

Research Article

The Slash Power Normal Distribution with Application to Pollution Data

Mingming Chen ¹, Jianghong Ma ² and Yee Leung³

¹School of Mathematics and Data Science, Shaanxi University of Science and Technology, Xi'an 710021, China

²Department of Mathematics and Information Science, Chang'an University, Xi'an 710064, China

³Department of Geography and Resource Management,

Institute of Future Cities Stanley Ho Big Data Decision Analytic Center The Chinese University of Hong Kong, Hong Kong

Correspondence should be addressed to Jianghong Ma; jhma@chd.edu.cn

Received 12 July 2021; Revised 12 November 2021; Accepted 7 January 2022; Published 3 February 2022

Academic Editor: Hector Gomez

Copyright © 2022 Mingming Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we introduce a new slash distribution via the power normal distribution and uniform distribution. This new distribution, called the slash power normal distribution, models more appropriately the skewness and kurtosis of data than the power normal distribution. Moreover, the probability density function and the cumulative distribution function of the slash power normal distribution are derived and the density function curves with different parameters are given. We also study the basic properties of the moments, derive the maximum likelihood estimation of the parameters, and substantiate our arguments with numerical simulations. Finally, we model hourly measurements of sulphur dioxide concentrations of a station in Hong Kong by the slash power normal distribution, power normal distribution, and skew normal distribution and use the K-S test to evaluate the model fittings. The result demonstrates that the slash power normal distribution gives a better fit to the data.

1. Introduction

With the advancement of research in statistics and its applications, there has been a growing interest in the construction of flexible parametric classes of distributions exhibiting skewness or heavy-tailed characteristics that are different from those of the normal distribution. Some skewed symmetric distributions have been studied, for example, by Arnold and Beaver [1], Gupta et al. [2], Wang et al. [3], Azzalini [4], and Arellano-Valle and Azzalini [5]. The motivation originates from the analysis of real-life data, including financial, environmental, and medical data whose distributions often do not follow the normal law (see as reported in Azzalini et al. [6], Bartoletti and Loperfido [7], Adcock et al. [8], Hossain and Beyene [9], and Cancela and Pires [10]). It is thus necessary to find and study from the theoretical and practical points of view distributions that can better model skewed or heavy-tailed data.

Kafadar [11] introduced the standard slash normal (SN) distribution which is defined as the distribution of the ratio $X = Z/U^{1/q}$ of a standard normal random variable Z and an independent uniform random variable U on the interval $(0,1)$ raised to the power $1/q, q > 0$. We write as $X \sim SN(q)$. When $q \rightarrow \infty$, it yields the standard normal distribution.

In fact, straightforward algebra yields the probability density function (p.d.f.) of a univariate SN distribution $SN(q)$,

$$\psi(x; q) = q \int_0^1 u^q \phi(xu) du, \quad (1)$$

and the corresponding cumulative distribution function (c.d.f.) of the univariate slash normal distribution,

$$\Psi(x; q) = q \int_0^1 u^q \Phi(xu) du, \quad (2)$$

where ϕ and Φ are the standard normal density and distribution function, respectively. Hence, the expectation and the variance of the standard $SN(q)$ distribution are $E(X) = 0$ for $q > 1$ and $\text{Var}(X) = q/(q-2)$ for $q > 2$.

A general slash normal distribution is obtained by performing scale multiplication and location shift of a standard slash normal random variable (see Rogers and Tukey [12] and Mosteller and Tukey [13] for further properties). Wang and Genton [14] generalized Kafadar's univariate slash normal distribution to the multivariate slash normal distribution and investigated its properties. Compared to the standard normal distribution, the standard slash normal distribution has heavier tails and is symmetric about the origin. Thus, it has been used to simulate heavy-tailed data (see, for example, Andrew et al. [15] and Gross [16]).

Extensions of the slash normal distribution have been considered by several authors; among them, Gui [17] introduced a new extension which is not only heavy-tailed but also unimodal and bimodal via the alpha skew normal distribution introduced by Elal-Olivero [18]. Furthermore, the structure of uniform slash and α -slash distributions under discrete and continuous is explored by Jones and Higuchi [19].

In this paper, we consider a generalization of the slash normal distribution, called the slash power normal (SPN) distribution, via the power normal (PN) distribution introduced by Durrans [20]. The $PN(\alpha)$ distribution is a special form of fractional order statistical distribution. Formally, we have $X \sim PN(\alpha)$ if the p. d. f. of X is

$$f_{PN}(x; \alpha) = \alpha [\Phi(x)]^{\alpha-1} \phi(x), \alpha > 0, -\infty < x < +\infty. \quad (3)$$

The main idea of this distribution was introduced in detail by Lehmann [21]. Gupta and Gupta [22] studied the fundamental properties of PN distribution. One usual extension that should be noted is the location-scale version of PN distributions. If $Y \sim PN(\alpha)$, then $X = \mu + \sigma Y$, where $-\infty < \mu < \infty$ and $\sigma > 0$, has the p.d.f. given by

$$f_{PN}(x; \alpha, \mu, \sigma) = \frac{\alpha}{\sigma} \left[\Phi\left(\frac{x-\mu}{\sigma}\right) \right]^{\alpha-1} \phi\left(\frac{x-\mu}{\sigma}\right), \quad (4)$$

and is denoted by $X \sim PN(\alpha, \mu, \sigma)$.

We derive the p.d.f. and c.d.f. of the SPN distributions. We also study their moments and the maximum likelihood estimation of their parameters.

The rest of this paper is organized as follows. In Section 2, we propose a generalized skewed slash distribution to model heavy-tailed and skewed data and study its properties. In Section 3, we give the maximum likelihood estimation of the parameters. We substantiate our theoretical arguments with a simulation study in Section 4 and a real-life application in Section 5. We then conclude our paper in Section 6.

2. Slash Power Normal Distribution

We give formally the definition of a slash power normal distribution in the stochastic representation by the following theorem.

2.1. Stochastic Representation. We derive the probability density function of the slash power normal distribution in the following theorem.

Theorem 1. *Let Y and U be two independent random variables, where $Y \sim PN(\alpha)$ and $U \sim U(0, 1)$. Then the p.d.f. of*

$$X = \frac{Y}{U^{1/q}}, \quad (5)$$

is

$$f(x; \alpha, q) = q\alpha \int_{-\infty}^{+\infty} [\Phi(xv)]^{\alpha-1} \phi(xv)v^q dv, \quad (6)$$

where $-\infty < x < +\infty$ and $\alpha, q > 0$.

Proof. Let $\begin{cases} x = y/v \\ v = u^{1/q} \end{cases}$ and the inverse function is $\begin{cases} y = xv \\ u = v^q \end{cases}$. The Jacobian determinant is

$$|J| = \begin{vmatrix} \frac{\partial y}{\partial x} & \frac{\partial y}{\partial v} \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial v} \end{vmatrix} = \begin{vmatrix} v & x \\ 0 & qv^{q-1} \end{vmatrix} = qv^q. \quad (7)$$

We have the joint p.d.f. of X and V

$$p(x, v) = p_Y(x, v)p_U(v)|J| = \alpha [\Phi(xv)]^{\alpha-1} \phi(xv)1_{[0,1]}(v)qv^q. \quad (8)$$

Now the marginal p.d.f. of X is given by

$$f(x) = \int_{-\infty}^{+\infty} p(x, v)dv = q\alpha \int_0^1 [\Phi(xv)]^{\alpha-1} \phi(xv)v^q dv. \quad (9)$$

This completes the proof of the theorem.

Since the p.d.f. in (6) is reduced to the SN distribution for $\alpha = 1$, we can say that the distribution given by (6) generalizes the SN distribution. Thus, we have the following definition. \square

Definition 1. If the p.d.f. of a random variable X is given by (6), that is,

$$f(x; \alpha, q) = q\alpha \int_0^1 [\Phi(xv)]^{\alpha-1} \phi(xv)v^q dv, \quad (10)$$

where $\alpha, q > 0$ are shape parameters, we say that X has the slash power normal (SPN) distribution, denoted by $X \sim \text{SPN}(\alpha, q)$. The c.d.f. of the SPN distribution is given as follows. The graph is given in Figure 1.

$$\begin{aligned}
 F(x; \alpha, q) &= \int_{-\infty}^x f(u; \alpha, q) du \\
 &= \int_0^1 \int_{-\infty}^x q\alpha [\Phi(uv)]^{\alpha-1} \varphi(uv) v^q du dv \\
 &= q\alpha \int_0^1 \int_{-\infty}^x [\Phi(uv)]^{\alpha-1} v^{q-1} d\Phi(uv) dv \quad (11) \\
 &= q\alpha \int_0^1 \int_{-\infty}^{\Phi(xv)} w^{\alpha-1} v^{q-1} dw dv, w = \Phi(uv) \\
 &= q \int_0^1 v^{q-1} [\Phi(xv)]^\alpha dv.
 \end{aligned}$$

In different types of statistical analysis, it is common to study the survival function, the hazard function, and the inverse risk function. For the SPN, the survival and hazard functions are given separately by

$$\begin{aligned}
 S(x) &= 1 - F(x) = 1 - q \int_0^1 [\Phi(xv)]^\alpha v^{q-1} dv, \\
 h(x) &= \frac{f(x)}{S(x)} = \frac{q\alpha \int_0^1 [\Phi(xv)]^{\alpha-1} \phi(xv) v^q dv}{1 - q \int_0^1 [\Phi(xv)]^\alpha v^{q-1} dv}. \quad (12)
 \end{aligned}$$

Furthermore, the inverse hazard function is

$$r(x) = \frac{f(x)}{S(x)} = \frac{q\alpha \int_0^1 [\Phi(xv)]^{\alpha-1} \phi(xv) v^q dv}{1 - q \int_0^1 [\Phi(xv)]^\alpha v^{q-1} dv}. \quad (13)$$

2.2. Density Shape and Special Cases. Similar to the power normal distribution, we graphically show a typical p.d.f. of the SPN distributions with various combinations of parameters.

By the graphical presentation in Figure 2, the p.d.f. of the SPN distribution has heavier tail and lower peak than the PN distribution.

For the ordinary slash normal distribution, it is well known that the thickness of the tail increases with decreasing values of q . From Figure 3, we can observe that for fixed α , the same is also true for the slash power normal distribution.

The SPN distribution has a unimodal density which is skewed to the left if $0 < \alpha < 1$ and to the right if $\alpha > 1$. Figure 4 depicts a few p.d.f. graphs for different values of α . When $\alpha = 1$, the distribution is symmetric since $\varphi(-u) = \varphi(u)$. In this case, the p.d.f. of the distribution is given by

$$f(x; 1, q) = q \int_0^1 \phi(xv) v^q dv. \quad (14)$$

This means that the SN distribution is a special case of the SPN distribution.

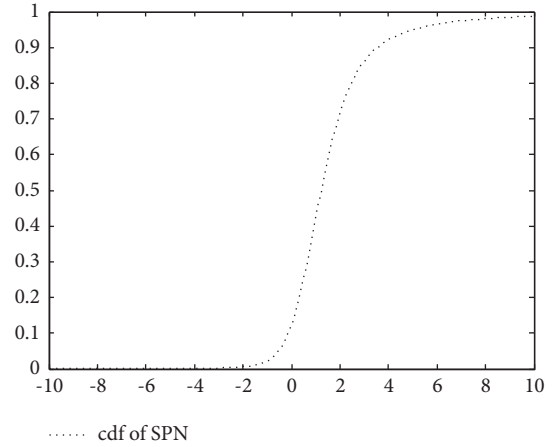


FIGURE 1: The c.d.f. of the SPN distribution.

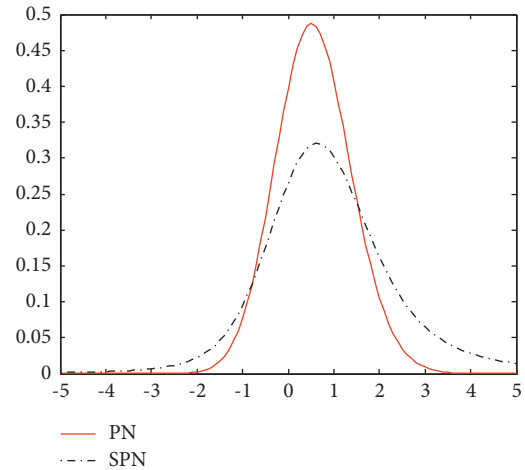


FIGURE 2: The p.d.f. of the PN and SPN distribution.

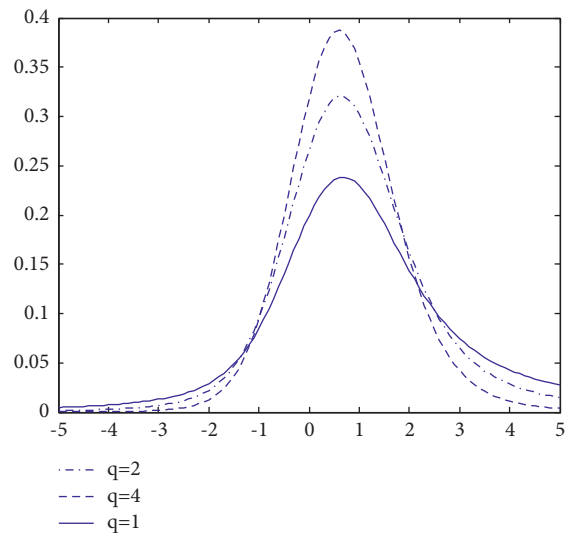


FIGURE 3: The p.d.f. of the SPN distribution with various q values.

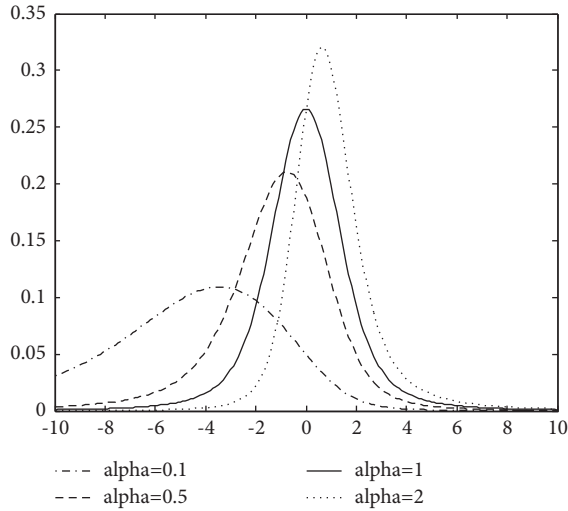


FIGURE 4: The p.d.f. of the SPN distribution with various α values.

When $\alpha = 2$, we have the standard univariate skew-slash distribution, which is introduced by Wang and Genton [14] with the p.d.f. $f(x; \lambda, q) = 2q \int_0^1 \Phi(xv) \phi(\lambda xv) v^q dv = \eta(x; 0, 1, q, 1)$ (see [14]).

When $\alpha = 1$ and $q = 1$, the distribution is reduced to

$$\begin{aligned} f(x; 1, 1) &= \int_0^1 \phi(xv) v dv = \frac{1}{x} \Phi(x) - \frac{1}{x^2} \int_0^x \Phi(u) du \\ &= \frac{1}{x} \Phi(x) - \frac{1}{x^2} [u\Phi(u) + \phi(u)]_0^x \\ &= \frac{1}{x^2} [\phi(0) - \phi(x)], \quad x \neq 0. \end{aligned} \quad (15)$$

That is,

$$f(x; 1, 1) = \begin{cases} \frac{1}{x^2} [\phi(0) - \phi(x)], & x \neq 0, \\ \frac{\phi(0)}{2}, & x = 0. \end{cases} \quad (16)$$

The above distribution is just the standard SN distribution $SN(1)$.

Also, in terms of Theorem 1, the SPN distribution approaches the PN distribution when q approaches infinity.

Proposition 1. Let $U \sim U(0, 1)$ and $X|U = u \sim PN(\alpha, 0, u^{-1/q})$, $\alpha > 0$, $q > 0$, then

$$X \sim SPN(\alpha, q). \quad (17)$$

Proof. According to the definition of $PN(\alpha, \mu, \sigma)$, we have

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f(x, u) du \\ &= \int_{-\infty}^{+\infty} f_{X|U}(x|u) f_U(u) du \\ &= \int_0^1 \frac{\alpha}{u^{-1/q}} \left[\Phi\left(\frac{x}{u^{-1/q}}\right) \right]^{\alpha-1} \phi\left(\frac{x}{u^{-1/q}}\right) du \quad (t = U^{1/q}) \\ &= q \int_0^1 \alpha t \Phi^{\alpha-1}(xt) \phi(xt) t^{q-1} dt \\ &= q\alpha \int_0^1 \Phi^{\alpha-1}(xt) \phi(xt) t^q dt, \end{aligned} \quad (18)$$

which is just the same as (6). Thus the proof is completed.

The proposition shows that the SPN distribution can be represented as a mixture of the PN distribution and uniform $U(0, 1)$ distribution. The result gives us an alternative way to generate random numbers from the SPN distribution.

2.3. Random Number Generation. Firstly, we employ the inverse function method to generate a random number y from the

$PN(\alpha)$ distribution, and the p.d.f. is $f(y) = \alpha [\Phi(y)]^{\alpha-1} \phi(y)$.

Secondly, we generate a random number u on $(0, 1)$. Then, the quotient $x = y/u^{1/q}$ will serve as a random variable from the $SPN(\alpha, q)$ distribution.

2.4. Moments of the Slash Power Normal Distribution. If a random variable $X \sim SPN(\alpha, q)$ and has the stochastic representation in Theorem 1, then the mean of X is given by

$$\begin{aligned} E[X] &= E[YU^{-1/q}] = E[Y]E[U^{-1/q}] \\ &= \frac{q}{q-1} E[Y], \quad q > 1. \end{aligned} \quad (19)$$

Let

$$C_n(\lambda) = \int_{-\infty}^{+\infty} [\Phi(\lambda y)]^n \phi(y) dy, \quad n = 0, 1, 2, \dots, \lambda \in R,$$

$$\begin{aligned} E[Y] &= \alpha \int_{-\infty}^{+\infty} y [\Phi(y)]^{\alpha-1} \phi(y) dy \\ &= \frac{\alpha(\alpha-1)}{2\sqrt{\pi}} \int_{-\infty}^{+\infty} \left[\Phi\left(\frac{y}{\sqrt{2}}\right) \right]^{\alpha-2} \phi(y) dy \\ &= \frac{\alpha(\alpha-1)}{2\sqrt{\pi}} C_{\alpha-2}\left(\frac{1}{\sqrt{2}}\right). \end{aligned} \quad (20)$$

Clearly, $C_0(\lambda) = 1$; then,

$$C_{2n+1}(\lambda) = \sum_{i=1}^{2n+1} (-1)^{i+1} \binom{2n+1}{i} \frac{1}{2^i} C_{2n+1-i}(\lambda). \quad (21)$$

See Arnold and Beaver [23] for the proof.

We can now easily obtain $C_1(\lambda) = 1/2, C_2(\lambda) = 1/\pi \arctan \sqrt{1+2\lambda^2}, C_3(\lambda) = 3/2\pi \arctan \sqrt{1+2\lambda^2} - 1/4$, etc.

Therefore,

$$\begin{array}{l} \alpha \quad : \quad 1 \quad 2 \quad 3 \quad 4 \quad \dots \\ E[Y] \quad : \quad 0 \quad \frac{1}{\sqrt{\pi}} \quad \frac{3}{2\sqrt{\pi}} \quad \frac{6 \arctan \sqrt{2}}{\pi^{3/2}} \quad \dots \\ E[X] \quad : \quad 0 \quad \frac{q}{\sqrt{\pi}(q-1)} \quad \frac{3q}{2\sqrt{\pi}(q-1)} \quad \frac{6q \arctan \sqrt{2}}{\pi^{3/2}(q-1)} \quad \dots \end{array} \quad (22)$$

In order to obtain the second moment of X , we need to compute the second moment of Y^2 ($t < 1/2$). Firstly, when $\alpha = 1$, it is obvious that

$$E[Y^2] = \int_{-\infty}^{+\infty} y^2 \phi(y) dy = 1 \quad (23)$$

Then, for $\alpha = 2, 3, \dots$, we have

$$\begin{aligned} E[Y^2] &= \alpha \int_{-\infty}^{+\infty} y^2 [\Phi(y)]^{\alpha-1} \phi(y) dy \\ &= -\alpha \int_{-\infty}^{+\infty} y [\Phi(y)]^{\alpha-1} d\phi(y) \\ &= -\alpha \{y[\Phi(y)]^{\alpha-1} \phi(y)\}_{-\infty}^{+\infty} + \alpha \int_{-\infty}^{+\infty} \{[\Phi(y)]^{\alpha-1} + y(\alpha-1)[\Phi(y)]^{\alpha-2} \phi(y)\} \phi(y) dy \\ &= [\Phi(y)]_{-\infty}^{+\infty} + \alpha(\alpha-1) \int_{-\infty}^{+\infty} y [\Phi(y)]^{\alpha-2} [\Phi(y)]^2 dy \\ &= 1 + \frac{\alpha(\alpha-1)}{2\pi} \int_{-\infty}^{+\infty} [\Phi(y)]^{\alpha-2} e^{-y^2} y dy \\ &= 1 + \frac{\alpha(\alpha-1)}{4\pi} \left(- \int_{-\infty}^{+\infty} [\Phi(y)]^{\alpha-2} d e^{-y^2} \right) \quad (\text{It can be seen that } E[Y^2] = 1 \text{ if } \alpha = 2) \\ &= 1 + \frac{\alpha(\alpha-1)(\alpha-2)}{4\pi} \int_{-\infty}^{+\infty} [\Phi(y)]^{\alpha-3} \phi(y) e^{-y^2} dy \\ &= 1 + \frac{\alpha(\alpha-1)(\alpha-2)}{4\pi} \int_{-\infty}^{+\infty} [\Phi(y)]^{\alpha-3} \phi(\sqrt{3} y) dy \\ &= 1 + \frac{\alpha(\alpha-1)(\alpha-2)}{4\pi\sqrt{3}} \int_{-\infty}^{+\infty} \left[\Phi\left(\frac{u}{\sqrt{3}}\right) \right]^{\alpha-3} \phi(u) du \\ &= 1 + \frac{\alpha(\alpha-1)(\alpha-2)}{4\pi\sqrt{3}} C_{\alpha-3}\left(\frac{1}{\sqrt{3}}\right). \end{aligned} \quad (24)$$

Therefore,

$$\begin{aligned}
 \alpha & : 1 & 2 & 3 & 4 & \dots \\
 E[Y^2] & : 1 & 1 & 1 + \frac{\sqrt{3}}{2\pi} & 1 + \frac{\sqrt{3}}{\pi} & \dots \\
 E[X^2] & : \frac{q}{q-2} & \frac{q}{q-2} & \frac{q}{q-2} \left(1 + \frac{\sqrt{3}}{2\pi}\right) & \frac{q}{q-2} \left(1 + \frac{\sqrt{3}}{\pi}\right) & \dots
 \end{aligned} \tag{25}$$

Then, $\text{Var}(X) = E(X^2) - [E(X)]^2$ can be obtained by using the abovementioned expressions.

For special cases, when $\alpha > 2$ is a positive integer, we have

From the definition of X , the general moments are given by

$$\begin{aligned}
 E[X^k] &= E[YU^{-1/q}]^k = E[Y^k]E[U^{-k/q}] \\
 &= \frac{q}{q-k} E[Y^k] \\
 &= \frac{q\alpha}{q-k} \int_{-\infty}^{+\infty} [\Phi(y)]^{\alpha-1} \phi(y), q > k.
 \end{aligned} \tag{26}$$

$$\begin{aligned}
 E[Y^k] &= \int_{-\infty}^{+\infty} y^k \alpha [\Phi(y)]^{\alpha-1} \phi(y) dy \\
 &= \alpha \int_{-\infty}^{+\infty} y^{k-1} [\Phi(y)]^{\alpha-1} y \phi(y) dy \\
 &= \alpha \int_{-\infty}^{+\infty} (k-1) y^{k-2} [\Phi(y)]^{\alpha-1} \phi(y) dy + \alpha \int_{-\infty}^{+\infty} (\alpha-1) y^{k-1} [\Phi(y)]^{\alpha-2} \phi^2(y) dy \\
 &= (k-1)E[Y^{k-2}] + \frac{\alpha(\alpha-1)}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^{k-1} [\Phi(y)]^{\alpha-2} \phi(\sqrt{2}y) dy \\
 &= (k-1)E[Y^{k-2}] + \frac{\alpha(\alpha-1)}{\sqrt{2\pi}\sqrt{2^k}} \int_{-\infty}^{+\infty} v^{k-1} \left[\Phi\left(\frac{v}{\sqrt{2}}\right)\right]^{\alpha-2} \phi(v) dv \quad (v = \sqrt{2}y) \\
 &= (k-1)E[Y^{k-2}] + \frac{\alpha(\alpha-1)C_{\alpha-2}(1/\sqrt{2})}{\sqrt{2\pi}\sqrt{2^k}} E(V^{k-1}), \quad k = 3, 4, \dots,
 \end{aligned} \tag{27}$$

where the p.d.f. of V is given by

$$f(v) = \frac{[\Phi(v/\sqrt{2})]^{\alpha-2} \phi(v)}{C_{\alpha-2}(1/\sqrt{2})}, \text{ where } C_{\alpha-2}\left(\frac{1}{\sqrt{2}}\right) = \int_{-\infty}^{+\infty} \left[\Phi\left(\frac{1}{\sqrt{2}}\right)\right]^{\alpha-2} \phi(v) dv. \tag{28}$$

And $f(v)$ is the p.d.f. of the generalized skew normal distribution (see Gupta and Gupta [24]). Then,

$$\begin{aligned}
 E[X^k] &= \frac{q}{q-k} E[Y^k] \\
 &= \frac{q(k-1)}{q-k} E[Y^{k-2}] + \frac{q\alpha(\alpha-1)C_{\alpha-2}(1/\sqrt{2})}{\sqrt{2\pi}\sqrt{2^k}(q-k)} E(V^{k-1}), \quad k = 3, 4, \dots
 \end{aligned}
 \tag{29}$$

3. Maximum Likelihood Estimation

In this section, we consider the location-scale form of the distribution and introduce an additional location parameter μ and a scale parameter σ to the density function. In this form, for $\alpha = 1$ and $q \rightarrow \infty$, the $N(\mu, \sigma^2)$ distribution is obtained. The density function of the general form is given by

$$f(x; \theta) = \frac{q\alpha}{\sigma} \int_0^1 \left[\Phi\left(\frac{x-\mu}{\sigma}v\right) \right]^{\alpha-1} \phi\left(\frac{x-\mu}{\sigma}v\right) v^q dv, \tag{30}$$

where $\theta = (\alpha, q, \mu, \sigma)$.

Suppose X_1, X_2, \dots, X_n is a random sample of size n from (30). Then, the log-likelihood function is expressed as

$$\begin{aligned}
 l(\theta) &= \log \prod_{i=1}^n f(x_i, \alpha, q, \mu, \sigma) \\
 &= n \log q + n \log \alpha - n \log \sigma + \sum_{i=1}^n \log \int_0^1 \left[\Phi\left(\frac{x_i-\mu}{\sigma}v\right) \right]^{\alpha-1} \phi\left(\frac{x_i-\mu}{\sigma}v\right) v^q dv.
 \end{aligned}
 \tag{31}$$

The maximum likelihood estimates (MLE) of the parameters maximize this likelihood function. Taking the partial derivatives of the log-likelihood function with respect

to α, q, μ , and σ , respectively, and equalizing the obtained expressions to zero yield the likelihood equations

$$\begin{aligned}
 l_\alpha &= \frac{n}{\alpha} + \frac{\sum_{i=1}^n \int_0^1 \Phi^{\alpha-1}(u_i) \log \Phi(u_i) \phi(u_i) v^q dv}{\int_0^1 \Phi^{\alpha-1}(u_i) \phi(u_i) v^q dv} = 0, \\
 l_q &= \frac{n}{q} + \frac{\sum_{i=1}^n \int_0^1 \Phi^{\alpha-1}(u_i) \phi(u_i) v^q \log v dv}{\int_0^1 \Phi^{\alpha-1}(u_i) \phi(u_i) v^q dv} = 0, \\
 l_\mu &= \frac{n}{\alpha} + \frac{\sum_{i=1}^n \int_0^1 \Phi^{\alpha-2}(u_i) \phi(u_i) v^{q+1} [(1-\alpha)\phi(u_i) + u_i \Phi(u_i)] dv}{\int_0^1 \Phi^{\alpha-1}(u_i) \phi(u_i) v^q dv} = 0, \\
 l_\sigma &= -n + \frac{\sum_{i=1}^n \int_0^1 \Phi^{\alpha-2}(u_i) \phi(u_i) u_i v^q [(1-\alpha)\phi(u_i) + u_i \Phi(u_i)] dv}{\int_0^1 \Phi^{\alpha-1}(u_i) \phi(u_i) v^q dv} = 0,
 \end{aligned}
 \tag{32}$$

where $u_i = (x_i - \mu)v/\sigma$.

It should be noted that the abovementioned maximum likelihood estimation equations are not in simple analytical form. Therefore, the estimates can be obtained via numerical procedures such as the Newton-Raphson method and the L-BFGS-B methods. The programs in software packages such as MATLAB and R provide computing routines for solving such nonlinear optimization problems. In this paper, we will use the nonlinear optimization routine optim in R which uses the L-BFGS-B method to compute the estimates.

4. Simulation Study

In this section, we employ a simulation experiment to illustrate the behavior of the maximum likelihood estimators of the parameters. It is well known that the maximum likelihood estimators are asymptotically normal and unbiased under some regularity conditions. However, we cannot derive the Fisher information matrix because the log-likelihood function is a complex transformation. It is thus difficult to obtain the theoretical properties of the maximum likelihood estimators. Hence, we only numerically study the properties of the estimators.

TABLE 1: Empirical means and SD for the MLE estimator of α , q , μ , and σ .

α	q	μ	σ	$n = 20$				$n = 20$			
				$\hat{\alpha}$ (SD)	\hat{q} (SD)	$\hat{\mu}$ (SD)	$\hat{\sigma}$ (SD)	$\hat{\alpha}$ (SD)	\hat{q} (SD)	$\hat{\mu}$ (SD)	$\hat{\sigma}$ (SD)
2	4	3	2	2.2424 (0.5508)	4.9144 (0.8583)	2.8099 (0.5481)	2.0757 (0.1916)	1.9917 (0.4190)	4.1486 (0.6273)	3.0395 (0.4654)	1.9972 (0.1649)
3	3	2	4	3.720 (0.6925)	3.2037 (0.3806)	1.5117 (0.9553)	4.1820 (0.3337)	3.1885 (0.6213)	3.0733 (0.3497)	1.7942 (0.8882)	4.0413 (0.3129)
1	5	2	3	1.1133 (0.2487)	5.1366 (0.9776)	1.7256 (0.7528)	3.1274 (0.2578)	1.0317 (0.2072)	4.9623 (0.7574)	1.9576 (0.7266)	3.0311 (0.2469)
2	3	4	2	2.1801 (0.4083)	3.6099 (0.4596)	3.8509 (0.4438)	2.0694 (0.1606)	2.0028 (0.3291)	3.0744 (0.3357)	4.0191 (0.3916)	2.0031 (0.1428)

TABLE 2: Descriptive statistics of hourly SO₂ concentrations.

Mean	SD	Median	Skewness	Kurtosis
8.595	6.935	7.000	3.375	21.113

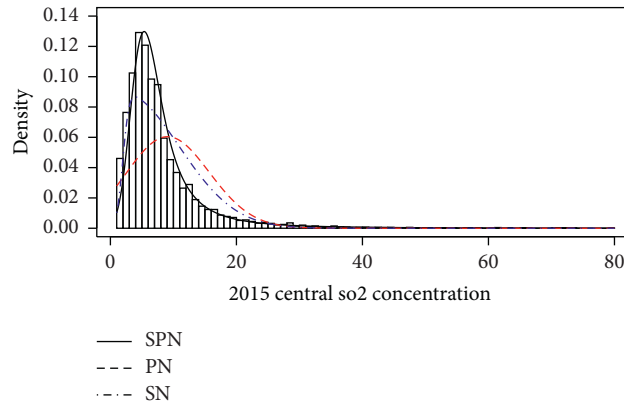


FIGURE 5: Histogram of the SO₂ concentration dataset and the graphs of the fitted models.

TABLE 3: MLE (with (SD)) for the SPN, SN, and PN distributions.

Model	$\hat{\alpha}$	\hat{q}	$\hat{\mu}$	$\hat{\sigma}$	D	AIC	BIC	CAIC
SPN	9.467 (0.694)	2.115 (0.039)	0.001 (0.199)	3.119 (0.040)	0.00932	50636.72	50665.01	50636.72
SN	11.202 (0.532)	--	1.960 (0.042)	8.921 (0.073)	0.0212	53267.2	53288.42	53267.2
PN	1.367 (0.092)	--	7.348 (0.439)	7.247 (0.079)	0.0914	58163.4	58184.62	58163.4

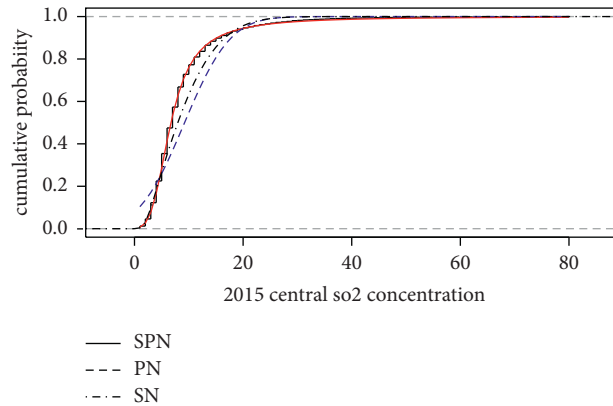


FIGURE 6: Empirical distribution of the SO₂ and the graphs of the fitted models.

We first generate 1000 samples of sizes 20 and 100 from the SPN distribution with fixed parameters. The estimates are computed by the optim function using the L-BFGS-B method in R. The empirical means and standard deviations of the estimates are given in Table 1.

As can be observed from Table 1, the estimates approach the true values as the sample size increases, which implies the consistency of the estimates.

5. Real-Life Data Analysis

We further demonstrate the existence of the slash power normal distribution in real life in this section. The dataset studied is about air pollution in Hong Kong, a high-density

city in China. The data contain the hourly measurements of sulphur dioxide ($\mu\text{g}/\text{m}^3$) concentrations of a station in Central, Hong Kong, in January–December 2015. Table 2 summarizes the basic descriptive statistics of the dataset. Figure 5 reveals the positive skewness intrinsic in the data, so the symmetrical distributions are no longer appropriate for describing such data.

We fit the dataset with the SPN(α, q, μ, σ), PN(α, μ, σ), and SN(α, μ, σ) and use the Kolmogorov–Smirnov (K-S) test to evaluate the model fittings. The K-S test quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution. The K-S test statistic is defined as

$$D = \sup_x |F_n(x) - F(x)|, \quad (33)$$

where $F_n(x)$ is the empirical distribution function and $F(x)$ is the theoretical cumulative distribution of the distribution being tested. The results are presented in Table 3, where the initial values of each parameter are $\alpha = 7, q = 5, \mu = 2$, and $\sigma = 2$.

Table 3 shows the maximum likelihood estimates together with the standard errors of different distributions. The statistic D indicates that the proposed SPN model is a best fit, which is consistent with Figure 6. The Akaike information criterion (AIC), Bayesian information criterion (BIC), and corrected Akaike information criterion (CAIC) are used to measure the goodness of fit of the models. $AIC = 2k - 2 \log L$, $BIC = k \log n - 2 \log L$, and $CAIC = 2kn/n - k - 1 - 2 \log L$, where n is the sample size, k is the number of parameters in the model, and L is the maximized value of the likelihood function for the estimated model. The best model is the one with the smallest AIC (or BIC or CAIC).

From Table 3, AIC, BIC, or CAIC shows that the proposed SPN model is a best fit. We can also observe that the SPN model successfully captures the kurtosis and skewness of the SO₂ concentration data of Central, Hong Kong (see Figure 5). This shows that the SPN model provides a good vehicle to deal with skewed and heavy-tailed data.

6. Conclusion

In this paper, we have proposed and analyzed a new slash version of the power normal distribution. It is defined as the ratio of two independent random variables, namely, the power normal and the power of the uniform distribution. The slash normal, skew-slash normal, and related distributions turn out to be special cases of the proposed slash power normal distribution. We have studied its basic properties, including the variance and the general moments.

The SPN distribution is fitted to a real-life dataset by the maximum likelihood approach and we have compared it with the power normal and skew normal distributions. The empirical result shows that the proposed SPN model better fits the dataset and it provides us a more appropriate model to deal with skewed and heavy-tailed data.

Data Availability

The data used to support the findings of this study have been deposited in the repository (https://www.epd.gov.hk/epd/sc_chi/environmentinhk/air/air_quality/air_quality.html).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

This research was funded by the earmarked grant CUHK 446213 of the Hong Kong Research Grant Council, the National Natural Science Foundation of China (Grant no.

11261044), and Scientific Research Program Funded by Shaanxi Provincial Education Department (Program no. 20JK0549).

References

- [1] B. C. Arnold and R. J. Beaver, "The skew-Cauchy distribution," *Statistics & Probability Letters*, vol. 49, no. 3, pp. 285–290, 2000.
- [2] A. K. Gupta, F. C. Chang, and W. J. Huang, "Some skew-symmetric models," *Random Operators and Stochastic Equations*, vol. 10, no. 2, pp. 133–140, 2002.
- [3] J. Wang, J. Boyer, and M. G. Genton, "A skew-symmetric representation of multivariate distribution," *Statistica Sinica*, vol. 14, pp. 1259–1270, 2002.
- [4] A. Azzalini, "The skew-normal distribution and related multivariate families," *Scandinavian Journal of Statistics*, vol. 32, pp. 159–188, 2010.
- [5] R. B. Arellano-Valle and A. Azzalini, "Some properties of the unified skew-normal distribution," *Statistical Papers*, vol. 2021, pp. 1–27, 2021.
- [6] A. Azzalini, T. D. Cappello, and S. Kotz, "Log-skew-normal and log-skew-t distributions as models for family income data," *Journal of Income Distribution*, vol. 11, pp. 12–20, 2002.
- [7] S. Bartoletti and N. Loperfido, "Modelling air pollution data by the skew-normal distribution," *Stochastic Environmental Research and Risk Assessment*, vol. 24, no. 4, pp. 513–517, 2010.
- [8] C. Adcock, M. Eling, and N. Loperfido, "Skewed distributions in finance and actuarial science: a review," *The European Journal of Finance*, vol. 21, pp. 1–29, 2012.
- [9] A. Hossain and J. Beyene, "Application of skew-normal distribution for detecting differential expression to micro-RNA data," *Journal of Applied Statistics*, vol. 42, no. 3, pp. 477–491, 2015.
- [10] L. G. C. Cancela and J. O. Pires, "Applying the skew-normal distribution to model coherent MPI and to evaluate its impact on PAM signals," *Optical Fiber Technology*, vol. 56, pp. 1–5, 2020.
- [11] K. Kafadar, *Slash Distribution*, Wiley, New York, 1988.
- [12] W. H. Rogers and J. W. Tukey, "Understanding some long-tailed symmetrical distributions," *Statistica Neerlandica*, vol. 26, no. 3, pp. 211–226, 1972.
- [13] F. Mosteller and J. W. Tukey, *Data Analysis and Regression*, M. Addison-Wesley, Reading, MA, 1977.
- [14] J. Wang and M. G. Genton, "The multivariate skew-slash distribution," *Journal of Statistical Planning and Inference*, vol. 136, no. 1, pp. 209–220, 2006.
- [15] D. F. Andrew, P. J. Bickel, F. R. Hampel, P. J. Huber, W. H. Rogers, and J. W. Tukey, *Robust Estimates of Location: Survey and Advances*, Princeton University Press, Princeton, N.J., 1972.
- [16] A. M. Gross, "A Monte Carlo swindle for estimators of location," *Applied Statistics*, vol. 22, no. 3, pp. 347–353, 1973.
- [17] W. Gui, "A generalization of the slashed distribution via alpha skew normal distribution," *Statistical Methods and Applications*, vol. 23, no. 4, pp. 547–563, 2014.
- [18] D. Elal-Olivero, "Alpha-skew-normal distribution," *Proyecciones Journal of Mathematics*, vol. 29, pp. 224–240, 2010.
- [19] M. C. Jones and T. Higuchi, "On univariate slash distributions, continuous and discrete," *Annals of the Institute of Statistical Mathematics*, vol. 72, no. 3, pp. 645–657, 2020.
- [20] S. R. Durrans, "Distributions of fractional order statistics in hydrology," *Water Resources Research*, vol. 28, no. 6, pp. 1649–1655, 1992.

- [21] E. L. Lehmann, "The power of rank tests," *The Annals of Mathematical Statistics*, vol. 24, no. 1, pp. 23–43, 1953.
- [22] R. D. Gupta and R. C. Gupta, "Analyzing skewed data by power normal model," *Test*, vol. 17, no. 1, pp. 197–210, 2008.
- [23] B. C. Arnold and R. J. Beaver, "Skewed multivariate models related to hidden truncation and/or selective reporting," *Test*, vol. 11, pp. 7–54, 2002.
- [24] R. C. Gupta and R. D. Gupta, "Generalized skew normal model," *Test*, vol. 13, pp. 501–524, 2004.