



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

Distribuição Generalizada de Valor Extremo Bimodal

por

Bianca Souza de Paiva

Brasília, 30 de Novembro de 2020

Distribuição Generalizada de Valor Extremo Bimodal

por

Bianca Souza de Paiva

Dissertação apresentada ao Departamento de Estatística da Universidade de Brasília, como requisito parcial para obtenção do título de Mestre em Estatística.

Orientadora: Prof. Dr. Cira Etheowalda Guevara Otiniano

Brasília, 30 de Novembro de 2020

Texto aprovado por:

Prof. Dr. Cira Etheowalda Guevara Otiniano
Orientadora, EST/UnB

Prof. Roberto Vila Gabriel
EST/UnB

Prof. Marcelo Bourguignon Pereira
EST/UFRN

Prof. Eduardo Yoshio Nakano
Suplente, EST/UnB

I am still learning.

(Michelangelo)

Para minha mãe, irmãs e sobrinhos.

Agradecimentos

Meus sinceros agradecimentos à professora Cira Etheowalda Guevara Otiniano do PPGEST/UnB por todo apoio, por ter acreditado na minha capacidade mesmo quando eu mesmo não acreditei, pelas palavras em momentos difíceis e por toda paciência e compreensão.

Agradeço também à minha família, em especial à minha mãe Elizabeth, às minhas irmãs Lorryne e Débora, meus sobrinhos amados Igor e Lia e meus avós Francisco e Minervina. Muito obrigada por serem meu porto seguro, minha base e por sempre me mostrarem a direção de um mundo melhor.

Obrigada também a todos aqueles que estiveram me incentivando durante este trabalho, com palavras encorajadoras, com acalentos e todos os gestos e estímulos que me fizeram chegar até o fim. Meus mais profundos agradecimentos.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

A distribuição de valor extremo generalizada, conhecida como GEV, é amplamente utilizada em hidrologia e finanças para modelar eventos extremos. Neste trabalho propomos o modelo bimodal gev, ou BGEV, que generaliza a distribuição GEV e apresenta bimodalidade para determinadas combinações dos parâmetros. Fazemos um estudo gráfico intenso do comportamento da BGEV, em seguida o estudo de suas medidas descritivas como: moda, quantis e momentos. Utilizamos o método de máxima verossimilhança (MV) para estimar os parâmetros da BGEV. Através de experimentos de simulação de Monte Carlo verificamos o bom desempenho dos estimadores de MV. Além disso, adicionamos uma aplicação à dados climáticos.

Sumário

1	Conceitos Preliminares	3
1.1	Distribuição Generalizada de Valor Extremo - GEV	3
1.2	Moda	6
1.3	Quantis	6
1.4	Momentos	7
2	Distribuição Bimodal GEV - BGEV	9
2.1	Definição	9
2.1.1	Função de distribuição acumulada BGEV	11
2.1.2	Estudo gráfico Bimodal GEV: $\xi \neq 0$	12
2.1.3	Estudo gráfico Bimodal GEV: $\xi = 0$	15
2.2	Monotonicidade	17
2.3	Moda	19
2.4	Momentos	19
2.4.1	Momentos: caso $\xi \neq 0$	19
2.4.2	Momentos: caso $\xi = 0$	24
2.5	Quantis	25
2.5.1	Mediana	26
3	Estimação	27

3.1	Estimação por máxima verossimilhança	27
3.2	Estatísticas de qualidade de ajuste	29
4	Simulação	31
5	Aplicação	45
6	Conclusão	53
6.1	Anexo A	57
6.2	Anexo B	73

Introdução

A grande aplicabilidade da distribuição generalizada de valores extremos decorre da sua capacidade de modelar acontecimentos raros (de baixa frequência), porém com grande impacto. A modelagem e previsão desse tipo de acontecimentos são de particular interesse nas áreas de finanças, hidrologia, atuária, bioinformática, entre outras.

A distribuição generalizada de valores extremos, popularmente conhecida por GEV, do inglês: *generalized extreme value*, baseia-se na modelagem de eventos extremos. Segundo a teoria dos valores extremos (TVE), as distribuições de valor extremo são distribuições limites das estatísticas extremas, máximo ou mínimo normalizado de variáveis aleatórias independentes e identicamente distribuídas (iid). A teoria clássica e aplicações da TVE podem ser encontradas nos livros de Galambos 1978, Resnick 1987 e de Haas e Ferreira 2006.

No ano de 1928, Fisher e Tippett, e em 1943 Gnedenko, os principais desenvolvedores da teoria, mostraram que uma distribuição de valor extremo pertence à um dos três tipos: Gumbel, Fréchet e Weibull, que também são conhecidas como distribuições de valor extremo tipo I, II e III, respectivamente. Posteriormente, Jenkinson [1955] mostrou que as três distribuições extremas mencionadas podem ser escritas em uma única distribuição, a distribuição generalizada de valores extremos.

Historicamente, as primeiras aplicações da TVE se concentraram na área de hidrologia, buscando analisar a frequência de fluídos em fenômenos como precipitações e inundações, e também em modelagens de terremotos e emissões radioativas, dentre outros. Porém, com o desenvolver da teoria, sua aplicabilidade tem se tornado cada vez mais ampla, algumas das

principais áreas são: atuária, para cálculo da probabilidade da ruína por eventos extremos, finanças para calcular a medida de risco VaR (Value at Risk), dentre outros exemplos que podem ser encontrados em Embrechts et al 1999 e Jondeau 2007.

Entretanto, conforme Otiniano et al. (2019), existem dados de eventos extremos que não seguem a distribuição GEV, porque eles requerem uma distribuição mais assimétrica ou com uma cauda mais pesada do que a distribuição GEV. Assim, foram desenvolvidas novas classes de distribuições de probabilidade que são mais gerais do que a distribuição GEV, tais como: distribuição dual gamma GEV (GGEV) e distribuição GEV exponenciada (EGEV) estudadas por Nascimento et al. (2016), distribuição transmutada GEV (TGEV) definida por Aryal e Tsokos (2009), e q-GEV definida por Provost et al. (2018).

Estudos empíricos com dados reais de extremos correspondentes à velocidade de vento, pontos de ovalho e radioatividade, entre outros, tem apresentado bimodalidade. Acredita-se que isso deve estar ocorrendo devido à incidência de mudanças climáticas que ocasionam eventos com mudanças de regime. Por outro lado, alguns dados de finanças como preço do petróleo no Brasil, também apresentam bimodalidade.

Motivados com a busca de uma distribuição que melhor se ajuste esse tipo de dados, neste trabalho propomos uma extensão da distribuição GEV que apresenta bimodalidade para determinados conjuntos de parâmetros. Aqui denominada de Bimodal GEV ou BGEV.

Para tanto, este trabalho está dividido em quatro capítulos. No Capítulo 1, apresentamos o conceito e propriedades básicas da distribuição GEV que será a base do estudo. No Capítulo 2, definimos a distribuição BGEV e investigamos algumas de suas medidas de probabilidade como momentos e quantis. Finalmente no Capítulo 3, são apresentadas expressões que devem ser resolvidas para obter os estimadores de máxima verossimilhança dos parâmetros da BGEV. O capítulo 4, apresenta um estudo de simulação da distribuição BGEV. Para analisar a distribuição desenvolvida, no Capítulo 5 a mesma será utilizada para modelar um banco de dados reais.

Capítulo 1

Conceitos Preliminares

O presente capítulo apresenta conceitos e propriedades importantes sobre a distribuição generalizada de valores extremos que serão úteis ao longo deste trabalho.

1.1 Distribuição Generalizada de Valor Extremo - GEV

Proposta por Jenkinson (1955) e advinda da teoria de valores extremos, a distribuição generalizada de valores extremos consiste em uma família de distribuições desenvolvidas por Fisher e Tippett (1928).

Uma distribuição de valor extremo é a distribuição limite não degenerada do máximo (ou mínimo) normalizado de variáveis aleatórias independentes e identicamente distribuídas (iid) X_1, X_2, \dots, X_n . Isto é

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} [F(a_n x + b_n)]^n = G(x),$$

sendo que M_n representa $\max\{X_1, X_2, \dots, X_n\}$ (ou o mínimo), $a_n > 0$ e $b_n \in \mathbb{R}$ são constantes de normalização, e F a função de distribuição acumulada da variável comum às variáveis $\{X_1, X_2, \dots, X_n\}$. A distribuição G pertence à um dos três tipos das seguintes distribuições:

Gumbel

$$G(x) = \exp\{-\exp[-x]\}, -\infty < x < \infty;$$

Fréchet

$$G(x) = \begin{cases} 0 & , x \leq 0, \\ \exp(-x^{-\alpha}) & , x > 0, \alpha > 0; \end{cases}$$

Weibull

$$G(x) = \begin{cases} \exp[-(-x)^\alpha] & , x < 0, \alpha > 0, \\ 1 & , x \geq 0; \end{cases}$$

sendo $\alpha = \xi^{-1}$.

Para propósitos estatísticos não é conveniente trabalhar com as três distribuições separadamente, é preferível adotar uma parametrização que unifique as três distribuições, Eastoe e Tawn (2007).

Von Mises (1954) e Jenkinson (1955) obtiveram a distribuição $GEV(\mu, \sigma, \xi)$ que reúne as distribuições extremais com função de distribuição acumulada caracterizada pela equação (1.1).

$$F_{GEV}(y; \mu, \sigma, \xi) = \begin{cases} \exp\left\{-\left[1 + \xi\left(\frac{y - \mu}{\sigma}\right)\right]^{-1/\xi}\right\} & , \xi \neq 0, \\ \exp\left\{-\exp\left(-\frac{y - \mu}{\sigma}\right)\right\} & , \xi = 0. \end{cases} \quad (1.1)$$

Uma variável aleatória Y tem distribuição generalizada de valor extremo (GEV) se sua função de densidade de probabilidade for definida segundo a equação (1.2) a seguir:

$$f_{GEV}(y; \mu, \sigma, \xi) = \begin{cases} \sigma^{-1} [1 + \xi w]^{(-1/\xi)-1} \exp\left\{-[1 + \xi w]^{(-1/\xi)}\right\} & , \xi \neq 0, \\ \sigma^{-1} \exp\{-w - \exp[-\exp(-w)]\} & , \xi = 0, \end{cases} \quad (1.2)$$

§1.2. Distribuição Generalizada de Valor Extremo - GEV

em que $w = \frac{y-\mu}{\sigma}$, válida para $y > \mu - \frac{\sigma}{\xi}$ no caso de $\xi > 0$, $y < \mu - \frac{\sigma}{\xi}$ para o caso $\xi < 0$ e $y \in \mathbb{R}$ quando $\xi = 0$. Sendo o parâmetro de forma ξ também denominado de índice caudal, $\sigma > 0$ o parâmetro de escala e μ o parâmetro de locação da distribuição.

A Figura 1.1 mostra o comportamento da distribuição GEV para diferentes valores de ξ . A densidade da GEV é concentrada no suporte positivo para $\xi > 0$, no suporte negativo para $\xi < 0$ e em toda reta dos reais para $\xi = 0$.

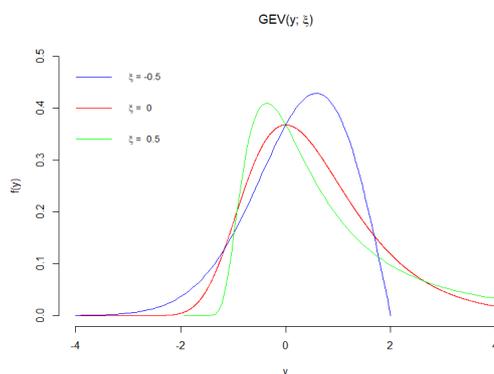


Figura 1.1: Gráficos da GEV para $\xi > 0$, $\xi < 0$ e $\xi = 0$.

Neste trabalho será utilizado um caso particular da distribuição GEV, com o parâmetro de escala constante e igual à 1 ($\sigma = 1$), definido aqui como sendo GEV^* , assim, uma variável aleatória Y^* segue a distribuição GEV^* se sua função de distribuição acumulada e sua função de densidade de probabilidade são caracterizadas por:

$$\begin{cases} F_{GEV^*}(y^*; \mu, \xi) = F_{GEV}(y^*; \mu, \sigma = 1, \xi), \\ f_{GEV^*}(y^*; \mu, \xi) = f_{GEV}(y^*; \mu, \sigma = 1, \xi). \end{cases} \quad (1.3)$$

As próximas sessões apresentarão algumas medidas estatísticas descritivas características da distribuição GEV com parâmetros de locação (μ), escala (σ) e forma (ξ).

1.2 Moda

A moda é uma medida de tendência central que representa o valor observado com maior frequência de uma distribuição, valor tal em que a derivada da densidade é igual à zero.

Se $Y \sim GEV(y; \mu, \sigma, \xi)$, a moda da distribuição, aqui denotada por y_M , tal que $f'_{GEV}(y_M; \mu, \sigma, \xi) = 0$ é dada por:

$$y_M = \begin{cases} \mu + \sigma \frac{(1+\xi)^{-\xi-1}}{\xi} & , \xi \neq 0, \\ \mu & , \xi = 0. \end{cases} \quad (1.4)$$

A moda portanto representa o valor de máximo local da função de densidade de probabilidade da distribuição.

1.3 Quantis

A função quantil é obtida a partir da inversa da função de distribuição acumulada e retorna o valor associado à uma dada probabilidade acumulada e também pode ser usada para simular dados da distribuição.

Se $Y \sim GEV(y; \mu, \sigma, \xi)$, cuja função de distribuição acumulada

$$P(Y \leq y_Q) = F_{GEV}(y_Q; \mu, \sigma, \xi) = Q$$

é dada em 1.1, uma função inversível, então o Q-quantil de Y , aqui escrito como y_Q , é dado por:

$$y_Q = \begin{cases} \mu + \frac{\sigma}{\xi} [(-\log(Q))^{-\xi} - 1] & , \xi \neq 0; \\ \mu - \sigma \log(-\log(Q)) & , \xi = 0, \end{cases} \quad (1.5)$$

em que Q pertence ao intervalo $(0, 1]$.

Em particular, quando $Q = 0,5$ a função quantil retorna a mediana da distribuição. Sendo assim, a mediana da distribuição GEV está representada em (1.6).

$$y_{0,5} = \begin{cases} \mu + \frac{\sigma}{\xi} [\log(2)^{-\xi} - 1] & , \xi \neq 0; \\ \mu - \sigma \log(\log(2)) & , \xi = 0. \end{cases} \quad (1.6)$$

A mediana é uma medida de tendência central que corresponde ao valor central de um conjunto de valores ordenados, ou seja, o valor que separa a metade maior e a metade menor de uma amostra, uma população ou uma distribuição de probabilidade.

1.4 Momentos

Os momentos da distribuição GEV são obtidos de diferentes maneiras ao considerar $\xi = 0$ e $\xi \neq 0$. Para $\xi = 0$, os momentos são obtidos a partir da função geradora de momentos $E(e^{tY})$ já para $\xi \neq 0$ os momentos inteiros $E(Y^k)$ tem forma fechada.

Quando $\xi = 0$, a função geradora de momentos de $Y \sim GEV(y; \mu, \sigma, \xi)$ é dada por

$$E(e^{ty}) = \Gamma(1 - \sigma t) e^{\mu t}. \quad (1.7)$$

Ao calcular as derivadas de (1.7) no ponto $t = 0$, obtem-se os momentos da distribuição.

$$\begin{aligned}
E(Y^1) &= \mu + \sigma\gamma; \\
E(Y^2) &= \sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^2; \\
E(Y^3) &= 2\sigma^3 \zeta(3) + 3(\mu + \sigma\gamma)\sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^3; \\
E(Y^4) &= \sigma^4 \frac{\pi^4}{15} + 4(\mu + \sigma\gamma)(2\sigma^3 \zeta(3) + 3(\mu + \sigma\gamma)\sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^3) \\
&\quad - (\sigma^2 \frac{\pi^2}{6} + 3(\mu + \sigma\gamma)^2)(\sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^2),
\end{aligned} \tag{1.8}$$

sendo γ a constante de Euler, definida como o limite da diferença entre a série harmônica e o logaritmo natural $\gamma = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \ln(n) \right)$ e $\zeta(\cdot)$ a função zeta de Riemann, uma função especial de variável complexa, definida pela série $\zeta(s) = \sum_{k=1}^{\infty} k^{-s}$.

Já para o caso em que $\xi \neq 0$ a fórmula fechada para os momentos é dada por

$$E(Y^k) = \sum_{i=1}^k \binom{k}{i} \left(\mu - \frac{\sigma}{\xi} \right)^{k-i} \left(\frac{\sigma}{\xi} \right)^i \Gamma(1 - \xi i). \tag{1.9}$$

Ao considerar $k = 1$ e $k = 2$ obtém-se a média e variância da distribuição GEV, obtidas ao calcular a integral $E(Y^k) = \int_{\mu - \frac{\sigma}{\xi}}^{\infty} y^k f_{GEV}(y; \mu, \sigma, \xi) dy$ para $\xi > 0$ e a integral $E(Y^k) = \int_{-\infty}^{\mu - \frac{\sigma}{\xi}} y^k f_{GEV}(y; \mu, \sigma, \xi) dy$ para $\xi < 0$.

$$E(Y) = \mu + \frac{\sigma}{\xi}(\Gamma(1 - \xi) - 1), \quad Var(Y) = \frac{\sigma^2}{\xi^2}(\Gamma(1 - 2\xi) - \Gamma^2(1 - \xi)). \tag{1.10}$$

Neste caso, existe uma restrição em relação a quantidade de momentos calculáveis. Os momentos somente são calculados para $\xi < 1/k$, sendo assim, para que seja viável, por exemplo, obter o segundo momento da distribuição GEV é necessário que $\xi < 1/2$, e assim por diante.

Capítulo 2

Distribuição Bimodal GEV - BGEV

Neste capítulo serão apresentados os resultados principais deste trabalho. É proposta uma modificação da distribuição GEV com o objetivo de torna-la eficiente na modelagem de dados extremos que podem apresentar bimodalidade.

2.1 Definição

Uma variável aleatória X tem distribuição bimodal generalizada de valor extremo, com parâmetros μ, σ, ξ e δ , denotada por $X \sim BGEV(x; \mu, \sigma, \xi, \delta)$ se sua função de densidade de probabilidade é dada por

$$f(x; \mu, \sigma, \xi, \delta) = \begin{cases} f_{GEV^*}(T_\delta(x); \mu, \xi) T'_\delta(x) & , \xi \neq 0, \delta > -1, \\ f_{GEV}(x; \mu, \sigma, \xi) \frac{1}{Z_\delta} ((1 - \delta x)^2 + 1) & , \xi = 0, \delta \in \mathbb{R}, \end{cases} \quad (2.1)$$

em que f_{GEV} é a distribuição GEV com parâmetros de localização (μ), escala (σ) e forma (ξ) como definido em 1.2 e f_{GEV^*} o caso particular com escala constante ($\sigma = 1$) apresentado em (1.3) .

A constante Z_δ é definida por:

$$Z_\delta = 1 + \delta^2 \sigma^2 \frac{\pi^2}{6} + (\delta\mu + \delta\sigma\gamma - 1)^2, \quad (2.2)$$

sendo γ a constante de Euler. A transformada dada por $T_\delta(x)$ é bijetiva de \mathbb{R} em \mathbb{R}

$$T_\delta(x) = x(\sigma|x|^\delta), \quad x \in \mathbb{R}, \sigma > 0 \quad (2.3)$$

é inversível, com inversa dada por:

$$T^{-1}(y) = \begin{cases} \left(\frac{y}{\sigma}\right)^{\frac{1}{1+\delta}} & , y > 0 \\ \left(\frac{y}{\sigma(-1)^\delta}\right)^{\frac{1}{1+\delta}} & , y < 0 \end{cases} \quad (2.4)$$

e com primeira e segunda derivadas, respectivamente:

$$T'_\delta(x) = \sigma(\delta + 1)|x|^\delta, \quad T''_\delta(x) = \sigma(\delta + 1)\delta \operatorname{sign}(x)|x|^{\delta-1}. \quad (2.5)$$

Em geral, a transformada T é de classe C^k com k -ésima derivada

$$T_\delta^{(k)}(x) = (\operatorname{sign}(x))^{k-1} \sigma \left(\prod_{i=1}^{k-2} (\delta - i) \right) |x|^{\delta-(k-1)}, \quad k \geq 2,$$

em que:

$$\operatorname{sign}(x) = \begin{cases} 1 & , x > 0 \\ 0 & , x = 0 \\ -1 & , x < 0 \end{cases}$$

O domínio da distribuição é definido pelas expressões $x > T_\delta^{-1}(\mu - 1/\xi)$ e $x < T_\delta^{-1}(\mu - 1/\xi)$ no caso de $\xi > 0$ e $\xi < 0$, respectivamente, e por $x \in \mathbb{R}$ para $\xi = 0$.

De fato, pode-se demonstrar que a distribuição BGEV dada por (2.1) é uma função de densidade de probabilidade (fdp), pois para $\xi > 0$ e $\delta > -1$ tem-se

$$\int_{T_\delta^{-1}(\mu - 1/\xi)}^{\infty} f(x; \mu, \sigma, \xi, \delta) dx = \int_{\mu - 1/\xi}^{\infty} f_{\text{GEV}^*}(T_\delta(x); \xi) dT_\delta(x) = 1.$$

Similarmente, também é verificado que, para $\xi < 0$, f é uma fdp.

Para $\xi = 0$ e $\delta \in \mathbb{R}$, considere uma variável aleatória $Y \sim \text{GEV}(\xi, \mu, \sigma, \xi = 0)$, então tem-se que $E(Y) = \mu + \sigma\gamma$ e $\text{Var}(Y) = \sigma^2 \frac{\pi^2}{6}$, note que

$$\begin{aligned} \int_{-\infty}^{\infty} f(x; \mu, \sigma, \xi, \delta) dx &= \int_{-\infty}^{\infty} \frac{1}{Z_\delta} ((1 - \delta x)^2 + 1) f_Y(x) dx \\ &= \frac{1}{Z_\delta} [2 - 2\delta E(Y) + \delta^2 E(Y^2)] \\ &= \frac{1}{Z_\delta} Z_\delta = 1. \end{aligned}$$

2.1.1 Função de distribuição acumulada BGEV

A função de distribuição acumulada associada à uma variável aleatória $X \sim \text{BGEV}(x; \mu, \sigma, \xi, \delta)$ é definida por:

$$F(x; \mu, \sigma, \xi, \delta) = \begin{cases} F_{\text{GEV}^*} \circ T_\delta(x) & , \xi \neq 0, \delta > -1, \\ \frac{1}{Z_\delta} \left[2F_{\text{GEV}}(x; \mu, \sigma, \xi) \right. \\ \left. - 2\delta E(Y I_{Y \leq x}) + \delta^2 E(Y^2 I_{Y \leq x}) \right] & , \xi = 0, \delta \in \mathbb{R}, \end{cases} \quad (2.6)$$

em que $x > T_\delta^{-1}(\mu - 1/\xi)$ para $\xi > 0$, $x < T_\delta^{-1}(\mu - 1/\xi)$ no caso de $\xi < 0$, e $x \in \mathbb{R}$ para $\xi = 0$.

O conhecimento da função de distribuição acumulada é suficiente para caracterizar o comportamento de uma variável aleatória.

Proposição 2.1. Seja $\xi \neq 0$ e $\delta > -1$.

1. Se $X \sim \text{BGEV}(\mu, \sigma, \xi, \delta)$, então $Y = T_\delta(X) \sim \text{GEV}^*(\mu, \xi)$.
2. Se $Y \sim \text{GEV}^*(\mu, \xi)$, então $X = T_\delta^{-1}(Y) \sim \text{BGEV}(\mu, \sigma, \xi, \delta)$.

Demonstração. Se $X \sim \text{BGEV}(\mu, \sigma, \xi, \delta)$, segue que

$$P(Y \leq y) = P(X \leq T_\delta^{-1}(y)) = F(T_\delta^{-1}(y)) = (F_{\text{GEV}} \circ T_\delta)(T_\delta^{-1}(y)) = F_{\text{GEV}}(y).$$

Isso prova o item (1). A prova do segundo item segue analogamente. □

2.1.2 Estudo gráfico Bimodal GEV: $\xi \neq 0$

Com o intuito de avaliar a maneira como os parâmetros modificam a distribuição BGEV, foram geradas diversas densidades com variações de cada parâmetro, mantendo os demais constantes. Para tanto, foram criados alguns cenários para melhor visualização do efeito dos parâmetros na distribuição. O primeiro apresenta variações da distribuição BGEV para $\xi < 0$, o segundo para $\xi > 0$, em ambos foi mantido constante o valor do parâmetro de escala $\sigma = 1$ e realizadas variações no parâmetro δ e no parâmetro μ . Há ainda um terceiro cenário com $\xi = 0$ no qual foram simuladas variações de δ para μ constante igual a 0 e variações de μ mantendo $\delta = 3$, também com valor de σ constante iguala 1, dado que σ é um parâmetro escalar.

Para os cenários em que $\xi \neq 0$, como indicados nas Figuras 2.1 e 2.2, em que tem-se o parâmetro delta variando negativamente dentro do intervalo $(-1, 0]$ para diferentes valores negativos de ξ e $\mu = 0$ (Figura 2.1) e para diferentes valores positivos de ξ e $\mu = 0$ (Figura 2.2), não observa-se bimodalidade da distribuição. No valor de $\delta = 0$ a distribuição coincide com a distribuição de origem GEV, com mesmo valor de ξ e $\sigma = 1$ e $\mu = 0$. Nota-se que a medida que

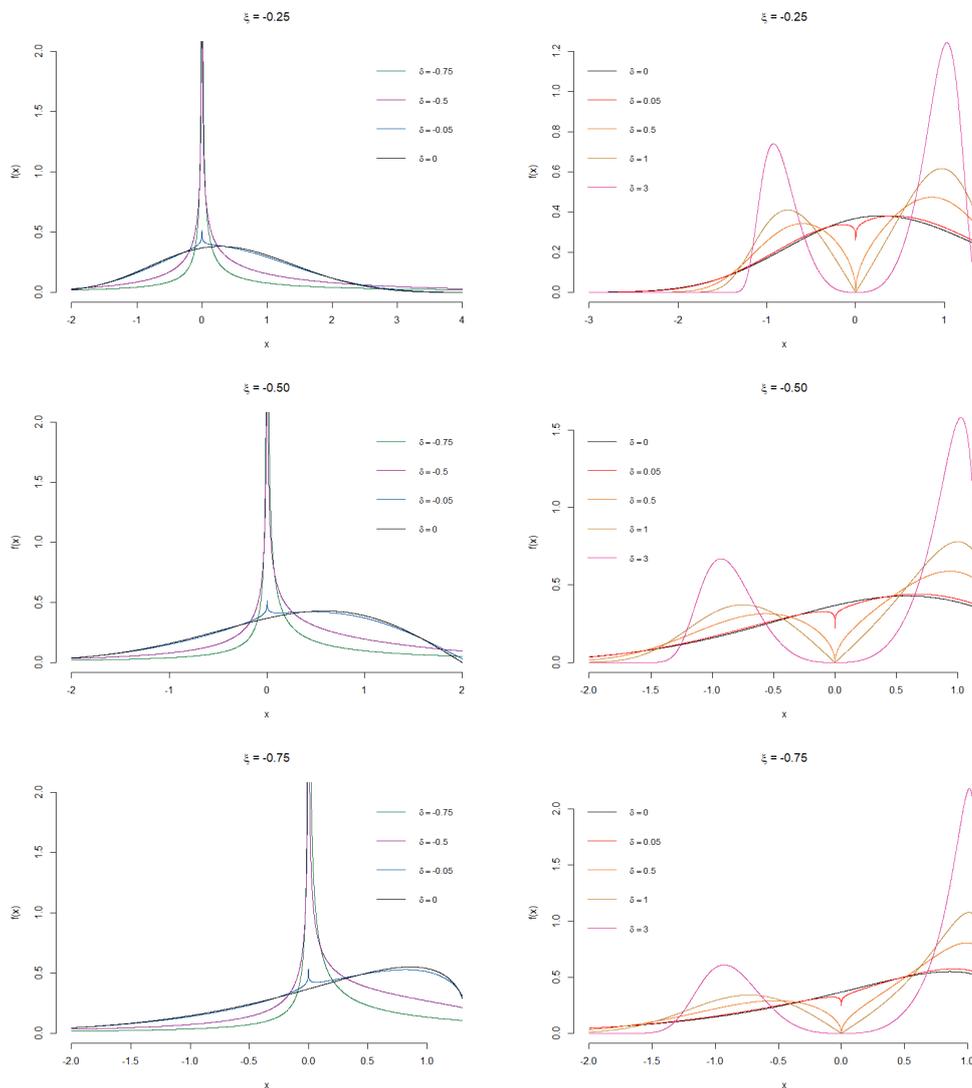


Figura 2.1: Densidade de $Y \sim BGEV(y; \xi, 0, 1, \delta)$, $\xi < 0$ e $\delta \in \{-0,75; -0,50; -0,05; 0; 0,05; 0,50; 1; 3\}$.

o parâmetro diminui, em direção a $\delta = -1$ a distribuição se mostra cada vez mais concentrada, apresentando um pico. Analisando as variações positivas de δ , percebe-se que quanto maior é o valor do parâmetro, mais evidente se mostra a bimodalidade.

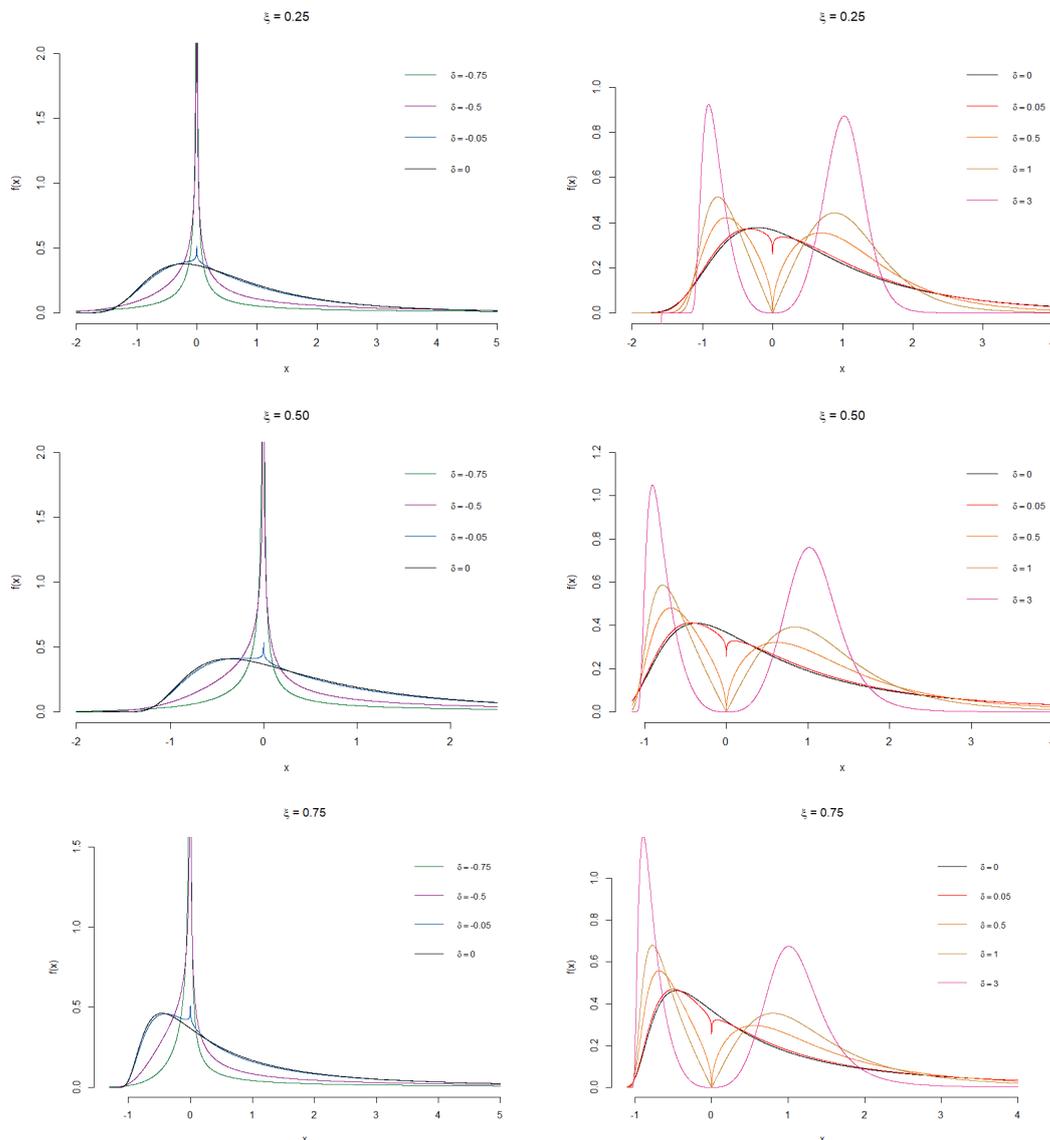


Figura 2.2: Densidade de $Y \sim BGEV(y; \xi, 0, 1, \delta)$, $\xi > 0$ e $\delta \in \{-0,75; -0,50; -0,05; 0; 0,05; 0,50; 1; 3\}$.

Ainda no caso em que $\xi \neq 0$, com o intuito de avaliar a maneira que o parâmetro μ modifica a distribuição BGEV, foram realizadas algumas variações de μ . A Figura 2.3 apresenta variações do parâmetro μ para os demais constantes sendo $\sigma = 1$, $\xi = 3$ e $\delta = 3$. Pode-se notar que a distribuição perde sua característica bimodal entre os valores 0,30, onde ainda existe bimodalidade, e 0,35 em que já observamos que a distribuição se torna unimodal.

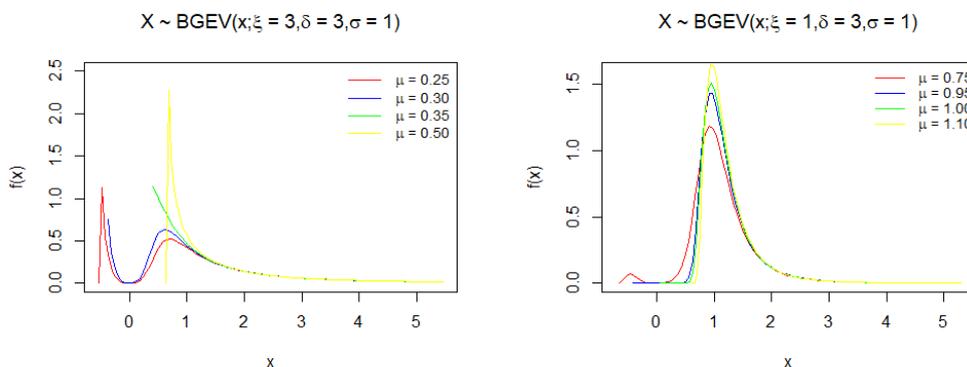


Figura 2.3: Densidade de $Y \sim BGEV$, com variações de μ

Outro exemplo de variação de μ , em que $\xi = 2$ e $\delta = 3$, nota-se que a distribuição também perde sua bimodalidade. Na Figura 2.3 em que as μ assume os valores 0,75, 0,95, 1,00 e 1,10, observa-se que a bimodalidade é nítida para $\mu = 0,75$, e para os outros valores ela é quase imperceptível ou nula.

2.1.3 Estudo gráfico Bimodal GEV: $\xi = 0$

No terceiro cenário, em que tem-se o valor de $\xi = 0$, a Figura 2.4 apresenta ilustrações para variações nos parâmetros δ e μ , sendo que para as realizações na variação de δ foi considerado $\mu = 0$ e para as variações de μ foi mantido δ constante igual a 3. Em todos os casos o parâmetro de escala é fixo, $\sigma = 1$.

Nota-se nas variações de valores negativos de δ que a bimodalidade torna-se evidente a partir de aproximadamente $\delta = -1$, sendo que para valores menores que este a bimodalidade é claramente visível. Em contrapartida, em suas variações positivas, o parâmetro δ gera bimodalidade na distribuição para valores próximos e maiores do que 1, como é visto na curva para $\delta = 0,9$, posto que, quanto maior seu valor, mais notória é a bimodalidade da distribuição.

Na análise das variações do parâmetro μ para $\xi = 0$ e $\delta = 3$ verifica-se que a atuação do parâmetro se dá de maneira que, para valores negativos, quanto menor o valor mais valores serão observados próximos a menor moda da distribuição até que ela se torne novamente unimodal.

O comportamento se dá de maneira semelhante para as variações positivas, porém no sentido inverso.

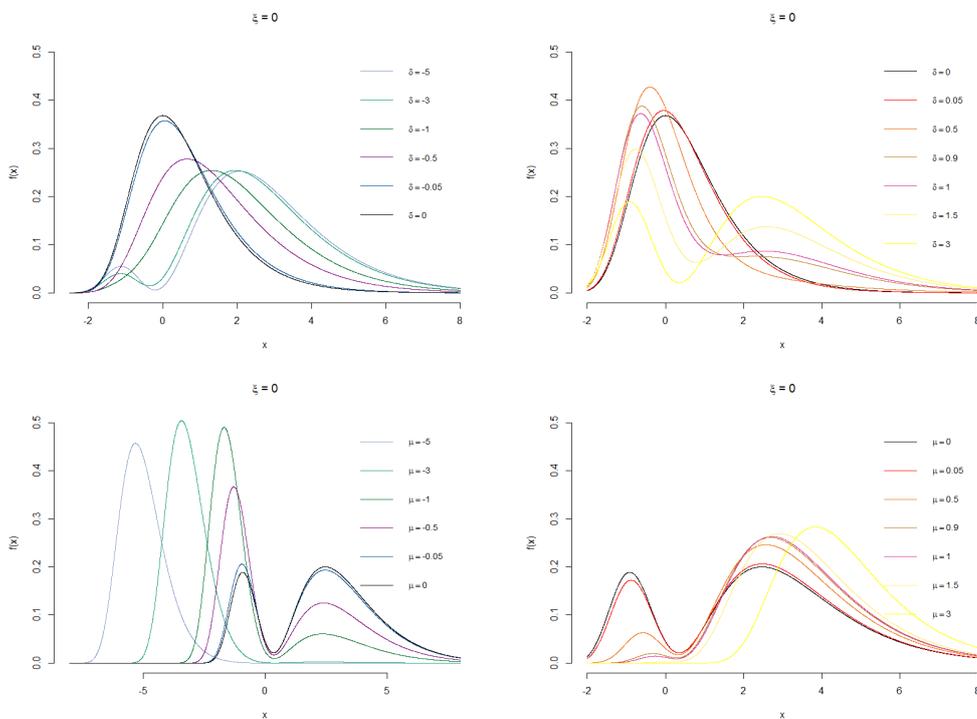


Figura 2.4: Densidade de $Y \sim BGEV(y; \xi = 0, 0, 1, \delta)$, com variações de δ e μ .

Como no caso em que $\xi \neq 0$, o modelo BGEV retorna a distribuição GEV (ou Gumbel, dado que $\xi = 0$) quando o parâmetro $\delta = 0$ sendo, portanto, uma generalização da mesma.

2.2 Monotonicidade

Para a análise da monotonicidade da distribuição, considere o parâmetro m_ξ definido por:

$$m_\xi = \mu \cdot \delta_{\xi,0} + \left[\mu + \frac{(1 + \xi)^{-\xi} - 1}{\xi} \right] \cdot (1 - \delta_{\xi,0}) \quad \text{para } \xi \in \mathbb{R} \text{ e } \mu \in \mathbb{R},$$

em que $\delta_{\xi,0}$ é o delta de Kronecker, tal que

$$\delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j \end{cases}$$

Seja ainda $x_\xi^{\min} = \min\{0, T_\delta^{-1}(m_\xi)\}$ e $x_\xi^{\max} = \max\{0, T_\delta^{-1}(m_\xi)\}$. Agora podemos afirmar os seguintes resultados:

Proposição 2.2. A densidade da distribuição BGEV, quando $\xi \neq 0$, dada em (2.1), tem as seguintes propriedades de monotonicidade:

- 1) Se $\xi > 0$ e $\delta < 0$, então $f(x; \mu, \sigma, \xi, \delta)$ está aumentando para $x < x_\xi^{\min}$;
- 2) Se $\xi > 0$ e $\delta > 0$, então $f(x; \mu, \sigma, \xi, \delta)$ está diminuindo para $x < x_\xi^{\min}$;
- 3) Se $\xi < 0$ e $\delta < 0$, então $f(x; \mu, \sigma, \xi, \delta)$ está diminuindo para $x > x_\xi^{\max}$;
- 4) Se $\xi < 0$ e $\delta > 0$, então $f(x; \mu, \sigma, \xi, \delta)$ está aumentando para $x > x_\xi^{\max}$.

Demonstração. Primeiramente, tendo $T'_\delta(x) = \sigma(\delta + 1)|x|^\delta$ como já mencionado na equação (2.5), observa-se que, para $\delta < 0$,

$$T'_\delta(x) \begin{cases} \text{é decrescente para } x > 0, \\ \text{é crescente para } x < 0, \end{cases}$$

no caso de $\delta > 0$,

$$T'_\delta(x) \begin{cases} \text{é crescente para } x > 0, \\ \text{é decrescente para } x < 0. \end{cases}$$

Por outro lado, é sabido que m_ξ é a única moda da distribuição $f_{\text{GEV}}(x; \xi, \mu)$ definida como (1.3). Consequentemente,

$$f(T_\delta(x); \xi, \mu) \begin{cases} \text{é crescente para } x < x_\xi^{\min} \text{ e } \xi > 0, \\ \text{é decrescente para } x > x_\xi^{\max} \text{ e } \xi < 0. \end{cases}$$

Quando $\xi > 0$ e $\delta < 0$, $f(T_\delta(x); \mu, \xi)$ é uma função crescente para $x < x_\xi^{\min}$. Sendo $T'_\delta(x)$ crescente e não negativa para $x < 0$, pela definição da densidade da distribuição BGEV (2.1), tem-se que $f(x; \mu, \sigma, \xi, \delta) = f_{\text{GEV}^*}(T_\delta(x); \mu, \xi) T'_\delta(x)$ é o produto de duas funções crescentes e não negativas. Assim, a função de densidade BGEV $f(x; \mu, \sigma, \xi, \delta)$ é crescente para $x < x_\xi^{\min}$. Isso completa a prova do primeiro item.

A comprovação dos demais itens segue o mesmo raciocínio do Item 1). □

Proposição 2.3. A densidade da distribuição BGEV, quando $\xi = 0$, dada em (2.1), tem as seguintes propriedades de monotonicidade:

- 1) Se $\delta < 0$, então $f(x; \mu, \sigma, \xi, \delta)$ é crescente para $x < x_\xi^{\min}$;
- 2) Se $\delta > 0$, então $f(x; \mu, \sigma, \xi, \delta)$ é decrescente para $x < x_\xi^{\min}$;
- 3) Se $\delta < 0$, então $f(x; \mu, \sigma, \xi, \delta)$ é decrescente para $x > x_\xi^{\max}$;
- 4) Se $\delta > 0$, então $f(x; \mu, \sigma, \xi, \delta)$ é crescente para $x > x_\xi^{\max}$.

Demonstração. A prova desta proposição segue o mesmo raciocínio da prova da Proposição 2.2. □

2.3 Moda

Uma característica descritiva de especial interesse neste trabalho é a moda da distribuição.

Proposição 2.4. Se um ponto x_M é a moda da distribuição BGEV com $\xi \neq 0$ e $\delta \geq -1$, então é solução da seguinte equação

$$\frac{T_\delta''(x_M)}{[T_\delta'(x_M)]^2} = \frac{1 + \xi - [1 + \xi(T_\delta(x_M) - \mu)]^{-1/\xi}}{1 + \xi(T_\delta(x_M) - \mu)}, \quad (2.7)$$

em que $T_\delta(x)$, $T_\delta'(x)$ e $T_\delta''(x)$ são dadas em 2.3 e 2.5.

Proposição 2.5. Se um ponto x_M é a moda da distribuição BGEV com $\xi = 0$ e $\delta \in \mathbb{R}$, então é solução da equação

$$\frac{2\delta(1 - \delta x_M)}{(1 - \delta x_M)^2 + 1} = \frac{f'_{\text{GEV}}(x_M; \xi, \mu, \sigma)}{f_{\text{GEV}}(x_M; \xi, \mu, \sigma)} = \frac{1}{\sigma} \left\{ \exp \left[- \left(\frac{x_M - \mu}{\sigma} \right) \right] - 1 \right\}. \quad (2.8)$$

Proposição 2.6. Se $\xi = 0$ e $\delta < 0$, então uma das modas da distribuição BGEV ocorre dentro do intervalo $[x_\xi^{\min}, x_\xi^{\max}]$, em que x_ξ^{\min} e x_ξ^{\max} são dados na Subseção 2.2.

2.4 Momentos

Nesta seção serão apresentados os momentos para distribuição BGEV. Os momentos são medidas capazes de caracterizar uma distribuição de probabilidade. Sendo os quatro primeiros momentos indicativos de tendência central, dispersão, assimetria e curtose nesta ordem.

2.4.1 Momentos: caso $\xi \neq 0$

Para a obtenção de uma fórmula geral para os momentos, calcula-se a seguir o k -ésimo momento de ordem ($E(X^k) = \int x^k f(x) dx$) da distribuição BGEV.

Proposição 2.7. Seja $X \sim \text{BGEV}(\xi, \mu, \sigma, \xi, \delta)$ com $\xi \neq 0$ e $\delta > -1$. Usando a variável $W = |X|^{\delta+1}$. Temos que

$$\mathbb{E}(W^k) = \frac{\text{sign}(\mu)}{\sigma^k} \sum_{i=1}^k \binom{k}{i} \left(\mu - \frac{1}{\xi}\right)^{k-i} \frac{1}{\xi^i} \Gamma(1 - \xi i) \cdot [1_{\{1/\mu < \xi < 1/k, \mu > 0\}} + 1_{\{\xi < 1/\mu, \mu < 0\}}],$$

em que $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ é a função Gamma. Isso é, para $k \leq \frac{\delta+1}{\xi}$ os momentos $\mathbb{E}(|X|^k)$ existem.

Demonstração. Da equação (2.3) que define a função $T_\delta(x)$ é possível concluir que

$$w = |x|^{\delta+1} = \text{sign}(x) \frac{T_\delta(x)}{\sigma}. \quad (2.9)$$

Por outro lado, a Proposição 2.1 mostra que $Y = T_\delta(X) \sim F_{\text{GEV}^*}(\cdot; \mu, \xi)$. Portanto os momentos de Y são conhecidos e estão apresentados pela equação (1.9), atentando-se ao fato de que $\sigma = 1$ para GEV^* , como definido em (1.3).

Então usando a equação (2.9) segue que

$$\mathbb{E}(W^k) = \frac{1}{\sigma^k} [\mathbb{E}(Y^k | X > 0) \mathbb{P}(X > 0) - \mathbb{E}(Y^k | X < 0) \mathbb{P}(X < 0)]. \quad (2.10)$$

Se $\xi > 1/\mu$ e $\mu > 0$ tem-se que $\mathbb{P}(X > 0) = 1$. Consequentemente, pela combinação de (1.9) e (2.10), segue a prova do primeiro item.

Para $\xi < 1/\mu$ e $\mu < 0$ a prova é similar. □

Corolário 2.1. Seja $X \sim BGEV(\cdot, \mu, \sigma, \xi, \delta)$, $W = |X|^{\delta+1}$ com $\xi \neq 0$ e $\delta \geq -1$.

- 1) Se $1/\mu < \xi < 1$ então, $\mathbb{E}(W) = \mu + \frac{g_1 - 1}{\sigma \xi}$;
- 2) Se $1/\mu < \xi < 1/2$ então, $\mathbb{E}(W^2) = \frac{g_2 - g_1^2 + (\mu \xi + g_1 - 1)^2}{\sigma^2 \xi^2}$;

em que g_k representa a função $\Gamma(1 - k\xi)$.

Aqui, é possível encontrar uma fórmula analítica fechada para o $k(\delta + 1)$ -ésimo momento de uma variável aleatória com distribuição BGEV.

Proposição 2.8. Seja $X \sim BGEV(x, \mu, \sigma, \xi, \delta)$ com $\xi \neq 0$ e $\delta \geq -1$. Então para $\xi < \frac{1}{k}$

$$\mathbb{E}(X^{k(\delta+1)}) = \frac{1}{\xi^k \sigma^k} \sum_{i=0}^k \binom{k}{i} \Gamma(1 - \xi i), \quad \mu - \frac{1}{\sigma} > 0 \quad (2.11)$$

e

$$\begin{aligned} \mathbb{E}(X^{k(\delta+1)}) &= \frac{(-1)^{k(\delta+2)}}{\xi^k \sigma^k} \sum_{i=0}^k \binom{k}{i} \gamma\left(1 - \xi i; (1 - \xi \mu)^{-1/\xi}\right) \\ &+ \frac{1}{\xi^k \sigma^k} \sum_{i=0}^k \binom{k}{i} \Gamma\left(1 - \xi i; (1 - \xi \mu)^{-1/\xi}\right), \quad \mu - \frac{1}{\sigma} < 0, \end{aligned} \quad (2.12)$$

em que $\gamma(a; x) = \int_x^\infty t^{a-1} e^{-t} dt$ e $\Gamma(a; x) = \int_0^x t^{a-1} e^{-t} dt$ são as funções gamma incompletas, inferior e superior respectivamente.

Demonstração. Para $\xi > 0$, pela definição de esperança e pela expressão 1.2, tem-se que

$$\begin{aligned} \mathbb{E}(X^{k(\delta+1)}) &= \int_{T_{\sigma, \delta}^{-1}(\mu-1/\xi)}^\infty x^{k(\delta+1)} f_X(x; \xi, \mu, \sigma, \delta) dx \\ &= \int_{T_{\delta}^{-1}(\mu-1/\xi)}^\infty x^{k(\delta+1)} f_{GEV}(T_{\delta}(x); \xi, \mu) dT_{\delta}(x) \\ &= \int_{\mu-1/\xi}^\infty [T_{\delta}^{-1}(y)]^{k(\delta+1)} f_{GEV}(y; \xi, \mu) dy. \end{aligned} \quad (2.13)$$

Ao substituir a expressão de T^{-1} dada em (2.4), em (2.13) obtemos dois casos:

(i) Se $\mu - \frac{1}{\sigma} > 0$

$$\begin{aligned} \mathbb{E}(X^{k(\delta+1)}) &= \frac{1}{\sigma^k} \int_{\mu-1/\xi}^\infty \left(\frac{y}{\sigma}\right)^k f_{GEV}(y; \xi, \mu) dy \\ &= \frac{1}{\sigma^k} E(Y^k), \quad Y \sim F_{GEV}, \end{aligned} \quad (2.14)$$

como em (1.9).

(ii) Se $\mu - \frac{1}{\sigma} < 0$

$$\begin{aligned}\mathbb{E}(X^{k(\delta+1)}) &= \frac{(-1)^{k(\delta+2)}}{\sigma^k} \int_{\mu-1/\xi}^0 y^k f_{\text{GEV}}(y; \xi, \mu) dy + \frac{1}{\sigma^k} \int_0^\infty y^k f_{\text{GEV}}(y; \xi, \mu) dy \\ &= \frac{(-1)^{k(\delta+2)}}{\sigma^k} I_1 + \frac{1}{\sigma^k} I_2.\end{aligned}\quad (2.15)$$

Usando a função de densidade de probabilidade dada na equação 1.3 e a substituição $w^{-1} = 1 + \xi(y - \mu)$, obtem-se

$$I_1 = \frac{1}{\xi^{k+1}} \int_{(1-\xi\mu)^{-1}}^\infty [w^{-1} + (\xi\mu - 1)]^{\frac{k}{1+\delta}} w^{\frac{1}{\xi}-1} e^{-w^{1/\xi}} dw. \quad (2.16)$$

Para resolver a integral de (2.16) utiliza-se o binômio de Newton, portanto

$$\begin{aligned}I_2 &= \frac{1}{\xi^{k+1}} \sum_{i=1}^k \binom{k}{i} (\xi\mu - 1)^{k-i} \int_{(1-\xi\mu)^{-1}}^\infty w^{i+\frac{1}{\xi}-\frac{k}{1+\delta}-1} e^{-w^{1/\xi}} dw \\ &= \frac{1}{\xi^k} \sum_{i=1}^k \binom{k}{i} (\xi\mu - 1)^{k-i} \int_{(1-\xi\mu)^{-1/\xi}}^\infty z^{-\xi i} e^{-z} dz,\end{aligned}\quad (2.17)$$

usando a nova substituição $z = w^{1/\xi}$.

Similarmente obtem-se

$$I_2 = \frac{1}{\xi^k} \sum_{i=1}^k \binom{k}{i} (\xi\mu - 1)^{k-i} \int_0^{(1-\xi\mu)^{-1/\xi}} z^{-\xi i} e^{-z} dz. \quad (2.18)$$

A prova é completada expressando como integrais de (2.17) e (2.18) em termos de funções gama incompletas e, em seguida, atualizando a equação (2.15).

Para o caso de $\xi < 0$ a prova é similar. □

Corolário 2.2. Seja $X \sim BGEV(., \mu, \sigma, \xi, \delta)$ com $\xi \neq 0$ e $\delta = 0$, então

$$\mathbb{E}(X^k) = \frac{1}{\xi^k \sigma^k} \sum_{i=0}^k \binom{k}{i} \Gamma(1 - \xi i). \quad (2.19)$$

Demonstração. Para $\mu - \frac{1}{\sigma} > 0$, o resultado segue diretamente quando considera-se $\delta = 0$ em (2.20).

Para $\mu - \frac{1}{\sigma} < 0$, a propriedade

$$\gamma(a; x) + \Gamma(a; x) = \Gamma(a)$$

é usada na equação (2.21). □

Corolário 2.3. Seja $X \sim BGEV(\xi, \mu, \sigma, \delta)$ com $\xi \neq 0$ e $\delta \leq 0$, então

$$\mathbb{E}(X) = \frac{1}{(\xi\sigma)^{1/1+\delta}} \sum_{i=0}^{1/1+\delta} \binom{1/1+\delta}{i} \Gamma(1 - \xi i), \quad \mu - \frac{1}{\sigma} > 0 \quad (2.20)$$

e

$$\begin{aligned} \mathbb{E}(X) &= \frac{(-1)^{k(\delta+2)}}{(\xi\sigma)^{1/1+\delta}} \sum_{i=0}^{1/1+\delta} \binom{1/1+\delta}{i} \gamma\left(1 - \xi i; (1 - \xi\mu)^{-1/\xi}\right) \\ &+ \frac{1}{(\xi\sigma)^{1/1+\delta}} \sum_{i=0}^{1/1+\delta} \binom{1/1+\delta}{i} \Gamma\left(1 - \xi i; (1 - \xi\mu)^{-1/\xi}\right), \quad \mu - \frac{1}{\sigma} < 0. \end{aligned} \quad (2.21)$$

Demonstração. O resultado segue diretamente da Proposição 3.11. por considerar $\delta = \frac{1}{k} - 1$ and $k \in \mathcal{Z}^+$. □

2.4.2 Momentos: caso $\xi = 0$

Proposição 2.9. Se $X \sim F_{BGEV}(\cdot; \mu, \sigma, \xi = 0, \delta)$ com $\delta \in \mathbb{R}$, então

$$\mathbb{E}(X^k) = \frac{1}{Z_\delta} [\delta^2 \mathbb{E}(Y^{k+2}) - 2\delta \mathbb{E}(Y^{k+1}) + 2 \mathbb{E}(Y^k)], \quad k \geq 1, \quad (2.22)$$

em que $Y \sim F_{GEV}(\cdot; \mu, \sigma, \xi = 0)$ como apresentado em (1.1) e Z_δ é expressa pela equação (2.2).

Demonstração. A partir da definição de esperança e usando a função de densidade de probabilidade da distribuição BGEV com $\xi = 0$, dada em (1.2), tem-se

$$\begin{aligned} \int_{-\infty}^{\infty} x^k f(x; \xi, \mu, \sigma, \delta) dx &= \frac{1}{Z_\delta} \mathbb{E}[Y^k(1 - \delta Y)^2 + Y^k] \\ &= \frac{1}{Z_\delta} [\delta^2 \mathbb{E}(Y^{k+2}) - 2\delta \mathbb{E}(Y^{k+1}) + 2 \mathbb{E}(Y^k)]. \end{aligned}$$

Portanto, a prova está completa. □

Corolário 2.4. Se $X \sim BGEV(\cdot; \mu, \sigma, \xi = 0, \delta)$ com $\delta \in \mathbb{R}$, então

$$\begin{aligned} 1) \mathbb{E}(X) &= \frac{1}{Z_\delta} \left\{ \delta^2 \left[\sigma^3 \zeta(3) + 3(\mu + \sigma\gamma) \sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^3 \right] \right. \\ &\quad \left. - 2\delta \left[\sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^2 \right] + 2(\mu + \sigma\gamma) \right\}, \\ 2) \mathbb{E}(X^2) &= \frac{1}{Z_\delta} \left(\delta^2 \left\{ \sigma^4 \frac{\pi^4}{15} + 4(\mu + \sigma\gamma) \left[2\sigma^3 \zeta(3) + 3(\mu + \sigma\gamma) \sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^3 \right] \right. \right. \\ &\quad \left. \left. - \left[\sigma^2 \frac{\pi^2}{2} + 3(\mu + \sigma\gamma)^2 \right] \left[\sigma^2 \frac{\pi^2}{3} + (\mu + \sigma\gamma)^2 \right] \right\} \right. \\ &\quad \left. - 2\delta \left[2\sigma^3 \zeta(3) + 3(\mu + \sigma\gamma) \sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^3 \right] + 2 \left[\sigma^2 \frac{\pi^2}{6} + (\mu + \sigma\gamma)^2 \right] \right), \end{aligned}$$

em que Z_δ está definido em (2.2), γ é a constante de Euler e $\zeta(\cdot)$ a função de Riemann.

Demonstração. Os momentos $Y \sim F_{GEV}(\cdot; \mu, \sigma, \xi = 0)$ estão apresentados em (1.8), o pri-

meiro, segundo, terceiro e quarto momentos. Por outro lado, pela Proposição 2.9 tem-se:

$$\begin{aligned}\mathbb{E}(X) &= \frac{1}{Z_\delta} [\delta^2 \mathbb{E}(Y^3) - 2\delta \mathbb{E}(Y^2) + 2 \mathbb{E}(Y)] ; \\ \mathbb{E}(X^2) &= \frac{1}{Z_\delta} [\delta^2 \mathbb{E}(Y^4) - 2\delta \mathbb{E}(Y^3) + 2 \mathbb{E}(Y^2)] .\end{aligned}$$

Finalmente, combinando as identidades acima, a prova segue. □

2.5 Quantis

A função quantil é obtida a partir da inversa da função de distribuição acumulada e retorna o valor associado à uma dada probabilidade acumulada e também é usada para simular dados que seguem a distribuição. Sendo assim, a seguir serão dados os passos para obtenção da função quantil da distribuição BGEV, onde por questão de notação $F(x)$ passa a ser denotado por Q .

Para o caso em que $\xi \neq 0$ tem-se que a função de distribuição acumulada de uma variável BGEV é dada por $Q = F(x) = F_{\text{GEV}^*}(T_\delta(x); \mu, \xi)$, como em (2.6), então a função quantil é definida por:

$$\begin{aligned}x_Q &= T_\delta^{-1}(F_{\text{GEV}^*}^{-1}(Q)) \\ T_\delta(x_Q) &= F_{\text{GEV}^*}^{-1}(Q) \\ F_{\text{GEV}^*}(T_\delta(x_Q)) &= Q.\end{aligned}$$

Sabendo-se que T_δ^{-1} é definida como em (2.4) e que a função de distribuição acumulada inversa do modelo GEV^* , quando $\xi \neq 0$, é dada por $F_{\text{GEV}^*}^{-1}(Q) = \mu + ((-\log(Q))^{-\xi} - 1)$. Tem-se que a função quantil para a distribuição BGEV com $\xi \neq 0$ é:

$$x_Q = \begin{cases} \left(\frac{\mu}{\sigma} + \frac{(-\log(Q))^{-\xi} - 1}{\sigma\xi} \right)^{\frac{1}{1+\delta}} & , \mu + \frac{1}{\xi} [(-\log(Q))^{-\xi} - 1] > 0, \\ - \left(-\frac{\mu}{\sigma} - \frac{(-\log(Q))^{-\xi} - 1}{\sigma\xi} \right)^{\frac{1}{1+\delta}} & , \mu + \frac{1}{\xi} [(-\log(Q))^{-\xi} - 1] < 0. \end{cases} \quad (2.23)$$

2.5.1 Mediana

Em particular, para $\xi \neq 0$, ao substituir o valor de Q por $0,5$ na função quantil obtém-se a mediana da distribuição BGEV que é expressa por:

$$T_{\delta}^{-1}\left(\mu + \frac{1}{\xi}[(\log 2)^{-\xi} - 1]\right).$$

Assim temos uma forma fechada para mediana da distribuição, que representa o valor que divide os dados ordenados pela metade.

Para o caso em que $\xi = 0$ devido a complexidade da função de distribuição acumulada da BGEV, não é trivial o cálculo do quantil, e portanto não será apresentado aqui.

Capítulo 3

Estimação

A metodologia que será utilizada para a estimação dos parâmetros da distribuição BGEV é de máxima verossimilhança.

3.1 Estimação por máxima verossimilhança

O método utilizado para a estimação dos parâmetros procura, baseado nos resultados obtidos por amostra, o parâmetro Θ (ou o conjunto de parâmetros, como é o caso) que maximiza a função de verossimilhança (L) de uma distribuição de probabilidades. Tal função é definida da seguinte forma, seja $\mathbf{x} = x_1, x_2, \dots, x_n$ os valores de uma amostra de uma variável aleatória $X \sim F$, com $F' = f$, então a função de verossimilhança da amostra é dada por:

$$L(\Theta; \mathbf{x}) = \prod_{i=1}^n f(x_i|\Theta).$$

Usualmente, utiliza-se a função conhecida como log-verossimilhança, que nada mais é que a aplicação da função logaritmo na função de máxima verossimilhança.

$$\mathcal{L}(\Theta; \mathbf{x}) = \log(L(\Theta; x_1, x_2, \dots, x_n)) = \sum_{i=1}^n \log(f(x_i|\Theta)).$$

Para a distribuico em estudo BGEV com conjunto de parmetros $\Theta = (\mu, \sigma, \xi, \delta)$, as funoes de mxima log-verossimilhana esto descritas em (3.1) e (3.2), para os casos de $\xi \neq 0$ e $\xi = 0$, respectivamente.

$$\mathcal{L}_1(\Theta; \mathbf{x}) = n \log(\sigma) + n \log(\delta + 1) + \sum_{i=1}^n \left\{ \delta \log(|x_i|) - \left(1 + \frac{1}{\xi}\right) \log[\Psi_i(\Theta)] - \Psi_i^{-1/\xi}(\Theta) \right\} \quad (3.1)$$

em que $\Psi_i(\Theta) = 1 + \xi(\sigma x_i |x_i|^\delta - \mu)$. Note ainda que, para todo $i = 1, \dots, n$,

$$\begin{aligned} \frac{\partial \Psi_i(\Theta)}{\partial \mu} &= -\xi \mu; & \frac{\partial \Psi_i(\Theta)}{\partial \sigma} &= \xi x_i |x_i|^\delta; \\ \frac{\partial \Psi_i(\Theta)}{\partial \delta} &= \xi \sigma x_i |x_i|^\delta \ln |x_i|; & \frac{\partial \Psi_i(\Theta)}{\partial \xi} &= \sigma x_i |x_i|^\delta - \mu. \end{aligned}$$

$$\mathcal{L}_0(\Theta; \mathbf{x}) = -n \log(Z_\delta) - n \log(\sigma) + \sum_{i=1}^n \log[(1 - \delta x_i)^2 + 1] - \sum_{i=1}^n \left[\left(\frac{x_i - \mu}{\sigma}\right) + e^{-\left(\frac{x_i - \mu}{\sigma}\right)} \right]. \quad (3.2)$$

Os estimadores de mxima verossimilhana so obtidos pela soluo do sistema de equacoes dado pelas derivadas da funo log-verossimilhana em relao a cada um dos parmetros a serem estimados. Ento, as estimativas de mxima verossimilhana de μ, σ, ξ, δ so as soluoes dos seguintes sistemas de equacoes:

$$\begin{cases} \frac{\partial \mathcal{L}_1(\Theta; x)}{\partial \mu} = \mu \sum_{i=1}^n \Omega_i(\Theta) = 0 \\ \frac{\partial \mathcal{L}_1(\Theta; x)}{\partial \sigma} = \frac{n}{\sigma} - \sum_{i=1}^n x_i |x_i|^\delta \Omega_i(\Theta) = 0 \\ \frac{\partial \mathcal{L}_1(\Theta; x)}{\partial \xi} = \xi^{-2} \sum_{i=1}^n (\sigma x_i |x_i|^\delta - \mu) \{ \ln[\Psi_i(\Theta)] - \xi \Omega_i(\Theta) \} = 0 \\ \frac{\partial \mathcal{L}_1(\Theta; x)}{\partial \delta} = \frac{n}{\delta + 1} + \sum_{i=1}^n [\ln |x_i| + \sigma x_i |x_i|^\delta \ln |x_i| \Omega_i(\Theta)] = 0, \end{cases} \quad (3.3)$$

em que $\Omega_i(\Theta) = \Psi_i^{-1}(\Theta) \left[1 + \xi - \Psi_i^{-1/\xi}(\Theta) \right]$, $i = 1, \dots, n$. Além disso,

$$\begin{cases} \frac{\partial \mathcal{L}_0(\Theta; x)}{\partial \mu} = -\frac{n}{Z_\delta} \frac{\partial Z_\delta}{\partial \mu} + \frac{n}{\sigma} - \frac{1}{\sigma} \sum_{i=1}^n e^{-(\frac{x_i - \mu}{\sigma})} = 0 \\ \frac{\partial \mathcal{L}_0(\Theta; x)}{\partial \sigma} = -\frac{n}{Z_\delta} \frac{\partial Z_\delta}{\partial \sigma} - \frac{n}{\sigma} - \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \left[1 - e^{-(\frac{x_i - \mu}{\sigma})} \right] = 0 \\ \frac{\partial \mathcal{L}_0(\Theta; x)}{\partial \delta} = -\frac{n}{Z_\delta} \frac{\partial Z_\delta}{\partial \delta} - \sum_{i=1}^n 2x_i(1 - \delta x_i) \left[(1 - \delta x_i)^2 + 1 \right] = 0. \end{cases} \quad (3.4)$$

Para encontrar os estimadores de máxima verossimilhança dos parâmetros da BGEV é utilizado um método de otimização de funções através do software estatístico R, desenvolvido por Nelder e Mead, no qual os sistemas de equações (3.3) e (3.4) são úteis para minimizar (3.1) e (3.2), respectivamente.

Para estimação é necessário inserir valores iniciais dos parâmetros para começar o algoritmo, a escolha desses valores na simulação se deu de forma que, como os valores reais dos parâmetros eram conhecidos, então os "chutes" iniciais, como são chamados, fora, definidos pelo proprio valor do parâmetro somado à uma variável aleatória uniforme com parâmetros 0 e 1.

3.2 Estatísticas de qualidade de ajuste

Para avaliar a adequação relativa de modelos concorrentes, é preciso confiar em certas estatísticas de qualidade do ajuste. Isso pode incluir a função de log-verossimilhança avaliada nos estimadores de máxima verossimilhança (MLE) indicados por $\hat{\mathcal{L}}$, critério de informação de Akaike (AIC), o critério de informação de Akaike corrigido (AICC) e Kolmogrov – Smirnov. Quanto menores essas estatísticas, melhor o ajuste.

As estatísticas AIC e AICC são dadas respectivamente por:

$$AIC = -2\mathcal{L}(\hat{\theta}) + 2p \quad e \quad AICC = AIC + \frac{2p(p+1)}{n-p-1}$$

, em que $\mathcal{L}(\hat{\theta})$ denota a funo de log-verossimilhana avaliada no MLE, p é o nmero de parâmetros estimados e n , o tamanho da amostra.

Quanto à estatística Kolmogrov–Smirnov, é definido por:

$$K - S = \max \left\{ \frac{i}{n} - z_i; z_i - \frac{i-1}{n} \right\}$$

, em que $z_i = F(y_i)$, sendo y_i a i -ésima estatística de ordem.

Capítulo 4

Simulação

Foram realizadas simulações de amostras aleatórias de uma variável com distribuição BGEV, através do método da transformada inversa. O método utiliza a função quantil da distribuição a ser simulada (equação (2.23)) tendo como domínio uma variável uniforme no intervalo de 0 a 1 ($U \sim U(0, 1)$) para gerar valores aleatórios, de modo que a expressão $X = F^{-1}(U)$ seja satisfeita.

Algumas combinações de parâmetros foram definidas conforme apresentado na Tabela 4.1, e para cada uma foram geradas $M = 100$ amostras de tamanhos $n = 50, n = 100, n = 250$ e $n = 1000$, sendo que os resultados apresentados neste capítulo se referem às simulações de $n = 1000$ e os demais resultados estão descritos no Anexo B. Tendo em vista que o parâmetro σ é um parâmetro de escala da distribuição, o mesmo foi mantido constante e igual a 1 em todas as combinações. Os demais parâmetros variaram da seguinte maneira: o parâmetro ξ variou entre os valores $-0.5, -0.25, 0.25, 0.5, 1$, os valores definidos para μ foram $-1, 0, 1$ e δ variou entre $0, 2, 4$. Com esses valores foi possível criar 45 combinações de parâmetros para as simulações da distribuição BGEV.

Para cada uma das amostras geradas os parâmetros ξ, μ e δ foram estimados pelo método da máxima verossimilhança (utilizando a função `optim` do software R). Sendo assim para cada combinação de parâmetros obteve-se 100 estimadores resultantes das simulações geradas. O

estimador de Monte Carlo para essas 100 replicações é representado pelo valor esperado dos estimadores encontrados, portanto, $\bar{\theta} = \sum_{i=1}^M \frac{\hat{\theta}_i}{M}$. Os resultados dos estimadores e seus reais valores estão apresentados na Tabela 4.1:

Tabela 4.1: Estimativas via Monte Carlo.

ξ	$\bar{\xi}$	μ	$\bar{\mu}$	δ	$\bar{\delta}$
1	1,005	1	0,996	0	0,012
1	1,011	1	0,995	2	2,024
1	1,002	1	1,001	4	3,998
1	1,01	0	0	0	0,005
1	1,011	0	-0,003	2	2,001
1	1,008	0	0,001	4	4,039
1	1,001	-1	-1,006	0	0,006
1	0,998	-1	-1,002	2	2,019
1	1,001	-1	-1,001	4	4,012
0,5	0,493	1	1,003	0	-0,003
0,5	0,512	1	0,992	2	2,041
0,5	0,506	1	0,999	4	4,023
0,5	0,507	0	-0,005	0	-0,001
0,5	0,501	0	0,003	2	2,006
0,5	0,505	0	-0,001	4	3,998
0,5	0,501	-1	-0,998	0	0
0,5	0,505	-1	-1,001	2	2,004
0,5	0,499	-1	-0,997	4	4,007
0,25	0,253	1	1,002	0	0,002
0,25	0,25	1	0,996	2	2

Continua na próxima página.

Tabela 4.1 – continuação

ξ	$\bar{\xi}$	μ	$\bar{\mu}$	δ	$\bar{\delta}$
0,25	0,256	1	0,996	4	4,018
0,25	0,254	0	-0,001	0	0,002
0,25	0,249	0	0,004	2	2,013
0,25	0,252	0	-0,002	4	4,012
0,25	0,248	-1	-1,001	0	0,002
0,25	0,25	-1	-0,998	2	2,004
0,25	0,253	-1	-1,002	4	4,015
-0,25	-0,257	1	1,003	0	-0,002
-0,25	-0,248	1	1,003	2	2,01
-0,25	-0,253	1	1	4	4,003
-0,25	-0,251	0	0	0	0,002
-0,25	-0,252	0	0	2	2,004
-0,25	-0,251	0	0,001	4	4,003
-0,25	-0,25	-1	-1,003	0	0,003
-0,25	-0,253	-1	-1	2	1,998
-0,25	-0,254	-1	-0,997	4	4,021
-0,5	-0,505	1	1,006	0	0
-0,5	-0,502	1	1,003	2	2,002
-0,5	-0,505	1	1	4	4
-0,5	-0,501	0	0,009	0	0,017
-0,5	-0,505	0	0,005	2	2,01
-0,5	-0,505	0	0,004	4	4,005
-0,5	-0,505	-1	-0,985	0	0,023
-0,5	-0,503	-1	-1	2	2,02

Continua na próxima página.

Tabela 4.1 – continuação

ξ	$\bar{\xi}$	μ	$\bar{\mu}$	δ	$\bar{\delta}$
-0,5	-0,505	-1	-0,996	4	4,006

Pode-se notar que as estimativas geradas estão sempre muito próximas aos reais valores dos parâmetros, em todas as combinações de parâmetros geradas. Para avaliação dessa discrepância entre a medida estatística calculada (estimador) e o valor real do parâmetro observamos os vieses gerados pelas estimativas ($\bar{\theta} - \theta$) descritos na Tabela 4.2.

Tabela 4.2: Vieses de Monte Carlo dos estimadores da Tabela 4.1.

ξ	viés ($\bar{\xi}$)	μ	viés ($\bar{\mu}$)	δ	viés ($\bar{\delta}$)
1	-0,005	1	0,004	0	-0,012
1	-0,011	1	0,005	2	-0,024
1	-0,002	1	-0,001	4	0,002
1	-0,010	0	0,000	0	-0,005
1	-0,011	0	0,003	2	-0,001
1	-0,008	0	-0,001	4	-0,039
1	-0,001	-1	0,006	0	-0,006
1	0,002	-1	0,002	2	-0,019
1	-0,001	-1	0,001	4	-0,012
0,5	0,007	1	-0,003	0	0,003
0,5	-0,012	1	0,008	2	-0,041
0,5	-0,006	1	0,001	4	-0,023
0,5	-0,007	0	0,005	0	0,001

Continua na próxima página.

Tabela 4.2 – continuação

ξ	viés ($\bar{\xi}$)	μ	viés ($\bar{\mu}$)	δ	viés ($\bar{\delta}$)
0,5	-0,001	0	-0,003	2	-0,006
0,5	-0,005	0	0,001	4	0,002
0,5	-0,001	-1	-0,002	0	0,000
0,5	-0,005	-1	0,001	2	-0,004
0,5	0,001	-1	-0,003	4	-0,007
0,25	-0,003	1	-0,002	0	-0,002
0,25	0,000	1	0,004	2	0,000
0,25	-0,006	1	0,004	4	-0,018
0,25	-0,004	0	0,001	0	-0,002
0,25	0,001	0	-0,004	2	-0,013
0,25	-0,002	0	0,002	4	-0,012
0,25	0,002	-1	0,001	0	-0,002
0,25	0,000	-1	-0,002	2	-0,004
0,25	-0,003	-1	0,002	4	-0,015
-0,25	0,007	1	-0,003	0	0,002
-0,25	-0,002	1	-0,003	2	-0,010
-0,25	0,003	1	0,000	4	-0,003
-0,25	0,001	0	0,000	0	-0,002
-0,25	0,002	0	0,000	2	-0,004
-0,25	0,001	0	-0,001	4	-0,003
-0,25	0,000	-1	0,003	0	-0,003
-0,25	0,003	-1	0,000	2	0,002
-0,25	0,004	-1	-0,003	4	-0,021
-0,5	0,005	1	-0,006	0	0,000

Continua na próxima página.

Tabela 4.2 – continuação

ξ	viés ($\bar{\xi}$)	μ	viés ($\bar{\mu}$)	δ	viés ($\bar{\delta}$)
-0,5	0,002	1	-0,003	2	-0,002
-0,5	0,005	1	0,000	4	0,000
-0,5	0,001	0	-0,009	0	-0,017
-0,5	0,005	0	-0,005	2	-0,010
-0,5	0,005	0	-0,004	4	-0,005
-0,5	0,005	-1	-0,015	0	-0,023
-0,5	0,003	-1	0,000	2	-0,020
-0,5	0,005	-1	-0,004	4	-0,006

Outra medida que indica a qualidade do estimador para M amostras é o erro quadrático médio, definido pela expressão $\left(EQM = \sum_{i=1}^M \frac{(\hat{\theta}_i - \bar{\theta})^2}{M} \right)$. Os resultados do viés e EQM das estimativas de Monte Carlo dos parâmetros são mostrados nas Tabelas 4.2 e 6.3. Ambos resultados indicam que os estimadores de máxima verossimilhança dos parâmetros da distribuição BGEV são eficientes.

Tabela 4.3: Erro quadrático médio dos estimadores da Tabela 4.1.

ξ	EQM ($\hat{\xi}$)	μ	EQM ($\hat{\mu}$)	δ	EQM ($\hat{\delta}$)
1	0,0028	1	0,0014	0	0,0025
1	0,0025	1	0,0014	2	0,0213
1	0,0029	1	0,0014	4	0,0578
1	0,002	0	4e-04	0	0,0010
1	0,0019	0	4e-04	2	0,0064

Continua na próxima página.

Tabela 4.3 – continuação

ξ	EQM ($\hat{\xi}$)	μ	EQM ($\hat{\mu}$)	δ	EQM ($\hat{\delta}$)
1	0,0015	0	4e-04	4	0,0216
1	0,0019	-1	9e-04	0	6e-040
1	0,0011	-1	9e-04	2	0,0059
1	0,0014	-1	6e-04	4	0,0140
0,5	0,0026	1	0,0015	0	0,0020
0,5	0,0023	1	0,0013	2	0,0167
0,5	0,0025	1	0,0011	4	0,0497
0,5	8e-04	0	6e-04	0	9e-04
0,5	0,0013	0	7e-04	2	0,0061
0,5	9e-04	0	7e-04	4	0,0135
0,5	0,0011	-1	0,0012	0	6e-04
0,5	9e-04	-1	9e-04	2	0,0045
0,5	0,001	-1	0,0013	4	0,0149
0,25	0,0015	1	0,001	0	0,0010
0,25	0,002	1	9e-04	2	0,0132
0,25	0,0018	1	0,0013	4	0,0331
0,25	9e-04	0	8e-04	0	6e-04
0,25	8e-04	0	8e-04	2	0,0066
0,25	8e-04	0	6e-04	4	0,0199
0,25	6e-04	-1	0,0014	0	5e-04
0,25	7e-04	-1	0,0011	2	0,0037
0,25	7e-04	-1	0,001	4	0,0106
-0,25	9e-04	1	0,0013	0	9e-04
-0,25	8e-04	1	0,0013	2	0,0061

Continua na próxima página.

Tabela 4.3 – continuação

ξ	EQM ($\hat{\xi}$)	μ	EQM ($\hat{\mu}$)	δ	EQM ($\hat{\delta}$)
-0,25	0,001	1	0,0012	4	0,0228
-0,25	4e-04	0	0,0013	0	7e-04
-0,25	5e-04	0	0,0012	2	0,0087
-0,25	4e-04	0	0,0011	4	0,0176
-0,25	2e-04	-1	0,003	0	0,0013
-0,25	3e-04	-1	0,0016	2	0,0033
-0,25	2e-04	-1	0,0012	4	0,0088
-0,5	6e-04	1	0,0012	0	8e-04
-0,5	5e-04	1	0,0014	2	0,0063
-0,5	7e-04	1	0,0012	4	0,0226
-0,5	0,001	0	0,0048	0	0,0114
-0,5	3e-04	0	0,0013	2	0,0065
-0,5	4e-04	0	0,0011	4	0,0223
-0,5	4e-04	-1	0,0049	0	0,0152
-0,5	2e-04	-1	0,0026	2	0,0204
-0,5	1e-04	-1	7e-04	4	0,0114

A análise gráfica é uma ótima ferramenta para observar como os parâmetros modificam a distribuição em suas variações. Os gráficos apresentados contêm a curva da distribuição com os parâmetros em seus valores reais, desenhada na cor azul, e a curva da distribuição com os parâmetros estimados, em vermelho. Dado que as estimativas obtidas foram todas muito próximas dos valores dos parâmetros, as curvas se tornam bastante semelhantes e se sobrepõem, dificultando a visualização. Sendo assim, além das curvas, foram acrescentados aos gráficos um

histograma de uma amostra aleatória de tamanho $n = 1000$ de uma distribuição BGEV com os valores de parâmetros reais. Com isso, pode-se confirmar que as estimativas de máxima verossimilhança dos parâmetros de uma BGEV fornecem bons resultados.

É fácil observar, por exemplo, que em todos os casos de $\delta = 0$ obtém-se uma distribuição unimodal, pois para esses casos a distribuição retorna a uma distribuição GEV de origem. A bimodalidade é sempre centrada no ponto zero, desta forma uma das modas será negativa e a outra será positiva. A medida que o valor de δ aumenta nota-se que a distribuição bimodal se torna mais concentrada em torno de suas modas e os valores próximos à zero ficam menos frequentes. O parâmetro μ influencia dando intensidade à moda, sendo que maiores valores de μ deixam a moda positiva mais constante e menores valores de μ dão maior peso para a moda negativa. De acordo com o valor de μ , pode acontecer da distribuição perder sua bimodalidade. É possível notar esse fato, no gráfico que ilustra a distribuição com parâmetros $\xi = 1, \mu = 1, \sigma = 1$ e δ assumindo os valores 2 ou 4, que mostra unimodalidade na distribuição.

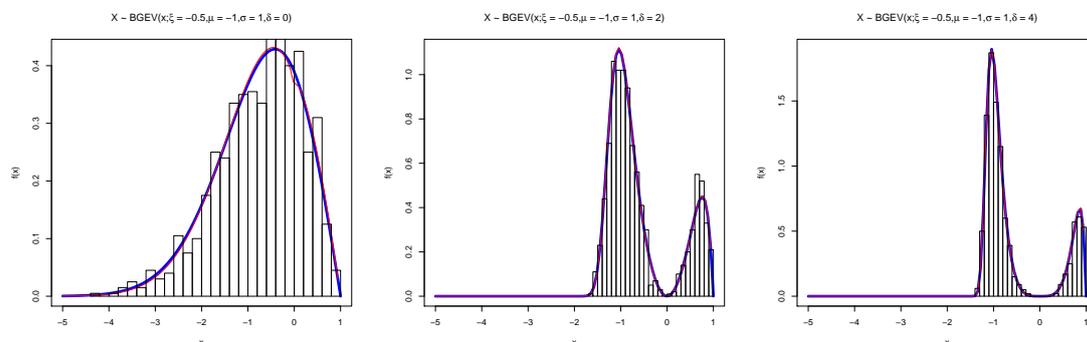


Figura 4.1: $BGEV(x; \mu = -1, \sigma = 1, \xi = -0.5, \delta)$, $\delta = 0, 2, 4$

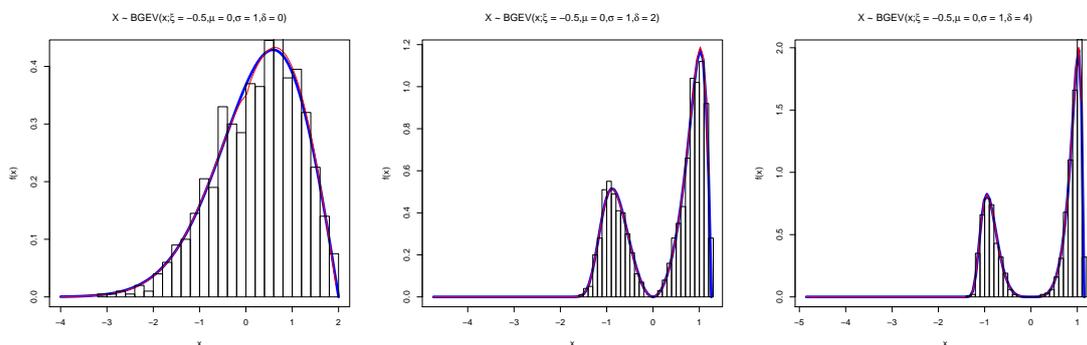


Figura 4.2: $BGEV(x; \mu = 0, \sigma = 1, \xi = -0.5, \delta), \delta = 0, 2, 4$

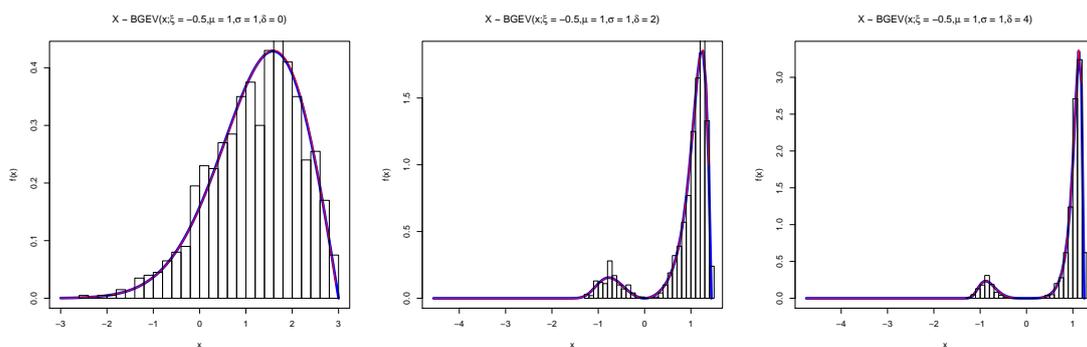


Figura 4.3: $BGEV(x; \mu = 1, \sigma = 1, \xi = -0.5, \delta), \delta = 0, 2, 4$

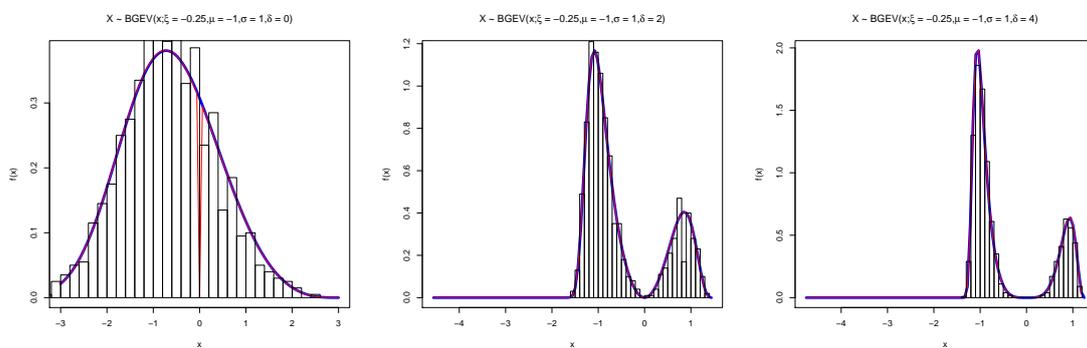


Figura 4.4: $BGEV(x; \mu = -1, \sigma = 1, \xi = -0.25, \delta), \delta = 0, 2, 4$

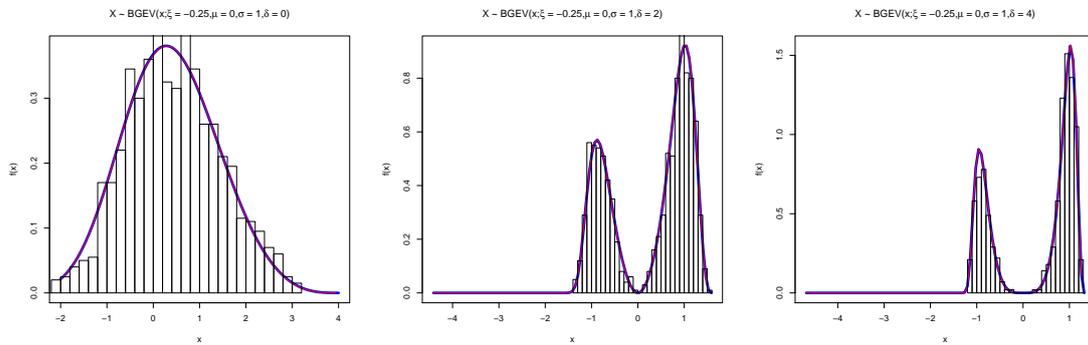


Figura 4.5: $BGEV(x; \mu = 0, \sigma = 1, \xi = -0.25, \delta), \delta = 0, 2, 4$

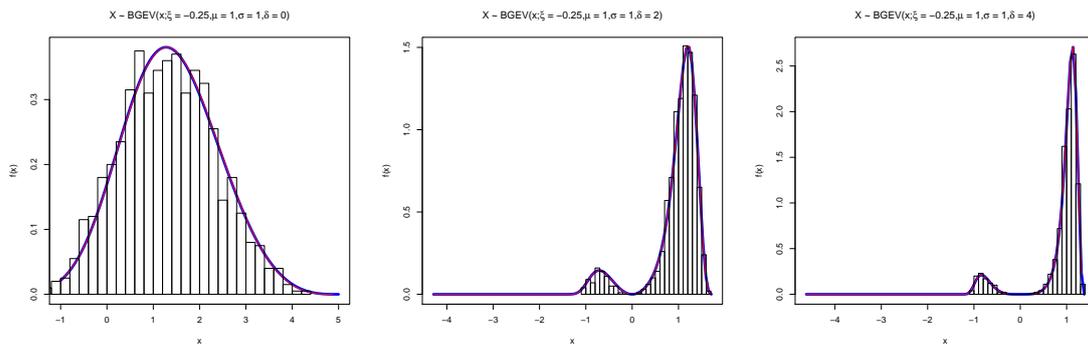


Figura 4.6: $BGEV(x; \mu = 1, \sigma = 1, \xi = -0.25, \delta), \delta = 0, 2, 4$

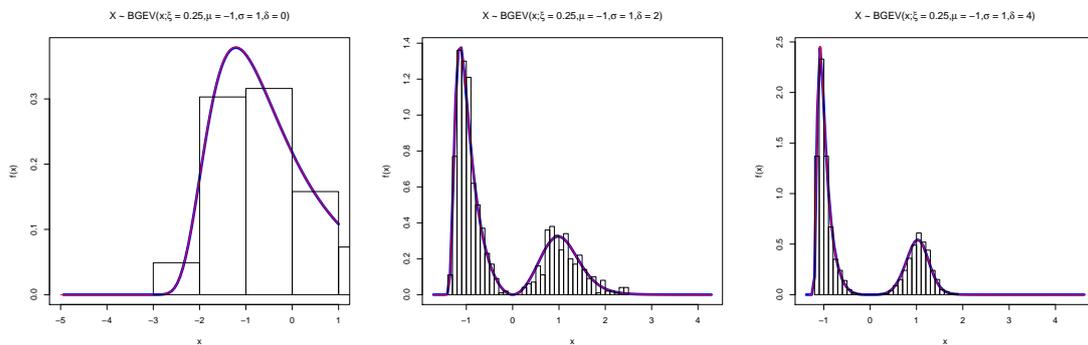


Figura 4.7: $BGEV(x; \mu = -1, \sigma = 1, \xi = 0.25, \delta), \delta = 0, 2, 4$

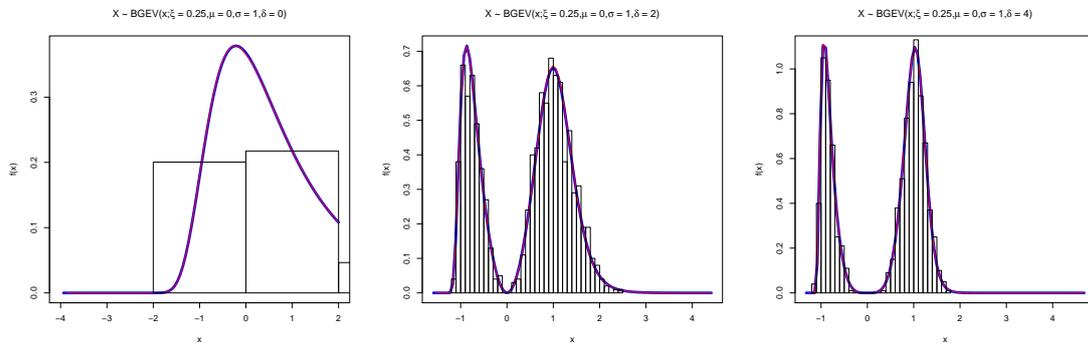


Figura 4.8: $BGEV(x; \mu = 0, \sigma = 1, \xi = 0.25, \delta), \delta = 0, 2, 4$

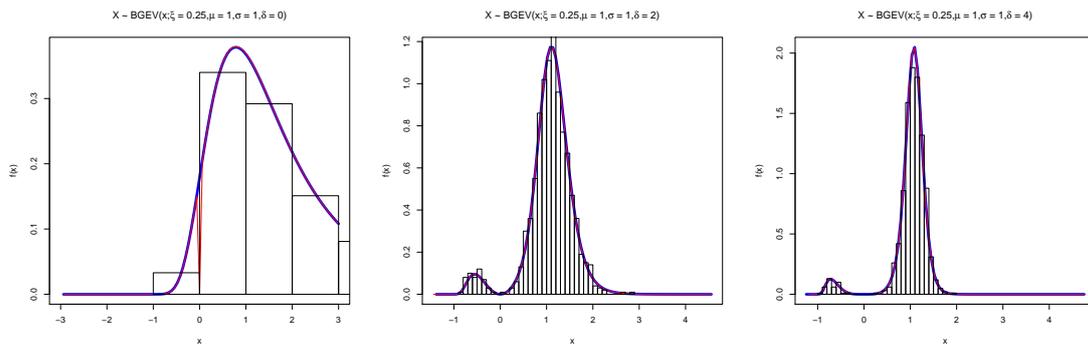


Figura 4.9: $BGEV(x; \mu = 1, \sigma = 1, \xi = 0.25, \delta), \delta = 0, 2, 4$

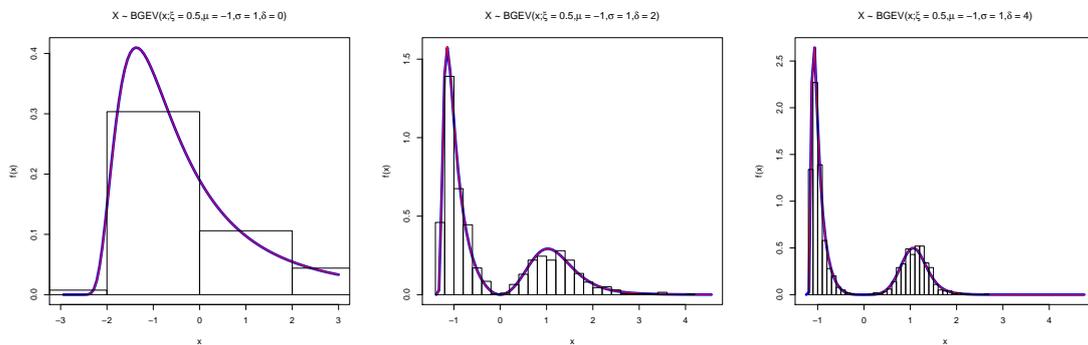


Figura 4.10: $BGEV(x; \mu = -1, \sigma = 1, \xi = 0.5, \delta), \delta = 0, 2, 4$

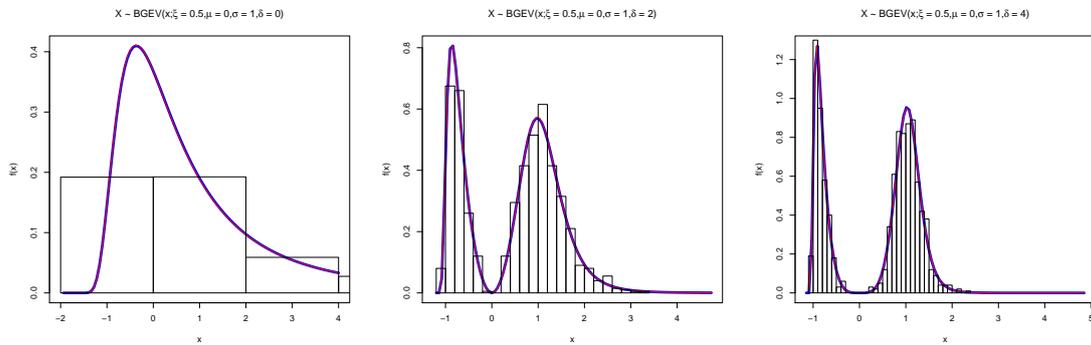


Figura 4.11: $BGEV(x; \mu = 0, \sigma = 1, \xi = 0.5, \delta)$, $\delta = 0, 2, 4$

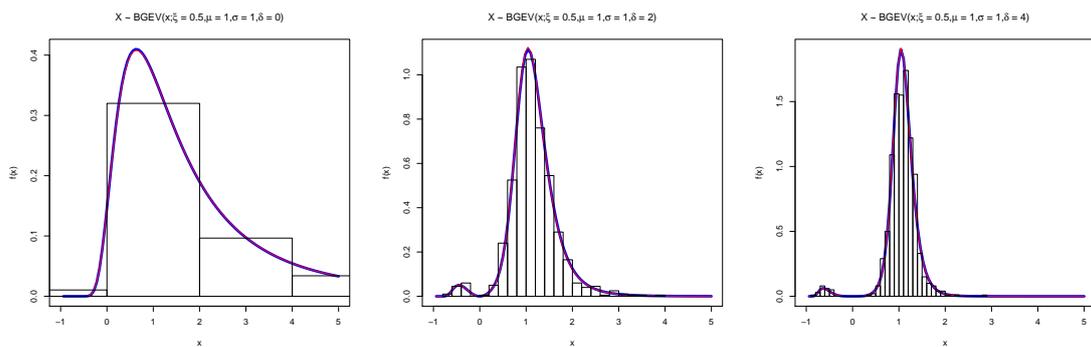


Figura 4.12: $BGEV(x; \mu = 1, \sigma = 1, \xi = 0.5, \delta)$, $\delta = 0, 2, 4$

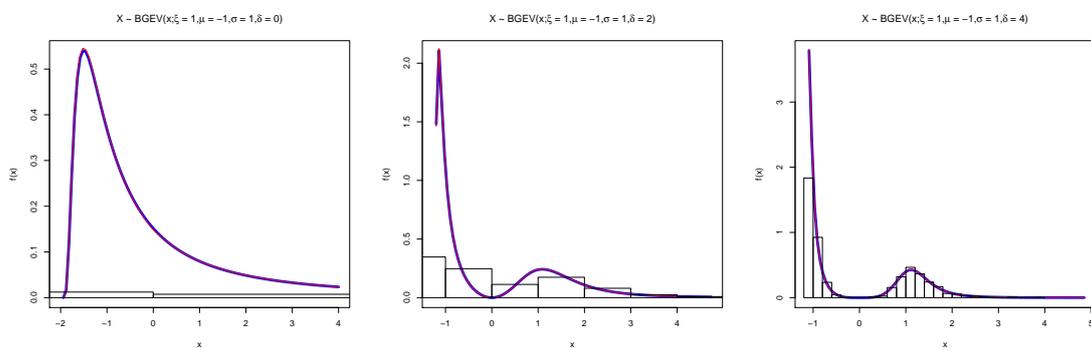


Figura 4.13: $BGEV(x; \mu = -1, \sigma = 1, \xi = 1, \delta)$, $\delta = 0, 2, 4$

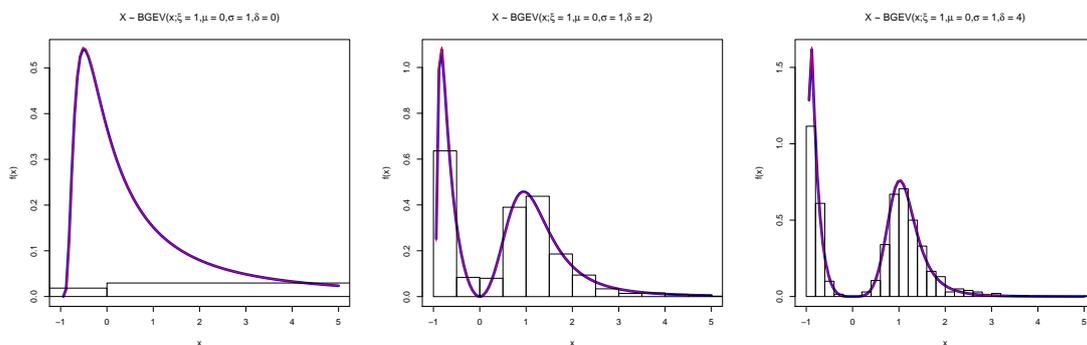


Figura 4.14: $BGEV(x; \mu = 0, \sigma = 1, \xi = 1, \delta)$, $\delta = 0, 2, 4$

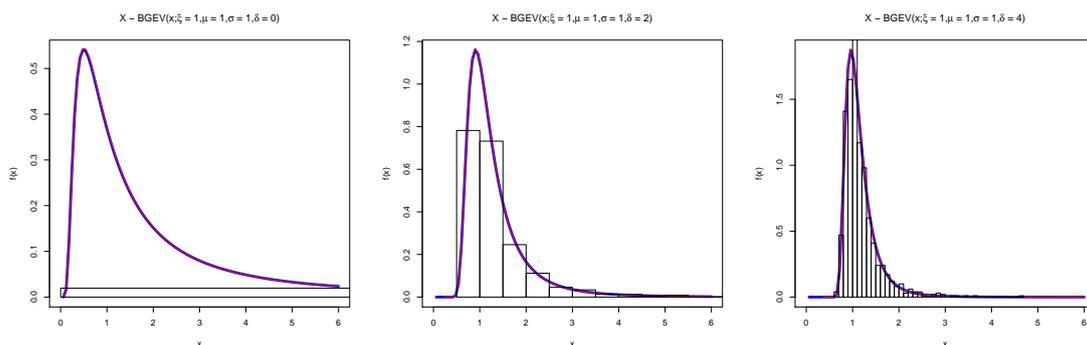


Figura 4.15: $BGEV(x; \mu = 0, \sigma = 1, \xi = 1, \delta)$, $\delta = 0, 2, 4$

Capítulo 5

Aplicação

O grande incentivo da realização deste trabalho se deu pelo encontro de dados de eventos raros que apresentaram bimodalidade. Em especial dados climatológicos que usualmente são modelados pela distribuição generalizada de eventos extremos - GEV.

Tais dados são coletados e disponibilizados pelo INMET - Instituto Nacional de Meteorologia. As estações meteorológicas espalhadas em todo o Brasil analisam diversas variáveis sobre as condições atmosféricas em determinado momento.

Os dados são abertos e acessíveis pelo site do INMET, <http://www.inmet.gov.br>. Foi selecionado, na aba referente a Estações de Dados > Estações Automáticas. Para cada hora do dia, que começa a 00:00 até as 23:00 horas, é feita uma observação no instante, máximo e mínimo para as variáveis temperatura (°C), Umidade (%), Ponto de orvalho (°C) e Pressão (hPa), já para a variável vento (m/s) e feita uma observação para direção, velocidade e rajada, as outras variáveis são Radiação (km/m^2) e chuva (mm). Os dados aqui analisados correspondem à velocidade do vento e ponto de orvalho nos dias 12/10/2018 até 12/10/2019 de Brasília - Distrito Federal. A escolha das estações automáticas está relacionada ao fato de possuírem maior volume de dados e por possuírem algumas variáveis a mais em relação às estações convencionais.

A Figura 5.1 apresenta o histograma da variável velocidade do vento, em que nitidamente é observada bimodalidade.

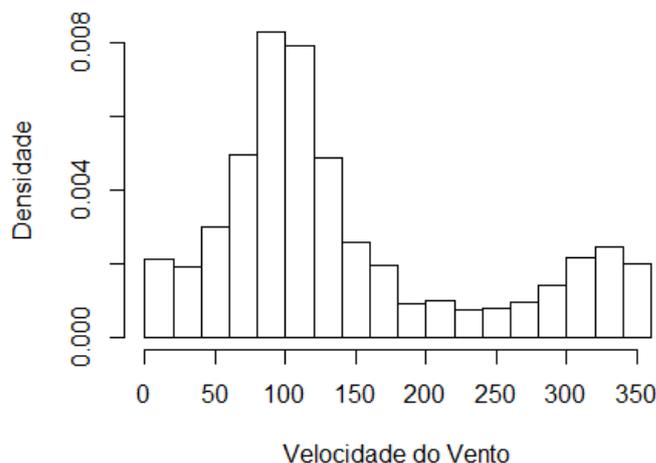


Figura 5.1: Histograma da velocidade do vento.

A Tabela 5.1 apresenta algumas medidas descritivas dos dados correspondentes à velocidade do vento. Nela se observa uma grande variabilidade da variável em estudo. O valor médio não pode ser interpretado devido a bimodalidade.

Min.	1° Quartil	Mediana	Média	3° Quartil	Max.
1.0	82.0	112.0	143.3	179.0	360.0

Tabela 5.1: Estatísticas descritivas para velocidade de vento.

Antes de estimar os parâmetros da distribuição BGEV que deve-se ajustar aos dados da velocidade do vento, primeiramente foi feita uma normalização nos dados. O primeiro passo foi subtrair um valor próximo à observação menos frequente (não exatamente igual para que não exista uma observação zero). Esse passo irá transladar os dados para que o ponto de mínimo local fique próximo à zero. Outra modificação realizada foi a divisão dos dados pelo seu desvio padrão para que diminuir a variabilidade dos dados. Após a normalização, a Figura 5.2 apresenta o novo histograma dos dados.

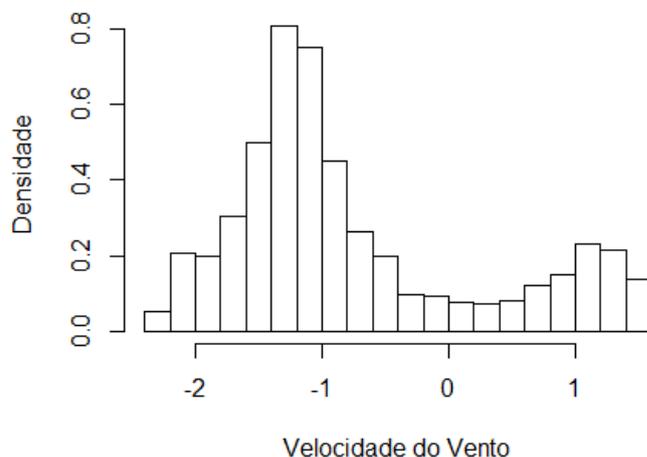


Figura 5.2: Histograma da velocidade do vento normalizado.

Min.	1° Quartil	Mediana	Média	3° Quartil	Max.
-2.26	-1.40	-1.08	-0.75	-0.38	1.54

Tabela 5.2: Estatísticas descritivas para velocidade de vento normalizado.

A partir dos dados normalizados ajusta-se a distribuição BGEV estimando os parâmetros pelo método de máxima verossimilhança, com a função 3.1. O algoritmo numérico de Nelder and Mead (1965) foi utilizado para a estimação. Os chutes iniciais para o algoritmo foram escolhidos de tal forma que atendam às restrições da função de máxima verossimilhança, atentando-se também ao domínio da distribuição, de forma que os chutes iniciais sejam de uma distribuição em que o domínio acolha os dados. O estudo gráfico realizado na seção 2.1.2 ajudou na escolha dos chutes iniciais para o vetor de parâmetros do modelo que deve-se ajustar aos dados. Aqui, os valores de chute inicial foram $\mu = 0$, $\sigma = 1$, $\xi = -0.5$ e $\delta = 0.5$. Sendo assim os parâmetros estimados da distribuição BGEV para os dados de velocidade de vento normalizado estão apresentados na Tabela 5.3.

$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\delta}$
-1.0430436	0.6516189	-0.3042093	0.9434247

Tabela 5.3: Estimativas dos parâmetros para os dados.

Para visualização do ajuste da distribuição aos dados a Figura 5.3 apresenta o histograma dos dados juntamente com a densidade BGEV ajustada. Também o gráfico do QQ-plot que é um método gráfico para comparar a distribuição de probabilidade empírica com a distribuição BGEV ajustada pelos parâmetros estimados.

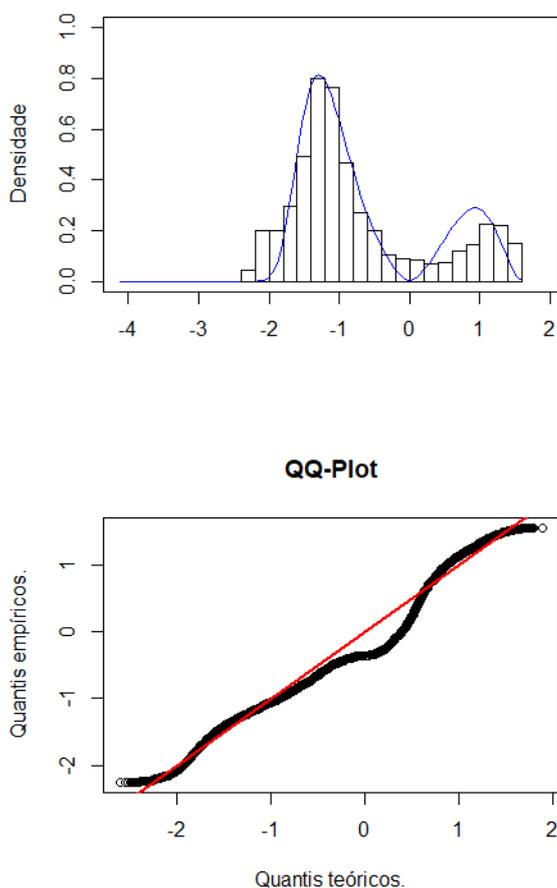


Figura 5.3: Histograma dos dados versus BGEV ajustada e QQ-plot empírico versus ajustado.

Graficamente pode-se observar que a distribuição BGEV tem um bom ajuste aos dados utilizados, atentando-se ao ajuste próximo ao zero, em que, pelo QQ-plot é possível visualizar um desvio. Para analisar numericamente o ajustamento da distribuição estimada, foi realizado o teste de Kolmogorov-Smirnov, este é um teste não paramétrico baseado na diferença absoluta máxima entre a função de distribuição acumulada estimada e a função de distribuição empírica dos dados, e também o teste de Anderson Darling. Ambos testam as seguintes hipóteses:

$$\begin{cases} H_0 : \text{Os dados seguem a distribuição em questão;} \\ H_A : \text{Os dados não seguem a distribuição.} \end{cases}$$

Estão apresentados na Tabela 5.8 os resultados para as estatísticas dos teste e seus respectivos p-valores.

D_n	p-valor	A^2	p-valor
0.05396	< 2.2e-16	47.802	6.919e-08

Tabela 5.4: Testes não paramétricos de ajustamento.

Abaixo estão apresentados os resultados da aplicação para a variável ponto de orvalho, a Figura 5.4 apresenta o histograma para esta variável. E a Tabela 5.5 as medidas descritivas dos dados correspondentes.

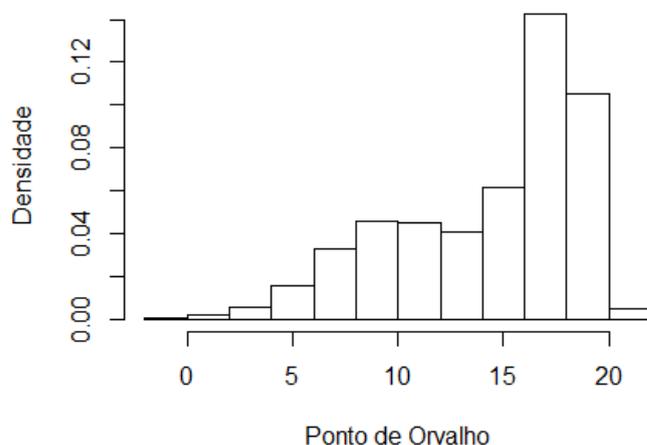


Figura 5.4: Histograma da ponto de orvalho.

Min.	1° Quartil	Mediana	Média	3° Quartil	Max.
-2.00	11.10	16.10	14.42	17.90	21.20

Tabela 5.5: Estatísticas descritivas para ponto de orvalho.

O mesmo processo de normalização dos dados foi realizado. Assim, foi subtraído um valor próximo à observação menos frequente (não exatamente igual para que não exista uma observação zero) e foi realizada a divisão dos dados pelo seu desvio padrão para que diminuir a variabilidade. Após a normalização, a Figura 5.5 apresenta o novo histograma dos dados.

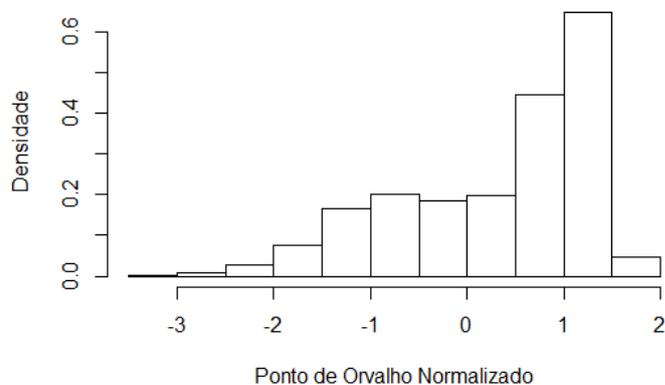


Figura 5.5: Histograma do ponto de orvalho normalizado.

Min.	1° Quartil	Mediana	Média	3° Quartil	Max.
-3.44	-0.43	0.71	0.32	1.12	1.8833

Tabela 5.6: Estatísticas descritivas para ponto de orvalho normalizado.

O método de máxima verossimilhança foi utilizado para ajustar a distribuição BGEV aos dados. Igualmente, os chutes iniciais foram escolhidos de tal forma que atendam às restrições da função de máxima verossimilhança, atentando-se também ao domínio da distribuição. Aqui, os valores de chute inicial foram $\mu = 0, \sigma = 1, \xi = -0.25$ e $\delta = 0.5$. Sendo assim os parâmetros estimados da distribuição BGEV para os dados de velocidade de vento normalizado estão apresentados na Tabela 5.7.

A Figura 5.6 apresenta o histograma dos dados juntamente com a densidade BGEV ajustada.

O gráfico indica bom ajuste da distribuição aos dados, na Tabela ?? estão apresentados testes estatísticos para verificação do ajuste.

$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\delta}$
0.1804285	0.8196716	-0.4839371	0.4201143

Tabela 5.7: Estimativas dos parâmetros para ponto de orvalho.

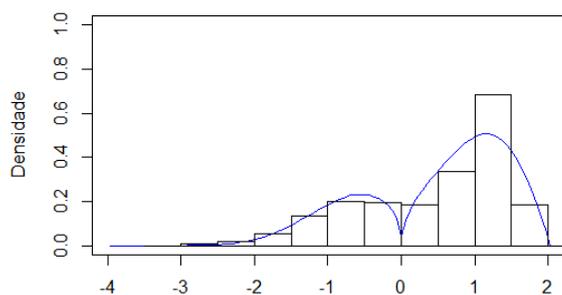


Figura 5.6: Histograma dos dados versus BGEV ajustada.

D_n	p-valor	A^2	p-valor
0.073807	< 2.2e-16	66.297	6.919e-08

Tabela 5.8: Testes não paramétricos de ajustamento.

Capítulo 6

Conclusão

A distribuição BGEV, ou Bimodal GEV, proposta neste trabalho é um modelo que generaliza a distribuição GEV, usualmente utilizada para modelar eventos raros. A generalização proposta buscou ajustar a distribuição tornando-a capaz de modelar bem dados extremos que apresentem bimodalidade. A generalização proposta é capaz de modelar bem dados extremos que apresentem bimodalidade. A BGEV se mostrou capaz de modelar dados heterogêneos independentemente distribuídos. O estudo de uma distribuição BGEV para dados temporais fica em aberto para ser tratado em trabalhos futuros.

Bibliografia

- [1] Fisher, R. A. and Tippett (1928), L. H. C., Limiting forms of the frequency distribution of the largest or smallest member of a sample, Proceedings of the Cambridge Philosophical Society, vol 24, pp. 180-290.

- [2] Jenkinson, A. F (1955). The frequency distribution of the annual maximum (or minimum) of meteorological elements, Quarterly Journal of the Royal Meteorological Society, London, v.81, n.551.501.45, p.158-171.

- [3] Reiss, R. D and Thomas M. (1997). Statistical Analysis of Extreme Values wuth Applications to Insurance, Finance, Hydrology and Other Fields, Birkhäuser Verlag, Basel.

- [4] E. Jondeau, S-H. Poon and M. Rockinger (2007). Financial Modeling under Non-Gaussian Distributions. Spring Verlag, New York .

- [5] P. Embrechts, C. Klüppelberg and T. Mikosch (1999). Modelling Extremal Events for Insurance and Finance. Applications of Mathematics. Springer, Ed. 2°.

- [6] J. Galambos (1978). The assymptotic Theory of Extreme Order Statistics. John Wiley Sons, New York.

- [7] B. V. Gnedenko (1943). Sur la distribution limité du terme d'une série aléatoire. *Ann Math.* 44:423-453.
- [8] S. I. Resnick (1987). *Extreme Values, Regular Variation and Weak Convergence*. Springer-Verlag, New York.
- [9] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction* Springer Science + Business Media. LLC, New York.
- [10] Kotz, Samuel and Nadarajah, Saralees (2000). *Extreme value distributions: Theory and Applications*. Imperial College Press, London.
- [11] Mendes, Beatriz V. M. (2004). *Introdução à Análise de Eventos Extremos*.
- [12] Serge B. Provost, Abdus Saboor, Gauss M. Cordeiro and Muhammad Mansoor (2018). *On the q-Generalized Extreme Value Distribution*.
- [13] Nelder, J. A. and Mead, R. (1965) A simplex algorithm for function minimization. *Computer Journal* 7, 308-313.
- [14] Towe, R., Eastoe, E., Tawn, J., Jonathan, P. (2017). Statistical downscaling for future extreme wave heights in the North Sea. *The Annals of Applied Statistics*, 11(4), 2375-2403.

Anexos

6.1 Anexo A

```
## ----- SIMULAÇÃO -----  
# gev ----  
gev.padrao <- function(y,E){  
  if(E != 0){  
    y <- (((1+(E*(y)))^((-1/E)-1))*  
          (exp(-(1+(E*y))^(-1/E))))  
  }else{  
    y <- exp(-y-exp(-y))  
  }  
  return(y)  
}  
  
gev. <- function(y,E,s,u){  
  y <- (gev.padrao(((y-u)/s),E)) / s  
  return(y)  
}
```

```

# gev BIMODAL ----
gev.B <- function(x,E,s,u,d){
  Td <- s * x * abs(x)^d
  Td1 <- s * (d+1) * abs(x)^d
  Zd <- 1 + d^2 * s^2 * ((pi^2)/6) +
    (d*u + d*s*(-digamma(1)) -1)^2

  if(E != 0 & d > -1){
    y <- gev.(Td,E,s=1,u) * Td1
  }else{
    y <- gev.(x,E,s,u) * (1/Zd) * ((1-dx)^2 + 1)
  }
  return(y)
}

dominio <- function(E,s,u,d){

  y <- u-(1/E)

  if(y>=0){ limite <- (y/s)^(1/(1+d))
  }else{ limite <- -(-y/s)^(1/(1+d)) }

  if(E>0){ x <- c(limite, limite+6)
  }else{ x <- c(limite-6,limite) }

  return(x)
}

```

```

}

# EXEMPLO
E <- -0.25; s <- 1; u <- -1; d <- 4
curve(gev.B(x,E,s,u,d), xlim = dominio(E,s,u,d))

## QUANTIL ----
quantil.gev.B <- function(Q,E,s,u,d){

  Fgev <- ( u + ((-log(Q))^-E) - 1)/(E) ) # pra gev s=1

  FgevMAIS <- Fgev[Fgev >= 0]
  yMAIS <- (FgevMAIS/s)^(1/(1+d))

  FgevMENOS <- Fgev[Fgev < 0]
  yMENOS <- -(-FgevMENOS/s)^(1/(1+d))

  return(c(yMAIS,yMENOS))
}

# gerando amostras
E <- -0.5 ; s <- 1 ; u <- 2.2 ; d <- 3

curve(gev.B(x,E,s,u,d), xlim = dominio(E,s,u,d))

```

```

amostra <- quantil.gev.B(Q = runif(1000), E,s,u,d)
hist(amostra, breaks=150, freq=F, add=T)

# VEROSSIMILHANÇA
vero<-function(para,x) {

  E<-para[1]
  s <- 1 # s<-para[2]
  u<-para[2]
  d<-para[3]

  if( (s > 0) && (d >= -1) &&
      ( (E<0 && u-(1/E)>0 && max(x) < ((u-(1/E))/s)^(1/(d+1)) ) |
        (E<0 && u-(1/E)<0 && max(x) < -(-(u-(1/E))/s)^(1/(d+1)) ) |
        (E>0 && u-(1/E)>0 && min(x) > ((u-(1/E))/s)^(1/(d+1)) ) |
        (E>0 && u-(1/E)<0 && min(x) > -(-(u-(1/E))/s)^(1/(d+1)) ) ) ) {

    l <- -((1/E)+1)*sum(log(1+(E*(s*x*(abs(x)^d)-u)))) -
      sum((1+(E*(s*x*(abs(x)^d)-u)))^(-1/E)) +
      length(x)*(log(s)+log(1+d)) +
      d*sum(log(abs(x)))

  } else{ l <- Inf }

  return(-l)
}

```

```

## SIMULAÇÃO

setwd("C:\\Users\\b3002848\\Downloads\\simulação")

for(i in c(0.25,0.5,1,-0.25,-0.5)){ # substituir o valor de E: i <-0.25
  for(j in c(-1, 0, 1)){          # substituir o valor de u: j <- -1
    for(k in c(0, 2, 4)){        # substituir o valor de d: k <- 0

      parametros.estimados<-data.frame()
      erro <- 0
      n=1

      system.time(
        while(n <= 100){

          x<-quantil.gev.B(Q= runif(1000),
                          E= i,s= 1,u= j,d= k)

          v1 <- tryCatch(optim(c(E= i+runif(1,-1,1),
                                u= j+runif(1,-1,1),
                                d= k+runif(1,-1,1)),
                                vero,
                                NULL, hessian=T, x=x),
                              error=function(e) "Erro")

          print(n)
        }
      )
    }
  }
}

```

```

    if(sum(v1 != "Erro") && v1$convergence == 0){

        parametros.estimados[n,"E"]<- v1$par[1]
        parametros.estimados[n,"u"]<- v1$par[2]
        parametros.estimados[n,"d"]<- v1$par[3]
        parametros.estimados[n,"Erro"] <- erro

        n=n+1

    }else{ erro <- erro+1 }
}
)

parametros.estimados$Erro <- c(parametros.estimados$Erro[1],diff(pa
nomes <- c(paste0("E=", gsub("\\.", "",i)),
           paste0("u=", gsub("\\.", "",j)),
           paste0("d=", gsub("\\.", "",k)),
           "Erro")

names(parametros.estimados) <- nomes
saveRDS(parametros.estimados, file = paste0(paste0(nomes[-4], colla

}

}

}

```

```

## MEDIDAS

for(M in c("", "(n=500)")){ # M <- ""

  if(M == ""){m="(n=1000)"}else{m=M}

  assign("theta_hat", data.frame())

  for(i in c(0.25,0.5,1,-0.25,-0.5)){ # substituir o valor de E: i <- 1
    for(j in c(-1, 0, 1)){          # substituir o valor de u: j <- 1
      for(k in c(0, 2, 4)){        # substituir o valor de d: k <- 3

        nomes <- c(paste0("E=", i),
                   paste0("u=", j),
                   paste0("d=", k),
                   "Erro")

        x <- readRDS(paste0(paste0(gsub("\\.", "", nomes[-4]), collapse = "."))

        y <- apply(x,2,function(x) round(mean(x),3))
        z <- round(c(i, j, k) - y[-4],3)
        w <- apply(x[-4],2, function(x) round(sum((x - mean(x))^2)/100,4))

        assign("theta_hat",
              rbind(get("theta_hat"),
                    c(i, j, k,

```

```

y[1],y[2],y[3],
z[1],z[2],z[3],
w[1],w[2],w[3])) )

## GRAFICOS

pdf(paste0(paste0(gsub("\\.", "", nomes[-4]), collapse = "_"), M, ".p

titulo <- substitute(paste("X ~ BGEV(x; ",
                           xi, " = ", i, ", ", ",
                           mu, " = ", j, ", ", ",
                           sigma, " = 1, ",
                           delta, " = ", k, ")"),
                      list(i = i, j = j, k = k))

# ggplot(data.frame(x=dominio(i,1,j,k)), aes(x)) +
#   stat_function(fun=function(x) gev.B(x,i,1,j,k)) +
#   labs(title = titulo) +
#   stat_function(fun=function(x) gev.B(x,y[1],1,y[2],y[3]), col
#   geom_histogram(x)

if(M == ""){tam=1000}else{tam=500}

x<-quantil.gev.B(Q= runif(tam), E= i,s= 1,u= j,d= k)

curve(gev.B(x,i,1,j,k), xlim = dominio(i,1,j,k),

```

```

        lwd=5, col= "blue", ylab = "f(x)", main= titulo )

curve(gev.B(x,y[1],1,y[2],y[3]), col = "red", add = T)

hist(x, breaks=30, freq=F, add=T)

dev.off()

    }
  }
}

colnames(theta_hat) <- c("PAR.E", "PAR.u", "PAR.d",
                        "EST.E", "EST.u", "EST.d",
                        "VIES.E", "VIES.u", "VIES.d",
                        "EQM.E", "EQM.u", "EQM.d")

write.csv2(theta_hat, paste0("MEDIDAS",m,".csv"), row.names = F)
rm(theta_hat,i,j,k,nomes,w,x,y,z)

}

rm(M,m)

# analisando por grid:
# install.packages("animation")

```

```

library(animation)

# E: i <- 3
# u: j <- -1
# d: k <- 3

saveGIF({
  for (v in seq(-3, 5, 0.25)){ # v <- -0.80

    titulo <- substitute(paste("X ~ BGEV(x;",
                                xi," = ",i," ",
                                # mu," = ",j," ",
                                delta," = ",k," ",
                                sigma," = 1)"),
                          list(i = i, j = j, k = k))

    curve(gev.B(x,i,1,v,k), xlim = dominio(i,1,v,k),
          lwd=5, col= "blue", ylim = c(0,2),
          ylab = "f(x)", main= titulo )

    legend("topright"
          , legend = substitute(paste(mu, "=", v), list(v=v))
          , lwd = 1, col = "black", cex=0.9, bty = "n")
  }
}, movie.name = "E=-05.gif", interval = 1, ani.width = 700, ani.height =

```

```

i <- -.5
k <- 3

titulo <- substitute(paste("X ~ BGEV(x;",
                           xi," = ",i,",",
                           delta," = ",k,",",
                           sigma," = 1)"),
                    list(i = i, j = j, k = k))

grid <- c(2.75, 3, 3.15, 3.3)
cores <- c("red", "blue", "green", "yellow")

v <- grid[1]

curve(gev.B(x,i,1,v,k), xlim = dominio(i,1,v,k),
      lty = 1, ylim = c(0,3.5), col = cores[1],
      ylab = "f(x)", main= titulo )

for(w in 2:length(grid)){
  v <- grid[w]
  curve(gev.B(x,i,1,v,k), col = cores[w],
        from = dominio(i,1,v,k), add = T)
}

grid

```

```
legend("topright"  
      , legend = c(expression(paste(mu, " = 0.75")),  
                    expression(paste(mu, " = 0.95")),  
                    expression(paste(mu, " = 1.00")),  
                    expression(paste(mu, " = 1.10")))  
      , lty=1, col = cores, cex=0.9, bty = "n")  
  
x <- read.csv("MEDIDAS (n=1000).csv", sep=";")  
  
# estimativas  
xtable(x[,c("PAR.E", "PAR.u", "PAR.d", "EST.E", "EST.u", "EST.d")])  
  
# vies  
xtable(x[,c("PAR.E", "VIES.E", "PAR.u", "VIES.u", "PAR.d", "VIES.d")])  
  
# EQM  
xtable(x[,c("PAR.E", "EQM.E", "PAR.u", "EQM.u", "PAR.d", "EQM.d")])  
  
  
# DADOS REAIS  
  
setwd("C:\\Users\\b3002848\\Downloads")  
dados1<-read.table("dadosINMET.txt", sep="," ,dec = ".", header = TRUE)
```

```
hist(dados1$vento_vel, xlab = "Velocidade do Vento", ylab = " Densidade",  
summary(dados1$vento_vel)
```

```
library(tidyverse)
```

```
dados1 %>% count(vento_vel) %>% top_n(-5)
```

```
sd(dados1$vento_vel)
```

```
# não posso subtrair exatamente o valor de menor frequencia
```

```
# porque não posso ter ZERO nos dados
```

```
dados <- (dados1$vento_vel - 214.5)/94.5
```

```
hist(dados,
```

```
  xlab = "Velocidade do Vento Normalizada",
```

```
  ylab = " Densidade", main = " ", prob=T) #ponto de orvalho
```

```
par(mfrow=c(1,2))
```

```
plot(density(dados),
```

```
  xlab = "Velocidade do Vento",
```

```
  ylab = " Densidade", main = " ")
```

```
hist(dados, freq = F, add=T)
```

```
E <- -1/2 ; s <- 1 ; u <- -1 ; d <- 2
```

```

curve(gev.B(x,E,s,u,d), xlim = c(-2.5,1))

par(mfrow=c(1,1))

vero<-function(para,x){

  E<-para[1]
  s<-para[2]
  u<-para[3]
  d<-para[4]

  if( (s > 0) && (d >= -1) &&
      ( (E<0 && u-(1/E)>0 && max(x) < ((u-(1/E))/s)^(1/(d+1)) ) |
        (E<0 && u-(1/E)<0 && max(x) < -(-(u-(1/E))/s)^(1/(d+1)) ) |
        (E>0 && u-(1/E)>0 && min(x) > ((u-(1/E))/s)^(1/(d+1)) ) |
        (E>0 && u-(1/E)<0 && min(x) > -(-(u-(1/E))/s)^(1/(d+1)) ) ) ) {

    l <- -((1/E)+1)*sum(log(1+(E*(s*x*(abs(x)^d)-u)))) -
      sum((1+(E*(s*x*(abs(x)^d)-u)))^(-1/E)) +
      length(x)*(log(s)+log(1+d)) +
      d*sum(log(abs(x)))

  } else{ l <- Inf }

  return(-l)
}

```

```
}

est <- tryCatch(optim(c(E= -0.5,
                      s=1,
                      u= 0,
                      d= 0.5),
                    vero,
                    NULL, hessian=T, x=dados),
               error=function(e) "Erro")

print(est)

est$par

E <- est$par[1]
s <- est$par[2]
u <- est$par[3]
d <- est$par[4]

curve(gev.B(x,E,s,u,d), xlim = dominio(E,s,u,d), ylim = c(0,1),
      xlab = "",
      ylab = "Densidade", main = " ", col="blue")
hist(dados, freq=F, add=T)
plot(density(dados), add=T)

# QQPlot
xo <- sort(dados)
```

```
pteo <- (1:length(x))/length(x) # probabilidades teóricas

qteo <- quantil.gev.B(Q = pteo,
                    E = -0.3042093,
                    s = 0.6516189,
                    u = -1.0430436,
                    d = 0.9434247)

# quantis teóricos sob a distribuição

plot(xo~sort(qteo),
     xlab = "Quantis teóricos.",
     ylab = "Quantis empíricos.",
     main = "QQ-Plot") # quantis observados ~ quantis teó
abline(a = 0, b = 1, lwd=2, col = "red")

FBGEV <- function(x,u,s,E,d){
  Tx <- x*s*((abs(x))^d)
  return( exp(-(1+E*(Tx-u))^(1/E)) )
}

est$par

require("gofstest")

ks.test(dados, "FBGEV", est$par[3], est$par[2], est$par[1], est$par[4])
gofstest::ad.test(dados, "FBGEV", est$par[3], est$par[2], est$par[1], est$
```

6.2 Anexo B

Resultados da simulação para tamanho de amostra $n = 50$.

Tabela 6.1: Resultados da simulação para tamanho de amostra $n = 50$.

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
0,25	-1	0	0,256	-1,008	0,04	-0,006	0,008	-0,04	0,0216	0,0322	0,0116
0,25	-1	2	0,244	-1,018	2,116	0,006	0,018	-0,116	0,0162	0,0323	0,0838
0,25	-1	4	0,239	-1,013	4,185	0,011	0,013	-0,185	0,0186	0,0262	0,2052
0,25	0	0	0,273	0,016	0,056	-0,023	-0,016	-0,056	0,0302	0,0155	0,0251
0,25	0	2	0,285	0,026	2,131	-0,035	-0,026	-0,131	0,035	0,0168	0,2199
0,25	0	4	0,275	0,02	4,213	-0,025	-0,02	-0,213	0,0247	0,0143	0,4635
0,25	1	0	0,267	1,001	0,037	-0,017	-0,001	-0,037	0,0803	0,029	0,0414
0,25	1	2	0,289	1,017	2,197	-0,039	-0,017	-0,197	0,0827	0,0283	0,5167
0,25	1	4	0,292	1,009	4,309	-0,042	-0,009	-0,309	0,0812	0,0226	1,3124
0,5	-1	0	0,501	-0,999	0,032	-0,001	-0,001	-0,032	0,0233	0,024	0,0129
0,5	-1	2	0,524	-1,032	2,15	-0,024	0,032	-0,15	0,0268	0,025	0,1033
0,5	-1	4	0,486	-1,019	4,215	0,014	0,019	-0,215	0,0254	0,023	0,2411
0,5	0	0	0,549	0	0,064	-0,049	0	-0,064	0,0375	0,0154	0,0282
0,5	0	2	0,562	0,015	2,118	-0,062	-0,015	-0,118	0,045	0,0175	0,1552
0,5	0	4	0,551	0,019	4,282	-0,051	-0,019	-0,282	0,0303	0,0124	0,4835
0,5	1	0	0,541	1,029	0,057	-0,041	-0,029	-0,057	0,0725	0,0255	0,0515
0,5	1	2	0,559	1,006	2,094	-0,059	-0,006	-0,094	0,0786	0,0254	0,3663
0,5	1	4	0,563	0,998	4,379	-0,063	0,002	-0,379	0,083	0,0238	1,7225
1	-1	0	1,018	-1,017	0,036	-0,018	0,017	-0,036	0,049	0,0214	0,0171

Continua na próxima página.

Tabela 6.1 – continuação

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
1	-1	2	1,038	-1,013	2,082	-0,038	0,013	-0,082	0,039	0,0206	0,1167
1	-1	4	1,034	-1,053	4,292	-0,034	0,053	-0,292	0,0449	0,0201	0,3219
1	0	0	1,085	0,007	0,048	-0,085	-0,007	-0,048	0,0527	0,0081	0,0229
1	0	2	1,085	0,009	2,126	-0,085	-0,009	-0,126	0,0433	0,007	0,1972
1	0	4	1,049	0,028	4,152	-0,049	-0,028	-0,152	0,0546	0,0078	0,4478
1	1	0	1,148	0,97	0,096	-0,148	0,03	-0,096	0,0797	0,029	0,0668
1	1	2	1,083	0,995	2,14	-0,083	0,005	-0,14	0,0867	0,0279	0,4675
1	1	4	1,125	0,973	4,452	-0,125	0,027	-0,452	0,0954	0,0295	2,3334
-0,25	-1	0	-0,266	-0,981	0,055	0,016	-0,019	-0,055	0,0071	0,0269	0,0197
-0,25	-1	2	-0,288	-1,012	2,151	0,038	0,012	-0,151	0,0082	0,0279	0,0849
-0,25	-1	4	-0,296	-0,997	4,288	0,046	-0,003	-0,288	0,0093	0,0259	0,2736
-0,25	0	0	-0,286	0,002	0,024	0,036	-0,002	-0,024	0,0159	0,0298	0,018
-0,25	0	2	-0,266	0,003	2,094	0,016	-0,003	-0,094	0,0125	0,0243	0,1833
-0,25	0	4	-0,283	0,031	4,079	0,033	-0,031	-0,079	0,0163	0,0252	0,439
-0,25	1	0	-0,256	1,01	0,028	0,006	-0,01	-0,028	0,0326	0,0312	0,0211
-0,25	1	2	-0,278	1,03	2,119	0,028	-0,03	-0,119	0,0332	0,0224	0,2411
-0,25	1	4	-0,257	1,037	4,165	0,007	-0,037	-0,165	0,0374	0,0359	0,6435
-0,5	-1	0	-0,553	-0,955	0,044	0,053	-0,045	-0,044	0,0102	0,0305	0,0123
-0,5	-1	2	-0,563	-0,948	2,229	0,063	-0,052	-0,229	0,0129	0,0319	0,1927
-0,5	-1	4	-0,579	-0,952	4,25	0,079	-0,048	-0,25	0,0077	0,0209	0,3505
-0,5	0	0	-0,545	0,04	0,017	0,045	-0,04	-0,017	0,012	0,0259	0,0197
-0,5	0	2	-0,565	0,032	2,005	0,065	-0,032	-0,005	0,0112	0,0224	0,1354
-0,5	0	4	-0,553	0,043	4,069	0,053	-0,043	-0,069	0,0127	0,0275	0,4623
-0,5	1	0	-0,562	1,021	-0,012	0,062	-0,021	0,012	0,0258	0,0248	0,0211
-0,5	1	2	-0,532	1,025	2,067	0,032	-0,025	-0,067	0,0247	0,0286	0,1658

Continua na próxima página.

Tabela 6.1 – continuação

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
-0,5	1	4	-0,555	1,007	4,035	0,055	-0,007	-0,035	0,0257	0,0225	0,5358

Resultados da simulação para tamanho de amostra $n = 100$.

Tabela 6.2: Resultados da simulação para tamanho de amostra $n = 100$.

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
0,25	-1	0	0,245	-1,021	0,037	0,005	0,021	-0,037	0,0096	0,0199	0,0054
0,25	-1	2	0,247	-1,007	2,071	0,003	0,007	-0,071	0,0075	0,0157	0,048
0,25	-1	4	0,241	-1,004	4,045	0,009	0,004	-0,045	0,007	0,0114	0,1047
0,25	0	0	0,255	0,027	0,028	-0,005	-0,027	-0,028	0,0104	0,0074	0,007
0,25	0	2	0,262	0,01	2,067	-0,012	-0,01	-0,067	0,0107	0,0087	0,0579
0,25	0	4	0,263	0,007	4,115	-0,013	-0,007	-0,115	0,0101	0,0085	0,1793
0,25	1	0	0,249	1,009	0,011	0,001	-0,009	-0,011	0,025	0,0108	0,0124
0,25	1	2	0,272	1	2,058	-0,022	0	-0,058	0,0325	0,0117	0,1798
0,25	1	4	0,241	1,021	4,052	0,009	-0,021	-0,052	0,0232	0,0104	0,404
0,5	-1	0	0,503	-1,011	0,024	-0,003	0,011	-0,024	0,0113	0,0142	0,0068
0,5	-1	2	0,503	-1,02	2,09	-0,003	0,02	-0,09	0,0088	0,0128	0,0469
0,5	-1	4	0,518	-1,017	4,145	-0,018	0,017	-0,145	0,0094	0,011	0,1269
0,5	0	0	0,525	0,005	0,025	-0,025	-0,005	-0,025	0,0115	0,0052	0,0084
0,5	0	2	0,511	0,021	2,067	-0,011	-0,021	-0,067	0,0105	0,0065	0,0663
0,5	0	4	0,52	0,012	4,136	-0,02	-0,012	-0,136	0,0131	0,0057	0,2212
0,5	1	0	0,53	1,012	0,011	-0,03	-0,012	-0,011	0,0281	0,01	0,0195
0,5	1	2	0,53	0,997	2,102	-0,03	0,003	-0,102	0,0519	0,0125	0,3862
0,5	1	4	0,516	1,002	4,093	-0,016	-0,002	-0,093	0,024	0,0084	0,3418

Continua na próxima página.

Tabela 6.2 – continuação

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
1	-1	0	1,018	-1,002	0,009	-0,018	0,002	-0,009	0,0195	0,0073	0,0069
1	-1	2	1,037	-1,007	2,025	-0,037	0,007	-0,025	0,0151	0,0103	0,0635
1	-1	4	1,022	-1,007	4,114	-0,022	0,007	-0,114	0,019	0,0092	0,1591
1	0	0	1,038	0,005	0,024	-0,038	-0,005	-0,024	0,0189	0,0027	0,0086
1	0	2	1,07	-0,005	2,079	-0,07	0,005	-0,079	0,021	0,0034	0,089
1	0	4	1,029	0,005	4,079	-0,029	-0,005	-0,079	0,0177	0,0039	0,2877
1	1	0	1,08	0,97	0,071	-0,08	0,03	-0,071	0,034	0,0142	0,0295
1	1	2	1,076	0,978	2,199	-0,076	0,022	-0,199	0,037	0,0124	0,2379
1	1	4	1,03	0,995	4,157	-0,03	0,005	-0,157	0,0348	0,0124	0,591
-0,25	-1	0	-0,264	-1,006	0,027	0,014	0,006	-0,027	0,0033	0,0149	0,0051
-0,25	-1	2	-0,264	-0,992	2,056	0,014	-0,008	-0,056	0,0032	0,0169	0,0435
-0,25	-1	4	-0,272	-1,002	4,156	0,022	0,002	-0,156	0,0041	0,0116	0,134
-0,25	0	0	-0,258	0,002	0,028	0,008	-0,002	-0,028	0,0053	0,0138	0,0095
-0,25	0	2	-0,271	0,015	2,02	0,021	-0,015	-0,02	0,0056	0,014	0,0709
-0,25	0	4	-0,255	-0,001	4,095	0,005	0,001	-0,095	0,005	0,0118	0,2409
-0,25	1	0	-0,256	1,018	0,004	0,006	-0,018	-0,004	0,0115	0,0118	0,0093
-0,25	1	2	-0,275	1,026	2,007	0,025	-0,026	-0,007	0,0082	0,013	0,0599
-0,25	1	4	-0,274	1,015	4,02	0,024	-0,015	-0,02	0,0103	0,0131	0,2
-0,5	-1	0	-0,536	-0,964	0,025	0,036	-0,036	-0,025	0,0036	0,0131	0,0052
-0,5	-1	2	-0,528	-0,982	2,076	0,028	-0,018	-0,076	0,003	0,0114	0,0457
-0,5	-1	4	-0,531	-0,984	4,182	0,031	-0,016	-0,182	0,0027	0,0093	0,1574
-0,5	0	0	-0,523	0,015	0,015	0,023	-0,015	-0,015	0,004	0,0113	0,008
-0,5	0	2	-0,532	0,026	2,053	0,032	-0,026	-0,053	0,0042	0,011	0,0643
-0,5	0	4	-0,513	0,008	4,153	0,013	-0,008	-0,153	0,0042	0,0101	0,2634
-0,5	1	0	-0,516	1,01	0,016	0,016	-0,01	-0,016	0,011	0,0113	0,0102

Continua na próxima página.

Tabela 6.2 – continuação

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
-0,5	1	2	-0,532	1,008	1,988	0,032	-0,008	0,012	0,0095	0,0116	0,0731
-0,5	1	4	-0,527	0,983	3,955	0,027	0,017	0,045	0,0089	0,0103	0,1963

Resultados da simulação para tamanho de amostra $n = 250$.

Tabela 6.3: Resultados da simulação para tamanho de amostra $n = 250$.

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
0,25	-1	0	0,255	-0,994	0,007	-0,005	-0,006	-0,007	0,0022	0,0043	0,0017
0,25	-1	2	0,243	-1,003	2,053	0,007	0,003	-0,053	0,0026	0,0046	0,0186
0,25	-1	4	0,254	-1,014	4,072	-0,004	0,014	-0,072	0,003	0,0052	0,0436
0,25	0	0	0,267	0,002	0,02	-0,017	-0,002	-0,02	0,0028	0,0033	0,0032
0,25	0	2	0,255	0,005	2,01	-0,005	-0,005	-0,01	0,0034	0,0026	0,0229
0,25	0	4	0,249	0,001	4,038	0,001	-0,001	-0,038	0,0046	0,0031	0,0772
0,25	1	0	0,236	1,01	-0,003	0,014	-0,01	0,003	0,0079	0,005	0,0038
0,25	1	2	0,241	1,001	1,997	0,009	-0,001	0,003	0,0068	0,0037	0,0469
0,25	1	4	0,269	0,998	4,121	-0,019	0,002	-0,121	0,0087	0,0049	0,195
0,5	-1	0	0,5	-1,002	0,011	0	0,002	-0,011	0,0042	0,0033	0,002
0,5	-1	2	0,505	-1,009	2,027	-0,005	0,009	-0,027	0,0051	0,004	0,0164
0,5	-1	4	0,501	-1,009	4,044	-0,001	0,009	-0,044	0,0042	0,0054	0,0571
0,5	0	0	0,511	0,009	0,01	-0,011	-0,009	-0,01	0,0033	0,0024	0,0035
0,5	0	2	0,505	0,008	2,035	-0,005	-0,008	-0,035	0,0053	0,0026	0,0315
0,5	0	4	0,514	-0,005	4,051	-0,014	0,005	-0,051	0,006	0,003	0,0654
0,5	1	0	0,537	0,988	0,023	-0,037	0,012	-0,023	0,0108	0,0045	0,0066
0,5	1	2	0,511	1,005	2,008	-0,011	-0,005	-0,008	0,0109	0,0047	0,0639

Continua na próxima página.

Tabela 6.3 – continuação

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
0,5	1	4	0,528	0,998	4,062	-0,028	0,002	-0,062	0,0108	0,0046	0,1906
1	-1	0	0,999	-1,003	0,013	0,001	0,003	-0,013	0,0048	0,0038	0,003
1	-1	2	1,006	-1,002	2,026	-0,006	0,002	-0,026	0,0046	0,0029	0,0243
1	-1	4	1,018	-1,01	4,051	-0,018	0,01	-0,051	0,0069	0,0036	0,0744
1	0	0	1,028	-0,001	0,019	-0,028	0,001	-0,019	0,0065	0,0011	0,0042
1	0	2	1,031	0,001	2,043	-0,031	-0,001	-0,043	0,0078	0,0012	0,0318
1	0	4	1,016	0,001	4,065	-0,016	-0,001	-0,065	0,0068	0,0014	0,1125
1	1	0	1,044	0,983	0,039	-0,044	0,017	-0,039	0,0181	0,0079	0,0165
1	1	2	1,009	1,005	2,045	-0,009	-0,005	-0,045	0,0138	0,0067	0,1171
1	1	4	1,031	0,991	4,059	-0,031	0,009	-0,059	0,0125	0,0057	0,2744
-0,25	-1	0	-0,261	-0,994	0,012	0,011	-0,006	-0,012	9e-04	0,0041	0,0014
-0,25	-1	2	-0,255	-0,985	2,03	0,005	-0,015	-0,03	0,0011	0,0043	0,0166
-0,25	-1	4	-0,259	-1,007	4,037	0,009	0,007	-0,037	0,0012	0,0058	0,0458
-0,25	0	0	-0,257	0	0,012	0,007	0	-0,012	0,0018	0,0045	0,0108
-0,25	0	2	-0,257	0,014	2,03	0,007	-0,014	-0,03	0,0021	0,0052	0,027
-0,25	0	4	-0,258	0,014	4,023	0,008	-0,014	-0,023	0,0022	0,0067	0,1011
-0,25	1	0	-0,258	1,012	0,006	0,008	-0,012	-0,006	0,0044	0,0046	0,0035
-0,25	1	2	-0,268	0,997	1,974	0,018	0,003	0,026	0,0031	0,0038	0,026
-0,25	1	4	-0,256	1,007	4,003	0,006	-0,007	-0,003	0,0035	0,0042	0,0895
-0,5	-1	0	-0,516	-0,985	0,018	0,016	-0,015	-0,018	7e-04	0,0033	0,002
-0,5	-1	2	-0,516	-0,982	2,049	0,016	-0,018	-0,049	0,0012	0,0096	0,0346
-0,5	-1	4	-0,51	-1,004	4,096	0,01	0,004	-0,096	8e-04	0,0035	0,0767
-0,5	0	0	-0,514	0,011	-0,003	0,014	-0,011	0,003	0,0012	0,0033	0,003
-0,5	0	2	-0,51	0,007	2,025	0,01	-0,007	-0,025	0,0015	0,0042	0,0352
-0,5	0	4	-0,509	-0,001	4,009	0,009	0,001	-0,009	0,0012	0,0038	0,0823

Continua na próxima página.

Tabela 6.3 – continuação

ξ	μ	δ	$\hat{\xi}$	$\hat{\mu}$	$\hat{\delta}$	viés($\hat{\xi}$)	viés($\hat{\mu}$)	viés($\hat{\delta}$)	EQM ($\hat{\xi}$)	EQM ($\hat{\mu}$)	EQM ($\hat{\delta}$)
-0,5	1	0	-0,51	1,011	0,006	0,01	-0,011	-0,006	0,0031	0,0068	0,0031
-0,5	1	2	-0,513	0,999	1,99	0,013	0,001	0,01	0,0028	0,004	0,0283
-0,5	1	4	-0,511	1,007	3,995	0,011	-0,007	0,005	0,0033	0,0061	0,0712