

Article

Bridging Extremes: The Invertible Bimodal Gumbel Distribution

Cira G. Otiniano ^{1,*}, Eduarda B. Silva ^{1,†}, Raul Y. Matsushita ^{1,†} and Alan Silva ^{2,†}

¹ Department of Statistics, University of Brasília, Brasília 70910-900, Brazil; eduarda.silva@brb.com.br (E.B.S.); raulmta@unb.br (R.Y.M.)

² Institute of Mathematics and Statistics, São Paulo University, São Paulo 05508-220, Brazil; alansilva@ime.usp.br

* Correspondence: cira@unb.br

† The authors contributed equally to this work.

Abstract: This paper introduces a novel three-parameter invertible bimodal Gumbel distribution, addressing the need for a versatile statistical tool capable of simultaneously modeling maximum and minimum extremes in various fields such as hydrology, meteorology, finance, and insurance. Unlike previous bimodal Gumbel distributions available in the literature, our proposed model features a simple closed-form cumulative distribution function, enhancing its computational attractiveness and applicability. This paper elucidates the behavior and advantages of the invertible bimodal Gumbel distribution through detailed mathematical formulations, graphical illustrations, and exploration of distributional characteristics. We illustrate using financial data to estimate Value at Risk (VaR) from our suggested model, considering maximum and minimum blocks simultaneously.

Keywords: Gumbel distribution; bimodality; extreme value theory; value at risk

1. Introduction

Bimodal heavy-tailed distributions are powerful analytical tools for capturing the complex nature of phenomena subject to extreme events in hydrology, meteorology, insurance, reliability, and finance, among other disciplines. Two main features characterize it. Firstly, it exhibits bimodality, meaning it has two distinct peaks or modes, indicating the presence of two prominent regimes within the overall data set. Secondly, it has heavy tails, which means a higher likelihood of occurrences of extreme values than with light tails distributions.

Bimodal heavy-tailed distributions are related to René Thom's catastrophe theory, focusing on systems characterized by sudden, dramatic changes and a propensity for extreme events, e.g., [1,2]. Catastrophe theory deals with situations where small parameter changes can lead to abrupt shifts in the system's state. This concept aligns with bimodal distributions, where a system may switch between two states or regimes. The heavy-tailed aspect of these distributions reflects the likelihood of rare, extreme events, mirroring the focus of catastrophe theory on significant, sudden changes. Both concepts encapsulate the unpredictability and uncertainty inherent in the systems they describe. Catastrophe theory provides a mathematical framework for understanding these dynamics, which can manifest statistically as bimodal, heavy-tailed distributions. This connection is especially relevant in economics and finance, where catastrophic shifts and heavy-tailed distributions are frequently observed. Essentially, the interplay between these concepts helps us to understand and model systems where small inputs or changes can lead to significant, unpredictable, and often extreme outputs or shifts.

Unlike standard bimodal distributions, these heavy-tailed versions give significant weight to extreme events, allowing for a more accurate representation of systems where outliers or "black swans" play a critical role. For example, one can perform inference over tails of financial returns by fitting an appropriate limiting distribution over data that



Citation: Otiniano, C.G.; Silva, E.B.; Matsushita, R.Y.; Silva, A. Bridging Extremes: The Invertible Bimodal Gumbel Distribution. *Entropy* **2023**, *25*, 1598. <https://doi.org/10.3390/e25121598>

Academic Editor: Sandro Vaienti

Received: 25 October 2023

Revised: 16 November 2023

Accepted: 24 November 2023

Published: 29 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

exceeds a fixed threshold, where the dual peaks in such a distribution can indicate states or types of behavior within the system [3]. In addition, typically, a bimodal distribution exhibits higher entropy compared to an unimodal distribution because two distinct modes add complexity and unpredictability to the system. In this way, the concept of entropy dovetails nicely with its inherent complexities, providing a quantitative lens through which to assess and strategize based on this kind of data.

Meanwhile, under certain conditions, statistics of extreme events are described by theoretical distributions. For example, the Gumbel, also known as the extreme value or Fisher–Tippet type I distribution, is a limiting distribution for the maximum (or the minimum) of a sufficiently large simple random sample. This result arises from the Fisher–Tippett–Gnedenko theorem, which states that the normalized maximum of a sequence of such random samples converges to one of three types of extreme value distribution: Gumbel, Fréchet, or Weibull. The Gumbel case is suitable for some typical families of populational distributions, such as logistic, Gaussian, and gamma.

Nevertheless, practical situations demand more, and therefore we find several generalizations of Gumbel to make it more flexible, for example, the two-component extreme value distribution or mixture of two Gumbel distributions [4], the exponentiated Gumbel [5], the transmuted extreme value [6], the generalized Gumbel [7], the generalized three-parameter Gumbel [8], the beta-Gumbel [9], the Kumaraswamy-Gumbel [10], and the exponentiated generalized Gumbel [11]. However, some lead to non-identifiable models because of observationally equivalent parameterizations [12]. There are other closely related models such as the exponentiated Gumbel Type-2 [13], the Kumaraswamy generalized exponentiated Gumbel type-2 [14], the bimodal generalized extreme value (GEV) [15], and a bimodal Gumbel distribution applied to environmental data [16]. However, the disadvantage of the latter model is that its cumulated distribution function does not have a simple closed form.

In this work, we put forward an invertible bimodal Gumbel distribution whose cumulated distribution function has a simple closed form, making it more attractive for computational procedures and more flexible in applications (Section 2). Our suggested distribution allows us to model both maximum and minimum simultaneously, while the classical Gumbel distribution describes only one of the extremes (maximum or minimum).

After discussing the maximum likelihood estimation of the parameters from simulated data (Section 3), we illustrate our approach using two financial data sets to estimate the value at risk (VaR) in Section 4. As we are interested in studying the probability distribution's tails, we perform the block maxima technique among the available tools to find the appropriate cutoff. Instead of the usual power law distribution [3,17,18], we suggest a bimodal Gumbel distribution as a candidate model to describe the tail behavior of financial returns.

2. Main Results

Also known as type I extreme value distribution, the Gumbel distribution is one of the limit distributions of normalized maximum (or minimum) statistics [19], belonging to the class of the GEV distribution [20]. We denote the Gumbel random variable Y with a location parameter $\mu \in \mathbb{R}$ and a scale parameter $\sigma > 0$ as $Y \sim G(\cdot; \mu, \sigma)$. The forms of its probability density function (PDF) and cumulative distribution function (CDF) are, respectively,

$$g(y; \mu, \sigma) = \frac{1}{\sigma} \exp \left\{ - \left(\frac{y - \mu}{\sigma} \right) - \exp \left[- \left(\frac{y - \mu}{\sigma} \right) \right] \right\}, \quad \forall y \in \mathbb{R} \quad (1)$$

and

$$G(y; \sigma, \mu) = \exp \left[- \exp \left(- \frac{y - \mu}{\sigma} \right) \right], \quad \forall y \in \mathbb{R}. \quad (2)$$

Let us introduce our suggested generalization of the Gumbel distribution left open by [15] in the following way. Considering the transformation,

$$T_{\mu,\delta}(x) = x|x|^\delta + \mu, \quad x \in \mathbb{R}, \quad \delta > -1, \quad \mu \in \mathbb{R}, \quad (3)$$

after plugging it into (1) and (2), we obtain the invertible bimodal Gumbel distribution X with CDF and PDF given by, respectively,

$$\begin{aligned} F_{\text{IBG}}(x; \mu, \sigma, \delta) &= G(T_{\mu,\delta}(x); \sigma) \\ &= \exp\left\{-\exp\left[-\frac{(x|x|^\delta) + \mu}{\sigma}\right]\right\}, \quad x \in \mathbb{R} \end{aligned} \quad (4)$$

and

$$f_{\text{IBG}}(x; \mu, \sigma, \delta) = \frac{1}{\sigma}(\delta + 1)|x|^\delta \exp\left\{-\left(\frac{(x|x|^\delta) + \mu}{\sigma}\right) - \exp\left[-\frac{(x|x|^\delta) + \mu}{\sigma}\right]\right\}, \quad (5)$$

where $\delta > 0$ and $\mu \in \mathbb{R}$ are shape parameters and $\sigma > 0$ is a scale parameter.

We shall denote it as $X \sim F_{\text{IBG}}(\cdot; \mu, \sigma, \delta)$ throughout this paper.

To illustrate the role of its parameters, Figure 1 depicts the effect of the shape parameter δ . When $\delta = 0$, the model (5) reduces to the unimodal Gumbel (1). The density becomes bimodal for $\delta > 0$, and the modes' separation rises as δ increases. Figure 2 contrasts the PDF shapes with negative and positive values of μ , illustrating its role as a location parameter (left) or shape parameter (right). Finally, Figure 3 shows that σ remains the scale parameter.

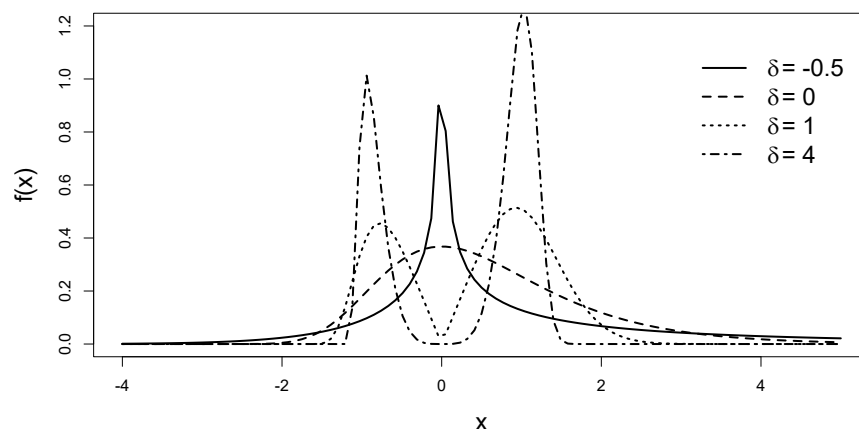


Figure 1. Density $f_{\text{IBG}}(x; 0, 1, \delta)$, with δ ranging from -0.5 to 4 . Bimodal distributions appear when $\delta > 0$, and we have the unimodal Gumbel if $\delta = 0$.

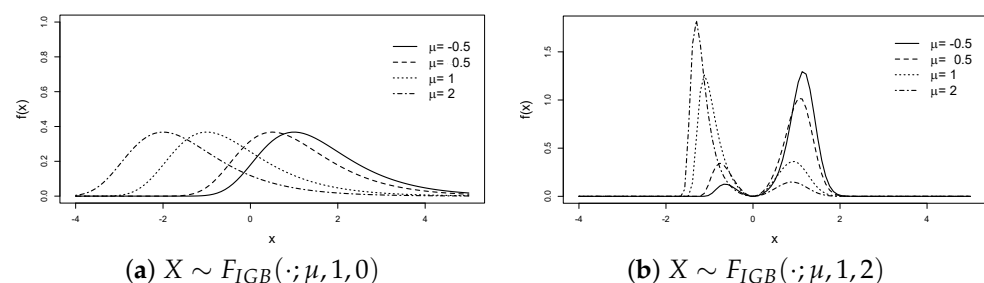


Figure 2. Density $f_{\text{IBG}}(\cdot; \mu, 1, \delta)$, with μ ranging from -0.5 to 2 , $\delta = 0$ or 1 , and μ acting as a location (a) or shape parameter (b).

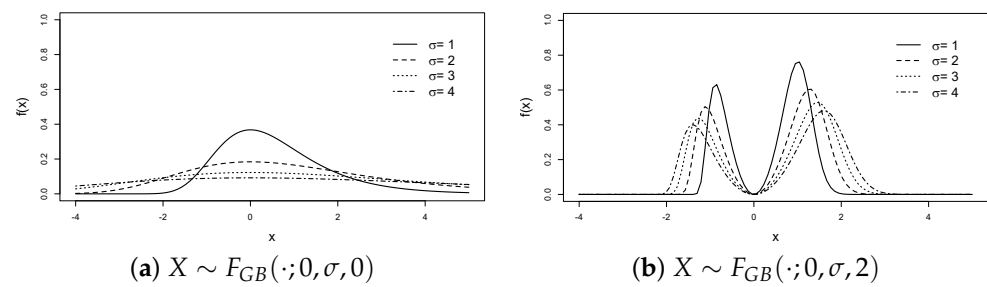


Figure 3. Density $f_{\text{IBG}}(\cdot; 0, \sigma, \delta)$, with σ ranging from 1 to 4 and $\delta = 0$ (unimodal, (a)) or 2 (bimodal, (b)). In both cases, σ represents the scale parameter.

Some Distributional Characteristics

Modes 1. Straightforwardly from the concept of the modes of $X \sim_{\text{GB}}(\cdot; \mu, \sigma, \delta)$, one can find that they are the solution of the differential equation

$$\frac{T''_{\mu,\delta}(x)}{[T'_{\mu,\delta}(x)]^2} = \frac{1}{\sigma} - \frac{e^{-\frac{T_{\mu,\delta}(x)}{\sigma}}}{\sigma}, \quad (6)$$

where

$$T'_{\mu,\delta}(x) = (\delta + 1)|x|^\delta \quad (7)$$

and

$$T''_{\mu,\delta}(x) = \text{sign}(x) (\delta + 1)\delta |x|^{\delta-1}, \quad (8)$$

with $\text{sign}(x) = x/|x|$ as the sign function.

Moments 2. We can write down the k th moment of X as

$$\begin{aligned} E(X^k) &= \int_{-\infty}^{\infty} x^k f_{\text{IBG}}(x) dx \\ &= \int_{-\infty}^{\infty} x^k g(T_{\mu,\delta}(x), 0, \mu) T'_{\mu,\delta}(x) dx. \end{aligned} \quad (9)$$

By substitution $y = T(x)$ and taking the inverse function $T^{-1}(y) = \text{sgn}(|y - \mu|)|y - \mu|^{\frac{1}{\delta+1}}$, we can express (9) in terms of a unimodal Gumbel Y , as defined in (2) as

$$E(X^k) = E\left(|Y - \mu|^{\frac{k}{\delta+1}} \cdot I_{[\mu, +\infty)}\right) + (-1)^k E\left(|Y - \mu|^{\frac{k}{\delta+1}} \cdot I_{(-\infty, \mu)}\right), \quad (10)$$

where I_A is the indicator function of an event A .

Moment-Generating Type Function 3. The moment-generating function (MGF) encapsulates information about the distributional moments, being a helpful tool to characterize an IBG random variable X . For our convenience, however, we consider its power transformation $X^{\delta+1}$ and derive its MGF shown in (13) as follows. From its definition,

$$\varphi_X(t) = E(e^{tX^{\delta+1}}) = \int_{-\infty}^{\infty} \frac{e^{tx^{1+\delta}}}{\sigma} e^{-\frac{T(x)}{\sigma}} e^{-e^{-\frac{T(x)}{\sigma}}} T'(x) dx. \quad (11)$$

Considering the substitution $y = e^{\frac{-T(x)}{\sigma}}$ and the expression $T^{-1}(\ln y^{-\sigma}) = \text{sgn}(|\ln y^{-\sigma} - \mu|) |\ln y^{-\sigma} - \mu|^{\frac{1}{\delta+1}}$, we can rewrite the integral (11) as being

$$\begin{aligned}\varphi_X(t) &= \int_0^{e^{\frac{\mu}{\sigma}}} e^{t(-\sigma \ln y - \mu)} e^{-y} dy + \int_{e^{\frac{\mu}{\sigma}}}^{\infty} e^{(-1)^{1+\delta} t(\mu + \sigma \ln y)} e^{-y} dy \\ &= e^{-t\mu} \int_0^{e^{\frac{\mu}{\sigma}}} y^{-t\sigma} e^{-y} dy + e^{(-1)^{1+\delta} t\mu} \int_{e^{\frac{\mu}{\sigma}}}^{\infty} y^{t\sigma(-1)^{1+\delta}} e^{-y} dy,\end{aligned}\quad (12)$$

that is,

$$\varphi_X(t) = e^{-t\mu} \Gamma(1 - t\sigma; e^{\frac{\mu}{\sigma}}) + e^{t\mu(-1)^{\delta+1}} \gamma(1 + (-1)^{\delta+1} t\sigma; e^{\frac{\mu}{\sigma}}), \quad (13)$$

where

$$\gamma(a; x) = \int_x^{\infty} t^{a-1} e^{-t} dt \quad \text{and} \quad \Gamma(a; x) = \int_0^x t^{a-1} e^{-t} dt, \quad (14)$$

are the upper and lower incomplete Gamma functions. Now, we can retrieve the moments of $X^{\delta+1}$ by taking derivatives of the cumulant-generating type function, $C_X = \ln \varphi_X(t)$. As usual, from the expansion $\ln z = (z - 1) - (z - 1)^2/2 + (z - 1)^3/3 - \dots$, we find

$$C_X(t) = (E(e^{tX^{\delta+1}}) - 1) - \frac{(E(e^{tX^{\delta+1}}) - 1)^2}{2} + \frac{(E(e^{tX^{\delta+1}}) - 1)^3}{3} - \dots \quad (15)$$

Thus, for example, we get the first two moments of $X^{\delta+1}$ by taking the derivatives

$$\begin{aligned}\frac{d}{dt} C_X(t) \big|_{t=0} &= E(X^{\delta+1}) \\ \frac{d^2}{dt^2} C_X(t) \big|_{t=0} &= E(X^{2(\delta+1)}) - [E(X^{\delta+1})]^2.\end{aligned}\quad (16)$$

To benchmark our result, if $\delta = 0$, from (13) and (16), we find $E(X) = \sigma\gamma - \mu$ and $\text{Var}(X) = \sigma^2\pi^2/6$, because $\Gamma(a; x) + \gamma(a; x) = \Gamma(a)$, where $\Psi(x) = d \ln(\Gamma(x))/dx = \Gamma'(x)/\Gamma(x)$, $\Psi(1) = \gamma$ is the Euler's constant, and $\Psi'(1) = \frac{\pi^2}{6}$. Thus, we get the mean and variance of the basic Gumbel model as expected, which confirms the accuracy of (13).

Quantiles 4. Sampling by the inverse transform is a basic method with which to generate a pseudo-random variate of X , based on its quantile function of F_{IBG} . While the bimodal Gumbel model introduced previously by [16] does not provide a simple way to perform this method, our suggested model (5) yields a simple expression for the quantile function. Since X is an absolutely continuous random variable, denoting the cumulative probability as the standard uniform random variable $F_{IBG}(x_q; \mu, \sigma, \delta) = q \sim U[0, 1]$, we obtain the random quantile function as

$$X_q = F_{IBG}^{-1}(q) = \begin{cases} \left(-\sigma \ln(-\ln(q)) - \mu \right)^{\frac{1}{1+\delta}}, & q > e^{-e^{-\frac{\mu}{\sigma}}}, \\ -\left(\mu + \sigma \ln(-\ln(q)) \right)^{\frac{1}{1+\delta}}, & q < e^{-e^{-\frac{\mu}{\sigma}}}. \end{cases} \quad (17)$$

Entropy 5. The differential entropy of the bimodal Gumbel distribution $X \sim F_{IBG}(\cdot; \mu, \sigma, \delta) = G(Y(T_{\mu, \delta}(\cdot), \sigma))$, where $Y \sim G(\cdot, \sigma)$ denotes the basic Gumbel distribution, is given by

$$\begin{aligned}
H(X) &= - \int_{-\infty}^{\infty} f_{IBG}(x; \mu, \sigma, \delta) \ln[f_{IBG}(x; \mu, \sigma, \delta)] dx \\
&= \int_{-\infty}^{\infty} g(T_{\mu, \delta}(x); \sigma) \left\{ \exp\left[-\frac{T_{\mu, \delta}(x)}{\sigma}\right] + \frac{T_{\mu, \delta}(x)}{\sigma} + \ln\left(\frac{T'_{\mu, \delta}(x)}{\sigma}\right) \right\} \frac{T'_{\mu, \delta}(x)}{\sigma} dx \\
&= \int_{-\infty}^{\infty} g(T_{\mu, \delta}(x); \sigma) \exp\left[-\frac{T_{\mu, \delta}(x)}{\sigma}\right] \frac{T'_{\mu, \delta}(x)}{\sigma} dx + \int_{-\infty}^{\infty} g(T_{\mu, \delta}(x); \sigma) \frac{T_{\mu, \delta}(x)}{\sigma} \frac{T'_{\mu, \delta}(x)}{\sigma} dx \\
&+ \int_{-\infty}^{\infty} g(T_{\mu, \delta}(x); \sigma) \ln\left(\frac{T'_{\mu, \delta}(x)}{\sigma}\right) \frac{T'_{\mu, \delta}(x)}{\sigma} dx, \tag{18}
\end{aligned}$$

where g is the PDF of Y , as defined in (2). By substituting $y = T(x)$ in (18), we obtain

$$\begin{aligned}
H(X) &= \frac{1}{\sigma} \int_{-\infty}^{\infty} \exp\left[-\frac{y}{\sigma}\right] g(y) dy + \frac{1}{\sigma^2} \int_{-\infty}^{\infty} y g(y) dy \\
&+ \frac{1}{\sigma} \int_{-\infty}^{\infty} \ln\left(\frac{\delta+1}{\sigma}\right) g(y) dy + \frac{1}{\sigma} \int_{-\infty}^{\infty} \ln\left(|y - \mu|^{\frac{\delta}{\delta+1}}\right) dy \\
&= 1 + \frac{\gamma}{\sigma} + \ln\left(\frac{\delta+1}{\sigma}\right)^{\frac{1}{\sigma}} + \ln\left(E|Y - \mu|^{\frac{\delta}{\delta+1}}\right). \tag{19}
\end{aligned}$$

3. Parameter Estimation

This section discusses the maximum likelihood (ML) estimation method to estimate the vector parameters $\Theta = (\mu, \sigma, \delta)$. Let $x_1, \dots, x_n$ be realizations independent copies of a random variable with PDF as defined in (5). The log-likelihood function is

$$\begin{aligned}
l(\Theta; x_1, x_2, \dots, x_n) &= \sum_{i=1}^n \ln f(x_i; \Theta) \\
&= n \ln(\delta+1) - n \ln \sigma + \delta \sum_{i=1}^n \ln |x_i| - \frac{\sum_{i=1}^n x_i |x_i|^{\delta} + \mu}{\sigma} \\
&- \sum_{i=1}^n e^{-\frac{x_i |x_i|^{\delta} + \mu}{\sigma}}. \tag{20}
\end{aligned}$$

This log-likelihood function is well-defined across the entire parameter space and is continuous and differentiable for the vector parameters. Additionally, the family of distributions F_{IBG} is identifiable, meaning different parameters should lead to distinct probability distributions, ensuring a unique maximum for the likelihood function.

Ahmad et al. (2010) [21] showed the identifiability of the finite mixture of Gumbel distributions; in particular, the family of a Gumbel component $\mathcal{F}_G = \{G : G = G(\cdot, \mu, \sigma) \text{ as (2)}\}$ is identifiable. Based on this, we have that the IBG family, $\mathcal{F}_{IBG} = \{F_{IBG} : F_{IBG}(\cdot, \mu, \sigma, \delta) \text{ as (4)}\}$, is identifiable. It must be proven that

$$F_{IBG}(x; \mu_1, \sigma_1, \delta_1) = F_{IBG}(x; \mu_2, \sigma_2, \delta_2) \text{ if and only if } \mu_1 = \mu_2, \sigma_1 = \sigma_2, \delta_1 = \delta_2.$$

Indeed, from (4)

$$\exp\left\{-\exp\left[-\frac{x|x|^{\delta_1} + \mu_1}{\sigma_1}\right]\right\} = \exp\left\{-\exp\left[-\frac{x|x|^{\delta_2} + \mu_2}{\sigma_2}\right]\right\}. \tag{21}$$

Since $\mathcal{F}_G$ is identifiable, then $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$. Thus, the Equation (21) is valid if and only if

$$|x|^{\delta_1} - |x|^{\delta_2} = 0,$$

which only happens when $\delta_1 = \delta_2$, for any $x \in \mathbb{R}$.

The ML estimates $\hat{\mu}, \hat{\sigma}, \hat{\delta}$ are the solution of the system of likelihood equations

$$\frac{\partial l(\Theta; x)}{\partial \mu} = -\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}} \sum_{i=1}^n e^{-\frac{x_i |x_i|^{\hat{\delta}}}{\hat{\sigma}}} = 0; \quad (22)$$

$$\frac{\partial l(\Theta; x)}{\partial \sigma} = -\frac{n}{\hat{\sigma}} - \sum_{i=1}^n \hat{\sigma}^{-2} (x_i |x_i|^{\hat{\delta}} + \hat{\mu}) - \sum_{i=1}^n \hat{\sigma}^{-2} (x_i |x_i|^{\hat{\delta}} + \hat{\mu}) e^{-(x_i |x_i|^{\hat{\delta}} + \hat{\mu}) \hat{\sigma}^{-1}} = 0; \quad (23)$$

$$\begin{aligned} \frac{\partial l(\Theta; x)}{\partial \delta} &= \frac{n}{\hat{\delta}+1} + \sum_{i=1}^n \ln |x_i| - \frac{1}{\hat{\sigma}} \sum_{i=1}^n x_i |x_i|^{\hat{\delta}} \ln \frac{\sum_{k=1}^n x_k |x_k|^{\hat{\delta}}}{\hat{\sigma}} \\ &+ e^{\frac{\hat{\mu}}{\hat{\sigma}}} \sum_{i=1}^n e^{-\frac{x_i |x_i|^{\hat{\delta}}}{\hat{\sigma}}} \left(\frac{\sum_{j=1}^n x_j |x_j|^{\hat{\delta}}}{\hat{\sigma}} \right) \ln \frac{\sum_{k=1}^n x_k |x_k|^{\hat{\delta}}}{\hat{\sigma}} = 0. \end{aligned} \quad (24)$$

After algebraic manipulations, we get the unique closed-form solution for estimating μ ,

$$\hat{\mu} = \hat{\sigma} \ln \frac{\sum_{i=1}^n e^{-\frac{x_i |x_i|^{\hat{\delta}}}{\hat{\sigma}}}}{n}. \quad (25)$$

However, the estimates $\hat{\sigma}$ and $\hat{\delta}$ must be obtained numerically.

Numerical Performance of ML Estimates

Now, we perform a Monte Carlo study to assess the performance of maximum likelihood estimators $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\delta}$ in terms of their means, mean squared errors (MSE), biases, and standard errors (SE). We defined a set of 9 parameter vectors, $\theta_1, \dots, \theta_9$, with $\mu \in \{-1, 0, 1\}$, $\delta \in \{0, 2, 4\}$, and $\sigma = 1$, for three sampling scenarios: $n = 50, 100$, and 1000 . For each of the 27 combinations between parameters and sample scenarios, we took 100 Monte Carlo replications using the software R (version 3.4.1) to get the empirical sampling distributions of $\hat{\mu}$, $\hat{\sigma}$, and $\hat{\delta}$. We generate the Monte Carlo variates of $X \sim F_{GB}(\cdot; \mu, \sigma, \delta)$ through the inverse transform method with the quantile Function (17).

Tables 1–3 depict the empirical expected values, bias, MSE, and SE of the ML estimators of the IBG model. Figures 4–6 illustrate the empirical behavior of the MSE vs n . Overall, the MSE decreases as the sample size increases, confirming the optimal properties of ML estimators from the statistical inference theory. In this study, we did not face numerical problems in estimating these parameters.

Table 1. Means, biases, mean squared errors (MSE), and standard errors (SE) of the estimated parameters from 100 Monte Carlo replications of samples with $n = 50$.

	θ	True	Mean	Bias	MSE	SE
θ_1	μ	−1	−1.00952	−0.00952	0.02163	0.1475
	σ	1	0.97492	−0.02507	0.02765	0.1653
	δ	0	0.01114	0.01114	0.01554	0.1248
θ_2	μ	−1	−1.03142	−0.03142	0.03419	0.1831
	σ	1	1.02366	0.02366	0.02592	0.1600
	δ	2	2.05863	0.05863	0.12062	0.3440
θ_3	μ	−1	−0.99788	0.00211	0.02748	0.1666
	σ	1	0.96513	−0.03486	0.02763	0.1634
	δ	4	3.91971	−0.08028	0.34339	0.5834

Table 1. *Cont.*

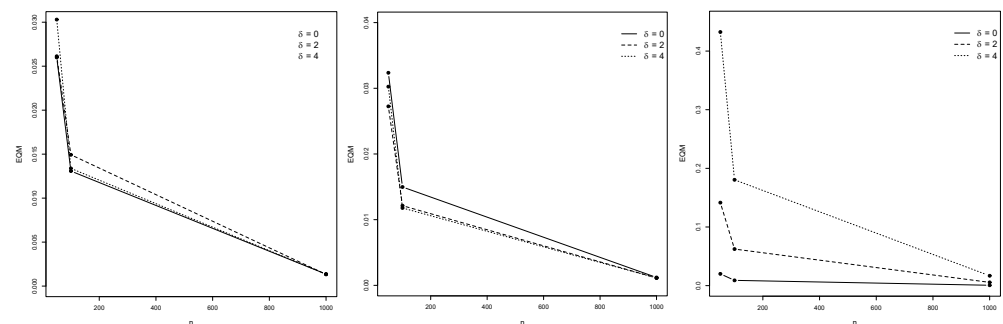
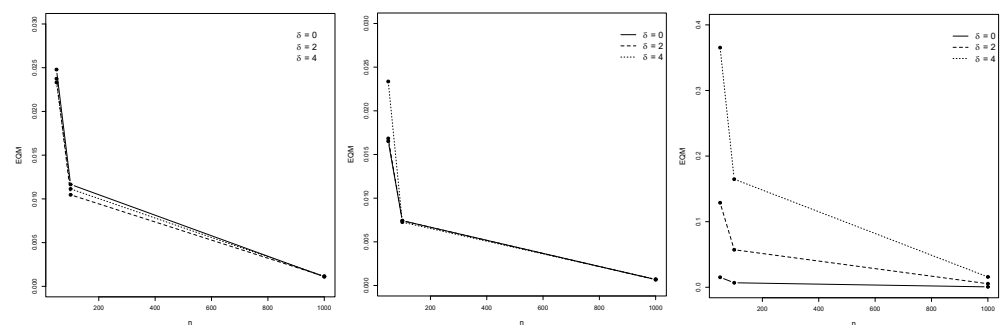
	θ	True	Mean	Bias	MSE	SE
θ_4	μ	0	0.00145	0.00145	0.02113	0.1461
	σ	1	0.99965	−0.00034	0.01646	0.1289
	δ	0	0.03047	0.03047	0.01316	0.1111
θ_5	μ	0	0.00339	0.00339	0.02363	0.1544
	σ	1	1.00572	0.00572	0.01401	0.1188
	δ	2	2.14749	0.14749	0.14933	0.3589
θ_6	μ	0	−0.02841	−0.02841	0.02129	0.1438
	σ	1	0.97091	−0.02908	0.01374	0.1141
	δ	4	4.00093	0.00093	0.26488	0.5172
θ_7	μ	1	0.95417	−0.04582	0.02859	0.1636
	σ	1	0.98530	−0.01469	0.01918	0.1384
	δ	0	0.00560	0.00560	0.01242	0.1118
θ_8	μ	1	0.96207	−0.03792	0.02218	0.1447
	σ	1	1.04081	0.04081	0.02338	0.1481
	δ	2	2.10659	0.10659	0.17084	0.4013
θ_9	μ	1	0.93990	−0.06009	0.05954	0.2376
	σ	1	0.97762	−0.02237	0.12103	0.3489
	δ	4	3.96419	−0.03580	0.91265	0.9594

Table 2. Means, biases, mean squared errors (MSE), and standard errors (SE) of the estimated parameters from 100 Monte Carlo replications of samples with $n = 100$.

	θ	True	Mean	Bias	MSE	SE
θ_1	μ	−1	−0.98025	0.01974	0.01280	0.1120
	σ	1	0.98180	−0.01819	0.01146	0.1060
	δ	0	−0.00898	−0.00898	0.00780	0.0883
θ_2	μ	−1	−0.99734	0.00265	0.01532	0.1243
	σ	1	0.99027	−0.00972	0.01188	0.1091
	δ	2	2.01504	0.01504	0.06759	0.2608
θ_3	μ	−1	−0.99937	0.00062	0.01558	0.1254
	σ	1	0.97695	−0.02304	0.01185	0.1069
	δ	4	3.9113	−0.08860	0.13284	0.3553
θ_4	μ	0	−0.02087	−0.02087	0.01232	0.1095
	σ	1	0.99485	−0.00514	0.00724	0.0854
	δ	0	0.02235	0.02235	0.00722	0.0824
θ_5	μ	0	−0.01413	−0.01413	0.01182	0.1083
	σ	1	0.98941	−0.01058	0.00694	0.0830
	δ	2	2.06691	0.06691	0.07194	0.2610
θ_6	μ	0	−0.00765	−0.00765	0.01211	0.1103
	σ	1	0.99801	−0.00198	0.00713	0.0848
	δ	4	4.05904	0.05904	0.16391	0.4025
θ_7	μ	1	0.94706	−0.05293	0.01350	0.1039
	σ	1	0.98666	−0.01333	0.00964	0.0977
	δ	0	−0.02825	−0.0282	0.00588	0.0717
θ_8	μ	1	0.96052	−0.03947	0.014831	0.1157
	σ	1	1.01628	0.01628	0.01126	0.1053
	δ	2	1.99650	−0.00349	0.06319	0.2526
θ_9	μ	1	0.98981	−0.01018	0.01102	0.1050
	σ	1	0.99397	−0.00602	0.00709	0.0844
	δ	4	4.02093	0.02093	0.11929	0.3464

Table 3. Means, biases, mean squared errors (MSE), and standard errors (SE) of the estimated parameters from 100 Monte Carlo replications of samples with $n = 1000$.

	θ	True	Mean	Bias	MSE	SE
θ_1	μ	−1	−0.99837	0.00162	0.00149	0.0388
	σ	1	0.99413	−0.00586	0.00104	0.0319
	δ	0	−0.00209	−0.00209	0.00072	0.0269
θ_2	μ	−1	−0.99906	0.00093	0.00130	0.0362
	σ	1	1.00507	0.00507	0.00088	0.0293
	δ	2	2.01457	0.01457	0.00540	0.0724
θ_3	μ	−1	−0.99417	0.00582	0.00110	0.0328
	σ	1	0.99574	−0.00425	0.00091	0.03016
	δ	4	3.99378	−0.00621	0.01433	0.1201
θ_4	μ	0	−0.00232	−0.00232	0.00134	0.0367
	σ	1	1.00226	0.00226	0.00069	0.02643
	δ	0	0.00303	0.00303	0.00076	0.0277
θ_5	μ	0	−0.00322	−0.00322	0.00094	0.0307
	σ	1	0.99633	−0.00366	0.00062	0.0248
	δ	2	2.01715	0.01715	0.00587	0.0751
θ_6	μ	0	−0.00824	−0.00824	0.00114	0.0330
	σ	1	1.00379	0.00379	0.00065	0.0254
	δ	4	4.01962	0.01962	0.01519	0.1223
θ_7	μ	1	0.99310	−0.00689	0.00136	0.0365
	σ	1	0.99649	−0.00350	0.00107	0.0327
	δ	0	−0.00116	−0.00116	0.00080	0.0284
θ_8	μ	1	0.95857	−0.04142	0.00253	0.0286
	σ	1	0.99516	−0.00483	0.00104	0.0320
	δ	2	1.96530	−0.03469	0.00929	0.0904
θ_9	μ	1	0.99638	−0.00361	0.00133	0.0365
	σ	1	1.00135	0.00135	0.00108	0.0330
	δ	4	4.00819	0.00819	0.01701	0.1308

**Figure 4.** MSE from Monte Carlo replications of samples with n ranging from 50 to 1000, for $\hat{\theta}_1$, $\hat{\theta}_2$, and $\hat{\theta}_3$.**Figure 5.** MSE behavior of $\hat{\theta}_4$, $\hat{\theta}_5$ and $\hat{\theta}_6$. MSE from Monte Carlo replications of samples with n ranging from 50 to 1000, for $\hat{\theta}_4$, $\hat{\theta}_5$, and $\hat{\theta}_6$.

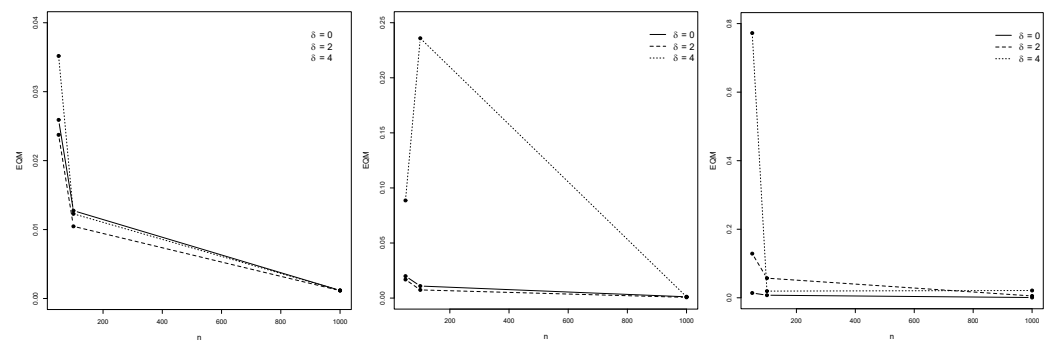


Figure 6. MSE from Monte Carlo replications of samples with n ranging from 50 to 1000, for $\hat{\theta}_7$, $\hat{\theta}_8$, and $\hat{\theta}_9$.

4. Application

We use two financial data sets taken from <https://finance.yahoo.com> (accessed on 1 December 2021) to illustrate the applicability of the invertible bimodal Gumbel model. The first is the daily stock prices of Petrobras (PETR4), quoted in US dollars, from 1 March 2000 to 10 January 2021, totaling 5,465 observations. The other is the daily exchange rate of the Brazilian real against the US dollar (USD/BRL) from 12 January 2003 to 15. October 2021, totaling 4,223 data points. We aim to get the value-at-risk (VaR) of these data, a common measure of financial risk. It denotes the maximum loss incurred on a portfolio over a specific time horizon with a given confidence level $1 - \alpha$ [22]. It is expressed in probabilistic terms as

$$\text{VaR}_\alpha(X_t) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}, \quad (26)$$

where $F(x)$ is the cumulative distribution function (CDF) of a real random variable X_t observed at time $t \in \{0, 1, 2, \dots\}$, and $0 < \alpha < 1$ is a small prespecified probability. Particularly, in our study, the time horizon comprises the totality of data in each discrete time series. Moreover, as X_t is an absolutely continuous random variable with an invertible CDF, we can write

$$x_\alpha = \text{VaR}_\alpha(X_t) = F^{-1}(1 - \alpha),$$

where F^{-1} denotes the inverse function of F , and x_α is the α -quantile of X_t . As usual, X_t means the log return of prices, that is,

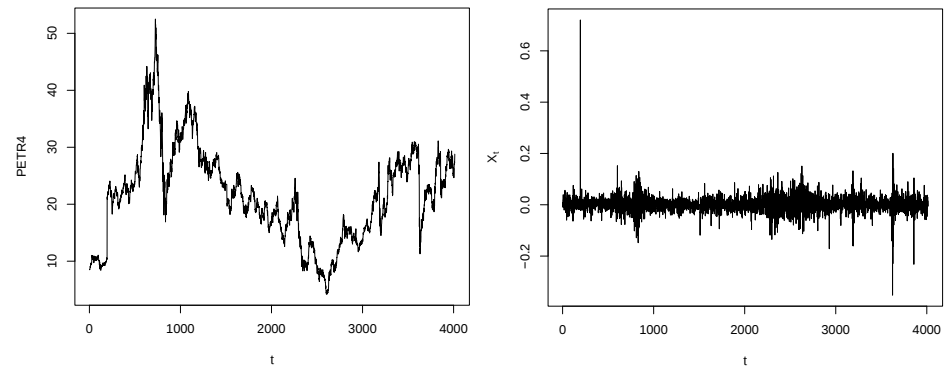
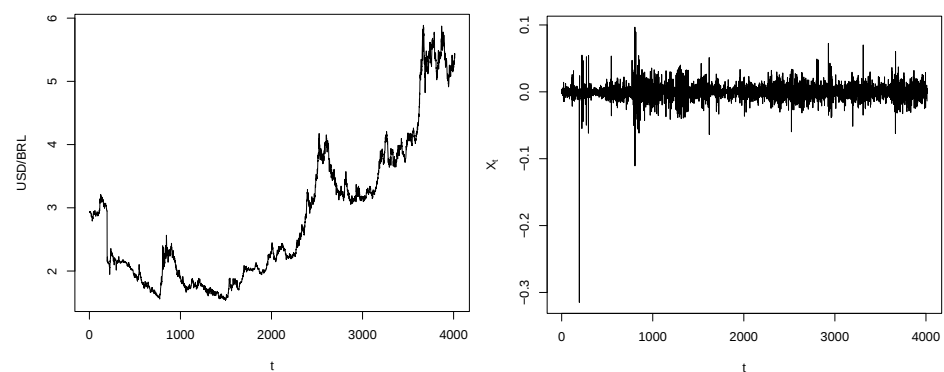
$$X_t = \ln P_t - \ln P_{t-1}, \quad (27)$$

where P_t is a price at time t .

Table 4 summarizes descriptive statistics for PETR4 and USD/BRL returns. The return averages are close to zero, and the proximity between the absolute values of the first and third quartiles indicates the possible symmetry of the data, except for the possible extreme values suggested by the maximum (PETR4) and minimum (USD/BRL) statistics. Indeed, Figures 7 and 8 depict extreme returns, some of them due to the COVID-19 event. In the critical period of the pandemic, Petrobras shares plummeted 57% due to low demand for petroleum products (Figure 7). As for the exchange rate USD/BRL, the effect was the opposite: the American dollar became more expensive than the Brazilian real because of various political and economic reasons. Furthermore, in these data sets, we observe extreme positive (PETR4) and negative (USD/BRL) values that stand out significantly from the rest of the observations.

Table 4. Descriptive statistics.

Stock	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
PETR4	−0.3523667	−0.0137843	0.0000000	0.0003036	0.0138190	0.7203695
USD/BRL	−0.3148314	−0.0056746	0.0000000	0.0001515	0.0060800	0.0966945

**Figure 7.** PETR4 Prices (left) and PETR4 log returns (right).**Figure 8.** USD/BRL Prices (left) and USD/BRL log returns (right).

Now, we perform the block maxima and minima method to extract extreme values from our data. Let $X_1, \dots, X_n$ be a random sample of log returns following $X \sim F_{IBG}$. Based on its realized values $\{x_t\}_{t=1}^n$, we organize it into T non-overlapping sub-samples of length N , where T means the integer part of n/N , resulting in T data blocks of size N . We choose N to cover natural periods (e.g., a week or month) so that the new sub-sample is IID. Now, we take the maximum and the minimum over each N -history. We define the j th sub-sample of maximum and minimum as

$$M_j = \max\{x_{(j-1)N+1}, \dots, x_{jN}\}, \quad (28)$$

and

$$m_j = \min\{x_{(j-1)N+1}, \dots, x_{jN}\}, \quad (29)$$

for $j = 1, \dots, T$. This results in a new sample of size $2T$, consisting of maxima and minima,

$$\{Y_t\}_{t=1}^{2T} = \{m_t, M_t\}_{t=1}^T. \quad (30)$$

For our case, $N = 15$ is a block length providing IID sub-samples based on the Ljung–Box test for serial independence with a significance level of 5%. The left panels of

Figures 9 and 10 depict the series of extreme PETR4 and USD/BRL returns extracted from the blocks, while the left ones show the bimodal form of their distributions.

Thus, we fit these empirical distributions of extreme returns using our suggested invertible bimodal Gumbel model, $F_{IBG}(x; \theta)$, with parameters μ , σ , and δ . Table 5 shows their maximum likelihood estimates. Figure 11 depicts the fitted model against the corresponding distribution of the extracted extreme returns, indicating that the IBG is suitable for simultaneously describing minima and maxima extreme returns.

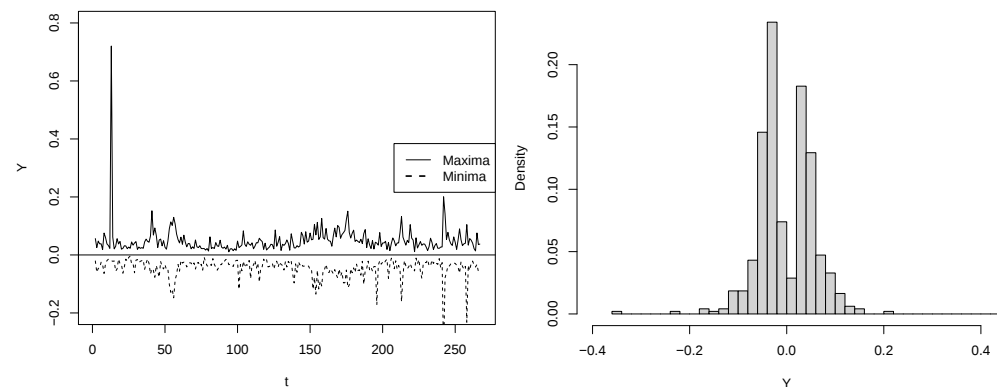


Figure 9. PETR4 sub-sample $\{Y_t\}_{t=1}^{70}$: Extremes obtained from the blocks (left) and histogram (right).

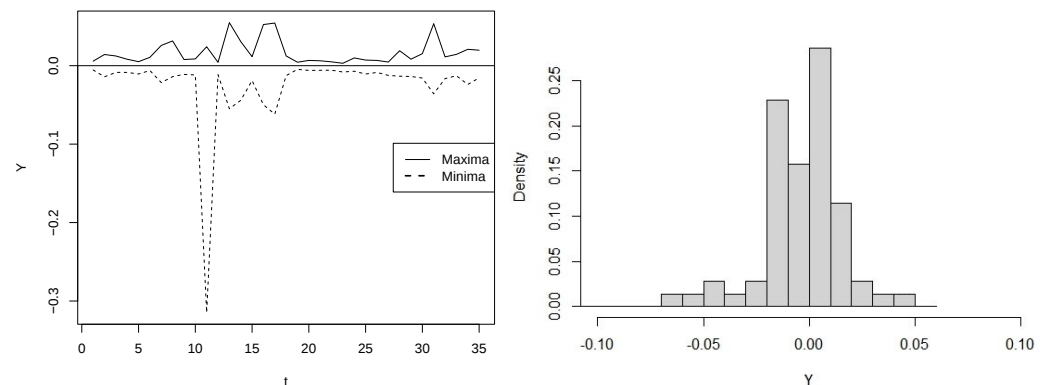


Figure 10. USD/BRL sub-sample $\{Y_t\}_{t=1}^{70}$: Extremes obtained from the blocks (left) and histogram (right).

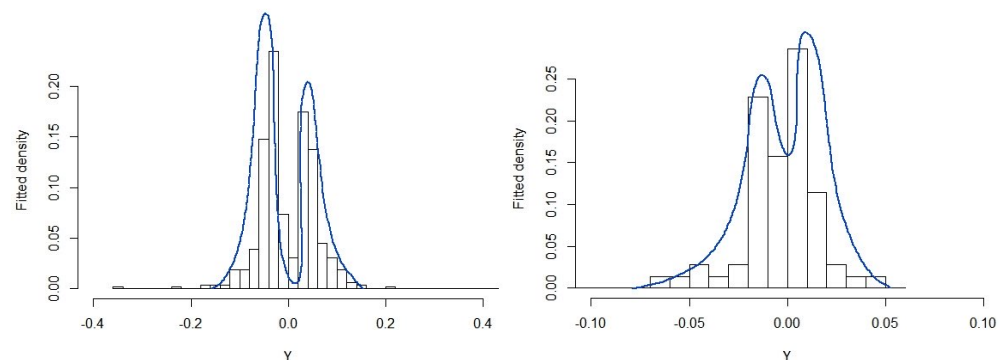


Figure 11. Histogram *versus* fitted distribution: PETR4 (left) and USD/BRL (right).

Finally, Table 6 presents the estimated VaR for $\alpha = 10\%$, 5% , and 1% . As we are dealing with the logarithmic returns, to make these VaR values more understandable, we may consider that the maximum return of the stock is $\exp \text{VaR} - 1$. Thus, for example, over a 15-day period, we do not expect a return greater than 7.4% for PETR4 and 2.6% for USD/BRL with a confidence of 90%.

Table 5. ML estimates of $\theta = (\mu, \sigma, \delta)$ and their respective standard error (SE).

Stock	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\delta}$
PETR4	0.000009962	0.000089877	1.246255323
SE	0.0000071	0.0000003	0.0000083
USD/BRL	0.000016486	0.000099908	1.31295954
SE	0.0000007	0.0000001	0.0000027

Table 6. Estimated VaR_α from (17).

Stock	10%	5%	1%
PETR4	0.07166153	0.07835825	0.09003984
USD/BRL	0.02594245	0.0295429	0.03612925

5. Conclusions

This paper introduced and examined the IBG distribution as an extension of the classical Gumbel distribution. We addressed the limitations of the unimodal Gumbel by proposing a model capable of simultaneously representing both maximum and minimum extremes, enhancing its applicability and versatility. The mathematical formulations accompanied by illustrative figures elucidate the characteristics and behavior of the proposed distribution, emphasizing its advantages in terms of computational efficiency and flexibility. We presented its distributional properties, including mode, moment-generating functions, and entropy. In our illustration, we performed the maximum and minimum blocks technique to obtain serial independent data for the VaR estimation through the ML method, offering a novel perspective on modeling extremes through the lens of the invertible bimodal Gumbel distribution.

Author Contributions: Conceptualization, C.G.O. and E.B.S.; methodology, C.G.O.; software, E.B.S. and A.S.; validation, E.B.S., R.Y.M. and A.S.; formal analysis, C.G.O.; investigation, C.G.O.; resources, E.B.S.; data curation, C.G.O.; writing—original draft preparation, C.G.O.; writing—review and editing, R.Y.M.; visualization, A.S.; supervision, R.Y.M.; project administration, C.G.O.; funding acquisition, C.G.O. and R.Y.M. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financed in part by the Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior—Brazil (CAPES)—Finance Code 001, Fundacao de Apoio a Pesquisa do Distrito Federal (FAPDF)—Finance Code 00193-00001860/2023-17, and National Council for Scientific and Technological Development (CNPq)—Finance Code 311548/2022-9.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: <https://finance.yahoo.com>, accessed on 1 December 2021.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Aubin, D. Forms of explanation in the catastrophe theory of René Thom: Topology, morphogenesis, and structuralism. In *Growing Explanations: Historical Perspectives on Recent Science*; Wise, M.N., Smith, B.H., Weintraub, E.R., Eds.; Duke University Press: New York, NY, USA, 2004; pp. 95–130.
2. Ebeling, W.; Schimansky-Geier, L. Stochastic dynamics of a bistable reaction system. *Phys. A Stat. Mech. Its Appl.* **1979**, *98*, 587–600. [[CrossRef](#)]
3. Smirnov, V.; Ma, Z.; Volchenkov, D. Invited article by M. Gidea Extreme events and emergency scales. *Commun. Nonlinear Sci. Numer. Simul.* **2020**, *90*, 105350. [[CrossRef](#)] [[PubMed](#)]
4. Rossi, F.; Fiorentino, M.; Versace, P. Two-component extreme value distribution for flood frequency analysis. *Water Resour. Res.* **1984**, *20*, 847–856. [[CrossRef](#)]
5. Nadarajah, S. The exponentiated Gumbel distribution with climate application. *Environmetrics* **2006**, *17*, 13–23. [[CrossRef](#)]
6. Aryal, G.R.; Tsokos, C.P. On the transmuted extreme value distribution with application. *Nonlinear Anal. Theory Methods Appl.* **2009**, *71*, 401–407. [[CrossRef](#)]

7. Cooray, K. Generalized Gumbel distribution. *J. Appl. Stat.* **2010**, *37*, 171–179. [[CrossRef](#)]
8. Jeong, B.Y.; Murshed, M.S.; Am Seo, Y.; Park, J.S. A three-parameter kappa distribution with hydrologic application: A generalized Gumbel distribution. *Stoch. Environ. Res. Risk Assess.* **2014**, *8*, 2063–2074. [[CrossRef](#)]
9. Nadarajah, S.; Kotz, S. The beta Gumbel distribution. *Math. Probl. Eng.* **2004**, *4*, 323–332. [[CrossRef](#)]
10. Cordeiro, G.M.; Nadarajah, S.; Ortega, E.M.M. The Kumaraswamy Gumbel distribution. *Stat. Methods Appl.* **2012**, *21*, 139–168. [[CrossRef](#)]
11. Cordeiro, G.M.; Ortega, E.M.M.; Da Cunha, D.C.C. The exponentiated generalized class of distributions. *J. Data Sci.* **2013**, *11*, 1–27. [[CrossRef](#)]
12. Pinheiro, E.C.; Ferrari, S.L.P. A comparative review of generalizations of the Gumbel extreme value distribution with an application to wind speed data. *J. Stat. Comput. Simul.* **2016**, *86*, 2241–2261. [[CrossRef](#)]
13. Okorie, I.E.; Akpanta, A.C.; Ohakwe, J. The Exponentiated Gumbel Type-2 Distribution: Properties and Application. *Int. J. Math. Math. Sci.* **2016**, *10*, 5898356. [[CrossRef](#)]
14. Okorie, I.E.; Akpanta, A.C.; Ohakwe, J.; Chikezie, D.C.; Obi, E.O. The Kumaraswamy G Exponentiated Gumbel type-2 distribution. *Afr. Stat.* **2017**, *12*, 1367–1396.
15. Otiniano, C.E.G.; Sousa, B.; Vila, R.; Bourguignon, M. A Bimodal Model for Extremes Data. *Environ. Ecol. Stat.* **2023**, *30*, 261–288. [[CrossRef](#)]
16. Otiniano, C.E.G.; Vila, R.; Brom, P.C.; Bourguignon, M. On the bimodal Gumbel model with application to environmental data. *Austrian J. Stat.* **2023**, *52*, 45–65. [[CrossRef](#)]
17. Maia, A.; Matsushita, R.; Da Silva, S. Earnings distributions of scalable vs. non-scalable occupations. *Phys. A* **2020**, *560*, 125192.
18. Newman, M.E.J. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **2005**, *46*, 323–351. [[CrossRef](#)]
19. Gumbel, E.J. Les valeurs extremes des distributions statistiques. *Ann. L'Institut Henri Poincaré* **1935**, *5*, 115–158.
20. Jenkinson, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.* **1995**, *81*, 158–171. [[CrossRef](#)]
21. Ahmad, K.E.; Jaheen, Z.; Modhesh, A.A. Estimation of a Discriminant Function Based on Small Sample Size from a Mixture of Two Gumbel Distributions. *Commun. Stat. Simul. Comput.* **2010**, *39*, 713–725. [[CrossRef](#)]
22. Longin, F.M. From value at risk to stress testing: The extreme value approach. *J. Bank. Financ.* **2000**, *24*, 1097–1130. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.