}\phi(x)[1+\Phi(x)]^{\alpha-1}\, dx \\
&= \frac{\alpha}{2^\alpha - 1}\int_0^1 [\Phi^{-1}(u)]^r (1+u)^{\alpha-1}\, du \\
&= \frac{\alpha}{2^\alpha - 1}a_r(\alpha). \ \square
\end{aligned}
$$

Corollary 1. *The mean and variance of X are given by*

$$
\begin{aligned}
\mathbb{E}(X) &= \frac{\alpha}{2^\alpha - 1}a_1(\alpha) \ \ \text{and} \\
\mathbb{V}ar(X) &= \frac{\alpha}{2^\alpha - 1}\left(a_2(\alpha) - \frac{\alpha}{2^\alpha - 1}a_1^2(\alpha)\right),
\end{aligned}
$$

respectively.

Corollary 2. *The skewness (β_1) and kurtosis (β_2) coefficients are, respectively, given by*

$$\beta_1 = \frac{a_3(\alpha) - \frac{3\alpha}{2^\alpha-1}a_1(\alpha)a_2(\alpha) + \frac{2\alpha^2}{(2^\alpha-1)^2}a_1^3(\alpha)}{\left(\frac{\alpha}{2^\alpha-1}\right)^{3/2}[a_2(\alpha) - \frac{\alpha}{2^\alpha-1}a_1^2(\alpha)]^{3/2}} \ \ \text{and}$$

$$\beta_2 = \frac{a_4(\alpha) - \frac{4\alpha}{2^\alpha-1}a_1(\alpha)a_3(\alpha) + \frac{6\alpha^2}{(2^\alpha-1)^2}a_1^2(\alpha)a_2(\alpha) - \frac{3\alpha^3}{(2^\alpha-1)^3}a_1^4(\alpha)}{\frac{\alpha}{2^\alpha-1}[a_2(\alpha) - \frac{\alpha}{2^\alpha-1}a_1^2(\alpha)]^2}.$$

Remark 4. *Observe that the integral in Equation (12) can be numerically approximated by using the built-in function* integrate *available in the software package* R. *Below, in Table 2, some approximations of the mean and*

variance for the $\mathcal{MPN}$ *distribution for different values of* α *are displayed. Figure 3 illustrates the behavior of the* $\mathbb{E}(X)$ *and* $\mathbb{V}ar(X)$ *of the* $\mathcal{MPN}$ *distribution for different values of* α*. It is observable that when* α *grows, the mean increases and the variance decreases.*

Figure 4 displays the curves associated with the coefficients of skewness (left panel) and kurtosis (right) of the $\mathcal{MPN}$ *and* $\mathcal{PN}$ *distributions. It is shown that, depending on the values of* α*, the* $\mathcal{MPN}$ *distribution exhibits equal, greater, or lesser values for these coefficients compared to the* $\mathcal{PN}$ *model. In general, the* $\mathcal{MPN}$ *distribution has a smaller range of skewness than the* $\mathcal{PN}$ *distribution. On the other hand, when* α < 13.05*, the* $\mathcal{MPN}$ *distribution has a greater kurtosis coefficient than the* $\mathcal{PN}$ *model.*

Table 2. Approximations of $\mathbb{E}(X)$ and $\mathbb{V}ar(X)$ of the $\mathcal{MPN}$ distribution for different values of α.

α	$\mathbb{E}(X)$	$\mathbb{V}ar(X)$
0.5	−0.097	1.006
1.0	0.000	1.000
5.0	0.659	0.770
10.0	1.119	0.521
100.0	2.247	0.218
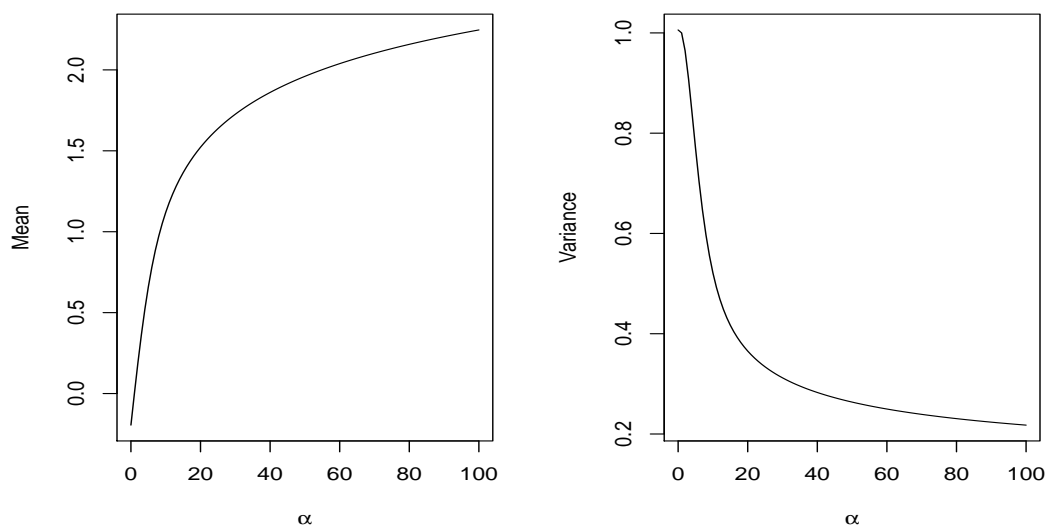
Figure 3. Plot of the $\mathbb{E}(X)$ and $\mathbb{V}ar(X)$ of the $\mathcal{MPN}$ distribution.
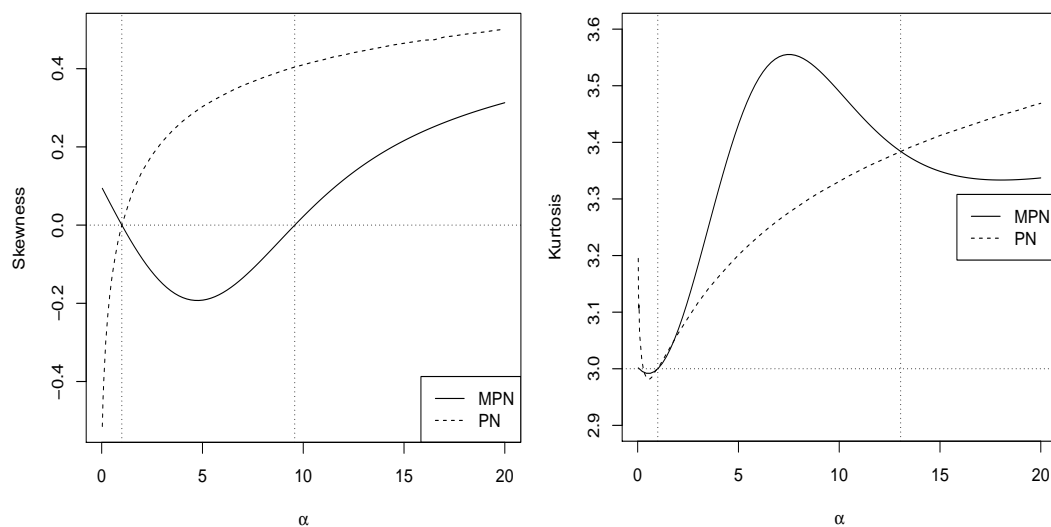
Figure 4. Graphs of the skewness and kurtosis coefficients for the $\mathcal{MPN}$ and $\mathcal{PN}$ distributions.

2.2.3. Stochastic Ordering

Stochastic ordering is an important tool to compare continuous random variables. It is well-known that random variable X_1 is smaller than random variable X_2 in stochastic ordering ($X_1 \leq_{st} X_2$) if $F_{X_1}(x) \geq F_{X_2}(x)$ for all x, and in likelihood ratio order ($X_1 \leq_{lr} X_2$) if $f_{X_1}(x)/f_{X_2}(x)$ decreases with x. Using Theorem 1.C.1 and Theorem 2.A.1 of Shaked and Shanthikumar [20], the above stochastic orders hold according to the following implications,

$$X_1 \leq_{lr} X_2 \Rightarrow X_1 \leq_{st} X_2. \tag{13}$$

The proposition shows that the members of the $\mathcal{MPN}$ family can be stochastically ordered according to parameters values.

Proposition 3. *Let $X_1 \sim \mathcal{MPN}(0,1,\alpha_1)$ and $X_2 \sim \mathcal{MPN}(0,1,\alpha_2)$. If $\alpha_1 > \alpha_2$, then $X_1 \leq_{lr} X_2$ and, therefore, $X_1 \leq_{st} X_2$.*

Proof. From the quotient of both densities, it follows that

$$\frac{f_{X_2}(x;\alpha_2)}{f_{X_1}(x;\alpha_1)} = \frac{\alpha_2}{\alpha_1}\left(\frac{2^{\alpha_1}-1}{2^{\alpha_2}-1}\right)[1+\Phi(x)]^{\alpha_2-\alpha_1},$$

is non-decreasing if and only if $\mu'(x) \geq 0$ for $x \in (-\infty,\infty)$, where

$$\mu(x) = [1+\Phi(x)]^{\alpha_2-\alpha_1}.$$

After some calculations, it is shown that

$$\mu'(x) = (\alpha_2-\alpha_1)\phi(x)[1+\Phi(x)]^{\alpha_2-\alpha_1-1}.$$

It is straightforward that for $\alpha_1 > \alpha_2$, then $\mu'(x) < 0$ for $x \in (-\infty,\infty)$. Therefore, $f_{X_2}(x;\alpha_2)/f_{X_1}(x;\alpha_1)$ is decreasing in x, and consequently $X_1 \leq_{lr} X_2$. The other implication follows immediately from (13). $\square$

3. Inference

In this section, parameters estimation for the $\mathcal{MPN}$ distribution is discussed by using the method of moments and ML estimation. Additionally, a simulation analysis is carried out to illustrate the behavior of the ML estimators.

3.1. Method of Moments

The following proposition illustrates the derivation of the moment estimates of the $\mathcal{MPN}$ distribution.

Proposition 4. *Let $x_1,\ldots,x_n$ be a random sample obtained from the random variable $X \sim \mathcal{MPN}(\mu,\sigma,\alpha)$, then the moment estimates $\widehat{\boldsymbol{\theta}}_M = (\widehat{\mu}_M, \widehat{\sigma}_M, \widehat{\alpha}_M)$ for $\boldsymbol{\theta} = (\mu,\sigma,\alpha)$ are given by*

$$\widehat{\sigma}_M = \frac{S_{\underline{x}}}{\sqrt{\frac{\widehat{\alpha}_M}{2^{\widehat{\alpha}_M}-1}\left(a_2(\widehat{\alpha}_M) - \frac{\widehat{\alpha}_M}{2^{\widehat{\alpha}_M}-1}a_1^2(\widehat{\alpha}_M)\right)}}, \quad \widehat{\mu}_M = \bar{x} - \widehat{\sigma}_M\frac{\widehat{\alpha}_M}{2^{\widehat{\alpha}_M}-1}a_1(\widehat{\alpha}_M) \tag{14}$$

and

$$\frac{a_3(\widehat{\alpha}_M) - \frac{3\widehat{\alpha}_M}{2^{\widehat{\alpha}_M}-1}a_1(\widehat{\alpha}_M)a_2(\widehat{\alpha}_M) + \frac{2\widehat{\alpha}_M^2}{(2^{\widehat{\alpha}_M}-1)^2}a_1^3(\widehat{\alpha}_M)}{\left(\frac{\widehat{\alpha}_M}{2^{\widehat{\alpha}_M}-1}\right)^{3/2}[a_2(\widehat{\alpha}_M) - \frac{\widehat{\alpha}_M}{2^{\widehat{\alpha}_M}-1}a_1^2(\widehat{\alpha}_M)]^{3/2}} - A_{\underline{x}} = 0, \tag{15}$$

where $\overline{x}$, $S_{\underline{x}}$ and $A_{\underline{x}}$ denote the sample mean, sample standard deviation and sample Fisher's skewness coefficient respectively.

Proof. As μ and σ are location and scale parameters respectively, the skewness coefficient does not depend on these parameters. Thus, the result in (15) is directly obtained from matching the sample skewness coefficient with population counterpart given in Corollary 2. In addition, by considering that $X = \sigma Z + \mu$, where $Z \sim \mathcal{MPN}(0, 1, \alpha)$, and again by equating sample mean and sample variance to the mean and variance respectively, it follows that

$$\overline{x} = \widehat{\sigma}_M \mathbb{E}(X) + \widehat{\mu}_M,$$

$$= \widehat{\sigma}_M \frac{\widehat{\alpha}_M}{2^{\widehat{\alpha}_M} - 1} a_1(\widehat{\alpha}_M) + \widehat{\mu}_M,$$

and

$$S_{\underline{x}}^2 = \widehat{\sigma}_M^2 \mathbb{V}ar(X)$$

$$= \widehat{\sigma}_M^2 \frac{\widehat{\alpha}_M}{2^{\widehat{\alpha}_M} - 1} \left(a_2(\widehat{\alpha}_M) - \frac{\widehat{\alpha}_M}{2^{\widehat{\alpha}_M} - 1} a_1^2(\widehat{\alpha}_M) \right),$$

where $\widehat{\alpha}_M$ satisfies expression (15). Then, (14) is obtained by solving the latter equations for $\widehat{\mu}_M$ and $\widehat{\sigma}_M$, respectively. $\square$

3.2. Maximum Likelihood Estimation

For a random sample $x_1, \ldots, x_n$ derived from the $\mathcal{MPN}(\mu, \sigma, \alpha)$ distribution, the log-likelihood function can be written as

$$\ell(\mu, \sigma, \alpha) = nc(\sigma, \alpha) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 + (\alpha - 1) \sum_{i=1}^{n} \log \left[1 + \Phi\left(\frac{x_i - \mu}{\sigma} \right) \right], \tag{16}$$

where $c(\sigma, \alpha) = \log(\alpha) - \log(2^\alpha - 1) - \log(\sigma) - \frac{1}{2}\log(2\pi)$.

The score equations are given by

$$n\mu + \sigma(\alpha - 1) \sum_{i=1}^{n} \kappa(x_i) = n\overline{x}, \tag{17}$$

$$n\sigma^2 + \sigma(\alpha - 1) \sum_{i=1}^{n} (x_i - \mu)\kappa(x_i) = \sum_{i=1}^{n} (x_i - \mu)^2, \tag{18}$$

$$\frac{n}{\alpha} + \sum_{i=1}^{n} \log \left[1 + \Phi\left(\frac{x_i - \mu}{\sigma} \right) \right] = \frac{2^\alpha \log(2)n}{2^\alpha - 1}, \tag{19}$$

where $\kappa(w) = \kappa(w; \mu, \sigma) = \frac{\phi\left(\frac{w - \mu}{\sigma} \right)}{1 + \Phi\left(\frac{w - \mu}{\sigma} \right)}$.

Solutions for these Equations (17)–(19) can be obtained by using numerical procedures such as Newton–Raphson algorithm. Alternatively, these estimates can be found by directly maximizing the log-likelihood surface given by (16) and using the subroutine "optim" in the software package [21].

3.3. Simulation Study

To examine the behavior of the proposed approach, a simulation study is carried out to assess the performance of the estimation procedure for the parameters μ, σ, and α in the $\mathcal{MPN}$ model. The simulation analysis is conducted by considering 1000 generated samples of sizes $n = 50, 100$, and 200 from the $\mathcal{MPN}$ distribution. The goal of this simulation is to study the behavior of the ML

estimators of the parameters by using our proposed procedure. To generate $X \sim \mathcal{MPN}(\mu, \sigma, \alpha)$, the following algorithm is used,

1.   Step 1: Generate $W \sim Uniform(0, 1)$.
2.   Step 2: Compute $X = \mu + \sigma \Phi^{-1} \left[(2^{\alpha} W - W + 1)^{1/\alpha} - 1 \right]$.

where $\mu \in \mathbb{R}$, $\sigma > 0$, $\alpha > 0$ and $\Phi^{-1}(\cdot)$ is the quantile function of the standard normal distribution. For each generated sample of the $\mathcal{MPN}$ distribution, the ML estimates and corresponding standard deviation (SD) were computed for each parameter. As it can be seen in Table 3, the performance of the estimates improves when n and α increases.

Table 3. Maximum likelihood (ML) estimates and standard deviation (SD) for the parameters μ, σ and α of the $\mathcal{MPN}$ model for different generated samples of sizes $n = 50$, 100, and 200.

μ	σ	α	$\widehat{\mu}$ (SD)	$\widehat{\sigma}$ (SD)	$\widehat{\alpha}$ (SD)
			$n = 50$		
0	1	0.1	−0.352478 (0.149214)	0.994441 (0.091321)	0.190243 (0.175202)
		0.5	−0.19534 (0.14501)	0.990622 (0.094550)	0.613052 (0.272096)
		0.8	−0.083183 (0.144587)	0.990286 (0.098669)	0.854338 (0.164924)
		1	−0.009586 (0.141691)	0.995312 (0.0997256)	1.007328 (0.122688)
		5	0.004225 (0.100001)	0.997408 (0.088254)	5.030272 (0.229064)
		10	0.001108 (0.066610)	0.999124 (0.068611)	10.060478 (0.475019)
		100	0.002171 (0.017362)	1.001152 (0.029604)	100.437990 (2.668190)
			$n = 100$		
0	1	0.1	−0.351446 (0.104552)	0.998513 (0.070831)	0.180054 (0.111930)
		0.5	−0.19268 (0.101786)	0.997622 (0.068806)	0.576957 (0.223378)
		0.8	−0.08140 (0.099360)	0.997674 (0.069451)	0.830318 (0.152995)
		1	0.002786 (0.097411)	0.996444 (0.069648)	1.002200 (0.088749)
		5	0.002014 (0.099305)	0.996788 (0.085987)	5.023032 (0.221756)
		10	0.002897 (0.046109)	1.000515 (0.050192)	10.032857 (0.339106)
		100	0.000623 (0.012137)	1.000185 (0.019759)	100.168752 (1.866302)
			$n = 200$		
0	1	0.1	−0.348177 (0.072732)	0.999433 (0.047548)	0.170978 (0.076165)
		0.5	−0.196617 (0.072015)	0.999142 (0.047896)	0.562935 (0.218890)
		0.8	−0.076657 (0.069510)	0.997719 (0.050718)	0.824700 (0.127661)
		1	0.001158 (0.06877)	0.998408 (0.050586)	1.003651 (0.058344)
		5	−0.000165 (0.053006)	1.000623 (0.044182)	5.005130 (0.115719)
		10	−0.000239 (0.033615)	1.000017 (0.035902)	10.014958 (0.246652)
		100	0.000514 (0.008452)	1.000491 (0.014599)	100.104380 (1.295144)

Fisher's Information Matrix

Let us now consider $X \sim \mathcal{MPN}(\mu, \sigma, \alpha)$ and $Z = \left(\frac{X - \mu}{\sigma} \right) \sim \mathcal{MPN}(0, 1, \alpha)$. For a single observation x of X, the log-likelihood function for $\boldsymbol{\theta} = (\mu, \sigma, \alpha)$ is given by

$$\ell(\boldsymbol{\theta}) = \log f_X(\boldsymbol{\theta}, x) = c(\sigma, \alpha) - \frac{1}{2\sigma^2} (x - \mu)^2 + (\alpha - 1) \log \left[1 + \Phi \left(\frac{x - \mu}{\sigma} \right) \right].$$

The corresponding first and second partial derivatives of the log-likelihood function are derived in the Appendix A. It can be shown that the Fisher's information matrix for the $\mathcal{MPN}$ distribution is provided by

$$I_F(\boldsymbol{\theta}) = \begin{pmatrix} I_{\mu\mu} & I_{\mu\sigma} & I_{\mu\alpha} \\ & I_{\sigma\sigma} & I_{\sigma\alpha} \\ & & I_{\alpha\alpha} \end{pmatrix}$$

with the following entries,

$$I_{\mu\mu} = \frac{1}{\sigma^2} + \left(\frac{\alpha-1}{\sigma^2}\right)(b_{11}+b_{02}),$$

$$I_{\mu\sigma} = \frac{2}{\sigma^2}\mathbb{E}(Z) - \left(\frac{\alpha-1}{\sigma^2}\right)(b_{01}-b_{21}-b_{12}),$$

$$I_{\mu\alpha} = \frac{1}{\sigma}b_{01},$$

$$I_{\sigma\sigma} = -\frac{1}{\sigma^2} + \frac{3}{\sigma^2}\mathbb{E}(Z^2) - \left(\frac{\alpha-1}{\sigma^2}\right)(2b_{11}-b_{31}-b_{22}),$$

$$I_{\sigma\alpha} = \frac{1}{\sigma}b_{11},$$

$$I_{\alpha\alpha} = \frac{1}{\alpha^2} - \frac{2^\alpha\log^2(2)}{(2^\alpha-1)^2},$$

where $b_{ij} = \mathbb{E}\left[Z^i\kappa^j(Z;0,1)\right]$ must be numerically computed.

The Fisher's (expected) information matrix can be obtained by computing the expected values of the above expressions. By taking in this matrix, $\alpha = 1$, we have that $Z \sim N(\mu,\sigma)$ and

$$I_F(\mu,\sigma,\alpha=1) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 & \frac{1}{\sigma}d_{02} \\ 0 & \frac{2}{\sigma^2} & \frac{1}{\sigma}d_{12} \\ \frac{1}{\sigma}d_{02} & \frac{1}{\sigma}d_{12} & 1-2\log^2(2) \end{pmatrix},$$

where $d_{ij} = \int_{-\infty}^{\infty}\frac{z^i\phi^j(z)}{1+\Phi(z)}dz$ must be numerically obtained.

The determinant of $I_F(\mu,\sigma,\alpha=1)$ is $(2-4\log^2(2)-b_{12}^2-2b_{02}^2)/\sigma^4 = 0.003357435/\sigma^4 \neq 0$, consequently, the Fisher's information matrix is nonsingular at $\alpha = 1$.

Therefore, for large samples, the ML estimators, $\widehat{\theta}$, of θ are asymptotically normal, that is,

$$\sqrt{n}\left(\widehat{\theta}-\theta\right) \xrightarrow{\mathcal{L}} N_3(\mathbf{0},\ I(\theta)^{-1}),$$

resulting in the asymptotic variance of the ML estimators $\widehat{\theta}$ being the inverse of Fisher's information matrix $I(\theta)$. As the parameters are unknown, the observed information matrix is usually considered, where the unknown parameters are estimated by ML.

4. Application

In this section, a numerical illustration based on a real dataset is presented. The goal of this application is to show empirical evidence that the $\mathcal{MPN}$ yields a better fit to data than the $\mathcal{PN}$, $\mathcal{SN}$, and t-student ($\mathcal{TS}$) with α degrees of freedom distributions. For that reason, we consider a set of 3848 observations of the variable "density" included in the dataset verb "POLLEN5.DA" available at http://lib.stat.cmu.edu/datasets/pollen.data. This variable measures a geometric characteristic of a specific type of pollen. This dataset was previously used by Pewsey et al. [9] to compare the $\mathcal{PN}$ and $\mathcal{SN}$ distributions. A summary of some descriptive statistics are displayed in Table 4 below.

Table 4. Summary of descriptive statistics for the pollen density dataset.

Mean	Median	Variance	Skewness	Kurtosis
0.000	−0.030	9.887	0.109	3.193

By using the results derived in Proposition 4, we have computed the moment estimates for the parameters (μ, σ, α) of the $\mathcal{MPN}$ distribution, obtaining $(-5.609, 4.576, 11.857)$. Then, by taking these numbers as initial values, the ML estimates are derived. In Table 5, the ML estimates for the parameters of the $\mathcal{MPN}$, $\mathcal{PN}$, $\mathcal{SN}$, and $\mathcal{TS}$ distributions. The figures between brackets are the asymptotic standard errors of the estimates obtained by inverting the Fisher's information matrices for the three models evaluated at their respective ML estimates. Additionally, for each model, the values of the maximum of the log-likelihood function (ℓ_{max}) are reported. The $\mathcal{MPN}$ distribution attains the largest value, and consequently provides a better fit to data.

Table 5. Parameter estimates; standard errors (SE); and maximum of the log-likelihood function, ℓ_{max}, for the $\mathcal{TS}$, $\mathcal{SN}$, $\mathcal{PN}$, and $\mathcal{MPN}$ corresponding to the pollen density dataset.

| Parameters	$\mathcal{TS}$ (SE)	$\mathcal{SN}$ (SE)	$\mathcal{PN}$ (SE)	$\mathcal{MPN}$ (SE)
μ	−0.010 (0.05)	−2.04 (0.24)	−1.74 (0.68)	−5.73 (0.43)
σ	3.037 (0.05)	3.75 (0.14)	3.69 (0.21)	4.62 (0.14)
α	29.995 (13.01)	0.93 (0.14)	1.77 (0.37)	12.13 (1.21)
ℓ_{max}	−9864.99	−9863.42	−9863.37	**−9861.98**

To compare the fit achieved by each distribution, the values of several measures of model selection, i.e., Akaike's information criterion (AIC) (see Akaike [22]) and Bayesian information criterion (BIC) (see Schwarz [23]) are reported in Table 6. A model with lower numbers in these measures of model selection is preferable. It can be seen that the $\mathcal{MPN}$ is preferable in terms of these two measures of model validation. In addition, the Kolmogorov–Smirnov test statistics and the corresponding p-values has been included in this table for all the models considered. It can be observed that none of the models is rejected at the usual significance levels. However, the $\mathcal{MPN}$ distribution has a higher p-value and is rejected later than the other two models. Alternative methods of model selection to the Kolmogorov–Smirnov test that can be applied here can be found in Jäntschi and Bolboacă [24] and Jäntschi [25]. Furthermore, the histogram associated to the empirical distribution of the variable "density" in the pollen dataset is illustrated in the left hand side of Figure 5. In addition, the densities of $\mathcal{TS}$, $\mathcal{SN}$, $\mathcal{PN}$, and $\mathcal{MPN}$, by using the maximum likelihood estimates of their parameters, have been superimposed. Similarly, on the right hand side of Figure 5, the fit in both tails is shown. It is observable that, for this dataset, the $\mathcal{MPN}$ has thicker tails than the other three distributions. Finally, the QQ-plots for each distribution considered have been illustrated in Figure 6. Here, note that the $\mathcal{MPN}$ distribution exhibits an almost perfect alignment with the 45° line, and therefore it provides a better fit for extreme quantiles. Finally, Figure 7 displays the profile log-likelihood of μ, σ, and α of the MPN distribution. It is noticeable that the estimates are unique.

Table 6. Akaike's information criterion (AIC), Bayesian information criterion (BIC), Kolmogorov–Smirnov (KSS) test, and the corresponding p-values for all the models considered.

Criteria	$\mathcal{TS}$	$\mathcal{SN}$	$\mathcal{PN}$	$\mathcal{MPN}$
AIC	19,735.98	19,732.84	19,732.74	**19,729.96**
BIC	19,754.74	19,751.61	19,751.50	**19,748.72**
KSS (p-value)	0.014 (0.516)	0.013 (0.559)	0.012 (0.627)	**0.010 (0.820)**
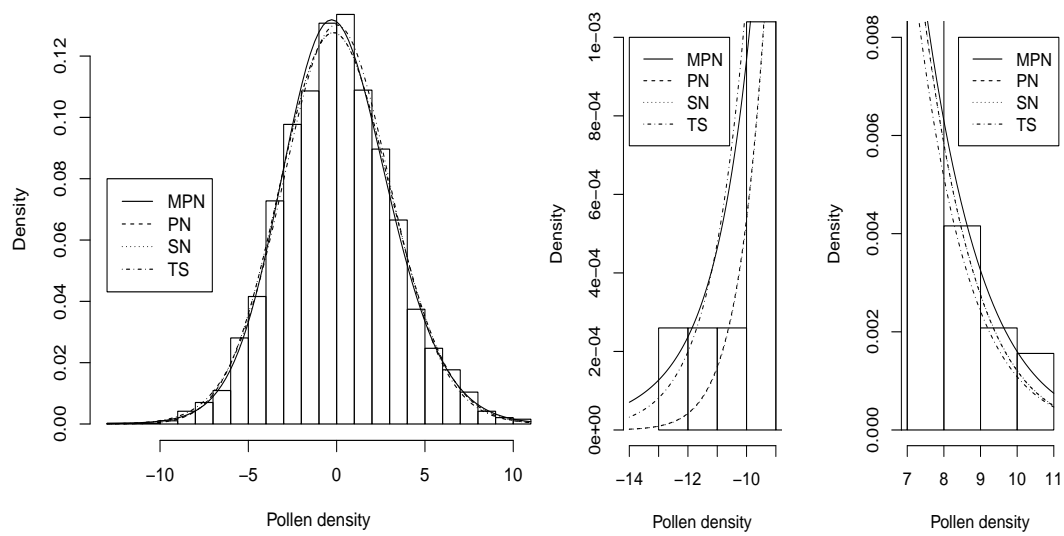
Figure 5. Left panel: Histogram of the empirical distribution and fitted densities by ML superimposed for pollen dataset. **Right** panel: Plots of the tails for the four models.
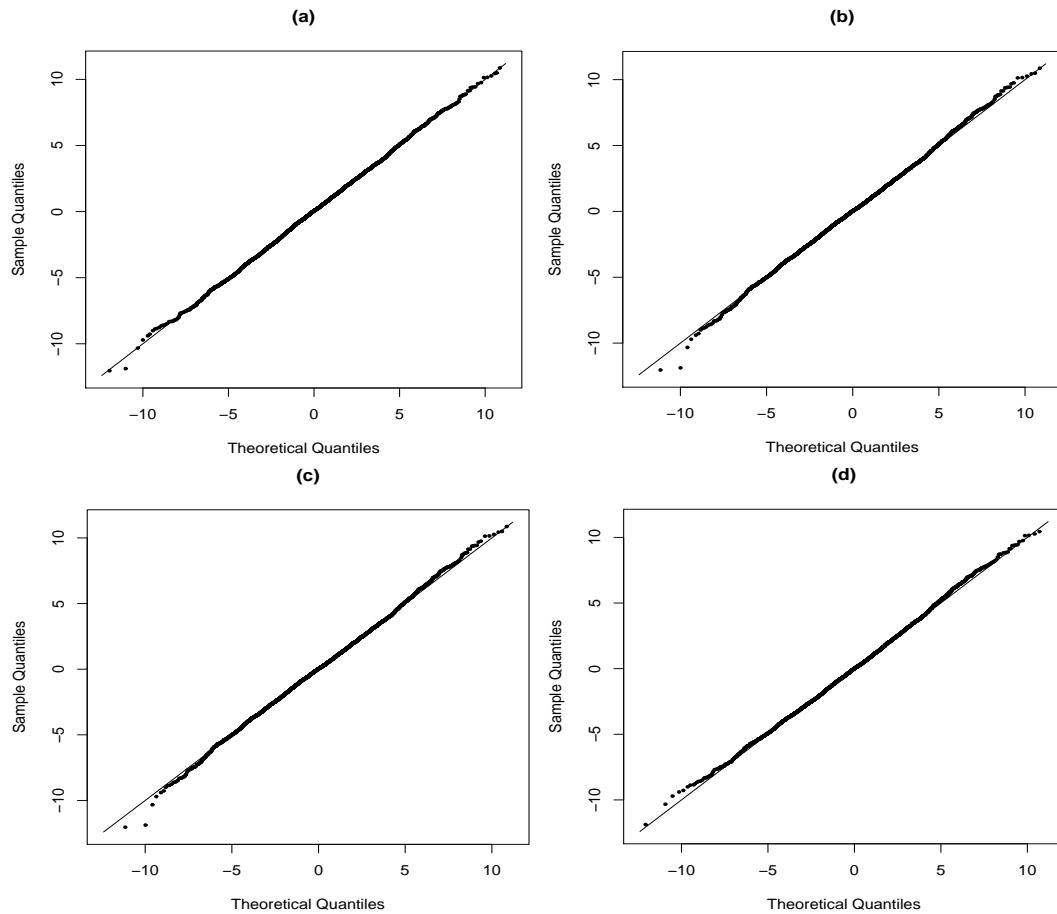
Figure 6. QQ-plots: (**a**) $\mathcal{MPN}$ model; (**b**) $\mathcal{PN}$ model; (**c**) $\mathcal{SN}$ model; (**d**) $\mathcal{TS}$ model.
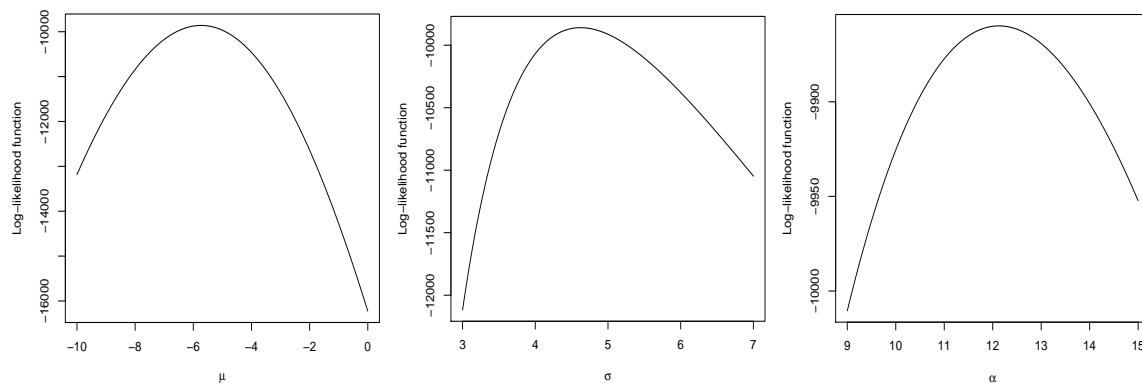
Figure 7. Profile log-likelihood of μ, σ and α for the $\mathcal{MPN}$ distribution.

5. Concluding Remarks

In this paper, a modification of the continuous symmetric-power distribution has been introduced. The particular case of the normal distribution the $\mathcal{MPN}$ distribution has been examined in detail. This distribution arises by modifying the distribution function of the symmetrical powers family. After carrying out this modification, a more flexible family of probability distributions is obtained, allowing for the kurtosis coefficient to take a certain range of values in the parameter space. For this model, its basic properties, different method of estimation and Fisher's information matrix were studied. By using a real dataset, we showed that the $\mathcal{MPN}$ distribution provides a better fit than other existing models in the literature such as the $\mathcal{TS}$, $\mathcal{SN}$, and $\mathcal{PN}$ distributions.

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

Appendix A

The first derivatives of $\ell(\boldsymbol{\theta})$ are given by

$$
\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu} &= \frac{1}{\sigma}\left[\frac{x-\mu}{\sigma}-(\alpha-1)\kappa(x)\right] \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \sigma} &= -\frac{1}{\sigma}\left[1-\left(\frac{x-\mu}{\sigma}\right)^2+(\alpha-1)\left(\frac{x-\mu}{\sigma}\right)\kappa(x)\right] \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \frac{1}{\alpha}-\frac{2^\alpha \log(2)}{2^\alpha-1}+\log\left[1+\Phi\left(\frac{x-\mu}{\sigma}\right)\right]
\end{aligned}
$$

The second derivatives of $l(\boldsymbol{\theta})$ are

$$
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu^2} = -\frac{1}{\sigma^2} - \left(\frac{\alpha-1}{\sigma^2}\right)\left[\left(\frac{x-\mu}{\sigma}\right)\kappa(x) + \kappa^2(x)\right]
$$

$$
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu \partial \sigma} = -\frac{2}{\sigma^2}\left(\frac{x-\mu}{\sigma}\right) + \left(\frac{\alpha-1}{\sigma^2}\right)\kappa(x)\left[1 - \left(\frac{x-\mu}{\sigma}\right)^2 - \left(\frac{x-\mu}{\sigma}\right)\kappa(x)\right]
$$

$$
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu \partial \alpha} = -\frac{k(x)}{\sigma}
$$

$$
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma^2} = \frac{1}{\sigma^2} - \frac{3}{\sigma^2}\left(\frac{x-\mu}{\sigma}\right)^2 + \left(\frac{\alpha-1}{\sigma^2}\right)\left(\frac{x-\mu}{\sigma}\right)\kappa(x)\left[2 - \left(\frac{x-\mu}{\sigma}\right)^2 - \left(\frac{x-\mu}{\sigma}\right)\kappa(x)\right]
$$

$$
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma \partial \alpha} = -\left(\frac{x-\mu}{\sigma^2}\right)k(x)
$$

$$
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} = -\frac{1}{\alpha^2} + \frac{2^\alpha \log^2(2)}{(2^\alpha - 1)^2}
$$

References

1. Azzalini, A. A class of distributions which includes the normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
2. Arellano-Valle, R.B.; Gómez, H.W.; Quintana, F.A. A New Class of Skew-Normal Distributions. *Commun. Stat. Theory Methods* **2004**, *33*, 1465–1480. [CrossRef]
3. Arellano-Valle, R.B.; Gómez, H.W.; Salinas, H.S. A note on the Fisher information matrix for the skew-generalized-normal model. *Stat. Oper. Res. Trans.* **2013**, *37*, 19–28.
4. Azzalini, A. *The Skew-Normal and Related Families*; IMS monographs; Cambridge University Press: New York, NY, USA 2014.
5. Martínez-Flórez, G.; Barranco-Chamorro, I.; Bolfarine, H.; Gómez, H.W. Flexible Birnbaum–Saunders Distribution. *Symmetry* **2019**, *11*, 1305. [CrossRef]
6. Contreras-Reyes, J.E.; Arellano-Valle, R.B. Kullback–Leibler Divergence Measure for Multivariate Skew-Normal Distributions. *Entropy* **2012**, *14*, 1606–1626. [CrossRef]
7. Chiogna, M. A note on the asymptotic distribution of the maximum likelihood estimator for the scalar skew-normal distribution. *Stat. Methods Appl.* **2005**, *14*, 331–341. [CrossRef]
8. Arellano-Valle, R.B.; Azzalini, A. The centred parametrization for the multivariate skew-normal distribution. *J. Multivar. Anal.* **2008**, *99*, 1362–1382. [CrossRef]
9. Pewsey, A.; Gómez, H.W.; Bolfarine, H. Likelihood-based inference for power distributions. *Test* **2012**, *21*, 775–789. [CrossRef]
10. Lehmann, E.L. The power of rank tests. *Ann. Math. Statist.* **1953**, *24*, 23–43. [CrossRef]
11. Durrans, S.R. Distributions of fractional order statistics in hydrology. *Water Resour. Res.* **1992**, *28*, 1649–1655. [CrossRef]
12. Gupta, D.; Gupta, R.C. Analyzing skewed data by power normal model. *Test* **2008**, *17*, 197–210. [CrossRef]
13. Martínez-Flórez, G.; Arnold, B.C.; Bolfarine, H.; Gómez, H.W. The Multivariate Alpha-power Model. *J. Stat. Plan. Inference* **2013**, *143*, 1236–1247. [CrossRef]
14. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. Asymmetric regression models with limited responses with an application to antibody response to vaccine. *Biom. J.* **2013**, *55*, 156–172. [CrossRef] [PubMed]
15. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. The log alpha-power asymmetric distribution with application to air pollution. *Environmetrics* **2014**, *25*, 44–56. [CrossRef]
16. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. Doubly censored power-normal regression models with inflation. *Test* **2015**, *24*, 265–286. [CrossRef]
17. Castillo, N.O.; Gallardo, D.I.; Bolfarine, H.; Gómez, H.W. Truncated Power-Normal Distribution with Application to Non-Negative Measurements. *Entropy* **2018**, *20*, 433. [CrossRef]

18. Maciak, M.; Peštová, B.; Pešta, M. Structural breaks in dependent, heteroscedastic, and extremal panel data. *Kybernetika* **2018**, *54*, 1106–1121. [CrossRef]

19. Pešta, M. Total least squares and bootstrapping with application in calibration. *Statistics* **2013**, *47*, 966–991. [CrossRef]

20. Shaked, M.; Shanthikumar, J.G. *Stochastic Orders*; Springer Science & Business Media: New York, NY, USA, 2007.

21. R Development Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.

22. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]

23. Schwarz, G. Estimating the dimension of a model. *Ann. Statist.* **1978**, *6*, 461–464. [CrossRef]

24. Jäntschi, L.; Bolboacă, S.D. Computation of Probability Associated with Anderson–Darling Statistic. *Mathematics* **2018**, *6*, 88. [CrossRef]

25. Jäntschi, L. A Test Detecting the Outliers for Continuous Distributions Based on the Cumulative Distribution Function of the Data Being Tested. *Symmetry* **2019**, *11*, 835. [CrossRef]