A general class of continuous asymmetric distributions with positive support

Felipe S. Quintino [*1], Pushpa N. Rathie [†1], Luan C. S. M. Ozelim [‡2], Tiago A. da Fonseca [§3], and Roberto Vila [¶1]

[1]Department of Statistics, University of Brasília, Brasília 70.910-900, Brazil
[2]Department of Civil and Environmental Engineering, University of Brasília, Brasília 70.910-900, Brazil
[3]Gama Engineering College, University of Brasília, Brasília 72.444-240, Brazil

December 11, 2025

Abstract

In order to better fit real-world datasets, studying asymmetric distribution is of great interest. In this work, we derive several mathematical properties of a general class of asymmetric distributions with positive support which shows up as a unified framework for Extreme Value Theory asymptotic results. The new model generalizes some well-known distribution models such as Generalized Gamma, Inverse Gamma, Weibull, Fréchet, Half-normal, Modified half-normal, Rayleigh, and Erlang. To highlight the applicability of our results, the performance of the analytical models is evaluated through real-life dataset modeling.

Keywords: Generalized classes of distributions, Skewed probability distributions, Extreme value $\mathbb{H}$-function, Maximum likelihood estimator, Statistical modeling of asymmetric data.

1 Introduction

Asymmetric distributions are common across various fields, such as epidemiology and finance, where they help model phenomena like disease spread and risk assessment [8]. In finance, heavy-tailed distributions are the most suitable tool for predicting stock performance and optimizing portfolios by quantifying probabilities and managing risk [20, 21, 16]. Our focus, however, is to study a new class of distributions with positive support that generalizes many existing asymmetric models.

Nadarajah [12] studied problems involving stress-strength reliability (SSR) for extreme value distributions. He wrote the probability $P(X < Y)$ in terms of generalized hypergeometric functions when X and Y were independent random variables with Weibull, Fréchet, or Gumbel distribution and relied on severe parameter restrictions to get analytical results.

Inspired by the work of [12], [22] observed that the parameter restrictions could be relaxed if a new class of special functions was introduced, called extreme value $\mathbb{H}$-function. Later, [14, 15, 16] showed that $\mathbb{H}$-functions could also be used to describe SSR for the distributions of Generalized Extreme Value (GEV) distributions, transmuted GEV and p-max stable laws families.

Departing from mathematical properties of $\mathbb{H}$-functions studied in [22], a new class of probability density function (PDF) with positive support was introduced in [22]. However, the mathematical and statistical properties and the applications of this model require a more in-depth study, which will be provided in this work. The PDF is given by:

$$g(y; \boldsymbol{\theta}) = \frac{1}{c(\boldsymbol{\theta})} y^{\theta_6} \exp\left(-\theta_1 y - (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5}\right), \quad y > 0, \tag{1}$$

---

[*]Corresponding author: felipes.quintino2@gmail.com

[†]pushpanrathie@yahoo.com

[‡]luanoz@gmail.com

[§]fonsecafga@unb.br

[¶]rovig161@gmail.com

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$ and $c(\boldsymbol{\theta})$ is a normalizing constant (more details about the parameter space for $\boldsymbol{\theta}$ will be given in the following sections). This new class of PDFs generalizes several well-known models in the literature, such as exponential, Gamma, Generalized Gamma [26], Inverse Gaussian [9], Inverse Generalized Gamma [19], Weibull [9], Fréchet, Half-normal, Modified half-normal [28], Rayleigh [9], Maxwell–Boltzmann [9] and Erlang [9] distributions, among others.

Although [22] introduced the PDF (1), they focused only on the mathematical properties of the extreme value $\mathbb{H}-$function, thus a complete characterization of the model needs to be discussed to fully expose its potential. For example, a topic that needs to be studied is the description of the incomplete extreme value $\mathbb{H}-$function, which is necessary for presenting the model's cumulative distribution function (CDF). The stochastic representation provides an alternative way to generate random variables, which was not addressed in the previous work. A more in-depth study of the modality and shapes, which are essential for a proper understanding of the model's characteristics, is also needed to be described, in addition to other important features such as moments and characteristic function. These aspects are fundamental not only for ensuring the internal consistency of the proposed probability model but also for enabling its practical use in statistical inference and simulation-based analysis. Without a complete characterization, including the CDF, stochastic representation, and moment properties, the applicability of the distribution to real-world problems remains limited. Therefore, this work aims to bridge this gap by providing both the theoretical development and the computational tools required for its implementation.

The aim of this paper is to study statistical properties of (1), such as particular cases, modality, shapes, finite mixtures, characteristic function, Mellin transform, moments, and entropy. Inferential properties are also addressed through maximum likelihood estimators and conditions for $g(y; \boldsymbol{\theta})$ to belong to the exponential family are derived. A second estimator is also proposed, based on the difference between the empirical distribution function and the theoretical cumulative distribution function.

In our study, we assess two estimators through real data modeling. We examine three distinct datasets to gain insights into their performance. The first dataset focuses on the minimum monthly water flows (measured in cubic meters per second, m$^{\mathbf{3}}$/s) of the Piracicaba River, situated in São Paulo state, Brazil. The second dataset investigates the strength of carbon fibers under stress. These fibers were tested to understand their behavior when subjected to tension. Finally, we analyze a dataset that captures the failure times of various machine parts. This information is important for maintenance and reliability considerations.

The paper is organized as follows: In Section 2 the model is presented and we study modality, shapes, and particular cases. In Section 3, properties of the new class of distributions are derived. Two classes of estimators are presented in Section 4, while Section 5 presents real data applications.

2   A class of continuous distributions with positive support

Let $\Theta \subset \mathbb{R}^6$ be a parameter space. For a random variable Y we define the PDF

$$g(y; \boldsymbol{\theta}) = \frac{1}{c(\boldsymbol{\theta})} y^{\theta_6} \exp\left(-\theta_1 y - (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5}\right), \quad y > 0, \tag{2}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6) \in \Theta$ is a parameter vector such that $\theta_1 \geq 0$, $\theta_2 \geq 0$ (θ_1 and θ_2 cannot be zero simultaneously), $\theta_4 \geq 0$ and $\theta_3, \theta_5 \in \mathbb{R}$. $\theta_6 > -1$ when $\theta_1 \neq 0$ or $\theta_1 = 0$ and $\text{sign}(\theta_3) = \text{sign}(\theta_5)$, $\theta_6 < -1$ when $\theta_1 = 0$ and $\text{sign}(\theta_3) \neq \text{sign}(\theta_5)$. These are general restrictions which could be further manipulated when $\theta_5 \in \mathbb{Z}$. In particular, an integer valued θ_5 allows one to relax the constraints on θ_1, θ_2 and θ_4, extending their possible domains to negative values. This should, on the other hand, be assessed in a case by case basis. In Subsection 2.1, for example, the case when $\theta_5 \in \mathbb{N}$ is discussed. In that case, if θ_5 is even, then $\theta_1 > 0$ and $\theta_2, \theta_4 \in \mathbb{R}$. On the other hand, if θ_5 is odd, then three possible scenarios arise: if $\theta_3 \theta_5 < 1$, then $\theta_2 > 0$ and $\theta_1 \in \mathbb{R}$; if $\theta_3 \theta_5 = 1$, then $(\theta_1 + \theta_2^{\theta_5}) > 0$; finally, if $\theta_3 \theta_5 < 1$, then $\theta_1 > 0$ and $\theta_2 \in \mathbb{R}$. The function $c(\boldsymbol{\theta})$ appearing in the definition of g is a normalizing factor written as

$$c(\boldsymbol{\theta}) = \mathbb{H}(\boldsymbol{\theta}) := \mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6) = \int_0^\infty y^{\theta_6} \exp\left(-\theta_1 y - (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5}\right) \mathrm{d}y. \tag{3}$$

We recall that extreme value $\mathbb{H}$-function was first defined in [22] who also proposed the PDF (2). Statistical and mathematical properties of $g(y; \boldsymbol{\theta})$ are addressed in this paper.

Note that, in general, the model (2) is not (globally) identifiable. For example, for $\boldsymbol{\theta} = (\theta, 0, 0, 0, 1, 0) \neq \boldsymbol{\theta}' = (0, \theta, 1, 0, 1, 0)$, we have $g(y; \boldsymbol{\theta}) = g(y; \boldsymbol{\theta}')$. By globally identifiable, we mean that the mapping $\Theta \ni \boldsymbol{\theta} \mapsto g(y; \boldsymbol{\theta})$ is one-to-one, for all $y > 0$, onto the entire parameter space Θ. It is often possible to achieve identifiability in a parametric model that fails to satisfy this property by introducing specific technical constraints, such as those discussed in Ref. [23], thereby facilitating meaningful inference.

The CDF corresponding to PDF $g(y; \boldsymbol{\theta})$ is given by

$$G(x) = G(x; \boldsymbol{\theta}) := \frac{1}{c(\boldsymbol{\theta})} \mathbb{H}(x; \boldsymbol{\theta}) = \frac{\mathbb{H}(x; \boldsymbol{\theta})}{\mathbb{H}(\boldsymbol{\theta})}, \quad x > 0, \tag{4}$$

where $\mathbb{H}(x; \boldsymbol{\theta}) = \int_0^x y^{\theta_6} \exp\left(-\theta_1 y - (\theta_2 y_3^\theta + \theta_4)^{\theta_5}\right) \mathrm{d}y$ is the incomplete extreme value $\mathbb{H}$-function (for more details see Subsection 2.1). Figure 1 presents the behavior the PDFs of this random variable for some arbitrary values of its parameters.
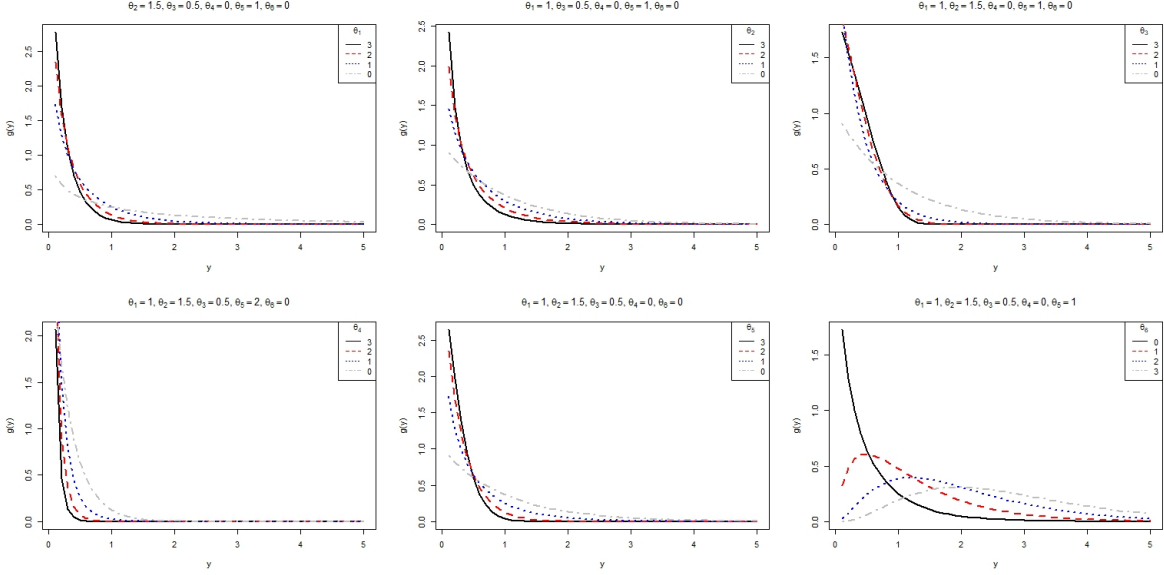

Figure 1: Plots of the PDF $g(y; \boldsymbol{\theta})$.

Remark 2.1. *Note that the PDF* (2) *of* $g(y; \boldsymbol{\theta})$ *defines a weighted gamma distribution by considering the gamma distribution with shape parameter* $\theta_6 + 1$ *and rate parameter* θ_1 *and the weight function* $w(y) = \exp(-(\theta_2 y^{\theta_3} + \theta_4)^{\theta_5})$, $y > 0$.

2.1 The incomplete extreme value $\mathbb{H}$-function

Consider the extreme value $\mathbb{H}$-function in (3). Define

$$h_y(\boldsymbol{\theta}) := y^{\theta_6} \exp\left(-\theta_1 y - (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5}\right), \; y > 0.$$

Then, the incomplete extreme value $\mathbb{H}-$function is defined by

$$\mathbb{H}(x; \boldsymbol{\theta}) := \int_0^x y^{\theta_6} \exp\left(-\theta_1 y - (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5}\right) \mathrm{d}y, \; x > 0. \tag{5}$$

Observe that $h_y(\boldsymbol{\theta}) = c(\boldsymbol{\theta}) g(y; \boldsymbol{\theta})$. Then, the CDF corresponding to $g(y; \boldsymbol{\theta})$ is given by

$$G(x; \boldsymbol{\theta}) = \frac{1}{c(\boldsymbol{\theta})} \int_0^x h_y(\boldsymbol{\theta}) \mathrm{d}y, \; x > 0.$$

When $\theta_5 = m \in \mathbb{N}$, it follows from Eq.(16) in [22] that

$$\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, m, \theta_6) = \sum_{n=0}^{\infty} \frac{(-1)^n \theta_4^{mn}}{n!} \sum_{k=0}^{mn} \binom{mn}{k} \left(\frac{\theta_2}{\theta_4}\right)^k \frac{\Gamma(\theta_6 + \theta_3 k + 1)}{\theta_1^{\theta_6 + \theta_3 k + 1}}. \tag{6}$$

thus, we can write (5) in terms of series as follows:

$$\mathbb{H}(x;\theta_1,\theta_2,\theta_3,\theta_4,m,\theta_6) = \sum_{n=0}^{\infty}\frac{(-1)^n\theta_4^{mn}}{n!}\sum_{k=0}^{mn}\left(\begin{array}{c}mn\\k\end{array}\right)\left(\frac{\theta_2}{\theta_4}\right)^k\frac{\gamma(\theta_6+\theta_3k+1,x)}{\theta_1^{\theta_6+\theta_3k+1}},\ x>0, \quad (7)$$

where $\gamma(p,x)$ is the incomplete gamma function $\gamma(p,x) = \int_0^x w^{p-1}\mathrm{e}^{-w}\mathrm{d}w$. Note that by taking $x\to\infty$ in (7) we get (6).

Another important function is $\int_x^\infty h_y(\boldsymbol{\theta})\mathrm{d}y$. It can be used, for example, to define the survival function and the Hazard function. Thus,

$$\int_x^\infty h_y(\boldsymbol{\theta})\mathrm{d}y = \mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,m,\theta_6) - \mathbb{H}(x;\theta_1,\theta_2,\theta_3,\theta_4,m,\theta_6),$$

which can be further manipulated, by considering (6) and (7), to obtain

$$\int_x^\infty h_y(\boldsymbol{\theta})\mathrm{d}y = \sum_{n=0}^{\infty}\frac{(-1)^n\theta_4^{mn}}{n!}\sum_{k=0}^{mn}\left(\begin{array}{c}mn\\k\end{array}\right)\left(\frac{\theta_2}{\theta_4}\right)^k\left[\frac{\Gamma(\theta_6+\theta_3k+1)-\gamma(\theta_6+\theta_3k+1,x)}{\theta_1^{\theta_6+\theta_3k+1}}\right],\ x>0.$$

2.2 Particular probability models

Some particular examples of densities $g(y;\boldsymbol{\theta})$ are presented in Table 1. Figure 2 presents the plots of the PDFs of such particular cases. There are other models that are not particular cases of (2), but it can be written as a finite mixture of $g(y;\boldsymbol{\theta})$ (see Subsection 2.5).

Table 1: Particular cases of PDF $g(y;\boldsymbol{\theta})$.

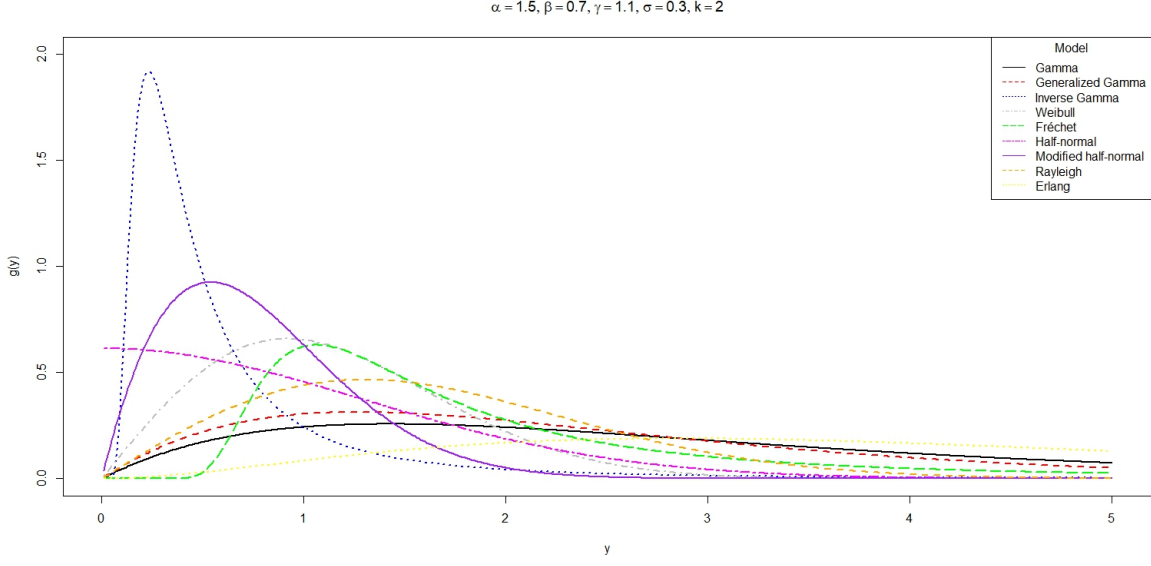| Distribution	PDF	$\boldsymbol{\theta}$
Gamma [9]	$\frac{\beta^\alpha}{\Gamma(\alpha)}y^{\alpha-1}\exp(-\beta y),\ \alpha,\beta>0$	$(\beta,0,1,0,1,\alpha-1)$
Generalized Gamma [26]	$\frac{\gamma\beta^{\alpha/\gamma}}{\Gamma(\alpha/\gamma)}y^{\alpha-1}\exp(-\beta y^\gamma),\ \alpha,\beta,\gamma>0$	$(0,\beta,\gamma,0,1,\alpha-1)$
Inverse Gamma [19]	$\frac{\beta^\alpha}{\Gamma(\alpha)}y^{-(\alpha+1)}\exp\left(-\beta\frac{1}{y}\right),\ \alpha,\beta>0$	$(0,\beta,-1,0,1,-\alpha-1)$
Weibull [9]	$\frac{\alpha}{\sigma}\left(\frac{y}{\sigma}\right)^{\alpha-1}\exp\left(-\left(\frac{y}{\sigma}\right)^\alpha\right),\ \alpha,\sigma>0$	$(0,1/\sigma,1,0,\alpha,\alpha-1)$
Fréchet [18]	$\frac{\alpha}{\sigma}\left(\frac{y}{\sigma}\right)^{-\alpha-1}\exp\left(-\left(\frac{y}{\sigma}\right)^{-\alpha}\right),\ \alpha,\sigma>0$	$(0,1/\sigma,1,0,-\alpha,-\alpha-1)$
Half-normal [9]	$\frac{\sqrt{2}}{\sigma\sqrt{\pi}}\exp\left(-\frac{y^2}{2\sigma^2}\right),\ \sigma>0$	$(0,\frac{1}{2\sigma^2},2,0,1,0)$
Modified half-normal [28]	$\frac{2\beta^{\alpha/2}}{\Psi(\alpha/2,\gamma/\sqrt{\beta})}y^{\alpha-1}\exp(\gamma x-\beta x^2),\ \alpha,\beta>0,\gamma<0$	$(-\gamma,\beta,2,0,1,\alpha-1)$
Rayleigh [9]	$\frac{1}{\sigma^2}y\exp\left(-\frac{1}{2\sigma^2}y^2\right),\ \sigma>0$	$(0,\frac{1}{2\sigma^2},2,0,1,1)$
Erlang [9]	$\frac{\beta^k}{(k-1)!}y^{k-1}\exp(-\beta y),\ \beta>0,\ k=1,2,\cdots$	$(\beta,0,1,1,1,k-1)$

Figure 2: Plots of the PDF $g(y; \boldsymbol{\theta})$ for particular models presented in Table 1.

It is interesting to highlight that Gumbel random variables are a limiting case of the new class of random variables hereby presented. In that case, one may notice that:

$$\exp(z) = \lim_{n \to \infty} \left(1 + \frac{z}{n}\right)^n \tag{8}$$

Thus, considering the PDF of Gumbel distribution with location parameter $\alpha \in \mathbb{R}$ and scale $\beta > 0$:

$$
\begin{aligned}
e^{-y/\beta + \alpha/\beta - e^{-y/\beta + \alpha/\beta}}/\beta &= e^{-y/\beta + \alpha/\beta - \lim_{n\to\infty}\left(1 + \frac{-y/\beta + \alpha/\beta}{n}\right)^n}/\beta \\
&= \lim_{n\to\infty} e^{-y/\beta + \alpha/\beta - \left(1 + \frac{\alpha}{n\beta} - \frac{y}{\beta n}\right)^n}/\beta \\
&= e^{\alpha/\beta} \lim_{n\to\infty} h_y(1/\beta, -1/\beta n, 1, 1 + \alpha/\beta n, n, 0)/\beta
\end{aligned}
\tag{9}
$$

2.3 Stochastic representation

Let Y be a random variable with PDF $g(\cdot; \boldsymbol{\theta})$ given in (2). In this subsection, we prove that Y has the following stochastic representation

$$Y \stackrel{d}{=} \sum_{n=0}^{\infty} X_n 1_{\{W_n = n\}}, \tag{10}$$

where "$\stackrel{d}{=}$" denotes equality in distribution, X_n and W_n are independent, X_n is an absolutely continuous random variable with CDF (for $x > 0$)

$$F_{X_n}(x) = \frac{E[(\theta_2 Z^{\theta_3} + \theta_4)^{n\theta_5} 1_{\{Z \le x\}}]}{E[(\theta_2 Z^{\theta_3} + \theta_4)^{n\theta_5}]}, \quad Z \sim \text{Gamma}(\theta_6 + 1, \theta_1) \tag{11}$$

and W_n is a discrete random variable with probability function

$$P(W_n = n) = \frac{(-1)^n}{n!} \frac{\theta_6 \Gamma(\theta_6) E[(\theta_2 Z^{\theta_3} + \theta_4)^{n\theta_5}]}{c(\boldsymbol{\theta}) \theta_1^{\theta_6 + 1}}, \quad n = 0, 1, \dots. \tag{12}$$

Indeed, by using the law of total probability and by (10) (for $y > 0$),

$$P(Y \leq y) = \sum_{k=0}^{\infty} P(Y \leq y | W_k = k) P(W_k = k)$$

$$= \sum_{k=0}^{\infty} P(X_k \leq y | W_k = k) P(W_k = k)$$

$$= \sum_{k=0}^{\infty} F_{X_k}(y) P(W_k = k), \tag{13}$$

where in the last line the independence of X_n and W_n was used. By employing (11) and (12) in (13) the CDF of Y is written as (for $y > 0$)

$$P(Y \leq y) = \frac{1}{c(\boldsymbol{\theta})} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} E[(\theta_2 Z^{\theta_3} + \theta_4)^{k\theta_5} 1_{\{Z \leq y\}}], \quad Z \sim \mathrm{Gamma}(\theta_6 + 1, \theta_1),$$

$$= \frac{\theta_6 \Gamma(\theta_6)}{c(\boldsymbol{\theta}) \theta_1^{\theta_6+1}} E\left[\exp\{-(\theta_2 Z^{\theta_3} + \theta_4)^{\theta_5}\} 1_{\{Z \leq y\}}\right]$$

$$= \frac{1}{c(\boldsymbol{\theta})} \int_0^y s^{\theta_6} \exp\left(-\theta_1 s - (\theta_2 s^{\theta_3} + \theta_4)^{\theta_5}\right) \mathrm{d}s = \frac{\mathbb{H}(y; \boldsymbol{\theta})}{\mathbb{H}(\boldsymbol{\theta})} = G(y),$$

where in the last identity we have used (4). We have thus completed the proof of (10).

Remark 2.2. *It is clear that when $\theta_5 \in \mathbb{N}$ the expectations in (11) and (12) become simple closed-forms depending on the CDF of the gamma distribution.*

2.4 Modality and shapes

A simple calculation shows that

$$\frac{d}{dy} g(y; \boldsymbol{\theta}) = \frac{g(y; \boldsymbol{\theta})}{c(\boldsymbol{\theta}) y} \left[\theta_6 - \theta_1 y - \theta_2 \theta_3 \theta_5 (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5 - 1} y^{\theta_3}\right] = 0$$

iff

$$\theta_6 - \theta_1 y = \theta_2 \theta_3 \theta_5 (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5 - 1} y^{\theta_3}. \tag{14}$$

Therefore, every critical point y of g is a positive solution of the above equation.

In what follows we analyze the number of roots of Equation (14) in the less trivial cases.

1. If $[\theta_2 = 0$ or $\theta_3 = 0$ or $\theta_5 = 0]$, $\theta_1 > 0$ and $\theta_6 > 0$ then $y = \theta_6/\theta_1$ is the unique solution of (14).

2. If $-1 < \theta_6 \leq 0$, $[\theta_1 \neq 0$ or $\theta_1 = 0]$ and $\mathrm{sign}(\theta_3) = \mathrm{sign}(\theta_5)$ then it is clear that (14) has no solution.

3. If $\theta_6 < -1$, $\theta_1 = 0$, $\mathrm{sign}(\theta_3) \neq \mathrm{sign}(\theta_5)$ and $\theta_2, \theta_4 > 0$ then (14) it is equivalent to

$$\frac{\theta_6}{\theta_2 \theta_3 \theta_5} = (\theta_2 z + \theta_4)^{\theta_5 - 1} z, \quad z = y^{\theta_3}. \tag{15}$$

First, let us assume that $\theta_5 > 0$. In this case, the function $z \mapsto r(z) = (\theta_2 z + \theta_4)^{\theta_5 - 1} z$ is increasing, and hence, (15) has a unique positive solution. On the other hand, assume that $\theta_5 < 0$. In this case, a routine calculation shows that $z \mapsto r(z)$ is unimodal, with maximum point $z_0 = -\theta_4/(\theta_2 \theta_5)$ and maximum value

$$r(z_0) = \left(\frac{1}{-\theta_5} + 1\right)^{\theta_5 - 1} \frac{\theta_4^{\theta_5}}{\theta_2(-\theta_5)}.$$

6

Imposing the condition $(1/(-\theta_5) + 1)^{\theta_5-1} \theta_4^{\theta_5} > -\theta_6/\theta_3$ we have $r(z_0) > \theta_6/(\theta_2\theta_3\theta_5)$, and therefore (15) has two positive solutions. Which implies that (14) also has two positive solutions. On the other hand, under condition $(1/(-\theta_5) + 1)^{\theta_5-1} \theta_4^{\theta_5} < -\theta_6/\theta_3$ we have $r(z_0) < \theta_6/(\theta_2\theta_3\theta_5)$, and therefore (15) has no positive solutions. Which implies that (14) also has no positive solutions.

4. If $\theta_6 > 0$, $[\theta_1 \neq 0$ or $\theta_1 = 0]$ and $\text{sign}(\theta_3) = \text{sign}(\theta_5)$ then the function $y \mapsto s(y) = (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5-1} y^{\theta_3}$ is increasing when $\theta_5 > 0$ and is unimodal when $\theta_5 < 0$. It is clear that in the case $\theta_5 > 0$ the equation (14) has a unique positive root. Following the steps in the previous item, it can be proven that, when $\theta_5 < 0$ the equation (14) has two or no positive solutions.

Since $\lim_{y\to\infty} g(y; \boldsymbol{\theta}) = 0$, the number of roots deduced in Items 1-4 above show that the density g in (2) has strictly decreasing, unimodal and decreasing-increasing-decreasing forms.

It is also possible to build a general solution to (14) where one of the roots can be expressed by using Lagrange's Inversion Theorem [10]. In that case, let y be defined as the following function of constant χ, function ϕ, and a parameter δ:

$$y = \chi + \delta\phi(y) \tag{16}$$

then any function $\zeta(y)$ is expressed as the following power series in δ:

$$\zeta(y) = \zeta(\chi) + \sum_{n=1}^{\infty} \frac{\delta^n}{n!} \frac{d^{n-1}}{dx^{n-1}} \left\{ \frac{d\zeta(x)}{dx} \phi^n(x) \right\} \Bigg|_{x=\chi}, \tag{17}$$

where suitable restrictions apply to make sure the series above converges. This way, equation (14) can be rearranged as:

$$y = \frac{\theta_6}{\theta_1} - \frac{\theta_2\theta_3\theta_5}{\theta_1}(\theta_2 y^{\theta_3} + \theta_4)^{\theta_5-1} y^{\theta_3}. \tag{18}$$

The direct comparison of (16), (17) and (18) reveals that, for $\zeta(y) = y$:

$$\begin{aligned}
y &= \frac{\theta_6}{\theta_1} + \sum_{n=1}^{\infty} \frac{(-1)^n \theta_2^n \theta_3^n \theta_5^n}{\theta_1^n n!} \frac{d^{n-1}}{dx^{n-1}} \left\{ (\theta_2 x^{\theta_3} + \theta_4)^{\theta_5 n - n} x^{\theta_3 n} \right\} \Bigg|_{x=\frac{\theta_6}{\theta_1}} \\
&= \frac{\theta_6}{\theta_1} - \frac{\theta_2\theta_3\theta_5\theta_6^{\theta_3}}{\theta_1^{\theta_3+1}} \left(\frac{\theta_2\theta_6^{\theta_3}}{\theta_1^{\theta_3}} + \theta_4 \right)^{\theta_5-1} + \cdots
\end{aligned} \tag{19}$$

under suitable constraints.

2.5 Finite mixtures

Some recently introduced distributions can be written as a finite mixture of g densities. These are the cases of a bimodal Weibull distribution–introduced by [31]– and the transmuted GEV distribution –introduced by [2].

It follows from (2) and Eq.(6) in [31] that the PDF of the bimodal Weibull can be written as:

$$\begin{aligned}
f(y; \alpha, \beta, \delta) &= c_0(\alpha, \beta, \delta)g(y; (0, 1/\beta, 1, 0, \alpha, \alpha-1)) + c_1(\alpha, \beta, \delta)g(y; (0, 1/\beta, 1, 0, \alpha, \alpha)) \\
&+ c_2(\alpha, \beta, \delta)g(y; (0, 1/\beta, 1, 0, \alpha, \alpha+1)),
\end{aligned}$$

where $c_j = cj(\alpha, \beta, \delta)$ are constants depending only on (α, β, δ) and $c_0 + c_1 + c_2 = 1$.

The PDF of TGEV model depends on $(\mu, \sigma, \gamma, \lambda) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R} \times [-1, 1]$. The support of the model is $(\mu - \sigma/\gamma, \infty)$, if $\gamma > 0$, or $(-\infty, \mu - \sigma/\gamma)$, if $\gamma < 0$, or $\mathbb{R}$, if $\gamma = 0$. We are interested in the case where TGEV is used to model positive data. In this case, we consider $\gamma > 0$ and $\mu = \sigma/\gamma$. It follows from (2) and Eq.(4) in [2] that the PDF of TGEV distribution is

$$f_2(y; \sigma, \gamma, \lambda) = c_0 g(y; (0, \gamma/\sigma, 1, 0, -1/\gamma, -1-1/\gamma)) + c_1 g(y; (0, 2^{-\gamma}\frac{\gamma}{\sigma}, 1, 0, -1/\gamma, -1-1/\gamma)),$$

where c_0 and c_1 are constants depending only on $(\sigma, \gamma, \lambda)$ and $c_0 + c_1 = 1$.

3 Some properties of the new class of distributions

In this section, we study several probabilistic properties of the density $g(y; \boldsymbol{\theta})$ given in (2).

3.1 Characteristic function, Mellin transform and moments

Let Y be a random variable with PDF $g(\cdot; \boldsymbol{\theta})$. The characteristic function of Y, denoted by $\varphi(t)$, is given by

$$\varphi(t) = \mathbb{E}_{\boldsymbol{\theta}}\left[e^{itY}\right] = \frac{\mathbb{H}(\theta_1 - it, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)}{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)}, \quad t \in \mathbb{R}.$$

where $\mathbb{E}_{\boldsymbol{\theta}}(\cdot)$ denotes the expected value under the density $g(\cdot; \boldsymbol{\theta})$.

The Mellin transform[1] is important in determining the distributions of the products and quotients of independent random variables. Furthermore, in the case of positive random variables, the Mellin transform of the PDF gives us the moments of the random variable. The Mellin transform of $g(\cdot; \boldsymbol{\theta})$ is given by

$$\{\mathcal{M}g\}(s) = \int_0^\infty y^{s-1} g(y; \boldsymbol{\theta}) \mathrm{d}y = \frac{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 + s - 1)}{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)}.$$

Then, real moments of Y of order r are given by

$$\mathbb{E}_{\boldsymbol{\theta}}[Y^r] = \{\mathcal{M}g\}(r+1) = \frac{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 + r)}{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)}. \tag{20}$$

3.2 Entropy

Let Y be a random variable with PDF $g(\cdot; \boldsymbol{\theta})$. We define the differential entropy of Y by

$$h(Y) := -\int_0^\infty g(y; \boldsymbol{\theta}) \log g(y; \boldsymbol{\theta}) \mathrm{d}y = -\mathbb{E}_{\boldsymbol{\theta}}[\log g(Y; \boldsymbol{\theta})], \tag{21}$$

where $\mathbb{E}_{\boldsymbol{\theta}}(\cdot)$ denotes the expected value under the density $g(\cdot; \boldsymbol{\theta})$. Substituting (2) in (21), we obtain

$$h(Y) = -\int_0^\infty \left[-\log c(\boldsymbol{\theta}) + \theta_6 \log y - \theta_1 y - (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5} \right] g(y; \boldsymbol{\theta}) \mathrm{d}y,$$

which implies

$$h(Y) = I_1 + I_2 + I_3 + I_4,$$

where

$$I_1 = \log c(\boldsymbol{\theta}),$$

$$I_2 = -\frac{\theta_6}{c(\boldsymbol{\theta})} \int_0^\infty (\log y) y^{\theta_6} \exp\left(-\theta_1 y - (\theta_2 y_3^\theta + \theta_4)^{\theta_5}\right) \mathrm{d}y,$$

$$I_3 = \frac{\theta_1}{c(\boldsymbol{\theta})} \int_0^\infty y^{\theta_6+1} \exp\left(-\theta_1 y - (\theta_2 y_3^\theta + \theta_4)^{\theta_5}\right) \mathrm{d}y = \theta_1 \frac{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 + 1)}{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)}$$

and

$$I_4 = \frac{1}{c(\boldsymbol{\theta})} \int_0^\infty (\theta_2 y^{\theta_3} + \theta_4)^{\theta_5} y^{\theta_6} \exp\left(-\theta_1 y - (\theta_2 y_3^\theta + \theta_4)^{\theta_5}\right) \mathrm{d}y.$$

It remains to find closed expressions for I_2 and I_4. For this, consider the case $\theta_5 = m \in \mathbb{N}$. Firstly, observe that $I_2 = -[\theta_6/c(\boldsymbol{\theta})]\partial_{\theta_6}\mathbb{H}(\boldsymbol{\theta})$, where ∂_{θ_6} denotes the partial derivative with respect to θ_6. It follows from Eq.(16) in [22] that

$$\partial_{\theta_6}\mathbb{H}(\boldsymbol{\theta}) = \frac{1}{\theta_1} \sum_{n=0}^\infty \frac{(-\theta_4^m)^n}{n!} \sum_{k=0}^{mn} \frac{(mn)!}{(mn-k)!k!} \left(\frac{\theta_2}{\theta_4 \theta_1^{\theta_3}}\right)^k \partial_{\theta_6}\left\{\theta_1^{-\theta_6}\Gamma(\theta_6 + 1 + \theta_3 k)\right\}.$$

---

[1] We may refer to [25] for further details on the properties of Mellin transforms in the context of the algebra of random variables.

Observe that, for $\theta_1 > 0$,

$$\partial_{\theta_6}\left\{\theta_1^{-\theta_6}\Gamma(\theta_6+1+\theta_3 k)\right\} = \theta_1^{-\theta_6}\Gamma(\theta_6+1+\theta_3 k)\left[\Psi(\theta_6+1+\theta_3 k)-\log\theta_1\right],$$

where $\partial_z\Gamma(z) = \Gamma(z)\Psi(z)$. Then,

$$I_2 = -\frac{\theta_6\theta_1^{-\theta_6-1}}{c(\boldsymbol{\theta})}\sum_{n=0}^{\infty}\frac{(-\theta_4^m)^n}{n!}\sum_{k=0}^{mn}\frac{(mn)!}{(mn-k)!k!}\left(\frac{\theta_2}{\theta_4\theta_1^{\theta_3}}\right)^k\Gamma(\theta_6+1+\theta_3 k)\left[\Psi(\theta_6+1+\theta_3 k)-\log\theta_1\right].$$

Now, for I_4, taking $\theta_5 = m \in \mathbb{N}$, it follows from Binomial expansion that

$$I_4 = \frac{1}{c(\boldsymbol{\theta})}\theta_4^m\sum_{k=0}^{m}\binom{m}{k}\left(\frac{\theta_2}{\theta_4}\right)^k\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,m,\theta_3 k+\theta_6).$$

3.3 Kullback-Leibler Divergence

The Kullback-Leibler divergence (see [5]) is useful to measure the difference between two probability distributions. If Y_1 and Y_2 are two random variables with PDFs $g_{Y_1}(y;\boldsymbol{\theta})$ and $g_{Y_2}(y;\boldsymbol{\theta}')$, respectively, where $\boldsymbol{\theta} = (\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6)$, $\boldsymbol{\theta}' = (\theta_1',\theta_2,\theta_3,\theta_4,\theta_5,\theta_6')$, $\theta_1 \neq \theta_1'$ and $\theta_6 \neq \theta_6'$, then their Kullback-Leibler divergence has the form

$$D_{\mathrm{KL}}(g_{Y_1}\|g_{Y_2}) = \int_0^{\infty}g_{Y_1}(y;\boldsymbol{\theta})\log\left(\frac{g_{Y_1}(y;\boldsymbol{\theta})}{g_{Y_2}(y;\boldsymbol{\theta}')}\right)\mathrm{d}y.$$

From (2) the above integral is

$$= \log(c(\boldsymbol{\theta}') - \log(c(\boldsymbol{\theta}) + (\theta_6-\theta_6')\int_0^{\infty}\log(y)g_{Y_1}(y;\boldsymbol{\theta})\mathrm{d}y + (\theta_1'-\theta_1)\int_0^{\infty}yg_{Y_1}(y;\boldsymbol{\theta})\mathrm{d}y$$

$$= \log(c(\boldsymbol{\theta}') - \log(c(\boldsymbol{\theta}) + (\theta_6-\theta_6')E_{\boldsymbol{\theta}}[\log(Y)] + (\theta_1'-\theta_1)E_{\boldsymbol{\theta}}[Y]$$

$$= \log(c(\boldsymbol{\theta}') - \log(c(\boldsymbol{\theta}) + (\theta_6-\theta_6')\xi_Y + (\theta_1'-\theta_1)\frac{\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6+1)}{\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6)},$$

where $\xi_Y := E_{\boldsymbol{\theta}}[\log(Y)]$ and in the last equality the formula (20) of $E_{\boldsymbol{\theta}}[Y]$ was used. That is,

$$D_{\mathrm{KL}}(g_{Y_1}\|g_{Y_2}) = \log(c(\boldsymbol{\theta}') - \log(c(\boldsymbol{\theta}) + (\theta_6-\theta_6')\xi_Y + (\theta_1'-\theta_1)\frac{\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6+1)}{\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6)}. \tag{22}$$

By using the well-known inequalities $1-x^{-1} < \log(x) \leq x-1$, for $x > 0$, from (20) we get the finiteness of ξ_Y, more precisely,

$$1 - \frac{\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6-1)}{\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6)} < \xi_Y \leq \frac{\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6+1)}{\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6)} - 1.$$

Remark 3.1. *By using the entropy formula (see Subsection 3.2) and Kullback-Leibler divergence formula in (22), we can obtain a closed-form expression for the cross entropy, denoted by $H(g_{Y_1},g_{Y_2})$, since $D_{\mathrm{KL}}(g_{Y_1}\|g_{Y_2}) = H(g_{Y_1},g_{Y_2}) - h(Y_1)$.*

4 Estimation

In this section, we present two classes of estimators for the model $g(y;\boldsymbol{\theta})$ (2).

4.1 Exponential family

Consider the case $\theta_5 = m \in \mathbb{N}$ being a known parameter and $\theta_3 = 1$. It follows from Binomial expansion that

$$g(y;\boldsymbol{\theta}) = \exp\left(-\sum_{k=0}^{m}\binom{m}{k}\theta_2^k\theta_4^{m-k}y^k - \theta_1 y + \theta_6\log y - \log c(\boldsymbol{\theta})\right)$$

$$= \exp\left(\sum_{j=0}^{m+2}a_j(\boldsymbol{\theta})T_j(y) + b(\boldsymbol{\theta})\right),$$

where

$$
a_j(\boldsymbol{\theta}) =
\begin{cases}
-\dbinom{m}{j}\theta_2^j\theta_4^{m-j}, & j = 0, 1, \cdots, m, \\
-\theta_1, & j = m+1, \\
\theta_6, & j = m+2,
\end{cases}
$$

$$
T_j(y) =
\begin{cases}
y^j, & j = 0, 1, \cdots, m, \\
y, & j = m+1, \\
\log y, & j = m+2,
\end{cases}
$$

and $b(\boldsymbol{\theta}) = -\log c(\boldsymbol{\theta})$. That means, $g(y;\boldsymbol{\theta})$ belongs to exponential family, provided that $\theta_3 = 1$ and $\theta_5 = m$.

4.2 Least squares estimation

Let $\mathbf{Y} = (Y_1, \cdots, Y_n)$ be a random sample of $g(\cdot; \boldsymbol{\theta})$. Consider the empirical CDF (ECDF) $\hat{G}(y)$ defined as

$$
\hat{G}(y) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{Y_i \leq y\}},
$$

where $\mathbb{1}_A$ denotes the indicator of the set A.

A widely used estimation method in linear models is the least squares estimator (LSE) (cf. [24]). In our case, we aim to estimate $\boldsymbol{\theta}$ by comparing the theoretical CDF $G(y; \boldsymbol{\theta})$, given in (4), with the ECDF $\hat{G}(y)$. To realize this, we minimize the quadratic loss function over a search space $\Theta_0 \subset \Theta$. Thus, the LSE, $\hat{\boldsymbol{\theta}} = \hat{\theta}_{LSE}(\mathbf{Y})$, is given by

$$
\hat{\boldsymbol{\theta}} \in \arg\min \left\{ \sum_{j=1}^{n} \left[G(Y_j; \boldsymbol{\theta}) - \hat{G}(Y_j) \right]^2 \right\}, \tag{23}
$$

with respect to $\boldsymbol{\theta} \in \Theta_0$.

4.3 Maximum likelihood estimation

Let $\mathbf{Y} = (Y_1, \cdots, Y_n)$ be a random sample of $g(\cdot; \boldsymbol{\theta})$. The (random) log-likelihood function is given by

$$
\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}; \mathbf{Y}) = -n \log c(\boldsymbol{\theta}) + \theta_6 \sum_{j=1}^{n} \log Y_j - \theta_1 \sum_{j=1}^{n} Y_j - \sum_{j=1}^{n} \left(\theta_2 Y_j^{\theta_3} + \theta_4 \right)^{\theta_5}. \tag{24}
$$

The maximum likelihood estimator (MLE) is given by

$$
\hat{\boldsymbol{\theta}}_n^{MLE} \in \arg\max_{\theta} \ell(\boldsymbol{\theta}). \tag{25}
$$

Observe that (24) is a continuous function of $\boldsymbol{\theta}$. Then, on a compact subset $\Theta_0 \subset \Theta$, there exists a maximum likelihood estimator. Computationally, (24) is enough to find the MLE. However, traditional methods of finding MLE via the gradient vector of $\ell(\boldsymbol{\theta})$ can be used. For a more detailed discussion on conditions under which the MLE exists, we refer the reader to [24].

4.3.1 Partial derivatives of the log-likelihood function

Possible candidates for the MLE are the vectors $\boldsymbol{\theta} \in \Theta$ that satisfy the likelihood equation (cf. [24])

$$
\partial_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}) = (\partial_{\theta_1} \ell(\boldsymbol{\theta}), \dots, \partial_{\theta_6} \ell(\boldsymbol{\theta})) = (0, \dots, 0).
$$

Since the log-likelihood function (24) depends on $c(\boldsymbol{\theta}) = \mathbb{H}(\boldsymbol{\theta})$, its derivatives, $\partial_{\boldsymbol{\theta}} c(\boldsymbol{\theta})$, are needed to explicitly express $\partial_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta})$. Moreover, power functions of the form x^{θ_5} appear in (24). To simplify the derivatives, we obtain the gradient of the likelihood function when $\theta_5 = m \in \mathbb{N}$. In this case, the MLE can be obtained using numerical procedures for solving the nonlinear system (cf. [4]):

$$
\partial_{\theta_i} \ell(\boldsymbol{\theta}) = 0, \quad i = 1, \cdots, 6. \tag{26}
$$

Remark 4.1. *In the derivatives $\partial_{\theta_i} \ell(\boldsymbol{\theta})$ below, we require $\theta_5 = m \in \mathbb{N}$. In computational terms of applying numerical methods to solve the system (26), there is no guarantee that $\hat{\theta}_5 \in \mathbb{N}$. We could overcome this problem by adding a projection step in the update rule of the numerical algorithm. That means, we could use the corrected estimator*

$$\tilde{\theta}_5 = \arg \inf_{n \in \mathbb{N}} \left| \hat{\theta}_5 - n \right|.$$

We have the following expressions:

$$\partial_{\theta_1} c(\boldsymbol{\theta}) = -\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 - 1), \tag{27}$$

$$\partial_{\theta_2} c(\boldsymbol{\theta}) = -\theta_5 \sum_{k=1}^{m-1} \binom{m-1}{k} \theta_2^k \theta_4^{m-1-k} \mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 + \theta_3(k+1)), \tag{28}$$

$$
\begin{aligned}
\partial_{\theta_3} c(\boldsymbol{\theta}) = & -\theta_5 \theta_1^{\theta_6 + \theta_3 - 1} \sum_{k=1}^{m-1} \binom{m-1}{k} \theta_2^{k+1} \theta_4^{m-1-k} \\
& \times \left[\sum_{l=0}^{\infty} \frac{(-\theta_4^m)^l}{l!} \sum_{r=0}^{ml} \binom{ml}{r} \frac{\theta_2^r}{\theta_4^r \theta_1^{3r}} \Gamma(\theta_6 + \theta_3 + 1 - \theta_3 r) \left(\Psi(\theta_6 + \theta_3 + 1 - \theta_3 r) - \log \theta_1 \right) \right]
\end{aligned} \tag{29}
$$

$$\partial_{\theta_4} c(\boldsymbol{\theta}) = -m \sum_{k=0}^{m-1} \binom{m-1}{k} \theta_2^k \theta_4^{m-1-k} \mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 + \theta_3 k). \tag{30}$$

A computable representation of the extreme value $\mathbb{H}$-function, when $\theta_1 > 0$ and $\theta_2 \geq 0$, is

$$\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6) \approx \frac{1}{\theta_1^{1+\theta_6}} \sum_{j=1}^{N} w_{j,\theta_6} \exp\{-(\theta_2 \theta_1^{-\theta_3} x_j^{\theta_3} + \theta_4)^{\theta_5}\}, \tag{31}$$

where x_j is the j-th root of the generalized Laguerre polynomials $L_N^{(\alpha)}(x)$ and the weight $w_{j,\alpha}$ is given by [1, 27]:

$$w_{j,\alpha} = \frac{\Gamma(N + \alpha + 1)x_j}{N!(N+1)^2 [L_{N+1}^{(\alpha)}(x_j)]^2}. \tag{32}$$

This way, the derivative with respect to θ_5 can be given as:

$$\partial_{\theta_5} c(\boldsymbol{\theta}) \approx -\frac{1}{\theta_1^{1+\theta_6}} \sum_{j=1}^{N} \frac{w_{j,\theta_6} (\theta_2 \theta_1^{-\theta_3} x_j^{\theta_3} + \theta_4)^{\theta_5} \log\left(\theta_2 \theta_1^{-\theta_3} x_j^{\theta_3} + \theta_4\right)}{\exp\{(\theta_2 \theta_1^{-\theta_3} x_j^{\theta_3} + \theta_4)^{\theta_5}\}}. \tag{33}$$

In particular, if $\theta_5 = m \in \mathbb{N}$,

$$\partial_{\theta_5} c(\boldsymbol{\theta}) \approx -\frac{1}{\theta_1^{1+\theta_6}} \sum_{j=1}^{N} \sum_{k=0}^{m} \binom{m}{k} \theta_2^k \theta_1^{-\theta_3 k} \theta_4^{m-k} x_j^k \frac{w_{j,\theta_6} \log\left(\theta_2 \theta_1^{-\theta_3} x_j^{\theta_3} + \theta_4\right)}{\exp\{(\theta_2 \theta_1^{-\theta_3} x_j^{\theta_3} + \theta_4)^m\}}. \tag{34}$$

Finally,

$$\partial_{\theta_6} c(\boldsymbol{\theta}) = \theta_1^{-\theta_6 - 1} \sum_{n=0}^{\infty} \frac{(-\theta_4^m)^n}{n!} \sum_{k=0}^{mn} \frac{(mn)!}{(mn-k)!k!} \left(\frac{\theta_2}{\theta_4 \theta_1^{\theta_3}}\right)^k \Gamma(\theta_6 + 1 + \theta_3 k) \left[\Psi(\theta_6 + 1 + \theta_3 k) - \log \theta_1\right]. \tag{35}$$

Now, we are able to describe the score functions:

$$\partial_{\theta_1} \ell(\boldsymbol{\theta}) = n \frac{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6 - 1)}{\mathbb{H}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)} - \sum_{j=1}^{n} Y_j,$$

11

$$\partial_{\theta_2}\ell(\boldsymbol{\theta}) = -\frac{n}{c(\boldsymbol{\theta})}\left[-\theta_5\sum_{k=0}^{m-1}\binom{m-1}{k}\theta_2^k\theta_4^{m-1-k}\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6+\theta_3(k+1))\right]$$

$$- \theta_5\sum_{j=1}^{n}\left(\theta_2 Y_j^{\theta_3}+\theta_4\right)^{\theta_5-1}Y_j^{\theta_3},$$

$$\partial_{\theta_3}\ell(\boldsymbol{\theta}) = -\frac{n}{c(\boldsymbol{\theta})}\partial_{\theta_3}c(\boldsymbol{\theta}) - m\theta_2\sum_{j=1}^{n}\sum_{k=0}^{m-1}\binom{m-1}{k}\theta_2^k\theta_4^{m-1-k}Y_j^{\theta_3(k+1)}\log Y_j,$$

where $\partial_{\theta_3}c(\boldsymbol{\theta})$ is given in (29).

$$\partial_{\theta_4}\ell(\boldsymbol{\theta}) = -m\sum_{k=0}^{m-1}\binom{m-1}{k}\theta_2^k\theta_4^{m-1-k}\mathbb{H}(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6+\theta_3 k),$$

$$\partial_{\theta_5}\ell(\boldsymbol{\theta}) \approx -\frac{n}{c(\boldsymbol{\theta})}\left[\frac{1}{\theta_1^{1+\theta_6}}\sum_{j=1}^{N}\sum_{k=0}^{m}\binom{m}{k}\theta_2^k\theta_1^{-\theta_3 k}\theta_4^{m-k}x_j^k\frac{w_{j,\theta_6}\log\left(\theta_2\theta_1^{-\theta_3}x_j^{\theta_3}+\theta_4\right)}{\exp\{(\theta_2\theta_1^{-\theta_3}x_j^{\theta_3}+\theta_4)^m\}}\right]$$

$$- \sum_{j=1}^{n}\sum_{k=0}^{m}\binom{m}{k}\theta_2^k\theta_4^{m-k}Y_j^{\theta_3 k}\log(\theta_2 Y_j^{\theta_3}+\theta_4),$$

and

$$\partial_{\theta_6}\ell(\boldsymbol{\theta}) = -\frac{n}{c(\boldsymbol{\theta})}\partial_{\theta_6}c(\boldsymbol{\theta}) + \sum_{j=1}^{n}\log Y_j,$$

where $\partial_{\theta_6}c(\boldsymbol{\theta})$ is given in (35).

Remark 4.2. *The gradient $\nabla\ell(\boldsymbol{\theta})$ can be computed for the general case $\theta_5 \in \mathbb{R}$ using approximation* (31).

4.4 Computational implementation

4.4.1 Choosing a good initial guess

While the methods used for estimation may be straightforward in the sense that they rely on standard estimation techniques, the presence of integral expressions in the PDF means that the computational implementation/optimization of models involving special functions is not always simple. Optimizers in languages such as Python [30] and R [17] can easily become inefficient if a well-planned strategy for implementing objective functions is not adopted. A good initial guess will be important for efficient convergence of the implemented methods.

Example 4.3. *Let $\mathbf{Y} = (Y_1,\ldots,Y_n)$ be a sample from the distribution $g(\cdot;\boldsymbol{\theta})$. Suppose we estimate $(\hat{\alpha},\hat{\beta})$ by modeling $\mathbf{Y}$ as a $Gamma(\alpha,\beta)$ distribution and find that this model fits the data well. Then, from Table 1, we obtain the initial estimate $\hat{\boldsymbol{\theta}}_0 = (\hat{\beta},0,1,1,\hat{\alpha}-1)$ for the parameter vector $\boldsymbol{\theta}$ in the general model $g(\cdot;\boldsymbol{\theta})$.*

In the next step, we can use $\hat{\boldsymbol{\theta}}_0$ as an initial guess to refine our estimation of $\boldsymbol{\theta}$ without being constrained to the specific Gamma model. This procedure can also be applied using other distributions listed in Table 1.

Although simple, Example 4.3 provides a powerful tool for obtaining initial guesses, as the distributions in Table 1 are mostly available in standard Python [30] and R [17] libraries. Another effective approach for obtaining an initial guess is to use the result of one estimation method (such as LSE, method of moments, methods based on characteristic function or Mellin transform) as the starting point for another method (such as MLE). This will be discussed in the next subsection.

Finally, a different option for choosing the initial guess involves defining a grid of values and testing various initial values within a reasonable range. The estimation method's performance is then evaluated for each of these values. The aim is to identify the value from the grid that leads to the most accurate final estimate, such as the one with the best convergence or smallest error. This process can be carried out empirically by observing which initial guess produces the most favorable result.

4.4.2 Algorithms for estimations

The parameter estimation was done by optimization procedures according to Equation (23) for the LSE and to Equation (25) for MLE.

The large number of parameters demanded by extreme value $\mathbb{H}$-function class of distributions lead to frequent convergence problems when the MLE approach was tried even using particular cases of Table 1 as initial guesses.

Thus, we decided to use LSE in Algorithm 1 as a first step in the parameter estimation due to its better convergence and try to refine the estimation by using a LSE-tuple as an initial guess to the MLE estimator of Algorithm 2.

Algorithm 1 Estimation of parameters $\boldsymbol{\theta}$ using the LSE

    Input: Data $(Y_1, \cdots, Y_n)$ sampled from the $g(\cdot; \boldsymbol{\theta})$.
    Output: Estimates $\widehat{\theta_Y}$.
1: Compute estimates $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$, and $\hat{\sigma}$ to fit the particular case chosen from Table 1.
2: Set the initial guess for the numerical optimization to be $\boldsymbol{\theta}_0 = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$ also according to Table 1.
3: Calculate $\widehat{\theta_Y}$ using (23).
4: **Return** $\widehat{\theta_Y}$.

Algorithm 2 Estimation of parameters $\boldsymbol{\theta}$ using the MLE

    Input: Data $(Y_1, \cdots, Y_n)$ sampled from the $g(\cdot; \boldsymbol{\theta})$.
    Output: Estimates $\hat{\theta}^{MLE}$.
1: Set the initial guess for the numerical optimization to be $\boldsymbol{\theta}_0 = \widehat{\theta_Y}$ from Algorithm 1.
2: Calculate $\hat{\theta}^{MLE}$ using (25).
3: **Return** $\hat{\theta}^{MLE}$.

In the next section, we apply Algorithms 1 and 2 to model real data. The algorithms were implemented using the Python language [30] and the optimizers `optimize.scipy` [32] and `hyperopt` [3]. Furthermore, the R software [17] was used to obtain descriptive statistics and the figures presented throughout this article (with the exception of Figure 4, which was generated in Python). All the code and a tutorial on how to use it are available at https://github.com/eip-unb/positive_support_asymmetric .

5 Applications

In order to show the versatility of extreme value $\mathbb{H}$-function class of distributions, we show that the model was able to fit three datasets with distinct characteristics. Until now, these datasets had been effectively modeled by three different and competing models [11, 18, 29]. Here, we unify them into the model (4). Even when accounting for penalties due to the increased number of parameters, the new model is at least as effective as its particular cases, demonstrating the versatility of the model.

The study by [18] shown that Frechét distributions fitted better the minimum monthly flows of water (m^3/s) on the Piracicaba Riber, located in São Paulo state, Brazil, from 1960 to 2014. We took the minimum monthly flow for September as presented below:

$$
\begin{aligned}
X \quad = \quad & (29.19, 8.49, 7.37, 82.93, 44.18, 13.82, 22.28, 28.06, 6.84, 12.14, \\
& 153.78, 17.04, 13.47, 15.43, 30.36, 6.91, 22.12, 35.45, 44.66, 95.81, \\
& 6.18, 10.00, 58.39, 24.05, 17.03, 38.65, 47.17, 27.99, 11.84, 9.60, \\
& 6.72, 13.74, 14.60, 9.65, 10.39, 60.14, 15.51, 14.69, 16.44).
\end{aligned}
$$

The strength of carbon fibers tested under stressing tensions is frequently used in the literature [29] and we take the strength data measured in GPa (Gigapascal) for single carbon fibers of 20 mm

length as presented below

$$
\begin{aligned}
Y \quad = \quad & (1.312, 1.314, 1.479, 1.552, 1.700, 1.803, 1.861, 1.865, 1.944, 1.958, 1.966, \\
& 1.977, 2.006, 2.021, 2.027, 2.055, 2.063, 2.098, 2.140, 2.179, 2.224, 2.240, \\
& 2.253, 2.270, 2.272, 2.274, 2.301, 2.301, 2.359, 2.382, 2.382, 2.426, 2.434, \\
& 2.435, 2.478, 2.490, 2.511, 2.514, 2.535, 2.554, 2.566, 2.570, 2.586, 2.629, \\
& 2.633, 2.642, 2.648, 2.684, 2.697, 2.726, 2.770, 2.773, 2.800, 2.809, 2.818, \\
& 2.821, 2.848, 2.880, 2.954, 3.012, 3.067, 3.084, 3.090, 3.096, 3.128, 3.233, \\
& 3.433, 3.585, 3.585)
\end{aligned}
$$

to be modelled as a Weibull RV.

Life data are sometimes modeled with the Gamma distribution and [11] or [13] discuss applications of the gamma distribution to life data. The following dataset represents failure times of machine parts, some of which are manufactured by two manufacturers and were merged after a previous test which rejected the hypothesis of significant differences in the part life for the two manufacturers.

$$
\begin{aligned}
Z \quad = \quad & (620, 470, 260, 89, 388, 242, 103, 100, 39, 460, 284, 1285, 218, 393, 106, \\
& 158, 152, 477, 403, 103, 69, 158, 818, 947, 399, 1274, 32, 12, 134, 660, 548, \\
& 381, 203, 871, 193, 531, 317, 85, 1410, 250, 41, 1101, 32, 421, 32, 343, 376, \\
& 1512, 1792, 47, 95, 76, 515, 72, 1585, 253, 6, 860, 89, 1055, 537, 101, 385, \\
& 176, 11, 565, 164, 16, 1267, 352, 160, 195, 1279, 356, 751, 500, 803, 560, 151, \\
& 24, 689, 1119, 1733, 2194, 763, 555, 14, 45, 776, 1, 1747, 945, 12, 1453, 14, \\
& 150, 20, 41, 35, 69, 195, 89, 1090, 1868, 294, 96, 618, 44, 142, 892, 1307, 310, \\
& 230, 30, 403, 860, 23, 406, 1054, 1935, 561, 348, 130, 13, 230, 250, 317, 304, \\
& 79, 1793, 536, 12, 9, 256, 201, 733, 510, 660, 122, 27, 273, 1231, 182, 289, \\
& 667, 761, 1096, 43, 44, 87, 405, 998, 1409, 61, 278, 407, 113, 25, 940, 28, \\
& 848, 41, 646, 575, 219, 303, 304, 38, 195, 1061, 174, 377, 388, 10, 246, 323, \\
& 198, 234, 39, 308, 55, 729, 813, 1216, 1618, 539, 6, 1566, 459, 946, 764, \\
& 794, 35, 181, 147, 116, 141, 19, 380, 609, 546).
\end{aligned}
$$

Descriptive statistics for the X, Y and Z RVs are presented in Table 2. For each dataset, the following summary statistics are given: the minimum value observed in the dataset (Min.); the first quartile value (1st Qu.); the median value (Median); the average value (Mean); the third quartile value (3rd Qu.); the maximum value observed (Max.); the standard deviation (Sd.); the coefficient of Symmetry (CS); the coefficient of Kurtosis (CK); the sample size (n).

Table 2: Descriptive statistics for the RVs X, Y and Z.

Dataset	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd	CS	CK	n
X	6.180	11.115	16.440	28.285	32.905	153.78	29.319	2.451	6.785	39
Y	1.312	2.098	2.478	2.451	2.773	3.585	0.495	-0.027	-0.148	69
Z	1	96	304	463.647	667	2194	476.956	1.358	1.250	201

The descriptive statistics (Table 2) and boxplots (Figure 3) show that the distributions of data X and Z are right-skewed. This supports the initial choice of using the Fréchet, Weibull, and Gamma models for each case, respectively.
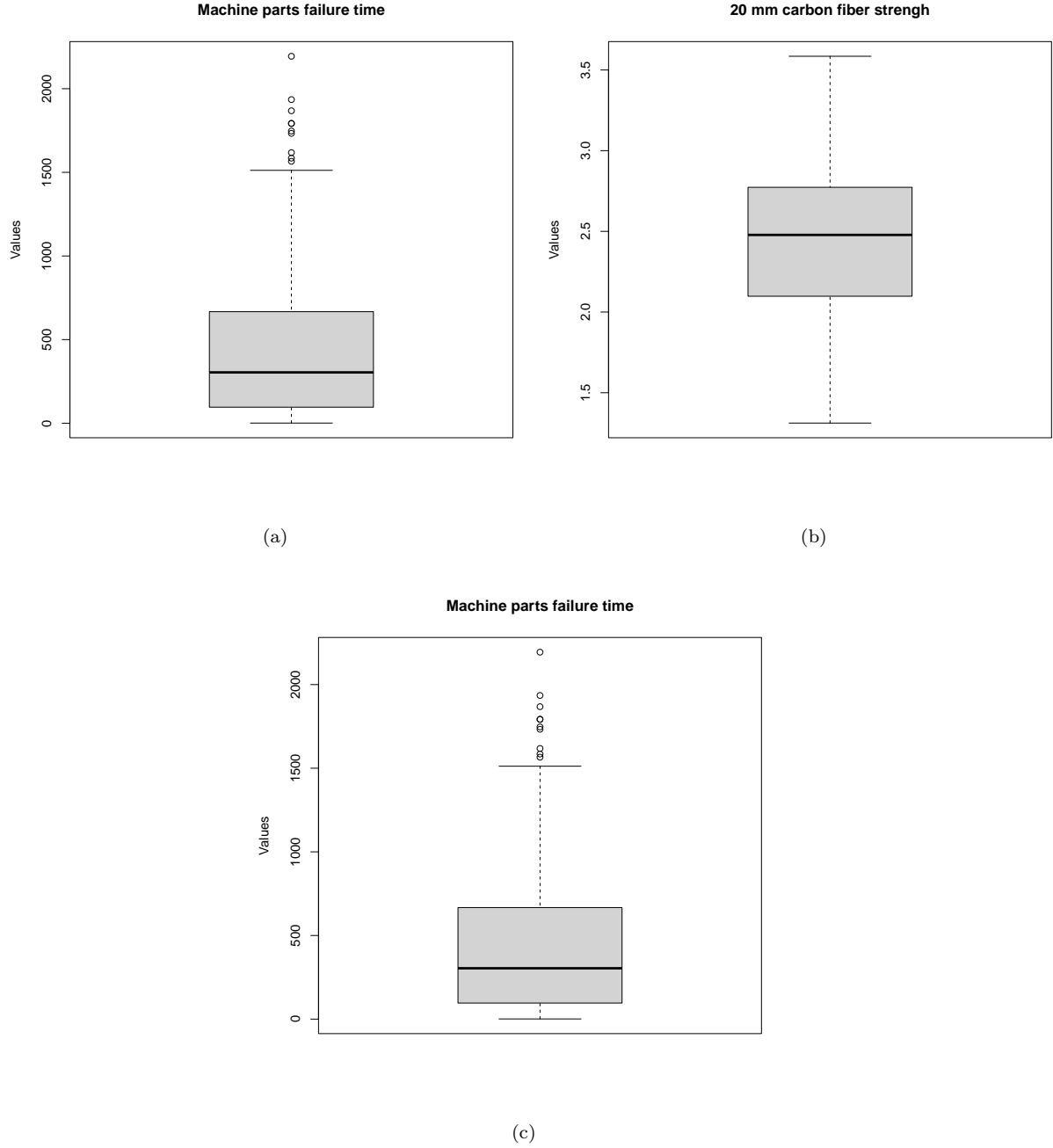
(a)

(b)

(c)

Figure 3: Boxplot for (a) X, (b) Y and (c) Z.

We aim to verify that the newly developed model in this study not only generalizes the existing models mentioned earlier but also provides a better fit to the data. Figure 4 shows the fit of distributions to dataset, meanwhile the estimated parameters of the previous modeling reported by [18, 29] were $(\hat{\alpha}, \hat{\sigma})$ equal to (1.564, 13.76) for X (Fréchet model), (5.505, 2.651) for Y (Weibull model) and (0.8274, 560.35) for Z (Gamma model).

Figure 4: Empirical and fitted extreme value $\mathbb{H}$-function distribution functions for (a) X, (b) Y and (c) Z.

According to Table 1, the Fréchet RV X can be expressed by a distribution of the extreme value $\mathbb{H}$-function class with the parameter set $\boldsymbol{\theta}_{X_0} = (0, 0.0727, 1, 0, -1.564, -2.564)$. We take $\boldsymbol{\theta}_{X_0}$ as a initial guess and estimate the parameter set $\widehat{\boldsymbol{\theta}_X}$ which better fits to data. The parameters estimation was done by means of an optimization procedure by minimizing the distance (mean squared error) between the empirical cumulative density function (ECDF) and the CDF (see Algorithm 1).

The parameter set $\hat{\theta}_X^{MLE} = (0.021, 0.0008, 3.973, 0.007, -9.51, -0.899)$ yields a better fit as supported by the information criteria such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Efficient Determination Criterion (EDC). These estimates were obtained using Algorithm 2. The same procedure were used to get $\widehat{\theta_Y}$ and $\widehat{\theta_Z}$ (respectively, $\hat{\theta}_Y^{MLE}$ and $\hat{\theta}_Z^{MLE}$), which are presented on Table 3.

Table 3: Estimated parameters and information criteria for model selection. $\mathbb{H}$-extreme distribution outperforms particular cases.

Dataset	PDF	$\widehat{\theta}$	AIC	BIC	EDC
X	Fréchet	$\boldsymbol{\theta}_{X_0} = (0, 0.0727, 1, 0, -1.564, -2.564)$	333.70	343.68	313.9
	$\mathbb{H}$-extreme	$\widehat{\boldsymbol{\theta}_X} = (0.013, 0.364, 0.517, 0.0006, -5.089, -1.446)$	332.34	342.32	312.55
		$\hat{\theta}_X^{MLE} = (0.021, 0.0008, 3.973, 0.007, -9.51, -0.899)$	**328.89**	**338.87**	**309.10**
Y	Weibull	$\boldsymbol{\theta}_{Y_0} = (0, 0.377, 1, 0, 5.50, 4.50)$	111.19	124.60	90.53
	$\mathbb{H}$-extreme	$\widehat{\boldsymbol{\theta}_Y} = (0.002, 0.5, 0.80, 0.03, 5.44, 6.12)$	110.35	123.75	89.69
		$\hat{\theta}_X^{MLE} = (0.0, 0.657, 0.621, 0.0, 5.394, 7.150)$	**109.75**	**123.15**	**89.09**
Z	Gamma	$\boldsymbol{\theta}_{Z_0} = (0.0018, 0, 1, 0, 1, -0.1726)$	2876.85	2896.67	2854.84
	$\mathbb{H}$-extreme	$\widehat{\boldsymbol{\theta}_Z} = (0.0015, 7.17, -6.55, 0.0018, 8.57, -0.29)$	2865.10	2884.92	2843.09
		$\hat{\theta}_Z^{MLE} = (0.0017, 7.17, -6.55, 0.005, 8.57, -0.195)$	**2863.10**	**2882.92**	**2841.09**

Table 3 summarizes the improvements achieved through estimations using both LSE (Algorithm 1) and MLE (Algorithm 2). For all the datasets, the modeling gains of using MLE instead of LSE were validated by the information criteria (AIC, BIC, EDC). Such findings are justified by the fact that MLE explicitly incorporates the probabilistic structure of the data through the likelihood function. By maximizing the likelihood, MLE identifies parameter values that make the observed data most probable under the assumed model, leading to estimators that are typically unbiased, consistent, and asymptotically efficient. In contrast, LSE focuses solely on minimizing the squared deviation between observed and predicted values, without considering the underlying distributional properties. As a result, MLE tends to provide more accurate and statistically meaningful estimates. Figure 5 shows the fit of the theoretical PDF to the histogram of the data.
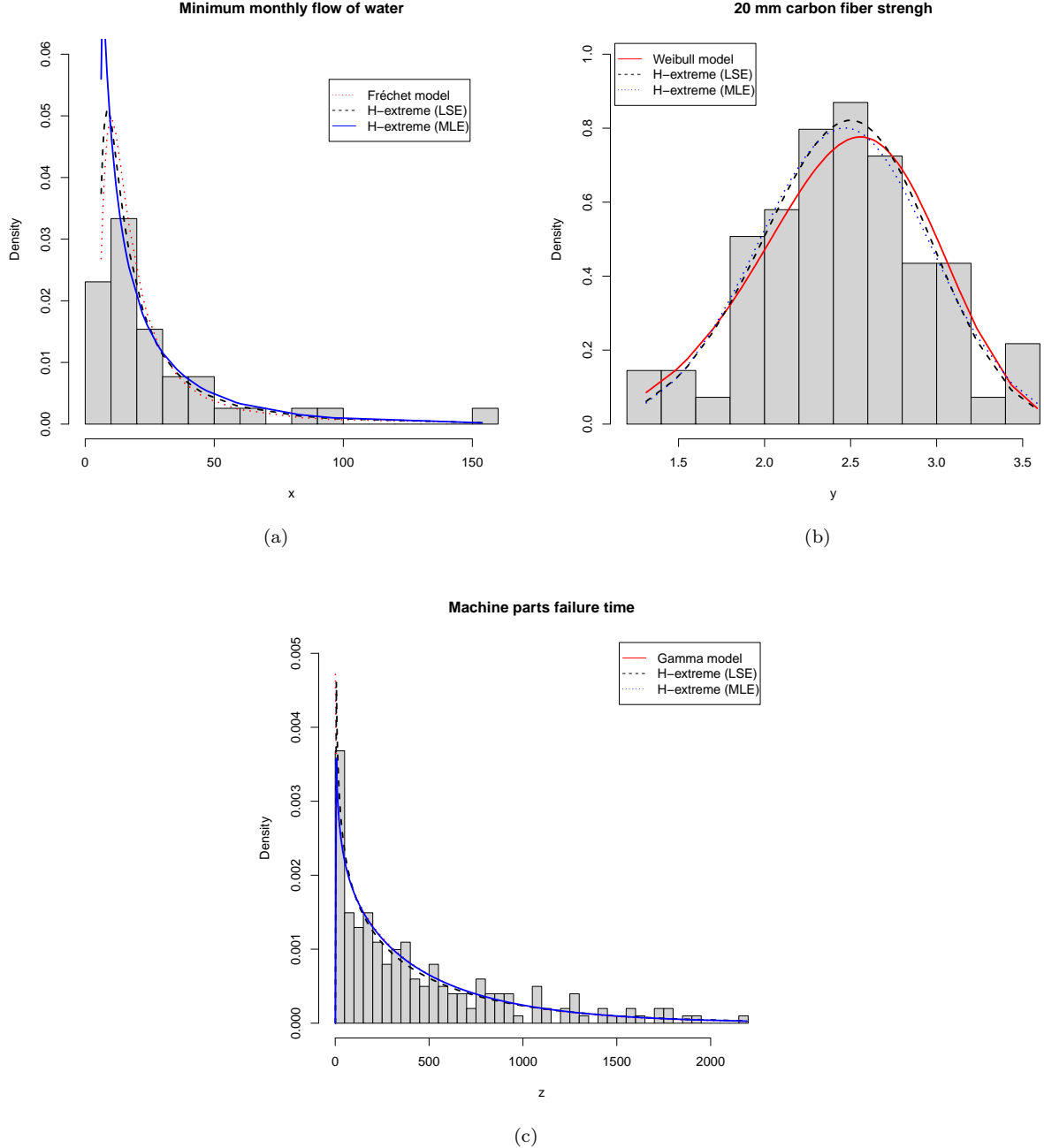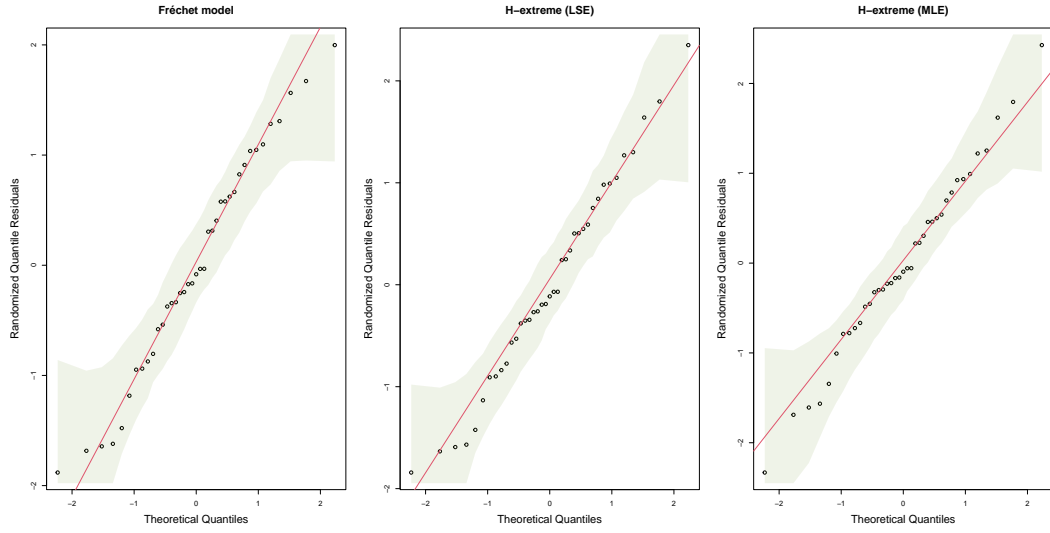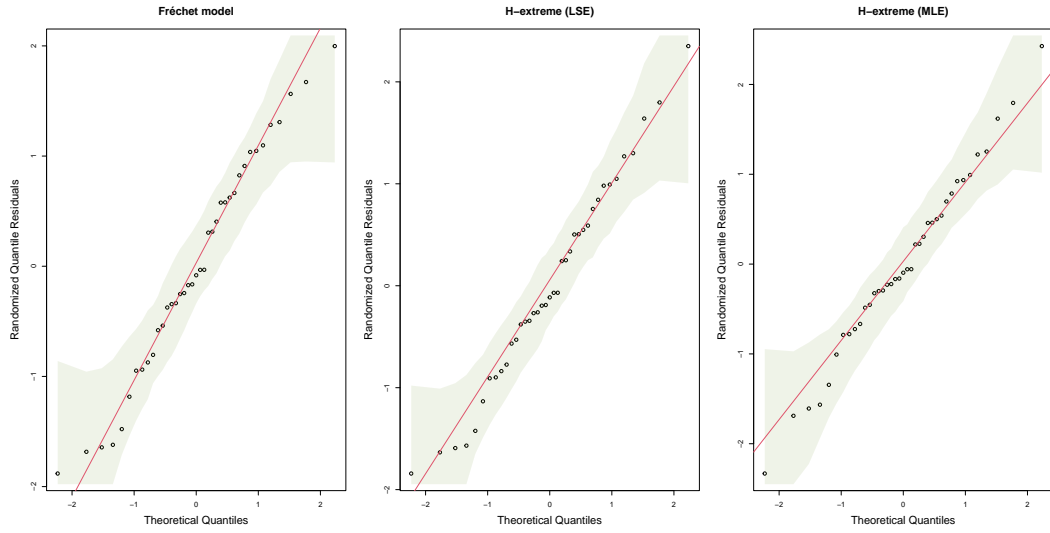

(a)

(b)

(c)

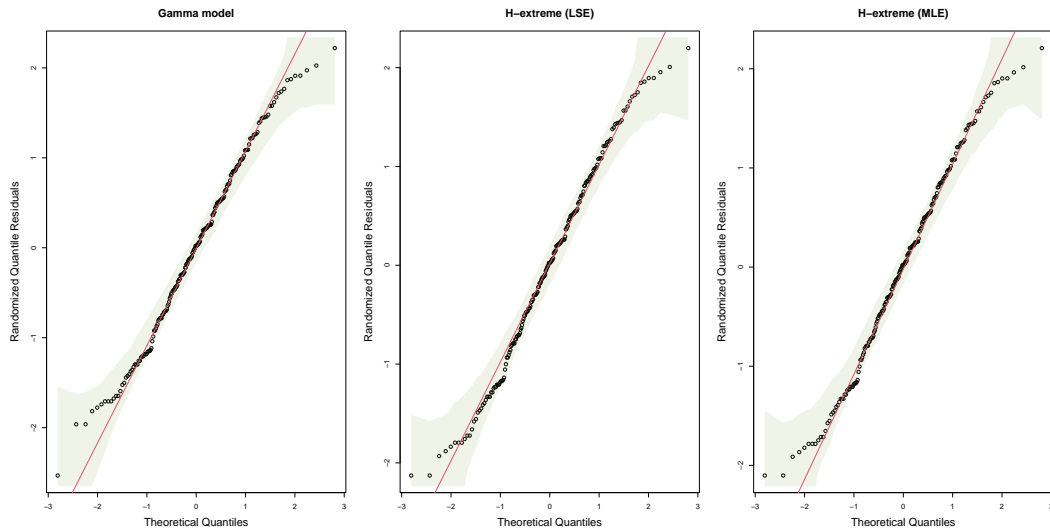Figure 5: Histograms and fitted probability density functions for (a) X, (b) Y and (c) Z.

To ensure statistical validity when model parameters are estimated from the data, a parametric bootstrap procedure was implemented to obtain reliable p-values for the fitted CDFs. Specifically, we employed the Kolmogorov–Smirnov (KS) and Cramér–von Mises (CVM) goodness-of-fit tests, combined with bootstrap-based resampling using $M = 1,000$ (see, e.g., Chapter 4.2.3 in [6]). This approach accounts for the composite nature of the null hypothesis and provides corrected significance levels. The resulting p-values indicate that the fitted models adequately describe the data: for dataset X, the KS and CVM p-values are 0.60 and 0.83, respectively; for Y, 0.99 and 0.99; and for Z, 0.99 and 0.95. All values are sufficiently high to suggest no significant departure between the empirical and fitted distributions.

Figure 6: Normal Quantile-Quantile plot displaying residuals from fitted models for (a) X, (b) Y and (c) Z.

To ensure a good model fit, we use Randomized Quantile (RQ) residuals, as defined by [7]. The RQ residuals are calculated using the formula $R_i = \Phi^{-1}(G(y_i; \hat{\boldsymbol{\theta}}))$, where $\hat{\boldsymbol{\theta}}$ is the estimated parameter, $G(y_i; \hat{\boldsymbol{\theta}})$ is the CDF (4) of the model fitted to each observation y_i, and Φ^{-1} represents the quantile of the standard normal distribution $N(0, 1)$. When F is continuous, the RQ residuals follow a standard normal distribution, excluding the impact of sampling variability in the estimated parameters.

For all datasets, the Quantile–Quantile (QQ) plots indicate that the $\mathbb{H}$-extreme distribution provides a superior fit compared to the original reference distributions. The improvement is particularly noticeable in the tails, where the residuals from the $\mathbb{H}$-extreme models exhibit a closer alignment with the theoretical reference lines. This behavior suggests that the proposed model more accurately captures the extreme-value behavior and tail dependence present in the data. In particular, visually the results from LSE seem a bit better than MLE, which is explained in the scenarios of small and moderate sample sizes. In those cases, MLE can suffer from high variance in the tail estimates because it relies heavily on the likelihood contributions of few extreme observations. LSE, which spreads the influence across all observations, can occasionally yield smoother (though less theoretically efficient, as demonstrated in Table 3) estimates in these regions. This does not affect our general conclusion that the $\mathbb{H}$-extreme distribution does provide a superior fit in the cases studied.

6   Conclusions

We studied a general class of probability distributions with positive support, which generalizes several well-known and established models in the literature. In this study, we described various theoretical properties of this class. Additionally, to define the model's distribution function, we introduced the concept of the incomplete extreme value $\mathbb{H}$-function.

Two estimation methods were proposed for the new model. These methods were implemented computationally, and their results were validated by modeling three real-world datasets. These datasets had previously been modeled in existing literature using particular cases of the new model—specifically, the Fréchet, Weibull, and Gamma distributions. Our study demonstrates that this new class of distributions, despite having more parameters than the basic models, provides a better fit to the data. This conclusion is supported by information criteria based on penalized likelihood (AIC, BIC, and EDC).

Despite its superior fit, the new model presents computational challenges. These arise from the need to estimate six parameters, and the potential non-identifiability of the model can sometimes hinder or limit the performance of numerical optimization algorithms. Therefore, using good initial guesses and carefully controlling the search space are essential for successful optimization.

Potential future work could involve enhancing the optimization methods employed in this context. For instance, one approach might be to implement global search metaheuristics that build upon local search heuristics while also exploring variations in the search regions.

Disclosure statement   There are no conflicts of interest to disclose.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables.* Dover, 1965.

[2] G. R. Aryal and C. P. Tsokos. On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods & Applications*, 71(12):1401–1407, 2009.

[3] J. Bergstra, D. Yamins, and D. Cox. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In *International Conference on Machine Learning*, volume 28, pages 115–123. PMLR, 2013.

[4] R. L. Burden and J. D. Faires. *Numerical Analysis.* Brooks Cole, 9 edition, 2010.

[5] I. Csiszar. I-divergence geometry of probability distributions and minimization problems. *Annals of Probability*, 3(1):146–158, 1975.

[6] A. C. Davison and D. V. Hinkley. *Bootstrap methods and their application.* Number 1. Cambridge University Press, 1997.

[7] P. K. Dunn and G. K. Smyth. Randomized quantile residuals. *Journal of Computational and graphical statistics*, 5(3):236–244, 1996.

[8] C.-Y. Huang, W.-C.-B. Chin, T.-H. Wen, Y.-H. Fu, and Y.-S. Tsai. Epirank: Modeling bidirectional disease spread in asymmetric commuting networks. *Scientific reports*, 9(1):5415, 2019.

[9] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions, volume 2*, volume 289. John wiley & sons, 1995.

[10] J. L. Lagrange. *Nouvelle méthode pour résoudre les équations littérales par le moyen des séries.* Chez Haude et Spener, Libraires de la Cour & de l'Académie royale, 1770.

[11] J. F. Lawless. *Statistical models and methods for lifetime data.* John Wiley & Sons, 2011.

[12] S. Nadarajah. Reliability for extreme value distributions. *Mathematical and computer modelling*, 37(9-10):915–922, 2003.

[13] W. B. Nelson. *Applied Life Data Analysis.* John Wiley & Sons, 2005.

[14] M. Oliveira, F. S. Quintino, D. Aguiar, P. N. Rathie, H. Saulo, T. A. d. Fonseca, and L. C. d. S. M. Ozelim. On the stress–strength reliability of transmuted gev random variables with applications to financial assets selection. *Entropy*, 26(6):441, 2024.

[15] F. S. Quintino, M. Oliveira, P. N. Rathie, L. Ozelim, and T. A. da Fonseca. Asset selection based on estimating stress-strength probabilities: The case of returns following three-parameter generalized extreme value distributions. *AIMS Math*, 9:2345–2368, 2024.

[16] F. S. Quintino, P. N. Rathie, L. C. d. S. M. Ozelim, and T. A. da Fonseca. Estimation of $P(X < Y)$ Stress–Strength Reliability Measures for a Class of Asymmetric Distributions: The Case of Three-Parameter p-Max Stable Laws. *Symmetry*, 16(7):837, 2024.

[17] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2024.

[18] P. L. Ramos, F. Louzada, E. Ramos, and S. Dey. The fréchet distribution: Estimation and application - an overview. *Journal of Statistics and Management Systems*, 23(3):549–578, 2020.

[19] P. L. Ramos, A. L. Mota, P. H. Ferreira, E. Ramos, V. L. Tomazella, and F. Louzada. Bayesian analysis of the inverse generalized gamma distribution using objective priors. *Journal of Statistical Computation and Simulation*, 91(4):786–816, 2021.

[20] P. Rathie and L. Ozelim. Exact and approximate expressions for the reliability of stable lévy random variables with applications to stock market modelling. *Journal of Computational and Applied Mathematics*, 321:314–322, Sept. 2017.

[21] P. Rathie, L. Ozelim, and C. Otiniano. Exact distribution of the product and the quotient of two stable lévy random variables. *Communications in Nonlinear Science and Numerical Simulation*, 36:204–218, July 2016.

[22] P. N. Rathie, L. C. d. S. M. Ozelim, F. Quintino, and T. A. d. Fonseca. On the Extreme Value H-Function. *Stats*, 6(3):802–811, 2023.

[23] T. J. Rothenberg. Identification in Parametric Models. *Econometrica*, 39(3):577–591, 1971.

[24] J. Shao. *Mathematical Statistics.* Springer Science & Business Media, 2008.

[25] M. D. Springer. *The Algebra of Random Variables.* John Wiley, New York, 1979.

[26] E. W. Stacy. A Generalization of the Gamma Distribution. *The Annals of Mathematical Statistics*, 33:1187–1192, 1962.

[27] A. Stroud and D. Secrest. *Gaussian quadrature formulas (without numerical tables)*. Prentice-Hall series in automatic computation. PH, 1966.

[28] J. Sun, M. Kong, and S. Pal. The modified-half-normal distribution: Properties and an efficient sampling scheme. *Communications in Statistics-Theory and Methods*, 52(5):1591–1613, 2023.

[29] R. Valiollahi, A. Asgharzadeh, and M. Z. Raqab. Estimation of $P(Y < X)$ for Weibull distribution under progressive Type-II censoring. *Communications in Statistics-Theory and Methods*, 42(24):4476–4498, 2013.

[30] G. Van Rossum and F. L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.

[31] R. Vila and M. N. Çankaya. A bimodal Weibull distribution: properties and inference. *Journal of Applied Statistics*, 49(12):3044–3062, 2022.

[32] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.