



Research article

The Gauss hypergeometric Gleser distribution with applications to flood peaks exceedance and income data

Neveka M. Olmos¹, Emilio Gómez-Déniz² and Osvaldo Venegas^{3,*}

¹ Departamento de Estadística y Ciencia de Datos, 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 Ciencias Matemáticas y Físicas, Facultad de Ingeniería, Universidad Católica de Temuco, Temuco 4780000, Chile

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

Abstract: We introduced the Gauss hypergeometric Gleser (GHG) distribution, a novel extension of the Gleser (G) distribution that unifies families of Gleser distributions. We studied their representations and some basic properties and showed that the GHG distribution is heavy-tailed. The maximum likelihood method is used for parameter estimation, and the Fisher information matrix derived. We assessed the performance of the maximum likelihood estimators via Monte Carlo simulations. Moreover, we present applications to two data sets in which the GHG distribution shows a better fit than other known distributions.

Keywords: Gleser distribution; heavy-tailed distribution; scale mixture; maximum likelihood; Gauss hypergeometric function

Mathematics Subject Classification: 62E15, 62E20, 62P12

1. Introduction

Distributions with heavy right tail are commonly used to model data in various areas of knowledge in which extreme observations generally occur, including environmental studies, geosciences, hydrology, actuarial sciences, economics, finance, product reliability assessment, and income distribution analysis. These datasets tend to be positive with positive asymmetry and heavy right tail (see [1]). Distributions with positive support and these characteristics are usually used to model such data. Areas where heavy-tailed distributions are applicable include engineering, physics, and finance; for other fields, we refer the reader to [2] and its references.

A frequently used distribution in statistical modeling is the Pareto model (see book by Arnold [3]), named after economist Vilfredo Pareto. A random variable X follows a Pareto distribution when its probability density function (pdf) is given by:

$$f_X(x; \alpha, \beta) = \frac{\beta \alpha^\beta}{x^{\beta+1}}, \quad x \geq \alpha, \quad (1.1)$$

where $\alpha > 0$ is the scale parameter and $\beta > 0$ the shape parameter. We denote this by $X \sim \text{Pareto}(\alpha, \beta)$. The corresponding cumulative distribution function (cdf) is:

$$F_X(x; \alpha, \beta) = 1 - \left(\frac{\alpha}{x}\right)^\beta, \quad x \geq \alpha. \quad (1.2)$$

Johnson et al. [4] classify several types of Pareto distributions, referring to the above model as Pareto Type I. Beyond this standard form, various extensions have been proposed in the literature. For instance, Pickands [5] examines the generalized Pareto distribution for making statistical inferences about the upper tail of a distribution. Choulakian and Stephens [6] develop goodness-of-fit tests for this generalized model, while Davison and Smith [7] apply it to hydrological data, analyzing river-flow exceedances over 35 years.

As a further extension, Gupta et al. [8] introduce a family of distributions by raising any positive support CDF to a positive power. A special case arises when using the Pareto cdf (1.2), leading to the exponentiated Pareto (EP) distribution. A random variable X follows an EP distribution if its pdf is defined as:

$$f_X(x; \alpha, \beta, \gamma) = \frac{\beta \gamma \alpha^\beta}{x^{\beta+1}} \left(1 - \left(\frac{\alpha}{x}\right)^\beta\right)^{\gamma-1}, \quad x \geq \alpha, \quad (1.3)$$

where $\alpha > 0$ is the scale parameter and $\beta, \gamma > 0$ are the shape parameters. This model, denoted $X \sim \text{EP}(\alpha, \beta, \gamma)$, was first studied by Stoppa [9] (see also [10]), and is a particular case of the family studied by Gupta et al. [8]. Another generalization is the beta-Pareto distribution introduced by Akinsete et al. [11], and Boumaraf et al. [12] perform several optimization methods on the beta-Pareto distribution. For a comprehensive review of Pareto models (see Arnold's book [3]).

Andrews and Mallows [13] conducted research on normal distribution scale mixtures, resulting in the creation of a class of heavy-tailed distributions (including notable examples like Student's t -distribution). These distribution families are extensively applied in robust statistical analysis for symmetrical datasets.

A scale mixture distribution is formed by combining a base probability density with a scaling distribution. The mathematical representation of this mixture takes the form:

$$f_Z(z) = \int f_{Z|U=u}(z; \kappa(u)) f_U(u) du,$$

where $f_{Z|U=u}$ is the conditional distribution of the random variable Z given $U = u$, while $\kappa(u)$ represents a positive function of the random variable U with density function f_U , which influences the scale of the random variable Z . A particularly noteworthy case is the Slash distribution (see [4]), where $Z|U = u \sim N(0, \kappa(u))$, $U \sim \text{Uniform}(0, 1)$ and $\kappa(u) = u^{1/q}$. This family's theoretical properties, inferential methods, and extensions have been extensively studied by Rogers and Tukey [14] and

Mosteller and Tukey [15], while maximum likelihood (ML) estimation of its location and scale parameters are addressed in [16–18]. The same idea has been applied when the conditional distribution is with positive support, for example, to extend the Birnbaum-Saunders distribution (see [19]), and to extend the generalized half-normal distribution (see [20]).

The beta function (B) plays a fundamental role in this paper. It is defined as:

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt = \int_0^\infty \frac{t^{a-1}}{(1+t)^{a+b}} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (1.4)$$

where $a > 0$, $b > 0$, and $\Gamma(\cdot)$ denotes the gamma function. A random variable Y follows a beta distribution with parameters a and b if its pdf is given by:

$$f_Y(y; a, b) = \frac{1}{B(a, b)} y^{a-1} (1-y)^{b-1}, \quad 0 < y < 1, \quad (1.5)$$

where $a > 0$ and $b > 0$.

The incomplete beta function, denoted by $B(y; a, b)$, is defined as:

$$B(y; a, b) = \int_0^y t^{a-1}(1-t)^{b-1} dt, \quad 0 < y < 1, \quad (1.6)$$

where $a > 0$ and $b > 0$. A related function is the regularized incomplete beta function, $I_y(a, b)$, expressed as $I_y(a, b) = B(y; a, b)/B(a, b)$.

Also used in this paper is the Gauss hypergeometric function (see [21]); it is denoted by ${}_2F_1$, and is given by

$${}_2F_1(a, b, c; x) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 v^{b-1}(1-v)^{c-b-1}(1-xv)^{-a} dv, \quad (1.7)$$

where $a, b, c > 0$ and $c > b$.

Lerch transcendent function (Φ), is also used in this paper, and can be expressed as:

$$\Phi(-z, s, a) = \frac{1}{\Gamma(s)} \int_0^1 \frac{x^{a-1} \log^{s-1} \left(\frac{1}{x} \right)}{1+zx} dx, \quad (1.8)$$

where $a > 0$, $s > 0$ and $z > -1$.

A random variable X follows a G distribution with parameters σ and α when its density function is expressed as:

$$f_X(x; \sigma, \alpha) = \frac{\sigma^\alpha x^{-\alpha}}{B(\bar{\alpha}, \alpha)(\sigma + x)}, \quad x > 0, \quad (1.9)$$

with scale parameter $\sigma > 0$, shape parameter $0 < \alpha < 1$, and $\bar{\alpha} = 1 - \alpha$. The notation $X \sim G(\sigma, \alpha)$ indicates this distribution, which possesses these key characteristics:

- The G distribution is unimodal, its mode located at zero.
- For $Y \sim \text{Beta}(1 - \alpha, \alpha)$, the transformed variable $\frac{\sigma Y}{1-Y}$ follows a $G(\sigma, \alpha)$ distribution.
- The cdf of X is expressed as:

$$F_X(x; \sigma, \alpha) = I_y(\bar{\alpha}, \alpha), \quad x > 0, \quad (1.10)$$

where $y = \frac{x}{\sigma+x}$ and $I_y(\cdot, \cdot)$ is the regularised incomplete beta function.

This distribution was first introduced by Gleser [22] and later studied by Olmos et al. [23]. More recently, Olmos et al. [24] proposed a scale mixture of the G and Beta distributions, naming it scale mixture of Gleser (SMG) distribution. They show that the SMG distribution exhibits a heavy right tail, making it a viable alternative to other heavy-tailed distributions such as the Pareto distribution.

A random variable X follows a SMG distribution with parameters σ and α when its density function is:

$$f_X(x; \sigma, \alpha) = \frac{\alpha \sigma^\alpha}{B(\bar{\alpha}, \alpha)} x^{-(\alpha+1)} \log\left(1 + \frac{x}{\sigma}\right), \quad x > 0, \quad (1.11)$$

where $\sigma > 0$ represents the scale parameter and $0 < \alpha < 1$ the shape parameter. We denote this by $X \sim SMG(\sigma, \alpha)$.

The main motivation of this work is to extend the study of the G distribution by pursuing the following objectives. First, we aim to introduce the GHG distribution, generated from a scale mixture of the G and Beta distributions, similar to how the SMG distribution was derived. Notably, both the SMG and G distributions emerge as special cases of the GHG distribution. Second, we conduct a comprehensive study of the GHG distribution, demonstrating its applicability as an alternative model for actuarial data (e.g., insurance claims, income, and expenses) and hydrological data (e.g., flood peak exceedances), among other fields.

The article is organized as follows: In Section 2, we give a representation and the density of the extension of the G distribution, and its basic properties. In Section 3, we carry out an inference by the ML method, with a simulation study and the Fisher information matrix. Section 4 contains two applications to real data. In Section 5, we offer some conclusions.

2. New density and properties

In this section, we present the mathematical representation, pdf, key properties, and graphical illustrations of the GHG distribution.

2.1. Density function

The pdf of the GHG distribution is formally defined below, including several special cases.

Definition 2.1. Let $Z \sim GHG(\sigma, \alpha, \beta)$. Then, the pdf of Z is given by

$$f_Z(z; \sigma, \alpha, \beta) = \frac{\beta B(\bar{\alpha} + \beta, 1)}{\sigma^{\bar{\alpha}} B(\bar{\alpha}, \alpha)} z^{-\alpha} {}_2F_1\left(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -\frac{z}{\sigma}\right), \quad z > 0, \quad (2.1)$$

where $\sigma > 0$ represents the scale parameter, $0 < \alpha < 1$ and $\beta > 0$ are shape parameters and ${}_2F_1$ is the Gauss hypergeometric function.

Remark 1. It is verified as the following:

(1) If $\beta < \alpha$, then the GHG distribution can be written as

$$f_Z(z; \sigma, \alpha, \beta) = \frac{\beta \sigma^\beta z^{-(\beta+1)}}{B(\bar{\alpha}, \alpha)} B\left(\frac{z}{\sigma + z}, \bar{\alpha} + \beta, \alpha - \beta\right), \quad z > 0. \quad (2.2)$$

- (2) If $\beta = \alpha$, the $SMG(\sigma, \alpha)$ distribution is obtained, introduced by Olmos et al. [24].
 (3) If $\frac{z}{\sigma} < 1$ y using the Lerch transcendent function we obtain

$$f_Z(z; \sigma, \alpha, \beta) = \frac{\beta \sigma^{-\bar{\alpha}} z^{-\alpha}}{B(\bar{\alpha}, \alpha)} \Phi\left(-\frac{z}{\sigma}, 1, \bar{\alpha} + \beta\right), \quad z > 0. \quad (2.3)$$

Figure 1 displays the density plots of the GHG distribution for $\sigma = 1$, $\alpha = 0.5$ and two values of β , while Table 1 presents the values of $P(Z > z)$ for various values of z in the distributions mentioned. It can be seen that the right tail of the GHG distribution is heavier as the value of parameter β decreases.

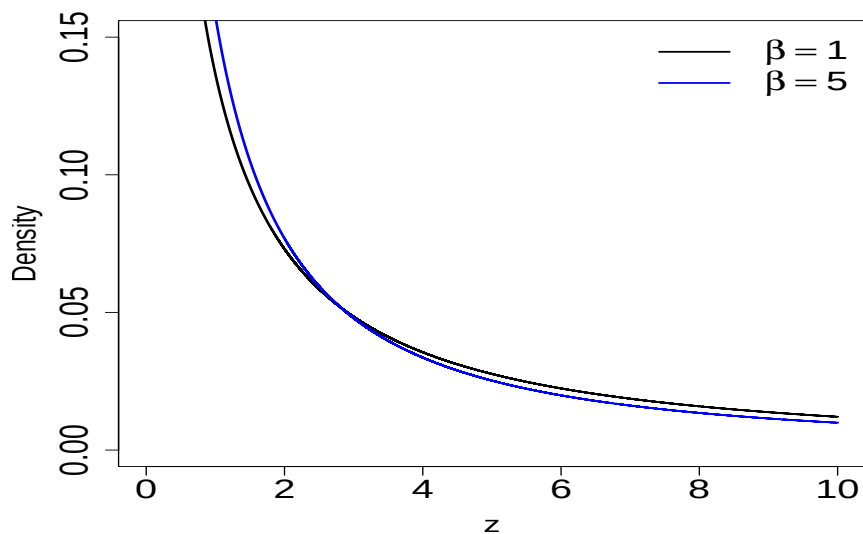


Figure 1. Examples of the $GHG(1, 0.5, \beta)$.

Table 1. Tails comparison for different GHG distributions.

Distribution	$P(Z > 4)$	$P(Z > 5)$	$P(Z > 6)$
GHG(1,0.5,1)	0.437	0.406	0.381
GHG(1,0.5,3)	0.341	0.311	0.288
GHG(1,0.5,10)	0.308	0.280	0.258

2.2. Properties

In this subsection, we show some properties of the GHG distribution.

Proposition 2.1. Let $Z|U = u \sim G(\sigma u^{-1}, \alpha)$ and $U \sim \text{Beta}(\beta, 1)$ then $Z \sim GHG(\sigma, \alpha, \beta)$.

Proof. Proceeding straight to the integral calculation, we obtain:

$$f_Z(z; \sigma, \alpha, \beta) = \int_0^1 f_{Z|U}(z) f_U(u) du = \int_0^1 \frac{(\sigma u^{-1})^\alpha z^{-\alpha}}{B(\bar{\alpha}, \alpha)(\sigma u^{-1} + z)} \beta u^{\beta-1} du = \frac{\beta z^{-\alpha}}{B(\bar{\alpha}, \alpha) \sigma^{\bar{\alpha}}} \int_0^1 \frac{u^{\beta-\alpha}}{1 + \frac{z}{\sigma} u} du.$$

The result follows from applying (1.7). $\square$

Proposition 2.2. Let $X \sim G(1, \alpha)$ and $Y \sim \text{Beta}(\beta, 1)$ be independent random variables. Then, $Z = \sigma \frac{X}{Y} \sim \text{GHG}(\sigma, \alpha, \beta)$.

Proof. Using the stochastic representation $Z = \sigma \frac{X}{Y}$ and applying the Jacobian transformation, we obtain the desired distribution. $\square$

Proposition 2.3. Let $Z \sim \text{GHG}(\sigma, \alpha, \beta)$. If $\beta \rightarrow \infty$ then Z converges in law to a random variable $Z \sim G(\sigma, \alpha)$.

Proof. Let $Z \sim \text{GHG}(\sigma, \alpha, \beta)$. Then, by the representation given in Proposition 2.2, we have that $Z = \sigma \frac{X}{Y}$. First, we study the probability convergence of Y .

When $Y \sim \text{Beta}(\beta, 1)$, we have that $E[(Y-1)^2] = \frac{2}{(\beta+1)(\beta+2)}$, if $\beta \rightarrow \infty$, then $E[(Y-1)^2] = \frac{2}{(\beta+1)(\beta+2)} \rightarrow 0$; this implies that $Y \xrightarrow{P} 1$. Applying Slutsky's Lemma (see the book by Lehmann [25]), we obtain the result. $\square$

Remark 2. The result of Proposition 2.1 shows that this distribution arises from a scale mixture of the G and Beta distributions, while Proposition 2.3 indicates that when $\beta \rightarrow \infty$ in the GHG distribution, the G distribution is obtained. Figure 2 shows the relationship between the three distributions; we observe that the GHG distribution contains the SMG and G distributions as particular cases.

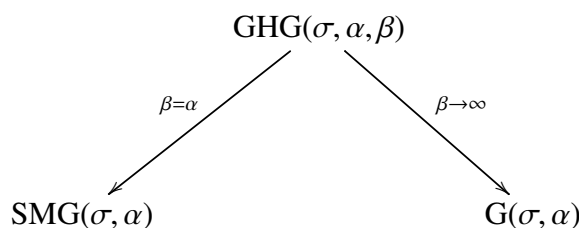


Figure 2. Particular cases for the GHG distribution.

Proposition 2.4. Let $Z \sim \text{GHG}(\sigma, \alpha, \beta)$. Then, the cdf of Z is given by

$$F_Z(z; \sigma, \alpha, \beta) = I_y(\bar{\alpha}, \alpha) - \frac{z}{\beta} f_Z(z; \sigma, \alpha, \beta), \quad z > 0, \quad (2.4)$$

where $\sigma > 0$, $0 < \alpha < 1$, $\beta > 0$, $y = \frac{z}{\sigma+z}$ and $I_y(\cdot, \cdot)$ is the regularised incomplete beta function.

Proof. Applying the definition of the cdf directly, we obtain

$$\begin{aligned} F_Z(z; \sigma, \alpha, \beta) &= \frac{\beta B(\bar{\alpha}+\beta, 1)}{\sigma^{\bar{\alpha}} B(\bar{\alpha}, \alpha)} \int_0^z t^{-\alpha} {}_2F_1\left(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -\frac{t}{\sigma}\right) dt, \\ &= \frac{\beta}{\sigma^{\bar{\alpha}} B(\bar{\alpha}, \alpha)} \int_0^z t^{-\alpha} \int_0^1 \frac{v^{\bar{\alpha}+\beta-1}}{1 + \frac{tv}{\sigma}} dv dt, \\ &= \frac{\beta}{B(\bar{\alpha}, \alpha)} \int_0^1 v^{\beta-1} \left(\int_0^{\frac{vz}{\sigma}} \frac{w^{-\alpha}}{1+w} dw \right) dv. \end{aligned}$$

Hence by integrating by parts and considering $u = \int_0^{\frac{vz}{\sigma}} \frac{w^{-\alpha}}{1+w} dw$, we obtain the result. $\square$

Corollary 2.1. Let $Z \sim GHG(\sigma, \alpha, \beta)$, the survival function $\bar{F}_X(t)$ defined by $\bar{F}_X(t) = 1 - F_X(t)$, is given by

$$\bar{F}_Z(t; \sigma, \alpha, \beta) = 1 - I_y(\bar{\alpha}, \alpha) + \frac{t}{\beta} f_Z(t; \sigma, \alpha, \beta).$$

The hazard function $h(t)$, is defined by $h(t) = \frac{f_X(t)}{\bar{F}_X(t)}$, for a random variable GHG is given by

$$h(t; \sigma, \alpha, \beta) = \frac{f_Z(t; \sigma, \alpha, \beta)}{1 - I_y(\bar{\alpha}, \alpha) + \frac{t}{\beta} f_Z(t; \sigma, \alpha, \beta)}, \quad t > 0,$$

where $y = \frac{t}{\sigma+t}$.

Figure 3 displays the hazard rate function's behavior for varying β parameters, with σ fixed at 1 and α at 0.5.

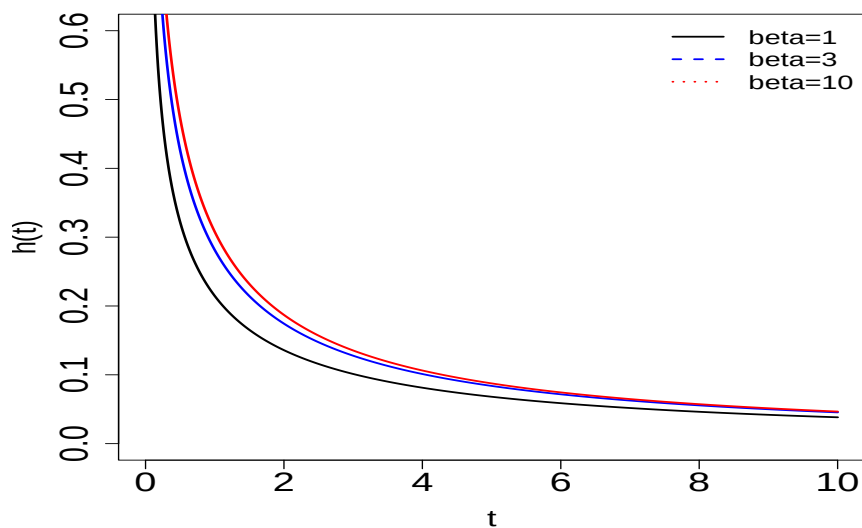


Figure 3. Plots of the hazard function $h(t)$ for $\beta = 1, 3, 10$.

In the simulation study and in parameter estimation by the ML method, we considered the GHG distribution with only two parameters, i.e., modifying Proposition 2.1 as follows: $Z|U = u \sim G(u^{-1}, \alpha)$ and $U \sim \text{Beta}(\beta, 1)$, and the following pdf of the random variable Z is the result.

$$f_Z(z; \beta, \alpha) = \frac{\beta B(\bar{\alpha} + \beta, 1)}{B(\bar{\alpha}, \alpha)} z^{-\alpha} {}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z), \quad z > 0, \quad (2.5)$$

where $\beta > 0$ and $0 < \alpha < 1$ are shape parameters. This reduction in the parametric space aims to obtain a two-parameter distribution that serves as an alternative to the Pareto distribution and other two-parameter heavy right-tailed distributions. The random variable U considers parameter β , which influences the scale in the G distribution since it is where the mixing takes place. When including parameter σ , we observe that it can lead to an overestimation of β for small sample sizes.

For generating random numbers from the GHG model, two approaches are available: Algorithm 1 uses Proposition 2.1 and the composition method (see [26]), whereas Algorithm 2 employs the representation from Proposition 2.2.

Algorithm 1 Simulates $Z \sim GHG(\beta, \alpha)$ as follows:

- Step 1: Generate $U \sim \text{Beta}(\beta, 1)$
 Step 2: Compute $Z|U = u \sim G(u^{-1}, \alpha)$
-

Algorithm 2 Simulates $Z \sim GHG(\beta, \alpha)$ as follows:

- Step 1: Generate $V \sim \text{Beta}(1 - \alpha, \alpha)$
 Step 2: Compute $X = \frac{V}{1-V}$
 Step 3: Generate $Y \sim \text{Beta}(\beta, 1)$
 Step 4: Compute $Z = \frac{X}{Y}$
-

Proposition 2.5. Let $T \sim GHG(\beta, \alpha)$. Then, the hazard function of T is decreasing for all $t > 0$.

Proof. Applying item b) of Glaser's Theorem [27], we defined the function:

$$\eta(t) = -\frac{f'(t)}{f(t)} = \frac{\alpha}{t} + \left(\frac{\bar{\alpha} + \beta}{\bar{\alpha} + \beta + 1} \right) {}_2F_1(2, \bar{\alpha} + \beta + 1, \bar{\alpha} + \beta + 2; -t),$$

where $f(t)$ is the pdf given in (2.5). Their first derivative and sign are:

$$\eta'(t) = -\left(\frac{\alpha}{t^2} + 2 \left(\frac{\bar{\alpha} + \beta}{\bar{\alpha} + \beta + 2} \right) {}_2F_1(3, \bar{\alpha} + \beta + 2, \bar{\alpha} + \beta + 3; -t) \right) < 0, \quad \forall t > 0.$$

This completes the proof by Glaser's Theorem. □

The following Proposition shows that no moments exist for the GHG distribution; this is quite common in heavy-tailed distributions.

Proposition 2.6. For the random variable $Z \sim GHG(\beta, \alpha)$, the r -th moments does not exist when $r \geq \alpha$.

Proof. We begin with the change of variable $t = zw$ and apply integrating by parts with $u = \int_0^z \frac{t^{\beta-\alpha}}{1+t} dt$. This yields:

$$\begin{aligned} \mathbb{E}(Z^r) &= \frac{\beta}{B(\bar{\alpha}, \alpha)} \int_0^\infty z^{r-\alpha} \int_0^1 \frac{w^{\beta-\alpha}}{1+zw} dw dz = \frac{\beta}{B(\bar{\alpha}, \alpha)} \int_0^\infty z^{r-\beta-1} \int_0^z \frac{t^{\beta-\alpha}}{1+t} dt dz \\ &= \frac{z^{r-\beta}}{r-\beta} \int_0^z \frac{t^{\beta-\alpha}}{1+t} dt \Big|_0^\infty - \frac{1}{r-\beta} \int_0^\infty \frac{z^{r-\alpha}}{1+z} dz. \end{aligned}$$

The left-hand side expression vanishes for $r < \beta$ and does not exist when $r \geq \beta$. As demonstrated in Proposition 1(e) of Olmos et al. [23], the right-hand integral diverges. Consequently, the expression diverges, proving that the r -th moments of $Z \sim GHG(\beta, \alpha)$ do not exist. □

Remark 3. By Proposition 2.2, the existence of the r -th moment of Z reduces to that of X , which does not exist (see Olmos et al. [23]).

2.3. Tail of the distribution

In approximate terms, heavy-tailed distributions are those whose tails decay at a slower than exponential rate. The exponential distribution is regarded as the limit between heavy and light tails. In the areas of knowledge where they are applied, for example in finance and actuarial sciences, the data are positive and have a heavy right tail. The Pareto distribution has been widely used to model financial datasets, but in many cases, it does not provide a good fit. New distributions with heavy right tail are therefore of interest. Such models are crucial for representing losses in insurance, reinsurance, and catastrophic risk. Formally, a distribution of probability with cdf $F(x)$, on the real line, is said to have a heavy right tail (see [28]) if $\limsup_{x \rightarrow \infty} (-\log(\bar{F}_X(x))/x) = 0$.

Regular variation theory plays a fundamental role in extreme value analysis (see for example [29–32], etc.). This concept is formalized as follows:

Definition 2.2. A distribution function is called regular varying at infinity with index $-\beta$ (where $\beta \geq 0$) if:

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_X(tx)}{\bar{F}_X(x)} = t^{-\beta}.$$

Here, β is termed the tail index.

The GHG distribution exhibits this important property:

Proposition 2.7. The survival function of $Z \sim GHG(\beta, \alpha)$, has regularly varying tails.

Proof. Applying the above definition and L'Hospital's Rule, we have that

$$\lim_{z \rightarrow \infty} \frac{\bar{F}_Z(tz)}{\bar{F}_Z(z)} = \lim_{z \rightarrow \infty} \frac{1 - F(tz)}{1 - F(z)} = t^{-\alpha+1} \lim_{z \rightarrow \infty} \frac{\int_0^1 \frac{w^{\beta-\alpha}}{1+tzw} dw}{\int_0^1 \frac{w^{\beta-\alpha}}{1+zw} dw} = t^{-\beta} \lim_{z \rightarrow \infty} \frac{\int_0^{tz} \frac{u^{\beta-\alpha}}{1+u} du}{\int_0^z \frac{u^{\beta-\alpha}}{1+u} du} = t^{-\beta},$$

and taking into account that $\beta > 0$ the result follows. $\square$

As the GHG distribution is of regularly varying tails, it is therefore heavy-tailed (see [29]).

3. Inference

In this section, we describe ML estimation for the GHG model and include a simulation study assessing the ML estimators' behavior.

3.1. ML estimation

Let $z_1, \dots, z_n$ be a random sample from the $GHG(\beta, \alpha)$ distribution. The log-likelihood is given by:

$$\ell(\beta, \alpha) = c(\beta, \alpha) - \alpha \sum_{i=1}^n \log(z_i) + \sum_{i=1}^n \log({}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z_i)), \quad (3.1)$$

where $c(\beta, \alpha) = n \log(\beta) - n \log(\bar{\alpha} + \beta) - n \log(\Gamma(\bar{\alpha})) - n \log(\Gamma(\alpha))$.

By equating the first partial derivatives (w.r.t. each parameter) to zero, we obtain the following system of equations:

$$(\bar{\alpha} + \beta) \sum_{i=1}^n \frac{\Phi(-z_i, 2, \bar{\alpha} + \beta)}{{}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z_i)} = \frac{n}{\beta}, \quad (3.2)$$

$$(\bar{\alpha} + \beta) \sum_{i=1}^n \frac{\Phi(-z_i, 2, \bar{\alpha} + \beta)}{{}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z_i)} - \sum_{i=1}^n \log(z_i) + n\psi(\bar{\alpha}) = n\psi(\alpha), \quad (3.3)$$

where $\psi(\cdot)$ is the digamma function. From (3.2) and (3.3) we derive:

$$\hat{\beta}(\hat{\alpha}) = \frac{n}{\sum_{i=1}^n \log(z_i) + n\psi(\hat{\alpha}) - n\psi(1 - \hat{\alpha})}. \quad (3.4)$$

The ML estimator for α ($\hat{\alpha}$) is obtained by numerically solving:

$$(1 - \hat{\alpha} + \hat{\beta}(\hat{\alpha})) \sum_{i=1}^n \frac{\Phi(-z_i, 2, 1 - \hat{\alpha} + \hat{\beta}(\hat{\alpha}))}{{}_2F_1(1, 1 - \hat{\alpha} + \hat{\beta}(\hat{\alpha}), 2 - \hat{\alpha} + \hat{\beta}(\hat{\alpha}); -z_i)} = \frac{n}{\hat{\beta}(\hat{\alpha})}, \quad (3.5)$$

The estimator $\hat{\alpha}$ corresponds to the solution of Eq (3.5), which when substituted into (3.4) yields $\hat{\beta}$. Equation (3.5) can be solved using numerical methods such as the Newton-Raphson algorithm. Alternatively, both estimates can be obtained by directly maximizing the log-likelihood surface defined in (3.1) using the `optim` subroutine in the R software package [33].

3.2. Fisher's information matrix

For a random variable $Z \sim GHG(\beta, \alpha)$, the log-likelihood function corresponding to a single observation z of Z and parameter vector $\theta = (\beta, \alpha)$ is expressed as:

$$\ell(\theta) = \log(\beta) - \log(\bar{\alpha} + \beta) - \log(B(\bar{\alpha}, \alpha)) - \alpha \log(z) + \log({}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z)).$$

The Appendix contains derivations of both first and second partial derivatives of this log-likelihood function. The Fisher information matrix, denoted $I_F(\cdot)$, for the G distribution takes the form:

$$I_F(\theta) = \begin{pmatrix} \frac{1}{\beta^2} - 2\eta_{31} + \eta_{22} & 2\eta_{31} - \eta_{22} \\ 2\eta_{31} - \eta_{22} & \psi'(\alpha) + \psi'(\bar{\alpha}) - \eta_{31} + \eta_{22} \end{pmatrix},$$

where $\psi'(\cdot)$ is the trigamma function and $\eta_{ij} = E\left(\frac{\kappa_0 \Phi(-Z, i, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -Z)}\right)^j$ are calculated numerically.

For large samples, the ML estimator $\hat{\theta}$ is asymptotically normal bivariate:

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{\mathcal{L}} N_2(\mathbf{0}, I_F(\theta)^{-1}).$$

Consequently, the asymptotic variance of $\hat{\theta}$ equals the inverse of $I_F(\theta)$. In practice, since the parameters are typically unknown, we employ the observed information matrix with ML-estimated parameters.

3.3. Simulation study

To examine the behavior of estimation by ML, we conducted a simulation study to assess the performance of the estimation method for the β and α parameters of the GHG distribution. Using Algorithm 2 given in Subsection 2.2 to generate random numbers from the GHG distribution, the simulation analysis was performed by generating 1000 samples of sizes $n = 20, 100$ and 300 from the GHG distribution. Table 2 shows the empirical bias (Bias), medium bias (MBias), mean of the standard errors (SE), and root of the empirical mean squared error (RMSE). As shown in Table 2, the estimation performance improves as n increases.

Table 2. Bias, MBias, SE, and RMSE for the estimators of α and β .

True values		Estimator	$n = 20$				$n = 100$				$n = 300$			
α	β		Bias	MBias	SE	RMSE	Bias	MBias	SE	RMSE	Bias	MBias	SE	RMSE
0.25	0.5	$\hat{\alpha}$	-0.0124	-0.0213	0.0606	0.0542	-0.0277	-0.0291	0.0233	0.0355	-0.0292	-0.0296	0.0124	0.0320
		$\hat{\beta}$	0.2411	0.1496	0.6230	0.3956	0.2930	0.2213	0.2248	0.3750	0.2615	0.2322	0.0747	0.2963
0.25	1.0	$\hat{\alpha}$	0.0230	0.0173	0.0723	0.0680	0.0002	-0.0013	0.0298	0.0289	-0.0031	-0.0039	0.0162	0.0183
		$\hat{\beta}$	0.0484	-0.2142	15.138	0.6845	0.2650	-0.0001	10.147	0.7274	0.2239	0.0406	0.4547	0.5718
0.5	0.5	$\hat{\alpha}$	-0.0346	-0.0318	0.1004	0.0986	-0.0447	-0.0432	0.0406	0.0728	-0.0473	-0.0428	0.0188	0.0618
		$\hat{\beta}$	0.2436	0.1512	0.6225	0.4308	0.2234	0.1677	0.2024	0.3440	0.1973	0.1543	0.0663	0.2879
0.5	1.0	$\hat{\alpha}$	0.0103	0.0032	0.1067	0.0893	-0.0051	-0.0067	0.0517	0.0502	-0.0101	-0.0097	0.0282	0.0328
		$\hat{\beta}$	0.2108	-0.0227	18.389	0.7911	0.2919	0.0633	10.861	0.7391	0.2592	0.0624	0.5843	0.6181
0.8	0.5	$\hat{\alpha}$	-0.0204	-0.0151	0.0554	0.0546	-0.0092	-0.0083	0.0230	0.0253	-0.0045	-0.0041	0.0129	0.0133
		$\hat{\beta}$	0.1773	0.0777	0.4606	0.3622	0.0733	0.0474	0.1234	0.1473	0.0395	0.0334	0.0622	0.0693
0.8	1.0	$\hat{\alpha}$	-0.0060	0.0028	0.0517	0.0484	-0.0023	-0.0001	0.0227	0.0241	-0.0012	-0.0010	0.0126	0.0136
		$\hat{\beta}$	0.2104	-0.0373	16.475	0.7766	0.2051	0.0264	0.6920	0.6181	0.1027	0.0168	0.2561	0.3444

4. Applications with real data

In this section, we analyze two real-data applications, comparing the fits of several distributions against the GHG distribution. Model selection is performed using four information criteria: AIC (Akaike Information Criterion; [34]), BIC (Bayesian Information Criterion; [35]), CAIC (Consistent AIC; [36]) and HQIC (Hannan-Quinn Information Criterion; [37]). These criteria are defined as: $AIC = -2\hat{\ell}(\cdot) + 2p$, $BIC = -2\hat{\ell}(\cdot) + \log(n)p$, $CAIC = -2\hat{\ell}(\cdot) + (\log(n) + 1)p = BIC + p$, and $HQIC = -2\hat{\ell}(\cdot) + 2 \log(\log(n))p$, where p denotes the number of parameters in the model.

4.1. Application to exceedances of flood peak data

The data correspond to the exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada. The data consist of 72 exceedances for the years 1958-1984, rounded to one decimal place (see [6, 11, 38]), and are listed in Table 3. Table 4 presents descriptive statistics, including the sample skewness coefficient (CS) and kurtosis coefficient (CK). We compare the fits of the EP and Pareto distributions, with the fit of the GHG distribution with two parameters given in (2.5).

Table 3. Data on exceedances of Wheaton River floods.

1.7	2.2	14.4	1.1	0.4	20.6	5.3	0.7	1.9	13.0	12.0	9.3
1.4	18.7	8.5	25.5	11.6	14.1	22.1	1.1	2.5	14.4	1.7	37.6
0.6	2.2	39.0	0.3	15.0	11.0	7.3	22.9	1.7	0.1	1.1	0.6
9.0	1.7	7.0	20.1	0.4	2.8	14.1	9.9	10.4	10.7	30.0	3.6
5.6	30.8	13.3	4.2	25.5	3.4	11.9	21.5	27.6	36.4	2.7	64.0
1.5	2.5	27.4	1.0	27.1	20.2	16.8	5.3	9.7	27.5	2.5	27.0

Table 4. Descriptive statistics for exceedances of flood peaks data.

n	Median	Mean	SD	CS	CK
72	9.5	12.204	12.297	1.473	5.890

The box plot in Figure 4 shows one very extreme datum. This outlier makes the right tail heavier. While most data points cluster around 10 flood peak exceedances (m^3/s), there is one notable outlier at 64 exceedances (m^3/s).

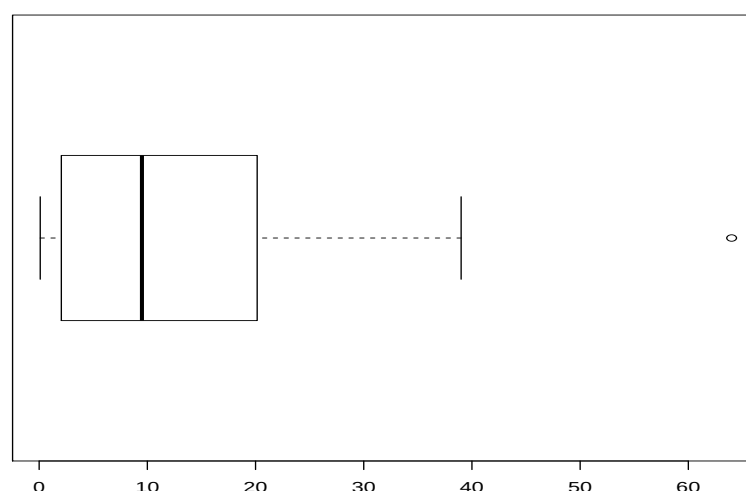
**Figure 4.** Boxplot for exceedances of flood peaks data.

Table 5 presents the ML estimates for the GHG, EP, and Pareto model parameters, along with their corresponding AIC, BIC, CAIC, and HQIC values.

Table 5. Exceedances of flood peaks data: model, ML estimates, AIC, BIC, CAIC, and HQIC values.

Model	ML estimates	AIC	BIC	CAIC	HQIC
GHG(β, α)	$\hat{\beta} = 1.236(0.465), \hat{\alpha} = 0.403(0.045)$	569.8	574.3	576.3	571.6
EP(α, β, γ)	$\hat{\alpha} = 0.1, \hat{\beta} = 0.424(0.046), \hat{\gamma} = 2.880(0.491)$	578.6	583.2	586.2	581.3
Pareto(α, β)	$\hat{\alpha} = 0.1, \hat{\beta} = 0.244(0.029)$	608.1	610.4	611.4	609.0

The GHG model demonstrates superior fit to the data, as evidenced by its consistently lower AIC, BIC, CAIC, and HQIC values compared to both EP and Pareto models.

In Figure 5, the left side displays the histogram of the data along with the curves of the respective fitted distributions. It is evident that the GHG distribution provides a better fit for the exceedances of flood peak data compared to the EP and Pareto distributions. On the right side, there is a zoomed-in view of the right tail of the histogram, which more conclusively demonstrates that the GHG distribution concentrates more probability in the right tail than the other two distributions for this data set.

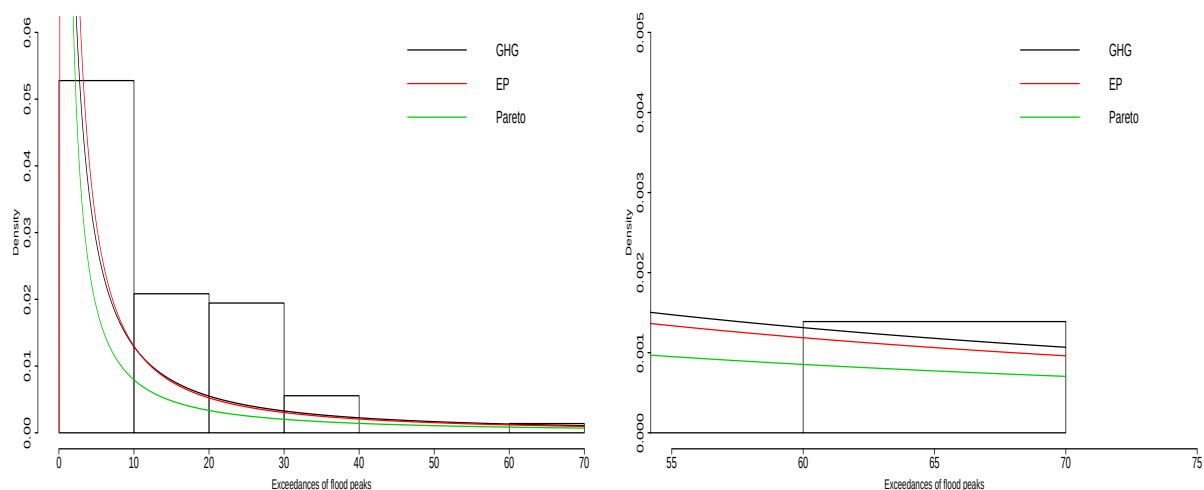


Figure 5. Plots of the GHG (black), EP (red) and Pareto (green) models.

4.2. Application to income data

The dataset used in this application comes from the Survey of Consumer Finances (SCF), a nationally representative sample containing extensive information on assets, liabilities, income, and demographic characteristics of respondents (potential U.S. customers). We analyzed a random sample of 500 income-positive U.S. households interviewed in the 2004 survey. The variable of interest is each household's annual income in US dollars. The data can be accessed from the Federal Reserve's webpage: <https://www.federalreserve.gov/econres/scfindex.htm>. Descriptive statistics for these data are presented in Table 6. We compared the fits of the SMG and Pareto distributions with that of the three-parameter GHG distribution given in (2.1).

Table 6. Descriptive statistics for income data.

n	Median	Mean	SD	CS	CK
500	54000	321021.9	3410936	21.127	461.575

Figure 6 displays a boxplot revealing several extreme values. These outliers create a heavy right tail in the distribution. Notably, while most observations cluster around 5,400 (representing typical household income), there is an extreme value of 75 million in family income.

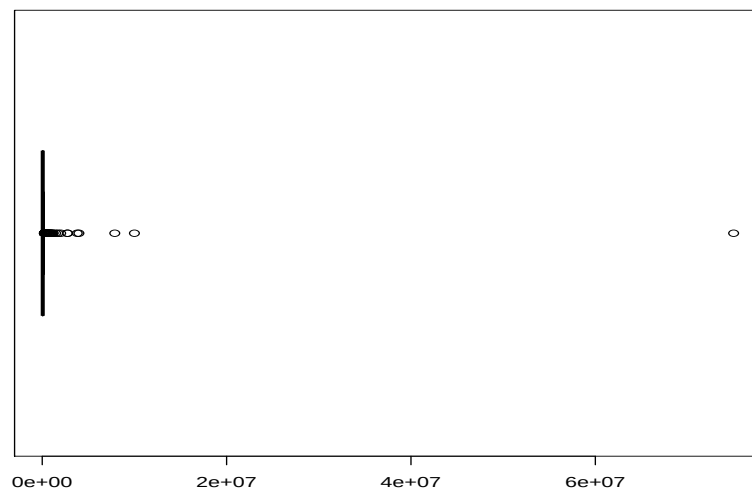


Figure 6. Boxplot for Descriptive statistics for income data.

Table 7 presents the ML estimates for the parameters and log-likelihood values of the GHG, SMG, and Pareto models. Table 8 displays the corresponding values for the AIC, BIC, CAIC, and HQIC criteria for each model.

Table 7. Model, ML estimates, and Log-likelihood.

Model	ML estimates	Log-likelihood
GHG(σ, α, β)	$\hat{\sigma} = 56417.5(2174.647), \hat{\alpha} = 0.5(0.015), \hat{\beta} = 756.154(62.525)$	-6466.192
SMG(σ, α)	$\hat{\sigma} = 22584.770(4027.346), \hat{\alpha} = 0.581(0.018)$	-6541.776
Pareto(α, β)	$\hat{\alpha} = 260, \hat{\beta} = 0.186(0.008)$	-6802.113

Table 8. Model, AIC, BIC, CAIC, and HQIC values.

Model	AIC	BIC	CAIC	HQIC
GHG(σ, α, β)	12938.38	12951.03	12954.03	12943.35
SMG(σ, α)	13087.55	13095.98	13097.98	13090.86
Pareto(α, β)	13606.23	13610.44	13611.44	13606.05

We observed that the GHG model yields the lowest values for the AIC, BIC, CAIC, and HQIC criteria, indicating that it provides a better fit to the data compared to both the SMG and Pareto models.

Figure 7 shows the empirical cdf with estimated GHG, SMG, and Pareto cdf's, which also shows a favorable agreement between the GHG model and the income data.

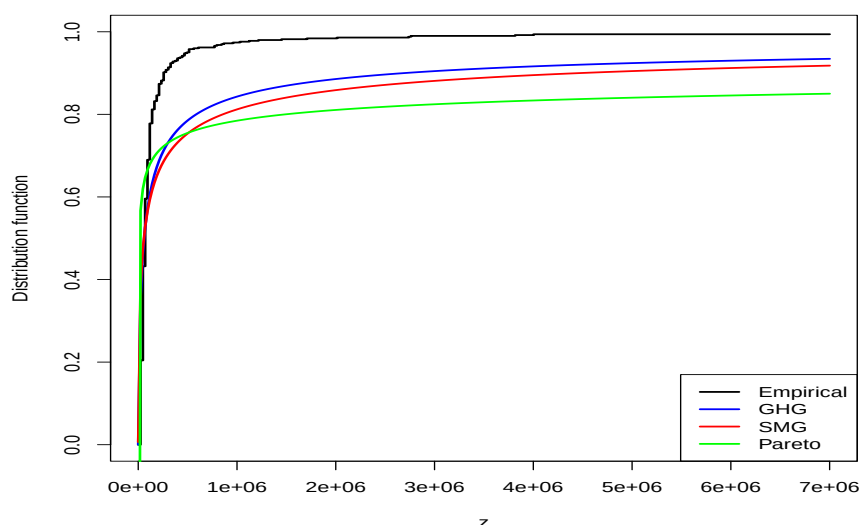


Figure 7. Plots of the empirical cdf, with estimated GHG cdf, estimated SMG, and estimated Pareto cdf models

5. Conclusions

In this work, we introduce a new heavy-tailed distribution through a scale mixture extension of the G distribution, combining it with the Beta distribution to create the GHG model. We study some properties, perform parameter estimation by the ML method, and show two applications to real data. The GHG distribution contains the G and SMG distributions as special cases. This is of interest, since it can be used in actuarial sciences and related fields. The GHG model is a viable alternative for fitting data with extreme observations. Some other properties of the GHG model are:

- The GHG distribution exhibits a heavy right tail.
- The GHG distribution has two representations, as presented in Propositions 2.1 and 2.2.
- The pdf, cdf, and hazard function have explicit forms expressed through known special functions, such as the Gauss hypergeometric function.
- The hazards function is monotonically decreasing, which provides information about the distribution's tail behavior (specifically, its heavy right tail; see [28]).
- In the first application, we demonstrate that the GHG distribution effectively models flood peak exceedance data, outperforming two well-known distributions: the EP and Pareto distributions.
- In the second application, we show that the three-parameter GHG distribution provides a good fit for income data, surpassing the SMG and Pareto distributions.
- A more comprehensive estimation study of the GHG distribution's three parameters will be addressed in future work, where we will analyze the effects and limitations of the σ and β parameters in the ML estimation method.

Author contributions

Conceptualization: N.M. Olmos and E. Gómez-Déniz; methodology: N.M. Olmos and E. Gómez-Déniz; software: N.M. Olmos and E. Gómez-Déniz; validation: N.M. Olmos, E. Gómez-Déniz, and O. Venegas; formal analysis: N.M. Olmos and O. Venegas; investigation: E. Gómez-Déniz and N.M. Olmos; writing—original draft preparation: N.M. Olmos and E. Gómez-Déniz; writing—review and editing: O. Venegas and E. Gómez-Déniz. 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 not 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. R. Ibragimov, A. Prokhorov, *Heavy tails and copulas: topics in dependence modelling in economics and finance*, Singapore: World Scientific, 2017. <https://doi.org/10.1142/9644>
2. I. B. Aban, M. M. Meerschaert, A. K. Panorska, Parameter estimation for the truncated Pareto distribution, *J. Am. Stat. Assoc.*, **101** (2006), 270–277. <https://doi.org/10.1198/016214505000000411>
3. B. C. Arnold, *Pareto distributions*, 2 Eds., Boca Raton: Chapman & Hall, 2015. <https://doi.org/10.1201/b18141>
4. N. L. Johnson, S. Kotz, N. Balakrishnan, *Continuous univariate distributions*, 2 Eds., Vol. 1, New York: Wiley, 1995.
5. J. Pickands, Statistical inference using extreme order statistics, *Ann. Statist.*, **3** (1975), 119–131. <https://doi.org/10.1214/aos/1176343003>
6. V. Choulakian, M. A. Stephens, Goodness-of-fit for the generalized Pareto distribution, *Technometrics*, **43** (2001), 478–484.

7. A. C. Davison, R. L. Smith, Models for exceedances over high thresholds, *J. R. Stat. Soc., Ser. B*, **52** (1990), 393–442.
8. R. C. Gupta, P. L. Gupta, R. D. Gupta, Modeling failure time data by Lehman alternatives, *Commun. Stat. Theory Methods*, **27** (1998), 887–904. <https://doi.org/10.1080/03610929808832134>
9. G. Stoppa, Proprieta campionarie di un nuovo modello Pareto generalizzato, *Proceedings of the Atti XXXV Riunione Scientifica della Societa Italiana di Statistica*, Cedam: Padova, Italy, 1990, 137–144.
10. C. Kleiber, S. Kotz, *Statistical size distributions in economics and actuarial sciences*, Hoboken: John Wiley & Sons, 2003. <https://doi.org/10.1002/0471457175>
11. A. Akinsete, F. Famoye, C. Lee, The Beta-Pareto distribution, *Statistics*, **42** (2008), 547–563.
12. B. Boumaraf, N. Seddik-Ameur, V. S. Barbu, Estimation of Beta-Pareto distribution based on several optimization methods, *Mathematics*, **8** (2020), 1055. <https://doi.org/10.3390/math8071055>
13. D. F. Andrews, C. L. Mallows, Scale mixtures of normal distributions, *J. Roy. Stat. Soc.: Ser. B*, **36** (1974), 99–102. <https://doi.org/10.1111/j.2517-6161.1974.tb00989.x>
14. 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>
15. F. Mosteller, J. W. Tukey, *Data analysis and regression*, Reading: Addison-Wesley, 1977.
16. K. Kafadar, A biweight approach to the one-sample problem, *J. Am. Stat. Assoc.*, **77** (1982), 416–424. <https://doi.org/10.1080/01621459.1982.10477827>
17. 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>
18. 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>
19. H. W. Gómez, J. F. Olivares-Pacheco, H. Bolfarine, An extension of the generalized Birnbaum-Saunders distribution, *Statist. Probab. Lett.*, **79** (2009), 331–338. <https://doi.org/10.1016/j.spl.2008.08.014>
20. N. M. Olmos, H. Varela, H. Bolfarine, H. W. Gómez, An extension of the generalized half-normal distribution, *Stat. Papers*, **55** (2014), 967–981. <https://doi.org/10.1007/s00362-013-0546-6>
21. M. Abramowitz, I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, 9 Eds., Dover Publications, 1965.
22. L. J. Gleser, The Gamma distribution as a mixture of exponential distributions, *Am. Stat.*, **43** (1989), 115–117. <https://doi.org/10.2307/2684515>
23. N. M. Olmos, E. Gómez-Déniz, O. Venegas, The heavy-tailed Gleser model: properties, estimation, and applications, *Mathematics*, **10** (2022), 4577. <https://doi.org/10.3390/math10234577>
24. N. M. Olmos, E. Gómez-Déniz, O. Venegas, Scale mixture of Gleser distribution with an application to insurance data, *Mathematics*, **12** (2024), 1397. <https://doi.org/10.3390/math12091397>

25. E. L. Lehman, *Elements of large-sample theory*, New York: Springer, 1999. <https://doi.org/10.1007/b98855>
26. M. A. Tanner, *Tools for statistical inference*, 3 Eds., New York: Springer, 1996. <https://doi.org/10.1007/978-1-4612-4024-2>
27. R. E. Glaser, Bathtub and related failure rate characterizations, *J. Am. Stat. Assoc.*, **75** (1980), 667–672. <https://doi.org/10.1080/01621459.1980.10477530>
28. T. Rolski, H. Schmidli, V. Schmidt, J. Teugels, *Stochastic processes for insurance and finance*, Hoboken: John Wiley & Sons, 1999. <https://doi.org/10.1002/9780470317044>
29. W. Feller, *An introduction to probability theory and its applications*, 2 Eds., Vol. 2, New York: John Wiley & Sons, 1968.
30. N. H. Bingham, C. M. Goldie, J. L. Teugels *Regular variation*, Cambridge: Cambridge University Press, 1987. <https://doi.org/10.1017/CBO9780511721434>
31. J. Beirlant, G. Matthys, G. Dierckx, Heavy-tailed distributions and rating, *ASTIN Bull.*, **31** (2001), 37–58. <https://doi.org/10.2143/AST.31.1.993>
32. D. G. Konstantinides, *Risk theory: a heavy tail approach*, World Scientific Publishing, 2017.
33. R Core Team, *Gradient estimates for solutions of nonlinear elliptic and parabolic equations*, R Foundation for Statistical Computing: Vienna, Austria, 2021. Available from: <https://www.R-project.org/>.
34. H. Akaike, A new look at the statistical model identification, *IEEE Trans. Automatic Control*, **19** (1974), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
35. G. Schwarz, Estimating the dimension of a model, *Ann. Stat.*, **6** (1978), 461–464.
36. H. Bozdogan, Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions, *Psychometrika*, **52** (1987), 345–370. <https://doi.org/10.1007/BF02294361>
37. E. J. Hannan, B. G. Quinn, The determination of the order of an autoregression., *J. R. Stat. Soc.: Ser. B*, **41** (1979), 190–195.
38. E. Mahmoudi, The beta generalized Pareto distribution with application to lifetime data, *Math. Comput. Simulat.*, **81** (2011), 2414–2430. <https://doi.org/10.1016/j.matcom.2011.03.006>

Appendix

The first-order partial derivatives of $\ell(\theta)$ with respect to θ are given by:

$$\begin{aligned}\frac{\partial \ell(\theta)}{\partial \beta} &= \frac{1}{\beta} - \frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)}, \\ \frac{\partial \ell(\theta)}{\partial \alpha} &= \psi(\bar{\alpha}) - \psi(\alpha) - \log(z) + \frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)}.\end{aligned}$$

The second-order partial derivatives of $\ell(\theta)$ with respect to θ are given by:

$$\frac{\partial^2 \ell(\theta)}{\partial \beta^2} = -\frac{1}{\beta^2} + \frac{2\kappa_0 \Phi(-z, 3, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} - \left(\frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} \right)^2,$$

$$\frac{\partial^2 \ell(\theta)}{\partial \beta \partial \alpha} = \left(\frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} \right)^2 - \frac{2\kappa_0 \Phi(-z, 3, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)},$$

$$\frac{\partial^2 \ell(\theta)}{\partial \alpha^2} = -\psi'(\bar{\alpha}) - \psi'(\alpha) + \frac{2\kappa_0 \Phi(-z, 3, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} - \left(\frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} \right)^2,$$

where $\kappa_i = \bar{\alpha} + \beta + i$, $\psi(\cdot)$ and $\psi'(\cdot)$ are the digamma and trigamma functions, respectively.



AIMS Press

© 2025 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)