AIMS *Mathematics*

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(10): 23849–23868.
DOI: 10.3934/math.20251060
Received: 06 September 2025
Revised: 06 October 2025
Accepted: 13 October 2025
Published: 21 October 2025

Research article

The heavy-tailed chi-square model: properties, estimation and application to wind speed data

Eliseo Martínez[1], Emilio Gómez-Déniz[2], Diego I. Gallardo[3], Osvaldo Venegas[4,*] and Héctor W. Gómez[5]

[1] Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile

[2] Department of Quantitative Methods in Economics and TIDES Institute, University of Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Spain

[3] Departamento de Estadística, Facultad de Ciencias, Universidad del Bío-Bío, Concepción 4081112, Chile

[4] Departamento de Ciencias Matemáticas y Físicas, Facultad de Ingeniería, Universidad Católica de Temuco, Temuco 4780000, Chile

[5] Departamento de Estadística y Ciencia de Datos, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile

* **Correspondence:** Email: ovenegas@uct.cl.

Abstract: In this article, we introduced an extension of the chi-square distribution by employing a slash-type methodology that enhanced the weight of the right tail, thereby producing a heavy-tailed distribution. We explored two different representations of the proposed distribution and examined several of its key properties, such as the mode, cumulative distribution function, reliability and hazard functions, moments, and the skewness and kurtosis coefficients. Additionally, we demonstrated that the classical chi-square distribution was a special case of our proposed model. Parameter estimation was carried out using both the method of moments and the maximum likelihood estimation, the latter via the expectation-maximization (EM) algorithm. A simulation study was conducted to evaluate the performance of parameter recovery. Finally, we applied the new distribution to a wind speed dataset, showing that it provided a good fit, particularly in the presence of extreme values.

Keywords: chi-square distribution; EM algorithm; heavy-tailed distribution; maximum likelihood; slash distribution
Mathematics Subject Classification: 62E15, 62E20, 62F10, 62P99

1. Introduction

The slash distribution is a heavy-tailed extension of the normal distribution, defined as the quotient of two independent random variables: a standard normal variable and a Beta-distributed variable with a single parameter. Formally, a random variable S follows a slash distribution if it can be expressed as:

$$S = X_1/X_2, \tag{1.1}$$

where $X_1 \sim N(0,1)$, $X_2 \sim Beta(\alpha, 1)$, and X_1 is independent of X_2. Its representation can be seen in Johnson et al. [1]. The principal feature of this distribution is its heavier tails and greater kurtosis compared to the normal distribution, making it more robust for modeling data with outliers. Early properties of this family were discussed by Rogers and Tukey [2] and Mosteller and Tukey [3]. Maximum likelihood (ML) estimators location and scale are discussed in Kafadar [4]. Wang and Genton [5] offer a multivariate version and a skew multivariate version of the slash distribution. Gómez et al. [6] and Gómez and Venegas [7] extend the slash distribution using the family of univariate and multivariate elliptical distributions. Genc [8] introduces a symmetric generalization of the slash distribution; Genc [9] extends the slash distribution with the beta-normal distribution, and Reyes and Iriarte [10] introduce a new family slash-type distribution using the Birnbaum–Saunders distribution. This methodology for increasing the weight of the tails has also been used in distributions with positive support, for example: Gui [11] in the Lindley distribution, Olmos et al. [12] in the generalised half-normal distribution, Iriarte et al. [13] in the Rayleigh distribution, and Castillo et al. [14] in the Fréchet distribution, among others.

A distribution that is very necessary in the present paper is the gamma distribution; the probability density function (pdf) of which is given by

$$g(t; a, b) = \frac{b^a}{\Gamma(a)} t^{a-1} e^{-bt}, \quad t > 0, \tag{1.2}$$

where $a, b > 0$, $\Gamma(\cdot)$ is the gamma function, and the corresponding cumulative distribution function (cdf) is denoted by:

$$G(z; a, b) = \int_0^z g(t; a, b) dt. \tag{1.3}$$

We can also say that a chi-square (CHI) distribution (see Johnson et al. [1]) has its pdf given by:

$$f_X(x; \sigma, \alpha) = \frac{x^{\frac{\alpha}{2}-1}}{(2\sigma)^{\alpha/2}\Gamma\left(\frac{\alpha}{2}\right)} \exp\left(-\frac{x}{2\sigma}\right), \quad x > 0, \tag{1.4}$$

where $\sigma > 0$ is the scale parameter and $\alpha > 0$ are degrees of freedom. It is denoted by $X \sim CHI(\sigma, \alpha)$. Some properties of this pdf are:

a) The cdf corresponding to (1.4) is

$$F_X(x; \alpha) = \frac{\gamma\left(\frac{\alpha}{2}, \frac{x}{2\sigma}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}, \quad x > 0, \tag{1.5}$$

where $\gamma(\cdot, \cdot)$ is the incomplete gamma function.

b) For $r = 1, 2, 3, \ldots$, the r-th moment of X is

$$\mathbb{E}(X^r) = \frac{(2\sigma)^r \Gamma\left(\frac{\alpha}{2} + r\right)}{\Gamma\left(\frac{\alpha}{2}\right)}. \tag{1.6}$$

c) The moment generating function ($M_X(t)$) is given by

$$M_X(t) = \mathbb{E}\left(\exp(tX)\right) = (1 - 2\sigma t)^{-\alpha/2}, \quad 2\sigma t < 1. \tag{1.7}$$

d) The coefficients of skewness ($\sqrt{\beta_1}$) and kurtosis (β_2), respectively, are

$$\sqrt{\beta_1} = \sqrt{\frac{8}{\alpha}}, \quad \beta_2 = 3 + \frac{12}{\alpha}. \tag{1.8}$$

We observe that when $\alpha = 1$, the CHI distribution has its mode at zero and a kurtosis coefficient of 15. However, we aim to develop a model with a more flexible mode, i.e., one that is not restricted to zero. This flexibility is achieved when $\alpha > 1$, though it comes at the cost of a decreasing kurtosis level. Based on this, the main motivation of this work is to increase the kurtosis and, consequently, to obtain a distribution with a heavier tail. From this perspective, the main object of this article is to introduce an extension of the CHI distribution making use of the slash-type methodology, which we call slash chi-square (SCHI). The SCHI distribution has a heavier right tail than the CHI distribution, allowing it to be used to model datasets with outliers.

The article is organized as follows. In Section 2 we give two representations of the SCHI distribution and some of its properties. In Section 3 we estimate the parameters of the SCHI distribution, estimating the moments and the ML using the expectation-maximization (EM) algorithm. In Section 4 we show an application to wind speed data, comparing it with two other distributions. Section 5 offers some conclusions.

2. SCHI distribution

In this section we introduce a representation of the SCHI distribution, with its pdf, cdf, and properties.

2.1. Stochastic representation

We can represent the SCHI distribution as:

$$Z = \frac{X}{Y}, \tag{2.1}$$

where $X \sim CHI(\sigma, \alpha)$, $Y \sim Beta(q, 1)$, and $\sigma, \alpha, q > 0$. We denote it by $Z \sim SCHI(\sigma, \alpha, q)$.

2.2. Density function

The following proposition shows the pdf of the SCHI distribution, generated using the representation given in (2.1).

Proposition 2.1. *Let $Z \sim SCHI(\sigma, \alpha, q)$. Then, the pdf of Z is given by*

$$f_Z(z; \sigma, \alpha, q) = \frac{q(2\sigma)^q}{\Gamma\left(\frac{\alpha}{2}\right)} z^{-(1+q)} \Gamma\left(\frac{\alpha}{2} + q\right) G\left(\frac{z}{2\sigma}, \frac{\alpha}{2} + q, 1\right), \quad z > 0, \tag{2.2}$$

where $\sigma > 0$ is the scale parameter and $\alpha > 0$ are degrees of freedom, $q > 0$ is a kurtosis parameter, and G is the cdf of the gamma distribution.

Proof. Using the representation given in (2.1) and the Jacobian method, we have:

$$\left.\begin{array}{l} Z = \frac{X}{Y} \\ V = Y \end{array}\right\} \Rightarrow \left.\begin{array}{l} X = ZV \\ Y = V \end{array}\right\} \Rightarrow J = \begin{vmatrix} \frac{\partial X}{\partial Z} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial Z} & \frac{\partial Y}{\partial V} \end{vmatrix} = \begin{vmatrix} v & z \\ 0 & 1 \end{vmatrix} = v,$$

$$\begin{aligned} f_{Z,V}(z, v) &= |J| f_{X,Y}(zv, v), \\ f_{Z,V}(z, v) &= v f_X(zv) f_Y(v)\,,\; z > 0\,,\; 0 < v < 1, \\ f_{Z,V}(z, v) &= \frac{q}{(2\sigma)^{\frac{\alpha}{2}} \Gamma\left(\frac{\alpha}{2}\right)} z^{\frac{\alpha}{2}-1} v^{\frac{\alpha}{2}+q-1} \exp\left(-\frac{zv}{2\sigma}\right),\; z > 0\,,\; 0 < v < 1, \end{aligned}$$

then marginalizing with respect to variable v,

$$f_Z(z; \sigma, \alpha, q) = \frac{q}{(2\sigma)^{\frac{\alpha}{2}} \Gamma(\frac{\alpha}{2})} z^{\frac{\alpha}{2}-1} \int_0^1 v^{\frac{\alpha}{2}+q-1} \exp\left(-\frac{zv}{2\sigma}\right) dv.$$

Making the following change to variable $u = \frac{zv}{2\sigma}$ and using (1.3), the distribution associated with Z is obtained. □

Figure 1 shows the shape of the SCHI distribution for some parameter values. We can observe that as the parameter q decreases, the right tail of the SCHI becomes heavier.

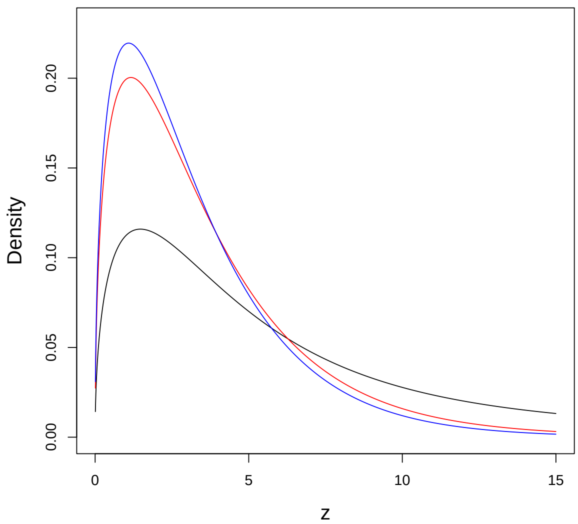

Figure 1. Densities SCHI(1,3,1) (black), SCHI(1,3,5) (red), and SCHI(1,3,10) (blue).

We perform a brief comparison illustrating that the tails of the SCHI distribution are heavier as the parameter q decreases. Table 1 shows $P(Z > z)$ for different values of z in the SCHI distribution. We also include the CHI distribution, enabling us to observe that as parameter q increases, the right tail of the SCHI distribution becomes closer to the right tail of the CHI distribution.

Table 1. Tails comparison.

Distribution	$P(Z > 8)$	$P(Z > 10)$	$P(Z > 12)$
SCHI(1,3,1)	0.362	0.296	0.249
SCHI(1,3,5)	0.096	0.050	0.027
SCHI(1,3,10)	0.067	0.031	0.014
CHI(1,3)	0.046	0.019	0.007

Proposition 2.2. *Let $Z \sim SCHI(\sigma, \alpha, q)$. Then, we obtain that:*

$$\lim_{z\to 0^+} f_Z(z;\sigma,\alpha,q) = \begin{cases} \infty & \text{if} \quad 0 < \alpha < 2, \\ \frac{q}{2\sigma(q+1)} & \text{if} \quad \alpha = 2, \\ 0 & \text{if} \quad \alpha > 2. \end{cases}$$

Proof. Calculating the limit as $z \to 0^+$ and applying L'Hopital's rule to the pdf given in (2.2), three results are obtained for each of the three cases of parameter α. □

Proposition 2.3. *Let $Z \sim SCHI(\sigma, \alpha, q)$. Then, the mode of Z is given as the solution of*

$$zg\left(\frac{z}{2\sigma}, \frac{\alpha}{2} + q, 1\right) = 2\sigma(1+q)G\left(\frac{z}{2\sigma}, \frac{\alpha}{2} + q, 1\right),$$

where g and G are given in (1.2) and (1.3), respectively.

Proof. Straightforward, studying the first derivative of (2.2) with respect to z. □

2.3. Properties

The following proposition shows the closed form of the cdf; it depends on G, which is the cdf of the gamma distribution given in (1.3).

Proposition 2.4. *Let $Z \sim SCHI(\sigma, \alpha, q)$. Then, the cdf of Z is given by*

$$F_Z(z;\sigma,\alpha,q) = G\left(\frac{z}{2\sigma}; \frac{\alpha}{2}, 1\right) - \left(\frac{2\sigma}{z}\right)^q \frac{\Gamma\left(\frac{\alpha}{2}+q\right)}{\Gamma\left(\frac{\alpha}{2}\right)} G\left(\frac{z}{2\sigma}; \frac{\alpha}{2} + q, 1\right), \tag{2.3}$$

where $\sigma, \alpha, q, z > 0$, and G is given in (1.3).

Proof. Using the definition of cdf and integration by parts, the result is obtained. □

2.3.1. Reliability analysis

The reliability function $r(t) = 1 - F(t)$ and the hazards function $h(t) = \frac{f(t)}{r(t)}$ of the SCHI distribution are given in the following corollary.

Corollary 2.1. *Let $T \sim SCHI(\sigma, \alpha, q)$. Then, the $r(t)$ and $h(t)$ of T are given by*

(1) $$r(t) = 1 - G\left(\tfrac{t}{2\sigma}; \tfrac{\alpha}{2}, 1\right) - \left(\frac{2\sigma}{t}\right)^q \frac{\Gamma\left(\frac{\alpha}{2} + q\right)}{\Gamma\left(\frac{\alpha}{2}\right)} G\left(\frac{t}{2\sigma}; \frac{\alpha}{2} + q, 1\right),$$

(2) $$h(t) = \frac{q(2\sigma)^q t^{-1} \Gamma\left(\frac{\alpha}{2} + q\right) G\left(\frac{t}{2\sigma}, \frac{\alpha}{2} + q, 1\right)}{t^q \Gamma\left(\frac{\alpha}{2}\right) - t^q \Gamma\left(\frac{\alpha}{2}\right) G\left(\frac{t}{2\sigma}; \frac{\alpha}{2}, 1\right) - (2\sigma)^q \Gamma\left(\frac{\alpha}{2} + q\right) G\left(\frac{t}{2\sigma}; \frac{\alpha}{2} + q, 1\right)},$$

where $\sigma, \alpha, q, t > 0$.

2.3.2. Heavy right tail of the SCHI distribution

A probability distribution with cdf $F(t)$ on the real numbers is said to have a heavy right tail (see Rolski et al. [15]) if

$$\limsup_{t \to \infty} \left(-\frac{\log r(t)}{t}\right) = 0.$$

The following result shows that the SCHI distribution has a heavy right tail.

Proposition 2.5. *The cdf of the random variable $T \sim SCHI(\sigma, \alpha, q)$ is a heavy-tailed distribution.*

Proof. Applying L'Hôpital's rule twice yields

$$\limsup_{t \to \infty} \left(-\frac{\log r(t)}{t}\right) = \limsup_{t \to \infty} \left(\frac{g\left(\frac{t}{2\sigma}; \frac{\sigma}{2} + q, 1\right)}{2\sigma G\left(\frac{t}{2\sigma}; \frac{\sigma}{2} + q, 1\right)} - \frac{1+q}{t}\right).$$

Calculating the righthand limit yields the result. □

Some recent studies on heavy-tailed distributions include those by Teamah et al. [16] and Afify et al. [17], among others.

Figure 2 shows the shape of the hazards function for different parameter values; we observe that its behavior is not monotonous.

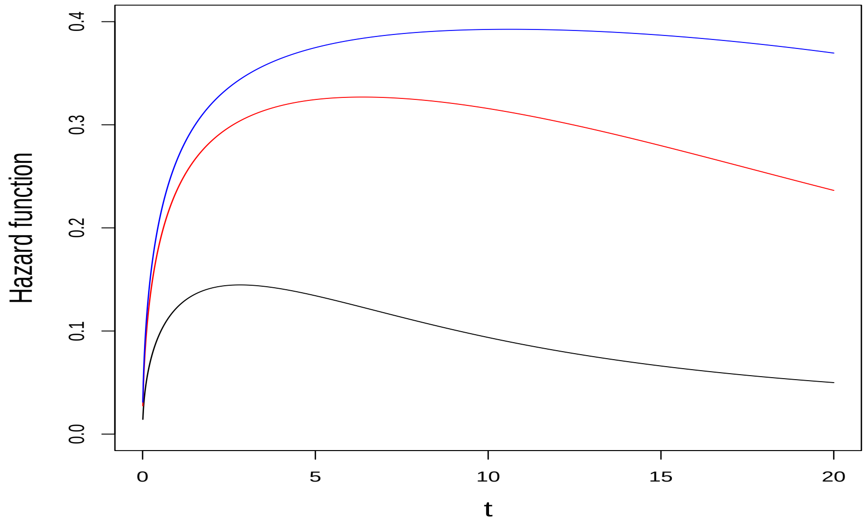

Figure 2. Hazard function for SCHI(1,3,1) (black), SCHI(1,3,5) (red), and SCHI(1,3,10) (blue).

The following Proposition shows that the SCHI distribution can also be represented as a scale mixture between the CHI and Beta distributions.

Proposition 2.6. *If* $Z|X = x \sim CHI(\sigma x^{-1}, \alpha)$ *and* $X \sim Beta(q, 1)$*, then* $Z \sim SCHI(\sigma, \alpha, q)$*.*

Proof. The marginal pdf of Z is given by

$$f_Z(z; \sigma, \alpha, q) = \int_0^1 f_{Z|X}(z) f_X(x)\, dx = \frac{q z^{\frac{\alpha}{2}-1}}{(2\sigma)^{\frac{\alpha}{2}} \Gamma\left(\frac{\alpha}{2}\right)} \int_0^1 x^{\frac{\alpha}{2}+q-1} \exp\left(-\frac{zx}{2\sigma}\right) dx.$$

Making the following change to variable $u = \frac{zx}{2\sigma}$ and using (1.3), the result is obtained. □

The following result shows that when parameter q tends to infinity in the SCHI distribution, the CHI distribution is obtained.

Proposition 2.7. *Let* $Z \sim SCHI(\sigma, \alpha, q)$*. If* $q \to \infty$*, then* Z *converges in law to a random variable* $X \sim CHI(\sigma, \alpha)$*.*

Proof. Using the representation $Z = \frac{X}{Y}$, we analyze the convergence of this quotient, where $X \sim CHI(\sigma, \alpha)$ and $Y \sim Beta(q, 1)$. In the Beta$(q, 1)$ distribution, we have that $Var[Y] = \frac{q}{(q+2)(q+1)^2}$. Then, applying Chebychev's inequality for Y, we have $\forall \epsilon > 0$

$$P\left[|Y - \mathbb{E}[Y]| > \epsilon\right] \leq \frac{Var(Y)}{\epsilon^2} = \frac{q}{(q+2)(q+1)^2 \epsilon^2}. \tag{2.4}$$

If $q \to \infty$, then the righthand side of (2.4) tends to zero, i.e., $Y - \mathbb{E}[Y]$ converges in probability to 0. Also, $\mathbb{E}[Y] = \frac{q}{1+q} \longrightarrow 1$, $q \to \infty$, then we have:

$$Y \xrightarrow{\mathcal{P}} 1, \ q \to \infty.$$

Since $Z \sim SCHI(\sigma, \alpha, q)$, applying Slutsky's lemma to $Z = \frac{X}{Y}$, we have:

$$Z \xrightarrow{\mathcal{L}} X \sim CHI(\sigma, \alpha), \qquad q \to \infty,$$

i.e., for increasing values of q, Z converges in law to a $CHI(\sigma, \alpha)$ distribution. □

2.4. Moments

In this subsection, the moments of the SCHI distribution and its asymmetry and kurtosis coefficients are given.

Proposition 2.8. *Let $Z \sim SCHI(\sigma, \alpha, q)$ with $\sigma, \alpha, q > 0$. For $r > 0$, $\mathbb{E}[Z^r]$ exists if and only if, $q > r$, and in this case,*

$$\mu_r = \mathbb{E}[Z^r] = \frac{q(2\sigma)^r \Gamma\left(\frac{\alpha}{2} + r\right)}{(q-r)\Gamma\left(\frac{\alpha}{2}\right)}. \tag{2.5}$$

Proof. Using the representation given in Proposition 2.6, we have:

$$\begin{aligned}
\mu_r = \mathbb{E}[Z^r] &= \mathbb{E}\left[\mathbb{E}\left(Z^r|X\right)\right] = \mathbb{E}\left[\frac{\left(2(\sigma X^{-1})\right)^r \Gamma\left(\frac{\alpha}{2} + r\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\right] \\
&= \frac{(2\sigma)^r \Gamma\left(\frac{\alpha}{2} + r\right)}{\Gamma\left(\frac{\alpha}{2}\right)} \mathbb{E}\left[X^{-r}\right] = \frac{(2\sigma)^r \Gamma\left(\frac{\alpha}{2} + r\right)}{\Gamma\left(\frac{\alpha}{2}\right)} \int_0^1 q x^{q-r-1} dx.
\end{aligned}$$

Solving the integral gives the result. □

From Proposition 2.8, the explicit expression of the noncentral moments, $\mu_r = \mathbb{E}[Z^r]$, for $r = 1, 2, 3, 4$ and the variance of $Z \sim SCHI(\sigma, \alpha, q)$, $Var(Z)$, follow.

Corollary 2.2. *Let $Z \sim SCHI(\sigma, \alpha, q)$ with $\sigma, \alpha, q > 0$.*

$$\mu_1 = \sigma\alpha\kappa_1, \quad q > 1, \qquad \mu_2 = \sigma^2\alpha(\alpha + 2)\kappa_2, \; q > 2,$$

$$\mu_3 = \sigma^3\alpha(\alpha + 2)(\alpha + 4)\kappa_3, \; q > 3, \quad \mu_4 = \sigma^4\alpha(\alpha + 2)(\alpha + 4)(\alpha + 6)\kappa_4, \; q > 4,$$

$$Var(Z) = \sigma^2\alpha\left[(\alpha + 2)\kappa_2 - \alpha\kappa_1^2\right], \; q > 2,$$

where $\kappa_i = \frac{q}{q-i}$, $q > i$.

Remark 1. *We observe that as $q \to \infty$, $Var(Z) \to 2\sigma^2\alpha$, which is the variance of the $CHI(\sigma, \alpha)$ distribution. Furthermore, as $q \to \infty$, $\kappa_i \to 1$ $\forall i$.*

Table 2 shows the expected values and the mode for some parameter values.

Table 2. Expected values and mode.

σ	α	q	Expected	Mode
1	3	2	6	1.3231
		3	9/2	1.2450
		4	4	1.1972
		5	15/4	1.1649
2	3	2	12	2.6462
		3	9	2.4901
		4	8	2.3945
		5	15/2	2.3299
3	4	2	24	7.8859
		3	18	7.4439
		4	16	7.1683
		5	15	6.9804
10	3	2	60	13.2310
		3	45	12.4507
		4	40	11.9726
		5	75/2	11.6496

The next corollary gives us the asymmetry coefficient, $\sqrt{\beta_1}$, of a $SCHI(\sigma, \alpha, q)$ distribution.

Corollary 2.3. *Let $Z \sim SCHI(\sigma, \alpha, q)$ with $q > 3$. Then, the asymmetry coefficient of Z is*

$$\sqrt{\beta_1} \quad = \quad \frac{(\alpha+2)(\alpha+4)\kappa_3 - 3\alpha(\alpha+2)\kappa_1\kappa_2 + 2\alpha^2\kappa_1^3}{\sqrt{\alpha}\left[(\alpha+2)\kappa_2 - \alpha\kappa_1^2\right]^{3/2}}.$$

Proof. Recall that

$$\sqrt{\beta_1} = \frac{\mathbb{E}[(Z-\mathbb{E}(Z))^3]}{(Var(Z))^{3/2}} = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2-\mu_1^2)^{3/2}},$$

where μ_1, μ_2 and μ_3 were given in Corollary 2.2. □

Next Corollary gives us the kurtosis coefficient, β_2, of a $SCHI(\sigma, \alpha, q)$ distribution.

Corollary 2.4. *Let $Z \sim SCHI(\sigma, \alpha, q)$ with $q > 4$. Then the kurtosis coefficient of Z is*

$$\beta_2 \quad = \quad \frac{(\alpha+2)(\alpha+4)(\alpha+6)\kappa_4 - 4\alpha(\alpha+2)(\alpha+4)\kappa_1\kappa_3 + 6\alpha^2(\alpha+2)\kappa_1^2\kappa_2 - 3\alpha^3\kappa_1^4}{\alpha\left[(\alpha+2)\kappa_2 - \alpha\kappa_1^2\right]^2}.$$

Proof. Recall that

$$\beta_2 = \frac{\mathbb{E}[(Z-\mathbb{E}(Z))^4]}{(Var(Z))^2} = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2-\mu_1^2)^2},$$

where μ_1, μ_2, μ_3, and μ_4 were given in Corollary 2.2. □

Remark 2. *It can be verified that as $q \to \infty$, the asymmetry and kurtosis coefficients converge to $\sqrt{\frac{8}{\alpha}}$ and $\frac{12}{\alpha} + 3$, respectively, which coincide with the corresponding coefficients for the CHI(σ, α) distribution.*

Table 3 shows that the values of the asymmetry and kurtosis coefficients depend on parameters α and q, and that as q diminishes, these two coefficients increase. At the same time, as q increases, the coefficients of asymmetry and kurtosis become those of the CHI(σ, α) distribution (see Proposition 2.7).

Table 3. Asymmetry and kurtosis values for the SCHI distribution for various shape parameters.

α	q	$\sqrt{\beta_1}$	β_2
1	5	3.555	31.909
3		2.397	19.256
1	6	3.224	21.750
3		2.047	11.611
1	7	3.079	18.766
3		1.893	9.480
1	10	2.926	16.273
3		1.732	7.791
1	100	2.829	15.008
3		1.634	7.004
1	∞	2.828	15
3		1.633	7

3. Inference

In this section, we show the moments estimators and ML estimators using the EM algorithm and a simulation study by which to observe the behavior of the ML estimators.

3.1. Moments estimators

Let $Z_1, ..., Z_n$ be a random sample with n terms distributed by $Z \sim SCHI(\sigma, \alpha, q)$ and $\overline{Z^r} = \frac{1}{n}\sum_{i=1}^{n} Z_i^r$ for the sample moments. Then, the moment estimators $(\hat{\sigma}_M, \hat{\alpha}_M, \hat{q}_M)$ of (σ, α, q) for $q > 3$ are obtained by solving the following system of nonlinear equations.

$$\begin{aligned} \overline{Z}(q-1) &= \sigma\alpha q, \\ \overline{Z^2}(q-2) &= \sigma^2\alpha(\alpha+2)q, \\ \overline{Z^3}(q-3) &= \sigma^3\alpha(\alpha+2)(\alpha+4)q. \end{aligned}$$

The solutions to these equations can be obtained using the nleqslv function in the R software version 4.0.5, [18].

3.2. ML Estimators

Let $Z_1, ..., Z_n$ be a random sample with n terms distributed $Z \sim SCHI(\sigma, \alpha, q)$, then the log-likelihood function for $\boldsymbol{\theta} = (\sigma, \alpha, q)$ can be expressed as:

$$\ell(\boldsymbol{\theta}) = c(\boldsymbol{\theta}) - (1+q)\sum_{i=1}^{n}\log(z_i) + \sum_{i=1}^{n}\log\left(G\left(\frac{z_i}{2\sigma}, \frac{\alpha}{2}+q, 1\right)\right), \tag{3.1}$$

where $c(\boldsymbol{\theta}) = n\log(q) + nq\log(2\sigma) - n\log\left(\Gamma\left(\frac{\alpha}{2}\right)\right) + n\log\left(\Gamma\left(\frac{\alpha}{2}+q\right)\right)$. When the log-likelihood function given in (3.1) is partially derived for σ, α, and q, bringing it to equal zero, three equations are obtained that are not easily resolved. One option for obtaining the ML estimators is to maximize equation (3.1) using the optim function of the R software version 4.0.5, [18]. However, to obtain a more robust estimation procedure, in the next subsection we will explore the use of the EM algorithm for this particular problem.

3.3. EM algorithm

The EM algorithm (see Dempster et al. [19]) is a widely-used tool for calculating ML estimators in scenarios with unobserved or latent data. Based on the stochastic representation in (2.1), we obtain that the complete log-likelihood function for $\boldsymbol{\theta}$ is given by

$$\begin{aligned}\ell_c(\boldsymbol{\theta}) = n\left[\log(q) - \frac{\alpha}{2}\log(2\sigma) - \log\Gamma\left(\frac{\alpha}{2}\right)\right] + \left(\frac{\alpha}{2}-1\right)\sum_{i=1}^{n}\log z_i \\ + \left(\frac{\alpha}{2}+q-1\right)\sum_{i=1}^{n}\log v_i - \frac{1}{2\sigma}\sum_{i=1}^{n} z_i v_i.\end{aligned}$$

Defining $\widehat{v_i} = \mathbb{E}(V_i \mid Z_i = z_i)$ and $\widehat{\log v}_i = \mathbb{E}(\log V_i \mid Z_i = z_i)$, we obtain that

$$\begin{aligned}Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}) = n\left[\log(q) - \frac{\alpha}{2}\log(2\sigma) - \log\Gamma\left(\frac{\alpha}{2}\right)\right] + \left(\frac{\alpha}{2}-1\right)\sum_{i=1}^{n}\log z_i \\ + \left(\frac{\alpha}{2}+q-1\right)\sum_{i=1}^{n}\widehat{\log v}_i - \frac{1}{2\sigma}\sum_{i=1}^{n} z_i\widehat{v_i}.\end{aligned}$$

On the other hand, note that the kernel of the distribution of $V_i \mid Z_i = z_i$ is given by

$$f(v_i \mid Z_i = z_i) \propto v_i^{\frac{\alpha}{2}+q-1}\exp\left(-\frac{z_i v_i}{2\sigma}\right), \quad 0 < v_i < 1,$$

i.e., $V_i \mid Z_i = z_i \sim TG_{(0,1)}(\alpha/2 + q, z_i/(2\sigma))$. Therefore, it is immediate that

$$\mathbb{E}(V_i \mid Z_i = z_i) = \frac{\sigma(\alpha+2q)G\left(\frac{z_i}{2\sigma}, \frac{\alpha}{2}+q+1, 1\right)}{z_i G\left(\frac{z_i}{2\sigma}, \frac{\alpha}{2}+q, 1\right)}, \quad \text{and} \tag{3.2}$$

$$\mathbb{E}(\log V_i \mid Z_i = z_i) = \frac{\left(\frac{z_i}{2\sigma}\right)^{\frac{\alpha}{2}+q}}{\Gamma\left(\frac{\alpha}{2}+q\right)G\left(\frac{z_i}{2\sigma}, \frac{\alpha}{2}+q, 1\right)}\int_0^1 \log(x)x^{\frac{\alpha}{2}+q-1}\exp\left\{-\frac{z_i x}{2\sigma}\right\}dx. \tag{3.3}$$

Details for these results are presented in the Appendix. We also argue the convergence of $\mathbb{E}(\log V_i \mid Z_i = z_i)$. Therefore, the EM algorithm is summarized as follows.

- E-step: given $\widehat{\alpha}^{(k-1)}$, $\widehat{q}^{(k-1)}$, and $\widehat{\sigma}^{(k-1)}$, the values for the parameters in the $(k-1)$-th step, update $\widehat{v_i}$ and $\widehat{\log v_i}$, for $i = 1, \ldots, n$, using Eqs (3.2) and (3.3), respectively.
- M1-step: update $\widehat{\alpha}^{(k)}$, the value of α as the solution for the nonlinear equation

$$\psi\left(\frac{\alpha}{2}\right) = \sum_{i=1}^{n} \log z_i + \sum_{i=1}^{n} \widehat{\log v_i} - n \log(2\widehat{\sigma}^{(k-1)}), \tag{3.4}$$

  where $\psi(\cdot)$ denotes the digamma function. Note that $\psi(u)$ is an increasing function in u such that $\lim_{u\to 0} \psi(u) = -\infty$ and $\lim_{u\to +\infty} \psi(u) = +\infty$. Therefore, Eq (3.4) has a unique solution.
- M2-step: update the value for q as follows

$$\widehat{q}^{(k)} = -\left(\frac{1}{n} \sum_{i=1}^{n} \widehat{\log v_i}\right)^{-1}.$$

- M3-step: update the value for σ as follows

$$\widehat{\sigma}^{(k)} = \frac{1}{n\widehat{\alpha}^{(k)}} \sum_{i=1}^{n} z_i \widehat{v_i}.$$

3.4. Simulation study

In this section, we present a simulation study in order to assess the performance of the ML estimators for the SCHI distribution. For this, we consider two values for σ: 1 and 5; two values for α: 2 and 10; three values for q: 1, 2, and 3 and; three values for the sample size: 100, 200, and 500. For each combination of σ, α, q, and n, we draw 5,000 samples of the corresponding SCHI model and compute the ML estimators based on the EM algorithm and their standard errors. Then, we compute the estimated bias (bias), mean of the estimated standard error (SE), root of the estimated mean squared error (RMSE), and the 95% coverage probability (CP). Table 4 summarizes these results. In general terms, note that the bias and the RMSE terms are reduced when the sample size is increased. On the other hand, the SE and RMSE terms are closer when the sample size is increased, suggesting that the variance of the ML estimators is well estimated even in finite samples. Finally, the CP terms are closer to the nominal value, suggesting that the normal distribution is reasonable as an approximation for the distribution of the ML estimators, again, even in finite samples.

Table 4. Estimated bias, SE, and RMSE for ML estimators in finite samples for the SCHI model.

				n = 100				n = 200				n = 500			
σ	α	q	estimator	bias	SE	RMSE	CP	bias	SE	RMSE	CP	bias	SE	RMSE	CP
1	2.00	1	$\widehat{\sigma}$	0.075	0.404	0.463	0.919	0.027	0.266	0.278	0.934	0.009	0.164	0.166	0.939
			$\widehat{\alpha}$	0.098	0.390	0.445	0.952	0.046	0.261	0.270	0.955	0.018	0.161	0.165	0.953
			$\widehat{q}$	0.086	0.282	0.391	0.957	0.031	0.151	0.165	0.951	0.011	0.091	0.094	0.954
		2	$\widehat{\sigma}$	0.031	0.378	0.343	0.929	0.036	0.252	0.255	0.943	0.016	0.149	0.153	0.953
			$\widehat{\alpha}$	0.087	0.340	0.372	0.961	0.035	0.228	0.241	0.946	0.013	0.140	0.141	0.955
			$\widehat{q}$	0.400	1.892	1.343	0.936	0.271	0.869	0.911	0.945	0.096	0.359	0.439	0.956
		3	$\widehat{\sigma}$	-0.049	0.366	0.280	0.918	-0.001	0.259	0.213	0.944	0.014	0.156	0.151	0.956
			$\widehat{\alpha}$	0.119	0.337	0.355	0.972	0.044	0.221	0.220	0.965	0.014	0.135	0.135	0.953
			$\widehat{q}$	0.282	3.748	1.680	0.891	0.413	2.504	1.532	0.923	0.315	1.190	1.122	0.944
	10	1	$\widehat{\sigma}$	0.014	0.415	0.444	0.892	0.004	0.288	0.293	0.927	0.001	0.181	0.184	0.939
			$\widehat{\alpha}$	1.951	4.325	12.469	0.948	0.737	2.516	2.923	0.951	0.265	1.470	1.558	0.953
			$\widehat{q}$	0.027	0.149	0.162	0.952	0.016	0.102	0.107	0.953	0.004	0.063	0.064	0.950
		2	$\widehat{\sigma}$	0.025	0.337	0.355	0.928	0.017	0.232	0.240	0.940	0.004	0.144	0.143	0.945
			$\widehat{\alpha}$	0.848	2.882	3.415	0.948	0.336	1.850	1.973	0.950	0.145	1.127	1.143	0.956
			$\widehat{q}$	0.154	0.511	0.615	0.953	0.064	0.299	0.330	0.959	0.023	0.178	0.183	0.955
		3	$\widehat{\sigma}$	0.015	0.312	0.305	0.930	0.016	0.214	0.222	0.943	0.007	0.132	0.136	0.941
			$\widehat{\alpha}$	0.702	2.530	2.873	0.961	0.278	1.647	1.764	0.947	0.104	1.002	1.029	0.948
			$\widehat{q}$	0.363	1.337	1.242	0.945	0.207	0.734	0.847	0.950	0.078	0.379	0.423	0.959
5	2.00	1	$\widehat{\sigma}$	0.287	1.981	2.187	0.915	0.141	1.330	1.398	0.933	0.049	0.823	0.830	0.950
			$\widehat{\alpha}$	0.103	0.392	0.459	0.953	0.044	0.261	0.273	0.946	0.018	0.161	0.163	0.951
			$\widehat{q}$	0.075	0.267	0.357	0.952	0.033	0.152	0.167	0.955	0.011	0.091	0.092	0.955
		2	$\widehat{\sigma}$	0.205	1.921	1.720	0.933	0.162	1.256	1.271	0.942	0.094	0.747	0.786	0.947
			$\widehat{\alpha}$	0.074	0.337	0.348	0.963	0.034	0.227	0.236	0.951	0.010	0.140	0.143	0.948
			$\widehat{q}$	0.450	2.054	1.419	0.942	0.265	0.899	0.943	0.944	0.094	0.349	0.398	0.953
		3	$\widehat{\sigma}$	-0.255	1.817	1.368	0.912	-0.013	1.294	1.076	0.944	0.080	0.786	0.753	0.956
			$\widehat{\alpha}$	0.119	0.336	0.352	0.971	0.045	0.221	0.221	0.965	0.014	0.135	0.136	0.950
			$\widehat{q}$	0.295	3.731	1.671	0.900	0.404	2.488	1.530	0.922	0.327	1.207	1.143	0.942
	10	1	$\widehat{\sigma}$	0.134	2.101	2.301	0.895	0.017	1.440	1.518	0.915	0.009	0.903	0.918	0.936
			$\widehat{\alpha}$	1.716	4.253	5.986	0.942	0.786	2.543	2.983	0.950	0.263	1.466	1.544	0.953
			$\widehat{q}$	0.030	0.150	0.161	0.958	0.012	0.102	0.104	0.952	0.005	0.063	0.064	0.947
		2	$\widehat{\sigma}$	0.131	1.680	1.800	0.920	0.057	1.150	1.185	0.932	0.028	0.720	0.740	0.940
			$\widehat{\alpha}$	0.875	2.890	3.475	0.945	0.400	1.867	2.054	0.952	0.132	1.126	1.160	0.943
			$\widehat{q}$	0.151	0.512	0.632	0.953	0.064	0.298	0.332	0.952	0.021	0.177	0.185	0.947
		3	$\widehat{\sigma}$	0.033	1.556	1.489	0.927	0.075	1.071	1.109	0.940	0.043	0.659	0.665	0.948
			$\widehat{\alpha}$	0.733	2.532	2.768	0.967	0.303	1.651	1.781	0.946	0.088	0.999	1.024	0.947
			$\widehat{q}$	0.342	1.340	1.240	0.946	0.215	0.734	0.840	0.952	0.075	0.375	0.393	0.964

4. Application

We apply the distribution to analyze 246 observations of monthly maximum wind speed (mph) in West Palm Beach, Florida (USA) for the months January 1984 to December 2005. The data is:

33,40,46,41,31,37,41,56,45,31,40,35,33,43,36,36,48,45,51,44,38,36,40,32,51,37,43,
33,35,44,41,41,33,45,38,43,62,45,51,39,35,58,48,35,43,49,43,39,39,40,39,45,48,43,
45,36,40,36,47,35,40,39,44,37,36,38,37,41,38,36,36,48,37,40,38,37,37,38,49,66,39,
45,37,35,39,52,66,51,39,64,59,36,36,36,41,41,39,45,40,37,33,66,38,59,38,41,45,35,
43,39,74,63,37,45,52,43,44,52,36,43,46,40,43,29,39,53,32,41,52,31,46,48,49,41,32,
37,29,43,40,47,45,38,28,30,40,36,37,38,37,33,30,34,38,45,40,31,39,31,31,38,32,34,

45,39,31,29,39,36,34,55,38,37,36,34,44,32,54,30,39,30,41,33,36,39,33,33,30,40,44,
61,34,26,38,26,34,36,28,36,43,35,43,37,40,35,36,28,41,30,31,48,43,43,49,36,38,30,
33,35,36,45,29,43,33,39,38,29,38,41,31,35,40,33,51,33,40,45,32,29,35,37,35,30,32,
39,32,39,38,39,83,30,33,39,33,36,39,44,31,43,44,43,41,101,37,33.

Figure 3 shows a scatterplot and a boxplot of the data. It may be noted that the majority of the observations are of a wind speed of around 40 mph, but there is one outlier (100 mph). Table 5 shows descriptive statistics for the wind speed data where b_1 and b_2 are sample asymmetry and kurtosis coefficients, respectively, with values of 2.344 and 13.030. In general, these values are higher than those of the CHI distribution, suggesting the need to use a distribution with a heavier right tail, as the kurtosis coefficient of the sample is very high, indicating the presence of outliers.

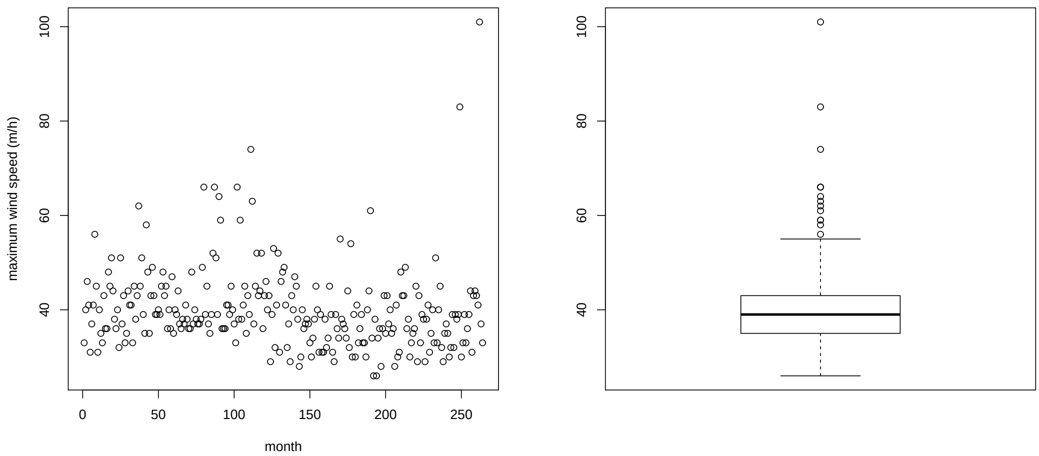

Figure 3. Plots of the scatterplot and boxplot of the maximum wind speed(mph).

Table 5. Summary statistics for the wind speed data.

n	$\bar{z}$	s^2	b_1	b_2
264	40.114	80.390	2.344	13.030

One of the distributions used to model wind speed data is the Gumbel (G) distribution (see Gumbel [20]). As this data presents a high sample kurtosis, Gómez et al. [21] extends this G distribution using the slash methodology, obtaining the slash-Gumbel (SG) distribution. We say that $Z \sim SG(\mu, \sigma, q)$, if the pdf of Z is given by

$$f_Z(z; \mu, \sigma, q) = \begin{cases} \dfrac{q\sigma^q}{(z-\mu)^{q+1}}\kappa(q; m(z,\mu,\sigma), 1) & \text{if} \quad z > \mu \\ \dfrac{q}{e\sigma(q+1)} & \text{if} \quad z = \mu \\ \dfrac{-q\sigma^q}{(z-\mu)^{q+1}}\kappa(q; 1, m(z,\mu,\sigma)) & \text{if} \quad z < \mu \end{cases} \tag{4.1}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$, $q > 0$, $z \in \mathbb{R}$, $\kappa(j; a, b) = (-1)^j \int_a^b (\log(v))^j \exp(-v)dv$, $m(z, \mu, \sigma) = \exp\left\{-\left(\frac{z-\mu}{\sigma}\right)\right\}$, $j = 1, 2, 3, ...,$ and $0 \leq a < b \leq \infty$.

In Nadarajah's [22] work, an extension of the G distribution was introduced, known as the *Exponentiated Gumbel (EG)* distribution. The pdf of the EG distribution is given by:

$$f_X(x; \mu, \sigma, \alpha) = \frac{\alpha}{\sigma} \exp\left\{-\frac{x-\mu}{\sigma} - \exp\left(-\frac{x-\mu}{\sigma}\right)\right\} \left[1 - \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}\right]^{\alpha-1},$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, and $\alpha > 0$ is a shape parameter. We denote this as $X \sim EG(\mu, \sigma, \alpha)$.

Note that when $\alpha = 1$, the standard G distribution is recovered as a particular case. In this work, we also consider the EG distribution for data fitting purposes. Using results from Section 3.1, moment estimators were computed, leading to the following values: $\hat{\sigma}_M = 0.278$, $\hat{\alpha}_M = 122.335$, and $\hat{q}_M = 6.588$, which were used as initial estimates for the ML approach.

For this data we first obtained ML estimates for the parameters of the SCHI distribution SCHI, and then compared them with other models such as the CHI, SG, and EG using the Akaike information criterion (AIC) (see Akaike [23]) and the Bayesian information criterion (BIC) (see Schwarz [24]), as shown in Table 6.

Table 6. ML estimates for the wind speed data with corresponding standard errors (in parentheses), AIC, and BIC values.

Estimates	CHI (SE)	SG (SE)	EG(SE)	SCHI (SE)
$\hat{\mu}$	–	36.346 (0.389)	34.692 (1.078)	–
$\hat{\sigma}$	0.819 (0.072)	4.577 (0.492)	4.889 (0.724)	0.246 (0.031)
$\hat{\alpha}$	49.000 (4.236)	–	0.694 (0.160)	138.378 (16.118)
$\hat{q}$	–	4.122 (1.380)	–	6.667 (0.698)
AIC	1850.221	1800.067	1801.324	**1794.629**
BIC	1857.373	1810.795	1812.052	**1805.356**

Figure 4 displays the quantile-quantile (QQ) plot for the best model. From them, we calculated the quantile residuals (QRs), which, if the model is appropriate for the data, should behave like a sample from the standard normal distribution (see Dunn and Smyth [25]). This can be validated by using traditional tests for normality, such as the Anderson-Darling (AD), Cramer-von Mises (CVM), and Shapiro-Wilks (SW) tests. The p-values for the normality tests (AD, CVM, and WS) of the QRs for the best model are shown below the QQ-plot in Figure 4. The QRs for the SCHI model seem to behave like normal standard deviations. This supports the claim made earlier that the SCHI model provides the best fit to the data.

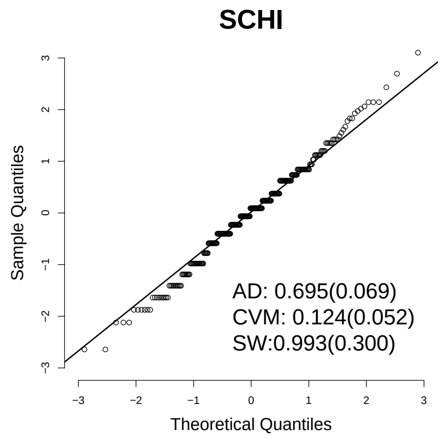

Figure 4. QQ-plot for QRs in the wind speed data for the SCHI model. The statistics AD, CVM, and SW (with the corresponding p-values in parentheses) are presented.

5. Conclusions

In this work we present the SCHI distribution, which is an extension of the CHI distribution in which we use the slash methodology to increase the weight of the right tail; it can thus be used for modeling positive datasets with positive asymmetry and high kurtosis. We report some properties of the SCHI distribution, and its parameters are estimated by the moments method and the ML using the EM algorithm. Below we highlight some of the most important characteristics of the SCHI distribution:

- The SCHI distribution has two stochastic representations, given in Eq (2.1) and Proposition 2.6. The mixed scale representation was important for implementing the EM algorithm.
- We obtained the expressions of the pdf, cdf, and hazard function in their closed form, and represented by the cdf of the gamma distribution.
- The asymmetry and kurtosis coefficients show that the SCHI distribution is more flexible than the CHI distribution. In Table 2, the right tail of the SCHI distribution becomes heavier as the parameter q diminishes.
- Using the EM algorithm, we obtained the ML estimators ML for the parameters of the SCHI distribution. A simulation study shows that as the sample size increases, the ML estimators remain consistent and stable.
- When applied to a set of wind speed data, the SCHI distribution is observed to provide a better fit than the CHI, SG and EG distributions. This is evidenced by lower AIC and BIC values, as well as higher p-values in the AD, CVM, and SW tests.
- Future directions for this model are related to reparametrize the model in terms of the mean, explore its use in survival analysis, and propose a methodology to bias reduction in the estimators, to name a few.

Author contributions

Conceptualization: E. Martínez and E. Gómez-Déniz; methodology: H.W. Gómez and E. Gómez-Déniz; software: D.I. Gallardo; validation: D.I. Gallardo, E. Gómez-Déniz, and O. Venegas; formal analysis: H.W. Gómez and O. Venegas; investigation: E. Martínez and D.I. Gallardo; writing—original draft preparation: E. Martínez and H.W. Gómez; writing—review and editing: O. Venegas and D.I. Gallardo. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

EGD was partially funded by grant PID2021-127989OB-I00 (Ministerio de Economía y Competitividad, Spain) and by grant TUR-RETOS2022-075 (Ministerio de Industria, Comercio y Turismo).

Conflict of interest

Prof. Emilio Gómez-Déniz is the Guest Editor of special issue "Advances in Probability Distributions and Social Science Statistics" for AIMS Mathematics. Prof. Emilio Gómez-Déniz was not involved in the editorial review and the decision to publish this article.

The authors declare no conflicts of interest.

References

1. N. L. Johnson, S. Kotz, N. Balakrishnan, *Continuous univariate distributions*, 2 Eds., New York: Wiley, 1994.
2. W. H. Rogers, J. W. Tukey, Understanding some long-tailed symmetrical distributions, *Stat. Neerl.*, **26** (1972), 211–226. https://doi.org/10.1111/j.1467-9574.1972.tb00191.x
3. F. Mosteller, J. W. Tukey, *Data analysis and regression*, Pearson, 1977.
4. K. Kafadar, A biweight approach to the one-sample problem, *J. Am. Stat. Assoc.*, **77** (1982), 416–424. https://doi.org/10.2307/2287262
5. J. Wang, M. G. Genton, The multivariate skew-slash distribution, *J. Stat. Plan. Infer.*, **136** (2006), 209–220. https://doi.org/10.1016/j.jspi.2004.06.023
6. H. W. Gómez, F. A. Quintana, F. J. Torres, A new family of slash distributions with elliptical contours, *Statist. Probab. Lett.*, **77** (2007), 717–725. https://doi.org/10.1016/j.spl.2006.11.006
7. H. W. Gómez, O. Venegas, Erratum to: A new family of slash-distributions with elliptical contours [Statist. Probab. Lett. 77 (2007) 717–725], *Statist. Probab. Lett.*, **78** (2008), 2273–2274. https://doi.org/10.1016/j.spl.2008.02.015
8. A. I. Genc, A generalization of the univariate slash by a scale–mixture exponential power distribution, *Commun. Stat. Simul. Comput.*, **36** (2007), 937–947. https://doi.org/10.1080/03610910701539161
9. A. I. Genc, A skew extension of the slash distribution via beta-normal distribution, *Stat. Pap.*, **54** (2013), 427–442. https://doi.org/10.1007/s00362-012-0439-0

10. J. Reyes, Y. A. Iriarte, A new family of modified slash distributions with applications, *Mathematics*, **11** (2023), 3018. https://doi.org/10.3390/math11133018

11. W. Gui, Statistical properties and applications of the Lindley slash distribution, *J. Appl. Stat. Sci.*, **20** (2012), 283–298.

12. N. M. Olmos, H. Varela, H. Bolfarine, H. W. Gómez, An extension of the generalized half-normal distribution, *Stat. Pap.*, **55** (2014), 967–981. https://doi.org/10.1007/s00362-013-0546-6

13. Y. A. Iriarte, F. Vilca, H. Varela, H. W. Gómez, Slashed generalized Rayleigh distribution, *Commun. Stat. Theory Methods*, **46** (2017), 4686–4699. https://doi.org/10.1080/03610926.2015.1066811

14. J. S. Castillo, M. A. Rojas, J. Reyes, A more flexible extension of the Fréchet distribution based on the incomplete gamma function and applications, *Symmetry*, **15** (2023), 1608. https://doi.org/10.3390/sym15081608

15. T. Rolski, H. Schmidli, V. Schmidt, J. Teugel, *Stochastic processes for insurance and finance*, New York: John Wiley & Sons, 1999. https://doi.org/10.1002/9780470317044

16. A. A. M. Teamah, A. A. Elbanna, A. M. Gemeay, Heavy-tailed log-logistic distribution: properties, risk measures and applications, *Stat. Optim. Inf. Comput.*, **9** (2021), 910–941. https://doi.org/10.19139/soic-2310-5070-1220

17. A. Z. Afify, A. M. Gemeay, N. A. Ibrahim, The heavy-tailed exponential distribution: risk measures, estimation, and application to actuarial data, *Mathematics*, **8** (2020), 1–28. https://doi.org/10.3390/math8081276

18. The R Core Team, *R: A Language and Environment for Statistical Computing*, 2025. Available from: `https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf`.

19. A. P. Dempster, N. M. Laird, D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. B: Stat. Methodol.*, **39** (1977), 1–38.

20. E. J. Gumbel, *Statistics of extremes*, New York: Columbia University Press, 1958.

21. Y. M. Gómez, H. Bolfarine, H. W. Gómez, Gumbel distribution with heavy tails and applications to environmental data, *Math. Comput. Simul.*, **157** (2019), 115–129. https://doi.org/10.1016/j.matcom.2018.10.003

22. S. Nadarajah, The exponentiated Gumbel distribution with climate application, *Environmetrics*, **17** (2006), 13–23. https://doi.org/10.1002/env.739

23. H. Akaike, A new look at the statistical model identification, *IEEE Trans. Autom. Control*, **19** (1974), 716–723. https://doi.org/10.1109/TAC.1974.1100705

24. G. Schwarz, Estimating the dimension of a model, *Ann. Statist.*, **6** (1978), 461–464. https://doi.org/10.1214/aos/1176344136

25. P. K. Dunn, G. K. Smyth, Randomized quantile residuals, *J. Comput. Graph. Stat.*, **5** (1996), 236–244. https://doi.org/10.2307/1390802

Appendix

Details about the truncated gamma model

The truncated gamma distribution in the interval (0, 1) has pdf given by

$$f(x; a, b) = \frac{b^a}{\Gamma(a)G(b, a)} x^{a-1} e^{-bx}, \quad x \in (0, 1), a, b > 0.$$

It is straightforward that

$$\mathbb{E}(X) = \frac{aG(b, a+1, 1)}{bG(b, a, 1)} \quad \text{and} \quad \mathbb{E}(\log X) = \int_0^1 \frac{b^a}{\Gamma(a)G(b, a)} \log(x) x^{a-1} e^{-bx} dx.$$

Note that the convergence of $\mathbb{E}(\log X)$ because

$$\begin{aligned}\mathbb{E}(\log X) &= \int_0^1 \frac{b^a}{\Gamma(a)G(b, a)} \log(x) x^{a-1} e^{-bx} dx \\ &\leq \int_0^{+\infty} \frac{b^a}{\Gamma(a)G(b, a)} \log(x) x^{a-1} e^{-bx} dx = \mathbb{E}(\log W),\end{aligned}$$

where $W \sim G(a, b)$ and then, $\mathbb{E}(\log W) = \psi(a) - \log(b)$, where $\psi(\cdot)$ denotes the digamma function.

Codes for the EM algorithm in R

```
x=c(33,40,46,41,31,37,41,56,45,31,40,35,33,43,36,36,48,45,51,44,38,36,40,
32,51,37,43,33,35,44,41,41,33,45,38,43,62,45,51,39,35,58,48,35,43,49,43,
39,39,40,39,45,48,43,45,36,40,36,47,35,40,39,44,37,36,38,37,41,38,36,36,
48,37,40,38,37,37,38,49,66,39,45,37,35,39,52,66,51,39,64,59,36,36,36,41,
41,39,45,40,37,33,66,38,59,38,41,45,35,43,39,74,63,37,45,52,43,44,52,36,
43,46,40,43,29,39,53,32,41,52,31,46,48,49,41,32,37,29,43,40,47,45,38,28,
30,40,36,37,38,37,33,30,34,38,45,40,31,39,31,31,38,32,34,45,39,31,29,39,
36,34,55,38,37,36,34,44,32,54,30,39,30,41,33,36,39,33,33,30,40,44,61,34,
26,38,26,34,36,28,36,43,35,43,37,40,35,36,28,41,30,31,48,43,43,49,36,38,
30,33,35,36,45,29,43,33,39,38,29,38,41,31,35,40,33,51,33,40,45,32,29,35,
37,35,30,32,39,32,39,38,39,83,30,33,39,33,36,39,44,31,43,44,43,41,101,
37,33)
EM.algorithm<-function(par, x, prec=1e-3, max.iter=1000)
{llike.schi2<-function(par,x){
 sigma=par[1];alpha=par[2];q=par[3]
 ll=log(q)+q*log(2*sigma)-lgamma(alpha/2)-(q+1)*log(x)+lgamma(alpha/2+q)
 +pgamma(x/(2*sigma),shape=alpha/2+q,scale=1, log.p=TRUE)
 -sum(ll)}
 par0<-par;i=1; dif=10
 while(i<=max.iter & dif>prec)
 {aux<-E.step(x, par0);v=aux$v; logv=aux$logv
 alpha.new=optim(llike.schi.c, par=par0[2], sigma=par0[1], x=x, v=v,
```

```
 logv=logv, method="Brent", lower=0, upper=10000)$par
 q.new=-1/(mean(logv));sigma.new=mean(x*v)/alpha.new
 par=c(sigma.new, alpha.new, q.new)
 dif=max(abs(par0-par));par0=par;i=i+1}
 se=sqrt(diag(solve(hessian(f=llike.schi2,x0=par0,x=x))))
 par0=cbind(par0, se);colnames(par0)=c("estimate","se")
    rownames(par0)=c("sigma","alpha","q")
 list(estimate=par0, logLik=-llike.schi2(par0, x))}
sigma=0.278; alpha=122.335; q=6.588 ##initial values
EM.algorithm(c(sigma,alpha,q),x)
```