



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

JOSÉ IRAPONIL COSTA LIMA

**Diagnóstico por influência no modelo de regressão Birnbaum-Saunders
valor extremo**

Recife

2022

JOSÉ IRAPONIL COSTA LIMA

**Diagnóstico por influência no modelo de regressão Birnbaum-Saunders
valor extremo**

Tese apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Estatística.

Área de Concentração: Estatística Aplicada.

Orientador: Dr. Raydonal Ospina Martínez

Coorientadora: Dr^a. Michelli K. Barros da Silva

Recife

2022

Catálogo na fonte
Bibliotecária Nataly Soares Leite Moro, CRB4-1722

L732d Lima, José Iraponil Costa
Diagnóstico por influência no modelo de regressão Birnbaum-Saunders
valor extremo / José Iraponil Costa Lima. – 2022.
209 f.: il., fig., tab.

Orientador: Raydonal Ospina Martínez.
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN,
Estatística, Recife, 2022.
Inclui referências e apêndices.

1. Estatística aplicada. 2. Birnbaum-Saunders. 3. Curvatura. 4. Precipitação
pluviométrica. I. Ospina Martínez, Raydonal (orientador). II. Título.

310 CDD (23. ed.) UFPE- CCEN 2022 - 46

JOSÉ IRAPONIL COSTA LIMA

Diagnóstico por influência no modelo de regressão Birnbaum-saunders valor extremo

Tese apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Estatística.

Aprovada em: 23 de fevereiro de 2022.

BANCA EXAMINADORA

PROF. DR. RAYDONAL OSPINA MARTÍNEZ
UFPE

PROF. DR. KLAUS LEITE PINTO VASCONCELLOS
UFPE

PROF^a. DR^a FRANCYELLE DE LIMA MEDINA
UFPE

PROF. DR. ANTÔNIO MURILO SANTOS MACEDO
UFPE

PROF^a DR^a. CAROLINA IVONNE MARCHANT FUENTES
UNIVERSIDAD CATÓLICA DEL MAULE

Dedico às minhas amáveis professoras e aos meus saudosos professores.

AGRADECIMENTOS

A Deus por me conceder um espírito e iluminar o caminho percorrido nesta jornada. E à Virgem Maria por me acompanhar, revelando-me a condição de servo do Senhor.

A todos os meus familiares. Em especial, à minha mãe Benedita Costa e ao meu pai Expedito Lima, pelo acolhimento e por me mostrarem que a educação e o trabalho dignificam as pessoas.

À minha esposa Andréia Bartolomeu Lima e às nossas filhas Júlia e Laura, pelo afeto, compreensão e equilíbrio.

À Prof^a Michelli Karinne B. da Silva e ao Prof^o Raydonal Ospina Martínez, pessoas pelas quais tenho uma profunda admiração e uma gratidão eterna, pelo acolhimento e direcionamentos.

Aos docentes que compuseram a banca de avaliação da tese pela disponibilidade na tarefa de avaliar nosso trabalho e pelas valiosas sugestões de como melhorá-lo. Às professoras Carolina Fluentes e Francielle Medina e aos professores Antônio Macedo e Klaus Vasconcellos.

Às professoras e aos professores que compõem o corpo docente do Departamento de Estatística da UFPE. Em particular, Betsabé Achic, Francisco Cribari, Francisco Cysneiros, Gauss Cordeiro, Patrícia Ospina e Roberto Manghi.

Aos meus colegas de trabalho da Unidade Acadêmica de Estatística da UFCG, pelo incentivo e colaboração.

Aos servidores técnicos administrativos da UAEst-UFCG e do Programa de Pós-Graduação do DE-UFPE.

Aos meus colegas discentes pela amizade e companheirismo. Em especial, César Diogo, Marcelo Santos e Maria Ioneris.

À UFCG por atender ao meu pedido de afastamento para capacitação.

À Fundação de Amparo à Ciência e Tecnologia de Pernambuco (FACEPE) pela concessão de suporte financeiro durante o desenvolvimento deste trabalho.

RESUMO

No desenvolvimento do presente trabalho derivamos ferramentas que possibilitam a implementação de uma análise de influência local sobre o modelo de regressão Birnbaum-Saunders valor extremo (EVBS), pois obtivemos expressões que possibilitam o cálculo das curvaturas normal ou conforme para três diferentes esquemas de perturbação. Paralelamente, expressamos uma condição necessária e suficiente para existência dos dois primeiros momentos da distribuição EVBS e deduzimos algumas características relativas à monotonicidade e à existência de máximos locais para a função densidade de probabilidade da distribuição log-EVBS. Além disso, abordamos o problema de estimação no modelo de regressão EVBS e acentuamos a problemática decorrente do fato de estarmos diante de um caso não-regular. Nesse contexto, obtivemos, por meio de simulações de Monte Carlo, resultados numéricos que corroboram com a hipótese de consistência assintótica do estimador de máxima verossimilhança (EMV) dos parâmetros do modelo de regressão EVBS. Em seguida, verificamos que esquemas de perturbações usualmente empregados sobre o modelo de regressão linear normal são formas apropriadas de perturbação sobre os modelos de regressão BS e EVBS. Prosseguimos, com a proposta de uma nova família de distribuições BS derivada da distribuição Pareto generalizada (GPD) para modelagem de dados extremos obtidos segundo a metodologia *Peaks over threshold* (POT). Além disso, deduzimos algumas propriedades desta distribuição e tratamos de estimação por máxima verossimilhança nesta nova família. Ao longo do texto encontram-se quatro aplicações em dados reais. Na primeira, investigamos os níveis de precipitação pluviométrica máxima anual registrados na estação meteorológica Recife-Curado, utilizando os modelos valor extremo generalizado (GEV) e o EVBS, combinados com o método dos máximos por bloco (BM), onde obtivemos estimativas de máxima verossimilhança para parâmetros de interesse em Estatística de Extremos, tais como níveis de retorno e períodos de retorno. Nas duas seguintes, interessou-nos a velocidade máxima mensal de ventos registrados na estação Itajaí-SC e as temperaturas máximas anuais na estação Recife-Curado, em ambas confirmamos a aplicabilidade da técnica de influência local em detectar observações influentes. Na última, modelamos os excessos anuais de precipitação pluviométrica, acima de um limiar elevado, registrados na estação Recife-Curado, usando a metodologia POT.

Palavras-chaves: Birnbaum-Saunders; curvatura; influência local; verossimilhança; valores extremos; precipitação pluviométrica.

ABSTRACT

In the development of the present work, we derived tools that allow the implementation of a local influence analysis on the extreme value Birnbaum-Saunders regression model (EVBS), because obtained expressions that allow the calculation of normal or conformal curvatures for three different perturbation schemes. At the same time, express a necessary and sufficient condition for the existence of the moments of the EVBS distribution and deduce some characteristics related to the monotonicity and the existence of local maxima for the probability density function of the log-EVBS distribution. Furthermore, approach the estimation problem in the EVBS regression model and emphasize the problem arising from the fact that we are dealing with a non-regular case. In this context, we obtained, through Monte Carlo simulations, numerical results that corroborate the hypothesis of asymptotic consistency of the maximum likelihood estimator of the parameters of the EVBS regression model. Next, verified that perturbation schemes usually employed on the normal linear regression model are appropriate forms of perturbation on the BS and EVBS regression models. In addition, we propose a new family of BS distributions derived from the generalized Pareto distribution (GPD) for modeling extreme values obtained according to the *Peaks Over Threshold* (POT) methodology. Next, deduce some properties from this distribution and deal with maximum likelihood estimation in this new family. Throughout the text there are four applications in real data. In the first one, investigated the levels of maximum annual rainfall recorded at the Recife-Curado meteorological station, using the generalized extreme value (GEV) and EVBS models, combined with the maximum per block method, where obtained maximum likelihood estimates for parameters of interest in Extreme Statistics, such as return levels and return periods. In the following two, we were interested in the maximum monthly wind speed recorded at the Itajaí-SC station and the maximum annual temperatures at the Recife-Curado station, in both we confirmed the applicability of the local influence technique in detecting influential observations. In the last one, we model the annual excess rainfall, above a high threshold, recorded at the Recife-Curado station, using the POT methodology.

Keywords: Birnbaum-Saunders; curvature; local influence; likelihood; extreme values; rainfall.

LISTA DE FIGURAS

Figura 1 – Gráficos da fdp da EVBS para $\alpha \in \{1, 0.75, 0.50, 0.25\}$, $\beta = 1$ e $\gamma = 0.625$ em (a) e para $\alpha = 0.25$, $\beta \in \{1, 0.90, 0.75, 0.50\}$ e $\gamma = -0.25$ em (b) . . .	43
Figura 2 – Gráficos da fdp da EVBS para $\alpha = 0.5$, $\beta = 1.5$ e $\gamma \in \{-1.25, -0.75, 0.75, 1.25\}$ em (a) e para $\alpha = 1$, $\beta \in \{1, 0.60, 0.30, 0.15\}$ e $\gamma = -1$ em (b)	44
Figura 3 – Gráfico de dispersão das precipitações pluviométricas diárias (mm) registradas na estação Recife-Curado (1962:2019) em (a) e o correspondente gráfico das precipitações máximas anuais em (b).	65
Figura 4 – Box-plot e histograma das precipitações pluviométricas máximas anuais registradas na estação Recife-Curado (1962:2019).	66
Figura 5 – Gráficos das funções de autocorrelação amostral, em (a), e autocorrelação parcial, em (b), para as precipitações pluviométricas máximas anuais na estação Recife-Curado (1962:2019).	67
Figura 6 – QQ-plot Gumbel em (a) e QQ-plot GEV em (b) das precipitações pluviométricas máximas anuais na estação Recife-Curado (1962:2019).	68
Figura 7 – Gráficos da qualidade do ajuste por meio do modelo GEV	71
Figura 8 – Gráficos da qualidade do ajuste por meio do modelo EVBS	74
Figura 9 – Gráficos da fdp da log-EVBS para $\alpha \in \{0.25, 0.50, 2, 4\}$, $\eta = 0$ e $\gamma = 0$ em (a) e para $\alpha = 0.5$, $\eta \in \{-4, -2, 2, 4\}$ e $\gamma \in \{-0.60, 0.60\}$ em (b). . .	77
Figura 10 – Gráficos da fdp da log-EVBS para $\alpha = 1$, $\eta = 0$ e $\gamma \in \{-1.05, -0.5, 0.5, 1.05\}$ em (a) e para $\alpha \in \{1, 2, 4, 5\}$, $\eta = 0$ e $\gamma = -1$ em (b).	78
Figura 11 – Gráfico de dispersão das velocidades máximas mensais do vento registradas na estação Itajaí-SC em (a) e o respectivo Box-plot em (b).	117
Figura 12 – Gráficos das funções de autocorrelação amostral em (a) e autocorrelação parcial em (b) das velocidades máximas mensais do vento registradas na estação Itajaí-SC.	117
Figura 13 – Gráfico normal de probabilidades, com envelope, para o resíduo quantílico em (a) e os resíduos quantílicos correspondentes às observações em (b). . .	119
Figura 14 – Autovalores normalizados em (a); Contribuição agregada dos vetores básicos de perturbação em (b), sob ponderação de casos	119

Figura 15 – Gráfico de dispersão em (a) e o Box-plot em (b) das temperaturas máximas anuais registradas na estação Recife-Curado (1962:2019).	138
Figura 16 – Gráficos das funções de autocorrelação amostral em (a) e autocorrelação parcial em (b) das temperaturas máximas anuais registradas na estação Recife-Curado (1962:2019).	139
Figura 17 – Gráfico normal de probabilidades, com envelope, para o resíduo quantílico em (a) e os resíduos quantílicos correspondentes às observações em (b). . .	140
Figura 18 – Autovalores normalizados em (a); Contribuição agregada dos vetores básicos de perturbação em (b), sob ponderação de casos.	141
Figura 19 – Autovalores normalizados em (a); Contribuição agregada dos vetores básicos de perturbação em (b), perturbação na resposta.	141
Figura 20 – Gráficos da fdp da GPD-BS para $\alpha \in \{0.25, 0.5, 1, 2\}$, $\beta = 1$ e $\gamma = 0$ em (a); para $\alpha = 1$, $\beta \in \{0.25, 0.5, 1, 2\}$ e $\gamma = 0.25$ em (b) e $\alpha = 1$, $\beta = 1$ e $\gamma \in \{-1.25, -0.75, 0.25, 1\}$ em (c).	147
Figura 21 – ME-plot das precipitações pluviométricas registradas na estação Recife-Curado (1962:2019).	154
Figura 22 – Estimativas dos parâmetros de escala e forma versus <i>threshold</i>	155
Figura 23 – Precipitações pluviométricas diárias registradas na estação Recife-Curado (1962:2019) em (a) e os respectivos excessos além do limiar $u = 96$ mm em (b).	156
Figura 24 – Histograma e Box-plot dos excessos, acima do limiar $u = 96$ mm, das precipitações pluviométricas na estação Recife-Curado (1962:2019).	157
Figura 25 – Gráficos das funções de autocorrelação amostral em (a) e parcial em (b) dos excessos, acima do limiar $u = 96$ mm, das precipitações pluviométricas na estação Recife-Curado (1962:2019).	157
Figura 26 – QQ-plot exponencial dos excessos, acima do limiar $u = 96$ mm, das precipitações pluviométricas na estação Recife-Curado (1962:2019).	158
Figura 27 – Gráficos da qualidade do ajuste por meio do modelo GPD.	161
Figura 28 – Gráficos da qualidade do ajuste por meio do modelo GPD-BS.	163

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas para os dados de precipitações pluviométricas máximas anuais (mm) na estação Recife-Curado (1962:2019).	66
Tabela 2 – Estimativas de MV dos parâmetros μ , σ e γ no modelo GEV, acompanhadas do respectivo erro padrão (erro padrão (se)).	68
Tabela 3 – Estimativas de MV dos quantis elevados no modelo GEV para a precipitação pluviométrica máxima anual na estação Recife-Curado.	69
Tabela 4 – Estimativas de MV dos níveis de retorno no modelo GEV para a precipitação pluviométrica máxima anual na estação Recife-Curado.	69
Tabela 5 – Estimativas pontuais da probabilidade de excedência do nível u e do período de retorno, em anos, associado ao nível u no modelo GEV.	70
Tabela 6 – Estimativas de MV dos parâmetros α , β e γ do modelo EVBS, acompanhadas do respectivo erro padrão.	72
Tabela 7 – Estimativas de MV dos quantis elevados no modelo EVBS para a precipitação pluviométrica máxima anual na estação Recife-Curado.	72
Tabela 8 – Estimativas de MV dos níveis de retorno no modelo EVBS para a precipitação pluviométrica máxima anual na estação Recife-Curado.	72
Tabela 9 – Estimativas de MV da probabilidade de excedência do nível u e do período de retorno, em anos, associado ao nível u no modelo EVBS.	73
Tabela 10 – Aproximações empíricas para a estimativa de MV dos parâmetros e o viés do EMV, sob o cenário 1.	83
Tabela 11 – Aproximações empíricas para o erro padrão e a raiz quadrada do erro quadrático médio do EMV, sob o cenário 1.	84
Tabela 12 – Aproximações empíricas para o erro padrão, oriundo da matriz de informação de Fisher observada (IFO), e a taxa de cobertura do intervalo de confiança, sob o cenário 1.	84
Tabela 13 – Aproximações empíricas para a estimativa de MV dos parâmetros e o viés absoluto do EMV, sob o cenário 2.	85
Tabela 14 – Aproximações empíricas para o erro padrão e o raiz quadrada do erro quadrático médio do EMV, sob o cenário 2.	86

Tabela 15 – Aproximações empíricas para o erro padrão, oriundo da inversa da matriz de informação de Fisher observada (IFO), e a taxa de cobertura do intervalo de confiança, sob o cenário 2.	87
Tabela 16 – Aproximações empíricas para a estimativa de MV dos parâmetros e o viés absoluto do EMV, sob o cenário 3A.	88
Tabela 17 – Aproximações empíricas para o erro padrão e a raiz quadrada erro quadrático médio do EMV, sob o cenário 3A.	89
Tabela 18 – Aproximações empíricas para o erro padrão, oriundo da matriz de informação de Fisher observada (IFO), e taxa de cobertura do intervalo de confiança, sob o cenário 3A.	89
Tabela 19 – Aproximações empíricas para a estimativa de MV dos parâmetros e o viés absoluto do EMV, sob o cenário 3B.	90
Tabela 20 – Aproximações empíricas para o erro padrão e a raiz quadrada erro quadrático médio do EMV, sob o cenário 3B.	90
Tabela 21 – Aproximações empíricas para o erro padrão, oriundo da matriz de informação de Fisher observada (IFO), e taxa de cobertura do intervalo de confiança, sob o cenário 3B.	91
Tabela 22 – Estatísticas descritivas para velocidade máxima mensal do vento, m/s, registrada pela Estação A-868 em Itajaí-SC-Brasil.	116
Tabela 23 – Estimativa de MV dos parâmetros do modelo acompanhada dos respectivos erro padrão e nível descritivo dos coeficientes da regressão.	118
Tabela 24 – Estimativa de MV dos parâmetros do modelo, taxa de variação, em porcentagem, e nível descritivo associado aos coeficientes da regressão.	120
Tabela 25 – Estatísticas descritivas para temperatura máxima anual (°C) na estação Recife-Curado (1962:2019).	138
Tabela 26 – Estimativa de MV dos coeficientes da regressão acompanhada do seu respectivo erro padrão e nível descritivo.	139
Tabela 27 – Estimativa de MV, taxa de variação, em porcentagem, e nível descritivo associado aos coeficientes da regressão.	142
Tabela 28 – Estatísticas descritivas para os excessos de precipitação pluviométrica (mm) acima do limiar $u = 96$ mm, na estação Recife-Curado (1962:2019).	158
Tabela 29 – Estimativas de MV dos parâmetros σ e γ do modelo GPD para o excesso acima do limiar $u = 96$, acompanhadas do respectivo erro padrão.	159

Tabela 30 – Estimativas de MV dos quantis elevados no modelo GPD, excedentes acima do limiar $u = 96$	159
Tabela 31 – Estimativas de MV dos níveis de retorno, excedentes, no modelo GPD. . .	159
Tabela 32 – Estimativas de MV da probabilidade de excedência do nível e e do período de retorno correspondente, em anos, no modelo GPD.	160
Tabela 33 – Estimativas de MV dos parâmetros α e γ no modelo GPD-BS, excedentes acima do limiar $u = 96$, acompanhadas do respectivo erro padrão.	161
Tabela 34 – Estimativas de MV dos quantis elevados no modelo GPD-BS, excedentes acima do limiar $u = 96$	162
Tabela 35 – Estimativas de MV dos níveis de retorno no modelo GPD-BS.	162
Tabela 36 – Estimativas de MV da probabilidade de excedência do nível e e do período de retorno correspondente, em anos, no modelo GPD-BS.	162

LISTA DE ABREVIATURAS E SIGLAS

BM	método dos máximos por bloco
EMQ	estimador de mínimos quadrados
emv	estimativa de máxima verossimilhança
EMV	estimador de máxima verossimilhança
EQM	erro quadrático médio
EVBS	Birnbaum-Saunders valor extremo
EVI	índice de valor extremo
fd_a	função de distribuição acumulada
fd_p	função densidade de probabilidade
GEV	valor extremo generalizado
GPD	distribuição Pareto generalizada
GPD-BS	Birnbaum-Saunders-Pareto generalizada
iid	independentes e identicamente distribuídas
INMET	Instituto Nacional de Meteorologia
MDA	máximo domínio de atração
MV	máxima verossimilhança
POT	<i>Peaks over threshold</i>
se	erro padrão

SUMÁRIO

1	INTRODUÇÃO	17
1.1	MOTIVAÇÕES E JUSTIFICATIVAS	17
1.2	OBJETIVOS	17
1.3	RESULTADOS ALCANÇADOS	18
1.4	ORGANIZAÇÃO DO TRABALHO	19
1.5	REVISÃO BIBLIOGRÁFICA	20
1.6	SUORTE COMPUTACIONAL	27
2	O MODELO BIRNBAUM-SAUNDERS E SUA VERSÃO VALOR EXTREMO.	28
2.1	INTRODUÇÃO	28
2.2	A DISTRIBUIÇÃO BIRNBAUM-SAUNDERS	28
2.2.1	Origem, pressupostos e derivação	28
2.2.2	Definição e algumas propriedades básicas	31
2.2.3	A função de sobrevivência e a função de risco	32
2.2.4	Média, variância, assimetria e curtose	33
2.2.5	Estimação por máxima verossimilhança	34
2.2.6	Uma relação entre as distribuições Birnbaum-Saunders e senh-normal	37
2.2.7	O modelo de regressão log-Birnbaum-Saunders	38
2.3	A DISTRIBUIÇÃO BIRNBAUM-SAUNDERS VALOR EXTREMO	41
2.3.1	Origem, definição e propriedades	41
2.3.2	A função quantílica e a geração de números aleatórios	44
2.3.3	Os momentos	46
2.3.4	Domínios de atração	48
2.4	ESTIMAÇÃO DO ÍNDICE DE VALORES EXTREMOS	53
2.4.1	Existência e consistência do EMV para o índice de valores extremos	54
2.4.2	Sobre a normalidade assintótica do EMV para o índice de valores extremos	58
2.5	ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA NA FAMÍLIA EVBS	62
2.6	APLICAÇÃO	63
2.6.1	Análise exploratória dos dados	64

2.6.2	Estimativas de máxima verossimilhança sob o modelo GEV	67
2.6.3	Estimativas de máxima verossimilhança sob o modelo EVBS	71
2.7	UM MODELO DE REGRESSÃO PARA DISTRIBUIÇÃO EVBS	74
2.7.1	Caracterização	75
2.7.2	Propriedades distribuição log-EVBS	76
2.7.3	Estimação no modelo de regressão EVBS	78
2.7.4	Simulações	82
2.7.4.1	<i>Cenário 1</i>	83
2.7.4.2	<i>Cenário 2</i>	85
2.7.4.3	<i>Cenário 3</i>	87
3	AVALIAÇÃO DE INFLUÊNCIA LOCAL	92
3.1	INTRODUÇÃO	92
3.2	A FUNÇÃO AFASTAMENTO PELA VEROSSIMILHANÇA E O GRÁFICO DE INFLUÊNCIA	93
3.3	AVALIAÇÃO DE INFLUÊNCIA LOCAL	94
3.3.1	Influência individual	94
3.3.2	Influência conjunta	95
3.4	RELAÇÃO ENTRE MEDIDAS DE INFLUÊNCIA	96
3.4.1	Influência individual e influência conjunta.	96
3.4.2	Medidas de influência individual e conjunta	98
3.4.3	Autovalores competitivos	99
3.5	A CURVATURA NORMAL CONFORME	100
3.5.1	Reparametrização e invariância	100
3.5.2	Invariância da curvatura normal conforme	101
3.5.3	Limitação e construção de valores de referência	102
3.5.4	Contribuição agregada de vetores básicos de perturbação	104
3.6	INFLUÊNCIA LOCAL NO MODELO DE REGRESSÃO EVBS	105
3.6.1	Ponderação de casos	105
3.6.2	Perturbação na resposta	108
3.6.3	Perturbação em uma covariável	111
3.7	APLICAÇÃO	116
4	PERTURBAÇÕES APROPRIADAS SOBRE O MODELO DE REGRESSÃO EVBS	122

4.1	INTRODUÇÃO	122
4.2	VARIEDADE ESTATÍSTICA DE PERTURBAÇÃO	123
4.3	MEDIDAS DE INFLUÊNCIAS E SUAS PROPRIEDADES	126
4.4	FORMAS APROPRIADAS DE PERTURBAÇÃO.	129
4.4.1	Modelo de regressão normal linear.	129
4.4.2	Modelo de regressão log-Birnbaum-Saunders	131
4.4.3	Modelo de regressão EVBS	134
4.4.4	Aplicação.	137
5	O MODELO BIRNBAUM-SAUNDERS-PARETO GENERALIZADO	144
5.1	INTRODUÇÃO	144
5.2	A DISTRIBUIÇÃO GPD-BS	145
5.2.1	Definição e algumas propriedades	145
5.3	ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA	151
5.4	APLICAÇÃO	153
5.4.1	Introdução	153
5.4.2	Análise exploratória dos dados	154
5.4.3	Estimação por meio do modelo GPD	157
5.4.4	Estimação via o modelo GPD-BS	160
6	CONCLUSÕES E TRABALHOS FUTUROS	164
	REFERÊNCIAS	166
	APÊNDICE A – A ESTATÍSTICA DE EXTREMOS UNIVARIADOS	174
	APÊNDICE B – GEOMETRIA EM SUPERFÍCIES QUE SÃO GRÁFICOS DE FUNÇÕES DIFERENCIÁVEIS	187
	APÊNDICE C – MÉTODOS DE GEOMETRIA DA INFORMAÇÃO	195
	APÊNDICE D – PROVAS	204

1 INTRODUÇÃO

1.1 MOTIVAÇÕES E JUSTIFICATIVAS

A implementação de ações de diagnóstico baseadas em avaliação de influência local é uma etapa relevante em uma análise estatística, pois consiste em um procedimento de controle que possibilita a identificação criteriosa de observações que, quando sujeitas a pequenas perturbações, causam variação desproporcional nas estimativas dos parâmetros fornecidas pelo modelo postulado, podendo alterar significativamente os resultados inferenciais. A execução dessas ações se torna ainda mais necessária quando o modelo postulado rege o comportamento de variáveis aleatórias associadas a eventos extremos. A Estatística de Extremos revela-se útil no estudo de fenômenos aleatórios em que o interesse é a ocorrência de eventos com probabilidade pequena, usualmente chamados por eventos extremos ou raros, mas que são potencialmente desastrosos, suas aplicações se estendem por áreas tais como economia (finanças e seguros), meteorologia (climatologia e hidrologia) e geofísica (sismologia). Logo, dada a potencialidade dos impactos socioeconômicos de eventos extremos tais como chuvas intensas e quedas bruscas em índices de bolsas de valores, é prudente verificar se as estimativas fornecidas pelo o modelo postulado apresentam certa estabilidade quando perturbações são impostas nos dados ou sobre o próprio modelo.

O modelo de regressão Birnbaum-Saunders valor extremo (EVBS), definido por (LEIVA et al., 2016), é o alvo do estudo de diagnóstico que desenvolvemos em nosso trabalho. Aqui, conduzimos uma análise de influência local sobre esse modelo, segundo a metodologia proposta por (COOK, 1986), agregando os aperfeiçoamentos obtidos por (POON; POON, 1999) e (ZHU et al., 2007).

1.2 OBJETIVOS

1. Apresentar algumas características da distribuição que regula o comportamento da componente estocástica do modelo de regressão EVBS;
2. Implementar um estudo de simulações de Monte Carlo com a finalidade de obter indícios de consistência do EMV dos parâmetros do modelo de regressão EVBS;
3. Estimar os parâmetros associados à precipitação pluviométrica máxima anual, total em

um dia, registrada na estação Recife-Curado, utilizando o método dos máximos por bloco (BM) ou método de Gumbel;

4. Desenvolver análise de influência local sobre o modelo de regressão EVBS, obtendo expressões que possibilitem o cálculo de curvaturas para diferentes esquemas de perturbação;
5. Verificar se esquemas usuais de perturbações são formas apropriadas de perturbação sobre o modelo de regressão EVBS;
6. Propor uma nova família de distribuições BS que deriva da distribuição Pareto generalizada (GPD);
7. Modelar os excessos de precipitação pluviométrica anual, acima de um limiar alto, registrados na estação Recife-Curado, usando a metodologia *Peaks Over Threshold* (POT), a fim de obter estimativas pontuais para o período de retorno de elevados níveis pluviométricos.

1.3 RESULTADOS ALCANÇADOS

1. Dedução de algumas propriedades, no que tange à monotonicidade e à existência de máximos locais, da função densidade de probabilidade (fdp) da variável aleatória que rege o comportamento da componente estocástica do modelo de regressão EVBS;
2. Derivação de uma condição necessária e suficiente para existência dos momentos centrais da distribuição EVBS;
3. Modelagem das precipitações pluviométricas máximas anuais, total em um dia, registradas na estação meteorológica Recife-Curado por meio dos modelos GEV e EVBS, usando o método dos máximos por bloco (BM) ou método de Gumbel;
4. Implementação de um estudo de simulação de Monte Carlo com o propósito de verificar a existência de resultados numéricos que corroborem com a hipótese de consistência do EMV dos parâmetros do modelo de regressão EVBS;
5. Determinação de expressões que possibilitam o cálculo das curvaturas normal e conforme para três diferentes esquemas de perturbações impostos sobre o modelo de regressão

EVBS, combinada com uma aplicação em dados reais na qual a variável resposta foi a velocidade máxima mensal dos ventos registrados na estação meteorológica Itajaí-Santa Catarina;

6. Verificação de que os esquemas de perturbações usualmente empregados sobre o modelo de regressão linear normal são apropriados tanto para o modelo de regressão BS quanto para o EVBS (caso regular). Agrega-se a isso, uma aplicação em dados meteorológicos na qual a resposta foi temperatura máxima anual registrada na estação Recife-Curado;
7. Proposta de uma nova distribuição de probabilidade, chamada Birnbaum-Saunders-Pareto generalizada (GPD-BS), na qual se usa a distribuição Pareto generalizada (GPD), em vez da normal padrão, para compor o núcleo da transformação que determina a distribuição BS;
8. Modelagem dos excessos anuais de precipitação pluviométrica, acima de um limiar elevado, registrados na estação meteorológica Recife-Curado por meio dos modelos GPD e GPD-BS, usando a metodologia POT.

1.4 ORGANIZAÇÃO DO TRABALHO

Este trabalho encontra-se estruturado em seis capítulos. No Capítulo 1, que é o presente, apresentamos as motivações e justificativas para o desenvolvimento da pesquisa, expomos os objetivos e os resultados alcançados e concluímos com uma ampla revisão bibliográfica sobre as distribuições BS e EVBS. No Capítulo 2, nos reportamos às distribuições EVBS e log-EVBS, onde apresentamos as definições e deduzimos algumas propriedades das distribuições no que tange à existência de momentos, à monotonicidade e à ocorrência de máximos e mínimos locais da fdp da variável log-EVBS, em termos dos parâmetros. Em seguida, discutiremos acerca do modelo de regressão EVBS, onde abordamos o problema de estimação conjunta dos parâmetros por máxima verossimilhança (MV) e derivamos expressões para a função score e a matriz Hessiana da log-verossimilhança, equivalentes às obtidas por (LEIVA et al., 2016). Exibimos uma aplicação onde modelamos as precipitações pluviométricas máximas anuais registradas na estação meteorológica Recife-Curado por meio das distribuições GEV e EVBS, usando o método BM. Concluímos o capítulo com a realização de simulações de Monte Carlo com o propósito de verificar a existência de evidências numéricas que corrobore

com a hipótese de consistência e normalidade assintótica do EMV dos parâmetros do modelo de regressão EVBS. No Capítulo 3, tratamos de diagnóstico por influência local sobre o modelo de regressão EVBS, segundo a metodologia desenvolvida por (COOK, 1986), onde consideramos três diferentes esquemas de perturbação. Além disso, analisamos um conjunto de dados reais em que o interesse recai sobre a velocidade máxima de ventos registrados na estação meteorológica localizada em Itajaí-Santa Catarina, com a finalidade de ilustrar a aplicabilidade da metodologia de influência local na identificação criteriosa de observações influentes. No Capítulo 4, mostramos que os esquemas de perturbações reportados no Capítulo 3 são formas apropriadas de perturbação sobre o modelo de regressão EVBS (caso regular), segundo os critérios estabelecidos por (ZHU et al., 2007). E a título de ilustração, analisamos um conjunto de dados de temperaturas máximas anuais registrados na estação Recife-Curado. No Capítulo 5, definimos a família de distribuições Birnbaum-Saunders-Pareto generalizada (GPD-BS) e deduzimos algumas propriedades que exprimem características do comportamento dessa distribuição. Em seguida, na Seção 5.3, lidamos com estimação por MV na família GPD-BS. E para concluirmos, na Seção 5.4, desenvolvemos uma aplicação em que se investigou os excessos de precipitações pluviométricas, acima de um limiar elevado, registrados na estação Recife-Curado, usando os modelos GPD e GPD-BS. Por fim, no Capítulo 6, expressamos as nossas conclusões e indicaremos alguns pontos que propiciam o desenvolvimento de pesquisas futuras.

1.5 REVISÃO BIBLIOGRÁFICA

A distribuição Birnbaum-Saunders (BS) teve origem em contextos diferentes e possui derivações distintas, conforme exprimem (BALAKRISHNAN; KUNDU, 2019), destacando-se a que foi proposta por (BIRNBAUM; SAUNDERS, 1969a) na qual ela é derivada de um fenômeno físico em que a variável de interesse é o tempo até a ocorrência de falha devido a fadiga em materiais sob cargas cíclicas e contínuas. A produção científica relativa à distribuição é rica e volumosa, estendendo-se por várias áreas da Estatística, pois são muitos e frutíferos os artigos relacionados diretamente com o modelo e que tratam de temáticas tais como: propriedades estatísticas e matemáticas do modelo; estimação pontual e por intervalo, com abordagens clássicas e bayesianas; testes de hipóteses; modelos de regressão log-linear e não-linear; desempenho de geradores de números aleatórios; desenvolvimento de pacotes para o *software* R; surgimento de generalizações, versões multivariadas e extensões tanto da distribuição BS quanto do seu

correspondente modelo de regressão.

Os autores (BALAKRISHNAN; KUNDU, 2019) apresentam três diferentes contextos para o surgimento da distribuição BS. Primeiro, relacionam a origem do modelo com o estudo do movimento Browniano ao citarem os trabalhos de (FLETCHER, 1911), (SHRÖEDINGER, 1915) e (KONSTANTINOWSKY, 1914). Em seguida, reportam-se a (FREUDENTHAL; M., 1961) e (BIRNBAUM; SAUNDERS, 1969a) para relacionar o surgimento da distribuição com outro fenômeno físico, a saber, a modelagem do tempo até a ocorrência de falha, devido à fadiga, em materiais sujeitos a cargas cíclicas. Por fim, citam (DESMOND, 1985) para indicar que a distribuição BS pode ser derivada num contexto mais geral, baseado em um modelo biológico. Além disso, ao citarem os trabalhos de (BHATTACHARYYA; FRIES, 1982) e (DESMOND, 1986) os autores informam que o modelo BS também pode ser derivado teoricamente da distribuição gaussiana inversa.

Do ponto de vista estatístico, o modelo BS é detentor de vantajosas propriedades. De início enumeramos os resultados obtidos por (BIRNBAUM; SAUNDERS, 1969a) : a distribuição BS pode ser obtida por meio de uma transformação monótona sobre a variável normal padrão; a mesma possui momentos finitos, condição necessária para a aplicação de vários resultados limites como, por exemplo, a Lei Forte dos Grandes Números e o Teorema Central do Limite; o parâmetro de escala da distribuição também cumpre o papel de parâmetro de locação, pois é a mediana; a família de distribuições BS é fechada por reciprocidade e proporcionalidade. Foram (BIRNBAUM; SAUNDERS, 1969a) que deram os primeiros passos na caracterização da função de risco, eles citam evidências empíricas de que tal função decresce lentamente a partir de certo valor real, quando ambos os parâmetros são iguais a 1. Os trabalhos de (KUNDU; KANNAN; BALAKRISHNAN, 2008) e (BEBBINGTON; C-D; ZITIKIS, 2008) também investigaram o comportamento da função de risco. No primeiro deles há uma justificativa, recorrendo a um resultado presente em (GLASER, 1980), de que a função de risco da BS cresce atinge um ponto de máximo e em seguida decresce monotonicamente; combinado a isso há uma condição, na forma de equação não-linear, que o ponto onde a função atinge o máximo deve cumprir e uma aproximação para tal valor. No segundo, há uma prova similar para a não-monotonicidade da função de risco da BS, porém não há indicação explícita do ponto onde tal função atinge o máximo nem sugestão de um valor aproximado. Por fim, ressaltamos o trabalho conjunto de (ATHAYDE et al., 2019) que reacende a temática no contexto de extensões da BS.

Em relação à estimação dos parâmetros, α e β , do modelo BS, (BIRNBAUM; SAUNDERS, 1969b) foram os primeiros a lidarem com a problemática. Eles trataram da estimação pontual

dos parâmetros via o método de máxima verossimilhança. No trabalho, constam as funções e as equações que levam à determinação dos estimadores, acompanhadas de uma argumentação da unicidade do EMV para β , denotado por $\hat{\beta}$, e uma fórmula que permite determinar a estimativa de α função da estimativa de β . Aliado a isso, há dois procedimentos numéricos que permitem o cálculo de $\hat{\beta}$, um deles baseado no método de Newton-Raphson e o outro em um método de aproximação sucessivas, ambos convergindo para $\hat{\beta}$. Figura naquele trabalho uma proposta de estimador para β , denotado por $\tilde{\beta}$, definido como sendo a média geométrica da média aritmética amostral e média harmônica dos dados e o mesmo é consistente para o β , quando α é pequeno. E sugerem que a estimativa dada por $\tilde{\beta}$ pode ser utilizada como chute inicial nos procedimentos iterativos para a obtenção de $\hat{\beta}$. No mais, são apresentados exemplos numéricos com o propósito de comparar a performance dos métodos iterativos propostos e as estimativas geradas por $\hat{\beta}$ e $\tilde{\beta}$.

Ainda no contexto de estimação dos parâmetros da BS, (ENGELHARDT; BAIN; WRIGHT, 1981) constroem intervalos de confiança para os parâmetros α e β , recorrendo à normalidade assintótica do estimador de máxima verossimilhança e diante do fato de que a família de distribuições BS satisfaz as condições de regularidades, fixadas, por exemplo, em (LEHMANN; CASELLA, 1998). (RIECK, 1995) é o primeiro a lidar com a estimação no caso de amostras censuradas; (DUPUIS; MILLS, 1998) tratam da estimação robusta dos parâmetros e quantis da BS; (NG; KUNDU; BALAKRISHNAN, 2006) lidam com estimação pontual e por intervalo, no caso de amostras sob censura tipo-II; (NG; KUNDU; BALAKRISHNAN, 2003) apresentam os estimadores de momentos modificados (EMM) para os dois parâmetros da BS e, posteriormente, (WANG; DESMOND; LU, 2006) desenvolvem os estimadores de momentos modificados no caso de observações censuradas; (FROM; LI, 2006) apresentam um leque de quatro novos estimadores para os parâmetros da BS. As primeiras abordagens que tratam de correção de viés dos estimadores de máxima verossimilhança são creditadas a (LEMONTE; CRIBARI-NETO; VASCONCELOS, 2007). Dentre as produções que aplicam métodos bayesianos nos procedimentos de estimação, temos (PADGETT, 1982), pioneiro na aplicação de tais métodos na estimação do parâmetro β da BS, mas sob a hipótese de que α é conhecido, com o uso de uma distribuição a priori não informativa; (ACHCAR, 1993), que trata da estimação, admitindo que ambos os parâmetros são desconhecidos, (ACHCAR; MOALA, 2010) retomam o mesmo problema, mas no caso de existirem observações censuradas na amostra; (WANG; SUN; PARK, 2016) realizam análise bayesiana via o método generalizado da razão de uniformes. Por fim, destaquemos (LILLO et al., 2018) que obtém os estimadores de L-Momentos dos parâmetros da distribuição

BS e de uma versão valor extremo da BS.

O fato teórico de que uma variável com distribuição BS está relacionada, por meio de uma transformação monótona, com a normal padrão instigou pesquisadores a desenvolverem generalizações, versões multivariadas e extensões da BS. A ideia chave é substituir a variável normal padrão por uma variável de interesse, de modo que a transformação resulte em outra distribuição com boas propriedades estatísticas. Dentre os trabalhos que estão nesta direção, destacamos aquele devido a (DÍAZ-GARCÍA; LEIVA, 2005) em que se faz a substituição do núcleo normal padrão por variáveis pertencentes a classe de distribuições de contornos elípticos, originando a distribuição Birnbaum-Saunders generalizada (BSG), que possui a BS com um caso particular; em (KUNDU; BALAKRISHNAN; JAMALIZADEH, 2010; KUNDU; BALAKRISHNAN; JAMALIZADEH, 2013) constam as versões bivariada e multivariada da BS, originadas mediante a substituição natural da variável normal padrão por suas respectivas versões bivariada e multivariada. Outras extensões da distribuição BS podem ser encontradas, por exemplo, nas produções de (OWEN, 2006), (VILCA; LEIVA, 2006), (GÓMEZ; OLIVARES; BOLFARINE, 2009), (GUIRAUD; LEIVA; FIERRO, 2009), (LEIVA; SANHUEZA; ANGULO, 2009), (VILCA et al., 2010), (CASTILLO; GÓMEZ; BOLFARINE, 2011), (FERREIRA; GOMES; LEIVA, 2012), (DESOUSA et al., 2017) e (BOURGUINGON et al., 2017).

No que tange à geração de números aleatórios, o artigo de (CHANG; TANG, 1994) apresenta a construção de dois geradores: um deles baseado na distribuição normal padrão, mas com uma inconveniência na forma de raízes múltiplas, que é resolvida por meio de um critério de decisão fundamentado na uniforme unitária; o outro se baseia em uma relação existente entre a distribuição BS e inversa gaussiana. Na mesma linha, (RIECK, 2003) compara um gerador baseado na relação entre a distribuição BS e a *senh-normal* com o primeiro dentre aqueles propostos por (CHANG; TANG, 1994) e verifica que o mesmo é um caso especial deste último, porém de mais simples execução. A geração de números aleatórios da BSG é tratada em (LEIVA et al., 2008). Sobre a disponibilidade de pacotes implementados no *software R* relacionados com o modelo BS e BSG temos as produções de (LEIVA; HERNÁNDEZ; RIQUELME, 2006) e (BARROS; PAULA; LEIVA, 2009) responsáveis pelos desenvolvimentos dos pacotes *bs* e *gbs*, respectivamente. Aqui merece destaque a enorme abrangência do pacote *gbs* cujas funções implementadas possibilitam lidar com geração de números aleatórios, estimação de parâmetros a partir de dados com ou sem censura, testes de bondade de ajuste e ferramentas de diagnóstico.

Os autores (RIECK; NEDELMAN, 1991) inauguraram uma próspera linha de pesquisa ci-

entífica ao desenvolver um modelo de regressão para a distribuição BS cuja componente estocástica possui distribuição log-BS, um caso particular da senh-normal, e o denominaram de modelo regressão log-linear Birnbaum-Saunders (MRL-log-BS). De fato, as produções científicas subsequentes percorrem ao menos duas variantes do modelo originalmente proposto. Essas variantes têm origem ora a partir de uma mudança na componente estocástica do modelo ora por meio de uma alteração na componente determinística. Antes de apresentarmos as principais referências bibliográficas que tratam das variantes, temos algumas produções relativas ao modelo original: (OWEN; PADGETT, 2000) derivam um modelo de tempo de vida acelerado para BS; (TSIOMAS, 2001) trata de inferência sobre o MRL-log-BS usando procedimentos bayesianos; (GALEA; LEIVA; A, 2004) desenvolvem procedimentos de diagnóstico sobre o modelo, com aplicação da técnica de influência local formulada por (COOK, 1986); (XIE; WEI, 2007) obtêm medidas de diagnósticos, baseando-se em exclusão de casos, e apresentam um teste de homogeneidade sobre o parâmetro α do modelo; (LEIVA; BARROS M. GILBERTO; GALEA, 2007) tratam de diagnóstico por influência no MRL-log-BS, sob a presença de observações censuradas na amostra; (DESMOND; RODRIGUEZ-YAM; LU, 2008) fazem uso do algoritmo EM (*Expectation-Maximization*) para tratarem de estimação, na presença de dados censurados; (LI; CHEN; F-C., 2012) desenvolvem um modelo de regressão heterocedástico para BS em que tanto o parâmetro de locação quanto o de forma dependem de covariáveis; (LEMONTE; FERRARI, 2011) derivam três estatísticas segundo o teste da razão de verossimilhanças sinalizada para o modelo; (LEMONTE, 2011) deriva uma fórmula para matriz de covariâncias dos estimadores de máxima verossimilhança; (BALAKRISHNAN; ZHU, 2015a) tratam de estimação e testes de hipóteses sobre o parâmetros do MRL-log-BS; (TSUYUGUCHI; PAULA; BARROS, 2019) tratam de inferência e diagnóstico sobre modelos de regressão Birnbaum-Saunders na hipótese de dados correlacionados, usando equações de estimação.

Agora descrevemos algumas das variantes do modelo formulado por (RIECK; NEDELMAN, 1991). A primeira delas, que é caracterizada por uma mudança na hipótese sobre a distribuição da componente aleatória do modelo original, está associada a uma vasta produção acadêmica, temos, por exemplos, os trabalhos de: (BARROS; PAULA; LEIVA, 2008) que desenvolvem um novo modelo de regressão (MRL-log-BS-t) ao assumirem a distribuição *t-Student* em vez da normal padrão para modelar o dano acumulativo; (CANCHO; ORTEGA; PAULA, 2010) tratam de estimação e diagnóstico sobre o MRL-log-BS-t por meio de procedimentos bayesianos; (PAULA et al., 2012) lidam com modelagem robusta e diagnóstico por influência sobre o MRL-log-BS-t; (LI; F-C., 2012) tratam de diagnóstico no modelo log-BS-t; (BALAKRISHNAN; ZHU, 2015b)

lidam com inferência sobre o modelo de regressão baseado na distribuição BS bivariada. (MARCHANT; LEIVA; CYSNEIROS, 2015) derivam uma distribuição BS generalizada multivariada e o seu correspondente modelo de regressão, incorporando o método de máxima verossimilhança e o algoritmo EM na estimação dos parâmetros. Em seguida, (MARCHANT; LEIVA; CYSNEIROS, 2016) desenvolvem análise de diagnóstico sobre o modelo de regressão linear log-BSG multivariado.

Atrelada a segunda variante, que tem como característica uma alteração na componente determinística do modelo original, destaquemos os trabalhos de: (LEMONTE; CORDEIRO, 2009) em que se formula um modelo de regressão não-linear para a BS (MRN-log-BS); (LEMONTE; CORDEIRO, 2010) obtêm uma fórmula assintótica para a assimetria dos estimadores de MV dos parâmetros do modelo; (LEMONTE; PATRIOTA, 2011) tratam de influência local sobre o MRN-log-BS; (FARIAS; LEMONTE, 2011) lidam com inferência bayesiana na estimação dos parâmetros do modelo; (LEMONTE; CORDEIRO; MORENO, 2012) tratam de correções de Bartlett sobre o MRN-log-BS.

Sobre outras versões do modelo original observemos que (VILLEGAS; PAULA; LEIVA, 2011) desenvolvem um modelo misto para BS; (LEMONTE, 2012) propõe um modelo de regressão baseado na distribuição *senh-normal skew*; (LEIVA et al., 2014) apresentam o modelo de regressão log-linear BS reparametrizado, baseando-se na reparametrização proposta por (SANTOS-NETO et al., 2012); (SANTANNA; VILCA; LEIVA, 2011) tratam de estimação, via o algoritmo EM, e análise de influência sobre os parâmetros do modelo de regressão *skew-Birnbaum-Saunders*; (LEMONTE, 2013b) apresenta uma extensão do modelo log-BS; (LEMONTE, 2013a) desenvolve uma versão do modelo log-BS na qual a resposta é um vetor multivariado; (ORTEGA; CORDEIRO; LEMONTE, 2013) propõem um novo modelo de regressão baseado na distribuição β -Birnbaum-Saunders; (GARCIA-PAPANI et al., 2017) derivam um modelo log-linear espacial para distribuição BS; (LEIVA et al., 2016) desenvolvem um modelo de regressão baseado-se na distribuição BS Valor extremo, definida por (FERREIRA; GOMES; LEIVA, 2012); (MARTÍNEZ-FLÓREZ; BOLFARINE; GÓMEZ, 2017) formulam um modelo potência log-linear BS; (TOMAZELLA et al., 2018) apresentam o modelo regressão BS *zero-adjusted*; (TOMAZELLA et al., 2019) derivam o modelo regressão BS *zero-adjusted* reparametrizado; (CRIBARI-NETO; FONSECA, 2019) propõem um modelo de regressão log-linear BS bimodal; (MARTÍNEZ; GIRALDO; LEIVA, 2019) desenvolvem um modelo de regressão funcional BS para dados espaciais.

A distribuição EVBS, definida por (FERREIRA; GOMES; LEIVA, 2012), surge a partir de uma transformação sobre a variável aleatória GEV. A transformação é semelhante a usada

por (BIRNBAUM; SAUNDERS, 1969a) quando expressam a distribuição Birnbaum-Saunders em termos da variável Normal padrão. Logo, a distribuição EVBS apresenta propriedades tanto da BS quanto da GEV. Em (FERREIRA; GOMES; LEIVA, 2012) encontra-se algumas das funções que caracterizam a EVBS, por exemplo, a fdp e a fda. Além disso, destaca-se naquele trabalho o fato de que os autores especificam o domínio de atração da EVBS, relacionando-o com o domínio de atração da variável que sobre a qual a transformação é imposta.

Credita-se a (LEIVA et al., 2016) a proposta de um modelo de regressão associado à distribuição EVBS. Na formulação do modelo, os autores seguem um caminho análogo ao trilhado por (RIECK; NEDELMAN, 1991) quando desenvolveram o modelo de regressão correspondente à BS. Mais precisamente, (LEIVA et al., 2016) desenvolvem o modelo, especificando a parte determinística e caracterizando a componente estocástica como tendo distribuição log-EVBS. Em seguida, abordam o problema de estimação dos parâmetros por máxima verossimilhança, conduzem uma análise de diagnóstico via resíduos e constatam, por meio de simulações de Monte Carlo, que o resíduo Cox-Snell e o resíduo quantílico apresentam bom desempenho, dentre cinco investigados, na verificação da qualidade do ajuste do modelo.

Duas outras publicações que tratam do modelo EVBS são (GOMES; FERREIRA; LEIVA, 2012) e (LILLO et al., 2018). Na primeira, os autores reportam-se a (VILCA; LEIVA, 2006) para destacar um resultado que possibilita a obtenção do r -ésimo momento da distribuição EVBS, com r inteiro, recorrendo aos momentos da GEV. Na segunda, são obtidas as equações que permitem determinar, numericamente, os estimadores de L-momentos das distribuições BS e EVBS. Além disso, os pesquisadores conduzem um estudo de simulação onde comparam o desempenho do estimador de L-momentos (LM) com o EMV em amostras pequenas e grandes. Por fim, fazem uso dos estimadores de L-Momentos para construir métodos de avaliação de bondade do ajuste.

Por fim, apresentaremos algumas áreas de aplicação do modelo EVBS e do seu correspondente modelo de regressão. Em (FERREIRA; GOMES; LEIVA, 2012) encontra-se uma aplicação na área ambiental em que a variável EVBS modela o nível diário de ozônio registrados na cidade de New York; em (GOMES; FERREIRA; LEIVA, 2012) consta uma aplicação em biometria onde se modela o tempo de sobrevivência de pacientes com mieloma múltiplo. Mais recentemente, em (LEIVA et al., 2016) encontra-se outra aplicação em meio ambiente onde se modela a concentração de ozônio máxima diária na cidade de La Calera-Chile, porém aqui a modelagem inclui a temperatura máxima diária como variável explicativa; já em (LILLO et al., 2018) consta uma aplicação em geologia na qual a variável EVBS modela a magnitude máxima anual de

terremotos ocorridos na zona sísmica Andean South America. Destaquemos que nestes dois últimos trabalhos a modelagem incide sobre dados extremos, obtidos segundo a metodologia do máximo por bloco (BM).

Em resumo, o fato da distribuição BS ter origem relacionada com um modelo físico combinado com as propriedades estatísticas que o modelo BS detém contribuíram para que a mesma seja alvo de estudo por vários pesquisadores, gerando uma literatura vasta, diversa e atual que se estende por várias áreas da Estatística. Para aqueles que buscam um estudo amplo e pormenorizado sobre a distribuição BS, recomendamos (LEIVA, 2016). Por outro lado, se o objetivo é se familiarizar com os principais resultados da literatura e, paralelamente, identificar alguns problemas abertos que podem ser considerados em pesquisas futuras, destaquemos o artigo *review* devido a (BALAKRISHNAN; KUNDU, 2019).

1.6 SUPORTE COMPUTACIONAL

Ao longo deste trabalho usamos o editor de textos \LaTeX para edição tipográfica e empregamos o *software* estatístico R, versão 4.0.5, na construção de gráficos e na obtenção dos resultados numéricos. Especialmente, o pacote *extRemes*.

2 O MODELO BIRNBAUM-SAUNDERS E SUA VERSÃO VALOR EXTREMO.

2.1 INTRODUÇÃO

No início deste capítulo, apresentamos a distribuição BS e uma síntese dos principais resultados existentes na literatura estatística que estão relacionados com o modelo BS. Mais precisamente, relatamos sobre a origem, os pressupostos, a definição e algumas propriedades do modelo; discutimos o problema de estimação dos parâmetros da distribuição BS por máxima verossimilhança e apresentamos o modelo de regressão Birnbaum-Saunders desenvolvido por (RIECK; NEDELMAN, 1991). Em seguida, discorremos sobre a versão valor extremo da distribuição BS, definida por (FERREIRA; GOMES; LEIVA, 2012), citamos algumas propriedades do modelo; propomos um gerador de números aleatórios; estabelecemos condições necessárias e suficientes para existência dos momentos; expomos alguns resultados obtidos por (FERREIRA; GOMES; LEIVA, 2012) sobre o domínio de atração da EVBS. Além disso, abordamos o problema de estimação dos parâmetros do modelo EVBS por máxima verossimilhança. Por fim, estudamos o modelo de regressão EVBS, desenvolvido por (LEIVA et al., 2016), deduzimos algumas propriedades da variável aleatória que caracteriza a componente estocástica do modelo; tratamos do problema de estimação e implementamos um estudo de simulação a fim de investigar propriedades assintóticas do estimador de máxima verossimilhança. Concluimos o capítulo com uma aplicação na qual modelamos a precipitação pluviométrica máxima anual registrada na estação meteorológica Recife-Curado.

2.2 A DISTRIBUIÇÃO BIRNBAUM-SAUNDERS

2.2.1 Origem, pressupostos e derivação

De acordo com (BIRNBAUM; SAUNDERS, 1969a), as distribuições bi-paramétricas usualmente empregadas na modelagem do tempo até a ocorrência de falha por causa de fadiga em materiais são a log-normal, a Weibull e a gama. Porém, chamam atenção para o fato de que embora esses modelos se ajustem de forma satisfatória na região central da distribuição do tempo, nas caudas os ajustes apresentam, em geral, resultados muito distintos, pois na estimação de quantis, de ordem $1/1000$, observa-se grande discrepância nos valores preditos segundo tais modelos. Diante dessa problemática, (BIRNBAUM; SAUNDERS, 1969a) propuseram

uma nova família de distribuições para modelar o tempo até a ocorrência de falha por fadiga. (LEIVA, 2016) reporta-se a (ASTM-INTERNATIONAL, 2013) para caracterizar a fadiga como sendo um dano acumulativo que ocorre em materiais sujeitos a diferentes níveis de tensão ou cargas que variam ciclicamente. Admite-se que o material está exposto a uma sequência de cargas durante períodos de tempo que correspondem a um ciclo e que a mesma sequência é aplicada repetidamente no material, possibilitando desgaste por fadiga. A falha ocorre quando a quantidade de dano acumulativo excede um limiar de resistência associado ao tipo de material.

Segundo (LEIVA, 2016), a teoria para fadiga em materiais baseia-se no pressuposto de que o surgimento e a propagação de uma rachadura dominante é a principal causa da falha. Além disso, cita que o processo de fadiga, em geral, passa pelas seguintes etapas:

- E1. O surgimento de uma fissura microscópica;
- E2. O crescimento e a propagação da fissura, devido às cargas cíclicas de tensão, provocam uma rachadura no material;
- E3. Ocorrência da falha do material por fadiga.

No desenvolvimento da família de distribuições, (BIRNBAUM; SAUNDERS, 1969a) fixam algumas considerações e admitem certas hipóteses. Consideram que cada unidade do material está exposta a uma sequência de m cargas que se repetem de modo cíclico, denotada por meio do seguinte esquema

$$\underbrace{l_1, l_2 \dots, l_m}_{1^\circ \text{ ciclo}}, \quad \underbrace{l_{m+1}, l_{m+2} \dots, l_{2m}}_{2^\circ \text{ ciclo}}, \quad \dots, \quad \underbrace{l_{jm+1}, l_{jm+2} \dots, l_{jm+m}}_{(j+1)\text{-ésimo ciclo}}, \quad \dots \quad (2.1)$$

em que $l_{jm+i} = l_{km+i}$, para $j \neq k$.

Além disso, (BIRNBAUM; SAUNDERS, 1969a) fazem várias suposições expressas por (LEIVA, 2016) nos seguintes termos:

- H1. O material está sujeito a uma sequência de cargas capazes de provocar fissuras em sua estrutura, ocasionando desgaste;
- H2. A falha ocorre no momento em que o tamanho total da fissura dominante excede um limiar crítico, denotado por ω ;
- H3. A sequência de cargas imposta sobre o material é mesma de um ciclo para outro;

- H4. Durante um ciclo de cargas, o incremento de extensão na fissura devido à i -ésima carga, denotado por X_i , é uma variável aleatória cuja distribuição de probabilidade depende de todas as cargas l_j para $j < i$ e da extensão da fissura no momento anterior;
- H5. O tamanho da fissura devido ao k -ésimo ciclo de cargas é uma variável aleatória, denotada por Y_k , cuja distribuição de probabilidade possui média μ e variância σ^2 , ambas finitas para todo $k = 1, 2, \dots$;
- H6. Se Y_j e Y_k , com $j \neq k$, são variáveis que indicam o tamanho de fissura devido a ciclos distintos de cargas, então admite-se que Y_j e Y_k são independentes.

Das suposições acima, conclui-se que o tamanho da fissura devido ao $(k + 1)$ -ésimo ciclo de cargas é

$$Y_{k+1} = X_{km+1} + X_{km+2} + \dots + X_{km+m}, \quad k, m \in \mathbb{N}.$$

Consequentemente, a extensão total da fissura imediatamente após o n -ésimo ciclo é a variável aleatória, S_n , dada por

$$S_n = \sum_{k=1}^n Y_k$$

Se N é a variável aleatória que indica o número de ciclos até a ocorrência da falha, tem-se

$$P(N \leq n) = P(S_n > \omega) = 1 - P\left(\frac{\sum_{k=1}^n Y_k - n\mu}{\sqrt{n} \sigma} \leq \frac{\omega}{\sqrt{n} \sigma} - \frac{\sqrt{n}\mu}{\sigma}\right).$$

Dado que as variáveis Y_k 's são independentes e identicamente distribuídas (iid) com variância finita, o Teorema Central do Limite assegura que

$$P(N \leq n) \approx 1 - \Phi\left(\frac{\omega}{\sqrt{n} \sigma} - \frac{\sqrt{n}\mu}{\sigma}\right),$$

onde $\Phi(\cdot)$ denota a função de distribuição acumulada (fda) da normal padrão. Ou equivalentemente,

$$P(N \leq n) \approx \Phi\left(\frac{\sqrt{n}\mu}{\sigma} - \frac{\omega}{\sqrt{n} \sigma}\right). \quad (2.2)$$

Baseando-se na equação (2.2), (BIRNBAUM; SAUNDERS, 1969a) propuseram uma família de distribuições na qual a variável aleatória T , vista pelo autores como sendo uma extensão contínua de N , possui função de distribuição dada por

$$F(t) = \Phi\left(\frac{1}{\alpha} \left[\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}} \right]\right), \quad t > 0, \quad (2.3)$$

em que $\alpha = \frac{\sigma}{\sqrt{\omega\mu}}$, $\beta = \frac{\omega}{\mu}$.

2.2.2 Definição e algumas propriedades básicas

Diz-se que uma variável aleatória T possui distribuição BS com parâmetros $\alpha > 0$ e $\beta > 0$, e denota-se $T \sim \text{BS}(\alpha, \beta)$, se a sua respectiva fda, denotada por F , pode ser escrita na forma

$$F(t) = \Phi \left(\frac{1}{\alpha} \left[\left(\frac{t}{\beta} \right)^{\frac{1}{2}} - \left(\frac{\beta}{t} \right)^{\frac{1}{2}} \right] \right), \quad t > 0. \quad (2.4)$$

Nota-se que α e β são parâmetros de forma e escala em (2.4), respectivamente.

Ao definirmos a função

$$\xi : (0, \infty) \rightarrow \mathbb{R}, \quad \xi(t) = t^{1/2} - t^{-1/2}, \quad (2.5)$$

é imediato verificarmos que ξ é contínua, derivável e estritamente crescente no intervalo $(0, \infty)$. Além disso, notemos que $\lim_{u \rightarrow 0^+} \xi(u) = -\infty$ e $\lim_{u \rightarrow +\infty} \xi(u) = +\infty$. Esses fatos combinados com as propriedades da função Φ nos asseguram que F é uma legítima função de distribuição e admite ser reescrita na seguinte forma:

$$F(t) = \Phi \left(\alpha^{-1} \xi(t/\beta) \right), \quad t > 0.$$

A partir de (2.4), constata-se que se $T \sim \text{BS}(\alpha, \beta)$, então a fdp associada a T , é

$$f(t) = \frac{1}{2\alpha\beta\sqrt{2\pi}} \left[\left(\frac{t}{\beta} \right)^{-\frac{1}{2}} + \left(\frac{t}{\beta} \right)^{-\frac{3}{2}} \right] \exp \left(-\frac{1}{2\alpha^2} \left[\frac{t}{\beta} + \frac{\beta}{t} - 2 \right] \right), \quad t > 0. \quad (2.6)$$

Em termos da função ξ , definida em (2.5) e cuja primeira derivada é igual a $\xi'(u) = \frac{1}{2}(u^{-1/2} + u^{-3/2})$, a densidade apresentada em (2.6) pode ser reescrita da seguinte forma:

$$f(t) = \frac{1}{\alpha\beta\sqrt{2\pi}} \xi'(t/\beta) \exp \left(-\frac{1}{2\alpha^2} \xi^2(t/\beta) \right), \quad t > 0, \alpha > 0, \beta > 0. \quad (2.7)$$

Recorrendo-se a (2.4), verifica-se que dados $\alpha > 0$ e $\beta > 0$, se $Z \sim \text{N}(0, 1)$, então

$$T = \beta \left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2} \right)^2 + 1} \right)^2 \sim \text{BS}(\alpha, \beta).$$

Reciprocamente, se $T \sim \text{BS}(\alpha, \beta)$, então

$$Z = \frac{1}{\alpha} \left(\sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right) \sim \text{N}(0, 1). \quad (2.8)$$

Por fim, é imediato verificar que se $T \sim \text{BS}(\alpha, \beta)$, então são válidas as seguintes propriedades básicas:

- i) $cT \sim \text{BS}(\alpha, c\beta)$ para todo número real $c > 0$.
- ii) $\frac{1}{T} \sim \text{BS}(\alpha, \frac{1}{\beta})$.
- iii) $V = \frac{1}{\alpha^2} \left(\frac{T}{\beta} + \frac{\beta}{T} - 2 \right) \sim \chi_{(1)}^2$.

2.2.3 A função de sobrevivência e a função de risco

Na caracterização de distribuições que mensuram o tempo até a ocorrência de falha ou tempo de sobrevivência é comum o emprego de outras funções além da fda e fdp. Por exemplo, a função de sobrevivência, ou de confiabilidade, e a função de risco, ou taxa de risco instantânea. Dada uma variável aleatória T , absolutamente contínua, cujas correspondentes fda e fdp são denotadas por F e f , usualmente, define-se

1. A função de sobrevivência, denotada por S , associa a cada número real, t , a probabilidade de que o tempo até a ocorrência de falha seja maior do que t . Mais precisamente,

$$S : \mathbb{R} \rightarrow [0, 1], \quad S(t) = P(T > t). \quad (2.9)$$

2. A função de risco, denotada por h , determina a razão instantânea de falha no tempo t , dado a não ocorrência de falha até o tempo t . Ou seja,

$$h : \mathbb{R}_+ \rightarrow \mathbb{R}, \quad h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}. \quad (2.10)$$

Algumas propriedades gerais das funções S e h , assim como, expressões matemáticas que relacionam F , f , S e h , são fáceis de serem derivadas e podem ser encontradas em (LAWLESS, 2003).

Se $T \sim \text{BS}(\alpha, \beta)$, então a função de sobrevivência, S , e a função de risco, h , são dadas, respectivamente, por:

$$S(t) = 1 - F(t; \alpha, \beta) = \Phi \left(-\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right), \quad t > 0, \quad (2.11)$$

e

$$h(t) = \frac{f(t; \alpha, \beta)}{S(t; \alpha, \beta)} = \frac{\frac{1}{\alpha} \phi \left(\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right) \xi' \left(\frac{t}{\beta} \right)}{\Phi \left(-\frac{1}{\alpha} \xi \left(\frac{t}{\beta} \right) \right)}, \quad t > 0 \quad (2.12)$$

onde ϕ denota a fdp da variável aleatória normal padrão. Particularmente, sobre a função de risco h de uma variável $T \sim \text{BS}(\alpha, \beta)$ são válidas as seguintes propriedades:

1ª) h possui um único ponto de máximo. Mais precisamente, h é crescente para $t < t_0$ e decrescente para $t > t_0$, sendo t_0 um único número real positivo em torno do qual a derivada primeira de h muda de sinal. No caso em que $\beta = 1$, (KUNDU; KANNAN; BALAKRISHNAN, 2008) obtém valores de t_0 para diferentes escolhas de α e sugerem que, quando $\alpha > 0.25$, pode-se usar $\tilde{t}_0 = \frac{1}{(-0.4604 + 1.8417\alpha)^2}$ como uma aproximação para t_0 . Caso $\beta \neq 1$, mostra-se que o ponto onde h atingirá o máximo é βt_0 .

2ª) Quando t tende ao infinito, $h(t)$ aproxima-se de $\frac{1}{2\alpha^2\beta}$. Ou seja, $\lim_{t \rightarrow \infty} h(t) = \frac{1}{2\alpha^2\beta}$.

Mais detalhes sobre o comportamento da função de risco h de uma variável BS podem ser encontrados nos trabalhos de (CHANG; TANG, 1993) e (KUNDU; KANNAN; BALAKRISHNAN, 2008).

2.2.4 Média, variância, assimetria e curtose

Usando o fato de que uma variável $T \sim \text{BS}(\alpha, \beta)$ está relacionada com a variável normal padrão, Z , por meio da expressão

$$T = \beta \left(\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1} \right)^2, \quad (2.13)$$

(LEIVA, 2016) mostra que para todo r inteiro positivo o r -ésimo momento central da variável $T \sim \text{BS}(\alpha, \beta)$ é

$$\mathbb{E}(T^r) = \beta^r \sum_{j=0}^r \binom{2r}{2j} \sum_{k=0}^j \binom{j}{k} \frac{(2(r-j+k))!}{2^{r-j+k}(r-j+k)!} \left(\frac{\alpha}{2}\right)^{2(r-j+k)}. \quad (2.14)$$

Consequentemente,

$$\mathbb{E}(T) = \beta \left(1 + \frac{\alpha^2}{2} \right) \quad (2.15)$$

e

$$\text{Var}(T) = \beta^2 \alpha^2 \left(1 + \frac{5}{4} \alpha^2 \right). \quad (2.16)$$

Além disso, os coeficientes de variação (CV), assimetria (CS) e curtose (CK) são dados, respectivamente, por

$$\text{CV}(T) = \alpha \left(\frac{\sqrt{5\alpha^2 + 4}}{\alpha^2 + 2} \right), \quad (2.17)$$

$$\text{CS}(T) = \frac{16\alpha^2(11\alpha^2 + 6)^2}{(5\alpha^2 + 4)^3}, \quad (2.18)$$

e

$$CK(T) = 3 + \frac{6\alpha^2(93\alpha^2 + 40)}{(5\alpha^2 + 4)^2}. \quad (2.19)$$

Mais informações, em particular, sobre os momentos inversos $E((1/T)^r)$, podem ser obtidas em (LEIVA, 2016).

Recorrendo à função geradora de momentos, (RIECK, 1999) obtém os momentos da BS, com a vantagem de incluir o caso em que r assume valores fracionários. A partir da fgm da variável $Y = \log T$, onde $T \sim BS(\alpha, \beta)$, ele conclui que

$$E(T^r) = \beta^r \left(\frac{K_{r+1/2}(1/\alpha^2) + K_{r-1/2}(1/\alpha^2)}{2K_{1/2}(1/\alpha^2)} \right),$$

onde $K_\lambda(\cdot)$ é a função de Bessel modificada do terceiro tipo e definida por

$$K_\lambda(u) = \frac{1}{2} \left(\frac{u}{2} \right)^\lambda \int_0^\infty v^{-\lambda-1} \exp\left(-v - \frac{u^2}{4v}\right) dv.$$

2.2.5 Estimação por máxima verossimilhança

A estimação dos parâmetros, α e β , do modelo BS, foi inicialmente tratada por (BIRNBAUM; SAUNDERS, 1969b) utilizando o método de máxima verossimilhança. No artigo, constam as funções e as equações que levam à determinação dos estimadores, acompanhadas de uma argumentação da unicidade do estimador de máxima verossimilhança para o parâmetro β , denotado por $\hat{\beta}$, e uma fórmula que permite determinar a estimativa de α em função da obtida para β . Aliado a isso, há dois procedimentos numéricos que permitem o cálculo de $\hat{\beta}$, um deles baseado no método de Newton-Raphson e o outro em um método de aproximação sucessivas, ambos convergindo para $\hat{\beta}$. A seguir, apresentaremos uma síntese dos resultados obtidos.

A média aritmética e a harmônica de um conjunto finito de números reais positivos, t_1, \dots, t_n , denotadas por s e r , respectivamente, são definidas por:

$$s = \frac{1}{n} \sum_{i=1}^n t_i \quad \text{e} \quad r = \left[\frac{1}{n} \sum_{i=1}^n t_i^{-1} \right]^{-1}.$$

Além disso, definamos as funções

$$K : \mathbb{R}_+ \mapsto \mathbb{R}, \quad K(x) = \left[\frac{1}{n} \sum_{i=1}^n (x + t_i)^{-1} \right]^{-1}$$

e

$$g : \mathbb{R}_+ \mapsto \mathbb{R}, \quad g(x) = x^2 - x[2r + K(x)] + r[s + K(x)]. \quad (2.20)$$

Teorema 2.1 (BIRNBAUM; SAUNDERS, 1969b) Se t_1, \dots, t_n são observações correspondentes a uma amostra de variáveis aleatórias independentes e identicamente distribuídas, T_1, T_2, \dots, T_n , da variável $T \sim \text{BS}(\alpha, \beta)$, então o estimador de máxima verossimilhança de β , denotado por $\hat{\beta}$, é a única solução positiva da equação $g(x) = 0$, ou seja,

$$x^2 - x[2r + K(x)] + r[s + K(x)] = 0,$$

onde $s = \frac{1}{n} \sum_{i=1}^n t_i$ e $r = \left[\frac{1}{n} \sum_{i=1}^n t_i^{-1} \right]^{-1}$. Além disso, $r < \hat{\beta} < s$ e a estimativa de máxima verossimilhança (emv) de α , denotada por $\hat{\alpha}$, é dada por

$$\hat{\alpha} = \left(\frac{s}{\hat{\beta}} + \frac{\hat{\beta}}{r} - 2 \right)^{1/2}.$$

Agora, serão apresentados dois métodos para encontrar a $\hat{\beta}$, ambos devido a (BIRNBAUM; SAUNDERS, 1969b).

Método I. Se os dados observados, t_1, \dots, t_n , satisfazem a desigualdade

$$2s < 3r + \min\{t_1, t_2, \dots, t_n\},$$

então o procedimento iterativo de Newton-Raphson com relação de recorrência dada por

$$\beta_{n+1} = \beta_n - \frac{g(\beta_n)}{g'(\beta_n)}, \quad n = 0, 1, \dots$$

converge para β qualquer que seja o valor inicial $\beta_0 \in (r, s)$, sendo $g(\cdot)$ a função definida em (2.20).

Método II. Sejam A e H funções definidas por

$$H : \mathbb{R}_+ \mapsto \mathbb{R}, \quad H(x) = \frac{1}{x}K(x)$$

e

$$A : \mathbb{R}_+ \mapsto \mathbb{R}, \quad A(x) = r + H(x) - [H^2(x) - r(s - r)]^{1/2}.$$

Se os dados observados, t_1, \dots, t_n , são tais que $2r > s$, então para qualquer valor inicial $\beta_0 \in (r, s)$, $A^{(n)}(\beta_0)$ converge para $\hat{\beta}$, quando $n \rightarrow \infty$.

Em (ENGELHARDT; BAIN; WRIGHT, 1981) argumenta-se que a família de distribuições BS é regular, com matriz de informação Fisher, $\mathcal{I}(\alpha, \beta)$, dada por

$$\mathcal{I}(\alpha, \beta) = - \begin{bmatrix} \frac{2n}{\alpha^2} & 0 \\ 0 & \frac{n}{\alpha^2 \beta^2} \left(1 + \alpha(2\pi)^{-1/2} \varphi(\alpha) \right) \end{bmatrix},$$

onde $\varphi(\alpha) = \alpha\sqrt{\frac{\pi}{2}} - \pi e^{2/\alpha^2}[1 - \Phi(2/\alpha)]$. Destaca-se a ortogonalidade dos parâmetros α e β . Conseqüentemente, usando propriedades assintóticas dos EMVs para parâmetros de classes de distribuições regulares, constata-se que a distribuição assintótica do EMV, $\hat{\theta} = (\hat{\alpha}, \hat{\beta})^\top$, é normal bivariada. Mais precisamente, conforme pode ser visto em (BALAKRISHNAN; KUNDU, 2019), tem-se

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \sim \mathbf{N}_2 \left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \frac{\alpha^2}{2n} & 0 \\ 0 & \frac{\beta^2}{n(0.25 + \alpha^{-2} + I(\alpha))} \end{pmatrix} \right], \quad (2.21)$$

onde $I(\alpha) = 2 \int_0^\infty [(1 + g(\alpha x))^{-1} - 1/2]^2 d\Phi(x)$, com $g(y) = 1 + \frac{y^2}{2} + y \left(1 + \frac{y^2}{4}\right)^{1/2}$. Neste ponto, (BALAKRISHNAN; KUNDU, 2019) destacam que os EMVs $\hat{\alpha}$ e $\hat{\beta}$ são assintoticamente independentes e que a partir do resultado assintótico expresso em (2.21) pode-se construir intervalos de confiança para os parâmetros α e β . O trabalho de (NG; KUNDU; BALAKRISHNAN, 2003) acentua que os EMVs $\hat{\alpha}$ e $\hat{\beta}$ têm um viés potencial em amostras pequenas e quando α é grande. Mais precisamente, os autores obtêm as seguintes aproximações para o viés $\hat{\alpha}$ e $\hat{\beta}$:

$$B(\hat{\alpha}) \approx -\frac{\alpha}{n} \quad \text{e} \quad B(\hat{\beta}) \approx -\frac{\alpha^2}{4n}.$$

Diante disso, eles impuseram uma correção de viés nos EMVs, resultando nos seguintes estimadores de viés-corrigido:

$$\hat{\alpha}^* = \left(\frac{n}{n-1}\right) \hat{\alpha} \quad (2.22)$$

e

$$\hat{\beta}^* = \left(1 + \frac{(\hat{\alpha}^*)^2}{4n}\right)^{-1} \hat{\beta}. \quad (2.23)$$

Por fim, baseando-se nos estimadores de viés-corrigido dados pelas equações (2.22) e (2.23) e na normalidade assintótica do EMV expressa em (2.21), (NG; KUNDU; BALAKRISHNAN, 2003) obtêm os seguintes intervalos de confiança para os parâmetros α e β da distribuição BS:

$$IC(\alpha, \gamma) = \left[\hat{\alpha}^* \left(\sqrt{\frac{n}{2}} \frac{z_{\gamma/2}}{(n-1)} + 1 \right)^{-1}, \hat{\alpha}^* \left(\sqrt{\frac{n}{2}} \frac{z_{1-\gamma/2}}{(n-1)} + 1 \right)^{-1} \right]$$

e

$$IC(\beta, \gamma) = \left[\hat{\beta}^* \left(\frac{n}{h_1(\hat{\alpha}^*)} \frac{4z_{\gamma/2}}{(4n + (\hat{\alpha}^*)^2)} + 1 \right)^{-1}, \hat{\beta}^* \left(\frac{n}{h_1(\hat{\alpha}^*)} \frac{4z_{1-\gamma/2}}{(4n + (\hat{\alpha}^*)^2)} + 1 \right)^{-1} \right]$$

em que $h_1(y) = 0.25 + y^{-2} + I(y)$ e $I(y) = 2 \int_0^\infty [(1 + g(yx))^{-1} - 1/2]^2 d\Phi(x)$, com $g(y) = 1 + \frac{y^2}{2} + y \left(1 + \frac{y^2}{4}\right)^{1/2}$.

2.2.6 Uma relação entre as distribuições Birnbaum-Saunders e senh-normal

A distribuição senh-normal (SN), definida por (RIECK, 1989), é dotada de boas propriedades estatísticas e se encontra estreitamente relacionada com a distribuição BS. A seguir, apresentaremos tal distribuição e citaremos algumas de suas propriedades.

Seja Y uma variável aleatória cuja fda admite ser escrita na forma

$$F(y) = \Phi\left(\frac{2}{\alpha}\sinh\left(\frac{y-\mu}{\sigma}\right)\right), \quad y \in \mathbb{R}, \quad (2.24)$$

onde Φ denota fda da normal padrão, $\alpha > 0$, $\mu \in \mathbb{R}$ e $\sigma > 0$. Neste caso, diz-se que Y possui distribuição senh-normal com parâmetros de forma, locação e escala iguais a α , μ e σ , respectivamente. E, denota-se por $Y \sim \text{SN}(\alpha, \mu, \sigma)$.

A partir da função exibida em (2.24), constata-se que as fdp's, de sobrevivência e de risco de uma variável $Y \sim \text{SN}(\alpha, \mu, \sigma)$ são dadas, respectivamente, por

$$f(y) = \frac{2}{\alpha\sigma} \cosh\left(\frac{y-\mu}{\sigma}\right) \phi\left(\frac{2}{\alpha}\sinh\left(\frac{y-\mu}{\sigma}\right)\right), \quad y \in \mathbb{R},$$

$$S(y) = \Phi\left(-\frac{2}{\alpha}\sinh\left(\frac{y-\mu}{\sigma}\right)\right), \quad y \in \mathbb{R}$$

e

$$h(y) = \frac{2}{\alpha\sigma} \cosh\left(\frac{y-\mu}{\sigma}\right) \frac{\phi\left(\frac{2}{\alpha}\sinh\left(\frac{y-\mu}{\sigma}\right)\right)}{\Phi\left(-\frac{2}{\alpha}\sinh\left(\frac{y-\mu}{\sigma}\right)\right)}, \quad y \in \mathbb{R},$$

onde Φ e ϕ denotam a fda e a fdp da normal padrão, respectivamente. Também decorre de (2.24) as seguintes implicações: se $Y \sim \text{SN}(\alpha, \mu, \sigma)$, então $Z = \frac{2}{\alpha}\sinh\left(\frac{Y-\mu}{\sigma}\right) \sim \text{N}(0, 1)$, e, reciprocamente, se $Z \sim \text{N}(0, 1)$, então $Y = \mu + \sigma \operatorname{arcsinh}\left(\frac{\alpha Z}{2}\right) \sim \text{SN}(\alpha, \mu, \sigma)$.

Agora, citemos algumas propriedades de uma variável $Y \sim \text{SN}(\alpha, \mu, \sigma)$ que foram obtidas por (RIECK, 1989):

1. A distribuição SN é simétrica em torno do parâmetro de locação μ ;
2. Se $\alpha \leq 2$, a distribuição SN é unimodal. Se $\alpha > 2$, a mesma é bimodal;
3. $\mathbb{E}(Y) = \mu$ e $\operatorname{Var}(Y) = \sigma^2 \omega(\alpha)$, onde $\omega(\alpha)$ é a variância no caso particular $\sigma = 1$;
4. Se $Y \sim \text{SN}(\alpha, \mu, \sigma)$, então a variável $S_\alpha = \frac{2(Y_\alpha - \mu)}{\alpha\sigma}$ converge para normal padrão quando α aproxima-se de zero.

Por fim, apresentemos uma relação entre a distribuição senh-normal e Birnbaum-Saunders sob uma transformação logarítmica.

Teorema 2.2 (RIECK; NEDELMAN, 1991) Se $T \sim \text{BS}(\alpha, \beta)$, então a variável aleatória $Y = \log(T) \sim \text{SN}(\alpha, \mu = \log(\beta), \sigma = 2)$.

2.2.7 O modelo de regressão log-Birnbaum-Saunders

Motivados por aplicações na caracterização do tempo de falha, devido à fadiga de materiais sujeitos a processos com diferentes padrões de forças cíclicas, (RIECK; NEDELMAN, 1991) desenvolveram um modelo de regressão log-linear para a distribuição BS. Ressaltaram que o modelo poder ser usado para comparar a mediana de várias populações ou avaliar o impacto de variáveis de controle em testes de vida acelerados. Ao tratarem do problema de estimação dos parâmetros do modelo, eles usaram os métodos de mínimos quadrados e o de máxima verossimilhança. Sobre a estimação do parâmetro de forma α , apresentaram uma breve discussão sobre a existência de soluções para as equações de máxima verossimilhança nos casos em que $\alpha > 2$ e $\alpha \leq 2$, destacando a unicidade no último caso. Mediante simulações de Monte Carlo, obtiveram indícios de que, para valores pequenos de α , o estimador de mínimos quadrados (EMQ) dos coeficientes da regressão apresentava eficiência relativa próxima de 1, porém tal eficiência decresce a medida que α cresce. Ainda em relação à estimação de α , comparam o EMQ com o EMV, segundo a eficiência relativa, e constataram, por meio de simulações, superioridade do último sobre o primeiro, mesmo em amostras pequenas ($n = 10$ ou $n = 20$).

Na caracterização do modelo, os autores admitem que T_1, T_2, \dots, T_n são variáveis aleatórias independentes tais que $T_i \sim \text{BS}(\alpha_i, \eta_i)$ para $i = 1, 2, \dots, n$. Além disso, supõem que a distribuição da variável resposta T_i depende de um vetor, $\mathbf{x}_i^\top = (1, x_{i2}, \dots, x_{ip})$, contendo $p - 1$ variáveis explicativas. Por fim, fixam as seguintes hipóteses:

- i) $\eta_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta})$ para $i = 1, 2, \dots, n$, sendo $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^\top$ um vetor de parâmetros desconhecidos;
- ii) Os parâmetros de forma α_i não dependem do vetor de covariáveis \mathbf{x}_i^\top , ou seja, $\alpha_i = \alpha$ para todo $i = 1, 2, \dots, n$.

Sob as suposições fixadas acima, (RIECK; NEDELMAN, 1991) admitem que a variável resposta

T_i pode ser expressa na seguinte forma:

$$T_i = \eta_i \zeta_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \zeta_i, \quad i = 1, \dots, n \quad (2.25)$$

em que a componente aleatória ζ_i é tal que $\zeta_i \sim \text{BS}(\alpha, 1)$ para $i = 1, \dots, n$. Ao considerarem, $Y_i = \log(T_i)$, obtiveram o seguinte modelo de regressão:

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \nu_i, \quad i = 1, \dots, n, \quad (2.26)$$

em que os erros, ν_i 's, são variáveis aleatórias iid tais que $\nu_i = \log(\zeta_i) \sim \text{SN}(\alpha, 0, 2)$ para todo $i = 1, \dots, n$, o vetor \mathbf{x}_i^\top é composto por valores reais fixos das variáveis explicativas e $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^\top$ é o vetor de coeficientes que deve ser estimado.

Dado que $\nu_i \sim \text{SN}(\alpha, 0, 2)$ para todo $i = 1, 2, \dots, n$, cada variável Y_i do modelo (2.26) possui a seguinte função densidade de probabilidade:

$$f_{Y_i}(y) = \frac{1}{\alpha \sqrt{2\pi}} \cosh\left(\frac{y - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right) \exp\left(-\frac{2}{\alpha^2} \sinh^2\left(\frac{y - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)\right), \quad y \in \mathbb{R}.$$

Ou seja, $Y_i \sim \text{SN}(\alpha, \mathbf{x}_i^\top \boldsymbol{\beta}, 2)$ para cada $i = 1, 2, \dots, n$.

Dada uma amostra $\mathbf{y} = (y_1, y_2, \dots, y_n)^\top$ de observações independentes associadas ao modelo (2.26), para fins de estimação, a função de log-verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$ pode ser expressa na forma

$$l(\boldsymbol{\theta}; \mathbf{y}) = \log \prod_{i=1}^n f(y_i) = \sum_{i=1}^n \log(\xi_{i1}) - \frac{1}{2} \sum_{i=1}^n \xi_{i2}^2, \quad (2.27)$$

em que $\xi_{i1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mu_i}{2}\right)$, $\xi_{i2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mu_i}{2}\right)$ e $\mu_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ para cada $i = 1, 2, \dots, n$.

Logo, obtemos o seguinte sistema de equações:

$$\begin{cases} \frac{\partial}{\partial \beta_j} l(\boldsymbol{\theta}; \mathbf{y}) = \frac{1}{2} \sum_{i=1}^n x_{ij} \left(\xi_{i1} \xi_{i2} - \frac{\xi_{i2}}{\xi_{i1}} \right) = 0, & \text{para } j = 1, 2, \dots, p, \\ \frac{\partial}{\partial \alpha} l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{n}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n \xi_{i2}^2 = 0. \end{cases} \quad (2.28)$$

Consequentemente, os EMVs de α e $\boldsymbol{\beta}$, denotados, respectivamente, por $\hat{\alpha}$ e $\hat{\boldsymbol{\beta}}$, devem satisfazer, como condição necessária, ao sistema de equações (2.28). Uma vez encontrado $\hat{\boldsymbol{\beta}}$, a última equação do sistema (2.28) possibilita determinar $\hat{\alpha}^2$, em função de $\hat{\boldsymbol{\beta}}$, por meio da seguinte equação:

$$\hat{\alpha}^2 = \frac{4}{n} \sum_{i=1}^n \sinh^2\left(\frac{y_i - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}}{2}\right). \quad (2.29)$$

Com o propósito de investigar o comportamento do estimador de máxima verossimilhança $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\alpha})^\top$, (RIECK; NEDELMAN, 1991) supõem que os dados amostrais são tais que quando

o tamanho da amostra cresce se verificam as seguintes hipóteses: (i) existe $B > 0$ tal que $|x_{ij}| \leq B$ quaisquer que sejam $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, p$; (ii) existe uma matriz \mathbf{M} , positiva definida, tal que $\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{\mathbf{x}_i \mathbf{x}_i^\top}{n} = \mathbf{M}$. Sob as hipóteses fixadas, (RIECK; NEDELMAN, 1991) chegam as seguintes conclusões: (a) Existe pelo menos uma solução para o sistema de equações de máxima verossimilhança exibido em (2.28) que é um estimador consistente para o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$; (b) Dentre todas as possíveis soluções do sistema (2.28) uma, e somente uma, converge em probabilidade para $\boldsymbol{\theta}$; (c) $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ converge em distribuição para normal $(p+1)$ -variada com vetor de médias $\mathbf{0}$ e matriz de covariância, $\boldsymbol{\Sigma}$, dada pela inversa da matriz

$$\begin{bmatrix} \frac{C(\alpha)\mathbf{M}}{4} & \mathbf{0} \\ \mathbf{0}^\top & \frac{2}{\alpha^2} \end{bmatrix},$$

onde $\mathbf{0}$ é um vetor nulo com p zeros e $C(\alpha) = 2 + \frac{4}{\alpha^2} - \left(\frac{2\pi}{\alpha^2}\right)^{1/2} [1 - \text{erf}\left(\left(\frac{2}{\alpha^2}\right)^{1/2}\right)] \exp\left(\frac{2}{\alpha^2}\right)$, sendo $\text{erf}(\cdot)$ é a função erro, cuja definição e propriedades podem ser encontradas em (ABRAMOWITZ; STEGUN, 1970). As justificativas das conclusões (a), (b) e (c) encontram-se em (RIECK; NEDELMAN, 1990).

Embasados nos resultados acima, (RIECK; NEDELMAN, 1991) concluem que a matriz $\boldsymbol{\Sigma}$ de covariância assintótica do estimador $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\alpha})^\top$ pode ser aproximada por

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 4 \frac{(\mathbf{X}^\top \mathbf{X})^{-1}}{C(\hat{\alpha})} & \mathbf{0} \\ \mathbf{0}^\top & \frac{\hat{\alpha}^2}{2n} \end{bmatrix}, \quad (2.30)$$

onde $\mathbf{X}^\top = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ possui posto p .

A multiplicidade de soluções do sistema de equações de máxima verossimilhança (2.28) é tratada por (RIECK, 1989). Inicia com o caso em que o parâmetro de forma α é conhecido, considerando, separadamente, as situações em que $\alpha < 2$, $\alpha = 2$ e $\alpha > 2$. As conclusões alcançadas fundamentam-se no sinal da forma Hessiana da função de log-verossimilhança $l(\boldsymbol{\theta}; \mathbf{y})$. Mais precisamente, (RIECK, 1989) constata que se $\alpha < 2$, então a matriz Hessiana de $l(\boldsymbol{\theta}; \mathbf{y})$ é definida negativa para todo $\boldsymbol{\beta}$, acarretando a existência de um único máximo global. Quando $\alpha = 2$, argumenta que $l(\boldsymbol{\theta}; \mathbf{y})$ tem um máximo global, que pode ser único ou não. No caso em que $\alpha > 2$, afirma que a Hessiana é semi-definida negativa para alguns valores de $\boldsymbol{\beta}$ e semi-definida positiva para outros, logo é possível que o sistema de equações possua solução que não seja um ponto de máximo da função de log-verossimilhança. Neste último caso, (RIECK, 1989) exhibe um exemplo, considerando $\alpha = 4$, em que o sistema (2.28) tem

três soluções, sendo duas delas pontos de máximo e a outra um mínimo da função (2.27). A hipótese de que a matriz \mathbf{X} tem posto p é assumida nas análises feitas. Já o caso em que α é desconhecido é tratado por (RIECK, 1989) de forma menos analítica. Mais especificamente, apresenta-se um exemplo que revela a multiplicidade de soluções para o sistema (2.28) e evidencia a existência de solução que não é ponto de máximo de função expressa em (2.27).

2.3 A DISTRIBUIÇÃO BIRNBAUM-SAUNDERS VALOR EXTREMO

2.3.1 Origem, definição e propriedades

A distribuição EVBS, definida (FERREIRA; GOMES; LEIVA, 2012), surge a partir de uma transformação monótona sobre uma variável GEV. Logo, ela preserva algumas propriedades da distribuição BS e outras da GEV, revelando-se como um modelo paramétrico promissor em Estatística de Extremos, conforme citam (GOMES; GUILLOU, 2015). Uma das principais características do modelo EVBS é a presença do parâmetro de forma, γ , proveniente da distribuição GEV. Por um lado, isso faz com o modelo agregue distribuições de caudas leves, pesadas ou curtas, a depender do sinal de γ , tornando-o flexível. Por outro lado, traz a inconveniência de que o suporte da distribuição depende dos parâmetros do modelo, dificultando o processo de estimação. Nesta seção apresentaremos a definição e algumas propriedades da EVBS, passando pela geração de números aleatórios e condições de existência dos momentos. Em seguida, exploraremos alguns resultados obtidos por (FERREIRA; GOMES; LEIVA, 2012) que especificam o máx-domínio de atração da distribuição EVBS. Por fim, discutiremos o problema de estimação por máxima verossimilhança na família EVBS.

Seja X uma variável aleatória que segue uma distribuição valor extremo generalizada com parâmetros de locação, escala e forma iguais a 0, 1 e $\gamma \in \mathbb{R}$, respectivamente. Logo, a fda de $X \sim \text{GEV}(0, 1, \gamma)$, denotada por G_γ , é dada por

$$G_\gamma(x) = \begin{cases} \exp(-\exp(-x)); & \gamma = 0, x \in \mathbb{R}, \\ \exp(-(1 + \gamma x)^{-1/\gamma}); & \gamma \neq 0, 1 + \gamma x > 0. \end{cases} \quad (2.31)$$

Dada uma variável $X \sim \text{GEV}(0, 1, \gamma)$, (FERREIRA; GOMES; LEIVA, 2012) definem uma variável aleatória EVBS com parâmetros $\alpha > 0$, $\beta > 0$ e $\gamma \in \mathbb{R}$, denotada por $T \sim \text{EVBS}(\alpha, \beta, \gamma)$,

por meio da transformação

$$T = \beta \left(\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2} \right)^2 + 1} \right)^2. \quad (2.32)$$

Inicialmente, notemos que qualquer que seja o valor real assumido pela variável X e para todo $\alpha > 0$ é válida a seguinte desigualdade:

$$\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2} \right)^2 + 1} \geq \frac{\alpha X}{2} + \left| \frac{\alpha X}{2} \right| \geq 0. \quad (2.33)$$

Logo, T é não-negativa e a sua fda, F , associa a cada número real t o número $F(t)$ dado por

$$F(t) = \begin{cases} \exp \left(- \exp \left[-\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right] \right), & t > 0, & \text{se } \gamma = 0. \\ \exp \left(- \left[1 + \frac{\gamma}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]^{-1/\gamma} \right), & 1 + \frac{\gamma}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) > 0, & \text{se } \gamma \neq 0. \end{cases} \quad (2.34)$$

Ao definirmos a função

$$a : (0, \infty) \rightarrow \mathbb{R}, \quad a(t) = a_t = \frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right), \quad (2.35)$$

a fda, F , admite ser reescrita da seguinte forma:

$$F(t) = \begin{cases} \exp(-\exp(-a_t)); & t > 0, & \text{se } \gamma = 0. \\ \exp(-(1 + \gamma a_t)^{-1/\gamma}); & 1 + \gamma a_t > 0, & \text{se } \gamma \neq 0. \end{cases} \quad (2.36)$$

Consequentemente, a fdp, f , de uma variável $T \sim \text{EVBS}(\alpha, \beta, \gamma)$ é dada por

$$f(t) = \begin{cases} \frac{1}{2\alpha\beta} \left[\left(\frac{t}{\beta} \right)^{-\frac{1}{2}} + \left(\frac{t}{\beta} \right)^{-\frac{3}{2}} \right] \exp \left\{ - \exp \left[-\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right] \right\} \cdot \\ \exp \left\{ -\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right\}; & t > 0, & \text{se } \gamma = 0. \\ \frac{1}{2\alpha\beta} \left[\left(\frac{t}{\beta} \right)^{-\frac{1}{2}} + \left(\frac{t}{\beta} \right)^{-\frac{3}{2}} \right] \left[1 + \frac{\gamma}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]^{-1-1/\gamma} \cdot \\ \exp \left\{ - \left[1 + \frac{\gamma}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]^{-1/\gamma} \right\}; & \left[1 + \frac{\gamma}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right] > 0, & \text{se } \gamma \neq 0. \end{cases} \quad (2.37)$$

Neste ponto, observemos que se $\gamma > 0$, a condição $1 + \frac{\gamma}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) > 0$ equivale a $t + \frac{\alpha\beta^{1/2}}{\gamma} t^{1/2} - \beta > 0$. Por outro lado, se $\gamma < 0$ a mesma condição equivale a $t + \frac{\alpha\beta^{1/2}}{\gamma} t^{1/2} - \beta < 0$. Isso nos dá condições para explicitarmos o suporte de f quando $\gamma \neq 0$. Mais precisamente, se $\gamma > 0$ então o suporte de f , denotado por S , é o subconjunto $S = \left\{ t \in \mathbb{R}_+; t > \frac{\beta}{2} \left(\left(\frac{\alpha}{\gamma} \right)^2 + 2 - \frac{\alpha}{\gamma} \sqrt{4 + \left(\frac{\alpha}{\gamma} \right)^2} \right) \right\}$. Porém, se $\gamma < 0$, então $S =$

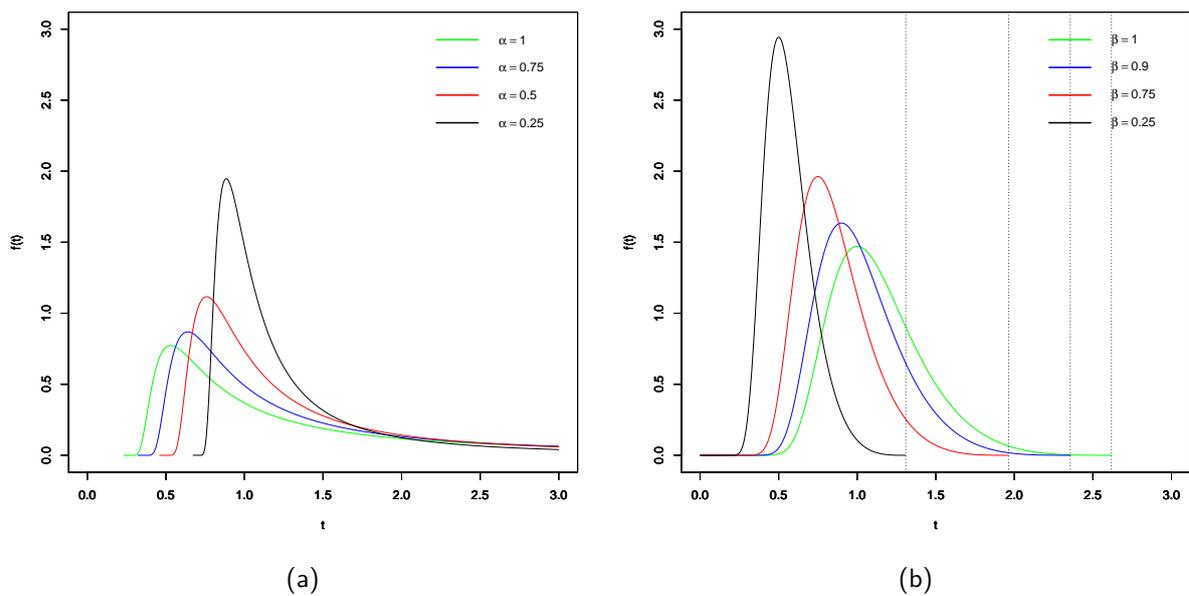
$\left\{ t \in \mathbb{R}_+; 0 < t < \frac{\beta}{2} \left(\left(\frac{\alpha}{\gamma} \right)^2 + 2 - \frac{\alpha}{\gamma} \sqrt{4 + \left(\frac{\alpha}{\gamma} \right)^2} \right) \right\}$. Daí, segue que o limite superior do suporte da distribuição EVBS, definido por $t^* = \sup\{t \in \mathbb{R}; F(t) < 1\}$, é finito se $\gamma < 0$, e infinito quando $\gamma \geq 0$.

Em termos de a_t , definido em (2.35), a função densidade admite ser reescrita na forma

$$f(t) = \begin{cases} A_t \exp(-\exp(-a_t) - a_t); & t > 0, & \text{se } \gamma = 0. \\ A_t (1 + \gamma a_t)^{-1-1/\gamma} \exp(-(1 + \gamma a_t)^{-1/\gamma}); & 1 + \gamma a_t > 0, & \text{se } \gamma \neq 0. \end{cases} \quad (2.38)$$

em que $A_t = \frac{d a(u)}{du} \Big|_{u=t} = \frac{1}{2\alpha\beta} \left[\left(\frac{t}{\beta} \right)^{-\frac{1}{2}} + \left(\frac{t}{\beta} \right)^{-\frac{3}{2}} \right]$.

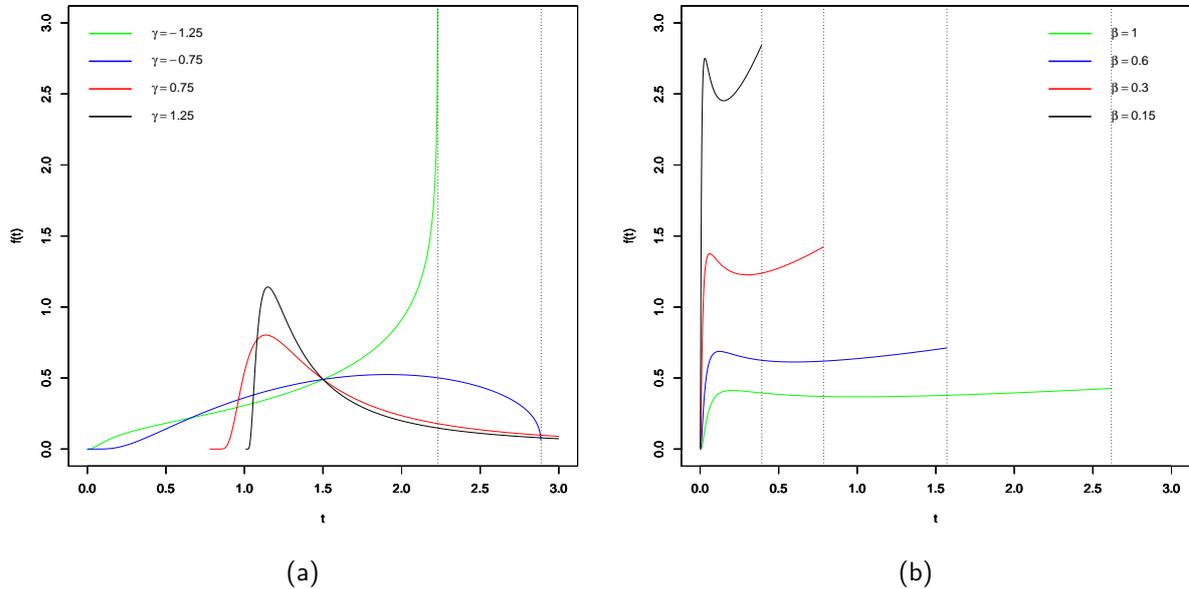
Figura 1 – Gráficos da fdp da EVBS para $\alpha \in \{1, 0.75, 0.50, 0.25\}$, $\beta = 1$ e $\gamma = 0.625$ em (a) e para $\alpha = 0.25$, $\beta \in \{1, 0.90, 0.75, 0.50\}$ e $\gamma = -0.25$ em (b)



Fonte: Elaborada pelo autor (2022)

A Figura 1 (a) apresenta gráficos de funções densidades associados às distribuições EVBS com caudas pesadas, ou seja, situações em que $\gamma > 0$, levando-nos a distribuições com limite superior do suporte infinito. Nota-se claramente um comportamento assimétrico positivo e indícios de que a distribuição é unimodal. A Figura 1 (b) ilustra as distribuições EVBS de caudas curtas ($\gamma < 0$) e limite superior do suporte finito. Particularmente, quando $\alpha = 0.25$ e $\gamma = -0.25$, nota-se uma ligeira simetria na distribuição à medida que β aproxima-se de zero; além disso, há evidências de diminuição na dispersão e no achatamento, permanecendo os indícios de unimodalidade. Na Figura 2 (a), observa-se que o parâmetro γ interfere bruscamente no comportamento da cauda da distribuição EVBS assim como no tipo de assimetria. Quando

Figura 2 – Gráficos da fdp da EVBS para $\alpha = 0.5$, $\beta = 1.5$ e $\gamma \in \{-1.25, -0.75, 0.75, 1.25\}$ em (a) e para $\alpha = 1$, $\beta \in \{1, 0.60, 0.30, 0.15\}$ e $\gamma = -1$ em (b)



Fonte: Elaborada pelo autor (2022)

$\gamma = -1.25$, há evidências que a função densidade assume valores arbitrariamente grandes próximo ao limite superior do suporte da distribuição, levando-nos a inexistência de um valor modal no interior do suporte. Por fim, a Figura 2 (b) descreve a situação na qual o parâmetro $\gamma = -1$, aqui temos comportamento distinto das distribuições anteriormente analisadas, pois a função densidade cresce e atinge um valor máximo; em seguida decresce e atinge um mínimo e, posteriormente, comporta-se como uma função crescente, porém este comportamento é suavizado à medida que o parâmetro β aproxima-se de zero.

2.3.2 A função quantílica e a geração de números aleatórios

Seja $0 < q < 1$ e t_q o quantil de ordem q na distribuição EVBS, isto é, o número real positivo tal que $F(t_q; \alpha, \beta) = q$, ou equivalentemente, $F^{-1}(t_q; \alpha, \beta) = q$. A partir da equação (2.34), constatamos que se $\gamma = 0$, t_q é a solução positiva da seguinte equação não-linear:

$$t_q + \beta^{1/2} \alpha \log(-\log(q)) t_q^{1/2} - \beta = 0. \quad (2.39)$$

Logo,

$$t_q = \beta \left(\frac{\alpha x_q}{2} + \sqrt{\left(\frac{\alpha x_q}{2} \right)^2 + 1} \right)^2. \quad (2.40)$$

Sendo, $x_q = G_{\gamma=0}^{-1}(q) = -\log(-\log(q))$ o quantil de ordem q da distribuição GEV no caso em que $\gamma = 0$.

Por outro lado, se $\gamma \neq 0$, então o quantil t_q satisfaz a equação

$$t_q - \frac{\alpha\beta^{1/2}}{\gamma} \left((-\log(q))^{-\gamma} - 1 \right) t_q^{1/2} - \beta = 0. \quad (2.41)$$

Logo,

$$t_q = \beta \left(\frac{\alpha x_q}{2} + \sqrt{\left(\frac{\alpha x_q}{2} \right)^2 + 1} \right)^2. \quad (2.42)$$

Porém, neste último caso tem-se $x_q = G_{\gamma \neq 0}^{-1}(q) = \frac{1}{\gamma} \left((-\log(q))^{-\gamma} - 1 \right)$, representando o quantil de ordem q da distribuição GEV quando $\gamma \neq 0$.

Seja T uma variável aleatória cuja fda é F . A função quantílica, associada à distribuição F , é aquela que a cada $u \in (0, 1)$ associa o número real $F^{\leftarrow}(u) = \inf\{t; F(t) \geq u\}$. No caso particular em que F é estritamente crescente, a função quantílica, F^{\leftarrow} , coincide com a inversa F^{-1} . Um dos métodos de geração de números aleatórios baseia-se na função F^{-1} e tem como fundamento o seguinte resultado: dada uma variável aleatória X com fda F , se F tem inversa F^{-1} e U é uma variável que segue uma distribuição uniforme sobre o intervalo $(0, 1)$, então $F^{-1}(U)$ e X possuem a mesma distribuição. Diante disso, tem-se o seguinte algoritmo para gerar números aleatórios com distribuição EVBS:

Algoritmo 1. Gerador de números aleatórios de uma variável

$$T \sim \text{EVBS}(\alpha, \beta, \gamma).$$

Passo 1. Gere um número aleatório, u , de uma variável $U \sim (0, 1)$.

Passo 2. Fixe os valores dos parâmetros α , β e γ .

Passo 3. Se $\gamma = 0$, considere $x_u = -\log(-\log(u))$ e calcule:

$$t = \beta \left(\frac{\alpha x_u}{2} + \sqrt{\left(\frac{\alpha x_u}{2} \right)^2 + 1} \right)^2;$$

Se $\gamma \neq 0$, considere $x_u = \frac{1}{\gamma} [(-\log(u))^{-\gamma} - 1]$ e calcule:

$$t = \beta \left(\frac{\alpha x_u}{2} + \sqrt{\left(\frac{\alpha x_u}{2} \right)^2 + 1} \right)^2.$$

Passo 4. Repita os passos 1, 2 e 3, n vezes, a fim de obter uma amostra de n números aleatórios.

2.3.3 Os momentos

Nesta subsecção investigaremos sob que suposições existem os momentos centrais da distribuição EVBS. Ao mesmo tempo, apresentaremos condições nas quais tais momentos não são finitos. A fim de atingirmos esse propósito, recorreremos à relação existente entre as distribuições EVBS e GEV, dada pela equação (2.32), e faremos uso de alguns resultados sobre a existência dos momentos da distribuição GEV.

Sabemos que se $X \sim \text{GEV}(0, 1, \gamma)$, então a variável

$$T = \beta \left(\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2} \right)^2 + 1} \right)^2, \quad (2.43)$$

segue uma distribuição EVBS(α, β, γ).

Ao desenvolvermos o binômio que figura na equação (2.43) e usarmos propriedades da função modular estabelecemos a primeira das desigualdades que figuram abaixo. A outra, advém do fato que a função raiz quadrada é estritamente crescente. Dessa forma, estabelecemos as seguintes desigualdades:

$$T \leq \frac{\beta}{2} \left(\alpha^2 X^2 + 2 + |\alpha X| \sqrt{\alpha^2 X^2 + 4} \right) < \frac{\beta}{2} \left(\alpha^2 X^2 + 2 + \sqrt{(\alpha^2 X^2)^2 + 4\alpha^2 X^2 + 4} \right).$$

Consequentemente,

$$T \leq \beta(\alpha^2 X^2 + 2). \quad (2.44)$$

De posse da desigualdade (2.44), podemos expressar uma condição suficiente para a existência da média de uma distribuição EVBS, impondo a condição de finitude do segundo momento de uma GEV. Sabemos que se $X \sim \text{GEV}(0, 1, \gamma)$ então

$$\mathbb{E}(X^2) = \begin{cases} \frac{\pi}{6} + (-\Psi(1))^2, & \text{se } \gamma = 0, \\ \frac{1}{\gamma^2} (1 + \Gamma(1 - 2\gamma) - 2\Gamma(1 - \gamma)), & \text{se } \gamma \neq 0 \text{ e } \gamma < 1/2 \\ \infty, & \text{se } \gamma \geq 1/2. \end{cases} \quad (2.45)$$

Logo, se $\gamma < \frac{1}{2}$, então $\mathbb{E}(T)$ é finita. De forma similar, obtém-se uma condição suficiente para existência do segundo momento da EVBS. Pois, se $X \sim \text{GEV}(0, 1, \gamma)$ e $\gamma < \frac{1}{4}$, então $\mathbb{E}(X^4) < \infty$. Consequentemente, decorre da desigualdade (2.44), que se $\gamma < \frac{1}{4}$, então $\mathbb{E}(T^2)$ é finita. De um modo geral, se r é um inteiro não-negativo, então $T^r \leq \beta^r \sum_{k=0}^r \binom{r}{k} \alpha^{2k} X^{2k} 2^{r-k}$. Portanto, se $\gamma < \frac{1}{2r}$, que implica $\mathbb{E}(X^{2r}) < \infty$, então $\mathbb{E}(T^r) < \infty$.

Agora, apresentemos situações em que os momentos da EVBS não existem ou são infinitos. A partir da desigualdade

$$\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2}\right)^2 + 1} > \frac{\alpha X}{2} + \left|\frac{\alpha X}{2}\right| \geq 0,$$

obtemos

$$T = \beta \left(\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2}\right)^2 + 1} \right)^2 > \beta \frac{\alpha^2}{2} (X^2 + X|X|). \quad (2.46)$$

Neste ponto, observemos que é válido o seguinte resultado: se Y é uma variável aleatória em um espaço de probabilidade $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$ e Y^2 não é integrável, então $Y^2 + Y|Y|$ não é integrável. De fato, dada uma variável aleatória Y , as partes positiva e negativa de Y são definidas por $Y^+(\omega) = \sup\{Y(\omega), 0\}$ e $Y^-(\omega) = \sup\{-Y(\omega), 0\}$, respectivamente. Além disso, sabe-se que Y é integrável se, e somente se, Y^+ e Y^- são integráveis. Logo, se Y^2 não é integrável, então $(Y^2)^+$ não é integrável ou $(Y^2)^-$ não é integrável. Ora, $(Y^2)^-$ é a função identicamente nula, que é integrável, sendo assim $(Y^2)^+$ não é integrável. Agora, observemos que

$$(Y^2 + Y|Y|)^+ = (2Y^2)^+ = 2(Y^2)^+$$

e

$$(Y^2 + Y|Y|)^- = 0.$$

Logo,

$$\int (Y^2)^+ dP = \infty \Rightarrow \int (Y^2 + Y|Y|)^+ dP = \infty \Rightarrow \int (Y^2 + Y|Y|) dP = \infty.$$

Portanto, $Y^2 + Y|Y|$ não é integrável.

Agora voltemos ao nosso contexto, sabemos que uma variável aleatória cuja distribuição é $GEV(0, 1, \gamma)$ possui segundo momento finito se, e somente se, o parâmetro $\gamma < \frac{1}{2}$. Portanto, se $\gamma \geq \frac{1}{2}$, então $\mathbb{E}(X^2) = \infty$ e, conseqüentemente, usando a desigualdade (2.46), combinada com o argumento exposto acima, tem-se $\mathbb{E}(T) = \infty$.

Um raciocínio similar ao usado na obtenção da desigualdade (2.46) permite-nos estabelecer a desigualdade

$$T^2 > \frac{1}{2} \alpha^4 \beta^2 (X^4 + X^3|X|).$$

Conseqüentemente, se $\gamma > 1/4$, então T^2 não é integrável. Os resultados obtidos acima podem ser resumidos no seguinte teorema:

Teorema 2.3 Seja T uma variável aleatória tal que $T \sim \text{EVBS}(\alpha, \beta, \gamma)$, são válidas as seguintes afirmações:

$$\mathbb{E}(T) < \infty \text{ se, e somente se, } \gamma < \frac{1}{2}. \quad (2.47)$$

$$\mathbb{E}(T^2) < \infty \text{ se, e somente se, } \gamma < \frac{1}{4}. \quad (2.48)$$

2.3.4 Domínios de atração

Nesta seção, apresentaremos alguns resultados obtidos por (FERREIRA; GOMES; LEIVA, 2012) que especificam o domínio de atração de uma variável aleatória T em termos do domínio da variável X , supondo que esta última possui distribuição pertencente a algum dos domínio de atração (Fréchet ou Weibull) e que as variáveis T e X estão relacionadas por meio igualdade

$$T = \beta \left(\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2} \right)^2 + 1} \right)^2, \quad \alpha > 0, \beta > 0. \quad (2.49)$$

Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas a uma variável X cuja fda é F . Nosso interesse recai sobre o comportamento limite do máximo parcial $M_n = \{X_1, \dots, X_n\}$, quando este for submetido a uma transformação afim positiva.

Associada à distribuição F definimos o limite superior do suporte de F , e indicamos x^* , por

$$x^* = \sup\{x : F(x) < 1\}. \quad (2.50)$$

Sem dificuldade pode-se verificar que para todo $x \in \mathbb{R}$, tem-se

$$P(M_n \leq x) = F^n(x). \quad (2.51)$$

Conseqüentemente, se $x < x^*$, então $F^n(x) \rightarrow 0$. E, se $x \geq x^*$, então $F^n(x) \rightarrow 1$, resultando em uma distribuição degenerada em x^* , quando $x^* < \infty$. Na verdade, é válida a seguinte convergência em probabilidade: $M_n \xrightarrow{P} x^*$, $n \rightarrow \infty$.

Do exposto acima, constata-se a necessidade de impormos uma transformação sobre M_n a fim obtermos uma distribuição limite não-degenerada. Suponhamos que a variável subjacente X possibilite a existência de sequências de constantes reais $a_n > 0$ e $b_n \in \mathbb{R}$ tais que

$\frac{M_n - b_n}{a_n}$ possua distribuição limite não-degenerada. Mais precisamente, admitamos que existem constantes $a_n > 0$ e $b_n \in \mathbb{R}$ tais que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x), \quad (2.52)$$

para todo ponto x no qual G é contínua, sendo G uma função de distribuição não-degenerada.

De início, interessa-nos obter a classe de funções de distribuição G que podem figurar como limite em (2.52), tal classe é denominada *distribuições de valores extremos*. Em seguida, apresentar condições necessárias e suficientes a serem satisfeitas pela distribuição F a fim de que o limite expresso na equação (2.52) seja válido. A classe de distribuições para as quais o limite (2.52) se verifica é denominada *domínio de atração de G* .

Agora, apresentamos algumas sentenças que são equivalentes ao limite expresso em (2.52).

- 1ª) $\lim_{n \rightarrow \infty} n \log F(a_n x + b_n) = \log G(x)$ para cada ponto x de continuidade de G tal que $0 < G(x) < 1$.
- 2ª) $\lim_{n \rightarrow \infty} n(1 - F(a_n x + b_n)) = -\log G(x)$ para cada ponto x de continuidade de G tal que $0 < G(x) < 1$.
- 3ª) $\lim_{n \rightarrow \infty} \frac{1}{n(1 - F(a_n x + b_n))} = -\frac{1}{\log G(x)}$ para cada ponto x de continuidade de G tal que $0 < G(x) < 1$.

Dada a F , definimos $U = \left(\frac{1}{1-F}\right)^{\leftarrow}$, a notação \leftarrow indica que U é a inversa generalizada de $\left(\frac{1}{1-F}\right)$, para mais detalhes veja a penúltima seção do Apêndice A. Assim, o limite (2.52) admite ser reescrito da seguinte forma equivalente:

$$\lim_{n \rightarrow \infty} \frac{U(nx) - b_n}{a_n} = G^{\leftarrow}(e^{-1/x}), \quad \text{para cada } x > 0. \quad (2.53)$$

As equivalências enunciadas anteriormente levam aos resultados estabelecidos no teorema abaixo.

Teorema 2.4 Sejam $a_n > 0$ e $b_n \in \mathbb{R}$ seqüências de constantes e G uma função de distribuição não-degenerada. As sentenças a seguir são equivalentes:

1. $\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = \log G(x)$ para cada ponto x no qual G é contínua.
2. $\lim_{t \rightarrow \infty} t(1 - F(a(t)x + b(t))) = -\log G(x)$ para cada ponto x de continuidade de G com $0 < G(x) < 1$, $a(t) =: a_{[t]}$ e $b(t) = b_{[t]}$, onde $[t]$ denota a parte inteira de t .

3. $\lim_{t \rightarrow \infty} \frac{U(tx) - b(t)}{a(t)} = D(x)$ para cada $x > 0$ ponto de continuidade de $D(x) = G^{\leftarrow}(e^{-1/x})$,
 $a(t) =: a_{[t]}$ e $b(t) = b_{[t]}$.

A seguir apresentaremos um resultado obtido por ((FISHER; TIPPETT, 1928)) e complementado por (GNEDENKO, 1943) que é fundamental na Teoria de Valores Extremos, pois especifica a classe das distribuições não-degeneradas que podem figurar como limite possível na equação (2.52).

Teorema 2.5 A classe das distribuições de valores extremos é $G_\gamma(ax + b)$, com $a > 0$, $b > 0$ e

$$G_\gamma(x) = \exp\left(- (1 + \gamma x)^{-\frac{1}{\gamma}}\right), \quad 1 + \gamma x > 0 \text{ e } \gamma \neq 0. \quad (2.54)$$

Se $\gamma = 0$, o lado direito da equação (2.54) é interpretado como $\exp(-e^{-x})$, $\forall x \in \mathbb{R}$. O parâmetro γ em (2.54) é denominado índice de valor extremo (EVI).

O Teorema (2.5) é um resultado fundamental na teoria de valores extremos, pois assegura que a família de distribuições limite, G_γ , possui forma explícita e é uniparamétrica, salvo por parâmetros de locação e escala. Além disso, nota-se que a classe G_γ contém distribuições com diferentes características, segundo o valor do parâmetro γ .

- i) Se $\gamma > 0$, então $G_\gamma(x) < 1$ para todo $x \in \{x : x > -\frac{1}{\gamma}\}$. Logo, G_γ possui limite superior do suporte infinito, isto é, $x_{G_\gamma} = \infty$. Ademais, quando $x \rightarrow \infty$ tem-se $1 - G_\gamma(x) \sim \gamma^{-1/\gamma} x^{-1/\gamma}$, ou seja, G_γ possui cauda direita pesada e os momentos de ordem maior ou igual a $1/\gamma$ não existem.
- ii) Se $\gamma = 0$, então $x_{G_0} = \infty$. A distribuição G_0 tem cauda direita equivalente a uma exponencial unitária, ou seja, $1 - G_0(x) \sim e^{-x}$, quando $x \rightarrow \infty$. Além disso, todos os momentos existem.
- iii) Se $\gamma < 0$, então limite superior do suporte de G_γ é finito e igual $-\frac{1}{\gamma}$. A distribuição G_γ possui cauda curta de modo que $1 - G_\gamma(-\gamma^{-1} - x) \sim (-\gamma x)^{-1/\gamma}$ quando $x \downarrow 0$.

Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias iid à variável X cuja função de distribuição é F . Diz-se que F é *máx-estável* se existem constantes $a_n > 0$ e $b_n \in \mathbb{R}$ tais que

$$P\left(\frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \leq x\right) = P(X \leq x), \quad \forall x \in \mathbb{R} \text{ e } \forall n \in \mathbb{N}. \quad (2.55)$$

Teorema 2.6 A classe das distribuições máx-estáveis coincide com a classe das distribuições de valores extremos.

O Teorema (2.5) possibilita uma reformulação do Teorema (2.4) expressa no resultado abaixo.

Teorema 2.7 Para $\gamma \in \mathbb{R}$, as afirmações a seguir são equivalentes:

1. Existem constantes reais $a_n > 0$ e $b_n \in \mathbb{R}$ tais que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_\gamma(x) = \exp\left(- (1 + \gamma x)^{-\frac{1}{\gamma}}\right), \quad (2.56)$$

para todo x , satisfazendo a condição $1 + \gamma x > 0$.

2. Existe uma função positiva $a(\cdot)$ tal que para todo $x > 0$, tem-se

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = D(x) = \frac{x^\gamma - 1}{\gamma}, \quad (2.57)$$

onde $U(x) = \left(\frac{1}{1-F(x)}\right)^\leftarrow$ e $D(x) = G^\leftarrow(e^{-1/x})$. No caso em $\gamma = 0$, o lado direito da equação acima deve ser interpretado como $\log(x)$.

3. Existe uma função positiva $a(\cdot)$ tal que

$$\lim_{t \rightarrow \infty} t(1 - F(a(t)x + U(t))) = (1 + \gamma x)^{-\frac{1}{\gamma}}, \quad (2.58)$$

para todo x , satisfazendo a desigualdade $1 + \gamma x > 0$.

4. Existe uma função positiva f tal que

$$\lim_{t \uparrow x^*} \frac{1 - F(t + xf(t))}{1 - F(t)} = (1 + \gamma x)^{-\frac{1}{\gamma}}, \quad (2.59)$$

para todo x tal que $1 + \gamma x > 0$.

Prova. Veja em (HAAN; FERREIRA, 2006), páginas 10 e 11.

Existe uma parametrização alternativa para a classe das distribuições valores extremos, G_γ , subdividindo-a em três famílias de distribuições. Mais precisamente,

- a) Se $\gamma > 0$, considere $G_\gamma\left(\frac{x-1}{\gamma}\right)$ e $\alpha = \frac{1}{\gamma} > 0$, dessa forma obtém-se

$$G_\gamma(x) = \Phi_\alpha(x) = \begin{cases} 0, & \text{se } x \leq 0, \\ \exp(-x^{-\alpha}), & \text{se } x > 0. \end{cases} \quad (2.60)$$

Esta família é denominada classe Fréchet de distribuições.

b) Para $\gamma = 0$, considere $G_0(x) = \lim_{\gamma \rightarrow 0} G_\gamma(x)$. Assim,

$$G_0(x) = \Lambda_\alpha(x) = \exp(-\exp(-x)), \quad x \in \mathbb{R}. \quad (2.61)$$

Neste caso, tem-se a distribuição Gumbel.

c) Se $\gamma < 0$, considere $G_\gamma\left(-\frac{(1+x)}{\gamma}\right)$ e $\alpha = -\frac{1}{\gamma} > 0$. Consequentemente,

$$G_\gamma(x) = \Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & \text{se } x < 0, \\ 1, & \text{se } x \geq 0. \end{cases} \quad (2.62)$$

Resultando na família Weibull reversa de distribuições, ou simplesmente, Weibull.

No Apêndice A deste trabalho exibimos alguns resultados em Teoria de Valores Extremos, levando em consideração à α -parametrização citada acima. Com atenção voltada para caracterização do máximo domínio de atração (MDA).

Agora, apresentaremos alguns resultados obtidos por (FERREIRA; GOMES; LEIVA, 2012) que especificam o domínio de atração de uma variável aleatória T . A exposição abordará apenas a situação em que se investiga o máx-domínio de atração de T . Iniciamos, supondo que a distribuição da variável X pertence ao max-domínio de atração Fréchet, ou seja, $F_X \in \text{MDA}(G_\gamma)$ com $\gamma > 0$. A condição necessária e suficiente que (FERREIRA; GOMES; LEIVA, 2012) usam na caracterização do máx-domínio de atração Fréchet é

$$F \in \text{MDA}(G_\gamma), \quad \gamma > 0 \iff \lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\frac{1}{\gamma}}, \quad (2.63)$$

para todo $x > 0$ e limite superior do suporte de X , denodado por x_F , necessariamente infinito. Essa caracterização foi obtida por (GNEDENKO, 1943) e figura em (HAAN; FERREIRA, 2006), que é referência consolidada em teoria de valores extremos.

Teorema 2.8 (FERREIRA; GOMES; LEIVA, 2012) Se uma variável aleatória X é tal que $F_X \in \text{MDA}(G_\gamma)$, com $\gamma > 0$, ou seja, a distribuição F_X pertence ao max-domínio de atração Fréchet, então a distribuição, F_T , da variável $T = \beta \left(\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2} \right)^2 + 1} \right)^2$, $\alpha > 0, \beta > 0$ também está no max-domínio de atração Fréchet com índice de valor extremo $\xi = 2\gamma$.

De modo similar, (FERREIRA; GOMES; LEIVA, 2012) especificam o domínio de atração da variável T quando X se encontra no domínio Weibull. Mais uma vez, recorrem a uma caracterização estabelecida por (GNEDENKO, 1943) para o máx-domínio de atração Weibull expressa

na seguinte equivalência:

$$F \in \text{MDA}(G_\gamma), \quad \gamma < 0 \iff \lim_{t \rightarrow \infty} \frac{1 - F(x_F - \frac{1}{tx})}{1 - F(x_F - \frac{1}{t})} = x^\gamma, \quad (2.64)$$

para todo $x > 0$ e limite superior do suporte de X , denotado por x_F , necessariamente finito.

Teorema 2.9 (FERREIRA; GOMES; LEIVA, 2012) Se uma variável aleatória X é tal que $F_X \in \text{MDA}(G_\gamma)$, com $\gamma < 0$, ou seja, a distribuição F_X pertence ao max-domínio de atração Weibull com índice de valor extremo $\gamma < 0$, então a distribuição, F_T , da variável $T = \beta \left(\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2} \right)^2 + 1} \right)^2$, $\alpha > 0, \beta > 0$ está no máx-domínio de atração Weibull com índice $\xi = \gamma$.

2.4 ESTIMAÇÃO DO ÍNDICE DE VALORES EXTREMOS

A estimação do índice de valores extremos, γ , é um tema de relevo na Estatística de Extremos. Existem na literatura diversos estimadores para γ , alguns de natureza paramétrica outros não-paramétrica. Por exemplo, o estimador não-paramétrico proposto por (HILL, 1975); o estimador paramétrico sugerido por (PICKANDS, 1975); o estimador de momentos ponderados, introduzido por (HOSKING; WALIIS; WOOD, 1985); e o estimador de momentos obtido por (DEKKERS; EINMAHL; HAAN, 1989). A estimação paramétrica de γ por máxima verossimilhança está estreitamente relacionada com a metodologia segundo a qual a amostra de máximos é obtida. Se fizermos uso do método dos máximos anuais ou do máximo por bloco (BM), o ajuste é feito utilizando a distribuição GEV. Por outro lado, caso façamos uso da metodologia POT (do termo Inglês *Peaks-Over-Threshold*), segundo a qual considera-se os valores acima de um limiar elevado, denominado *threshold*, faz-se o ajuste por meio do modelo GPD. Em ambas as situações, nos deparamos com a inconveniência de que o suporte da distribuição depende dos valores dos parâmetros, dificultando a dedução de propriedades assintóticas do estimador de máxima verossimilhança. Esse problema foi inicialmente enfrentado por (SMITH, 1985) que justifica a normalidade assintótica do EMV em certa classe não-regular de funções densidades de probabilidade; em seguida foi tratado por (DRESS H., 2004) que prova normalidade assintótica do EVM, baseando-se na distribuição dos excessos sobre um limiar alto; a existência, a consistência e a normalidade assintótica do EMV também foi objeto de estudo em (ZHOU, 2009; ZHOU, 2010). Mais recentemente, as publicações de (DOMBRY, 2015), (BUCHER; SEGERS, 2017) (DOMBRY; FERREIRA, 2019) e (ZHANG; SHABY, 2020) abordam novamente

a temática e estabelecem as propriedades assintóticas do EMV para os parâmetros da família GEV, quando a metodologia empregada para obter a amostra de máximos é a de máximos por bloco (BM, do Inglês *block method*). Fazemos uma descrição dos resultados contidos nesses últimos trabalhos a fim de que ela possa nos apresentar uma diretriz para lidarmos com estimação por máxima verossimilhança no modelo EVBS.

2.4.1 Existência e consistência do EMV para o índice de valores extremos

Nesta seção exploraremos os resultados obtidos por (DOMBRY, 2015) e (ZHANG; SHABY, 2020) relativos à existência, consistência, unicidade e à otimalidade global de um estimador de máxima verossimilhança para o parâmetro γ , índice de valores extremos (EVI, do Inglês *extreme value index*), segundo o método BM. A estimação do parâmetro EVI é de fundamental importância na Estatística de Extremos, pois γ reflete o comportamento da cauda direita da distribuição subjacente.

Em (DOMBRY, 2015) trata-se da existência e consistência do estimador de máxima verossimilhança para o parâmetro γ , segundo o método BM. Destaca-se que a técnica de máxima verossimilhança combinada com o método BM oferecem um caminho natural para estimar γ e as constantes apropriadas para padronização do máximo amostral, assumindo a condição existente no teorema de tipos extremos obtido por (FISHER; TIPPETT, 1928). Porém, chamam atenção para o fato de que uma vez que os estimadores de máxima verossimilhança são soluções de um sistema de equações para o qual não se obtém, em geral, uma solução analítica, a investigação sobre a existência e a consistência torna-se um problema delicado e de complexidade equiparada ao da normalidade assintótica.

A fda de uma variável aleatória $X \sim \text{GEV}(\mu = 0, \sigma = 1, \gamma)$, com $\gamma \neq 0$, é dada por

$$G_\gamma(x) = \exp\left(- (1 + \gamma x)^{-\frac{1}{\gamma}}\right), \quad 1 + \gamma x > 0. \quad (2.65)$$

Dizemos que uma fda F satisfaz a condição de valor extremo de primeira ordem com índice γ , ou equivalentemente, que F pertence ao MDA de G_γ , se existem constantes reais $a_m > 0$ e $b_m \in \mathbb{R}$ tais que

$$\lim_{m \rightarrow \infty} F^m(a_m x + b_m) = G_\gamma(x), \quad x \in \mathbb{R}. \quad (2.66)$$

E, usamos a notação $F \in \text{MDA}(G_\gamma)$. A condição citada é essencialmente a mesma que figura no teorema de tipos extremos. Dentre as caracterizações usadas para assegurar que

$F \in \text{MDA}(G_\gamma)$, conforme podem ser vistas em (HAAN; FERREIRA, 2006), usaremos a expressa no seguinte teorema:

Teorema 2.10 *Seja $U = \left(\frac{1}{1-F}\right)^\leftarrow$ a inversa generalizada da função $\frac{1}{1-F}$. A distribuição $F \in \text{MDA}(G_\gamma)$ se, e somente se, existe uma função $a(\cdot)$ positiva tal que*

$$\lim_{t \rightarrow +\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \quad \forall x > 0. \quad (2.67)$$

Uma escolha possível para a função $a(\cdot)$ em termos da função U é

$$a(t) = \begin{cases} \gamma U(t), & \text{se } \gamma > 0, \\ -\gamma(U(\infty) - U(t)), & \text{se } \gamma < 0, \\ U(t) - t^{-1} \int_0^t U(s) ds, & \text{se } \gamma = 0. \end{cases} \quad (2.68)$$

E para as constantes de padronização necessárias em (2.66), considera-se

$$a_m = a(m) \quad \text{e} \quad b_m = U(m). \quad (2.69)$$

O arcabouço para o uso do método dos blocos é o seguinte: consideramos uma sequência $\{X_i\}_{i \in \mathbb{N}}$ de variáveis aleatórias iid com função de distribuição comum $F \in \text{MDA}(G_{\gamma_0})$, para algum $\gamma_0 \in \mathbb{R}$, e constantes de normalização como sugeridas no Teorema 2.10. Particionamos a sequência $\{X_i\}_{i \in \mathbb{N}}$ em blocos de comprimento $m \geq 1$ e definimos o máximo no k -ésimo bloco por

$$M_{k,m} = \max\{X_{(k-1)m+1}, \dots, X_{km}\}. \quad (2.70)$$

Assim, $\frac{M_{k,m} - b_m}{a_m}$ tem função de distribuição dada por $F^m(a_m x + b_m)$ e $M_{k,m}$ terá, aproximadamente, uma distribuição valor extremo generalizada com parâmetros, $\mu = b_m$, $\sigma = a_m$ e γ_0 , que podem estimados pelo método de máxima verossimilhança.

Dada uma amostra $(M_{1,m}, \dots, M_{n,m})$, para fins estimação, a função de log-verossimilhança de $\theta = (\mu, \sigma, \gamma)$, denotada aqui por L_n , pode ser tomada na forma

$$L_n(\mu, \sigma, \gamma) = \frac{1}{n} \sum_{k=1}^n \ell_{(\mu, \sigma, \alpha)}(M_{k,n}), \quad (2.71)$$

ende $\ell_{(\mu, \sigma, \alpha)}(x) = -\left(1 + \frac{1}{\gamma}\right) \log \left[1 + \gamma \left(\frac{x-\mu}{\sigma}\right)\right] - \left(1 + \gamma \left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\gamma}} - \log \sigma$ e $1 + \gamma \left(\frac{x-\mu}{\sigma}\right) > 0$.

(DOMBRY, 2015) comenta que nem sempre a função L_n possuirá um máximo global, levando-o a definir uma noção mais fraca de estimador de máxima verossimilhança na qual se

exige apenas que o ponto seja um máximo local. Mais precisamente, diz-se que $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma}, \hat{\gamma})$ é um estimador de pseudo-máxima verossimilhança de $\boldsymbol{\theta}$ se L_n tem um máximo local em $\hat{\boldsymbol{\theta}}$.

Sabe-se que uma solução $\hat{\boldsymbol{\theta}}$ do sistema de equações

$$\nabla L_n(\boldsymbol{\theta}) = \mathbf{0}, \quad (2.72)$$

ou equivalentemente,

$$\begin{cases} \frac{\partial}{\partial \mu} L_n(\boldsymbol{\theta}) = 0, \\ \frac{\partial}{\partial \sigma} L_n(\boldsymbol{\theta}) = 0, \\ \frac{\partial}{\partial \gamma} L_n(\boldsymbol{\theta}) = 0, \end{cases} \quad (2.73)$$

para a qual a matriz Hessiana de L_n , em $\hat{\boldsymbol{\theta}}$, é definida negativa, é um estimador de máxima verossimilhança de $\boldsymbol{\theta}$. O teorema a seguir, obtido por (DOMBRY, 2015), estabelece condições que asseguram a existência e a consistência de um estimador de máxima verossimilhança para $\boldsymbol{\theta} = (\mu, \sigma, \gamma)$, segundo a metodologia dos máximos anuais ou do máximo por bloco.

Teorema 2.11 (DOMBRY, 2015) *Se $F \in MDA(G_{\gamma_0})$ com $\gamma_0 > -1$ e $\lim_{n \rightarrow \infty} \frac{m(n)}{\log(n)} = +\infty$, então existe uma sequência $(\hat{\mu}_n, \hat{\sigma}_n, \hat{\gamma}_n)$ e um inteiro aleatório $N \geq 1$ tal que*

$$P((\hat{\mu}_n, \hat{\sigma}_n, \hat{\gamma}_n) \text{ seja um EMV para todo } n \geq N) = 1. \quad (2.74)$$

Além disso,

$$\hat{\gamma}_n \xrightarrow{q.c.} \gamma_0, \quad \frac{\hat{\mu} - b_m}{a_m} \xrightarrow{q.c.} 0 \quad \text{e} \quad \frac{\hat{\sigma}}{a_m} \xrightarrow{q.c.} 1, \quad \text{quando } n \rightarrow \infty. \quad (2.75)$$

Observações:

1ª) A hipótese $\lim_{n \rightarrow \infty} \frac{m(n)}{\log(n)} = +\infty$ revela a necessidade de controle sobre o tamanho dos blocos, pois exige que o número, m , de unidades em cada um deles seja função do tamanho da amostra, n , e cresça mais rápido do que $\log(n)$.

2ª) A função de log-verossimilhança $L_n(\mu, \sigma, \gamma)$ é finita se, e somente se,

$$\min_{1 \leq k \leq n} \left(1 + \gamma \frac{M_{k,m} - \mu}{\sigma} \right) > 0.$$

Logo, qualquer estimador de máxima verossimilhança $(\hat{\mu}_n, \hat{\sigma}_n, \hat{\gamma}_n)$ deve satisfazer, quase certamente, a seguinte desigualdade:

$$\min_{1 \leq k \leq n} \left(1 + \hat{\gamma}_n \frac{M_{k,m} - \hat{\mu}_n}{\hat{\sigma}_n} \right) > 0.$$

3ª) A condição é $\gamma > -1$ se faz, de fato, necessária. Com efeito, observemos que

$$\frac{\partial}{\partial \mu} L_n(\mu, \sigma, \gamma; z) = \frac{1}{\sigma n} \sum_{k=1}^n \frac{1}{1 + \gamma z} \left(1 + \gamma - (1 + \gamma z)^{-\frac{1}{\gamma}} \right), \quad z = \frac{M_{k,m} - \mu}{\sigma} \text{ e } 1 + \gamma z > 0.$$

Conseqüentemente, se $\gamma < -1$, então $\frac{\partial}{\partial \mu} L_n(\mu, \sigma, \gamma; z) < 0$. Daí, segue que o sistema de equações (2.73) não possuirá solução real. Portanto, não existe um estimador de máxima verossimilhança no interior de Θ quando $\gamma < -1$.

As publicações de (DOMBRY, 2015), (BUCHER; SEGERS, 2017) e, mais recentemente, (ZHANG; SHABY, 2020) agregam resultados teóricos indispensáveis à Estatística de Extremos quando almejamos fazer inferência, sobre o parâmetro γ , baseada no método de máxima verossimilhança combinado com a metodologia do máximo por bloco. O primeiro trabalho, assegura a existência local e consistência de um EMV para o parâmetro γ , quando a distribuição subjacente dos dados pertence a algum domínio de atração G_{γ_0} , com $\gamma_0 > -1$. O segundo, garante normalidade assintótica do EMV fundamentando-se no conceito de diferenciabilidade em média quadrática, falaremos sobre isso na próxima seção. Em relação ao último, façamos uma breve discussão.

O trabalho (ZHANG; SHABY, 2020) complementa satisfatoriamente as duas primeiras publicações, pois os resultados teóricos contidos no artigo asseguram que o EMV local estudado por (DOMBRY, 2015) e (BUCHER; SEGERS, 2017) atinge um único máximo global da função de log-verossimilhança da distribuição GEV. (ZHANG; SHABY, 2020) supõem que o verdadeiro valor do parâmetro $\theta_0 = (\mu_0, \sigma_0, \gamma_0) \in \mathbb{R} \times (0, \infty) \times (-\frac{1}{2}, \infty)$. A hipótese $\gamma_0 > -\frac{1}{2}$ vai ao encontro do resultado obtido por (BUCHER; SEGERS, 2017) que assegura normalidade assintótica do EMV. A estratégia usada pelos autores para justificar o principal resultado do trabalho inclui dois passos cujas descrições são dadas a seguir:

1ª) Constrói-se um conjunto compacto \tilde{K} contendo $\theta_0 = (\mu_0, \sigma_0, \gamma_0)$ em seu interior e mostra-se que para todo n , suficientemente grande, a função de log-verossimilhança da distribuição GEV, denotada por $L_n(\theta)$, é estritamente côncava em \tilde{K} e atinge um ponto de máximo único;

2º) Obtém-se um outro conjunto compacto K , explicitamente definido em termos de $\theta_0 = (\mu_0, \sigma_0, \gamma_0)$, tal que $\tilde{K} \subset K$. Em seguida, mostra-se que para todo n , suficientemente grande, o máximo global de $L_n(\theta)$ é atingido em K , ou seja, $\arg \max_{\theta \in \Theta} L_n(\theta) \in K$.

Teorema 2.12 (ZHANG; SHABY, 2020) *Seja $\{X_i\}_{i \geq 1}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas a variável $X \sim \text{GEV}(\mu_0, \sigma_0, \gamma_0)$, com $\theta_0 = (\mu_0, \sigma_0, \gamma_0) \in \mathbb{R} \times (0, \infty) \times (-\frac{1}{2}, \infty)$. Se $\hat{\theta}_n$ é uma sequência local de estimadores de máxima verossimilhança cujos pontos pertencem a uma vizinhança compacta de θ_0 , então*

$$P \left(\exists N > 0 \text{ tal que para todo } n > N, \arg \max_{\theta \in \Theta_n} L_n(\theta) \text{ é único e igual a } \hat{\theta}_n \right) = 1, \quad (2.76)$$

onde $\Theta_n = \{\theta \in \Theta : \gamma < n - 1\}$.

2.4.2 Sobre a normalidade assintótica do EMV para o índice de valores extremos

A normalidade assintótica do EMV na família GEV é algo que foi inicialmente tratado, de forma parcial, por (SMITH, 1985). Recentemente, a temática voltou a ser discutida por (BUCHER; SEGERS, 2017) e (DOMBRY; FERREIRA, 2019), elucidando o assunto quase que plenamente. A normalidade assintótica do EMV em modelos regulares é um assunto consolidado. Mas, quando o suporte da distribuição depende dos valores dos parâmetros, violando as condições clássicas de regularidades fixadas por (CRAMÉR, 1946), a matemática necessária para fundamentar o comportamento assintótico do estimador é mais complexa. A família de distribuição GEV é um exemplo em que isso ocorre.

Levando-se em consideração o valor assumido pelo parâmetro γ , a função de distribuição de uma variável valor extremo generalizada, G_{θ} , indexada pelo parâmetro $\theta = (\mu, \sigma, \gamma) \in \mathbb{R} \times (0, \infty) \times \mathbb{R}$, pode ser escrita na forma

$$G_{\mu, \sigma, \gamma}(x) = G_{0,1,\gamma} \left(\frac{x - \mu}{\sigma} \right), \quad x \in \mathbb{R}, \quad (2.77)$$

em que

$$G_{0,1,\gamma}(z) = \begin{cases} \exp(-\exp(-z)), & \text{se } \gamma = 0, \\ \exp(-(1 + \gamma z)^{-1/\gamma}), & \text{se } \gamma \neq 0 \text{ e } 1 + \gamma z > 0, \\ 0, & \text{se } \gamma > 0 \text{ e } z \leq -1/\gamma, \\ 1, & \text{se } \gamma < 0 \text{ e } z \geq -1/\gamma. \end{cases} \quad (2.78)$$

Logo, o suporte de G_{θ} é o intervalo $S_{\theta} = \{x \in \mathbb{R} : \sigma + \gamma(x - \mu) > 0\}$, que depende de $\theta = (\mu, \sigma, \gamma)$.

O trabalho de (SMITH, 1985) foi pioneiro em lidar com a estimação por máxima verossimilhança em uma ampla classe de distribuições não-regulares. Mais precisamente, ele estudou a família uni-variada de distribuições cujas densidades podem ser escritas na forma

$$f(x; \mu, \phi) = (x - \mu)^{\alpha-1} g(x - \mu; \phi), \quad x \in (\mu, \infty), \quad (2.79)$$

em que μ é um parâmetro de locação, ϕ é um vetor de parâmetros, $\alpha = \alpha(\phi)$ é uma função diferenciável de ϕ e g é uma função conhecida. Segundo (BUCHER; SEGERS, 2017), essa formulação é geral o suficiente para incluir versões de locação das distribuições Weibull, Gamma, Beta e log-Gamma. Porém, ela não inclui, de forma plena, a distribuição GEV. Em (BUCHER; SEGERS, 2017) os autores sustentam que, até mesmo depois de uma reparametrização, a família estudada por (SMITH, 1985) inclui apenas o caso no qual o parâmetro γ pertence ao intervalo $(-1/2, 0)$. Além disso, destacam que a argumentação feita por (SMITH, 1985) apresenta algumas lacunas, pois este assume certas convergências uniformes sem justificá-las. Por fim, citam que o caso $\gamma = 0$ não é mencionado.

A contribuição dada por (BUCHER; SEGERS, 2017) vem na forma de um resultado geral sobre normalidade assintótica do EMV em um modelo cujo suporte depende dos parâmetros. Os autores fundamentam sua argumentação ao justificarem que certas relações limites são válidas uniformemente sobre subconjuntos específicos do espaço paramétrico. Paralelamente, usam a teoria dos processos empíricos, recorrendo a (VAART, 1998). Eles citam que a abordagem requer controle da entropia de certas classes de funções e que isto é feito, cuidadosamente, mediante verificação da condição de Lipschitz. O resultado obtido é aplicado sobre as famílias GEV e GPD. A seguir, apresentemos uma síntese dos resultados obtidos por (BUCHER; SEGERS, 2017).

Seja $(P_{\theta} : \theta \in \Theta)$ uma família de distribuições sobre o mesmo espaço mensurável $(\mathcal{X}, \mathcal{A})$, onde $\Theta \subset \mathbb{R}^k$. Admitamos que P_{θ} tem densidade p_{θ} com respeito a alguma medida μ . Diz-se que o modelo $(P_{\theta} : \theta \in \Theta)$ é diferenciável em média quadrática em um ponto θ_0 , pertencente ao interior de Θ , se existe uma função mensurável $\dot{\ell}_{\theta_0} : \mathcal{X} \rightarrow \mathbb{R}^k$ tal que

$$\int_{\mathcal{X}} \left\{ \sqrt{p_{\theta_0+h}} - \sqrt{p_{\theta_0}} - \frac{1}{2} h^{\top} \dot{\ell}_{\theta_0} \sqrt{p_{\theta_0}} \right\}^2 d\mu = o(\|h\|^2), \quad h \rightarrow \mathbf{0}. \quad (2.80)$$

A função $\dot{\ell}_{\theta_0}$ é o vetor escore. Admite-se que suas componentes são L_2 -integráveis com respeito a P_{θ_0} e têm média igual zero. (BUCHER; SEGERS, 2017) comentam que a diferenciabilidade em

média quadrática combinada com não-singularidade da matriz de informação de Fisher, $I_{\theta_0} = P_{\theta_0} \dot{\ell}_{\theta_0} \dot{\ell}_{\theta_0}^\top$, possibilitam uma expansão assintótica da estatística da razão de verossimilhança e acarretam a normalidade assintótica local. Para confirmação dessas consequências e outras implicações estatísticas resultantes da combinação dessas propriedades, (BUCHER; SEGERS, 2017) se reportam aos trabalhos de (VAART, 1998) e (CAM, 1986).

Sejam X_1, \dots, X_n uma amostra aleatória iid de variáveis cujas distribuições pertencem ao modelo P_{θ_0} . Por definição, qualquer máximo global, $\hat{\theta}_n$, da função $M_n: \Theta \rightarrow \overline{\mathbb{R}}$, $M_n(\theta) = \mathbb{P}_n \log(p_\theta) = \sum_{i=1}^n \log(p_\theta(X_i))$ é um estimador de máxima verossimilhança para θ_0 . Neste ponto, (BUCHER; SEGERS, 2017) assumem a hipótese de que o conjunto de maximizadores globais de M_n é não-vazio e declaram que isto é comumente satisfeito em muita situações onde Θ é um compacto em \mathbb{R}^k . Senão, pode-se usar os argumentos de compacidade presentes nos trabalhos de (VAART, 1998) ou, em vez disso, restringir a atenção a máximos locais de M_n .

Há na literatura estatística conjuntos de condições suficientes para garantir que a distribuição assintótica de $\sqrt{n}(\hat{\theta}_n - \theta_0)$ é $N_k(\mathbf{0}, I_{\theta_0}^{-1})$. Por exemplo, as condições clássicas fixadas por (CRAMÉR, 1946) requerem existência e limitação das derivadas, até a terceira ordem, da função $\theta \mapsto \ell_\theta(x) = \log p_\theta(x)$. Já as condições estabelecidas por (VAART, 1998) demandam diferenciabilidade em média quadrática do modelo no ponto θ_0 mais uma condição de Lipschitz sobre $\theta \mapsto \ell_\theta = \log p_\theta(x)$, para todo x no suporte de P_{θ_0} e todo θ em uma vizinhança de θ_0 . Entretanto, se o suporte, $S_\theta = \{x : p_\theta(x) > 0\}$, depende de θ , (BUCHER; SEGERS, 2017) afirmam que uma abordagem baseada apenas na suavidade de $\theta \mapsto \ell_\theta(x)$ não é possível, pois para qualquer vizinhança, V , de θ_0 pode-se encontrar $\theta \in V$ e $x \in \mathcal{X}$ tal que $p_{\theta_0}(x) > 0$, mas $p_\theta(x) = 0$, e daí $\ell_\theta(x) = -\infty$. Por essa mesma razão, (BUCHER; SEGERS, 2017) sustentam que os processos empíricos indexados por θ , na vizinhança de θ_0 , terão trajetórias ilimitadas com probabilidade tendendo a 1. Mas, a convergência fraca de tais processos empíricos no espaço das funções limitadas é um elemento crucial nas provas de muitos teoremas sobre comportamento assintótico de M-estimadores.

O caminho traçado por (BUCHER; SEGERS, 2017) para enfrentar a problemática inicia com a substituição da função critério ℓ_θ pela função a valores reais, m_θ , definida por

$$m_\theta(x) = 2 \log \left(\frac{p_\theta(x) + p_{\theta_0}(x)}{2p_{\theta_0}(x)} \right), \quad S_{\theta_0} = \{x : p_{\theta_0}(x) > 0\}, \quad (2.81)$$

satisfazendo $m_\theta(x) \geq -2 \log(2)$ para todo $x \in S_{\theta_0}$. Em seguida, citam resultados estabelecidos por (VAART, 1998) assegurando que a função m_θ pode ser utilizada para garantir que a razão de convergência do EMV para o parâmetro θ da GEV é da ordem de magni-

tude $O_{\mathbb{P}}(1/\sqrt{n})$. Por fim, argumentam que, uma vez estabelecida a razão de convergência em $O_{\mathbb{P}}(1/\sqrt{n})$, o principal resultado do trabalho, expresso no Teorema (2.13), apresenta condições alternativas que devem ser satisfeitas por ℓ_{θ} a fim de se possa garantir a normalidade assintótica do estimador de máxima verossimilhança.

Neste ponto, apresentemos algumas breves notações que auxiliam no manuseio das hipóteses do Teorema (2.13). Primeiro, relembremos que o suporte de P_{θ} , ou seja, o conjunto $S_{\theta} = \{x : p_{\theta}(x) > 0\} \subset \mathcal{X}$ pode depender de $\theta \in \Theta$. Usa-se a notação $U_{\varepsilon}(\theta)$ para denotar o conjunto $\{\theta' \in \Theta : \|\theta' - \theta\| < \varepsilon\}$, ou seja, a bola aberta centrada em θ e de raio ε contida $\Theta \subset \mathbb{R}^k$. Além disso, considera-se conjunto $\bar{S}(\varepsilon)$ das intersecções dos suportes de todas as distribuições em P_{θ} para todo $\theta \in U_{\varepsilon}(\theta_0)$, isto é,

$$\bar{S}(\varepsilon) = \bar{S}(\theta_0, \varepsilon) = \bigcap_{\theta \in U_{\varepsilon}(\theta_0)} \{x : p_{\theta}(x) > 0 \text{ para todo } \theta \text{ tal que } \|\theta - \theta_0\| < \varepsilon\}.$$

Agora, apresentemos o resultado obtido por (BUCHER; SEGERS, 2017) que expressa condições suficientes para normalidade assintótica do estimador de máxima verossimilhança, mesmo que o suporte da distribuição dependa dos valores paramétricos. Em seguida, exibimos outros resultados, obtidos pelos mesmos autores, que asseguram a normalidade assintótica do EMV na família GEV.

Teorema 2.13 (BUCHER; SEGERS, 2017) Dados um modelo paramétrico ($P_{\theta} : \theta \in \Theta$) sobre o espaço mensurável $(\mathcal{X}, \mathcal{A})$, com $\Theta \subset \mathbb{R}^k$, e θ_0 pertencente ao interior de Θ , sejam X_1, X_2, \dots uma sequência de variáveis aleatórias independentes e identicamente distribuídas em \mathcal{X} com distribuição comum P_{θ_0} e $\hat{\theta}_n$ um estimador de máxima verossimilhança baseado em uma amostra aleatória X_1, \dots, X_n . Nesses termos e sob as condições

H1) O modelo é diferenciável em média quadrática em θ_0 , com vetor escore $\dot{\ell}_{\theta_0}$ e matriz de informação de Fisher, I_{θ_0} , não-singular;

H2) $P_{\theta_0}(\mathcal{X} \setminus \bar{S}(\varepsilon)) = o(\varepsilon^2)$, $\varepsilon \downarrow 0$;

H3) Existe $\dot{\ell} \in L_2(P_{\theta_0})$ e $\varepsilon_0 > 0$ tal que, para qualquer $0 < \varepsilon < \varepsilon_0$, tem-se $|\ell_{\theta_1}(x) - \ell_{\theta_2}(x)| \leq \dot{\ell}(x)\|\theta_1 - \theta_2\|$ para todo $x \in \bar{S}(2\varepsilon)$ e $\theta_1, \theta_2 \in U_{\varepsilon}(\theta_0)$;

H4) $\sqrt{n}(\hat{\theta}_n - \theta_0) = O_{\mathbb{P}}(1)$ quando $n \rightarrow \infty$.

Conclui-se que

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = I_{\theta_0}^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \dot{\ell}_{\theta_0}(X_i) + o_{\mathbb{P}}(1) \rightsquigarrow \mathbf{N}_k(0, I_{\theta_0}^{-1}), \quad n \rightarrow \infty. \quad (2.82)$$

Teorema 2.14 (BUCHER; SEGERS, 2017) Seja $\{X_i\}_{i \geq 1}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas a variável $X \sim \text{GEV}(\mu_0, \sigma_0, \gamma_0)$, com $\theta_0 = (\mu_0, \sigma_0, \gamma_0) \in \mathbb{R} \times (0, \infty) \times (-1, \infty)$. Para qualquer conjunto compacto $\Theta \subset \mathbb{R} \times (0, \infty) \times (-1, \infty)$ tal que θ_0 é um ponto pertencente ao interior de Θ e para qualquer sequência de estimadores $\hat{\theta}_n$ tal que $\mathbb{P}_n \ell_{\hat{\theta}_n} = \max_{\theta \in \Theta} \mathbb{P}_n \ell_{\theta}$, supondo que o maximizador existe, conclui-se

$$\hat{\theta}_n \longrightarrow \theta_0, \text{quase certamente, quando } n \rightarrow \infty. \quad (2.83)$$

Teorema 2.15 (BUCHER; SEGERS, 2017) A família de distribuições valor extremo generalizada, $\{G_{(\mu, \sigma, \gamma)} : (\mu, \sigma, \gamma) \in \Theta_0 = \mathbb{R} \times (0, \infty) \times \mathbb{R}\}$, é diferenciável em média quadrática em $\theta_0 = (\mu_0, \sigma_0, \gamma_0)$ se, e somente se, $\gamma_0 > -\frac{1}{2}$. Neste caso, o vetor escore $\dot{\ell}_{\theta_0}(x)$ é igual ao gradiente da aplicação $\theta \mapsto \ell_{\theta}(x) = \log p_{\theta}(x)$ em $\theta = \theta_0$ para $x \in \{x : \sigma_0 + \gamma_0(x - \mu_0) > 0\}$, e igual ao vetor nulo, caso contrário.

Teorema 2.16 (BUCHER; SEGERS, 2017) Seja $\{X_i\}_{i \geq 1}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas a variável $X \sim \text{GEV}(\mu_0, \sigma_0, \gamma_0)$, com $\theta_0 = (\mu_0, \sigma_0, \gamma_0) \in \mathbb{R} \times (0, \infty) \times (-\frac{1}{2}, \infty)$. Então, para todo subconjunto compacto $\Theta \subset \mathbb{R} \times (0, \infty) \times (-\frac{1}{2}, \infty)$, qualquer sequência, $\hat{\theta}_n$ de estimadores de máxima verossimilhança definida sobre Θ é fortemente consistente e assintoticamente normal, ou seja,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = I_{\theta_0}^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \dot{\ell}_{\theta_0}(X_i) + o_{\mathbb{P}}(1) \rightsquigarrow N_3(0, I_{\theta_0}^{-1}), \quad n \rightarrow \infty. \quad (2.84)$$

2.5 ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA NA FAMÍLIA EVBS

No modelo EVBS proposto por (FERREIRA; GOMES; LEIVA, 2012) consta o parâmetro de forma, γ , proveniente da distribuição GEV, logo o processo de estimação por máxima verossimilhança do vetor paramétrico $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^{\top}$ do modelo EVBS apresentará dificuldades similares as encontradas na família GEV. Especialmente, as decorrentes do fato que o suporte dessas distribuições depende dos valores dos parâmetros, violando uma das condições de regularidades fixadas em (CRAMÉR, 1946).

Dada uma amostra de variáveis aleatórias iid, T_1, T_2, \dots, T_n , a variável $T \sim \text{EVBS}(\alpha, \beta, \gamma)$, com base nas equações (??) e (2.38), a função de log-verossimilhança de $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^{\top}$, dado

$\mathbf{t} = (t_1, t_2, \dots, t_n)^\top$, pode ser escrita na forma:

$$l(\theta; \mathbf{t}) = \begin{cases} \sum_{i=1}^n \log(A_{t_i}) - \sum_{i=1}^n \exp(-a_{t_i}) - \sum_{i=1}^n a_{t_i}, & \text{se } \gamma = 0, \\ \sum_{i=1}^n \log(A_{t_i}) - \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^n \log(1 + \gamma a_{t_i}) - \sum_{i=1}^n (1 + \gamma a_{t_i})^{-1/\gamma}, & \text{se } \gamma \neq 0. \end{cases} \quad (2.85)$$

Sendo que sobre a segunda sentença há a condição adicional $1 + \gamma a_{t_i} > 0$ para todo $i = 1, 2, \dots, n$, onde $a_{t_i} = \frac{1}{\alpha} \left(\sqrt{\frac{t_i}{\beta}} - \sqrt{\frac{\beta}{t_i}} \right)$.

No caso em que $\gamma \neq 0$, um estimador de máxima verossimilhança de $\theta = (\alpha, \beta, \gamma)^\top$, denotado por $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})^\top$, deve satisfazer, como condição necessária, ao seguinte sistema de equações não-lineares:

$$\begin{cases} -\frac{n}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n \left\{ \frac{a_{t_i}}{1 + \gamma a_{t_i}} \left[1 + \gamma - (1 + \gamma a_{t_i})^{-\frac{1}{\gamma}} \right] \right\} \Big|_{\theta = \hat{\theta}} = 0. \\ \frac{1}{2\alpha\beta} \sum_{i=1}^n \left\{ \frac{b_{t_i}}{1 + \gamma a_{t_i}} \left[1 + \gamma - (1 + \gamma a_{t_i})^{-\frac{1}{\gamma}} \right] - \frac{\alpha^2 a_{t_i}}{b_{t_i}} \right\} \Big|_{\theta = \hat{\theta}} = 0. \\ \sum_{i=1}^n \left\{ \frac{a_{t_i}}{\gamma(1 + \gamma a_{t_i})} \left[1 + \gamma - (1 + \gamma a_{t_i})^{-\frac{1}{\gamma}} \right] - \frac{\log(1 + \gamma a_{t_i})}{\gamma^2} \left[1 - (1 + \gamma a_{t_i})^{-\frac{1}{\gamma}} \right] \right\} \Big|_{\theta = \hat{\theta}} = 0. \end{cases} \quad (2.86)$$

Por outro lado, se $\gamma = 0$, então $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ deve ser uma solução do seguinte sistema:

$$\begin{cases} -\frac{n}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n a_{t_i} (1 - \exp(-a_{t_i})) \Big|_{\theta = \hat{\theta}} = 0. \\ \frac{1}{2\alpha\beta} \sum_{i=1}^n \left\{ b_{t_i} [1 - \exp(-a_{t_i})] - \frac{\alpha^2 a_{t_i}}{b_{t_i}} \right\} \Big|_{\theta = \hat{\theta}} = 0. \end{cases} \quad (2.87)$$

Sendo que $a_{t_i} = \frac{1}{\alpha} \left(\sqrt{\frac{t_i}{\beta}} - \sqrt{\frac{\beta}{t_i}} \right)$ e $b_{t_i} = \left(\sqrt{\frac{t_i}{\beta}} + \sqrt{\frac{\beta}{t_i}} \right)$.

Neste ponto, fazemos a seguinte consideração: dada a relação existente entre as distribuições EVBS e GEV, expressa pela equação (2.32), o problema de estimação no modelo EVBS terá barreiras similares àsquelas encontradas por (DOMBRY, 2015), (BUCHER; SEGERS, 2017) e (ZHANG; SHABY, 2020) quando justificaram as propriedades de consistência e normalidade assintótica para o EMV na família GEV. Logo, determinar as condições as serem satisfeitas a fim de que se possa assegurar que o EMV dos parâmetros o modelo EVBS seja consistente e assintoticamente normal é um problema que tangencia os objetivos do nosso trabalho, fazendo jus a uma investigação futura.

2.6 APLICAÇÃO

Dado que o modelo EVBS deriva do GEV, é natural conjecturar se o mesmo modela satisfatoriamente dados extremos obtidos segundo a metodologia BM. Nossa aplicação corrobora

com essa hipótese ao evidenciar um bom ajuste a um conjunto de dados meteorológicos registrados na estação Recife-Curado, código 82900, latitude -8.05917 , longitude -34.95917 e altitude 11.3 m, localizada em Recife-Pernambuco-Brasil. Os dados foram obtidos junto ao Instituto Nacional de Meteorologia (INMET), no site <https://portal.inmet.gov.br>. Nosso interesse consiste em modelar a variável aleatória que mensura a precipitação pluviométrica (mm) máxima anual, total por dia, durante o período que se inicia no ano de 1962 e se estende até 2019. Iniciamos com uma análise exploratória dos dados e culminamos com a estimação de parâmetros relevantes em Estatística de Extremos tais como quantis elevados, níveis de retorno em T-anos, probabilidade de excedência e período de retorno de níveis elevados. Obtemos estimativas provenientes de dois modelos paramétricos: o GEV, que é o candidato natural para modelagem, e o EVBS, o candidato cuja potencialidade deseja-se inspecionar.

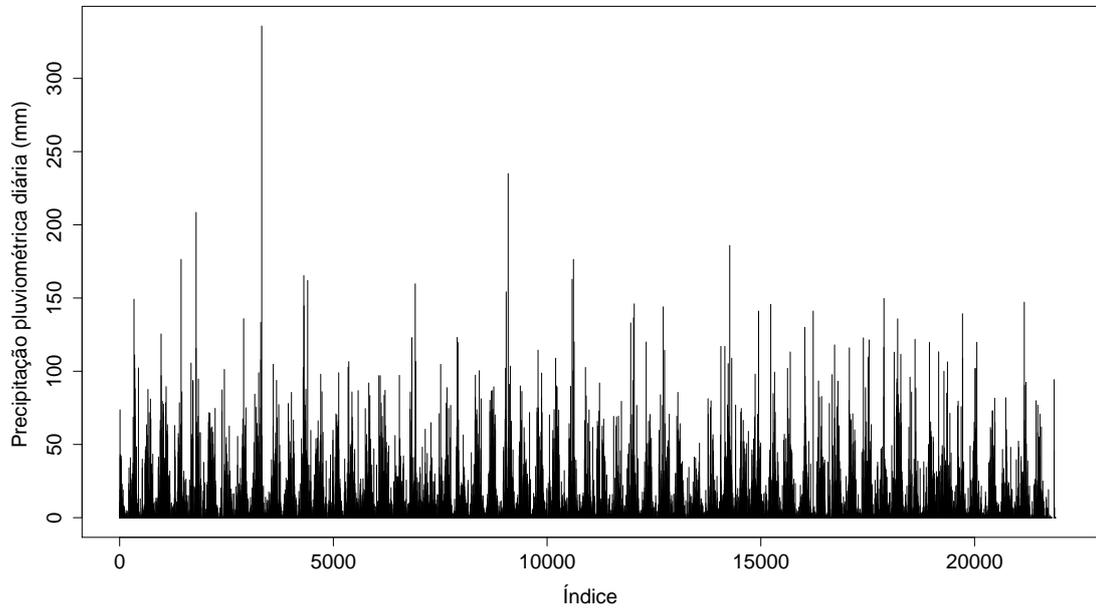
A importância de estudos sobre a variável precipitação pluviométrica reside no fato de que, uma vez obtido uma distribuição que a modele de modo satisfatório, é possível obtermos estimativas para os parâmetros fundamentais em acontecimentos extremos, possibilitando o estabelecimento de diretrizes e o planejamento de ações de longo prazo que minimizem as consequências de eventos extremos potencialmente catastróficos como chuvas intensas, inundações, deslizamentos e rompimento de barragens.

2.6.1 Análise exploratória dos dados

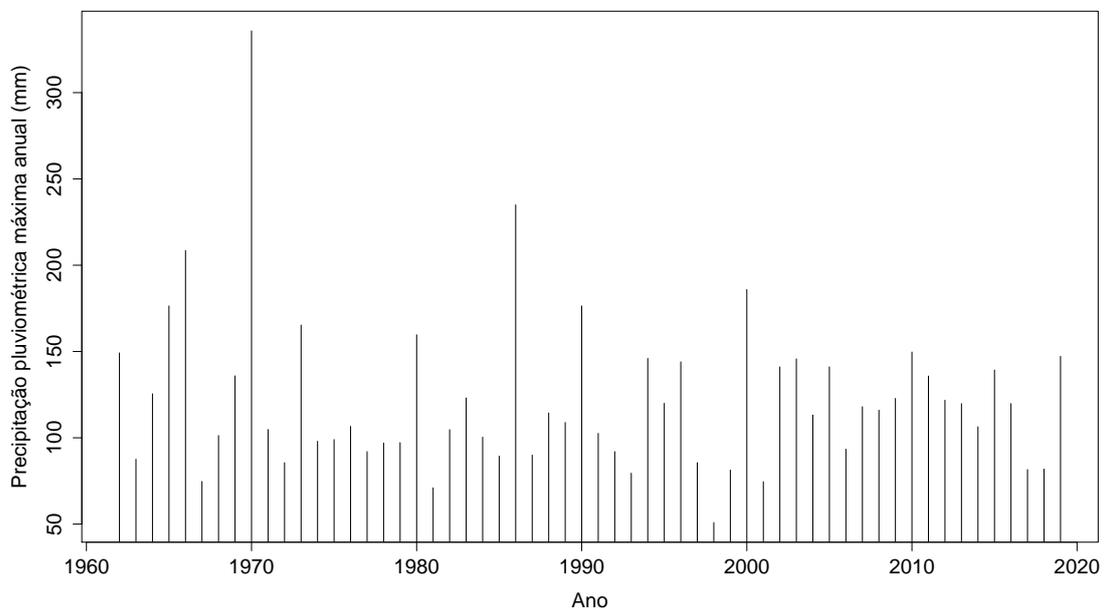
O propósito desta análise descritiva é obter informações prévias sobre o comportamento da cauda direita da distribuição subjacente associada aos dados de precipitação pluviométrica máxima anual registrados na estação Recife-Curado. Procedemos conforme as diretrizes estabelecidas em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997). Os dados de precipitações pluviométricas, total por dia, encontram-se representados na Figura 3 (a) e os correspondentes máximos anuais na Figura 3 (b), resultando em uma amostra de tamanho $n = 58$. A partir deste ponto, focaremos atenção na amostra precipitações pluviométricas máximas anuais.

Dentre as observações amostrais destaca-se a existência de um ponto cujo registro de precipitação pluviométrica diária atingiu 335.8 mm. Trata-se de um evento extremo ocorrido em 11 de agosto de 1970, noticiado pelo Diário de Pernambuco, Edição Nº 188 disponível no endereço eletrônico memoria.bn.br, que desabrigou mais de 15 mil pessoas e ocasionou a morte de outras 84. O Box-plot, Figura 4 (a), classifica tal ponto como sendo um *outlier*. Além disso, corrobora com a hipótese de que a distribuição que rege o comportamento da cauda

Figura 3 – Gráfico de dispersão das precipitações pluviométricas diárias (mm) registradas na estação Recife-Curado (1962:2019) em (a) e o correspondente gráfico das precipitações máximas anuais em (b).



(a)

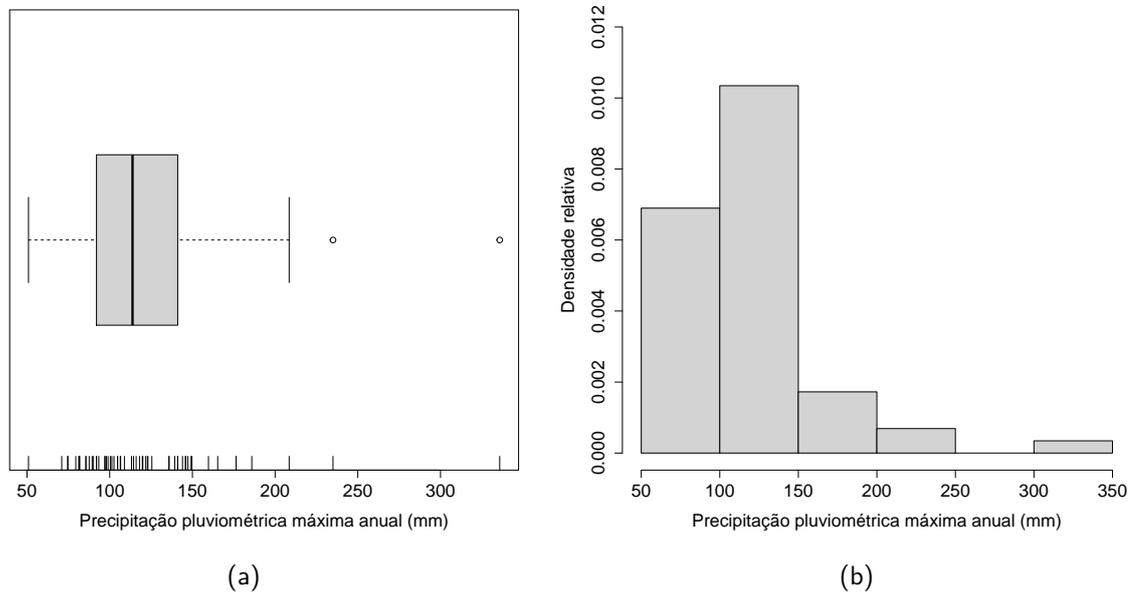


(b)

Fonte: Elaborada pelo autor (2022)

direita é assimétrica. Já o histograma, Figura 4 (b), revela indícios de que tal distribuição pertence a um máx-domínio que agrega as distribuições de cauda pesada da família GEV, talvez impulsionada pela existência do *outlier* capturado pela última classe de valores.

Figura 4 – Box-plot e histograma das precipitações pluviométricas máximas anuais registradas na estação Recife-Curado (1962:2019).



Fonte: Elaborada pelo autor (2022)

As características numéricas referentes à amostra, presentes na Tabela 1, dão suporte empírico às afirmações levantadas anteriormente. Em particular, os valores obtidos para assimetria e curtose, por meio do pacote *moments* do R, iguais a 2.17 e 10.27, respectivamente, reforçam que a distribuição possui assimetria positiva e tem curtose elevada, quando comparada com o modelo normal.

Tabela 1 – Estatísticas descritivas para os dados de precipitações pluviométricas máximas anuais (mm) na estação Recife-Curado (1962:2019).

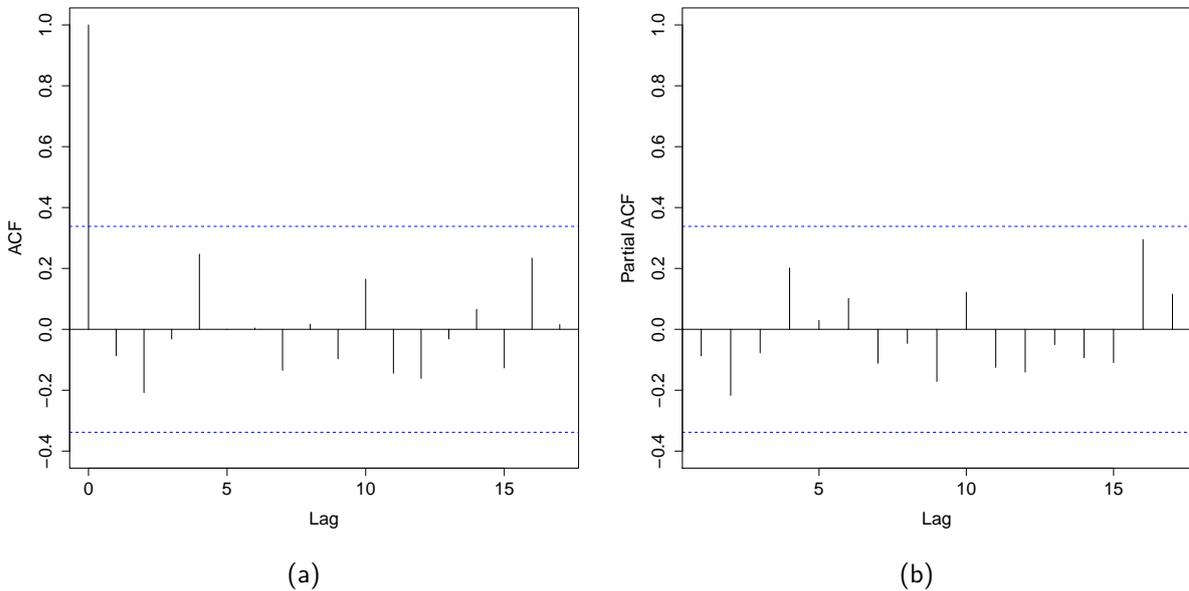
n	Mínimo	Q_1	Mediana	Média	Q_3	Máximo	Desvio	Assimetria	Curtose
58	50.90	92.35	113.80	121.90	141.10	335.80	45.29	2.17	10.27

Fonte: Elaborada pelo autor (2022)

A seguir, observemos que os gráficos das funções de autocorrelação amostral e parcial, exibidos na Figura 5, sugerem que as observações de precipitação pluviométrica máxima anual registradas na Estação Recife-Curado (1962:2019) não possuem correlação linear significativa. Diante disso, admitiremos a hipótese de independência entre as observações.

Uma distribuição natural para modelagem de valores máximos anuais é a GEV, que possui a Gumbel como um caso particular. Inicialmente, utilizaremos o modelo GEV no ajuste. Em seguida, usaremos o modelo EVBS. Antes, faremos uso de algumas ferramentas gráficas que irão nos auxiliar na implementação dos ajustes que pretendemos conduzir. Um desses ins-

Figura 5 – Gráficos das funções de autocorrelação amostral, em (a), e autocorrelação parcial, em (b), para as precipitações pluviométricas máximas anuais na estação Recife-Curado (1962:2019).



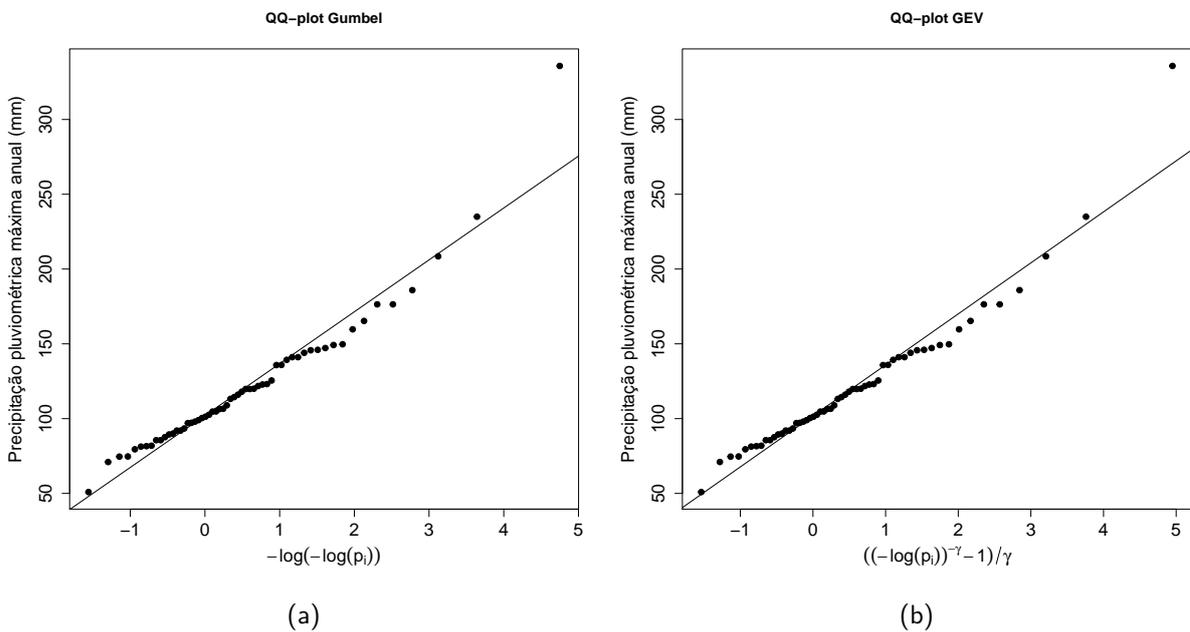
Fonte: Elaborada pelo autor (2022)

trumentos é o *quantile plot* (QQ-plot), formado pelos pontos da forma $(x_{i:n}, F^{\leftarrow}(p_i))$ onde $p_i = \frac{n-i+0.5}{n}$, $i = 1, 2, \dots, n$. Mais detalhes sobre a sua construção e propriedades podem ser encontrados em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997). Os QQ-plots ilustrados na Figura 6 revelam indícios de um ajuste satisfatório dos modelos Gumbel ou GEV à amostra de precipitações pluviométricas máximas anuais. Porém, é evidente a existência de um ponto amostral que possui distância relativamente grande da reta ajustada. Além disso, não evidências do comportamento da cauda direita, pois não se observa característica côncava ou convexa associada a nuvem de pontos em ambos os gráficos.

2.6.2 Estimativas de máxima verossimilhança sob o modelo GEV

Primeiramente, interessa-nos obter as estimativas de MV dos parâmetro de locação (μ), escala (σ) e forma (γ) do modelo GEV, sob a hipótese de que a distribuição GEV modela a precipitação pluviométrica máxima anual registrada na estação Recife-Curado (1962:2019). Em seguida, determinaremos as estimativas dos outros parâmetros interesse em Estatística de Extremos: quantis elevados, níveis de retorno, períodos de retorno e limite superior do suporte, este último somente quando $\gamma < 0$. Para tanto, utilizaremos as ferramentas disponíveis no pacote *extRemes* do R, em particular a função `fevd()`.

Figura 6 – QQ-plot Gumbel em (a) e QQ-plot GEV em (b) das precipitações pluviométricas máximas anuais na estação Recife-Curado (1962:2019).



Fonte: Elaborada pelo autor (2022)

Tabela 2 – Estimativas de MV dos parâmetros μ , σ e γ no modelo GEV, acompanhadas do respectivo erro padrão (se).

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\gamma}$
emv	102.1967	29.2958	0.0865
se	4.2353	3.1157	0.0809

Fonte: Elaborada pelo autor (2022)

Sob ajuste GEV, o sinal da estimativa de MV do parâmetro γ é positivo. Logo, baseando-se apenas no sinal desta estimativa pontual, somos levados a classificar a distribuição da precipitação pluviométrica máxima anual como sendo detentora de cauda pesada, compondo o máx-domínio de atração das distribuições Fréchet. Porém, nota-se que a estimativa do desvio padrão do EMV do parâmetro γ é relativamente grande quando comparada com a estimativa do próprio parâmetro.

Agora, determinaremos estimativas para os outros parâmetros de interesse em Estatística de Extremos. Iniciemos pela estimação de quantis elevados. No caso da distribuição GEV, os mesmos são obtidos por $q_p = \mu + \sigma \frac{((- \log(p))^{-\gamma} - 1)}{\gamma}$ para $0 < p < 1$. Ao usarmos a propriedade de invariância do EMV obtemos as estimativas expressas na Tabela 3. De imediato, nota-se que a precipitação pluviométrica máxima anual corresponde ao ano de 1970 está à direita da estimativa do quantil $q_{0.99}$, ratificando a atipicidade do evento ocorrido.

Tabela 3 – Estimativas de MV dos quantis elevados no modelo GEV para a precipitação pluviométrica máxima anual na estação Recife-Curado.

	$\hat{Q}_{0.95}$	$\hat{Q}_{0.99}$	$\hat{Q}_{0.999}$	$\hat{Q}_{0.9999}$
Estimativas	201.4068	267.7034	379.0486	514.7137

Fonte: Elaborada pelo autor (2022)

A seguir, passemos a estimação dos níveis de retorno. O nível de retorno T -anos, denotado por $U(T)$, é o quantil correspondente à probabilidade $1 - \frac{1}{T}$. No caso sob estudo, ele pode ser interpretado como a precipitação pluviométrica (mm) que é excedida, em média, uma vez a cada T anos. Antes, ressaltamos que a função `is.fixedfevd()`, presente no pacote `extRemes` do R, quando aplicada ao ajuste por meio do modelo GEV, indica estacionariedade do mesmo. A Tabela 4 exibe as estimativas de MV dos níveis de retorno para 10, 20, 50, 100 e 200 anos. De onde se infere que aproximadamente 175 mm é o nível de retorno que é excedido, em média uma vez, a cada período de 10 anos. Observemos que este nível foi superado no ano de 2000, mais precisamente no 01/08/2000 foi registrada uma precipitação pluviométrica diária igual 185.9 mm, levando as autoridades competentes a decretar situação de emergência no município do Recife (Decreto 18.595, de 1º de agosto de 2000). Se o período é alongado para um século, o nível de retorno é aproximadamente 267.7 mm, sendo 1986 a última vez que este nível foi excedido.

Tabela 4 – Estimativas de MV dos níveis de retorno no modelo GEV para a precipitação pluviométrica máxima anual na estação Recife-Curado.

	$\hat{U}(10)$	$\hat{U}(20)$	$\hat{U}(50)$	$\hat{U}(100)$	$\hat{U}(200)$
Estimativas	174.9746	201.4068	238.1570	267.7034	298.9706

Fonte: Elaborada pelo autor (2022)

Dois outros parâmetros relevantes são a probabilidade de excedência de um determinado nível u e o período de retorno associado ao evento $\{X > u\}$, denotado por $EL(u)$, e definido como o número médio de anos de espera até que o nível u seja ultrapassado. Para mais detalhes acerca da definição do período de retorno associado ao nível u , veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997). Na Tabela 5 encontram-se as estimativas de MV das probabilidades de excedência associadas aos quatro maiores níveis pluviométricos registrados na estação Recife-Curado no período 1962:2019, paralelamente são exibidas as estimativas dos períodos de retorno associados aos respectivos níveis. Destaca-se que a estimativa para o período de retorno correspondente ao maior nível pluviométrico registrado em 2000 é 13.3566

anos. Já a estimativa do período de retorno associada ao maior nível pluviométrico registrado no período de 1962:2019 é aproximadamente 431 anos.

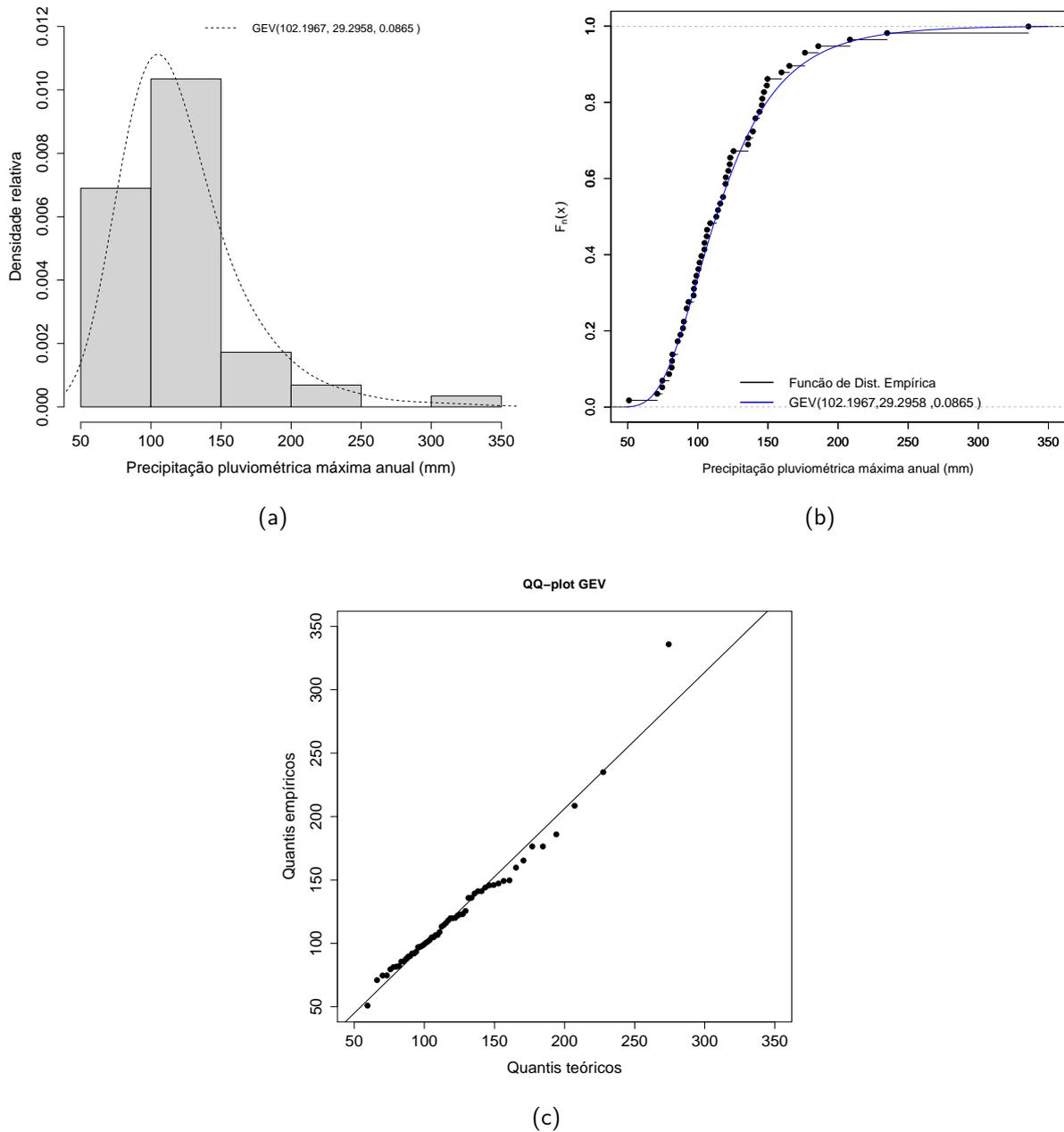
Tabela 5 – Estimativas pontuais da probabilidade de excedência do nível u e do período de retorno, em anos, associado ao nível u no modelo GEV.

Nível	$\hat{P}(X > u)$	$\hat{E}L(u)$
$u = 185.9$	0.0749	13.3566
$u = 208.5$	0.0417	23.9795
$u = 235.0$	0.0216	46.3246
$u = 335.8$	0.0023	430.9865

Fonte: Elaborada pelo autor (2022)

Por fim, apresentamos na Figura 7 os gráficos que evidenciam a qualidade do ajuste obtido por meio do modelo GEV. Particularmente, no QQ-plot, Figura 7 (c), a nuvem de pontos dispõe-se de modo satisfatório em torno da reta bissetriz, exceto o ponto correspondente ao ano de 1970.

Figura 7 – Gráficos da qualidade do ajuste por meio do modelo GEV



Fonte: Elaborada pelo autor (2022)

2.6.3 Estimativas de máxima verossimilhança sob o modelo EVBS

A Tabela 6 exibe as estimativas de MV dos parâmetros do modelo EVBS. Destaque-se que o sinal da estimativa de γ é negativo, levando-nos a classificar a distribuição da precipitação pluviométrica máxima anual na estação Recife-Curado como tendo cauda curta dentre aquelas que compõem a família EVBS, possibilitando a estimação do limite superior do suporte da distribuição.

Tabela 6 – Estimativas de MV dos parâmetros α , β e γ do modelo EVBS, acompanhadas do respectivo erro padrão.

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
emv	0.3004	101.3813	-0.1384
se	0.0285	4.3177	0.0605

Fonte: Elaborada pelo autor (2022)

O cálculo das estimativas de MV dos quantis elevados, sob o modelo EVBS, dar-se-á por meio da equação (2.40). A Tabela 7 exhibe as estimativas de alguns dos quantis. Nota-se que as estimativas dos quantis $q_{0.95}$ e $q_{0.99}$ são maiores do que àquelas obtidas sob o ajuste GEV. Por outro lado, as estimativas dos quantis $q_{0.999}$ e $q_{0.9999}$ são inferiores.

Tabela 7 – Estimativas de MV dos quantis elevados no modelo EVBS para a precipitação pluviométrica máxima anual na estação Recife-Curado.

	$\hat{q}_{0.95}$	$\hat{q}_{0.99}$	$\hat{q}_{0.999}$	$\hat{q}_{0.9999}$
Estimativas	207.5216	270.7911	354.8846	426.7782

Fonte: Elaborada pelo autor (2022)

A Tabela 8 apresenta as estimativas dos níveis de retorno para os períodos de 10, 20, 50, 100 e 200 anos obtidas pelo modelo EBVS. Nota-se que elas estão próximas daquelas provenientes do ajuste GEV.

Tabela 8 – Estimativas de MV dos níveis de retorno no modelo EVBS para a precipitação pluviométrica máxima anual na estação Recife-Curado.

	$\hat{U}(10)$	$\hat{U}(20)$	$\hat{U}(50)$	$\hat{U}(100)$	$\hat{U}(200)$
Estimativas	179.8275	207.5216	243.8053	270.7911	297.1454

Fonte: Elaborada pelo autor (2022)

Por outro lado, as estimativas das probabilidades de excedência associadas aos 4 maiores níveis pluviométricos máximos anuais obtidas pelo ajuste EVBS, exibidas na Tabela 9, são diferentes daquelas advindas do modelo GEV. Em particular, as estimativas da probabilidade de excedência associada a maior observação amostral, 335.8 mm, estão relativamente afastadas uma da outra. Isso faz com que as estimativas para o período de retorno sejam distintas. De modo que a diferença absoluta entre as estimativas do período de retorno associado ao nível 335.8mm é de aproximadamente 147 anos. Mas, acentuamos que tanto o QQ-plot GEV, Figura 8(c), quanto o QQ-plot EVBS, Figura 9(c) não acomodam satisfatoriamente o ponto associado a esta observação.

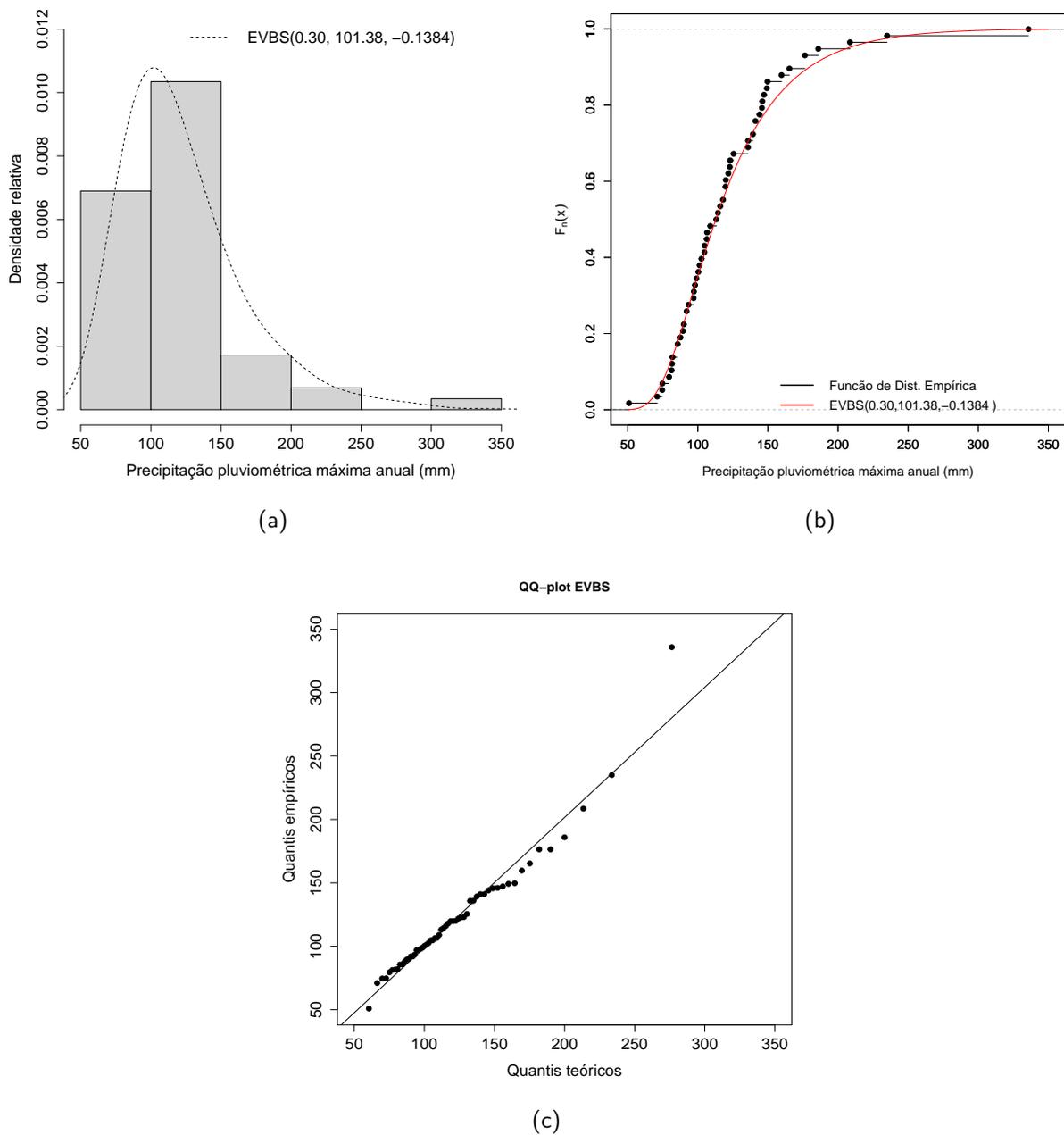
Tabela 9 – Estimativas de MV da probabilidade de excedência do nível u e do período de retorno, em anos, associado ao nível u no modelo EVBS.

Níveis	$\hat{P}(X > u)$	$\hat{E}L(u)$
$u = 185.9$	0.0859	11.6374
$u = 208.5$	0.0488	20.4974
$u = 235.0$	0.0250	39.9812
$u = 335.8$	0.0017	577.3839

Fonte: Elaborada pelo autor (2022)

Por fim, destaquemos que, para o caso sob estudo, o ajuste por meio do modelo EVBS possibilitou a obtenção de uma estimativa para o limite superior do suporte da distribuição, dada por $\hat{x}^* = 665.4258$ mm. Sendo assim, o valor real 665.4258 é uma estimativa para o supremo dos níveis pluviométricos máximos anuais (mm), por dia, a serem registrados estação Recife-Curado.

Figura 8 – Gráficos da qualidade do ajuste por meio do modelo EVBS



Fonte: Elaborada pelo autor (2022)

2.7 UM MODELO DE REGRESSÃO PARA DISTRIBUIÇÃO EVBS

Credita-se a (LEIVA et al., 2016) a proposta de um modelo de regressão associado à distribuição EVBS. Na formulação do modelo, eles seguem um caminho análogo ao trilhado por (RIECK; NEDELMAN, 1991). Nesta seção, apresentaremos o modelo de regressão Birnbaum-Saunders valor extremo, deduziremos algumas das propriedades correspondentes à distribuição do erro associado ao modelo, lidaremos com o problema de estimação via máxima verossimilhança e

realizaremos um estudo de simulação com o propósito de obter indícios que corroboraram com a hipótese de consistência para o estimador de máxima verossimilhança.

2.7.1 Caracterização

A fim de obter um modelo de regressão para distribuição EVBS, (LEIVA et al., 2016) fazem uso de ideias similares às usadas por (RIECK; NEDELMAN, 1991). Em termos mais precisos, consideram o modelo

$$T_i = \varrho_i \delta_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}) \delta_i, \quad i = 1, \dots, n, \quad (2.88)$$

onde T_i é a variável resposta, $\mathbf{x}_i^\top = (x_{i0}, x_{i1}, \dots, x_{i(p-1)})$ são observações fixas de p variáveis explicativas, $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^\top$ é um vetor de parâmetros desconhecidos, que apresenta os coeficientes a serem estimados, e $\delta_i \sim \text{EVBS}(\alpha, 1, \gamma)$ é a componente aleatória do modelo.

Após considerarem o logaritmo natural da variável T_i , denotando-o por Y_i , e definirem $\varepsilon_i = \log(\delta_i)$, obtiveram o modelo

$$Y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \dots, n, \quad (2.89)$$

denominando-o de modelo de regressão Birnbaum-Saunders valor extremo.

A seguir, apresentamos algumas características do modelo (2.88), comparando-o com o modelo (2.25):

- i) O preditor $\eta_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta})$, no modelo (2.25), é a mediana da variável T_i para $i = 1, \dots, n$. Isso não se verifica para o preditor $\varrho_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta})$ do modelo (2.88);
- ii) Em ambos os modelos, admite-se que o parâmetro α independe do vetor de variáveis explicativas $\mathbf{x}_i^\top = (x_{i0}, x_{i1}, \dots, x_{i(p-1)})$. No modelo (2.88), há a hipótese adicional de que o parâmetro γ detém essa mesma propriedade em relação ao vetor de covariáveis;
- iii) Há o pressuposto de homocedasticidade em ambos os modelos. Além disso, admite-se que os erros são independentes;
- iv) No modelo (2.25), verifica-se que $\nu_i = \log(\zeta_i) \sim \text{SN}(\alpha, 0, 2)$ para todo $i = 1, \dots, n$. Isso não ocorre com a distribuição dos erros, $\varepsilon_i = \log(\delta_i)$, no modelo (2.88).

Se usarmos a hipótese $\delta_i \sim \text{EVBS}(\alpha, 1, \gamma)$, concluímos que, para cada $i = 1, 2, \dots, n$, a variável Y_i presente no modelo (2.88) possui fdp dada por

$$f_{Y_i}(y_i) = \begin{cases} \frac{1}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right) \exp\left[-\frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right) - \exp\left(-\frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)\right)\right]; \\ y_i \in \mathbb{R}, \text{ se } \gamma = 0. \\ \frac{1}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right) \left(1 + \frac{2\gamma}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)\right)^{-1 - \frac{1}{\gamma}} \exp\left[-\left(1 + \frac{2\gamma}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)\right)^{-\frac{1}{\gamma}}\right]; \\ y_i \in B, \text{ se } \gamma \neq 0. \end{cases} \quad (2.90)$$

Sendo que o conjunto B depende dos valores dos parâmetros $\alpha, \eta_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ e γ . Mais precisamente, se $\gamma > 0$, então $B = \{y_i \in \mathbb{R}; y_i > \eta_i + 2\text{arcsenh}(-\alpha/2\gamma)\}$. Por outro lado, se $\gamma < 0$, então $B = \{y_i \in \mathbb{R}; y_i < \eta_i + 2\text{arcsenh}(-\alpha/2\gamma)\}$. Neste caso, dizemos que Y_i segue uma distribuição log-EVBS com parâmetros $\alpha > 0, \eta_i \in \mathbb{R}$ e $\gamma \in \mathbb{R}$, e denotamos $Y_i \sim \text{log-EVBS}(\alpha, \eta_i, \gamma)$.

2.7.2 Propriedades distribuição log-EVBS

Neste subseção, agregamos algumas contribuições ao modelo de regressão EVBS expressas como propriedades da distribuição log-EVBS. Se uma variável $Y \sim \text{log-EVBS}(\alpha, \eta, \gamma)$, ao considerarmos $\xi_1(y) = \frac{2}{\alpha} \cosh\left(\frac{y-\eta}{2}\right)$ e $\xi_2(y) = \frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)$, então a fdp de Y admite ser escrita na seguinte forma:

$$f(y) = \begin{cases} \frac{1}{2} \xi_1(y) \exp(-\xi_2(y) - \exp(-\xi_2(y))); & y \in \mathbb{R}, \quad \text{se } \gamma = 0. \\ \frac{1}{2} \xi_1(y) (1 + \gamma \xi_2(y))^{-1 - \frac{1}{\gamma}} \exp(-(1 + \gamma \xi_2(y))^{-\frac{1}{\gamma}}); & 1 + \gamma \xi_2(y) > 0, \quad \text{se } \gamma \neq 0. \end{cases} \quad (2.91)$$

Agora, citemos algumas propriedades da fdp da variável $Y \sim \text{log-EVBS}(\alpha, \eta, \gamma)$.

Propriedade 1. Seja f a função densidade de uma variável que segue uma distribuição log-EVBS com parâmetros $\alpha > 0, \eta \in \mathbb{R}$ e $\gamma = 0$. Se $0 < \alpha < 2$, então η é um máximo local da função f . Além disso, f é estritamente crescente no intervalo $(-\infty, \eta)$. Por outro lado, se $\alpha > 2$, então η é um mínimo local de f .

Propriedade 2. Se f é a função densidade de uma variável $Y \sim \text{log-EVBS}(\alpha, \eta, \gamma)$ e o parâmetro $\gamma < -1$, então f é ilimitada superiormente em $(-\infty, \eta + 2\text{arcsenh}(-\frac{\alpha}{2\gamma}))$.

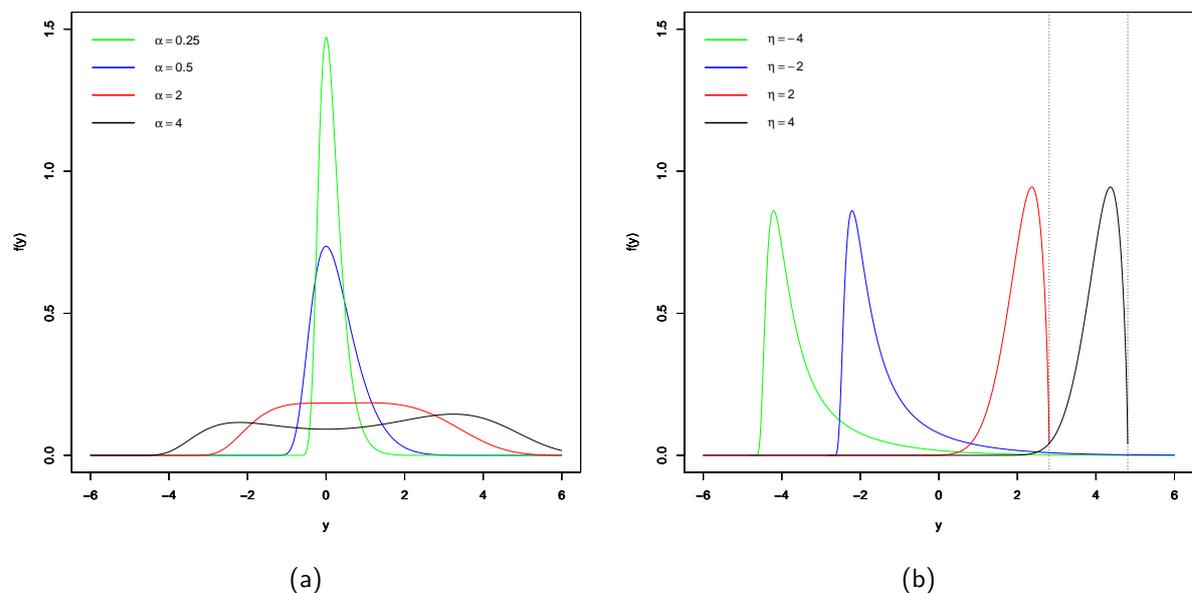
Propriedade 3. Seja f a densidade de uma variável $Y \sim \text{log-EVBS}(\alpha, \eta, \gamma = -1)$. Se $\alpha \geq 4$, então f é crescente nos intervalos $(-\infty, \eta + 2\text{arcsenh}(-\frac{\alpha}{4} - \frac{1}{4}\sqrt{\alpha^2 - 16}))$ e

$(\eta + 2\operatorname{arcsenh}(-\frac{\alpha}{4} + \frac{1}{4}\sqrt{\alpha^2 - 16}), \eta + 2\operatorname{arcsinh}(-\alpha/2\gamma))$. E, decrescente no intervalo $(\eta + 2\operatorname{arcsenh}(-\frac{\alpha}{4} - \frac{1}{4}\sqrt{\alpha^2 - 16}), \eta + 2\operatorname{arcsenh}(-\frac{\alpha}{4} + \frac{1}{4}\sqrt{\alpha^2 - 16}))$. Por outro lado, se $0 < \alpha < 4$, então f é estritamente crescente em $(-\infty, \eta + 2\operatorname{arcsenh}(\frac{\alpha}{2}))$.

Propriedade 4 Se $X \sim \operatorname{GEV}(\mu = 0, \sigma = 1, \gamma)$, então para $\eta \in \mathbb{R}$ e $\alpha > 0$ a variável $Y = \eta + 2\operatorname{arcsenh}(\frac{\alpha X}{2}) \sim \operatorname{log-EVBS}(\alpha, \eta, \gamma)$.

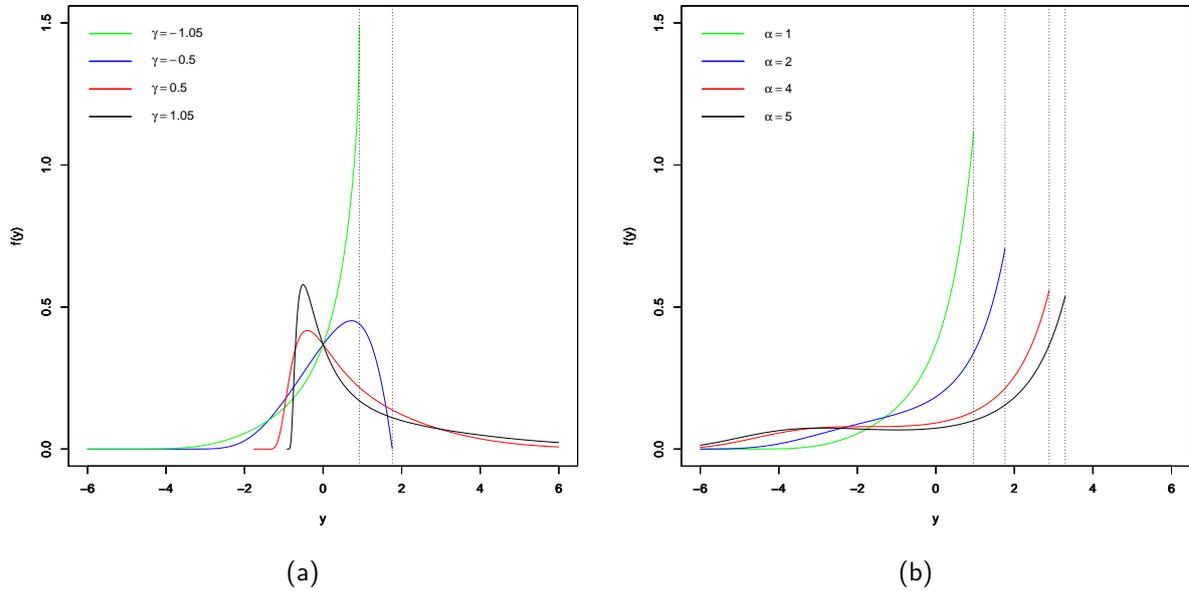
A seguir, apresentaremos algumas ilustrações que descrevem o comportamento de uma função densidade da família de distribuições log-EVBS. Primeiramente, na Figura 9 (a) constata-se evidências de que a distribuição é unimodal quando $\gamma = 0$ e α aproxima-se de zero; mas há indícios claros de bimodalidade quando $\alpha > 2$, indo ao encontro do resultado contido na Propriedade 1. A Figura 9 (b) descreve o fato de que η é um parâmetro de locação; além disso dá indícios de assimetria positiva e limite superior do suporte infinito no caso em que $\gamma > 0$; ao mesmo tempo dá evidências de assimetria negativa e limite superior do suporte finito quando $-1 < \gamma < 0$. Por fim, na Figura 10 (a) destaca-se o fato da densidade ser ilimitada quando $\gamma < -1$; paralelamente, a Figura 10 (b) reforça os indícios de não-unimodalidade da distribuição quando $\alpha \geq 4$, corroborando com o resultado citado na Propriedade 3.

Figura 9 – Gráficos da fdp da log-EVBS para $\alpha \in \{0.25, 0.50, 2, 4\}$, $\eta = 0$ e $\gamma = 0$ em (a) e para $\alpha = 0.5$, $\eta \in \{-4, -2, 2, 4\}$ e $\gamma \in \{-0.60, 0.60\}$ em (b).



Fonte: Elaborada pelo autor (2022)

Figura 10 – Gráficos da fdp da log-EVBS para $\alpha = 1$, $\eta = 0$ e $\gamma \in \{-1.05, -0.5, 0.5, 1.05\}$ em (a) e para $\alpha \in \{1, 2, 4, 5\}$, $\eta = 0$ e $\gamma = -1$ em (b).



Fonte: Elaborada pelo autor (2022)

2.7.3 Estimação no modelo de regressão EVBS

CASO: $\gamma \neq 0$. Dada uma amostra $\mathbf{y} = (y_1, y_2, \dots, y_n)^\top$ de observações independentes associadas ao modelo (2.89), a função de log-verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha, \gamma)^\top$ pode ser expressa como

$$l(\boldsymbol{\theta}; \mathbf{y}) = -n \log 2 + \sum_{i=1}^n \log(\xi_{i1}) - \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^n \log(1 + \gamma \xi_{i2}) - \sum_{i=1}^n (1 + \gamma \xi_{i2})^{-1/\gamma}, \quad (2.92)$$

onde $\xi_{i1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$ e $\xi_{i2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$, para cada $i = 1, 2, \dots, n$.

Consequentemente, obtemos as seguintes derivadas parciais:

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \beta_j} l(\boldsymbol{\theta}; \mathbf{y}) = \frac{1}{2} \sum_{i=1}^n x_{ij} \left[\frac{\xi_{i1}}{1 + \gamma \xi_{i2}} \left(1 + \gamma - (1 + \gamma \xi_{i2})^{-1/\gamma}\right) - \frac{\xi_{i2}}{\xi_{i1}} \right], \quad j = 0, 1, \dots, p-1. \\ \frac{\partial}{\partial \alpha} l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{n}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n \frac{\xi_{i2}}{1 + \gamma \xi_{i2}} \left(1 + \gamma - (1 + \gamma \xi_{i2})^{-1/\gamma}\right). \\ \frac{\partial}{\partial \gamma} l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{1}{\gamma} \sum_{i=1}^n \frac{\xi_{i2}}{1 + \gamma \xi_{i2}} \left[1 + \gamma - (1 + \gamma \xi_{i2})^{-1/\gamma}\right] + \\ \quad + \frac{1}{\gamma^2} \sum_{i=1}^n \log(1 + \gamma \xi_{i2}) \left[1 - (1 + \gamma \xi_{i2})^{-1/\gamma}\right]. \end{array} \right. \quad (2.93)$$

No caso sob análise, a função escore de θ é dada por

$$\mathbf{U}(\theta, \mathbf{y}) = \frac{\partial l(\theta, \mathbf{y})}{\partial \theta} = \begin{pmatrix} \frac{\partial}{\partial \beta} l(\theta, \mathbf{y}) \\ \frac{\partial}{\partial \alpha} l(\theta, \mathbf{y}) \\ \frac{\partial}{\partial \gamma} l(\theta, \mathbf{y}) \end{pmatrix}_{(p+2) \times 1}. \quad (2.94)$$

onde as entradas da matriz acima são as equações que compõem o sistema (2.93). Logo um estimador de MV de θ , denotado por $\hat{\theta} = (\hat{\beta}, \hat{\alpha}, \hat{\gamma})$ deve ser tal que $\mathbf{U}(\hat{\theta}) = \mathbf{0}$.

Uma vez que o sistema de equações $\mathbf{U}(\theta) = \mathbf{0}$, em geral, possui solução definida implicitamente, que não possibilita a determinação de uma solução analítica, é natural o uso de procedimentos numéricos de otimização não-linear a fim de obter as estimativas de máxima verossimilhança.

A matriz de primeiras derivadas da função escore avaliada em $\hat{\theta}$ e com sinal negativo, denotada por $\ddot{\mathbf{L}}_{\hat{\theta}}$, é dada por

$$\ddot{\mathbf{L}}_{\hat{\theta}} = -\frac{\partial}{\partial \theta} \mathbf{U}(\theta) \Big|_{\theta=\hat{\theta}} = -\frac{\partial^2}{\partial \theta \partial \theta^T} l(\theta) \Big|_{\theta=\hat{\theta}}$$

é denominada matriz de informação observada. Afim de que $\hat{\theta}$ seja um máximo local de $l(\cdot)$, as condições $\mathbf{U}(\hat{\theta}) = \mathbf{0}$ e $\ddot{\mathbf{L}}_{\hat{\theta}}$ ser definida positiva são suficientes. No caso sob estudo, temos

$$\ddot{\mathbf{L}}_{\hat{\theta}} = - \begin{pmatrix} \ddot{\mathbf{L}}_{11} & \ddot{\mathbf{L}}_{12} & \ddot{\mathbf{L}}_{13} \\ \ddot{\mathbf{L}}_{21} & \ddot{\mathbf{L}}_{22} & \ddot{\mathbf{L}}_{23} \\ \ddot{\mathbf{L}}_{31} & \ddot{\mathbf{L}}_{32} & \ddot{\mathbf{L}}_{33} \end{pmatrix}_{(p+2) \times (p+2)}. \quad (2.95)$$

onde as entradas das matrizes que compõem $\ddot{\mathbf{L}}_{\hat{\theta}}$ são dadas por:

$$\begin{aligned} \ddot{\mathbf{L}}_{11} &= \left(\frac{\partial^2 l(\theta)}{\partial \beta_k \partial \beta_j} \Big|_{\beta=\hat{\beta}, \alpha=\hat{\alpha}, \gamma=\hat{\gamma}} \right) = \frac{1}{4} \sum_{i=1}^n x_{ij} x_{ik} \left[1 - \left(\frac{\hat{\xi}_{12}}{\hat{\xi}_{i1}} \right)^2 \right] + \\ &+ \left(-\frac{1}{4} \right) \sum_{i=1}^n x_{ij} x_{ik} \left(\frac{\hat{\xi}_{i2}}{1 + \hat{\gamma} \hat{\xi}_{i2}} \right) \left[1 + \hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{i2})^{-1/\hat{\gamma}} \right] + \\ &+ \frac{1}{4} \sum_{i=1}^n x_{ij} x_{ik} \left(\frac{\hat{\xi}_{i1}}{1 + \hat{\gamma} \hat{\xi}_{i2}} \right)^2 (1 + \hat{\gamma}) \left[\hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{i2})^{-1/\hat{\gamma}} \right], \text{ onde } j, k = 0, \dots, p-1. \end{aligned}$$

$$\begin{aligned} \ddot{\mathbf{L}}_{12} &= \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \beta_j} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \alpha=\hat{\alpha}, \gamma=\hat{\gamma}} \right) = -\frac{1}{2\hat{\alpha}} \sum_{i=1}^n x_{ij} \left(\frac{\hat{\xi}_{i1}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right) \left[1 + \hat{\gamma} - (1 + \hat{\gamma}\hat{\xi}_{i2})^{-1/\hat{\gamma}} \right] + \\ &+ \frac{1}{2\hat{\alpha}} \sum_{i=1}^n x_{ij} \frac{\hat{\xi}_{i1}\hat{\xi}_{i2}}{(1 + \hat{\gamma}\hat{\xi}_{i2})^2} (1 + \hat{\gamma}) \left[\hat{\gamma} - (1 + \hat{\gamma}\hat{\xi}_{i2})^{-1/\hat{\gamma}} \right], \text{ onde } j = 0, \dots, p-1. \end{aligned}$$

$$\begin{aligned} \ddot{\mathbf{L}}_{13} &= \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma \partial \beta_j} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \alpha=\hat{\alpha}, \gamma=\hat{\gamma}} \right) = -\frac{1}{2} \sum_{i=1}^n x_{ij} \frac{\hat{\xi}_{i1}\hat{\xi}_{i2}}{(1 + \hat{\gamma}\hat{\xi}_{i2})^2} (1 + \hat{\gamma}) \left[1 - \frac{1}{\hat{\gamma}} (1 + \hat{\gamma}\hat{\xi}_{i2})^{-1/\hat{\gamma}} \right] + \\ &+ \frac{1}{2} \sum_{i=1}^n x_{ij} \left(\frac{\hat{\xi}_{i1}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right) \left[1 - \frac{1}{\hat{\gamma}^2} \log(1 + \hat{\gamma}\hat{\xi}_{i2}) \cdot (1 + \hat{\gamma}\hat{\xi}_{i2})^{-1/\hat{\gamma}} \right], j = 0, \dots, p-1. \end{aligned}$$

$$\begin{aligned} \ddot{\mathbf{L}}_{22} &= \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \alpha} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \alpha=\hat{\alpha}, \gamma=\hat{\gamma}} \right) = \frac{n}{\hat{\alpha}^2} - \frac{2}{\hat{\alpha}^2} \sum_{i=1}^n \left(\frac{\hat{\xi}_{i2}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right) \left[1 + \hat{\gamma} - (1 + \hat{\gamma}\hat{\xi}_{i2})^{-1/\hat{\gamma}} \right] + \\ &+ \frac{1}{\hat{\alpha}^2} \sum_{i=1}^n \left(\frac{\hat{\xi}_{i2}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right)^2 (1 + \hat{\gamma}) \left[\hat{\gamma} - (1 + \hat{\gamma}\hat{\xi}_{i2})^{-1/\hat{\gamma}} \right]. \end{aligned}$$

$$\begin{aligned} \ddot{\mathbf{L}}_{23} &= \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma \partial \alpha} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \alpha=\hat{\alpha}, \gamma=\hat{\gamma}} \right) = -\frac{1}{\hat{\alpha}} \sum_{i=1}^n \left(\frac{\hat{\xi}_{i2}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right)^2 (1 + \hat{\gamma}) \left[1 - \frac{1}{\hat{\gamma}} (1 + \hat{\gamma}\hat{\xi}_{i2})^{-1/\hat{\gamma}} \right] + \\ &+ \frac{1}{\hat{\alpha}} \sum_{i=1}^n \left(\frac{\hat{\xi}_{i2}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right) \left[1 - \frac{1}{\hat{\gamma}^2} \log(1 + \hat{\gamma}\hat{\xi}_{i2}) \cdot (1 + \hat{\gamma}\hat{\xi}_{i2})^{-1/\hat{\gamma}} \right]. \end{aligned}$$

$$\begin{aligned} \ddot{\mathbf{L}}_{33} &= \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma \partial \gamma} \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \alpha=\hat{\alpha}, \gamma=\hat{\gamma}} \right) = \sum_{i=1}^n \left(\frac{\hat{\xi}_{i2}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right)^2 + \\ &+ \sum_{i=1}^n \left[\frac{1}{\hat{\gamma}} \left(\frac{\hat{\xi}_{i2}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right)^2 + \frac{1}{\hat{\gamma}^2} \left(\frac{\hat{\xi}_{i2}}{1 + \hat{\gamma}\hat{\xi}_{i2}} \right) - \frac{2}{\hat{\gamma}^3} \log(1 + \hat{\gamma}\hat{\xi}_{i2}) \right] g(\hat{\gamma}) \\ &+ \sum_{i=1}^n \left(\frac{1}{\hat{\gamma}^2} \log(1 + \hat{\gamma}\hat{\xi}_{i2}) - \frac{1}{\hat{\gamma}} \frac{\hat{\xi}_{i2}}{(1 + \hat{\gamma}\hat{\xi}_{i2})} \right) g'(\hat{\gamma}). \end{aligned}$$

$$\ddot{\mathbf{L}}_{21} = \ddot{\mathbf{L}}_{12}^\top, \ddot{\mathbf{L}}_{31} = \ddot{\mathbf{L}}_{13}^\top \text{ e } \ddot{\mathbf{L}}_{32} = \ddot{\mathbf{L}}_{23}^\top. \quad (2.96)$$

Sendo que $\hat{\xi}_{i1} = \frac{2}{\hat{\alpha}} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}}{2}\right)$, $\hat{\xi}_{i2} = \frac{2}{\hat{\alpha}} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}}{2}\right)$, para cada $i = 1, 2, \dots, n$, e $g(r) = 1 - (1 + r\hat{\xi}_{i2})^{-1/r}$, $r \neq 0$.

CASO: $\gamma = 0$. Dada uma amostra $\mathbf{y} = (y_1, y_2, \dots, y_n)^\top$ de observações independentes associadas ao modelo (2.89), a função de log-verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$ pode ser expressa na forma

$$l(\boldsymbol{\theta}; \mathbf{y}) = -n \log 2 + \sum_{i=1}^n \log(\xi_{i1}) - \sum_{i=1}^n \xi_{i2} - \sum_{i=1}^n \exp(-\xi_{i2}), \quad (2.97)$$

em que $\xi_{i1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$ e $\xi_{i2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$, para cada $i = 1, 2, \dots, n$.

Uma vez que a função $l(\cdot)$, dada pela equação (2.97) é diferenciável, os possíveis estimadores de máxima verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$ devem satisfazer, como condição necessária, ao seguinte sistema de equações de log-verossimilhança:

$$\begin{cases} \frac{\partial}{\partial \beta_j} l(\boldsymbol{\theta}; \mathbf{y}) = \frac{1}{2} \sum_{i=1}^n x_{ij} \left(\xi_{i1} (1 - \exp(-\xi_{i2})) - \frac{\xi_{i2}}{\xi_{i1}} \right) = 0, & \text{para } j = 0, 1, \dots, p-1. \\ \frac{\partial}{\partial \alpha} l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{n}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^n \xi_{i2} (1 - \exp(-\xi_{i2})) = 0. \end{cases} \quad (2.98)$$

A matriz de informação observada $\ddot{\mathbf{L}}_{\hat{\boldsymbol{\theta}}}$, para o caso $\gamma = 0$, se reduz a

$$\ddot{\mathbf{L}}_{\hat{\boldsymbol{\theta}}} = - \begin{pmatrix} \ddot{\mathbf{L}}_{11} & \ddot{\mathbf{L}}_{12} \\ \ddot{\mathbf{L}}_{21} & \ddot{\mathbf{L}}_{22} \end{pmatrix}_{(p+1) \times (p+1)}. \quad (2.99)$$

onde cada matriz bloco componente é dada por

$$\begin{aligned} \ddot{\mathbf{L}}_{11} &= \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_k \partial \beta_j} \Big|_{\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}, \alpha = \hat{\alpha}} \right) = -\frac{1}{4} \sum_{i=1}^n x_{ij} x_{ik} \left(\hat{\xi}_{i1}^2 \exp(-\hat{\xi}_{i2}) \right) + \\ &\quad - \frac{1}{4} \sum_{i=1}^n x_{ij} x_{ik} \left[\hat{\xi}_{i2} (1 - \exp(-\hat{\xi}_{i2})) + 1 - \left(\frac{\hat{\xi}_{i1}}{\hat{\xi}_{i2}} \right)^2 \right] \text{ para } j, k = 0, \dots, p-1. \end{aligned}$$

$$\ddot{\mathbf{L}}_{12} = \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \beta_j} \Big|_{\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}, \alpha = \hat{\alpha}} \right) = -\frac{1}{2\hat{\alpha}} \sum_{i=1}^n x_{ij} \hat{\xi}_{i1} \left[1 - \exp(-\hat{\xi}_{i2}) \cdot (1 - \hat{\xi}_{i2}) \right] \text{ com } j = 0, \dots, p-1. \quad (2.100)$$

$$\ddot{\mathbf{L}}_{21} = \ddot{\mathbf{L}}_{12}^\top. \quad (2.101)$$

$$\ddot{\mathbf{L}}_{22} = \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha \partial \alpha} \Big|_{\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}, \alpha = \hat{\alpha}} \right) = \frac{n}{\hat{\alpha}^2} - \frac{2}{\hat{\alpha}^2} \sum_{i=1}^n \hat{\xi}_{i2} (1 - \exp(-\hat{\xi}_{i2})) - \frac{1}{\hat{\alpha}^2} \sum_{i=1}^n \hat{\xi}_{i2}^2 \exp(-\hat{\xi}_{i2}). \quad (2.102)$$

Neste momento, cabe uma consideração análoga àquela citada no final da Seção 2.5. Porém, aqui, faremos um estudo de simulação a fim de investigar se o estimador MV dos parâmetros do modelo de regressão EVBS detém a propriedade de consistência. Vimos na Subseção 2.4.1 que no modelo GEV há algumas hipóteses sobre o parâmetro γ para que a consistência do EMV seja assegurada, manteremos essas suposições, admitindo que a transformação monótona sobre a variável GEV não vai relaxá-las. Além disso, consideraremos os pressupostos sobre o parâmetro α admitidos por (RIECK; NEDELMAN, 1991) a fim de que não tenhamos multiplicidade de soluções. A razão de fazermos essas restrições é que o estudo de influência local que será desenvolvido no próximo capítulo exige, implicitamente, que o EMV seja único e consistente.

2.7.4 Simulações

Nesta subseção, faremos uma análise empírica baseada em dados simulados a fim de investigarmos o comportamento assintótico do EMV do vetor paramétrico $\theta = (\beta, \alpha, \gamma)^T$ associado ao modelo de regressão EVBS. Nas simulações, consideraremos 5.000 réplicas de Monte Carlo para diferentes tamanhos de amostras, $n \in \{60, 120, 180\}$, e almejamos identificar se existe suporte numérico que corrobora com a hipótese de que o EMV é consistente. Consideraremos o modelo de regressão no qual a resposta $T_i = \exp(\beta_0 + \beta_1 x_i) \delta_i$, $i = 1, 2, \dots, n$, onde $\delta_i \sim \text{EVBS}(\alpha, 1, \gamma)$, ou equivalentemente, $Y_i = \eta_i + \varepsilon_i$, $i = 1, \dots, n$, onde $Y_i = \log(T_i)$, $\eta_i = \exp(\beta_0 + \beta_1 x_i)$ e $\varepsilon_i \sim \text{log-EVBS}(\alpha, \eta_i, \gamma)$. Exploraremos três cenários distintos. No primeiro, investigamos a consistência do EMV; no segundo, a variável explicativa, X , cujos valores originalmente assumidos pertencem ao intervalo $(0,1)$, é submetida a uma situação de alavancagem imposta sobre 10% de seus valores; no último, a variável resposta agrega valores oriundos de distribuições EVBS distintas em que se considera diferentes valores para o parâmetro α . Em cada cenário, calculamos aproximações empíricas, baseadas nas 5.000 réplicas, para a estimativa de MV dos parâmetros, o viés, o erro padrão, a raiz quadrada do erro quadrático médio (EQM), a estimativa para o desvio padrão do estimador, oriunda da inversa da matriz de informação de Fisher observada, e a taxa de cobertura empírica para o intervalo de confiança com nível nominal teórico igual a 95%. Além disso, admitimos que o parâmetro γ , que regula o peso da cauda, assume cinco valores distintos, $-0.20, -0, 10, 0, 0.10$ e 0.20 , os dois primeiros associam distribuições de caudas curtas a componente estocástica do modelo, $\gamma = 0$ correspondente a caudas do tipo exponencial, e os dois últimos implicam em caudas

pesadas. Por fim, ressaltamos que, com o intuito de obter numericamente o ponto de máximo da log-verossimilhança, utilizamos a função `optim()`, considerando o método BFGS, contida no *software* estatístico R, por meio do gradiente e Hessiana analíticos.

2.7.4.1 Cenário 1

Neste caso, o propósito é investigar se os resultados numéricos corroboram com a hipótese de que o EMV é consistente. Os parâmetros do modelo foram fixados em $\beta_0 = 0.5$, $\beta_1 = 0.5$, $\alpha = 0.5$ e $\gamma \in \{-0.20, -0.10, 0, 0.10, 0.20\}$. Além disso, assumimos que os valores da covariável X pertencem ao intervalo unitário $(0, 1)$ e estão fixos ao longo das réplicas.

Tabela 10 – Aproximações empíricas para a estimativa de MV dos parâmetros e o viés do EMV, sob o cenário 1.

Estimativa de MV						Viés					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.508	0.505	0.488	-0.220	60	-0.20	0.008	0.005	-0.012	-0.020
	-0.10	0.508	0.503	0.486	-0.111		-0.10	0.008	0.003	-0.014	-0.011
	0	0.508	0.501	0.485	-0.004		0	0.008	0.001	-0.015	-0.004
	0.10	0.507	0.499	0.484	0.103		0.10	0.007	-0.001	-0.016	0.003
	0.20	0.507	0.497	0.483	0.209		0.20	0.007	-0.003	-0.017	0.009
120	-0.20	0.506	0.500	0.495	-0.211	120	-0.20	0.006	0.000	-0.005	-0.011
	-0.10	0.505	0.500	0.494	-0.106		-0.10	0.005	0.000	-0.006	-0.006
	0	0.504	0.500	0.493	-0.002		0	0.004	0.000	-0.007	-0.002
	0.10	0.504	0.499	0.493	0.101		0.10	0.004	-0.001	-0.007	0.001
	0.20	0.503	0.499	0.492	0.204		0.20	0.003	-0.001	-0.008	0.004
180	-0.20	0.504	0.500	0.496	-0.207	180	-0.20	0.004	0.000	-0.004	-0.007
	-0.10	0.503	0.500	0.496	-0.104		-0.10	0.003	0.000	-0.004	-0.004
	0	0.502	0.501	0.495	-0.001		0	0.002	0.001	-0.005	-0.001
	0.10	0.502	0.501	0.495	0.101		0.10	0.002	0.001	-0.005	0.001
	0.20	0.501	0.501	0.495	0.203		0.20	0.001	0.001	-0.005	0.003

Fonte: Elaborada pelo autor (2022)

As três primeiras tabelas agregam os resultados obtidos sob o cenário 1. Observemos que na Tabela 10 existem claras evidências de que o EMV é assintoticamente não-viesado, pois há indícios de que o viés do estimador tende a zero à medida que o tamanho da amostra cresce. A partir dos dados presentes na Tabela 11, nota-se que o erro padrão do EMV também aproxima-se que zero quando a amostra aumenta de tamanho. Essas observações, combinadas com o fato de que a distribuição EVBS possui segundo momento finito quando $\gamma < \frac{1}{4}$, corroboram com a hipótese de que o EMV dos parâmetros do modelo de regressão EVBS é

Tabela 11 – Aproximações empíricas para o erro padrão e a raiz quadrada do erro quadrático médio do EMV, sob o cenário 1.

Erro padrão						Raiz do EQM					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.152	0.225	0.052	0.110	60	-0.20	0.152	0.225	0.054	0.112
	-0.10	0.153	0.226	0.053	0.112		-0.10	0.153	0.226	0.055	0.112
	0	0.151	0.221	0.054	0.116		0	0.151	0.221	0.056	0.116
	0.10	0.146	0.212	0.057	0.122		0.10	0.146	0.212	0.059	0.122
	0.20	0.141	0.201	0.059	0.129		0.20	0.141	0.201	0.062	0.129
120	-0.20	0.093	0.150	0.036	0.066	120	-0.20	0.093	0.150	0.036	0.067
	-0.10	0.093	0.151	0.036	0.069		-0.10	0.094	0.151	0.037	0.070
	0	0.092	0.148	0.037	0.073		0	0.092	0.148	0.038	0.073
	0.10	0.090	0.142	0.039	0.078		0.10	0.090	0.142	0.039	0.078
	0.20	0.086	0.134	0.041	0.083		0.20	0.086	0.134	0.042	0.083
180	-0.20	0.074	0.124	0.029	0.052	180	-0.20	0.074	0.124	0.029	0.052
	-0.10	0.075	0.126	0.029	0.055		-0.10	0.075	0.126	0.029	0.055
	0	0.074	0.123	0.030	0.058		0	0.074	0.123	0.030	0.058
	0.10	0.072	0.118	0.031	0.062		0.10	0.072	0.118	0.032	0.062
	0.20	0.069	0.111	0.033	0.066		0.20	0.069	0.111	0.033	0.066

Fonte: Elaborada pelo autor (2022)

Tabela 12 – Aproximações empíricas para o erro padrão, oriundo da matriz de informação de Fisher observada (IFO), e a taxa de cobertura do intervalo de confiança, sob o cenário 1.

Erro Padrão-IFO						Taxa de cobertura					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.142	0.209	0.051	0.099	60	-0.20	0.927	0.922	0.918	0.924
	-0.10	0.143	0.212	0.051	0.103		-0.10	0.928	0.929	0.915	0.926
	0	0.141	0.208	0.053	0.108		0	0.926	0.926	0.913	0.933
	0.10	0.136	0.198	0.055	0.114		0.10	0.927	0.923	0.910	0.936
	0.20	0.131	0.186	0.058	0.122		0.20	0.923	0.920	0.909	0.938
120	-0.20	0.092	0.148	0.036	0.063	120	-0.20	0.944	0.943	0.941	0.935
	-0.10	0.093	0.150	0.036	0.067		-0.10	0.945	0.945	0.938	0.937
	0	0.092	0.147	0.037	0.071		0	0.947	0.944	0.933	0.939
	0.10	0.089	0.140	0.039	0.076		0.10	0.946	0.943	0.935	0.940
	0.20	0.085	0.131	0.041	0.081		0.20	0.946	0.940	0.934	0.940
180	-0.20	0.073	0.123	0.029	0.050	180	-0.20	0.946	0.947	0.947	0.942
	-0.10	0.074	0.125	0.029	0.053		-0.10	0.946	0.944	0.943	0.943
	0	0.073	0.122	0.030	0.056		0	0.945	0.944	0.940	0.944
	0.10	0.071	0.116	0.032	0.060		0.10	0.944	0.944	0.940	0.944
	0.20	0.068	0.109	0.033	0.064		0.20	0.944	0.940	0.939	0.943

Fonte: Elaborada pelo autor (2022)

consistente, mantidas as restrições análogas àquelas impostas sobre o modelo GEV. Além disso, os resultados contidos na Tabela 12 dão mais suporte numérico a hipótese de consistência do

EMV, pois as estimativas para o desvio padrão do EMV, capturadas na inversa matriz de informação de Fisher observada, tende a zero à medida que n cresce; paralelamente notemos que as taxas de coberturas empíricas para o intervalos de confiança dos parâmetros, calculadas sob a hipótese de normalidade para o EMV, se aproximam do nível nominal teórico, fixado em 95%.

2.7.4.2 Cenário 2

Diferentemente do primeiro caso, aqui o cenário é caracterizado por uma alavancagem imposta sobre os valores da variável explicativa, X , de modo que 10% de seus valores são fixados de forma aleatória no intervalo (5,10). Os demais parâmetros estão fixos em $\beta_0 = 0.5$, $\beta_1 = 0.5$, $\alpha = 0.5$ e $\gamma \in \{-0.20, -0.10, 0, 0.10, 0.20\}$. O propósito é investigar o comportamento das estimativas em situações nas quais valores extremos para variável resposta são induzidos por alavancagem nos valores das covariáveis.

Tabela 13 – Aproximações empíricas para a estimativa de MV dos parâmetros e o viés absoluto do EMV, sob o cenário 2.

Estimativa de MV						Viés					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.509	0.502	0.488	-0.219	60	-0.20	0.009	0.002	-0.012	-0.019
	-0.10	0.504	0.504	0.486	-0.111		-0.10	0.004	0.004	-0.014	-0.011
	0	0.500	0.506	0.485	-0.005		0	0.000	0.006	-0.015	-0.005
	0.10	0.497	0.507	0.484	0.102		0.10	-0.003	0.007	-0.016	0.002
	0.20	0.494	0.509	0.483	0.207		0.20	-0.006	0.009	-0.017	0.007
120	-0.20	0.506	0.500	0.494	-0.210	120	-0.20	0.006	0.000	-0.006	-0.010
	-0.10	0.503	0.501	0.494	-0.106		-0.10	0.003	0.001	-0.006	-0.006
	0	0.501	0.502	0.493	-0.003		0	0.001	0.002	-0.007	-0.003
	0.10	0.499	0.503	0.493	0.101		0.10	-0.001	0.003	-0.007	0.001
	0.20	0.498	0.504	0.492	0.204		0.20	-0.002	0.004	-0.008	0.004
180	-0.20	0.503	0.501	0.496	-0.207	180	-0.20	0.003	0.001	-0.004	-0.007
	-0.10	0.501	0.501	0.496	-0.104		-0.10	0.001	0.001	-0.004	-0.004
	0	0.501	0.502	0.495	-0.002		0	0.001	0.002	-0.005	-0.002
	0.10	0.499	0.503	0.495	0.101		0.10	-0.001	0.003	-0.005	0.001
	0.20	0.498	0.503	0.495	0.203		0.20	-0.002	0.003	-0.005	0.003

Fonte: Elaborada pelo autor (2022)

Os resultados contidos nas Tabelas 13, 14 e 15, correspondentes ao cenário 2, corroboram com a hipótese de consistência do EMV e as evidências são similares àquelas citadas na discussão relativa ao cenário 1. Os valores observados sugerem que a presença de observações

Tabela 14 – Aproximações empíricas para o erro padrão e o raiz quadrada do erro quadrático médio do EMV, sob o cenário 2.

Erro padrão						Raiz do EQM					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.085	0.032	0.052	0.108	60	-0.20	0.085	0.032	0.054	0.109
	-0.10	0.086	0.032	0.053	0.111		-0.10	0.086	0.032	0.055	0.111
	0	0.086	0.032	0.054	0.115		0	0.086	0.032	0.056	0.115
	0.10	0.085	0.031	0.057	0.121		0.10	0.085	0.032	0.059	0.121
	0.20	0.085	0.030	0.059	0.127		0.20	0.085	0.031	0.062	0.127
120	-0.20	0.058	0.024	0.036	0.066	120	-0.20	0.059	0.024	0.036	0.067
	-0.10	0.059	0.024	0.036	0.070		-0.10	0.059	0.024	0.037	0.070
	0	0.059	0.024	0.037	0.073		0	0.059	0.024	0.038	0.073
	0.10	0.059	0.023	0.039	0.078		0.10	0.059	0.023	0.040	0.078
	0.20	0.058	0.022	0.041	0.083		0.20	0.058	0.022	0.042	0.083
180	-0.20	0.047	0.019	0.028	0.052	180	-0.20	0.047	0.019	0.029	0.052
	-0.10	0.047	0.019	0.029	0.054		-0.10	0.047	0.019	0.029	0.055
	0	0.047	0.017	0.030	0.058		0	0.047	0.017	0.031	0.058
	0.10	0.047	0.018	0.031	0.062		0.10	0.047	0.018	0.032	0.062
	0.20	0.046	0.017	0.033	0.066		0.20	0.046	0.018	0.033	0.066

Fonte: Elaborada pelo autor (2022)

extremas na resposta, induzidas por alavancagem na covariável, são captadas satisfatoriamente pelo modelo, não interferindo de forma potencial no viés do EMV. Já em relação ao erro padrão do EMV do coeficiente da variável explicativa submetida à perturbação, observemos que há indícios de diminuição, quando comparado com o mesmo erro no cenário 1. Fato similar ocorre com a estimativa para o desvio padrão do EMV, oriunda da inversa da matriz de informação de Fisher observada, ocasionando uma ligeira diminuição na taxa de cobertura empírica do intervalo de confiança do coeficiente da covariável.

Tabela 15 – Aproximações empíricas para o erro padrão, oriundo da inversa da matriz de informação de Fisher observada (IFO), e a taxa de cobertura do intervalo de confiança, sob o cenário 2.

Erro Padrão-IFO						Taxa de cobertura					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.081	0.030	0.051	0.098	60	-0.20	0.936	0.932	0.918	0.919
	-0.10	0.082	0.030	0.051	0.102		-0.10	0.936	0.927	0.915	0.925
	0	0.082	0.029	0.053	0.107		0	0.936	0.918	0.913	0.931
	0.10	0.081	0.028	0.055	0.114		0.10	0.934	0.908	0.915	0.934
	0.20	0.081	0.026	0.058	0.121		0.20	0.933	0.896	0.913	0.937
120	-0.20	0.057	0.024	0.035	0.063	120	-0.20	0.948	0.945	0.940	0.931
	-0.10	0.058	0.024	0.036	0.066		-0.10	0.946	0.943	0.938	0.934
	0	0.058	0.023	0.037	0.071		0	0.947	0.936	0.936	0.936
	0.10	0.058	0.022	0.039	0.075		0.10	0.946	0.929	0.934	0.939
	0.20	0.057	0.021	0.041	0.081		0.20	0.944	0.926	0.933	0.942
180	-0.20	0.046	0.018	0.029	0.050	180	-0.20	0.949	0.940	0.946	0.941
	-0.10	0.047	0.019	0.029	0.053		-0.10	0.949	0.940	0.942	0.945
	0	0.046	0.016	0.030	0.057		0	0.945	0.939	0.943	0.945
	0.10	0.046	0.017	0.032	0.060		0.10	0.949	0.934	0.941	0.945
	0.20	0.046	0.016	0.033	0.064		0.20	0.948	0.930	0.940	0.946

Fonte: Elaborada pelo autor (2022)

2.7.4.3 Cenário 3

Aqui, impomos alterações na componente estocástica do modelo de modo que a variável resposta agrega valores oriundos de distribuições EVBS em que se considera diferentes valores para o parâmetro α . No caso 3A, o verdadeiro valor do parâmetro α é 1.2, porém fizemos com que 10% dos valores amostrais estivessem associados a $\alpha = 0.9$. O caso 3B é similar, mas com um aumento na distância entre o valor verdadeiro α e o que se usa na perturbação, assumindo o valor 0.7. O objetivo é verificar se o modelo é sensível a esta duplicidade imposta sobre o valor do parâmetro α . Tanto no caso 3A quanto no 3B os demais parâmetros foram fixados em $\beta_0 = 1$, $\beta_1 = 2$ e $\gamma \in \{-0.20, -0.10, 0, 0.10, 0.20\}$, este último percorrendo valores associados a distribuições de caudas diferentes.

Da décima sexta à vigéssima primeira tabela, encontramos os resultados relativos aos cenários 3A e 3B. Nota-se que a duplicidade de valores imposta sobre o parâmetro α foi captada pelo EMV. Os resultados expressos nas Tabelas 16 e 19 refletem indícios de aumento substancial, em termos relativos, no viés do EMV do parâmetro α quando comparado com o viés nos dois cenários anteriores. Particularmente, no cenário 3B, situação na qual se amplia a distância entre o verdadeiro valor de α e o usado na perturbação, o aumento é mais acentuado.

Tabela 16 – Aproximações empíricas para a estimativa de MV dos parâmetros e o viés absoluto do EMV, sob o cenário 3A.

Estimativa de MV						Viés					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	1.044	1.976	1.142	-0.227	60	-0.20	0.044	-0.024	-0.058	-0.027
	-0.10	1.034	1.988	1.139	-0.118		-0.10	0.034	-0.012	-0.061	-0.018
	0	1.024	1.998	1.136	-0.011		0	0.024	-0.002	-0.064	-0.011
	0.10	1.015	2.006	1.134	0.096		0.10	0.015	0.006	-0.066	-0.004
	0.20	1.008	2.012	1.132	0.202		0.20	0.008	0.012	-0.068	0.002
120	-0.20	1.016	1.999	1.158	-0.211	120	-0.20	0.016	-0.001	-0.042	-0.011
	-0.10	1.011	2.004	1.156	-0.108		-0.10	0.011	0.004	-0.044	-0.008
	0	1.007	2.008	1.155	-0.005		0	0.007	0.008	-0.045	-0.005
	0.10	1.004	2.011	1.155	0.098		0.10	0.004	0.011	-0.045	-0.002
	0.20	1.001	2.013	1.155	0.200		0.20	0.001	0.013	-0.045	0.000
180	-0.20	1.008	2.004	1.163	-0.207	180	-0.20	0.008	0.004	-0.037	-0.007
	-0.10	1.005	2.007	1.162	-0.105		-0.10	0.005	0.007	-0.038	-0.005
	0	1.004	2.009	1.162	-0.004		0	0.004	0.009	-0.038	-0.004
	0.10	1.002	2.010	1.162	0.097		0.10	0.002	0.010	-0.038	-0.003
	0.20	1.001	2.010	1.162	0.198		0.20	0.001	0.010	-0.038	-0.002

Fonte: Elaborada pelo autor (2022)

Além disso, os dados evidenciam que tanto o erro padrão quanto a raiz do erro quadrático médio do EVM são sensíveis a esse tipo de perturbação, pois nas Tabelas 17 e 20 encontram-se indícios de aumento do erro padrão do EMV de α quando comparado com os erros presentes nas Tabelas 11 e 14 que se mantiveram no mesmo patamar. Por fim, os resultados contidos nas Tabelas 18 e 21, onde constam as taxas de cobertura empíricas associadas ao intervalo de confiança do parâmetro α , também denunciam a duplicidade incidente sobre α , pois a taxa de cobertura do intervalo correspondente ao α é afetadas nos cenários 3A e 3B, entretanto as taxas associadas aos intervalos dos demais parâmetros mantêm o mesmo nível nominal obtido nos dois primeiros cenários.

Tabela 17 – Aproximações empíricas para o erro padrão e a raiz quadrada erro quadrático médio do EMV, sob o cenário 3A.

Erro padrão						Raiz do EQM					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.322	0.499	0.124	0.127	60	-0.20	0.325	0.499	0.137	0.130
	-0.10	0.328	0.510	0.125	0.130		-0.10	0.330	0.510	0.140	0.131
	0	0.327	0.505	0.129	0.134		0	0.328	0.505	0.144	0.135
	0.10	0.321	0.489	0.133	0.139		0.10	0.321	0.489	0.149	0.139
	0.20	0.311	0.466	0.139	0.145		0.20	0.311	0.466	0.155	0.145
120	-0.20	0.214	0.334	0.085	0.076	120	-0.20	0.215	0.334	0.095	0.076
	-0.10	0.219	0.344	0.086	0.078		-0.10	0.219	0.344	0.097	0.079
	0	0.219	0.342	0.089	0.082		0	0.219	0.342	0.100	0.082
	0.10	0.214	0.331	0.093	0.086		0.10	0.214	0.331	0.103	0.086
	0.20	0.207	0.315	0.098	0.091		0.20	0.207	0.315	0.107	0.091
180	-0.20	0.167	0.267	0.069	0.059	180	-0.20	0.167	0.267	0.078	0.059
	-0.10	0.171	0.276	0.070	0.061		-0.10	0.171	0.276	0.080	0.062
	0	0.172	0.275	0.072	0.065		0	0.172	0.275	0.082	0.065
	0.10	0.169	0.267	0.075	0.069		0.10	0.169	0.267	0.085	0.069
	0.20	0.164	0.254	0.079	0.072		0.20	0.164	0.254	0.088	0.072

Fonte: Elaborada pelo autor (2022)

Tabela 18 – Aproximações empíricas para o erro padrão, oriundo da matriz de informação de Fisher observada (IFO), e taxa de cobertura do intervalo de confiança, sob o cenário 3A.

Erro Padrão-IFO						Taxa de cobertura					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.295	0.459	0.121	0.113	60	-0.20	0.918	0.918	0.885	0.918
	-0.10	0.303	0.475	0.122	0.117		-0.10	0.924	0.925	0.884	0.919
	0	0.304	0.474	0.125	0.122		0	0.928	0.927	0.881	0.923
	0.10	0.298	0.460	0.130	0.128		0.10	0.925	0.929	0.879	0.928
	0.20	0.289	0.437	0.137	0.135		0.20	0.923	0.924	0.883	0.930
120	-0.20	0.210	0.324	0.084	0.072	120	-0.20	0.941	0.941	0.892	0.940
	-0.10	0.215	0.335	0.085	0.075		-0.10	0.944	0.940	0.894	0.940
	0	0.215	0.333	0.087	0.079		0	0.942	0.941	0.900	0.940
	0.10	0.211	0.322	0.091	0.083		0.10	0.942	0.940	0.902	0.941
	0.20	0.204	0.305	0.096	0.088		0.20	0.942	0.938	0.902	0.943
180	-0.20	0.163	0.262	0.068	0.056	180	-0.20	0.942	0.940	0.897	0.941
	-0.10	0.168	0.271	0.069	0.059		-0.10	0.942	0.939	0.895	0.941
	0	0.168	0.269	0.071	0.063		0	0.941	0.939	0.897	0.941
	0.10	0.164	0.259	0.074	0.066		0.10	0.941	0.937	0.901	0.943
	0.20	0.159	0.245	0.078	0.070		0.20	0.941	0.934	0.903	0.945

Fonte: Elaborada pelo autor (2022)

Tabela 19 – Aproximações empíricas para a estimativa de MV dos parâmetros e o viés absoluto do EMV, sob o cenário 3B.

Estimativa de MV						Viés					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	1.041	1.968	1.125	-0.218	60	-0.20	0.041	-0.032	-0.075	-0.018
	-0.10	1.029	1.984	1.122	-0.113		-0.10	0.029	-0.016	-0.078	-0.013
	0	1.018	2.000	1.120	-0.009		0	0.018	0.000	-0.080	-0.009
	0.10	1.008	2.013	1.119	0.095		0.10	0.008	0.013	-0.081	-0.005
	0.20	1.001	2.023	1.118	0.199		0.20	0.001	0.023	-0.082	-0.001
120	-0.20	1.010	1.997	1.141	-0.203	120	-0.20	0.010	-0.003	-0.059	-0.003
	-0.10	1.006	2.003	1.140	-0.103		-0.10	0.006	0.003	-0.060	-0.003
	0	1.003	2.009	1.140	-0.003		0	0.003	0.009	-0.060	-0.003
	0.10	1.000	2.014	1.140	0.097		0.10	0.000	0.014	-0.060	-0.003
	0.20	0.998	2.017	1.141	0.196		0.20	-0.002	0.017	-0.059	-0.004
180	-0.20	1.001	2.004	1.147	-0.200	180	-0.20	0.001	0.004	-0.053	0.000
	-0.10	1.000	2.007	1.147	-0.101		-0.10	0.000	0.007	-0.053	-0.001
	0	0.999	2.009	1.147	-0.003		0	-0.001	0.009	-0.053	-0.003
	0.10	0.999	2.011	1.147	0.096		0.10	-0.001	0.011	-0.053	-0.004
	0.20	0.999	2.012	1.149	0.194		0.20	-0.001	0.012	-0.051	-0.006

Fonte: Elaborada pelo autor (2022)

Tabela 20 – Aproximações empíricas para o erro padrão e a raiz quadrada erro quadrático médio do EMV, sob o cenário 3B.

Erro padrão						Raiz do EQM					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.323	0.499	0.123	0.125	60	-0.20	0.326	0.500	0.144	0.126
	-0.10	0.329	0.509	0.124	0.127		-0.10	0.330	0.509	0.146	0.128
	0	0.328	0.505	0.127	0.131		0	0.328	0.505	0.150	0.132
	0.10	0.321	0.490	0.131	0.136		0.10	0.321	0.490	0.154	0.137
	0.20	0.312	0.469	0.137	0.142		0.20	0.312	0.469	0.159	0.142
120	-0.20	0.213	0.335	0.085	0.074	120	-0.20	0.214	0.335	0.103	0.074
	-0.10	0.218	0.344	0.086	0.077		-0.10	0.218	0.344	0.105	0.077
	0	0.218	0.342	0.088	0.080		0	0.218	0.342	0.107	0.080
	0.10	0.214	0.332	0.092	0.085		0.10	0.214	0.333	0.109	0.085
	0.20	0.207	0.317	0.096	0.089		0.20	0.207	0.318	0.113	0.089
180	-0.20	0.166	0.268	0.069	0.057	180	-0.20	0.166	0.268	0.087	0.057
	-0.10	0.171	0.277	0.070	0.060		-0.10	0.171	0.277	0.088	0.060
	0	0.171	0.276	0.072	0.063		0	0.171	0.277	0.089	0.064
	0.10	0.169	0.269	0.075	0.067		0.10	0.169	0.269	0.091	0.067
	0.20	0.164	0.256	0.078	0.071		0.20	0.164	0.256	0.094	0.071

Fonte: Elaborada pelo autor (2022)

Tabela 21 – Aproximações empíricas para o erro padrão, oriundo da matriz de informação de Fisher observada (IFO), e taxa de cobertura do intervalo de confiança, sob o cenário 3B.

Erro Padrão-IFO						Taxa de cobertura					
n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$	n	γ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
60	-0.20	0.294	0.458	0.118	0.111	60	-0.20	0.916	0.918	0.856	0.923
	-0.10	0.302	0.473	0.119	0.115		-0.10	0.923	0.923	0.856	0.925
	0	0.302	0.471	0.123	0.119		0	0.925	0.927	0.857	0.925
	0.10	0.296	0.457	0.128	0.125		0.10	0.925	0.924	0.860	0.929
	0.20	0.287	0.436	0.134	0.132		0.20	0.922	0.921	0.863	0.930
120	-0.20	0.208	0.323	0.082	0.071	120	-0.20	0.940	0.939	0.856	0.946
	-0.10	0.214	0.333	0.083	0.074		-0.10	0.941	0.939	0.853	0.943
	0	0.213	0.331	0.086	0.077		0	0.940	0.938	0.860	0.942
	0.10	0.209	0.320	0.090	0.082		0.10	0.940	0.936	0.869	0.942
	0.20	0.202	0.304	0.094	0.086		0.20	0.939	0.934	0.877	0.942
180	-0.20	0.162	0.260	0.067	0.055	180	-0.20	0.941	0.938	0.845	0.946
	-0.10	0.166	0.268	0.068	0.058		-0.10	0.939	0.934	0.849	0.943
	0	0.166	0.266	0.070	0.061		0	0.939	0.935	0.856	0.944
	0.10	0.162	0.257	0.073	0.065		0.10	0.939	0.933	0.864	0.945
	0.20	0.157	0.243	0.077	0.069		0.20	0.940	0.929	0.873	0.946

Fonte: Elaborada pelo autor (2022)

3 AVALIAÇÃO DE INFLUÊNCIA LOCAL

3.1 INTRODUÇÃO

Neste capítulo, desenvolveremos análise de diagnóstico por influência local sobre o modelo de regressão EVBS, definido por (LEIVA et al., 2016). Nas Seções (3.2), (3.3) e (3.4), introduziremos o estudo de influência local segundo a metodologia desenvolvida por (COOK, 1986). Em seguida, na Seção (3.5), trataremos de influência local sob a ótica da curvatura normal conforme, abordagem dada por (POON; POON, 1999). Por fim, lidaremos com análise de influência local no modelo de regressão EVBS, sobre o qual consideraremos três esquemas de perturbação: ponderação de casos, perturbação na variável resposta e perturbação em uma covariável. A fim de ilustrar a eficácia da metodologia de influência local em detectar observações influentes, consideramos um conjunto de dados de velocidade máxima mensal de ventos registrados pela estação Itajaí-SC(A-868). No Apêndice B consta um estudo geométrico-diferencial de superfícies que são gráficos de funções diferenciáveis, fundamentando à análise de influência por meio da função afastamento pela verossimilhança e contribuindo para uma compreensão mais profunda da metodologia.

A técnica de influência local, proposta por (COOK, 1986), é um método de diagnóstico que permite avaliar, mediante o uso de uma medida de influência, a estabilidade das estimativas fornecidas por um modelo estatístico quando impomos pequenas perturbações nos dados ou no próprio modelo. Ela baseia-se no conceito geométrico de curvatura para investigar o comportamento local do gráfico de uma função denominada afastamento pela verossimilhança. Mais precisamente, a curvatura normal, C_l , em um ponto ω_0 , numa dada direção l , é calculada e usada como indicador de possíveis problemas locais. (COOK, 1986) sugere que examinemos as direções em que a curvatura normal é máxima, pois a partir dessas podemos identificar observações que podem interferir de forma desproporcional nos resultados inferenciais. Porém, (POON; POON, 1999) chamam atenção para o fato de que a curvatura normal pode assumir qualquer valor na reta real e não é invariante sob mudanças de escala, ocasionando perda de objetividade no julgamento da grandeza da curvatura. Com o objetivo de resolver esse problema e conseqüentemente aperfeiçoar a técnica, eles propõem o uso da curvatura normal conforme, que está relacionada com a curvatura normal e assume valores em um intervalo limitado da reta, possibilitando a construção de valores de referência que nos permite julgar a magnitude dessa curvatura. Além disso, mostram que a curvatura normal conforme é invariante

sob uma classe de reparametrizações chamadas de reparametrizações conformes.

3.2 A FUNÇÃO AFASTAMENTO PELA VEROSSIMILHANÇA E O GRÁFICO DE INFLUÊNCIA

Seja $\ell(\boldsymbol{\theta})$ a função de verossimilhança associada a um modelo estatístico postulado, onde $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^\top \in \Theta \subset \mathbb{R}^p$ é um vetor $p \times 1$ de parâmetros desconhecidos. A inserção de perturbações sobre o modelo é feita por meio de um vetor $\boldsymbol{\omega} = (\omega_1, \dots, \omega_q)^\top$ pertencente a um conjunto aberto $U \subset \mathbb{R}^q$. Segundo (COOK, 1986), o vetor $\boldsymbol{\omega}$ representa, em geral, um esquema de perturbação bem definido, logo deve ser cuidadosamente selecionado a fim de que o esquema correspondente seja passível de interpretação nas aplicações. Uma vez fixado o esquema de perturbação, seja $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ a função de log-verossimilhança associada ao modelo estatístico perturbado por um vetor qualquer $\boldsymbol{\omega} \in U$. Aqui, admite-se que $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ é diferenciável em $(\boldsymbol{\theta}^\top, \boldsymbol{\omega}^\top)$, ou seja, possui derivadas parciais de todas as ordens e as mesmas são contínuas. Além disso, supõe-se a existência de um vetor $\boldsymbol{\omega}_0$ tal que $\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0)$ para todo $\boldsymbol{\theta} \in \Theta$, chamado vetor de não-perturbação, e admitimos $q = n$. Prosseguindo, sejam $\hat{\boldsymbol{\theta}}$ e $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ os estimadores de máxima verossimilhança para $\boldsymbol{\theta}$ segundo os modelos postulado e perturbado, respectivamente. A fim de avaliar a influência quando $\boldsymbol{\omega}$ varia em U , (COOK, 1986) considera, inicialmente, a função afastamento pela verossimilhança definida por

$$f : U \mapsto \mathbb{R}, \quad f(\boldsymbol{\omega}) = 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}|\boldsymbol{\omega})]. \quad (3.1)$$

Em seguida, cita que o gráfico de $f(\boldsymbol{\omega})$ versus $\boldsymbol{\omega}$ possui informações relevantes sobre o esquema de perturbação utilizado e destaca a importância de identificá-lo como uma superfície $G \subset \mathbb{R}^{n+1}$ formada pelos pontos

$$\boldsymbol{\phi}(\boldsymbol{\omega}) = \begin{pmatrix} \boldsymbol{\omega} \\ f(\boldsymbol{\omega}) \end{pmatrix}, \quad \boldsymbol{\omega} \in U, \quad (3.2)$$

denominando-a gráfico de influência.

Suponhamos que o vetor $\boldsymbol{\theta}$ seja particionado de modo que $\boldsymbol{\theta}^\top = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)$, em que $\boldsymbol{\theta}_1 \in \Theta_1$ é $p_1 \times 1$ e $\boldsymbol{\theta}_2 \in \Theta_2$ é $p_2 \times 1$, com $p_1 + p_2 = p$. Se estivermos interessados apenas em avaliar a influência sobre os coeficientes de $\boldsymbol{\theta}_1$, então (COOK, 1986) afirma que o gráfico de influência

é

$$\phi_s(\omega) = \begin{pmatrix} \omega \\ f_s(\omega) \end{pmatrix}, \quad \omega \in U. \quad (3.3)$$

Neste caso, o afastamento pela verossimilhança é dada por

$$f_s(\omega) = 2[\ell(\hat{\theta}_1, \hat{\theta}_2) - \ell(\hat{\theta}_{1\omega}, g(\hat{\theta}_{1\omega}))], \quad (3.4)$$

onde $g : \Theta_1 \mapsto \Theta_2$ é uma função que maximiza $\ell(\theta_1, \theta_2)$ para cada θ_1 fixo em Θ_1 , $\hat{\theta}_{1\omega}$ é determinado por meio da partição $\hat{\theta}_\omega^\top = (\hat{\theta}_{1\omega}^\top, \hat{\theta}_{2\omega}^\top)$ e $\ell(\theta_1, g(\theta_1))$ é a função de verossimilhança perfilada para θ_1 .

3.3 AVALIAÇÃO DE INFLUÊNCIA LOCAL

3.3.1 Influência individual

O ponto de partida para uma análise de influência local é a inserção de perturbações sobre o modelo estatístico adotado, que é feita por meio de um vetor de perturbação $\omega = (\omega_1, \dots, \omega_n)^\top \in U$, em que U é um aberto em \mathbb{R}^n . Fixado o esquema de perturbação, ele determina a forma como o vetor ω atua na função de log-verossimilhança do modelo postulado, levando-nos a função afastamento pela verossimilhança, denotada por $f : U \rightarrow \mathbb{R}$. Daí, voltamos nossa atenção para a superfície G correspondente ao gráfico f . Particularmente, interessa-nos investigar a curvatura de G no ponto ω_0 ao longo de direções arbitrárias l pertencentes ao espaço tangente de G . Aqui há um pequeno abuso de notação, onde está escrito "curvatura de G no ponto ω_0 ", entende-se por "curvatura de G no ponto $\phi(\omega_0)$ ".

Na situação particular em que apenas a j -ésima componente do vetor de perturbação ω varia, a função f afastamento pela verossimilhança comporta-se exclusivamente como função de ω_j . Neste caso, o gráfico de f coincide com a seção normal $G(e_j)$, definida pela equação (B.14), e denotamos por C_j a curvatura da seção normal $G(e_j)$ no ponto ω_0 e na direção do vetor e_j . Segundo (POON; POON, 2010), o valor de C_j mensura a sensibilidade do modelo ao perturbarmos uma única componente de ω e o denominam de influência individual devido à perturbação na j -ésima componente do vetor ω . Se ω_0 é um ponto crítico de f , segue que o gradiente de f no ponto ω_0 é o funcional linear nulo. Consequentemente,

$$C_j = C_{e_j} = e_j^\top \ddot{F}(\omega_0) e_j = \frac{\partial^2 f(\omega_0)}{\partial \omega_j^2}. \quad (3.5)$$

É natural considerarmos uma extensão da ideia de influência individual associada ao vetor básico, e_j , para um vetor l que é combinação linear de vetores básicos. Neste caso, a influência individual correspondente à direção l é dada pela curvatura da seção normal $G(l)$ em ω_0 na direção do vetor l . Se ω_0 é ponto crítico de f e l é um vetor unitário, tem-se

$$C_l = l^\top \ddot{F}(\omega_0)l. \quad (3.6)$$

A partir das equações (3.1) e (3.6), (COOK, 1986) deduziu uma fórmula operacional para o cálculo da curvatura normal de G em um ponto crítico ω_0 da função afastamento pela verossimilhança. Mais precisamente, ele mostrou que

$$C_l = -2l^\top (\Delta^\top \ddot{L}^{-1} \Delta) l \Big|_{\theta=\hat{\theta}, \omega=\omega_0}, \quad (3.7)$$

em que $\Delta = (\Delta_{ij})_{p \times n}$ é a matriz cujos termos são dados por

$$\Delta_{ij} = \frac{\partial^2 \ell(\theta | \omega)}{\partial \theta_i \partial \omega_j}, \quad i = 1, \dots, p \text{ e } j = 1, 2, \dots, n \quad (3.8)$$

avaliados em $\theta = \hat{\theta}$ e $\omega = \omega_0$. E, $\ddot{L} = (\ddot{L}_{ij})_{p \times p}$ é a matriz cujas entradas são dadas por

$$\ddot{L}_{ij} = \frac{\partial^2 \ell(\theta | \omega)}{\partial \theta_i \partial \theta_j}, \quad i = 1, \dots, p \text{ e } j = 1, \dots, p \quad (3.9)$$

avaliadas em $\theta = \hat{\theta}$ e $\omega = \omega_0$.

De acordo com (COOK, 1986), existem diversas formas pelas quais a curvatura normal, dada pela equação (3.7), pode ser usada como meio para investigar o comportamento local da superfície gráfico de influência. Em particular, cita que podemos proceder a análise nos atentando à direção em que a curvatura normal é máxima, que corresponde ao autovetor de \ddot{F} associado ao maior autovalor. (COOK, 1986) sustenta que o autovetor l_{\max} associado ao $C_{\max} = \max_l C_l$ indica como perturbar o modelo postulado de modo a obter grandes mudanças locais na função afastamento pela verossimilhança e sugere que investiguemos a causa potencial desta instabilidade. Além disso, comenta que os autovetores associadas aos autovalores intermediários também podem ser utilizados para investigar o comportamento do gráfico de influência nas direções correspondente as curvaturas menos extremas.

3.3.2 Influência conjunta

Aqui, consideremos pequenas perturbações em apenas duas componentes distintas, ω_i e ω_j , do vetor de perturbação ω . Neste caso, a seção transversal $G(e_i, e_j)$ é o gráfico da função

f afastamento pela log-verossimilhança. Além disso, a matriz Hessiana de f no vetor de não-perturbação $\boldsymbol{\omega}_0$, denotada por $\ddot{\mathbf{F}}_{ij}(\boldsymbol{\omega}_0)$, é simétrica e dada por

$$\ddot{\mathbf{F}}_{ij}(\boldsymbol{\omega}_0) = \begin{pmatrix} \frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_i^2} & \frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_i \partial \omega_j} \\ \frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_j \partial \omega_i} & \frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_j^2} \end{pmatrix}. \quad (3.10)$$

O máximo da curvatura normal, C_{\max} , em $\boldsymbol{\omega}_0$, ao longo de direções arbitrárias no plano gerado pelos vetores \mathbf{e}_i e \mathbf{e}_j é o maior autovalor da matriz $\ddot{\mathbf{F}}_{ij}(\boldsymbol{\omega}_0)$, denotado por λ_{\max} . Segundo (POON; POON, 2010), o λ_{\max} é dado por

$$\lambda_{\max} = J_{ij} = \frac{1}{2} \left(\frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_i^2} + \frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_j^2} \right) + \sqrt{\frac{1}{4} \left(\frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_i^2} - \frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_j^2} \right)^2 + \left(\frac{\partial^2 f(\boldsymbol{\omega}_0)}{\partial \omega_i \partial \omega_j} \right)^2}. \quad (3.11)$$

Além disso, consideram que J_{ij} deve ser entendida como uma medida de influência local conjunta associada às componentes ω_i e ω_j do vetor de perturbação $\boldsymbol{\omega}$.

3.4 RELAÇÃO ENTRE MEDIDAS DE INFLUÊNCIA

Ao implementarmos uma análise de influência buscamos identificar as componentes do vetor de perturbação, $\boldsymbol{\omega}$, que exercem grande influência individual ou conjunta dentre as direções básicas $\mathbf{e}_1, \dots, \mathbf{e}_n$. No primeiro caso, isso pode ser feito por meio da obtenção do valor máximo dentre as curvaturas, C_1, \dots, C_n , associadas às direções dos vetores básicos, ou equivalentemente, mediante a determinação do maior autovalor da matriz Hessiana $\ddot{\mathbf{F}}$ da função afastamento pela log-verossimilhança. No caso de avaliação da influência conjunta, onde se perturba k componentes do vetor de perturbação, $\omega_{i_1}, \dots, \omega_{i_k}$, procedemos a fim de identificar o máximo de J_{i_1, \dots, i_k} . Porém, dado a dificuldade existente em obter fórmulas analíticas para J_{i_1, \dots, i_k} e o enorme número de combinações possíveis quando k é grande, a determinação do máximo de J_{i_1, \dots, i_k} revela-se pouco operacional. Conforme veremos ao longo desta seção, (POON; POON, 2010) contornaram essa dificuldade ao estabelecerem relações entre as medidas de influência.

3.4.1 Influência individual e influência conjunta.

Sejam $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ a base canônica do \mathbb{R}^n e $\mathcal{F} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ a coleção formada pelos autovetores ortonormais de $\ddot{\mathbf{F}}$. Dado $\mathbf{v}_i = \sum_{j=1}^n a_{ji} \mathbf{e}_j$, consideremos a matriz $\mathbf{A}_{n \times n}$ definida

por

$$\mathbf{A} = (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \quad (3.12)$$

Uma vez que a coleção de vetores $\{\mathbf{v}_i : 1 \leq i \leq n\}$ é uma base ortonormal, a matriz \mathbf{A} é ortogonal. Logo, $\mathbf{A}^{-1} = \mathbf{A}^\top$. É claro que $\mathbf{A}\mathbf{e}_j = \mathbf{v}_j$, daí são válidas as seguintes implicações:

$$\mathbf{A}\mathbf{e}_j = \mathbf{v}_j \Rightarrow \mathbf{A}^{-1}\mathbf{A}\mathbf{e}_j = \mathbf{A}^{-1}\mathbf{v}_j \Rightarrow \mathbf{e}_j = \mathbf{A}^{-1}\mathbf{v}_j \Rightarrow \mathbf{e}_j = \mathbf{A}^\top\mathbf{v}_j.$$

Logo,

$$\mathbf{e}_j = \sum_{i=1}^n a_{ji}\mathbf{v}_i. \quad (3.13)$$

Seja $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ a matriz diagonal formada pelos autovalores de $\ddot{\mathbf{F}}$ cujos autovetores correspondentes são $\mathbf{v}_1, \dots, \mathbf{v}_n$. Dado que $\ddot{\mathbf{F}}\mathbf{v}_j = \lambda_j\mathbf{v}_j$, tem-se $\ddot{\mathbf{F}}\mathbf{A} = \mathbf{A}\mathbf{\Lambda}$. Consequentemente, $\ddot{\mathbf{F}} = \mathbf{A}\mathbf{\Lambda}\mathbf{A}^{-1}$, ou seja,

$$\frac{\partial^2 f}{\partial \omega_i \partial \omega_j}(\boldsymbol{\omega}_0) = \sum_{k=1}^n \lambda_k a_{ik} a_{jk}. \quad (3.14)$$

Em particular,

$$C_j = \frac{\partial^2 f}{\partial \omega_j^2}(\boldsymbol{\omega}_0) = \sum_{k=1}^n \lambda_k a_{jk}^2. \quad (3.15)$$

Neste ponto, consideremos a hipótese de que a matriz Hessiana $\ddot{\mathbf{F}}$ da função afastamento pela verossimilhança, f , no ponto $\boldsymbol{\omega}_0$ possui um autovalor dominante associado a um único autovetor, no sentido de que é muito maior do que os demais, e fixemos a seguinte ordenação:

$$\lambda_{\max} = \lambda_1 \gg \lambda_2 \geq \cdots \geq \lambda_n = \lambda_{\min}.$$

Logo, quanto maior for λ_1 , comparado com os demais λ'_k s, mais próximo de zero estará a razão $\frac{\lambda_k}{\lambda_1}$. Nessas circunstâncias, teremos

$$C_1 + C_2 + \cdots + C_n = \text{tr}(\ddot{\mathbf{F}}) \approx \lambda_{\max}.$$

Consequentemente, dado que $C_j = \sum_{k=1}^n \lambda_k a_{jk}^2$, obtemos a relação

$$\frac{C_j}{\lambda_{\max}} \approx a_{j1}^2.$$

Teorema 3.1 (POON; POON, 2010). Se a matriz Hessiana, \ddot{F} , da função afastamento pela verossimilhança no ponto ω_0 possui um autovalor λ_{\max} dominante, então a influência individual relaciona-se com as coordenadas do autovetor, $\mathbf{v}_{\max} = \sum_{j=1}^n a_{ji} \mathbf{e}_j$, por meio da seguinte aproximação:

$$\frac{C_j}{\lambda_{\max}} \approx a_{j1}^2. \quad (3.16)$$

3.4.2 Medidas de influência individual e conjunta

Prosseguiremos com a análise ainda sob a hipótese de que \ddot{F} possui um autovalor dominante. O propósito nesta subseção é estabelecermos uma relação entre a medida de influência individual C_j e a medida de influência conjunta J_{ij} e, de um modo geral, entre C_j e $J_{i1,i2,\dots,ik}$.

Sob o pressuposto citado e usando a equação (3.14), constata-se a validade da seguinte aproximação:

$$\frac{\partial^2 f}{\partial \omega_i \partial \omega_j}(\omega_0) = \sum_{k=1}^n \lambda_k a_{ik} a_{jk} \approx \lambda_{\max} a_{i1} a_{j1}$$

Substituindo esse valor aproximado na equação (3.11), obtemos

$$J_{ij} = \frac{1}{2} (\lambda_{\max} a_{i1}^2 + \lambda_{\max} a_{j1}^2) + \sqrt{\frac{1}{4} (\lambda_{\max} a_{i1}^2 - \lambda_{\max} a_{j1}^2)^2 + \lambda_{\max}^2 a_{i1}^2 a_{j1}^2}.$$

Ou equivalentemente,

$$J_{ij} = \frac{1}{2} \lambda_{\max} (a_{i1}^2 + a_{j1}^2) + \frac{1}{2} \lambda_{\max} \sqrt{(a_{i1}^2 + a_{j1}^2)^2}.$$

Logo,

$$J_{ij} = \lambda_{\max} (a_{i1}^2 + a_{j1}^2) \approx C_i + C_j.$$

Portanto, se a matriz \ddot{F} possui um autovalor dominante, a influência conjunta, J_{ij} , devido a perturbações em exatamente duas componentes, ω_i e ω_j , do vetor ω é aproximadamente a soma das influências individuais C_i e C_j . Esse resultado admite uma generalização expressa no seguinte teorema:

Teorema 3.2 (POON; POON, 2010). Se a matriz Hessiana \ddot{F} da função afastamento pela verossimilhança no ponto ω_0 possui um autovalor λ_{\max} dominante, então a influência conjunta, $J_{i1,i2,\dots,ik}$, devido a perturbações em k componentes do vetor ω é aproximadamente a soma das influências individuais $C_{i1}, C_{i2}, \dots, C_{ik}$.

3.4.3 Autovalores competitivos

Neste ponto, nos afastamos da hipótese de que a matriz Hessiana $\ddot{\mathbf{F}}$ possui um autovalor dominante e passaremos a considerar a existência de autovalores competitivos ou autovetores concorrentes, no sentido de que \mathbf{v}_1 e \mathbf{v}_2 são autovetores distintos de $\ddot{\mathbf{F}}$ cujos autovalores correspondentes, λ_1 e λ_2 , são tais que $\lambda_{\max} = \lambda_1 = \lambda_2$. Mesmo que essa igualdade não se verifique para autovetores distintos de $\ddot{\mathbf{F}}$, é prudente levarmos em consideração as situações nas quais a distância entre λ_1 e λ_2 é pequena. Segundo (POON; POON, 2010), isso provoca a necessidade de estabelecermos um critério para decidirmos quais autovetores, além do \mathbf{v}_{\max} , devem ser investigados na análise de influência local.

Ressaltamos que na Subseção B.0.5 fica estabelecida a igualdade entre média aritmética dos autovalores da matriz $\ddot{\mathbf{F}}$ e a curvatura média, $\bar{\lambda}$, do gráfico da função afastamento pela verossimilhança no ponto $\boldsymbol{\omega}_0$, ou seja,

$$\bar{\lambda} = \frac{1}{n} (\lambda_1 + \cdots + \lambda_n) = \frac{1}{n} \text{tr}(\ddot{\mathbf{F}}).$$

Logo, se todos os autovalores, $\lambda_{k's}$, forem idênticos, então cada um deles é igual a $\bar{\lambda}$. Dessa forma, pode-se usar $2\bar{\lambda}$ ou $3\bar{\lambda}$ ou ainda $q\bar{\lambda}$ como valores de referência para determinarmos quais autovetores serão investigados.

Consideremos a ordenação

$$\lambda_1 \geq \cdots \geq \lambda_k \geq q\bar{\lambda} > \lambda_{k+1} \geq \cdots \geq \lambda_n. \quad (3.17)$$

Vimos que a influência individual C_j pode ser expressa em termos das coordenadas de cada autovetor por meio da relação

$$C_j = \frac{\partial^2 f}{\partial \omega_j^2}(\boldsymbol{\omega}_0) = \sum_{i=1}^n \lambda_i a_{ji}^2. \quad (3.18)$$

Suponha que, baseando-se na expressão (3.17), decidimos incluir os autovetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ como relevantes na análise de influência, e não somente o \mathbf{v}_{\max} . Neste caso, (POON; POON, 2010) usam a terminologia q -influyente para denominar os autovetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$. Assim, considerando apenas os autovetores q -influyentes, tem-se

$$C_j(q) = \sum_{i=1}^k \lambda_i a_{ji}^2. \quad (3.19)$$

Dentro desse ponto de vista, (POON; POON, 2010) consideram que a influência individual C_j é a contribuição agregada do j -ésimo vetor básico de perturbação, \mathbf{e}_j , para todos os autovetores q -influyentes. Além disso, a partir da equação (3.19), constatam que

$$\sum_{j=1}^n C_j(q) = \sum_{j=1}^n \left(\sum_{i=1}^k \lambda_i a_{ij}^2 \right) = \sum_{i=1}^k \lambda_i \left(\sum_{j=1}^n a_{ij}^2 \right) = \sum_{i=1}^k \lambda_i. \quad (3.20)$$

Em particular, quando $q = 0$, tem-se

$$\sum_{j=1}^n C_j(0) = \sum_{i=1}^k \lambda_i = \text{tr}(\ddot{\mathbf{F}})$$

Por fim, (POON; POON, 2010) sugerem que a curvatura média $\bar{\lambda}$ pode ser usada para construir valores de referência que permitem julgar a magnitude de C_j . Mas, acentuam que, embora C_j seja limitada superiormente por λ_{\max} , o autovalor λ_{\max} pode assumir um valor arbitrariamente grande, ocasionando perda de objetividade no julgamento da grandeza de C_j .

3.5 A CURVATURA NORMAL CONFORME

Do exposto nas seções anteriores, constata-se que a curvatura normal $C_j = C_{e_j}$ apresenta uma leve inconveniência quando temos por objetivo realizar uma análise de diagnóstico. Pois, uma vez que o λ_{\max} pode assumir um valor arbitrariamente grande, C_j assumirá qualquer valor real menor do que ou igual ao λ_{\max} , ocasionando perda de objetividade no julgamento da grandeza dessa curvatura. Além disso, (POON; POON, 1999) argumentam que a curvatura normal não é invariante por mudança de escala imposta no esquema de perturbação. Com intuito de solucionar essas inconveniências e, conseqüentemente, aperfeiçoar a técnica de influência local (POON; POON, 1999) propõem o uso da curvatura normal conforme, que está estreitamente relacionada com a curvatura normal, mas assume valores em um intervalo limitado e é invariante sob certa classe de reparametrizações.

3.5.1 Reparametrização e invariância

Esperamos que em algumas situações, caracterizadas por certas transformações sobre o vetor de perturbação ω , a análise de influência local não seja afetada. Em geral, uma transformação sobre um vetor de perturbação ou, equivalentemente, sobre um dado esquema de perturbação, vem na forma de uma aplicação bijetora $\varphi : V \rightarrow U$, onde V e U são abertos em \mathbb{R}^q . Assim, teríamos dois caminhos possíveis para proceder com a análise de influência: um deles baseado na função afastamento pela verossimilhança original, denotada por $f : U \rightarrow \mathbb{R}$; o outro fundado na função de afastamento pela verossimilhança modificada, decorrente da

reparametrização $\varphi : U \rightarrow V$, dada por $g = f \circ \varphi : V \rightarrow \mathbb{R}$. Interessa-nos expressar condições que φ deve satisfazer a fim de que as análises baseadas em f e $g = f \circ \varphi$ sejam indiferentes. Ou equivalentemente, obter uma medida de análise de influência que não é alterada quando o esquema de perturbação é reparametrizado por $\varphi : V \rightarrow U$.

Seja $f : U \rightarrow \mathbb{R}$ a função afastamento pela verossimilhança associada a determinado esquema de perturbação cujo vetor de não-perturbação é ω_0 . Agora, ao considerarmos uma aplicação $\varphi : V \rightarrow U$ bijetora com inversa denotada por $\varphi^{-1} : U \rightarrow V$, onde $U, V \subset \mathbb{R}^q$ e $\varphi(\mathbf{t}_0) = \omega_0$, a composta $g = f \circ \varphi$ é a função de verossimilhança correspondente ao esquema de perturbação reparametrizado, tendo \mathbf{t}_0 como vetor de não-perturbação. De acordo com (POON; POON, 1999), dizemos que uma aplicação diferenciável $\varphi : V \rightarrow U$ é uma reparametrização conforme em \mathbf{t}_0 , se a matriz jacobiana de φ no ponto \mathbf{t}_0 é uma matriz conforme, isto é, existe uma constante real $c > 0$ tal que

$$d\varphi^\top(\mathbf{t}_0)d\varphi^\top(\mathbf{t}_0) = c\mathbf{I}_q, \quad (3.21)$$

onde \mathbf{I}_q denota a matriz identidade de ordem q . Como veremos adiante, as reparametrizações conformes desempenham um papel crucial quando se objetiva que a análise de influência por meios das funções f e $g = f \circ \varphi$ seja indiferentes.

3.5.2 Invariância da curvatura normal conforme

Seja $f : U \rightarrow \mathbb{R}$, $f(\omega) = 2(\ell(\hat{\theta}) - \ell(\hat{\theta}_\omega))$ a função afastamento pela verossimilhança associada a um esquema de perturbação no qual $\omega_0 \in U \subset \mathbb{R}^q$ é o vetor de não-perturbação. (POON; POON, 1999) definem a curvatura normal do gráfico de f , em um ponto ω_0 , na direção do vetor l por

$$B_l = \frac{1}{\sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}} \frac{l^\top \ddot{\mathbf{F}} l}{l^\top l} \Big|_{\omega=\omega_0} = \frac{1}{\sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}} C_l \Big|_{\omega=\omega_0}. \quad (3.22)$$

Em termos das matrizes $\ddot{\mathbf{L}}$ e Δ dadas pela equações (3.9) e (3.8), respectivamente, uma fórmula operacional para o cálculo de B_l , em um ponto crítico ω_0 de f , direção do vetor unitário l é dada por:

$$B_l = - \frac{l^\top (\Delta^\top \ddot{\mathbf{L}}^{-1} \Delta) l}{\sqrt{\text{tr}[(\Delta^\top \ddot{\mathbf{L}}^{-1} \Delta)^2]}} \Big|_{\theta=\hat{\theta}, \omega=\omega_0}. \quad (3.23)$$

Neste ponto, ao compararmos as equações (3.7) e (3.23) constatamos que, em relação ao esforço computacional, o cálculo de B_l equipara-se ao de C_l . No entanto, a curvatura normal

conforme possui propriedades vantajosas quando comparada com a curvatura normal. Por exemplo, a curvatura normal conforme é invariante quando o esquema de perturbação usado é submetido a uma reparametrização conforme, segundo definição apresentada na Subseção 3.5.1. Esta propriedade é assegurada pelo seguinte teorema:

Teorema 3.3 (POON; POON, 1999). Sejam $f : U \rightarrow \mathbb{R}$ a função afastamento pela log-verossimilhança associada a um dado esquema de perturbação e ω_0 um ponto crítico de f . Se $\varphi : V \rightarrow U$ é uma reparametrização conforme, então a curvatura normal conforme no ponto ω_0 , em qualquer direção unitária \mathbf{l} , é invariante sob essa reparametrização.

3.5.3 Limitação e construção de valores de referência

Na Subseção 3.4.3, sugeriu-se o uso da curvatura média como valor de referência para julgar a magnitude da curvatura normal. Agora, trabalharemos com intuito de estabelecer valores análogos quando usamos a curvatura normal conforme como medida de influência. Sejam $\mathcal{F} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ uma base ortonormal formada por autovetores de $\ddot{\mathbf{F}}$ e λ_i o autovalor correspondente ao autovetor \mathbf{v}_i para cada $i = 1, 2, \dots, n$. Se $\mathbf{l} \in \mathbb{R}^n$ uma combinação linear dos autovetores da base \mathcal{F} , temos $\mathbf{l} = \sum_{i=1}^n l_i \mathbf{v}_i$, com $l_i \in \mathbb{R}$ para todo $i = 1, \dots, n$.

Observemos que $\mathbf{v}_i^\top \ddot{\mathbf{F}} \mathbf{v}_i = \mathbf{v}_i^\top (\lambda_i \mathbf{v}_i) = \lambda_i$. Logo, é válida a seguinte igualdade:

$$\sum_{i=1}^n (l_i \mathbf{v}_i)^\top \ddot{\mathbf{F}} (l_i \mathbf{v}_i) = \sum_{i=1}^n l_i^2 \lambda_i. \quad (3.24)$$

Dado que $(l_i \mathbf{v}_i)^\top \ddot{\mathbf{F}} (l_i \mathbf{v}_i) = \langle \ddot{\mathbf{F}} (l_i \mathbf{v}_i), l_i \mathbf{v}_i \rangle$, teremos

$$\sum_{i=1}^n (l_i \mathbf{v}_i)^\top \ddot{\mathbf{F}} (l_i \mathbf{v}_i) = \left\langle \sum_{i=1}^n \ddot{\mathbf{F}} (l_i \mathbf{v}_i), \sum_{i=1}^n l_i \mathbf{v}_i \right\rangle = \mathbf{l}^\top \ddot{\mathbf{F}} \mathbf{l}. \quad (3.25)$$

Portanto,

$$\mathbf{l}^\top \ddot{\mathbf{F}} \mathbf{l} = \sum_{i=1}^n l_i^2 \lambda_i. \quad (3.26)$$

Isso nos permite obter uma cota superior para o valor absoluto da curvatura $B_{\mathbf{l}}$. De fato, substituindo (3.26) na definição dada em (3.22), obtemos

$$B_{\mathbf{l}} = \left| \frac{1}{\sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}} \frac{\mathbf{l}^\top \ddot{\mathbf{F}} \mathbf{l}}{\mathbf{l}^\top \mathbf{l}} \right| = \frac{1}{\sqrt{\sum_{i=1}^n \lambda_i^2} \sqrt{\sum_{i=1}^n l_i^2}} \left| \sum_{i=1}^n l_i^2 \lambda_i \right|.$$

Usando a desigualdade de Cauchy-Schwartz, teremos

$$\left| \sum_{i=1}^n l_i^2 \lambda_i \right| = \left| \langle (l_1^2, \dots, l_n^2), (\lambda_1, \dots, \lambda_n) \rangle \right| \leq \sqrt{l_1^4 + \dots + l_n^4} \cdot \sqrt{\lambda_1^2 + \dots + \lambda_n^2}.$$

Logo,

$$|B_{\mathbf{l}}| \leq \frac{\sqrt{\sum_{i=1}^n l_i^4}}{\sum_{i=1}^n l_i^2}.$$

Agora, usando o fato de que

$$\left(\sum_{i=1}^n l_i^2\right)^2 = \sum_{i=1}^n l_i^4 + 2 \sum_{i<j} l_i^2 l_j^2,$$

ou equivalentemente,

$$\sum_{i=1}^n l_i^2 = \sqrt{\sum_{i=1}^n l_i^4 + 2 \sum_{i<j} l_i^2 l_j^2},$$

concluimos

$$|B_{\mathbf{l}}| \leq \frac{\sqrt{\sum_{i=1}^n l_i^4}}{\sqrt{\sum_{i=1}^n l_i^4 + 2 \sum_{i<j} l_i^2 l_j^2}} \leq 1.$$

Os argumentos apresentados acima sustentam que a curvatura normal conforme $B_{\mathbf{l}}$ assume valores no intervalo $[0, 1]$ e isso nos possibilita ter um melhor juízo da magnitude de $B_{\mathbf{l}}$.

Teorema 3.4 (POON; POON, 1999). A curvatura normal conforme $B_{\mathbf{l}}$ é tal que $|B_{\mathbf{l}}| \leq 1$ para qualquer direção $\mathbf{l} \in \mathbb{R}^n$.

Consideremos uma direção $\mathbf{v}_i \in \mathcal{F} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, que é base ortonormal formada pelos autovetores de $\ddot{\mathbf{F}}$, segue que

$$B_{\mathbf{v}_i} = \frac{1}{\sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}} \frac{\mathbf{v}_i^\top \ddot{\mathbf{F}} \mathbf{v}_i}{\mathbf{v}_i^\top \mathbf{v}_i} = \frac{1}{\sqrt{\sum_{k=1}^n \lambda_k^2}} \frac{\mathbf{v}_i^\top \lambda_i \mathbf{v}_i}{\mathbf{v}_i^\top \mathbf{v}_i} = \frac{\lambda_i}{\sqrt{\sum_{k=1}^n \lambda_k^2}}. \quad (3.27)$$

Ou seja, $B_{\mathbf{v}_i}$ é igual ao autovalor normalizado $\lambda_i^* = \frac{\lambda_i}{\sqrt{\sum_{k=1}^n \lambda_k^2}}$. Consequentemente,

$$\sum_{i=1}^n B_{\mathbf{v}_i}^2 = \sum_{i=1}^n \lambda_i^{*2} = 1. \quad (3.28)$$

Isso acarreta que se $B_{\mathbf{v}_i} = B$ para todo autovetor $\mathbf{v}_i \in \mathcal{F}$, então $B = \frac{1}{\sqrt{n}}$. Utilizando $\frac{1}{\sqrt{n}}$ como valor de referência para as curvaturas associadas aos autovetores, tornamos nossa decisão de julgar a magnitude de $B_{\mathbf{v}_i}$ mais objetiva. Nesse contexto, (POON; POON, 1999) estabelecem o seguinte critério: dado um número inteiro q , com $1 \leq q < \sqrt{n}$, diz-se que um autovetor \mathbf{v} é q -influyente se $|B_{\mathbf{v}}| \geq \frac{q}{\sqrt{n}}$. Equivalentemente, a direção \mathbf{v}_i é dita q -influyente se $\lambda_i^* \geq \frac{q}{\sqrt{n}}$.

Dada a existência de um inteiro positivo k tal que

$$\lambda_1^* \geq \dots \geq \lambda_k^* \geq \frac{q}{\sqrt{n}} > \lambda_{k+1}^* \geq \dots \geq \lambda_n^*, \quad (3.29)$$

tem-se exatamente k autovetores q -influentes. Se $k = 1$, então λ_1^* é um autovalor dominante e \mathbf{v}_1 é uma direção principal. Nesse caso, a análise da relação entre a influência individual e a conjunta se reduz a discussão apresentada na Subseção 3.4.3. Se $k > 1$, desenvolve-se um estudo análogo aquele exibido para situação na qual existem autovalores competitivos, exploraremos isso na Subseção 3.5.4.

Ao final da Subseção 3.4.3, sugeriu-se a curvatura média $\bar{\lambda} = \frac{1}{n}(\lambda_1 + \dots + \lambda_n)$ como valor de referência para julgarmos a magnitude da influência individual, C_j , associada ao j -ésimo vetor básico de perturbação. Podemos proceder de modo similar em relação a B_{e_j} . De fato, primeiramente observemos que

$$\sum_{j=1}^n B_{e_j} = \frac{1}{\sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}} \sum_{j=1}^n C_{e_j} = \frac{\text{tr}(\ddot{\mathbf{F}})}{\sqrt{\text{tr}(\ddot{\mathbf{F}}^2)}} = \frac{\lambda_1 + \dots + \lambda_n}{\sqrt{\lambda_1^2 + \dots + \lambda_n^2}}. \quad (3.30)$$

Logo, se $B_{e_j} = b$ para todo $j = 1, 2, \dots, n$, então $b = \frac{1}{n} \frac{\lambda_1 + \dots + \lambda_n}{\sqrt{\lambda_1^2 + \dots + \lambda_n^2}}$. Sendo assim, b pode ser usado como valor de referência para julgarmos a influência individual, B_{e_j} , correspondente ao j -ésimo vetor básico de perturbação.

3.5.4 Contribuição agregada de vetores básicos de perturbação

O estudo das relações entre influência individual e conjunta apresentado ao longo da Seção 3.4 revela a importância de obtermos a influência associada a cada vetor básico de perturbação e decidirmos quais autovetores são influentes. Nesta subseção, revisitaremos esta temática sob a ótica da curvatura normal conforme. Adotaremos uma notação similar a usada anteriormente, onde as relações entre os vetores básicos de perturbação, \mathbf{e}_j , e os autovetores \mathbf{v}_i da matriz Hessiana de $\ddot{\mathbf{F}}$ eram dadas por

$$\mathbf{v}_i = \sum_{j=1}^n a_{ji} \mathbf{e}_j \quad \text{e} \quad \mathbf{e}_j = \sum_{i=1}^n a_{ji} \mathbf{v}_i. \quad (3.31)$$

Na Subseção 3.5.3 sugeriu-se o uso do número real $\frac{q}{\sqrt{n}}$ como valor de referência para selecionarmos quais autovetores são q -influentes. A fim de detectar quais são os vetores básicos de perturbação que contribuem para que as direções associadas aos autovetores q -influentes sejam alvo de investigação na análise, (POON; POON, 2010) definem a contribuição agregada

do j -ésimo vetor básico de perturbação por

$$B_j(q) = \sum_{i=1}^k \lambda_i^* a_{ji}^2, \quad (3.32)$$

onde o inteiro positivo k é determinado pela da ordenação expressa em (3.29).

Agora, observemos que são válidas as seguintes igualdades:

$$\sum_{j=1}^n B_j(q) = \sum_{j=1}^n \left(\sum_{i=1}^k \lambda_i^* a_{ji}^2 \right) = \sum_{i=1}^k \lambda_i^* \left(\sum_{j=1}^n a_{ji}^2 \right) = \sum_{i=1}^k \lambda_i^*.$$

Logo, se $B_j(q) = b(q)$ para todo vetor básico de perturbação e_j , então

$$b(q) = \frac{1}{n} \sum_{i=1}^k \lambda_i^*. \quad (3.33)$$

(POON; POON, 2010) propõem o uso de $b(q)$ na construção de valores de referência para julgarmos a significância dos vetores básicos na determinação dos autovetores q -influentes.

Por fim, notemos que se fizéssemos q assumir o valor zero, teríamos $b(0) = b$, que está de acordo com o valor de b obtido no final da Subseção (3.5.3). Além disso,

$$B_j(0) = \sum_{i=1}^n \lambda_i^* a_{ji}^2 = \frac{1}{\sqrt{\sum_{i=1}^n \lambda_i^2}} \sum_{i=1}^n \lambda_i a_{ji}^2 = \frac{1}{\sqrt{\sum_{i=1}^n \lambda_i^2}} C_j = B_j. \quad (3.34)$$

Nesse contexto, $B_j(0)$ pode ser interpretada como a contribuição agregada do j -ésimo vetor básico de perturbação para todos os autovetores.

3.6 INFLUÊNCIA LOCAL NO MODELO DE REGRESSÃO EVBS

Nesta seção, derivaremos as expressões que possibilitam a condução de uma análise de influência local sobre o modelo de regressão EVBS sob três diferentes esquemas de perturbação: ponderação de casos, perturbação na resposta e perturbação em uma variável explicativa. Ao final, analisaremos um conjunto de dados meteorológicos com o propósito de ilustrar o potencial da metodologia.

3.6.1 Ponderação de casos

Dadas as variáveis aleatórias contínuas Y_1, \dots, Y_n unidimensionais e independentes tais que, para cada $i = 1, \dots, n$, Y_i possui fdp denotada por $p(y_i; \boldsymbol{\theta})$, sendo $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^\top$ o vetor de parâmetros desconhecidos a serem estimados. Dada uma amostra observada $\mathbf{y} =$

$(y_1, \dots, y_n)^\top$, a função de log-verossimilhança de θ correspondente ao modelo adotado pode ser tomada na forma

$$l(\theta) = \sum_{i=1}^n \log(p(y_i; \theta)) \quad (3.35)$$

No esquema de ponderação de casos, o vetor de perturbação ω tem o número de coordenadas igual ao tamanho da amostra observada, isto é, $q = n$. Além disso, a função de log-verossimilhança associada ao modelo perturbado assume a forma

$$l(\theta|\omega) = \sum_{i=1}^n \omega_i \log(p(y_i; \theta)) \quad (3.36)$$

Assim, o vetor de não-perturbação $\omega_0 = \mathbf{1}_n$, onde $\mathbf{1}_n$ é o vetor $n \times 1$ com todas as coordenadas iguais a 1.

Consideremos o modelo de regressão EVBS, definido em (2.89), as variáveis Y_1, \dots, Y_n são independentes e, para cada $i = 1, \dots, n$, $Y_i \sim \text{log-EVBS}(\alpha, \mathbf{x}_i^\top \beta, \gamma)$. Nossas ações abordarão dois casos distintos. Inicialmente, lidaremos com a situação na qual $\gamma = 0$. Em seguida, trataremos do caso mais geral em que $\gamma \neq 0$. Aqui admitiremos alguns pressupostos que viabilizam o processo de estimação e agregam propriedades ao EMV dos parâmetros do modelo. A primeira hipótese é $\gamma > -1$, isso assegura existência e consistência do EMV na família GEV. A segunda é uma condição imposta sobre o parâmetro α a fim de que possa assegurar unicidade do EMV de θ no modelo definido em (2.89), mais precisamente, assumiremos que $\alpha < 2$. Por fim, com o propósito de que a matriz de informação de Fisher assumira valores finitos, admitiremos a hipótese auxiliar $\gamma < 1/4$.

1. CASO $\gamma = 0$ e α desconhecido, mas supostamente menor do que 2. Neste caso, o vetor de parâmetros desconhecidos é $\theta = (\beta^\top, \alpha)^\top$ e, para fins de estimação, dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ de observações do vetor $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, a função de log-verossimilhança perturbada de θ correspondente ao esquema de ponderação de casos pode ser tomada na forma

$$l(\theta|\omega) = (-\log 2) \sum_{i=1}^n \omega_i + \sum_{i=1}^n \omega_i \log(\xi_{i1}) - \sum_{i=1}^n \omega_i \xi_{i2} - \sum_{i=1}^n \omega_i \exp(-\xi_{i2}), \quad (3.37)$$

onde $\xi_{i1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \beta}{2}\right)$, $\xi_{i2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \beta}{2}\right)$ e $\omega = (\omega_1, \dots, \omega_n)^\top \in \Omega \subset \mathbb{R}^n$

A partir da função de log-verossimilhança perturbada acima, obtemos as seguintes derivadas parciais:

$$\frac{\partial l(\theta|\omega)}{\partial \omega_j} = -\log(2) + \log(\xi_{j1}) - \xi_{j2} - \exp(-\xi_{j2}), \quad \text{para } j = 1, \dots, n. \quad (3.38)$$

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_i \partial \omega_j} = \frac{1}{2} x_{ji} \left(\xi_{j1} (1 - \exp(-\xi_{j2})) - \frac{\xi_{j2}}{\xi_{j1}} \right), \quad i = 0, 1, \dots, p-1 \text{ e } j = 1, \dots, n. \quad (3.39)$$

e

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \alpha \partial \omega_j} = -\frac{1}{\alpha} (1 - \xi_{j2} (1 - \exp(-\xi_{j2}))), \quad j = 1, \dots, n. \quad (3.40)$$

As derivadas parciais acima permite-nos determinar as entradas da matriz $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{ij})_{(p+1) \times n}$; $\boldsymbol{\Delta}_{ij} = \frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_i \partial \omega_j}$ $i = 0, 1, \dots, p$ e $j = 1, 2, \dots, n$, avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, que é dada por $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\beta, \boldsymbol{\Delta}_\alpha)^\top$ cujas entradas são

$$\boldsymbol{\Delta}_\beta = \mathbf{X}^\top \text{diag}\{\hat{b}_{01}, \hat{b}_{02}, \dots, \hat{b}_{0n}\} \quad (3.41)$$

onde $\hat{b}_{0j} = \frac{1}{2} \left(\hat{\xi}_{j1} (1 - \exp(-\hat{\xi}_{j2})) - \frac{\hat{\xi}_{j2}}{\hat{\xi}_{j1}} \right)$, $j = 1, 2, \dots, n$. E

$$\boldsymbol{\Delta}_\alpha = \mathbf{c}_0^\top = (\hat{c}_{01}, \hat{c}_{02}, \dots, \hat{c}_{0n}), \quad (3.42)$$

com $\hat{c}_{0j} = -\frac{1}{\hat{\alpha}} \left(1 - \hat{\xi}_{j2} (1 - \exp(-\hat{\xi}_{j2})) \right)$, para $j = 1, 2, \dots, n$.

2. CASO $\gamma \neq 0$. Dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ de observações do vetor $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, a função de log-verossimilhança perturbada de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha, \gamma)^\top$ correspondente ao esquema de ponderação de casos é

$$l(\boldsymbol{\theta} | \boldsymbol{\omega}) = -\log 2 \sum_{i=1}^n \omega_i + \sum_{i=1}^n \omega_i \log(\xi_{i1}) - (1 + 1/\gamma) \sum_{i=1}^n \omega_i \log(1 + \gamma \xi_{i2}) - \sum_{i=1}^n \omega_i (1 + \gamma \xi_{i2})^{-1/\gamma}, \quad (3.43)$$

onde $\xi_{i1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$, $\xi_{i2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$ e $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top \in \Omega \subset \mathbb{R}^n$

A partir da função de log-verossimilhança, dada pela equação (3.37), obtemos as seguintes derivadas parciais:

$$\frac{\partial l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \omega_j} = -\log(2) + \log(\xi_{j1}) - \left(1 + \frac{1}{\gamma}\right) \log(1 + \gamma \xi_{j2}) - (1 + \gamma \xi_{j2})^{-\frac{1}{\gamma}}, \quad j = 1, 2, \dots, n. \quad (3.44)$$

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_i \partial \omega_j} = \frac{1}{2} x_{ji} \left[\frac{\xi_{j1}}{1 + \gamma \xi_{j2}} \left(\gamma + 1 - (1 + \gamma \xi_{j2})^{-\frac{1}{\gamma}} \right) - \frac{\xi_{j2}}{\xi_{j1}} \right], \quad (3.45)$$

com $i = 0, 1, \dots, p-1$ e $j = 1, 2, \dots, n$.

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \alpha \partial \omega_j} = -\frac{1}{\alpha} + \frac{1}{\alpha} \frac{\xi_{j2}}{1 + \gamma \xi_{j2}} \left(\gamma + 1 - (1 + \gamma \xi_{j2})^{-\frac{1}{\gamma}} \right), \text{ para } j = 1, \dots, n. \quad (3.46)$$

e

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \gamma \partial \omega_j} = \frac{1}{\gamma^2} \log(1 + \gamma \xi_{j2}) \left(1 - (1 + \gamma \xi_{j2})^{-\frac{1}{\gamma}} \right) - \frac{1}{\gamma} \frac{\xi_{j2}}{1 + \gamma \xi_{j2}} \left(\gamma + 1 - (1 + \gamma \xi_{j2})^{-\frac{1}{\gamma}} \right), \quad (3.47)$$

para $j = 1, 2, \dots, n$.

Consequentemente, ao considerarmos o esquema de ponderação de casos, a matriz $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{ij})_{(p+2) \times n}$; $\boldsymbol{\Delta}_{ij} = \frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_i \partial \omega_j}$ $i = 0, 1, \dots, p+1$ e $j = 1, 2, \dots, n$, avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, é dada por $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\beta, \boldsymbol{\Delta}_\alpha, \boldsymbol{\Delta}_\gamma)^\top$ cujas entradas são:

$$\boldsymbol{\Delta}_\beta = \mathbf{X}^\top \text{diag}\{\hat{b}_1, \hat{b}_2, \dots, \hat{b}_n\} \quad (3.48)$$

onde $\hat{b}_j = \frac{1}{2} \left[\frac{\hat{\xi}_{j1}}{1 + \hat{\gamma} \hat{\xi}_{j2}} \left(\hat{\gamma} + 1 - (1 + \hat{\gamma} \hat{\xi}_{j2})^{-1/\hat{\gamma}} \right) - \frac{\hat{\xi}_{j2}}{\hat{\xi}_{j1}} \right]$, $j = 1, 2, \dots, n$.

$$\boldsymbol{\Delta}_\alpha = \mathbf{c}^\top = (\hat{c}_1, \hat{c}_2, \dots, \hat{c}_n), \quad (3.49)$$

com $\hat{c}_j = -\frac{1}{\hat{\alpha}} + \frac{1}{\hat{\alpha}} \frac{\hat{\xi}_{j2}}{(1 + \hat{\gamma} \hat{\xi}_{j2})} \left(\hat{\gamma} + 1 - (1 + \hat{\gamma} \hat{\xi}_{j2})^{-\frac{1}{\hat{\gamma}}} \right)$, para $j = 1, 2, \dots, n$. E,

$$\boldsymbol{\Delta}_\gamma = \mathbf{d}^\top = (\hat{d}_1, \hat{d}_2, \dots, \hat{d}_n), \quad (3.50)$$

onde $\hat{d}_j = \frac{1}{\hat{\gamma}^2} \log(1 + \hat{\gamma} \hat{\xi}_{j2}) \left(1 - (1 + \hat{\gamma} \hat{\xi}_{j2})^{-\frac{1}{\hat{\gamma}}} \right) - \frac{1}{\hat{\gamma}} \frac{\hat{\xi}_{j2}}{(1 + \hat{\gamma} \hat{\xi}_{j2})} \left(\hat{\gamma} + 1 - (1 + \hat{\gamma} \hat{\xi}_{j2})^{-\frac{1}{\hat{\gamma}}} \right)$, para $j = 1, 2, \dots, n$.

3.6.2 Perturbação na resposta

Neste caso, sujeitamos a resposta y_i a uma perturbação da forma $y_{i\omega} = y_i + \omega_i s_y$, em que $\omega_i \in \mathbb{R}$ para todo $i = 1, 2, \dots, n$ e s_y é um fator de escala. A exposição que faremos nesta subsecção segue o mesmo roteiro da anterior e os pressupostos fixados são similares.

1. Caso $\gamma = 0$. Ao considerarmos o modelo de regressão linear definido em (2.89) e o tipo de perturbação descrito anteriormente, a função de log-verossimilhança perturbada de

$\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$, dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ de observações, é dada por

$$l(\boldsymbol{\theta} | \boldsymbol{\omega}) = -n \log 2 + \sum_{i=1}^n \log(\xi_{i1\omega_r}) - \sum_{i=1}^n \xi_{i2\omega_r} - \sum_{i=1}^n \exp(-\xi_{i2\omega_r}), \quad (3.51)$$

onde $\xi_{i1\omega_r} = \frac{2}{\alpha} \cosh\left(\frac{y_i + \omega_i s_y - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$, $\xi_{i2\omega_r} = \frac{2}{\alpha} \sinh\left(\frac{y_i + \omega_i s_y - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$ e $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top \in \Omega \subset \mathbb{R}^n$. Levando-nos as seguintes derivadas parciais:

$$\frac{\partial l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \omega_j} = \frac{1}{2} s_y \left[\frac{\xi_{j2\omega_r}}{\xi_{j1\omega_r}} - \xi_{j1\omega_r} (1 - \exp(-\xi_{j2\omega_r})) \right], \quad \text{para } j = 1, \dots, n. \quad (3.52)$$

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_i \partial \omega_j} = -\frac{1}{4} s_y x_{ji} \left[1 - \left(\frac{\xi_{j2\omega_r}}{\xi_{j1\omega_r}} \right)^2 - \xi_{j2\omega_r} (1 - \exp(-\xi_{j2\omega_r})) - \xi_{j1\omega_r}^2 \exp(-\xi_{j2\omega_r}) \right], \quad (3.53)$$

com $i = 0, 1, \dots, p-1$ e $j = 1, \dots, n$.

e

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \alpha \partial \omega_j} = \frac{1}{2\alpha} s_y \xi_{j1\omega_r} [1 - \exp(-\xi_{j2\omega_r})(1 - \xi_{j2\omega_r})], \quad \text{com } j = 1, \dots, n. \quad (3.54)$$

Dessa forma, ao perturbarmos a resposta y_i conforme o esquema descrito, a matriz $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{ij})_{(p+1) \times n}$; $\boldsymbol{\Delta}_{ij} = \frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_i \partial \omega_j}$ $i = 0, 1, \dots, p$ e $j = 1, 2, \dots, n$, avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, 0, \dots, 0)^\top \in \mathbb{R}^n$, é dada por $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\beta, \boldsymbol{\Delta}_\alpha)^\top$ cujas entradas são

$$\boldsymbol{\Delta}_\beta = \mathbf{X}^\top \text{diag}\{\hat{m}_{01}, \hat{m}_{02}, \dots, \hat{m}_{0n}\} \quad (3.55)$$

onde $\hat{m}_{0j} = -\frac{1}{4} s_y \left[1 - \left(\frac{\hat{\xi}_{j2\omega_r}}{\hat{\xi}_{j1\omega_r}} \right)^2 - \hat{\xi}_{j2\omega_r} (1 - \exp(-\hat{\xi}_{j2\omega_r})) - \hat{\xi}_{j1\omega_r}^2 \exp(-\hat{\xi}_{j2\omega_r}) \right]$, $j = 1, 2, \dots, n$.

$$\boldsymbol{\Delta}_\alpha = \mathbf{z}_0^\top = (\hat{z}_{01}, \hat{z}_{02}, \dots, \hat{z}_{0n}), \quad (3.56)$$

com $\hat{z}_{0j} = \frac{1}{2\hat{\alpha}} s_y \hat{\xi}_{j1\omega_r} [1 - \exp(-\hat{\xi}_{j2\omega_r})(1 - \hat{\xi}_{j2\omega_r})]$, para $j = 1, 2, \dots, n$.

2. CASO $\gamma \neq 0$. Dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ de realizações do vetor $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, a função de log-verossimilhança perturbada de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha, \gamma)^\top$ correspondente ao esquema de perturbação na resposta é

$$l(\boldsymbol{\theta} | \boldsymbol{\omega}) = -n \log 2 + \sum_{i=1}^n \log(\xi_{i1\omega_r}) - (1 + 1/\gamma) \sum_{i=1}^n \log(1 + \gamma \xi_{i2\omega_r}) - \sum_{i=1}^n (1 + \gamma \xi_{i2\omega_r})^{-1/\gamma}, \quad (3.57)$$

onde $\xi_{i1\omega_r} = \frac{2}{\alpha} \cosh\left(\frac{y_i + \omega_i s_y - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$, $\xi_{i2\omega_r} = \frac{2}{\alpha} \sinh\left(\frac{y_i + \omega_i s_y - \mathbf{x}_i^\top \boldsymbol{\beta}}{2}\right)$ e $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top \in \Omega \subset \mathbb{R}^n$.

A partir da função de log-verossimilhança, dada pela equação (3.57), obtemos

$$\frac{\partial l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \omega_j} = \frac{1}{2} s_y \left[\frac{\xi_{j1\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \left(1 + \gamma - (1 + \gamma \xi_{j2\omega_r})^{-1/\gamma}\right) - \frac{\xi_{j2\omega_r}}{\xi_{j2\omega_r}} \right], \quad j = 1, 2, \dots, n. \quad (3.58)$$

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_i \partial \omega_j} = \frac{1}{4} s_y x_{ji} \left[1 - \frac{\xi_{j1\omega_r}^2}{\xi_{j2\omega_r}^2} - \frac{\xi_{j2\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \left(1 + \gamma - (1 + \gamma \xi_{j2\omega_r})^{-\frac{1}{\gamma}}\right) + \right. \quad (3.59)$$

$$\left. + (1 + \gamma) \left(\frac{\xi_{j1\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \right)^2 \left(\gamma - (1 + \gamma \xi_{j2\omega_r})^{-\frac{1}{\gamma}} \right) \right], \quad (3.60)$$

com $i = 0, 1, \dots, p-1$ e $j = 1, 2, \dots, n$.

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \alpha \partial \omega_j} = -\frac{1}{2\alpha} s_y \left(\frac{\xi_{j1\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \right) \left(1 + \gamma - (1 + \gamma \xi_{j2\omega_r})^{-\frac{1}{\gamma}}\right) + \quad (3.61)$$

$$+ \frac{1}{2\alpha} s_y (1 + \gamma) \left(\frac{\xi_{j1\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \right) \left(\frac{\xi_{j2\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \right) \left(\gamma - (1 + \gamma \xi_{j2\omega_r})^{-\frac{1}{\gamma}} \right), \quad (3.62)$$

para $j = 1, \dots, n$. E,

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \gamma \partial \omega_j} = \frac{1}{2} s_y \left(\frac{\xi_{j1\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \right) \left(1 - \frac{1}{\gamma^2} \log(1 + \gamma \xi_{j2\omega_r}) \right) + \quad (3.63)$$

$$- \frac{1}{2} s_y (1 + \gamma) \left(\frac{\xi_{j1\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \right) \left(\frac{\xi_{j2\omega_r}}{1 + \gamma \xi_{j2\omega_r}} \right) \left(1 - \frac{1}{\gamma} (1 + \gamma \xi_{j2\omega_r})^{-\frac{1}{\gamma}} \right), \quad (3.64)$$

para $j = 1, \dots, n$. Logo, a matriz $\boldsymbol{\Delta} = (\Delta_{ij})_{(p+2) \times n}$; $\Delta_{ij} = \frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_i \partial \omega_j}$ $i = 0, 1, \dots, p+1$ e $j = 1, 2, \dots, n$, avaliada em em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, é dada por $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\beta, \boldsymbol{\Delta}_\alpha, \boldsymbol{\Delta}_\gamma)^\top$ cujas entradas são:

$$\boldsymbol{\Delta}_\beta = \mathbf{X}^\top \text{diag}\{\hat{m}_1, \hat{m}_2, \dots, \hat{m}_n\}, \quad (3.65)$$

onde

$$\hat{m}_j = \frac{1}{4} s_y \left[1 - \frac{\hat{\xi}_{j1\omega_r}^2}{\hat{\xi}_{j2\omega_r}^2} - \frac{\hat{\xi}_{j2\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} (1 + \hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_r})^{-\frac{1}{\hat{\gamma}}}) + \right. \\ \left. + (1 + \hat{\gamma}) \left(\frac{\hat{\xi}_{j1\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right)^2 (\hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_r})^{-\frac{1}{\hat{\gamma}}}) \right], \text{ para } j = 1, \dots, n.$$

$$\Delta_\alpha = \mathbf{z}^\top = (\hat{z}_1, \dots, \hat{z}_n), \quad (3.66)$$

com

$$\hat{z}_j = -\frac{1}{2\hat{\alpha}} s_y \left(\frac{\hat{\xi}_{j1\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) (1 + \hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_r})^{-\frac{1}{\hat{\gamma}}}) + \\ + \frac{1}{2\hat{\alpha}} s_y (1 + \hat{\gamma}) \left(\frac{\hat{\xi}_{j1\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) \left(\frac{\hat{\xi}_{j2\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) (\hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_r})^{-\frac{1}{\hat{\gamma}}}),$$

$j = 1, 2, \dots, n$. E, por fim temos

$$\Delta_\gamma = \mathbf{v}^\top = (\hat{v}_1, \dots, \hat{v}_n), \quad (3.67)$$

onde

$$\hat{v}_j = \frac{1}{2} s_y \left(\frac{\hat{\xi}_{j1\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) \left(1 - \frac{1}{\hat{\gamma}^2} \log(1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}) \right) + \\ - \frac{1}{2} s_y (1 + \hat{\gamma}) \left(\frac{\hat{\xi}_{j1\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) \left(\frac{\hat{\xi}_{j2\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) \left(1 - \frac{1}{\hat{\gamma}} (1 + \hat{\gamma} \hat{\xi}_{j2\omega_r})^{-\frac{1}{\hat{\gamma}}} \right),$$

para $j = 1, 2, \dots, n$.

3.6.3 Perturbação em uma covariável

Admitamos que no modelo definido pela equação (2.89) figure uma variável contínua e a denotemos por \mathbf{x}_t . Impomos sobre \mathbf{x}_t uma perturbação aditiva da forma $x_{it\omega} = x_{it} + \omega_i s_x$, onde $t \in \{0, 1, \dots, p-1\}$ e s_x é um fator de escala, por exemplo o desvio padrão correspondente às observações relativas à variável \mathbf{x}_t . Em (COOK, 1986), acentua-se que pequenas perturbações em covariáveis de um modelo de regressão linear podem influenciar gravemente os resultados de uma análise inferencial baseada em mínimos quadrados quando a colinearidade entre covariáveis está presente. Logo, se uma dada covariável é perturbada e

constatarmos que não ocorrem mudanças inferenciais significativas, então há indícios de que tal covariável não se relaciona, linearmente, com nenhuma das outras que figuram no modelo postulado. A seguir, obteremos as expressões que possibilitam o cálculo da curvatura normal neste esquema de perturbação.

1. Caso $\gamma = 0$. Ao impormos uma perturbação sobre uma única variável explicativa contínua conforme o esquema descrito acima, a função de log-verossimilhança perturbada de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$, dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ de observações, é

$$l(\boldsymbol{\theta} | \boldsymbol{\omega}) = -n \log 2 + \sum_{i=1}^n \log(\xi_{i1\omega_c}) - \sum_{i=1}^n \xi_{i2\omega_c} - \sum_{i=1}^n \exp(-\xi_{i2\omega_c}), \quad (3.68)$$

onde $\xi_{i1\omega_c} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \beta_i \omega_i s_x}{2}\right)$, $\xi_{i2\omega_c} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \beta_i \omega_i s_x}{2}\right)$ e $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top \in \Omega \subset \mathbb{R}^n$.

Logo,

$$\frac{\partial l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \omega_j} = -\frac{1}{2} s_x \beta_t \left[\frac{\xi_{j2\omega_c}}{\xi_{j1\omega_c}} - \xi_{j1\omega_c} (1 - \exp(-\xi_{j2\omega_c})) \right], \quad \text{para } j = 1, \dots, n. \quad (3.69)$$

Se $i \neq t$, temos

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_i \partial \omega_j} = \frac{1}{4} s_x \beta_t x_{ji} \left[1 - \left(\frac{\xi_{j2\omega_c}}{\xi_{j1\omega_c}} \right)^2 - \xi_{j2\omega_c} (1 - \exp(-\xi_{j2\omega_c})) - \xi_{j1\omega_c}^2 \exp(-\xi_{j2\omega_c}) \right], \quad (3.70)$$

com $i = 0, 1, \dots, p-1$ e $j = 1, \dots, n$. Caso $i = t$, obtemos

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_i \partial \omega_j} &= \frac{1}{4} s_x \beta_t x_{ji} \left[1 - \left(\frac{\xi_{j2\omega_c}}{\xi_{j1\omega_c}} \right)^2 - \xi_{j2\omega_c} (1 - \exp(-\xi_{j2\omega_c})) - \xi_{j1\omega_c}^2 \exp(-\xi_{j2\omega_c}) \right] + \\ &\quad - \frac{1}{2} s_x \left[\frac{\xi_{j2\omega_c}}{\xi_{j2\omega_c}} - \xi_{j1\omega_c} (1 - \exp(-\xi_{j2\omega_c})) \right], \end{aligned}$$

com $i = 0, 1, \dots, p-1$ e $j = 1, \dots, n$.

E,

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \alpha \partial \omega_j} = -\frac{1}{2\alpha} s_x \beta_t \xi_{j1\omega_c} [1 - \exp(-\xi_{j2\omega_c}) (1 - \xi_{j2\omega_c})], \quad \text{com } j = 1, \dots, n. \quad (3.71)$$

Assim, a matriz $\Delta = (\Delta_{ij})_{(p+1) \times n}$; $\Delta_{ij} = \frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_i \partial \omega_j}$ $i = 0, 1, \dots, p$ e $j = 1, 2, \dots, n$, correspondente ao esquema de perturbação na covariável contínua \mathbf{x}_t , avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, 0, \dots, 0)^\top \in \mathbb{R}^n$, é $\Delta = (\Delta_\beta, \Delta_\alpha)^\top$ cujas entradas são

$$\Delta_\beta = \mathbf{X}^\top \text{diag}\{\hat{\kappa}_{01}, \hat{\kappa}_{02}, \dots, \hat{\kappa}_{0n}\} \quad (3.72)$$

De modo que se $i \neq t$, então

$$\hat{\kappa}_{0j} = \frac{1}{4} s_x \hat{\beta}_t \left[1 - \left(\frac{\hat{\xi}_{j2\omega_c}}{\hat{\xi}_{j1\omega_c}} \right)^2 - \hat{\xi}_{j2\omega_c} (1 - \exp(-\hat{\xi}_{j2\omega_c})) - \hat{\xi}_{j1\omega_c}^2 \exp(-\hat{\xi}_{j2\omega_c}) \right], \quad (3.73)$$

para $j = 1, 2, \dots, n$. Por outro lado, se $i = t$, então

$$\begin{aligned} \hat{\kappa}_{0j} = & \frac{1}{4} s_x \hat{\beta}_t \left[1 - \left(\frac{\hat{\xi}_{j2\omega_c}}{\hat{\xi}_{j1\omega_c}} \right)^2 - \hat{\xi}_{j2\omega_c} (1 - \exp(-\hat{\xi}_{j2\omega_c})) - \hat{\xi}_{j1\omega_c}^2 \exp(-\hat{\xi}_{j2\omega_c}) \right] + \\ & - \frac{1}{2} s_x \left[\frac{\hat{\xi}_{j2\omega_c}}{\hat{\xi}_{j2\omega_c}} - \hat{\xi}_{j1\omega_c} (1 - \exp(-\hat{\xi}_{j2\omega_c})) \right], \end{aligned}$$

para $j = 1, 2, \dots, n$.

E,

$$\Delta_\alpha = \boldsymbol{\tau}_0^\top = (\hat{\tau}_{01}, \hat{\tau}_{02}, \dots, \hat{\tau}_{0n}), \quad (3.74)$$

com $\hat{\tau}_{0j} = -\frac{1}{2\hat{\alpha}} s_x \hat{\xi}_{j1\omega_c} [1 - \exp(-\hat{\xi}_{j2\omega_c})(1 - \hat{\xi}_{j2\omega_c})]$, para $j = 1, 2, \dots, n$.

2. CASO $\gamma \neq 0$. Dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ de realizações do vetor $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, a função de log-verossimilhança perturbada de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha, \gamma)^\top$ correspondente ao esquema de perturbação na covariável é

$$l(\boldsymbol{\theta} | \boldsymbol{\omega}) = -n \log 2 + \sum_{i=1}^n \log(\xi_{i1\omega_c}) - (1 + 1/\gamma) \sum_{i=1}^n \log(1 + \gamma \xi_{i2\omega_c}) - \sum_{i=1}^n (1 + \gamma \xi_{i2\omega_c})^{-1/\gamma}, \quad (3.75)$$

onde $\xi_{i1\omega_c} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \beta_t \omega_i s_x}{2}\right)$, $\xi_{i2\omega_c} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \beta_t \omega_i s_x}{2}\right)$ e $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top \in \Omega \subset \mathbb{R}^n$.

A partir da função de log-verossimilhança, dada pela equação (3.75), obtemos

$$\frac{\partial l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \omega_j} = \frac{1}{2} s_x \beta_t \left[\frac{\xi_{j1\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \left(1 + \gamma - (1 + \gamma \xi_{j2\omega_c})^{-1/\gamma} \right) - \frac{\xi_{j2\omega_c}}{\xi_{j2\omega_c}} \right], \quad j = 1, 2, \dots, n. \quad (3.76)$$

Ademais, se $i \neq t$, então

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_i \partial \omega_j} = \frac{1}{4} s_x \beta_t x_{ji} \left[1 - \frac{\xi_{j1\omega_c}^2}{\xi_{j2\omega_c}^2} - \frac{\xi_{j2\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \left(1 + \gamma - (1 + \gamma \xi_{j2\omega_c})^{-\frac{1}{\gamma}} \right) + \right. \quad (3.77)$$

$$\left. + (1 + \gamma) \left(\frac{\xi_{j1\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \right)^2 \left(\gamma - (1 + \gamma \xi_{j2\omega_c})^{-\frac{1}{\gamma}} \right) \right], \quad (3.78)$$

com $i = 0, 1, \dots, p-1$ e $j = 1, 2, \dots, n$. Por outro lado, se $i = t$, então

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \beta_i \partial \omega_j} = \frac{1}{4} s_x \beta_t x_{ji} \left[1 - \frac{\xi_{j1\omega_c}^2}{\xi_{j2\omega_c}^2} - \frac{\xi_{j2\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \left(1 + \gamma - (1 + \gamma \xi_{j2\omega_c})^{-\frac{1}{\gamma}} \right) + \right. \quad (3.79)$$

$$\left. + (1 + \gamma) \left(\frac{\xi_{j1\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \right)^2 \left(\gamma - (1 + \gamma \xi_{j2\omega_c})^{-\frac{1}{\gamma}} \right) \right] + \quad (3.80)$$

$$- \frac{1}{2} s_x \left[\frac{\xi_{j2\omega_c}}{\xi_{j1\omega_c}} - \frac{\xi_{j1\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \left(1 + \gamma - (1 + \gamma \xi_{j2\omega_c})^{-1/\gamma} \right) \right], \quad (3.81)$$

com $i = 0, 1, \dots, p-1$ e $j = 1, 2, \dots, n$.

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \alpha \partial \omega_j} = -\frac{1}{2\alpha} s_x \beta_t \left(\frac{\xi_{j1\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \right) \left(1 + \gamma - (1 + \gamma \xi_{j2\omega_c})^{-\frac{1}{\gamma}} \right) + \quad (3.82)$$

$$+ \frac{1}{2\alpha} s_x \beta_t (1 + \gamma) \left(\frac{\xi_{j1\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \right) \left(\frac{\xi_{j2\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \right) \left(\gamma - (1 + \gamma \xi_{j2\omega_c})^{-\frac{1}{\gamma}} \right), \quad (3.83)$$

para $j = 1, \dots, n$. E,

$$\frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \gamma \partial \omega_j} = \frac{1}{2} s_x \beta_t \left(\frac{\xi_{j1\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \right) \left(1 - \frac{1}{\gamma^2} \log(1 + \gamma \xi_{j2\omega_c}) \right) + \quad (3.84)$$

$$- \frac{1}{2} s_x \beta_t (1 + \gamma) \left(\frac{\xi_{j1\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \right) \left(\frac{\xi_{j2\omega_c}}{1 + \gamma \xi_{j2\omega_c}} \right) \left(1 - \frac{1}{\gamma} (1 + \gamma \xi_{j2\omega_c})^{-\frac{1}{\gamma}} \right), \quad (3.85)$$

para $j = 1, \dots, n$. Logo, a matriz $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_{ij})_{(p+2) \times n}$; $\boldsymbol{\Delta}_{ij} = \frac{\partial^2 l(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \theta_i \partial \omega_j}$ $i = 0, 1, \dots, p+1$ e $j = 1, 2, \dots, n$, avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, é

$$\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\beta, \boldsymbol{\Delta}_\alpha, \boldsymbol{\Delta}_\gamma)^\top,$$

cujas entradas são

$$\Delta_{\beta} = \mathbf{X}^{\top} \text{diag}\{\hat{\kappa}_1, \hat{\kappa}_2, \dots, \hat{\kappa}_n\}, \quad (3.86)$$

De modo que se $i \neq t$, então

$$\begin{aligned} \hat{\kappa}_j = & \frac{1}{4} s_x \hat{\beta}_t \left[1 - \frac{\hat{\xi}_{j1\omega_c}^2}{\hat{\xi}_{j2\omega_c}^2} - \frac{\hat{\xi}_{j2\omega_c}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_c}} (1 + \hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_c})^{-\frac{1}{\hat{\gamma}}}) + \right. \\ & \left. + (1 + \hat{\gamma}) \left(\frac{\hat{\xi}_{j1\omega_c}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_c}} \right)^2 (\hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_c})^{-\frac{1}{\hat{\gamma}}}) \right]. \end{aligned}$$

Porém, se $i = t$, então

$$\begin{aligned} \hat{\kappa}_j = & \frac{1}{4} s_x \hat{\beta}_t \left[1 - \frac{\hat{\xi}_{j1\omega_c}^2}{\hat{\xi}_{j2\omega_c}^2} - \frac{\hat{\xi}_{j2\omega_c}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_c}} (1 + \hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_c})^{-\frac{1}{\hat{\gamma}}}) + \right. \\ & \left. + (1 + \hat{\gamma}) \left(\frac{\hat{\xi}_{j1\omega_c}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_c}} \right)^2 (\hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_c})^{-\frac{1}{\hat{\gamma}}}) \right] + \\ & - \frac{1}{2} s_x \left[\frac{\hat{\xi}_{j2\omega_c}}{\hat{\xi}_{j1\omega_c}} - \frac{\hat{\xi}_{j1\omega_c}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_c}} (1 + \hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_c})^{-1/\hat{\gamma}}) \right], \end{aligned}$$

com $i = 0, 1, \dots, p-1$ e $j = 1, 2, \dots, n$.

Além disso,

$$\Delta_{\alpha} = \tau^{\top} = (\hat{\tau}_1, \dots, \hat{\tau}_n), \quad (3.87)$$

com

$$\begin{aligned} \hat{\tau}_j = & -\frac{1}{2\hat{\alpha}} s_x \hat{\beta}_t \left(\frac{\hat{\xi}_{j1\omega_c}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_c}} \right) (1 + \hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_c})^{-\frac{1}{\hat{\gamma}}}) + \\ & + \frac{1}{2\hat{\alpha}} s_x \hat{\beta}_t (1 + \hat{\gamma}) \left(\frac{\hat{\xi}_{j1\omega_c}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_c}} \right) \left(\frac{\hat{\xi}_{j2\omega_c}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_c}} \right) (\hat{\gamma} - (1 + \hat{\gamma} \hat{\xi}_{j2\omega_c})^{-\frac{1}{\hat{\gamma}}}), \end{aligned}$$

para $j = 1, \dots, n$.

E, por fim temos

$$\Delta_{\gamma} = \mathbf{u}^{\top} = (\hat{u}_1, \dots, \hat{u}_n), \quad (3.88)$$

onde

$$\hat{u}_j = \frac{1}{2} s_x \hat{\beta}_t \left(\frac{\hat{\xi}_{j1\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) \left(1 - \frac{1}{\hat{\gamma}^2} \log(1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}) \right) + \\ - \frac{1}{2} s_x \hat{\beta}_t (1 + \hat{\gamma}) \left(\frac{\hat{\xi}_{j1\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) \left(\frac{\hat{\xi}_{j2\omega_r}}{1 + \hat{\gamma} \hat{\xi}_{j2\omega_r}} \right) \left(1 - \frac{1}{\hat{\gamma}} (1 + \hat{\gamma} \hat{\xi}_{j2\omega_r})^{-\frac{1}{\hat{\gamma}}} \right),$$

para $j = 1, 2, \dots, n$.

3.7 APLICAÇÃO

A fim de ilustrar a aplicabilidade da metodologia de influência local em detectar observações influentes e o potencial do modelo de regressão EVBS na modelagem de dados extremos, consideramos um conjunto de dados meteorológicos registrados pela Estação A-868, Latitude -26.95083, Longitude -48.76194 e Altitude 9.76 m, localizada em Itajaí-Santa Catarina-Brasil. Os dados foram obtidos junto ao Instituto Nacional de Meteorologia (INMET), no site <<https://portal.inmet.gov.br>>, e exprimem a velocidade máxima mensal do vento (m/s) em termos da pressão atmosférica média diária (mb), correspondente ao dia em que se registrou a rajada máxima, de julho-2010 até outubro-2020, resultando em uma amostra de tamanho $n = 124$, representada na Figura 11 (a). As estatísticas descritivas contidas na Tabela 22 e o Box-plot, exibido na Figura 11 (b), evidenciam que a distribuição da variável aleatória velocidade máxima mensal do vento é assimétrica positiva.

Tabela 22 – Estatísticas descritivas para velocidade máxima mensal do vento, m/s, registrada pela Estação A-868 em Itajaí-SC-Brasil.

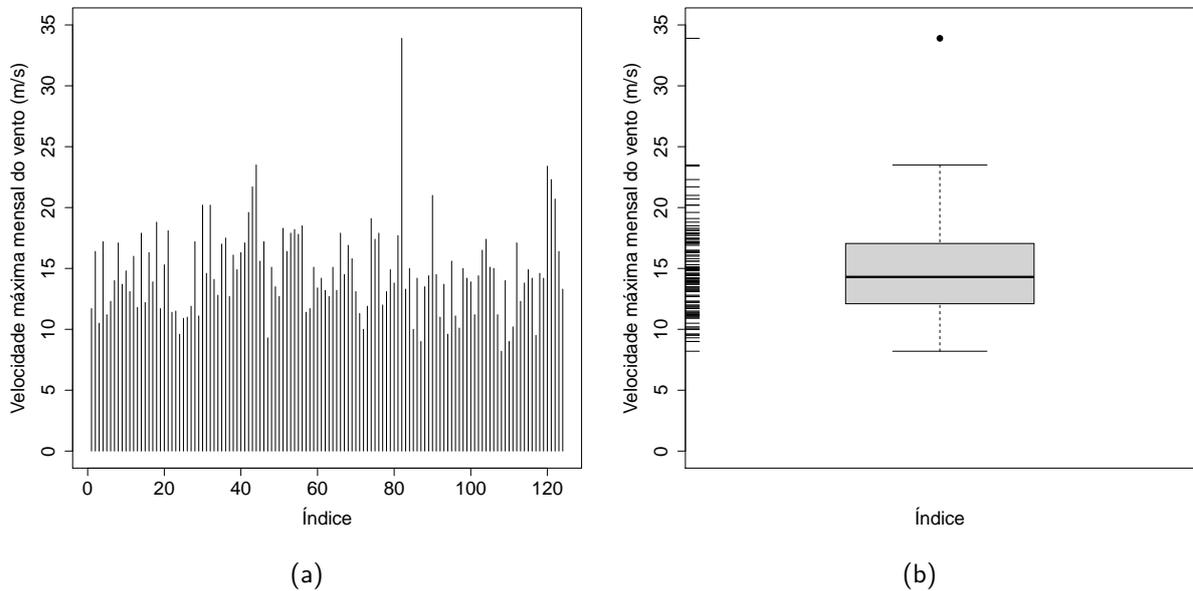
n	Mínimo	Mediana	Média	Máximo	Desvio	Assimetria	Curtose
124	8.20	14.30	14.73	33.90	3.63	1.41	5.01

Fonte: Elaborada pelo autor (2022)

Além disso, as Figuras 12 (a) e 12 (b), que exibem os gráficos das funções de autocorrelação amostral e parcial, respectivamente, corroboram com a hipótese de leve dependência linear entre as observações. De modo que admitiremos a hipótese de independência entre os valores registrados.

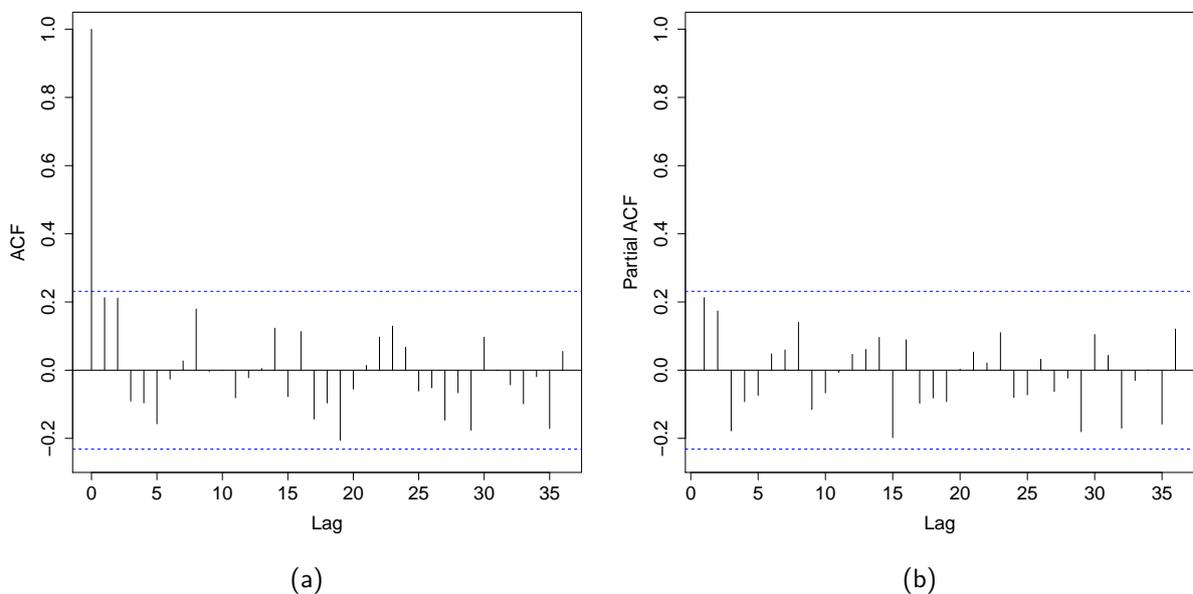
Em nossa análise, postulamos um modelo de regressão EVBS no qual a variável resposta, Y , é o logaritmo natural da velocidade máxima mensal do vento, T , e a variável explicativa é

Figura 11 – Gráfico de dispersão das velocidades máximas mensais do vento registradas na estação Itajaí-SC em (a) e o respectivo Box-plot em (b).



Fonte: Elaborada pelo autor (2022)

Figura 12 – Gráficos das funções de autocorrelação amostral em (a) e autocorrelação parcial em (b) das velocidades máximas mensais do vento registradas na estação Itajaí-SC.



Fonte: Elaborada pelo autor (2022)

a pressão atmosférica média diária, X . Mais precisamente, admitimos o modelo

$$Y_i = \log(T_i) = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, 124, \quad (3.89)$$

em que $\varepsilon_i = \log(\delta_i) \sim \log\text{-EVBS}(\alpha, 0, \gamma)$. Ao realizarmos um ajuste preliminar, obtivemos as

estimativas de MV para os parâmetros, acompanhadas do erro-padrão do EMV e o nível descritivo (p -valor), exibidos na Tabela 23. Destaquemos que o coeficiente associado à variável explicativa é significativo ao nível nominal de 5%.

Tabela 23 – Estimativa de MV dos parâmetros do modelo acompanhada dos respectivos erro padrão e nível descritivo dos coeficientes da regressão.

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
emv	25.5148	-0.0227	0.1857	-0.1551
se	3.3338	0.0033	0.0127	0.0472
p -valor	0.0000	0.0000	—	—

Fonte: Elaborada pelo autor (2022)

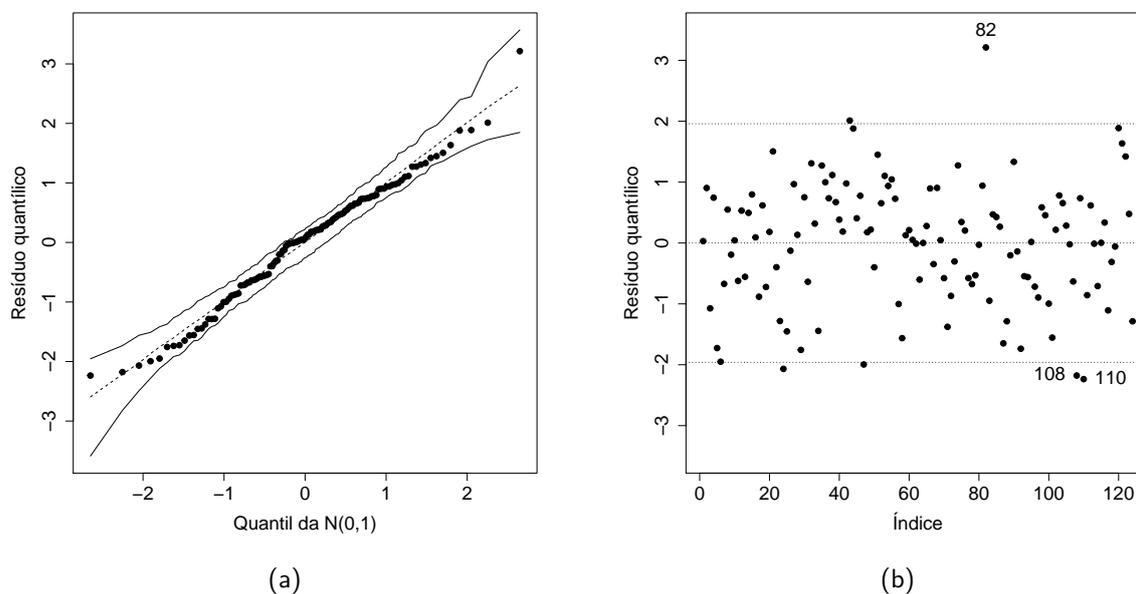
No que tange à adequação do modelo de regressão EVBS aos dados sob análise, investigamos o comportamento do resíduo quantílico aleatorizado, que segundo (LEIVA et al., 2016) mostra-se eficaz para essa finalidade. A Figura 13 (a) exibe o gráfico normal de probabilidades com envelope para o resíduo quantílico, sugerindo que não há indícios que nos leve ao afastamento da hipótese de que a resposta segue uma distribuição log-EVBS. Além disso, ao realizarmos os testes Shapiro-Wilk e Kolmogorov-Smirnov sobre os resíduos, ilustrados na 13 (b), obtivemos os p -valores 0.3021 e 0.4525, respectivamente. Logo, decidimos pela não rejeição da hipótese de normalidade dos resíduos quantílicos, ratificando a adequação do modelo de regressão EVBS aos dados sob estudo.

Com o objetivo de conduzirmos uma análise de influência local e, conseqüentemente, identificarmos observações que são influentes, perturbamos o modelo segundo o esquema de ponderação de casos. Sob este esquema de perturbação, os autovalores normalizados são 0.69678, 0.49511, 0.44547 e 0.26630, representados na Figura 14 (a). Logo, constatamos que associado ao maior autovalor tem-se um o autovetor 7-influente, indicando a direção na qual a curvatura conforme é máxima.

Ao fazermos uso da equação (3.32), obtemos a contribuição agregada dos vetores básicos de perturbação para o autovetor 7-influente, ou seja, a contribuição de cada direção canônica ao valor máximo da curvatura, isso é ilustrado na Figura 14 (b). Além disso, se usarmos $b(q = 7) = \frac{1}{124}(0.69678) = 0.0056$ na construção de um valor de referência, como sugere (POON; POON, 2010), para julgarmos a contribuição agregada de cada direção canônica de perturbação, classificamos as observações correspondentes aos números 82 e 108 como sendo potencialmente influentes, conforme descreve a Figura 14 (b).

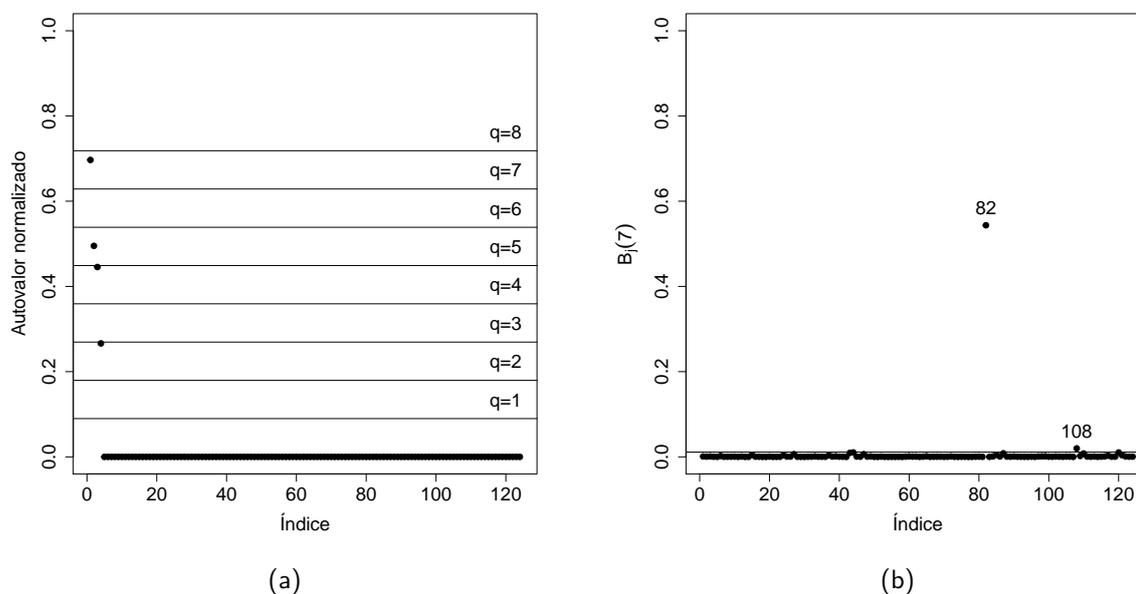
Agora, com o objetivo de verificar o impacto que as observações correspondentes aos nú-

Figura 13 – Gráfico normal de probabilidades, com envelope, para o resíduo quantílico em (a) e os resíduos quantílicos correspondentes às observações em (b).



Fonte: Elaborada pelo autor (2022)

Figura 14 – Autovalores normalizados em (a); Contribuição agregada dos vetores básicos de perturbação em (b), sob ponderação de casos



Fonte: Elaborada pelo autor (2022)

meros 82 e 108 exercem sobre as estimativas dos parâmetros, reajustamos o modelo, excluindo individualmente cada uma delas. Na Tabela 24, a seguir, constam as novas estimativas de MV, a respectiva taxa de variação da grandeza, em porcentagem, denotada por R_{θ_j} e definida por

$R_{\theta_j} = \left(\frac{\hat{\theta}_{j(i)} - \hat{\theta}_j}{|\hat{\theta}_j|} \right) \cdot 100$, em que $\hat{\theta}_{j(i)}$ denota a estimativa de $\hat{\theta}_j$ obtida após exclusão da i -ésima observação, e o nível descritivo (p -valor) associado aos coeficientes da regressão.

Tabela 24 – Estimativa de MV dos parâmetros do modelo, taxa de variação, em porcentagem, e nível descritivo associado aos coeficientes da regressão.

Observação excluída		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$	$\hat{\gamma}$
82 ^a	emv	24.6878	-0.0219	0.1861	-0.2694
	R_{θ_j}	-3.24%	+3.64%	+0.25%	-73.67%
	p -valor	0.0000	0.0000	—	—
108 ^a	emv	24.0750	-0.0213	0.1822	-0.1458
	R_{θ_j}	-5.64%	+6.29%	-1.89%	+6.00%
	p -valor	0.0000	0.0000	—	—

Fonte: Elaborada pelo autor (2022)

Como já era esperado, a exclusão da 82^a observação, considerada potencialmente influente, sob o esquema de ponderação de casos, acarreta uma taxa de variação expressiva nas estimativas de alguns dos parâmetros quando comparadas com aquelas obtidas no ajuste inicial e exibidas na Tabela 23. Particularmente, os resultados apresentados na Tabela 24 revelam que a exclusão da 82^a observação causa uma variação de -73.67% na estimativa de γ , parâmetro que regular comportamento da cauda da distribuição. Porém, tal exclusão implica em uma taxa de variação inexpressiva na estimativa de α e dos coeficientes da regressão, β_0 e β_1 , sem alterar a significância do coeficiente associado a variável explicativa. Por outro lado, a exclusão da 108^a observação, considerada influente, não causa variação expressiva em nenhum dos parâmetros do modelo de regressão, atendendo as expectativas. Ao investigarmos características associadas à 82^a observação constatamos que se trata de um dado registrado em 26 abril de 2017 onde o vento atingiu a velocidade máxima 33.9 m/s, aproximadamente 122 km/h, e a pressão atmosférica média diária foi de 1004.33 mb, classificado como sendo uma dia de baixa pressão.

Após a análise de influência local combinada com a investigação do comportamento do resíduo quantílico, conclui-se pela adequação do modelo de regressão postulado, dado pela equação (3.89), para explicar a velocidade máxima mensal do vento em termos da covariável pressão atmosférica na cidade de Itajaí-Santa Catarina-Brasil. Em outros termos, decidimos pela adequação do modelo de regressão EVBS dado por

$$Y_i = \log(T_i) = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, 124. \quad (3.90)$$

em que a estimativa de máxima verossimilhança de cada um dos parâmetros e seu respectivo erro-padrão, entre parêntesis, são $\hat{\beta}_0 = 25.5148(3.3338)$, $\hat{\beta}_1 = -0.0227(0.0033)$, $\hat{\alpha} = 0.1857(0.0127)$ e $\hat{\gamma} = -0.1551(0.0472)$, resultando em um modelo de predição para velocidade máxima mensal do vento (T) em função da pressão atmosférica média diária (X) dado por $\hat{T} = \exp(25.5148 - 0.0227x)$.

4 PERTURBAÇÕES APROPRIADAS SOBRE O MODELO DE REGRESSÃO EVBS

4.1 INTRODUÇÃO

Em (ZHU et al., 2007) destaca-se a necessidade de uma abordagem rigorosa para solucionar dois problemas relativos à metodologia de influência local: a seleção de uma forma apropriada de perturbação e o desenvolvimento de novas medidas de influência associadas a funções objetivo cujo gradiente é não-nulo. Para tanto, os autores associaram uma estrutura geométrica-diferencial ao modelo perturbado, recorrendo aos conceitos e resultados da Geometria Estatística da Informação, presentes em (AMARI, 1985) e (MURRAY; RICE, 1993). Isso resultou na noção de variedade estatística de perturbação e na identificação de que certas características associadas a esse objeto geométrico contém informações sobre como identificar uma forma apropriada de perturbação sobre o modelo postulado.

Neste capítulo mostramos que os esquemas de perturbações considerados no Capítulo 3, Seção 3.6, são formas apropriadas de perturbação sobre o modelo de regressão EVBS (caso regular), segundo os critérios estabelecidos por (ZHU et al., 2007). Nas Seções 4.2 e 4.3 discursamos acerca do conceito de variedade estatística de perturbação e das medidas de influência de primeira e segunda ordem. Na Seção 4.4 constatamos que certos esquemas de perturbação, usualmente empregados sobre o modelo de regressão normal linear, são apropriados tanto para o modelo de regressão BS, estabelecido por (RIECK; NEDELMAN, 1991), quanto para o modelo de regressão EVBS ($\gamma = 0$), definido por (LEIVA et al., 2016). Para atingir esta finalidade, mostramos que a matriz de informação esperada associada à variedade estatística de perturbação atende aos critérios fixados por (ZHU et al., 2007). Restringimos nosso estudo ao caso em que $\gamma = 0$, isso é imperativo porque o arcabouço estabelecido por (ZHU et al., 2007) supõe que as distribuições que compõem a variedade estatística de perturbação possuem suporte comum, não dependendo de parâmetros do modelo. Esse mesmo pressuposto também é admitido na estrutura geométrica-diferencial desenvolvida por (AMARI, 1985). Por fim, na Subseção 4.4.4, analisamos um conjunto de dados de temperaturas máximas anuais registrados na estação meteorológica Recife-Curado com o propósito de ilustrar a aplicabilidade da metodologia de influência local em detectar observações influentes.

4.2 VARIÉDADE ESTATÍSTICA DE PERTURBAÇÃO

Seja $p(\mathbf{Y}|\boldsymbol{\theta})$ a fdp conjunta de um vetor aleatório $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, indexada por um vetor de parâmetros desconhecidos $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)^\top \in \Theta$, onde Θ é um conjunto aberto em \mathbb{R}^q . Paralelamente, seja $\boldsymbol{\omega} = (\omega_1, \dots, \omega_p)^\top$ um vetor de perturbação pertencente a um aberto $\Omega \subset \mathbb{R}^p$.

No Capítulo 3, discursamos sobre a metodologia de influência local desenvolvida por (COOK, 1986) e vimos que a mesma baseia-se na curvatura da superfície gráfico de influência definida por

$$\phi(\boldsymbol{\omega}) = (\boldsymbol{\omega}^\top, f(\boldsymbol{\omega}))^\top, \quad (4.1)$$

onde $f : \Omega \mapsto \mathbb{R}$ é a função afastamento pela verossimilhança definida em (3.1).

Dados um ponto $\boldsymbol{\omega}_0 \in \Omega \subset \mathbb{R}^p$ e um vetor não-nulo $\mathbf{h} \in \mathbb{R}^p$, consideremos o caminho $\boldsymbol{\omega} : (-\varepsilon, \varepsilon) \mapsto \Omega$, $\boldsymbol{\omega}(t) = \boldsymbol{\omega}_0 + t\mathbf{h}$. Conseqüentemente, a curva, $\phi(\boldsymbol{\omega}(t))$, projetada sobre o gráfico de influência é tal que o vetor tangente e o vetor normal à superfície no ponto $\boldsymbol{\omega}_0$ são, respectivamente, dados pelas transformações

$$\begin{pmatrix} \mathbf{I}_p \\ \nabla_f^\top \end{pmatrix} \text{ e } (1 + \nabla_f^\top \nabla_f)^{-1/2} \begin{pmatrix} \mathbf{I}_p \\ \nabla_f^\top \end{pmatrix},$$

onde o gradiente $\nabla_f = \left(\frac{\partial f(\boldsymbol{\omega})}{\partial \omega_i} \right)$ é calculado em $\boldsymbol{\omega}_0$ e \mathbf{I}_p denota a matriz identidade de ordem p . Logo, as curvaturas normal, $C_{\mathbf{h}}$, e conforme, $B_{\mathbf{h}}$, em $\boldsymbol{\omega}_0$ e na direção do vetor não-nulo \mathbf{h} são dadas, respectivamente, por

$$C_{\mathbf{h}} = \frac{1}{(1 + \nabla_f^\top \nabla_f)^{1/2}} \frac{\mathbf{h}^\top \mathbf{H}_f \mathbf{h}}{\mathbf{h}^\top (\mathbf{I}_p + \nabla_f \nabla_f^\top) \mathbf{h}}, \quad (4.2)$$

e

$$B_{\mathbf{h}} = \frac{1}{\|\mathbf{H}_f\|} \frac{\mathbf{h}^\top \mathbf{H}_f \mathbf{h}}{\mathbf{h}^\top (\mathbf{I}_p + \nabla_f \nabla_f^\top) \mathbf{h}}, \quad (4.3)$$

onde \mathbf{H}_f denota a matriz Hessiana de f calculada em $\boldsymbol{\omega}_0$ e $\|\cdot\|$ a norma definida por $\|\mathbf{H}_f\| = \sqrt{\text{tr}(\mathbf{H}_f^2)}$.

Sabe-se que os valores máximos da curvaturas, $C_{\mathbf{h}}$ e $B_{\mathbf{h}}$, e a correspondente direção são instrumentos para avaliar o efeito da perturbação $\boldsymbol{\omega}(t) = \boldsymbol{\omega}_0 + t\mathbf{h}$ sobre o modelo estatístico postulado. Além disso, no caso em que $\nabla_f \neq \mathbf{0}$ tanto $C_{\mathbf{h}}$ quanto $B_{\mathbf{h}}$ não são invariantes por escala, no sentido de que os gráficos de influência das funções objetivo f e kf tem curvaturas

distintas, podendo-nos levar a conclusões ambíguas quando k varia, segundo destaca (ZHU et al., 2007), reportando-se a (FUNG; KWAN, 1997).

Dado um vetor de perturbação ω , denota-se por $p(\mathbf{Y}|\boldsymbol{\theta}, \omega)$ a fdp correspondente ao modelo perturbado. A fim de avaliar a influência local da perturbação ω sobre o modelo postulado, interessa-nos o comportamento de $p(\mathbf{Y}|\boldsymbol{\theta}, \omega)$, vista como uma função de ω , em torno do ponto de não-perturbação, representado por ω_0 . A partir daqui, assume-se que $p(\mathbf{Y}|\boldsymbol{\theta}, \omega)$ cumpre as quatro condições fixadas por (AMARI, 1985). Em (ZHU et al., 2007), considera-se que o modelo perturbado $\mathcal{M} = \{p(\mathbf{Y}|\boldsymbol{\theta}, \omega); \omega \in \Omega\}$ pode ser visto como uma variedade diferenciável p -dimensional.

Seja $T_{\omega}^{(1)}$ o espaço tangente à variedade \mathcal{M} em ω que é gerado pelas p funções $\partial_i \ell(\omega|\mathbf{Y}, \boldsymbol{\theta}) = \log(p(\mathbf{Y}|\boldsymbol{\theta}, \omega))$, $i = 1, \dots, p$. Para qualquer vetor $\mathbf{h} = \sum_{i=1}^p h^i \mathbf{e}_i \in \mathbb{R}^p$, tem-se a 1-representação de \mathbf{h} em $T_{\omega}^{(1)}$ dada por $\mathbf{h}(\mathbf{Y}) = \sum_{i=1}^p h^i \partial_i \ell(\omega|\mathbf{Y}, \boldsymbol{\theta})$, onde $\partial_i = \frac{\partial}{\partial \omega_i}$. Com o propósito de mensurar a contribuição que as componentes do vetor ω introduzem ao perturbar o modelo, (ZHU et al., 2007) definem o seguinte produto interno:

$$g_{ij}(\omega) = \langle \partial_i, \partial_j \rangle = \mathbb{E}_{\omega} (\partial_i \ell(\omega|\mathbf{Y}, \boldsymbol{\theta}) \cdot \partial_j \ell(\omega|\mathbf{Y}, \boldsymbol{\theta})), \quad i, j = 1, \dots, p, \quad (4.4)$$

onde \mathbb{E}_{ω} denota a esperança tomada com respeito a $p(\mathbf{Y}|\boldsymbol{\theta}, \omega)$. Dessa forma, $\mathbf{G}(\omega) = (g_{ij}(\omega))_{p \times p}$ é uma matriz de informação esperada de Fisher com respeito ao vetor de perturbação ω , de modo que os elementos de $\mathbf{G}(\omega)$ mensuram a quantidade de perturbação que as componentes do vetor ω introduzem sobre o modelo. Particularmente, $g_{ii}(\omega)$, por se só, indica a quantidade de perturbação introduzida pela i -ésima componente de ω . Já os elementos fora da diagonal representam a associação entre diferentes componentes de ω . Sugere-se o uso da grandeza $r_{ij} = \frac{g_{ij}(\omega)}{\sqrt{g_{ii}(\omega)g_{jj}(\omega)}}$ para medir o grau de associação linear entre as componentes i e j do vetor ω . Assim, se $\mathbf{G}(\omega)$ é uma matriz diagonal, então as componentes do vetor ω são ortogonais. Por outro lado, se $\mathbf{G}(\omega)$ não é positiva definida para algum esquema de perturbação, então os p operadores ∂_i não são linearmente independentes, indicando redundância em algumas componentes do vetor ω .

Segundo (ZHU et al., 2007), uma forma de perturbação apropriada para um modelo estatístico postulado deve satisfazer duas condições:

- i) A matriz $\mathbf{G}(\omega)$ deve ser positiva definida em uma vizinhança do vetor de não-perturbação ω_0 ;

- ii) Os elementos fora da diagonal principal de $\mathbf{G}(\boldsymbol{\omega})$, avaliada em $\boldsymbol{\omega}_0$, devem ser tão pequenos quanto possível.

Nestes termos, a condição (i) é colocada para evitar redundância nas componentes do vetor $\boldsymbol{\omega}$. Já a (ii) é fixada para que se possa localizar a causa potencial de uma grande instabilidade nos resultados inferenciais. Se componentes distintas de $\boldsymbol{\omega}$ tem alto grau de associação, então é difícil inferir se a instabilidade é causada por uma única componente de $\boldsymbol{\omega}$ ou por várias, exigindo-se que $G(\boldsymbol{\omega}_0)$ seja uma matriz diagonal. Porém, (ZHU et al., 2007) acentuam que sempre podemos escolher uma nova perturbação, $\tilde{\boldsymbol{\omega}}$, definida por

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega}_0 + c^{-1/2}[\mathbf{G}(\boldsymbol{\omega}_0)]^{1/2}(\boldsymbol{\omega} - \boldsymbol{\omega}_0) \quad (4.5)$$

de modo que $\mathbf{G}(\tilde{\boldsymbol{\omega}})$, avaliada em $\boldsymbol{\omega}_0$, é igual a $c\mathbf{I}_p$, com $c > 0$. Dessa forma, e sem perda de generalidade, assumem que uma perturbação apropriada $\boldsymbol{\omega}$ satisfaz $\mathbf{G}(\boldsymbol{\omega}_0) = c\mathbf{I}_p$. Entretanto, alertam que nem sempre é possível encontrar uma forma de perturbação $\boldsymbol{\omega}$ tal $\mathbf{G}(\boldsymbol{\omega}) = c\mathbf{I}_p$ para todo $\boldsymbol{\omega} \in \Omega$.

A partir deste ponto, associa-se algumas quantidades geométricas ao modelo perturbado \mathcal{M} . Primeiro, interessa-nos calcular a norma de vetores tangentes e o comprimento de curvas contidas em \mathcal{M} . Se $\mathbf{h} \in T_{\boldsymbol{\omega}}$, então

$$\|\mathbf{h}\|^2 = \langle \mathbf{h}, \mathbf{h} \rangle = \sum_{i,j} h^i h^j g_{ij}(\boldsymbol{\omega}) = \mathbf{h}^\top \mathbf{G}(\boldsymbol{\omega}) \mathbf{h}. \quad (4.6)$$

Além disso, se $\boldsymbol{\omega} : I \mapsto \mathcal{M}$, $\boldsymbol{\omega}(t) = (\omega_1(t), \dots, \omega_p(t))$ é uma curva diferenciável em \mathcal{M} , ligando os pontos $\boldsymbol{\omega}^1 = \boldsymbol{\omega}(t_1)$ e $\boldsymbol{\omega}^2 = \boldsymbol{\omega}(t_2)$, então o comprimento de $\boldsymbol{\omega}^1$ a $\boldsymbol{\omega}^2$ ao longo da curva, denotado por $S(\boldsymbol{\omega}^1, \boldsymbol{\omega}^2)$, é dado por

$$S(\boldsymbol{\omega}^1, \boldsymbol{\omega}^2) = \int_{t_1}^{t_2} \sqrt{\sum_{i,j} g_{ij}(\boldsymbol{\omega}(t)) \frac{d\omega_i(t)}{dt} \frac{d\omega_j(t)}{dt}} dt. \quad (4.7)$$

Agora, expressemos as definições de tensor simétrico, \mathbf{T} , e de uma família de conexões afins, Γ^α para algum $\alpha \in \mathbb{R}$, dadas, respectivamente, por

$$T_{ijk}(\boldsymbol{\omega}) = \mathbb{E} [\partial_i \ell(\boldsymbol{\omega} | \mathbf{Y}, \boldsymbol{\theta}) \partial_j \ell(\boldsymbol{\omega} | \mathbf{Y}, \boldsymbol{\theta}) \partial_k \ell(\boldsymbol{\omega} | \mathbf{Y}, \boldsymbol{\theta})] \quad (4.8)$$

e

$$\Gamma_{ijk}^\alpha = \mathbb{E} [\partial_i \partial_j \ell(\boldsymbol{\omega} | \mathbf{Y}, \boldsymbol{\theta}) \partial_k \ell(\boldsymbol{\omega} | \mathbf{Y}, \boldsymbol{\theta})] + 0.5(1 - \alpha) T_{ijk}(\boldsymbol{\omega}). \quad (4.9)$$

Mostra-se que

$$\Gamma_{ijk}^\alpha(\boldsymbol{\omega}) = \Gamma_{ijk}^0(\boldsymbol{\omega}) - \frac{\alpha}{2} T_{ijk}(\boldsymbol{\omega},) \quad (4.10)$$

onde $\Gamma_{ijk}^0(\boldsymbol{\omega})$ é o símbolo de Christoffel para a conexão Levi-Civita do tensor métrico e

$$\Gamma_{ijk}^0(\boldsymbol{\omega}) = \frac{1}{2}[\partial_i g_{jk}(\boldsymbol{\omega}) + \partial_j g_{ik}(\boldsymbol{\omega}) - \partial_k g_{ij}(\boldsymbol{\omega})]. \quad (4.11)$$

O arcabouço geométrico-diferencial exposto acima deram suporte para que (ZHU et al., 2007) estabelecessem as seguintes definições:

Definição. Uma variedade estatística de perturbação, $(\mathcal{M}, \mathbf{G}(\boldsymbol{\omega}), \mathbf{T}(\boldsymbol{\omega}))$, é uma variedade diferenciável \mathcal{M} munida de uma métrica $\mathbf{G}(\boldsymbol{\omega})$ e um 3-tensor covariante $\mathbf{T}(\boldsymbol{\omega})$.

Definição. Uma curva $\boldsymbol{\omega} : I \mapsto \mathcal{M}$ suave em $(\mathcal{M}, \mathbf{G}(\boldsymbol{\omega}), \mathbf{T}(\boldsymbol{\omega}))$ é dita α -geodésica com respeito a conexão afim $\Gamma_{ijk}^\alpha(\boldsymbol{\omega})$ se a mesma satisfaz a seguinte equação:

$$\frac{d^2 \omega_i(t)}{dt^2} + \sum_{s,j,k} g^{is}(\boldsymbol{\omega}(t)) \Gamma_{jks}^\alpha(\boldsymbol{\omega}(t)) \frac{d\omega_j(t)}{dt} \frac{d\omega_k(t)}{dt} = 0, \quad (4.12)$$

onde $g^{is}(\boldsymbol{\omega})$ é o elemento (i, s) da matriz $(\mathbf{G}(\boldsymbol{\omega}))^{-1}$.

4.3 MEDIDAS DE INFLUÊNCIAS E SUAS PROPRIEDADES

Sejam $f : \Omega \subset \mathbb{R}^p \mapsto \mathbb{R}$ a função objetivo usada na análise de influência, por exemplo a afastamento pela verossimilhança, e $\boldsymbol{\omega}(t)$ uma curva suave em \mathcal{M} tal que $\boldsymbol{\omega}(0) = \boldsymbol{\omega}_0$ e $\frac{d\boldsymbol{\omega}(t)}{dt}|_{t=0} = \mathbf{h} \in T_{\boldsymbol{\omega}_0}$. Logo, $f(\boldsymbol{\omega}(t))$ é uma função de $\boldsymbol{\omega}(t)$ definida na variedade de perturbação \mathcal{M} . A partir de uma expansão em série de Taylor em torno do ponto $t = 0$, obtém-se

$$f(\boldsymbol{\omega}(t)) = f(\boldsymbol{\omega}(0)) + \dot{f}_{\mathbf{h}}(0)t + \frac{1}{2}\ddot{f}_{\mathbf{h}}(0)t^2 + o(t^2), \quad (4.13)$$

onde a primeira e a segunda derivadas de $f(\boldsymbol{\omega}(t))$ em $t = 0$ são dadas, respectivamente, por:

$$\dot{f}_{\mathbf{h}}(0) = \sum_{j=1}^p \frac{\partial f(\boldsymbol{\omega}_0)}{\partial \omega_j} h_j = \nabla_f^\top \mathbf{h} \quad (4.14)$$

e

$$\ddot{f}_{\mathbf{h}}(0) = \mathbf{h}^\top \mathbf{H}_f \mathbf{h} + \nabla_f^\top \frac{d^2 \boldsymbol{\omega}(0)}{dt^2}. \quad (4.15)$$

Consequentemente, se $\nabla_f \neq \mathbf{0}$, então o termo de primeira ordem $\dot{f}_{\mathbf{h}}(0)$ carrega característica da influência do vetor de perturbação $\boldsymbol{\omega}$ sobre o modelo. Por outro lado, se $\nabla_f = \mathbf{0}$, segue que as equações (4.14) e (4.15) se reduzem a $\mathbf{0}$ e $\mathbf{h}^\top \mathbf{H}_f \mathbf{h}$, respectivamente. Logo, neste último caso, devemos usar exclusivamente o termo de segunda ordem, $\ddot{f}_{\mathbf{h}}(0)$ para avaliar o comportamento local da função objetivo.

A questão de como avaliar a influência local de pequenas perturbações sobre o modelo quando $\nabla_f \neq \mathbf{0}$ é solucionada por (ZHU et al., 2007) mediante o uso da grandeza denominada medida de influência de primeira-ordem na direção do vetor $\mathbf{h} \in T_{\omega_0}$ definida por

$$FI_{f,\mathbf{h}} = FI_{f(\omega_0),\mathbf{h}} = \frac{\mathbf{h}^\top \nabla_f \nabla_f^\top \mathbf{h}}{\mathbf{h}^\top \mathbf{G} \mathbf{h}}, \quad (4.16)$$

onde $\mathbf{G} = \mathbf{G}(\omega_0)$.

Teorema 4.1 (ZHU et al., 2007) A medida de influência de primeira ordem, FI , possui as seguintes propriedades:

(i)

$$FI_{f,\mathbf{h}} = \lim_{t \rightarrow 0} \frac{[f(\omega(t)) - f(\omega(0))]^2}{[S(\omega(0), \omega(t))]^2} \quad (4.17)$$

(ii) Se ϕ é um difeomorfismo de ω , então $FI_{f,\mathbf{h}}$ é invariante com respeito à reparametrização dada por ϕ . Além disso, a igualdade $FI_{k f, \mathbf{h}} = k^2 FI_{f, \mathbf{h}}$ é válida para todo k .

O teorema acima relaciona a medida $FI_{f,\mathbf{h}}$ com a primeira derivada da função $f(\omega)$, dando uma interpretação geométrica a medida. Além disso, mesmo que ω não seja uma perturbação apropriada, podemos usar \mathbf{G} para obtermos uma perturbação apropriada, $\tilde{\omega}$, dada por pela equação (4.5), que resulta em

$$FI_{f(\tilde{\omega}),\mathbf{h}}|_{\tilde{\omega}=\omega_0} = \frac{\mathbf{h}^\top \mathbf{G}^{-1/2} \nabla_f \nabla_f^\top \mathbf{G}^{-1/2} \mathbf{h}}{\mathbf{h}^\top \mathbf{h}}. \quad (4.18)$$

A abordagem em que se considera as informações contidas em $\ddot{f}_{\mathbf{h}}(0)$ para se conduzir a análise de influência local, mesmo quando $\nabla_f \neq \mathbf{0}$, é denominada abordagem de segunda ordem. A fim de que $\ddot{f}_{\mathbf{h}}(0)$ possua as propriedades geométricas convenientes para o desenvolvimento da metodologia, supõe-se que $\omega(t)$ é a 0-geodésica associada à conexão Levi-Civita do tensor métrico, isso acarreta que $\omega(t)$ está bem definida em um intervalo real contendo a origem, $\omega(t) = \omega_0$ e $\frac{d\omega(t)}{dt} = \mathbf{h} \in T_{\omega_0}$. Consequentemente, usando expansão de Taylor, segue que

$$f(\omega(t)) = f(\omega_0) + t \nabla_f^\top \mathbf{h} + \frac{1}{2} t^2 \mathbf{h}^\top \tilde{\mathbf{H}}_f^0 \mathbf{h} + o(t^2), \quad (4.19)$$

onde $\tilde{\mathbf{H}}_f^0 = \tilde{\mathbf{H}}_{f(\omega_0)}^0$ e o termo (i, j) de $\tilde{\mathbf{H}}_{f(\omega)}^0$ é dado por

$$\left(\tilde{\mathbf{H}}_{f(\omega)}^0\right)_{(i,j)} = \partial_i \partial_j f(\omega) - \sum_{s,r} g^{sr}(\omega) \Gamma_{ijs}^0(\omega) \partial_r f(\omega). \quad (4.20)$$

A matriz $\tilde{\mathbf{H}}_f^0(\boldsymbol{\omega})$ é denominada Hessiana covariante de $f(\boldsymbol{\omega})$. Dado que $\Gamma_{ijk}^0(\boldsymbol{\omega})$ é simétrica com respeito aos índices i e j , segue que $\tilde{\mathbf{H}}_f^0(\boldsymbol{\omega})$ é simétrica. Além disso, se ϕ é um difeomorfismo de $\boldsymbol{\omega}$ com matriz Jacobiana $\Phi = \partial\boldsymbol{\omega}/\partial\phi$, então $\tilde{\mathbf{H}}_{f(\phi)}^0 = \Phi^\top \tilde{\mathbf{H}}_f^0(\boldsymbol{\omega}) \Phi$.

Definição. A medida de influência de segunda ordem, SI , na direção do vetor $\mathbf{h} \in T\boldsymbol{\omega}_0$ é definida como

$$SI_{f,\mathbf{h}} = SI_{f(\boldsymbol{\omega}_0),\mathbf{h}} = \frac{\mathbf{h}^\top \tilde{\mathbf{H}}_f^0 \mathbf{h}}{\mathbf{h}^\top \mathbf{G} \mathbf{h}}. \quad (4.21)$$

De modo que a medida de influência de segunda ordem padronizada, SSI , na direção do vetor $\mathbf{h} \in T\boldsymbol{\omega}_0$ é dada por

$$SSI_{f,\mathbf{h}} = SI_{f(\boldsymbol{\omega}_0),\mathbf{h}} = \frac{1}{\|\mathbf{G}^{-1} \tilde{\mathbf{H}}_f^0\|} \frac{\mathbf{h}^\top \tilde{\mathbf{H}}_f^0 \mathbf{h}}{\mathbf{h}^\top \mathbf{G} \mathbf{h}}. \quad (4.22)$$

Teorema 4.2 (ZHU et al., 2007). As medidas de influência de segunda ordem SI e SSI detêm das seguintes propriedades:

i)

$$SI_{f,\mathbf{h}} = \lim_{t \rightarrow 0} \frac{2[f(\boldsymbol{\omega}(t)) - f(\boldsymbol{\omega}(0)) - t \nabla_f^\top \mathbf{h}]}{[S(\boldsymbol{\omega}(0), \boldsymbol{\omega}(t))]^2}. \quad (4.23)$$

ii) Se ϕ é um difeomorfismo de $\boldsymbol{\omega}$, então $SI_{f(\boldsymbol{\omega}_0),\mathbf{h}}$ e $SSI_{f(\boldsymbol{\omega}_0),\mathbf{h}}$ são invariantes com respeito à reparametrização dada por ϕ em $\boldsymbol{\omega}_0$. Além disso,

$$SI_{kf(\boldsymbol{\omega}),\mathbf{h}} = k SI_{f(\boldsymbol{\omega}),\mathbf{h}} \quad \text{e} \quad SSI_{kf(\boldsymbol{\omega}),\mathbf{h}} = SSI_{f(\boldsymbol{\omega}),\mathbf{h}}$$

para qualquer $k \neq 0$ e $\boldsymbol{\omega} \in \Omega$.

iii) Se $\{(\lambda_i, \mathbf{u}_i), i = 1, \dots, p\}$ é o conjunto de pares formados pelos autovalores e os correspondentes autovetores de $\tilde{\mathbf{H}}_f^0$, então para qualquer direção \mathbf{h} , tem-se

$$SI_{f,\mathbf{u}_i} = \lambda_i, \quad SSI_{f,\mathbf{u}_i} = \frac{\lambda_i}{\sqrt{\sum_{j=1}^p \lambda_j^2}} = \hat{\lambda}_i \quad \text{e} \quad 0 \leq SSI_{f,\mathbf{h}} \leq 1.$$

O teorema acima tem duas implicações imediatas. Primeiro, se $\boldsymbol{\omega}$ é uma perturbação apropriada e $\nabla_f = \mathbf{0}$, então $SI_{f,\mathbf{h}} = C_{\mathbf{h}}$ e $SSI_{f,\mathbf{h}} = B_{\mathbf{h}}$. Segundo, mesmo quando $\nabla_f \neq \mathbf{0}$, as medidas $SI_{f(\boldsymbol{\omega}),\mathbf{h}}$ e $SSI_{f(\boldsymbol{\omega}),\mathbf{h}}$ são invariantes por difeomorfismos sobre $\boldsymbol{\omega}$.

Assim, a condução de uma análise de influência local, incorporando os aperfeiçoamentos obtidos por (ZHU et al., 2007), deve ser implementada mediante o desenvolvimento dos seguintes passos:

1ª) Escolha um esquema de perturbação tal que $\int p(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\omega}) d\mathbf{Y} = 1$;

- 2º) Dado o modelo perturbado , $p(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{\omega})$, calcule as quantidades geométricas $g_{ij}(\boldsymbol{\omega})$, $T_{ijk}(\boldsymbol{\omega})$ e $\Gamma_{ijk}^\alpha(\boldsymbol{\omega})$ correspondentes à variedade estatística de perturbação;
- 3º) Verifique se o esquema de perturbação é apropriado, ou seja, se $\mathbf{G}(\boldsymbol{\omega}_0) = c\mathbf{I}_p$. Em caso afirmativo, prossegue-se com o passo seguinte. Caso contrário, encontre um novo esquema de perturbação e volte ao 2º passo;
- 4º) Escolha uma função objetivo $f(\boldsymbol{\omega})$. Se $\nabla_f = \mathbf{0}$, então use SI e SSI para avaliar a influência local de pequenas perturbações sobre o modelo. Porém, se $\nabla_f \neq \mathbf{0}$, então use, conjuntamente, as medidas FI , SI e SSI .

4.4 FORMAS APROPRIADAS DE PERTURBAÇÃO.

Ao longo desta seção, mostraremos que certos esquemas usuais de perturbação empregados sobre o modelo de regressão normal linear se revelam apropriados tanto para o modelo de regressão BS quanto para o modelo de regressão EVBS ($\gamma = 0$). Para atingir tal objetivo, justificaremos que a matriz de informação esperada associada à variedade estatística de perturbação atende aos critérios estabelecidos por (ZHU et al., 2007).

4.4.1 Modelo de regressão normal linear.

Exemplo 1. Ponderação de casos. Consideremos o modelo de regressão normal linear no qual cada componente do vetor de respostas $\mathbf{Y} = (y_1, y_2, \dots, y_n)^\top$ é tal que

$$y_i = \sum_{j=1}^p x_{ij}\beta_j + \varepsilon_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \dots, n, \quad (4.24)$$

onde as componentes ε_i do vetor de erros $\boldsymbol{\varepsilon}$ são variáveis aleatórias independentes e identicamente distribuídas a variável aleatória normal com média zero e variância σ^2 . Em notação matricial,

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}; \quad \boldsymbol{\varepsilon} \sim \mathbf{N}_p(\mathbf{0}, \sigma^2\mathbf{I}_p) \quad (4.25)$$

onde $\mathbf{x}_i^\top = (x_{i1}, \dots, x_{ip})$, $\mathbf{X} = (x_{ij})_{n \times p}$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ e \mathbf{I}_p denota a matriz identidade de ordem p . No caso em que o parâmetro σ^2 é conhecido, dada uma amostra de observações $\mathbf{y} = (y_1, \dots, y_n)^\top$ e um vetor de perturbação $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$, a função densidade conjunta do modelo perturbado é dada por

$$p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\omega}) = \prod_{i=1}^n \exp \left\{ \omega_i \left(-\frac{1}{2\sigma^2} \right) (y_i - \mathbf{x}_i^\top \boldsymbol{\beta})^2 \right\} \cdot [c_i(\omega_i; \boldsymbol{\beta})]^{-1}, \quad (4.26)$$

onde $c_i(\omega_i; \boldsymbol{\beta}) = \int \exp \left\{ \omega_i \left(-\frac{1}{2\sigma^2} \right) (y_i - \mathbf{x}_i^\top \boldsymbol{\beta})^2 \right\} dy_i$ para todo $i = 1, \dots, n$. Daí, segue que a log-verossimilhança perturbada é

$$\ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\beta}) = \log p(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \left(-\frac{1}{2\sigma^2} (y_i - \mathbf{x}_i^\top \boldsymbol{\beta})^2 \right) + \sum_{i=1}^n \log [c_i(\omega_i; \boldsymbol{\beta})]^{-1}, \quad (4.27)$$

levando-nos a obter a seguinte derivada parcial:

$$\frac{\partial \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\beta})}{\partial \omega_k} = -\frac{1}{2\sigma^2} (y_k - \mathbf{x}_k^\top \boldsymbol{\beta})^2 - \partial_k \log [c_k(\omega_k; \boldsymbol{\beta})], \quad k = 1, 2, \dots, n. \quad (4.28)$$

Isso acarreta que a matriz $\mathbf{G}(\boldsymbol{\omega})$, cujos termos são dados pela equação (4.4), terá entradas (i, j) dadas por

$$g_{ij}(\boldsymbol{\omega}) = \mathbb{E} \left(\left[-\frac{1}{2\sigma^2} (y_i - \mathbf{x}_i^\top \boldsymbol{\beta})^2 - \partial_i \log c_i(\omega_i; \boldsymbol{\beta}) \right] \left[-\frac{1}{2\sigma^2} (y_j - \mathbf{x}_j^\top \boldsymbol{\beta})^2 - \partial_j \log c_j(\omega_j; \boldsymbol{\beta}) \right] \right). \quad (4.29)$$

Neste ponto, ao usarmos as propriedades da função $c_i(\omega_i, \boldsymbol{\beta})$, contidas no Teorema 3 que figura em (ZHU et al., 2007), e as característica das componentes do vetor de erros $\boldsymbol{\varepsilon}$ segue que se $i \neq j$, então $g_{ij}(\boldsymbol{\omega}) = 0$; por outro lado se $i = j$, então $g_{ii}(\boldsymbol{\omega}) = \frac{1}{4\sigma^4} \text{Var}[(y_i - \mathbf{x}_i^\top \boldsymbol{\beta})^2] = \frac{1}{4\sigma^4} 2\sigma^4 = \frac{1}{2}$. Logo, $\mathbf{G}(\boldsymbol{\omega}) = \frac{1}{2} \mathbf{I}_n$ e, portanto, tal forma de perturbação revela-se apropriada, segundo os critérios estabelecidos por (ZHU et al., 2007). Além disso, uma vez que a função afastamento pela verossimilhança possui um mínimo local em $\boldsymbol{\omega}_0 = (1, 1, \dots, 1)^\top$, isto é $\nabla_f|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = 0$, segue do Teorema 4.2 que a análise de influência local usando a curvatura conforme, exemplificada em (POON; POON, 1999), e aquela por meio da medida de influência de segunda ordem padronizada, SSI , são numericamente equivalentes, levando-nos as mesmas conclusões no que tange à identificação de observações influentes.

Exemplo 2. Perturbação em uma variável explicativa. Suponhamos que no modelo definido pela equação (4.24) figure uma variável contínua e a denotemos por \mathbf{x}_t . Impomos sobre \mathbf{x}_t uma perturbação aditiva da forma $x_{it\omega} = x_{it} + \omega_i s_x$, onde $t \in \{1, \dots, p\}$ e s_x é um fator de escala associado à variável \mathbf{x}_t . Dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ e um vetor de perturbação $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$, a função de log-verossimilhança perturbada é

$$\ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\beta}) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \sum_{i=1}^n \frac{1}{2\sigma^2} (y_i - \mathbf{x}_i^\top \boldsymbol{\beta} - \beta_i \omega_i s_x)^2. \quad (4.30)$$

Conseqüentemente,

$$\frac{\partial \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\beta})}{\partial \omega_k} = -\frac{1}{\sigma^2} (\beta_t s_x) (y_k - \mathbf{x}_k^\top \boldsymbol{\beta} - \beta_t \omega_k s_x), \quad k = 1, \dots, n. \quad (4.31)$$

Logo,

$$g_{ij}(\boldsymbol{\omega}) = \mathbb{E} \left(\frac{1}{\sigma^4} \beta_t^2 s_x^2 (\varepsilon_i - \beta_t \omega_i s_x) (\varepsilon_j - \beta_t \omega_j s_x) \right), \quad \text{com } i, j = 1, \dots, n, \quad (4.32)$$

onde ε_i denota a i -ésima componente do vetor de erros ε presente em (4.24). Ou seja,

$$g_{ii}(\boldsymbol{\omega}) = \frac{1}{\sigma^4} \beta_t^2 s_x^2 (\sigma^2 + \beta_t^2 \omega_i^2 s_x^2) \quad \text{e} \quad g_{ij}(\boldsymbol{\omega}) = \frac{1}{\sigma^4} (\beta_t^4 s_x^4 \omega_i \omega_j), \quad \text{se } i \neq j. \quad (4.33)$$

Observemos que quando perturbamos uma única variável explicativa, o vetor de não-perturbação é $\boldsymbol{\omega}_0 = \mathbf{0}$. Logo, $g_{ii}(\boldsymbol{\omega}_0) = \frac{1}{\sigma^2} \beta_t^2 s_x^2$ e $g_{ij}(\boldsymbol{\omega}_0) = 0$ quando $i \neq j$, assegurando-nos que $\mathbf{G}(\boldsymbol{\omega}_0) = c\mathbf{I}_p$, com $c > 0$. Portanto, o esquema em que se perturba uma única covariável no modelo de regressão normal linear é apropriado.

4.4.2 Modelo de regressão log-Birnbaum-Saunders

Exemplo 1. Ponderação de casos. Consideremos o modelo de regressão log-BS, definido na Seção 2.26, no qual as componentes do vetor de respostas $\mathbf{Y} = (y_1, y_2, \dots, y_n)^\top$ são tais que

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \nu_i = \eta_i + \nu_i, \quad i = 1, \dots, n, \quad (4.34)$$

onde os erros ν_i 's são variáveis aleatórias independentes e identicamente distribuídas tais que $\nu_i \sim \text{SN}(\alpha, 0, 2)$ para todo $i = 1, 2, \dots, n$. Nesse caso, a função densidade de probabilidade y_i é dada por

$$p(y_i; \alpha, \boldsymbol{\beta}) = \frac{1}{\alpha \sqrt{2\pi}} \cosh\left(\frac{y_i - \eta_i}{2}\right) \exp\left[-\frac{2}{\alpha^2} \sinh^2\left(\frac{y_i - \eta_i}{2}\right)\right], \quad (4.35)$$

com $y_i \in \mathbb{R}$ e $\eta_i = \mathbf{x}_i^\top \boldsymbol{\beta}$. Consequentemente,

$$\ell(y_i; \alpha, \boldsymbol{\beta}) = \log p(y_i; \alpha, \boldsymbol{\beta}) = -\frac{1}{2} \log(8\pi) + \log(\xi_{i1}) - \frac{1}{2} \xi_{i2}^2, \quad (4.36)$$

onde $\xi_{i1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \eta_i}{2}\right)$ e $\xi_{i2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \eta_i}{2}\right)$.

Sob o esquema de ponderação de casos, a fdp conjunta do modelo perturbado é dada por

$$p(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{i=1}^n \exp\{\omega_i \ell(y_i; \boldsymbol{\theta})\} \cdot [c_i(\omega_i; \boldsymbol{\theta})]^{-1}, \quad (4.37)$$

onde $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ e $c_i(\omega_i; \boldsymbol{\theta}) = \int \exp\{\omega_i \ell(y_i; \boldsymbol{\theta})\} dy_i$ para todo $i = 1, \dots, n$. Daí, segue que a log-verossimilhança perturbada é

$$\ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta}) = \sum_{i=1}^n \omega_i \left\{ -\frac{1}{2} \log(8\pi) + \log(\xi_{i1}) - \frac{1}{2} \xi_{i2}^2 \right\} - \sum_{i=1}^n \log[c_i(\omega_i; \boldsymbol{\theta})], \quad (4.38)$$

Logo,

$$\frac{\partial \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta})}{\partial \omega_k} = -\frac{1}{2} \log(8\pi) + \log(\xi_{k1}) - \frac{1}{2} \xi_{k2}^2 - \partial_k \log c_k(\omega_k, \boldsymbol{\theta}), \quad k = 1, 2, \dots, n. \quad (4.39)$$

Neste ponto observemos que, sob o esquema de ponderação de casos,

$$\ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta}) = \sum_{i=1}^n \{\omega_i \ell(y_i; \boldsymbol{\theta}) - \log c_i(\omega_i; \boldsymbol{\theta})\}. \quad (4.40)$$

Logo,

$$\mathbb{E} \left(\frac{\partial}{\partial \omega_k} \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta}) \right) = \mathbb{E}[\ell(y_k; \boldsymbol{\theta})] - \partial_k \log c_k(\omega_k; \boldsymbol{\theta}). \quad (4.41)$$

Se o modelo associado à variável y_k é tal que as quatro suposições presentes em (AMARI, 1985) estão satisfeitas, implicando em $\mathbb{E}(\partial_k \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta})) = 0$, então

$$\mathbb{E}(\ell(y_k; \boldsymbol{\theta})) = \partial_k \log c_k(\omega_k; \boldsymbol{\theta}), \text{ para todo } k = 1, 2, \dots, n. \quad (4.42)$$

Daí, se $i \neq j$, então a hipótese de independência entre as variáveis aleatórias y_i e y_j acarreta

$$g_{ij}(\boldsymbol{\omega}) = \mathbb{E}(\partial_i \ell(y_i; \boldsymbol{\theta})) \cdot \mathbb{E}(\partial_j \ell(y_j; \boldsymbol{\theta})) = 0. \quad (4.43)$$

Por outro lado, se $i = j$, então

$$g_{ii}(\boldsymbol{\omega}) = \mathbb{E} \left[(\ell(y_i; \boldsymbol{\theta}) - \partial \log c_i(\omega_i; \boldsymbol{\theta}))^2 \right] = \mathbb{E}[(\ell(y_i; \boldsymbol{\theta}) - \mathbb{E}(\ell(y_i; \boldsymbol{\theta})))^2] = \text{Var}(\ell(y_i; \boldsymbol{\theta})) \quad (4.44)$$

Logo, usando a equação (4.36), obtemos

$$g_{ii}(\boldsymbol{\omega}) = \text{Var} \left(-\frac{1}{2} \log(8\pi) + \log(\xi_{i1}) - \frac{1}{2} \xi_{i2}^2 \right) = \text{Var} \left(\log(\xi_{i1}) - \frac{1}{2} \xi_{i2}^2 \right). \quad (4.45)$$

Por meio da identidade $\xi_{i1}^2 = \xi_{i2}^2 + (4/\alpha^2)$, temos

$$g_{ii}(\boldsymbol{\omega}) = \text{Var} \left(\frac{1}{2} \log(\xi_{i1}^2) - \frac{1}{2} \xi_{i2}^2 \right) = \frac{1}{4} \text{Var} \left(\log \left(\frac{4}{\alpha^2} + \xi_{i2}^2 \right) - \xi_{i2}^2 \right). \quad (4.46)$$

Neste ponto, destaquemos que a variável aleatória $Z_i = \xi_{i2} = \frac{2}{\alpha} \sinh \left(\frac{y_i - \mathbf{x}_i \boldsymbol{\beta}}{2} \right)$ é tal que $Z_i = \xi_{i2} \sim \text{N}(0, 1)$ para todo $i = 1, \dots, n$. De fato, dado que $y_i \sim \text{SN}(\alpha, \eta_i, 2)$, as funções de distribuição das variáveis Z_i e y_i são tais que

$$F_{Z_i}(z) = F_{y_i} \left(\eta_i + 2 \operatorname{arcsinh} \left(\frac{\alpha z}{2} \right) \right). \quad (4.47)$$

Logo, usando da equação (4.35), constatamos que a fdp da variável Z_i é

$$f_{Z_i}(z) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} z^2 \right), \quad z \in \mathbb{R}. \quad (4.48)$$

Consequentemente, a variância dada pela equação (4.46) é a mesma para todo $i = 1, \dots, n$. Isso implica que ao impormos sobre o modelo de regressão log-BS clássico uma

perturbação na forma de ponderação de casos, concluímos que $g_{ii}(\boldsymbol{\omega}_0)$ assume o mesmo valor para todo $i = 1, \dots, n$, sendo função apenas de α , e quando $i \neq j$ tem-se $g_{ij}(\boldsymbol{\omega}_0) = 0$. Portanto, $G(\boldsymbol{\omega}_0) = c\mathbf{I}_p$ com $c > 0$, assegurando-nos que a ponderação de casos é uma forma apropriada de perturbação sobre o modelo de regressão log-BS.

Exemplo 2. Perturbação na variável resposta. Mais uma vez, consideremos o modelo de regressão log-BS, definido na Seção 2.26 e sujeitemos a resposta y_i a uma perturbação da forma $y_{i\omega} = y_i + \omega_i s_y$, em que $\omega_i \in \mathbb{R}$ para todo $i = 1, 2, \dots, n$ e s_y é um fator de escala. Dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ e um vetor de perturbação $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$, a função de log-verossimilhança perturbada é

$$\ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta}) = -\frac{n}{2} \log(8\pi) + \sum_{i=1}^n \log(\xi_{i1\omega_1}) - \frac{1}{2} \sum_{i=1}^n \xi_{i2\omega_2}^2, \quad (4.49)$$

onde $\xi_{i1\omega_1} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta} + \omega_i s_y}{2}\right)$ e $\xi_{i2\omega_1} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \mathbf{x}_i^\top \boldsymbol{\beta} + \omega_i s_y}{2}\right)$. Consequentemente,

$$\frac{\partial \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\beta})}{\partial \omega_k} = \frac{s_y}{2} \begin{pmatrix} \xi_{k2\omega_1} \\ \xi_{k1\omega_1} \end{pmatrix} - \xi_{k2\omega_1} \xi_{k1\omega_1}. \quad (4.50)$$

No caso em que $i = j$, as entradas na \mathbf{G} são dadas por

$$g_{ii}(\boldsymbol{\omega}) = \frac{1}{4} s_y^2 \mathbb{E} \left(\left[\frac{\xi_{i2\omega_1}}{\xi_{i1\omega_1}} - \xi_{i2\omega_1} \xi_{i1\omega_1} \right]^2 \right). \quad (4.51)$$

Dado que $\cosh^2(z) = 1 + \sinh^2(z)$, segue a relação $\xi_{i1\omega_1}^2 = \frac{4}{\alpha^2} + \xi_{i2\omega_1}^2$. Daí, obtemos

$$g_{ii}(\boldsymbol{\omega}) = \frac{1}{4} s_y^2 \mathbb{E} \left(\frac{\xi_{i2\omega_1}^2}{\frac{4}{\alpha^2} + \xi_{i2\omega_1}^2} - 2\xi_{i2\omega_1}^2 + \frac{4}{\alpha^2} \xi_{i2\omega_1}^2 + \xi_{i2\omega_1}^4 \right). \quad (4.52)$$

Usando a fórmula de adição $\sinh(z_1 + z_2) = \sinh(z_1) \cosh(z_2) + \cosh(z_1) \sinh(z_2)$, segue que

$$\begin{aligned} \xi_{i2\omega_1}^2 &= \frac{4}{\alpha^2} \left[\sinh^2\left(\frac{y_i - \eta_i}{2}\right) \cosh^2\left(\frac{\omega_i s_y}{2}\right) + \cosh^2\left(\frac{y_i - \eta_i}{2}\right) \sinh^2\left(\frac{\omega_i s_y}{2}\right) + \right. \\ &\quad \left. + 2\sinh\left(\frac{y_i - \eta_i}{2}\right) \cosh\left(\frac{y_i - \eta_i}{2}\right) \cosh\left(\frac{\omega_i s_y}{2}\right) \sinh\left(\frac{\omega_i s_y}{2}\right) \right]. \end{aligned}$$

Sob o esquema de perturbação na resposta tem-se $\boldsymbol{\omega}_0 = (0, 0, \dots, 0)^\top$. Logo,

$$\mathbb{E}(\xi_{i2\omega_1}^2) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{4}{\alpha^2} \mathbb{E} \left(\sinh^2\left(\frac{y_i - \eta_i}{2}\right) \right) = \mathbb{E}(\xi_{i2}^2) = \mathbb{E}(Z^2) = 1, \quad (4.53)$$

pois, $\xi_{i2} = (2/\alpha) \sinh((y_i - \eta_i)/2) \sim N(0, 1)$. De modo análogo, usando o Binômio de Newton e as relações entre as funções hiperbólicas, constatamos que

$$\mathbb{E}(\xi_{i2\omega_1}^4) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \frac{16}{\alpha^4} \mathbb{E} \left(\sinh^4\left(\frac{y_i - \eta_i}{2}\right) \right) = \mathbb{E}(\xi_{i2}^4) = \mathbb{E}(Z^4) = 3. \quad (4.54)$$

Portanto,

$$g_{ii}(\boldsymbol{\omega}_0) = \frac{1}{4} s_y^2 \left[\mathbb{E} \left(\frac{\xi_{i2}^2}{(4/\alpha^2) + \xi_{i2}^2} \right) + 1 + \frac{4}{\alpha^2} \right] > 0. \quad (4.55)$$

Observemos que os momentos centrais $\mathbb{E}(\xi_{i2}^2 | \boldsymbol{\omega} = \boldsymbol{\omega}_0)$ não dependem dos valores associados à i -ésima observação. Logo, $g_{ii}(\boldsymbol{\omega}_0)$ assume o mesmo valor para todo $i = 1, 2, \dots, n$. Por outro lado, a hipótese de independência entre as variáveis y_i e y_j , para $i \neq j$, resulta que $g_{ij}(\boldsymbol{\omega}) = 0$. Portanto, $G(\boldsymbol{\omega}_0) = cI_p$ com $c > 0$, ou seja, a perturbação na variável resposta, conforme descrita acima, é adequada para o modelo de regressão BS.

4.4.3 Modelo de regressão EVBS

Exemplo 1. Ponderação de casos. Consideremos o modelo de regressão EVBS, definido na Seção 2.7, no qual as componentes do vetor de respostas $\mathbf{Y} = (y_1, y_2, \dots, y_n)^\top$ são tais que

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \varepsilon_i = \eta_i + \varepsilon_i, \quad i = 1, \dots, n, \quad (4.56)$$

onde os erros $\varepsilon_{i's}$ são variáveis aleatórias independentes e identicamente distribuídas tais que $\varepsilon_i \sim \text{log-EVBS}(\alpha, 0, \gamma)$ para todo $i = 1, 2, \dots, n$. No caso particular em que $\gamma = 0$, tem-se que $y_i \sim \text{log-EVBS}(\alpha, \eta_i)$ para cada $i = 1, \dots, n$. Logo, a função densidade de probabilidade y_i é dada por

$$p(y_i; \alpha, \eta_i) = \frac{1}{\alpha} \cosh \left(\frac{y_i - \eta_i}{2} \right) \exp \left[-\frac{2}{\alpha} \sinh \left(\frac{y_i - \eta_i}{2} \right) - \exp \left(-\frac{2}{\alpha} \sinh \left(\frac{y_i - \eta_i}{2} \right) \right) \right], \quad (4.57)$$

onde $y_i \in \mathbb{R}$ e $\eta_i = \mathbf{x}_i^\top \boldsymbol{\beta}$.

Sob o esquema de ponderação, a função de densidade conjunta do modelo perturbado é dada por

$$p(\mathbf{y} | \boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{i=1}^n \exp \{ \omega_i \ell(y_i; \boldsymbol{\theta}) \} \cdot [c_i(\omega_i; \boldsymbol{\theta})]^{-1}, \quad (4.58)$$

onde $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$, $\ell(y_i, \boldsymbol{\theta}) = \log p(y_i; \boldsymbol{\theta})$ e $c_i(\omega_i; \boldsymbol{\theta}) = \int \exp \{ \omega_i \ell(y_i; \boldsymbol{\theta}) \} dy_i$ para todo $i = 1, \dots, n$. Daí, segue que a log-verossimilhança perturbada é

$$\ell(\boldsymbol{\omega} | \mathbf{y}, \boldsymbol{\theta}) = \sum_{i=1}^n \omega_i \{ \log(1/2) + \log(\xi_{i1}) - \xi_{i2} - \exp(-\xi_{i2}) \} - \sum_{i=1}^n \log [c_i(\omega_i; \boldsymbol{\theta})], \quad (4.59)$$

onde $\xi_{i1} = \frac{2}{\alpha} \cosh \left(\frac{y_i - \eta_i}{2} \right)$ e $\xi_{i2} = \frac{2}{\alpha} \sinh \left(\frac{y_i - \eta_i}{2} \right)$ para $i = 1, 2, \dots, n$. Consequentemente,

$$\frac{\partial \ell(\boldsymbol{\omega} | \mathbf{y}, \boldsymbol{\theta})}{\partial \omega_k} = \log(1/2) + \log(\xi_{k1}) - \xi_{k2} - \exp(-\xi_{k2}) - \partial_k \log c_k(\omega_k, \boldsymbol{\theta}), \quad k = 1, 2, \dots, n. \quad (4.60)$$

Uma vez que estamos lidando com ponderação de casos sobre um modelo onde se verifica as suposições reportadas por (ZHU et al., 2007) no Teorema 3, tem-se que $g_{ii}(\boldsymbol{\omega}) = \text{Var}(\ell(y_i; \boldsymbol{\theta}))$. Logo,

$$g_{ii}(\boldsymbol{\omega}) = \text{Var}[\log(\xi_{i1}) - \xi_{i2} - \exp(-\xi_{i2})] = \text{Var} \left[\frac{1}{2} \log \left(\frac{4}{\alpha^2} + \xi_{i2}^2 \right) - \xi_{i2} - \exp(-\xi_{i2}) \right]. \quad (4.61)$$

Observemos que a variável aleatória $\xi_{i2} = \frac{2}{\alpha} \sinh \left(\frac{y_i - \eta_i}{2} \right)$ possui função densidade de probabilidade dada por $f_{\xi_{i2}}(\xi) = \exp(-\xi - \exp(-\xi))$, $\xi \in \mathbb{R}$. Ou seja $\xi_{i2} \sim \text{Gumbel}(\mu = 0, \sigma = 1)$ para todo $i = 1, \dots, n$. Logo, a função geradora de momentos da variável $\xi_{i2} \sim \text{Gumbel}(\mu = 0, \sigma = 1)$ é $M_{\xi_{i2}}(t) = \Gamma(1 - t)$, $t < 1$, isso acarreta $\mathbb{E}(\xi_{i2}^k) = (-1)^k \Gamma^{(k)}(1)$, $\forall i = 1, 2, \dots, n$ e k um inteiro positivo.

Agora, por meio e uma argumentação similar à apresentada quando lidamos com ponderação de casos no modelo de regressão log-BS, constata-se que se $i \neq j$, então a hipótese de independência entre as variáveis aleatórias y_i e y_j acarreta $g_{ij}(\boldsymbol{\omega}_0) = 0$. Além disso, a variância presente na equação (4.61) é a mesma para todo $i = 1, \dots, n$, ou seja, $g_{ii}(\boldsymbol{\omega}_0)$ assume o mesmo valor real para todo $i = 1, \dots, n$. Isso implica que $G(\boldsymbol{\omega}_0) = c\mathbf{I}_p$ com $c > 0$, assegurando-nos que a ponderação de casos é uma forma apropriada de perturbação sobre o modelo de regressão EVBS($\gamma = 0$).

Exemplo 2. Perturbação na variável resposta. Consideremos o modelo de regressão linear definido em (4.56) e sujeitemos a resposta y_i a uma perturbação da forma $y_{i\omega} = y_i + \omega_i s_y$, em que $\omega_i \in \mathbb{R}$ para todo $i = 1, 2, \dots, n$ e s_y é um fator de escala. Neste caso, dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ e um vetor de perturbação $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$, a função de densidade conjunta do modelo perturbado é dada por

$$p(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{i=1}^n \left\{ \frac{1}{2} \xi_{i1\omega_1} \exp(-\xi_{i2\omega_1} - \exp(-\xi_{i2\omega_1})) \right\}, \quad (4.62)$$

onde $\xi_{i1\omega_1} = \frac{2}{\alpha} \cosh \left(\frac{y_i - \eta_i - \omega_i s_y}{2} \right)$ e $\xi_{i2\omega_1} = \frac{2}{\alpha} \sinh \left(\frac{y_i - \eta_i - \omega_i s_y}{2} \right)$. Consequentemente, a função de log-verossimilhança perturbada é

$$\ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta}) = -n \log(2) + \sum_{i=1}^n \log(\xi_{i1\omega_1}) - \sum_{i=1}^n \xi_{i2\omega_1} - \sum_{i=1}^n \exp(-\xi_{i2\omega_1}). \quad (4.63)$$

Logo,

$$\frac{\partial \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta})}{\partial \omega_k} = -\frac{s_y}{2} \left(\frac{\xi_{i2\omega_1}}{\xi_{i1\omega_1}} - \xi_{i1\omega_1} (1 - \exp(-\xi_{i2\omega_1})) \right). \quad (4.64)$$

Para o caso em que $i = j$, teremos

$$g_{ii}(\boldsymbol{\omega}) = \frac{s_y^2}{4} \mathbb{E} \left[\left(\frac{\xi_{i2\omega_1}}{\xi_{i1\omega_1}} - \xi_{i1\omega_1} (1 - \exp(-\xi_{i2\omega_1})) \right)^2 \right]. \quad (4.65)$$

Ou seja,

$$g_{ii}(\boldsymbol{\omega}) = \frac{s_y^2}{4} \mathbb{E} \left[\frac{\xi_{i2\omega_1}^2}{(4/\alpha^2) + \xi_{i2\omega_1}^2} + 2\xi_{i2\omega_1}(1 - \exp(-\xi_{i2\omega_1})) + \left(\frac{4}{\alpha^2} + \xi_{i2\omega_1}^2 \right) (1 - \exp(-\xi_{i2\omega_1}))^2 \right]. \quad (4.66)$$

Usando o Binômio de Newton, temos

$$\xi_{i2\omega_1}^n = \left(\frac{2}{\alpha} \right)^n \sum_{k=0}^n \binom{n}{k} \cosh^k \left(-\frac{w_i s_y}{2} \right) \sinh^k \left(\frac{y_i - \eta_i}{2} \right) \sinh^{n-k} \left(-\frac{w_i s_y}{2} \right) \cosh^{n-k} \left(\frac{y_i - \eta_i}{2} \right). \quad (4.67)$$

Sob o esquema de perturbação na resposta o vetor de não-perturbação é $\boldsymbol{\omega}_0 = (0, 0, \dots, 0)^\top$.

Logo, usando propriedades das funções hiperbólicas e do operador esperança

$$\mathbb{E}(\xi_{i2\omega_1}^k) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \mathbb{E}(\xi_{i2}^k) \quad \text{para todo } k = 1, \dots, n. \quad (4.68)$$

Dado que a função geradora de momentos da variável $\xi_{i2} \sim \text{Gumbel}(\mu = 0, \sigma = 1)$ é $M_{\xi_{i2}}(t) = \Gamma(1 - t)$, $t < 1$, obtemos

$$\mathbb{E}(\xi_{i2\omega_1}^k) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = (-1)^k \Gamma^{(k)}(1), \quad \forall i = 1, 2, \dots, n \text{ e } k \text{ um inteiro positivo.} \quad (4.69)$$

Ademais, as funções elementares de $\xi_{i2\omega_1}$, que figuram na expressão equação (4.66), possuem representação em série de Taylor em termos de $\xi_{i2\omega_1}^n$. Conseqüentemente,

$$g_{ii}(\boldsymbol{\omega}_0) = \frac{s_y^2}{4} \mathbb{E} \left[\frac{\xi_{i2}^2}{(4/\alpha^2) + \xi_{i2}^2} + 2\xi_{i2} - 2\xi_{i2} \exp(-\xi_{i2}) + \left(\frac{4}{\alpha^2} + \xi_{i2}^2 \right) (1 - \exp(-\xi_{i2}))^2 \right], \quad (4.70)$$

de modo que o valor esperado presente na equação acima é o mesmo para todo $i = 1, \dots, n$.

Por outro lado, se $i \neq j$, então y_i e y_j são variáveis aleatórias independentes. Logo,

$$g_{ij}(\boldsymbol{\omega}) = \mathbb{E} \left[\frac{\partial \ell(\boldsymbol{\omega} | \mathbf{y}, \theta)}{\partial \omega_i} \frac{\partial \ell(\boldsymbol{\omega} | \mathbf{y}, \theta)}{\partial \omega_j} \right] = \mathbb{E} \left[\frac{\partial \ell(\boldsymbol{\omega} | \mathbf{y}, \theta)}{\partial \omega_i} \right] \mathbb{E} \left[\frac{\partial \ell(\boldsymbol{\omega} | \mathbf{y}, \theta)}{\partial \omega_j} \right]. \quad (4.71)$$

Usando a Regra de Leibniz (Derivação sob o sinal de integral), obtemos

$$\mathbb{E} \left[\frac{\partial \ell(\boldsymbol{\omega} | \mathbf{y}, \theta)}{\partial \omega_i} \right] = \int \frac{\partial \ell(\boldsymbol{\omega} | \mathbf{y}, \theta)}{\partial \omega_i} p(y_i; \boldsymbol{\theta}, \boldsymbol{\omega}) dy_i = \int \frac{\partial p(\boldsymbol{\omega} | \mathbf{y}, \boldsymbol{\theta})}{\partial \omega_i} dy_i = 0, \quad (4.72)$$

para todo $i = 1, \dots, n$. Isso acarreta que $g_{ij}(\boldsymbol{\omega}) = 0$ para todo $i \neq j$.

Portanto, $G(\boldsymbol{\omega}_0) = cI_p$ com $c > 0$, ou seja, a forma de perturbação $y_{i\omega} = y_i + \omega_i s_y$ é adequada para o modelo de regressão EVBS (caso em que $\gamma = 0$).

Exemplo 3. Perturbação em covariável. Suponhamos que no modelo regressão EVBS, definido em (4.56), figure uma variável explicativa contínua e a denotemos por \mathbf{x}_t . Impomos sobre \mathbf{x}_t uma perturbação aditiva da forma $x_{it\omega} = x_{it} + \omega_i s_x$, onde $t \in \{1, \dots, p\}$ e s_x é um

fator de escala associado à variável x_t . Dada uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ e um vetor de perturbação $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^\top$, a função de log-verossimilhança perturbada é

$$\ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta}) = -n \log(2) + \sum_{i=1}^n \log(\xi_{i1\omega_2}) - \sum_{i=1}^n \xi_{i2\omega_2} - \sum_{i=1}^n \exp(-\xi_{i2\omega_2}), \quad (4.73)$$

onde $\xi_{i1\omega_2} = \frac{2}{\alpha} \cosh\left(\frac{y_i - \eta_i - \beta_t \omega_i s_x}{2}\right)$ e $\xi_{i2\omega_2} = \frac{2}{\alpha} \sinh\left(\frac{y_i - \eta_i - \beta_t \omega_i s_x}{2}\right)$. Consequentemente,

$$\frac{\partial \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta})}{\partial \omega_k} = -\frac{1}{2} \beta_t s_x \left(\frac{\xi_{k2\omega_2}}{\xi_{k1\omega_2}} - \xi_{k1\omega_2} (1 - \exp(-\xi_{k2\omega_2})) \right). \quad (4.74)$$

Usando uma argumentação similar àquela empregada para o caso em que se perturba a variável resposta, obtemos

$$\mathbb{E}(\xi_{i2\omega_2}^k) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \mathbb{E}(\xi_{i2}^k) = (-1)^k \Gamma^{(k)}(1), \forall i = 1, 2, \dots, n \text{ e } k \text{ um inteiro positivo.} \quad (4.75)$$

Mais uma vez, observemos que os momentos centrais $\mathbb{E}(\xi_{i2\omega_2}^k) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$ não dependem dos valores associados à i -ésima observação. Logo,

$$g_{ii}(\boldsymbol{\omega}_0) = \frac{1}{4} (\beta_t s_x)^2 \mathbb{E} \left[\frac{\xi_{i2}^2}{(4/\alpha^2) + \xi_{i2}^2} + 2\xi_{i2}(1 - \exp(-\xi_{i2})) + \left(\frac{4}{\alpha^2} + \xi_{i2}^2 \right) (1 - \exp(-\xi_{i2}))^2 \right], \quad (4.76)$$

assume o mesmo valor para todo $i = 1, 2, \dots, n$. Por outro lado, a hipótese de independência entre as variáveis y_i e y_j , para $i \neq j$, resulta que

$$g_{ij}(\boldsymbol{\omega}) = \mathbb{E} \left[\frac{\partial \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta})}{\partial \omega_i} \frac{\partial \ell(\boldsymbol{\omega}|\mathbf{y}, \boldsymbol{\theta})}{\partial \omega_j} \right] = 0. \quad (4.77)$$

Portanto, $G(\boldsymbol{\omega}_0) = cI_p$ com $c > 0$, ou seja, a perturbação em uma covariável, conforme descrita acima, é adequada para o modelo de regressão EVBS (caso em que $\gamma = 0$).

4.4.4 Aplicação.

Com a finalidade de exemplificar a aplicabilidade da técnica de influência local em detectar observações influentes, consideramos um conjunto de dados registrados pela estação Recife-Curado, código 82900, latitude -8.05917, longitude -34.95917 e altitude: 11.3 m, localizada em Recife-Pernambuco-Brasil. Os dados foram obtidos junto ao Instituto Nacional de Meteorologia (INMET), no site <<https://portal.inmet.gov.br>>, e indicam as temperaturas máximas anuais, em graus Celsius, relacionando-as com a umidade relativa do ar média, correspondente ao dia em que se registrou o valor máximo anual da temperatura na estação. As temperaturas máximas anuais estão representadas na Figura 15 (a) e correspondem ao período com início

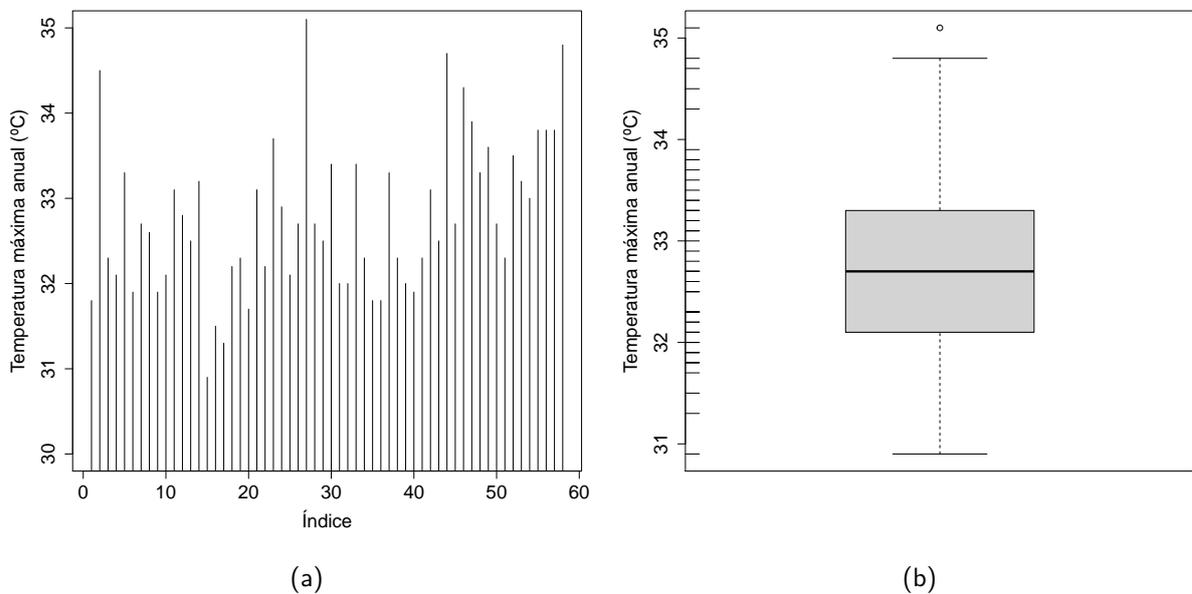
no ano de 1962 e término em 2019, resultando em uma amostra de tamanho $n = 58$. As estatísticas descritivas contidas na Tabela 25 e o Box-plot exibido na Figura 15 (b) corroboram com a hipótese de que a distribuição da variável temperatura possui uma leve assimetria.

Tabela 25 – Estatísticas descritivas para temperatura máxima anual (°C) na estação Recife-Curado (1962:2019).

n	Mínimo	Mediana	Média	Máximo	Desvio	Assimetria	Curtose
58	30.90	32.70	32.78	35.10	0.90	0.54	2.92

Fonte: Elaborada pelo autor (2022)

Figura 15 – Gráfico de dispersão em (a) e o Box-plot em (b) das temperaturas máximas anuais registradas na estação Recife-Curado (1962:2019).



Fonte: Elaborada pelo autor (2022)

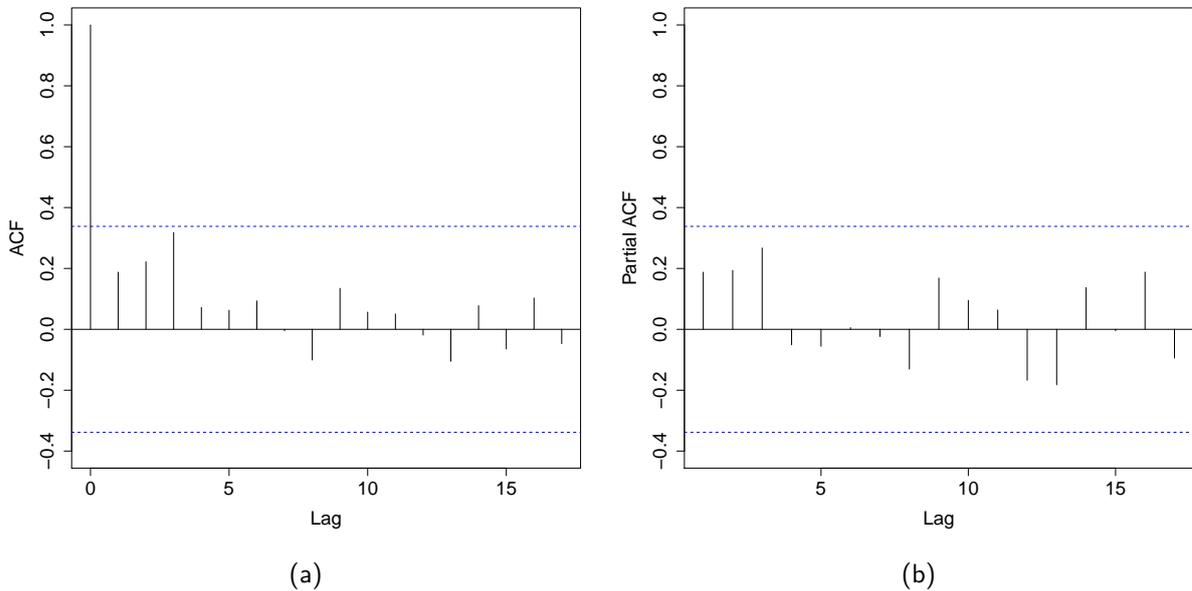
Os gráficos das funções de autocorrelação amostral e parcial das temperaturas máximas anuais registradas na estação, exibidos nas Figuras 16 (a) e 16 (b), respectivamente, trazem evidências de que a correlação existente entre as observações não é significativa. Sendo assim, admitiremos a hipótese de independência entre as mesmas.

Em nossa análise, postulamos um modelo de regressão EVBS no qual a variável resposta, Y , é o logaritmo natural da temperatura máxima anual, T , e a variável explicativa é a umidade relativa do ar média, representada por X . Em outros termos, admitimos que

$$Y_i = \log(T_i) = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, 58, \quad (4.78)$$

onde $\varepsilon_i = \log(\delta_i) \sim \log\text{-EVBS}(\alpha, 0, \gamma = 0)$. Ao realizarmos um ajuste preliminar, obtivemos as estimativas de máxima verossimilhança para os parâmetros, acompanhadas do erro-padrão

Figura 16 – Gráficos das funções de autocorrelação amostral em (a) e autocorrelação parcial em (b) das temperaturas máximas anuais registradas na estação Recife-Curado (1962:2019).



Fonte: Elaborada pelo autor (2022)

do EMV e o nível descritivo (p -valor), exibidos na Tabela 26. Destaquemos que o coeficiente associado à variável explicativa é significativo ao nível nominal de 5%.

Tabela 26 – Estimativa de MV dos coeficientes da regressão acompanhada do seu respectivo erro padrão e nível descritivo.

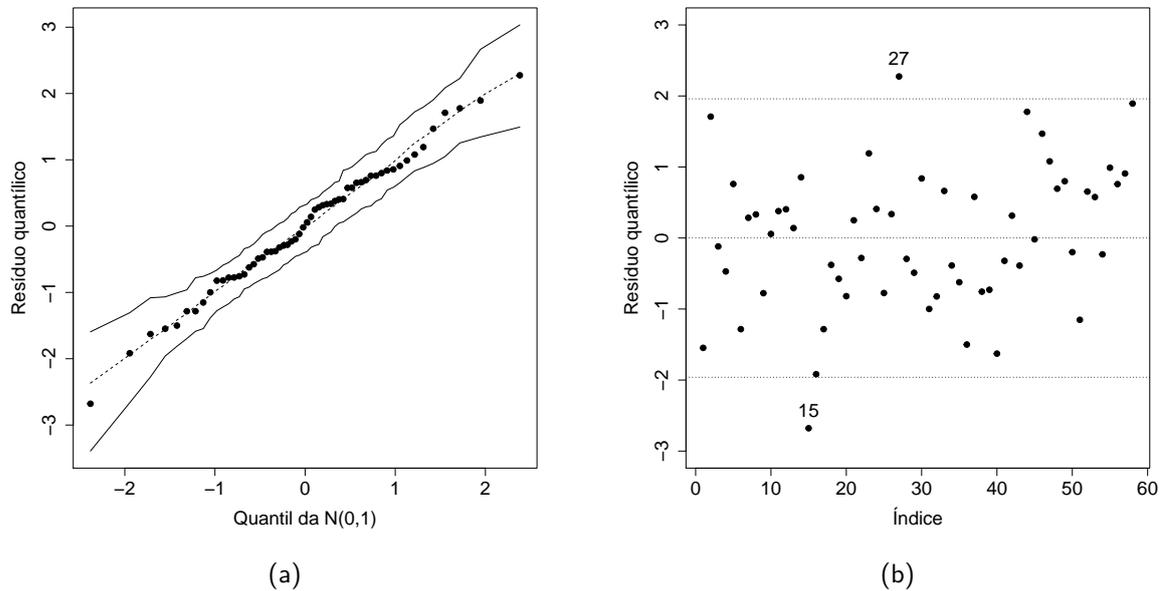
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$
emv	3.5831	-0.1460	0.0219
se	0.0400	0.0549	0.0022
p -valor	0.0000	0.0078	—

Fonte: Elaborada pelo autor (2022)

No que tange à adequação do modelo de regressão EVBS aos dados sob análise, a Figura 17 (a) exhibe o gráfico normal de probabilidades com envelope para o resíduo quantílico, sugerindo que não há indícios que nos a rejeitar a hipótese de que a resposta segue uma distribuição log-EVBS. Além disso, ao realizarmos os testes Shapiro-Wilk e Kolmogorov-Smirnov sobre os resíduos quantílicos, ilustrados na Figura 17 (b), obtivemos os p -valores 0.9821 e 0.9598, respectivamente. Logo, decidimos pela não rejeição da hipótese de normalidade dos resíduos quantílicos, ratificando a adequação do modelo de regressão EVBS ($\gamma = 0$).

A partir deste ponto, conduzimos uma análise de influência local sobre o modelo de regressão postulado para dois diferentes esquemas apropriados de perturbação: ponderação de

Figura 17 – Gráfico normal de probabilidades, com envelope, para o resíduo quantílico em (a) e os resíduos quantílicos correspondentes às observações em (b).



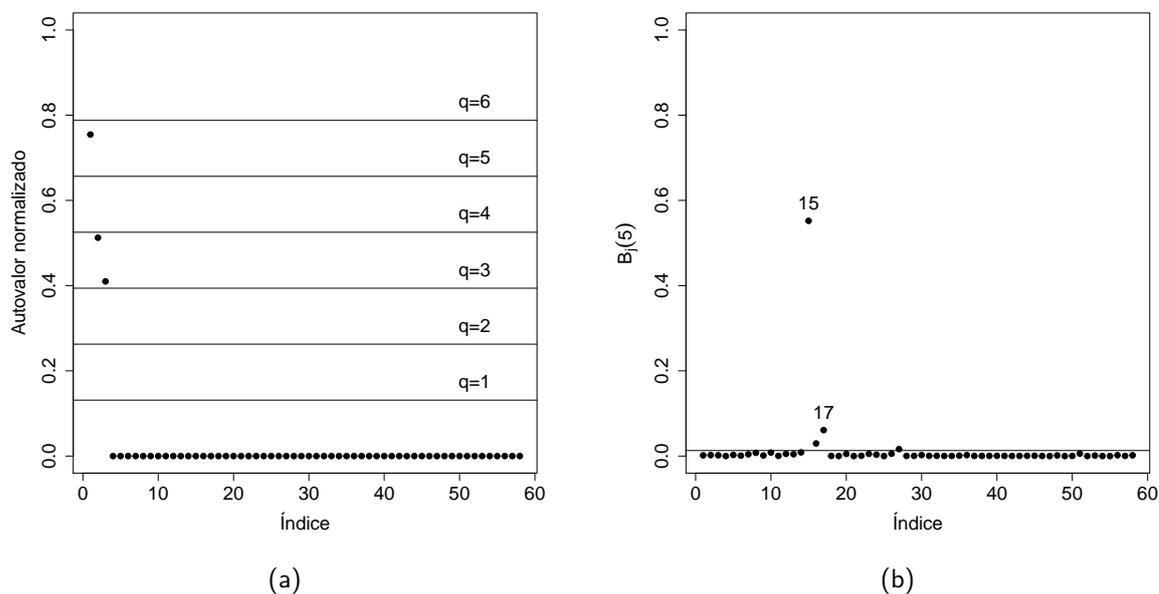
Fonte: Elaborada pelo autor (2022)

casos e perturbação na resposta.

Sob o esquema de ponderação de casos, os autovalores normalizados são 0.7546, 0.5124 e 0.4099, representados na Figura 18 (a). Logo, constatamos que associado ao maior autovalor tem-se um o autovetor 5-influente, indicando a direção na qual a curvatura conforme é máxima. Por meio da equação (3.32), obtemos a contribuição agregada dos vetores básicos de perturbação para o autovetor 5-influente, ou seja, a contribuição de cada direção canônica ao valor máximo da curvatura, isso é ilustrado na Figura 18 (b). Além disso, se usarmos $b(q = 5) = \frac{1}{58}(0.7546) = 0.0130$ na construção de um valor de referência, como sugere (POON; POON, 2010), para julgarmos a contribuição agregada de cada direção canônica de perturbação, classificamos as observações correspondentes ao números 15, 16 e 17 como sendo potencialmente influentes, conforme ilustra a Figura18 (b).

Quando perturbamos a variável resposta, os autovalores normalizados são 0.8958, 0.4182 e 0.1507, e estão representados na Figura 19 (a). Logo, o autovalor correspondente ao valor máximo da curvatura está associado a um autovetor 6-influente. A contribuição agregada dos vetores básicos de perturbação para o autovetor 6-influente é ilustrada na Figura 18 (b). Se usarmos $b(q = 6) = \frac{1}{58}(0.8958) = 0.0154$ na construção de um valor de referência para julgarmos a contribuição de cada direção canônica ao valor máximo da curvatura, mais uma vez constatamos que as observações correspondentes aos números 15, 16 e 17 são potencialmente

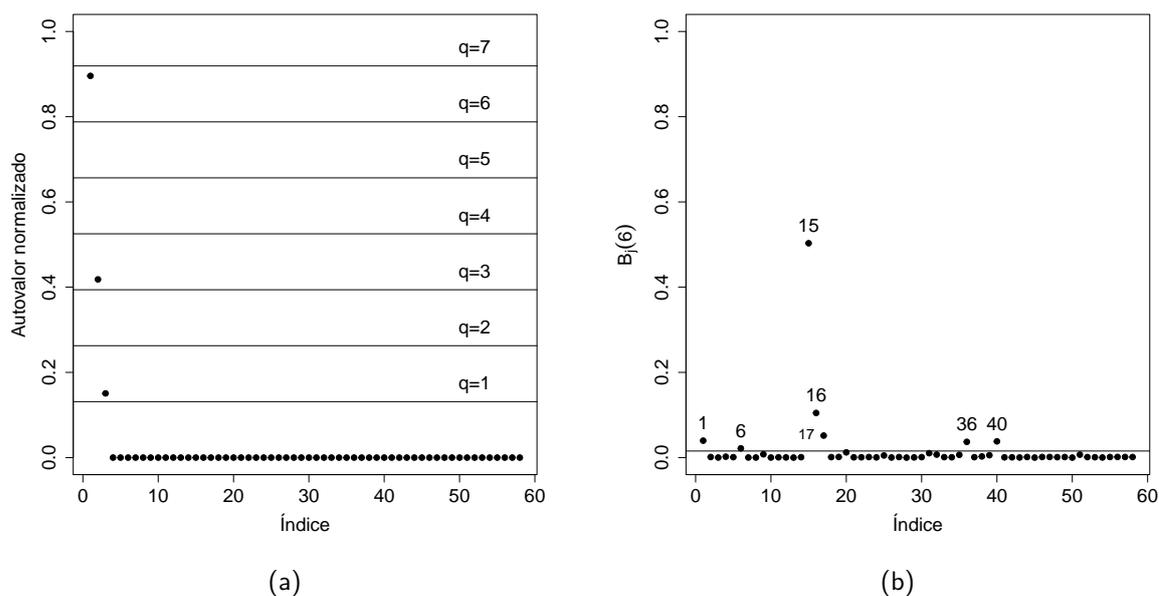
Figura 18 – Autovalores normalizados em (a); Contribuição agregada dos vetores básicos de perturbação em (b), sob ponderação de casos.



Fonte: Elaborada pelo autor (2022)

influentes. Agregando-se ao grupo de influentes as observações associadas aos números 1, 6, 36 e 40. Isso é ilustrado na Figura 19 (b).

Figura 19 – Autovalores normalizados em (a); Contribuição agregada dos vetores básicos de perturbação em (b), perturbação na resposta.



Fonte: Elaborada pelo autor (2022)

Com o objetivo de verificar o impacto que as observações correspondentes aos números

15, 16 e 17, em particular, exercem sobre as estimativas dos parâmetros, reajustamos o modelo, excluindo individualmente cada uma delas. Na Tabela 27, a seguir, constam as novas estimativas MV, a respectiva taxa de variação da grandeza, em porcentagem, denotada por R_{θ_j} e definida por $R_{\theta_j} = \left(\frac{\hat{\theta}_{j(i)} - \hat{\theta}_j}{|\hat{\theta}_j|} \right) \times 100$, em que $\hat{\theta}_{j(i)}$ denota a estimativa de $\hat{\theta}_j$ obtida após remoção da i -ésima observação, e o nível descritivo (p -valor) associado aos coeficientes da regressão.

Tabela 27 – Estimativa de MV, taxa de variação, em porcentagem, e nível descritivo associado aos coeficientes da regressão.

Observação excluída		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\alpha}$
15 ^a	emv	3.5542	-0.1042	0.0204
	R_{θ_j}	-0.81%	+28.62 %	-6.84%
	p -valor	0.0000	0.0431	—
16 ^a	emv	3.5862	-0.1490	0.0216
	R_{θ_j}	0.09%	+2.04%	-1.44%
	p -valor	0.0000	0.0048	—
17 ^a	emv	3.5673	-0.1234	0.0220
	R_{θ_j}	-0.44%	+15.48%	0.60%
	p -valor	0.0000	0.0382	—

Fonte: Elaborada pelo autor (2022)

A análise confirmatória não contradiz o fato de que a 15^a observação é influente, tanto sob o esquema de ponderação de casos quanto sob o de perturbação na resposta, pois a exclusão dessa observação acarreta uma taxa de variação expressiva na estimativa do coeficiente associado à covariável. Mais precisamente, os resultados apresentados na Tabela 27 revelam que a exclusão da 15^a observação causa uma variação de +28.62% na estimativa de β_1 . Porém, tal exclusão não altera a significância do coeficiente associado à variável explicativa. Além disso, a análise evidencia a influência correspondentes as observações de número 16 e 17. Particularmente, a exclusão da 16^a implica em uma taxa de variação inexpressiva na estimativa de α e dos coeficientes da regressão, β_0 e β_1 . Por outro lado, a exclusão da 17^a observação causa variação moderada na estimativa do coeficiente da covariável, sem mudar a significância deste coeficiente. Uma inspeção sobre a base empírica indica que os valores associados às observações influentes compõem um conglomerado de dados onde as temperaturas máximas anuais atingiram os menores patamares.

Após a análise de influência local combinada com a investigação do comportamento do resíduo quantílico, conclui-se pela adequação do modelo de regressão postulado, dado pela

equação (4.78), para explicar a logaritmo da temperatura máxima anual em termos da unidade relativa do ar, registradas na Estação Recife-Curado. Em outros termos, decidimos pela adequação do modelo de regressão EVBS dado por

$$Y_i = \log(T_i) = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, 58. \quad (4.79)$$

em que a estimativa de máxima verossimilhança de cada um dos parâmetros e seu respectivo erro-padrão, entre parêntesis, são $\hat{\beta}_0 = 3.5831(0.0400)$, $\hat{\beta}_1 = -0.1460(0.0549)$ e $\hat{\alpha} = 0.0219(0.0022)$, resultando no seguinte modelo de ajustado $\hat{T} = \exp(3.5831 - 0.1460x)$. Logo, uma diminuição de 1/10 na unidade relativa média diária acarreta uma variação no valor esperado da temperatura máxima anual de aproximadamente $(\exp(0.0146) - 1) \cdot 100 = 1.4707\%$.

5 O MODELO BIRNBAUM-SAUNDERS-PARETO GENERALIZADO

5.1 INTRODUÇÃO

O uso do modelo GEV em Estatística de Extremos é usual, pois o Teorema de Tipos Extremos, (GNEDENKO, 1943), estabelece a distribuição GEV como uma versão unificada de todas as leis limites não-degeneradas do máximo amostral, devidamente padronizado, de variáveis aleatórias iid. No método BM admite-se que é possível obter o máximo amostral associado a cada um dos blocos de mesma dimensão e a classe GEV é a família de distribuições utilizada para modelar os máximos dessas sub-amostras. Ao definirem o modelo EVBS, (FERREIRA; GOMES; LEIVA, 2012) acentuam que o propósito é obter uma versão da distribuição BS, não só do ponto de vista probabilístico, mas que se revele como um modelo paramétrico útil na análise estatística de eventos extremos. De fato, a substituição da distribuição normal pela GEV no núcleo da transformação que determina a distribuição BS resultou em um modelo próspero, conforme evidencia os trabalhos de (GOMES; GUILLOU, 2015), (LEIVA et al., 2016), (LILLO et al., 2018) e a aplicação exibida no Capítulo 2, Seção 2.6.

E se o critério para determinação das observações extremas for alterado, e empregarmos, por exemplo, a metodologia POT, o modelo EVBS ainda terá a mesma desenvoltura? Sabe-se que no método POT a atenção é voltada para as observações acima de um limiar elevado u , chamado *threshold*, e a distribuição Pareto generalizada (GPD) surge como distribuição limite dos excessos $Y_i = X_i - u | X_i > u, i = 1, \dots, k$ para u suficientemente elevado, conforme asseguram (BALKEMA; HANN, 1974). Sendo assim, propomos o uso da GPD padrão, em vez da distribuição normal padrão, para compor o núcleo da transformação obtida por (BIRNBAUM; SAUNDERS, 1969a) para determinar a distribuição BS. Com isso, derivamos um modelo que herda características das distribuições BS e GPD e esperamos que o mesmo modele satisfatoriamente dados extremos obtidos quando direcionamos nossa atenção às observações que excedem um limiar elevado.

Na Seção 5.2, deste capítulo, definimos uma nova família de distribuições, denominada Birnbaum-Saunders-Pareto generalizada (GPD-BS), e deduzimos algumas propriedades que exprimem características do comportamento dessa família. Em seguida, na Seção 5.3, lidamos com estimação por máxima verossimilhança na família GPD-BS, onde derivamos as expressões correspondentes às entradas da função escore, garantindo que algumas dessas assumem valor

esperado igual a zero. Para concluir, na Seção 5.4, desenvolvemos uma aplicação em que se investigou os excessos de precipitações pluviométricas, acima de um limiar elevado, registrados na estação meteorológica Recife-Curado, utilizando os modelos GPD e GPD-BS.

5.2 A DISTRIBUIÇÃO GPD-BS

5.2.1 Definição e algumas propriedades

Diz-se que uma variável aleatória X segue uma distribuição Pareto generalizada padrão (GPD) com parâmetro $\gamma \in \mathbb{R}$, se a fda de X admite ser escrita na forma

$$G(x) = \begin{cases} 1 - \exp(-x), & \text{se } \gamma = 0 \text{ e } x > 0. \\ 1 - (1 + \gamma x)^{-1/\gamma}, & \text{se } \gamma \neq 0 \text{ e } x \in S_\gamma. \\ 0, & \text{se } x \leq 0. \\ 1, & \text{se } x \geq -1/\gamma. \end{cases}$$

De modo que $S_\gamma = (0, \infty)$ para $\gamma \geq 0$ e $S_\gamma = (0, -1/\gamma)$ para $\gamma < 0$.

Dada a variável aleatória $X \sim \text{GPD}(0, 1, \gamma)$, consideremos a variável T resultante da seguinte transformação sobre X :

$$T = \beta \left(\frac{\alpha X}{2} + \sqrt{\left(\frac{\alpha X}{2} \right)^2 + 1} \right)^2, \quad (5.1)$$

em que $\alpha > 0$ e $\beta > 0$. Uma vez que a transformação dada pela equação (5.1) é monótona crescente, segue que a variável aleatória T é positiva e sua fda, segundo o valor do parâmetro γ , assume a forma a seguir:

1. Se $\gamma > 0$, então

$$F(t) = \begin{cases} 0, & \text{se } t \leq \beta. \\ 1 - \left[1 + \frac{\gamma}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]^{-1/\gamma}, & \text{se } t > \beta. \end{cases} \quad (5.2)$$

2. Se $\gamma < 0$, então

$$F(t) = \begin{cases} 0, & \text{se } t \leq \beta. \\ 1 - \left[1 + \frac{\gamma}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right]^{-1/\gamma}, & \text{se } \beta < t < t^*. \\ 1, & \text{se } t \geq t^*, \end{cases} \quad (5.3)$$

$$\text{onde } t^* = \frac{\beta}{2} \left[\left(\frac{\alpha}{\gamma} \right)^2 + 2 - \frac{\alpha}{\gamma} \sqrt{\left(\frac{\alpha}{\gamma} \right)^2 + 4} \right].$$

3. O comportamento limite da função expressa em (5.2) quando γ tende a zero deve ser interpretado como sendo a situação em que $\gamma = 0$. Neste caso, obtemos

$$F(t) = \begin{cases} 0, & \text{se } t \leq \beta. \\ 1 - \exp \left[-\frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right) \right], & \text{se } t > \beta. \end{cases} \quad (5.4)$$

Em resumo, a fda de T admite ser reescrita na forma

$$F(t) = G \left(\frac{1}{\alpha} \varphi \left(\frac{t}{\beta} \right) \right), \quad \alpha > 0, \beta > 0 \text{ e } \gamma \in \mathbb{R}, \quad (5.5)$$

onde $\varphi(t) = t^{1/2} - t^{-1/2}, t > \beta$. Neste caso, dizemos que a variável aleatória T segue uma distribuição Birnbaum-Saunders-Pareto generalizada com parâmetros $\alpha > 0, \beta > 0, \gamma \in \mathbb{R}$ e a denotamos por $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$. Conseqüentemente, a fdp da variável $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ é dada por

$$f(t) = \begin{cases} \frac{1}{2\alpha\beta} \left[\left(\frac{t}{\beta} \right)^{-1/2} + \left(\frac{t}{\beta} \right)^{-3/2} \right] \exp \left(-\frac{1}{\alpha} \left(\left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{t}{\beta} \right)^{-1/2} \right) \right); & t \in J, \text{ se } \gamma = 0. \\ \frac{1}{2\alpha\beta} \left[\left(\frac{t}{\beta} \right)^{-1/2} + \left(\frac{t}{\beta} \right)^{-3/2} \right] \left(1 + \frac{\gamma}{\alpha} \left(\left(\frac{t}{\beta} \right)^{1/2} - \left(\frac{t}{\beta} \right)^{-1/2} \right) \right)^{-1-\frac{1}{\gamma}}; & t \in J, \text{ se } \gamma \neq 0. \end{cases} \quad (5.6)$$

No caso em que $\gamma \geq 0$, o suporte da fdp é o intervalo real $J = (\beta, \infty)$; por outro lado se $\gamma < 0$, o suporte é o intervalo limitado $J = \left(\beta, \frac{\beta}{2} \left[\left(\frac{\alpha}{\gamma} \right)^2 + 2 - \frac{\alpha}{\gamma} \sqrt{\left(\frac{\alpha}{\gamma} \right)^2 + 4} \right] \right)$. A função expressa na equação (5.6) admite ser reescrita na forma

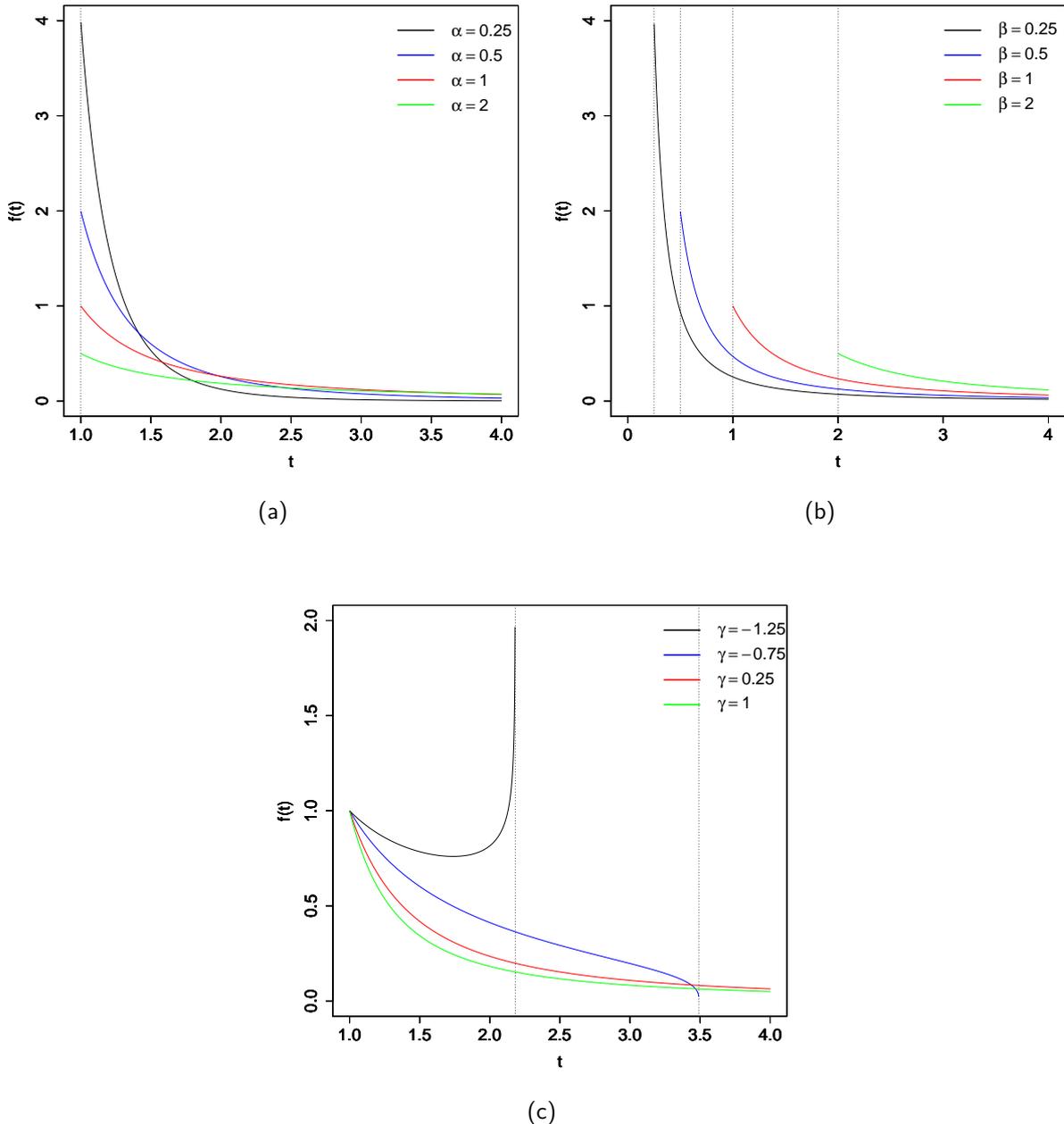
$$f(t) = \begin{cases} \frac{1}{\alpha\beta} \varphi' \left(\frac{t}{\beta} \right) \exp \left(-\frac{1}{\alpha} \varphi \left(\frac{t}{\beta} \right) \right); & t \in J, \text{ se } \gamma = 0. \\ \frac{1}{\alpha\beta} \varphi' \left(\frac{t}{\beta} \right) \left[1 + \frac{\gamma}{\alpha} \varphi \left(\frac{t}{\beta} \right) \right]^{-1-\frac{1}{\gamma}}; & t \in J, \text{ se } \gamma \neq 0. \end{cases} \quad (5.7)$$

onde $\varphi'(t) = \frac{1}{2}(t^{-1/2} + t^{-3/2}), t > \beta$.

A seguir, exibimos algumas ilustrações que descrevem o comportamento do gráfico da fdp de uma variável $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$. No caso em que $\gamma = 0$, a Figura 20 (a) revela evidências de que a fdp é monótona, assimétrica e não possui moda. Além disso, à medida que α aproxima-se de zero, os valores próximos de β se tornam cada vez mais densos. A Figura 20 (b) revela a natureza de β como sendo o ínfimo do suporte da distribuição. Na Figura 20 (c), caso em γ assume valores ora positivos ora negativos e os demais parâmetros estão fixos, acentua-se que quando $\gamma > 0$ a distribuição tem cauda pesada dentre as que compõem a família GPD-BS, preservando a monotonicidade vista no caso anterior. Por outro lado, se

$\gamma = -1.25 < 0$ nota-se que a distribuição possui cauda curta com limite superior do suporte finito. Ademais, a monotonicidade da fdp não se verifica e há indícios de que a mesma é ilimitada.

Figura 20 – Gráficos da fdp da GPD-BS para $\alpha \in \{0.25, 0.5, 1, 2\}$, $\beta = 1$ e $\gamma = 0$ em (a); para $\alpha = 1$, $\beta \in \{0.25, 0.5, 1, 2\}$ e $\gamma = 0.25$ em (b) e $\alpha = 1$, $\beta = 1$ e $\gamma \in \{-1.25, -0.75, 0.25, 1\}$ em (c).



Fonte: elaborada pelo autor (2022)

Agora, deduziremos algumas propriedades da variável $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$.

Propriedade 1. Se $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ e $\gamma \geq -1$, então a função densidade de probabilidade, f , da variável aleatória T é estritamente decrescente no suporte J .

Prova: suponhamos que $\gamma \geq -1$ e $\gamma \neq 0$. Logo, ao definirmos a função de uma variável real $\psi : J \mapsto \mathbb{R}$, $\psi(t) = \log(f(t))$, obtemos

$$\psi'(t) = \frac{1}{\beta} \left[\frac{\varphi''(t/\beta)}{\varphi'(t/\beta)} - \frac{(1+\gamma)}{\alpha} \varphi'(t/\beta) \left(1 + \frac{\gamma}{\alpha} \varphi(t/\beta) \right) \right] \quad (5.8)$$

Se $\gamma \geq -1$, então $-(1+\gamma) \leq 0$. Logo, uma vez que $\alpha > 0$, $\beta > 0$, $(1 + (\gamma/\alpha)\varphi(t/\beta)) > 0$, $\varphi'(t/\beta) > 0, \forall t \in J$ e $\varphi''(t/\beta) = -\frac{1}{4}(t^{-3/2} + 3t^{-5/2}) < 0, \forall t \in J$, segue que $\psi'(t) < 0, \forall t \in J$.

No caso $\gamma = 0$, a função f assume a forma dada pela primeira sentença da equação (5.6) e a análise do sinal de ψ' é similar ao caso já desenvolvido. Portanto, a fdp f é estritamente decrescente em J , se $\gamma \geq -1$.

Observação: em geral, se $\gamma < -1$, a monotonicidade de f não é assegurada. De fato, se $\alpha = \beta = 1$ e $\gamma = -\frac{3}{2}$, então a fdp não é monótona, pois para $t_0 = 100/81$, $t_1 = 64/49$ e $t_2 = 25/16$ tem-se $t_0 < t_1 < t_2$, mas $f(t_0) \approx 0.925 > f(t_1) \approx 0.917$ e $f(t_1) \approx 0.917 < f(t_2) \approx 0.954$. Isso corrobora com hipótese de que para $\gamma < -1$, a fdp possui a forma de um gancho.

Propriedade 2. Se $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ e $c > 0$, então $cT \sim \text{GPD-BS}(\alpha, c\beta, \gamma)$.

Prova: a justificativa é imediata por meio da equação (5.5).

Propriedade 3. Sejam T_1 e T_2 variáveis aleatórias tais que $T_1 \sim \text{EVBS}(\alpha, \beta, \gamma)$ e $T_2 \sim \text{GPD-BS}(\alpha, \beta, \gamma)$. As funções de distribuição de T_1 e T_2 se relacionam analiticamente por meio da expressão

$$F_{T_2}(t) = 1 + \log(F_{T_1}(t)), \text{ se } \log(F_{T_1}(t)) > -1.$$

Prova: o resultado é imediato a partir da composição da função logarítmica com a fda da variável EVBS, dada pela equação (2.36).

Propriedade 4. Se $\gamma \geq 0$, então a função de risco, h , da variável $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ é decrescente no intervalo (β, ∞) .

Prova: se $\gamma > 0$, então

$$h(t) = \left(\frac{1}{\alpha\beta} \right) \frac{\varphi'(t/\beta)}{1 + \frac{\gamma}{\alpha} \varphi(t/\beta)}, \quad t > \beta.$$

Ao realizarmos o estudo do sinal da primeira derivada da função $\psi : (\beta, \infty) \rightarrow \mathbb{R}$, $\psi(t) = \log(h(t))$, concluímos que

$$\psi'(t) = \frac{1}{\beta} \left(\frac{\varphi''(t/\beta)}{\varphi'(t/\beta)} - \frac{\gamma}{\alpha} \frac{\varphi'(t/\beta)}{1 + \frac{\gamma}{\alpha} \varphi(t/\beta)} \right) < 0,$$

pois, $\varphi''(t/\beta) < 0$ e os demais fatores são positivos. A prova do caso $\gamma = 0$ é análoga. Portanto, h é decrescente para todo $t \in (\beta, \infty)$ quando $\gamma \geq 0$.

Propriedade 5. Se $\gamma = 0$ e T_1, \dots, T_n são variáveis aleatórias independentes e identicamente distribuídas tais que $T_i \sim \text{GPD-BS}(\alpha, \beta, \gamma = 0)$ para todo $i = 1, \dots, n$, então $T_{(1)} \sim \text{GPD-BS}\left(\frac{\alpha}{n}, \beta, 0\right)$.

Prova: se T_1, \dots, T_n são variáveis aleatórias iid com fda comum F , então

$$P(T_{(1)} \leq t) = 1 - [1 - F(t)]^n.$$

Consequentemente, a fdp da variável $T_{(1)}$ é dada por

$$g(t) = nf(t)[1 - F(t)]^{n-1}, \quad t > \beta,$$

sendo f a fdp associada a F . Se $\gamma = 0$, então F e f assumem as formas das pelas equações (5.4) e (5.7), respectivamente. Logo,

$$g(t) = \frac{1}{(\alpha/n)\beta} \varphi'(t/\beta) \left[\exp\left(-\frac{1}{(\alpha/n)} \varphi(t/\beta)\right) \right], \quad t > \beta.$$

Portanto, $T_{(1)} \sim \text{GPD-BS}\left(\frac{\alpha}{n}, \beta, 0\right)$.

Propriedade 6. $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ se, e somente se, $\frac{1}{\alpha} \varphi\left(\frac{T}{\beta}\right) \sim \text{GPD}(0, 1, \gamma)$.

Prova: inicialmente, suponhamos que $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ com $\alpha > 0$, $\beta > 0$ e $\gamma > 0$. Isso acarreta que a fda da variável T é dada por

$$F_T(t) = 1 - \left(1 + \frac{\gamma}{\alpha} \varphi(t/\beta)\right)^{-\frac{1}{\gamma}}; \quad t > \beta. \quad (5.9)$$

Agora, consideremos a variável $X = \frac{1}{\alpha} \varphi\left(\frac{T}{\beta}\right)$, onde $\varphi(t) = t^{1/2} - t^{-1/2}$, $t > \beta$. Notemos que a função φ é estritamente crescente no intervalo (β, ∞) , logo possui inversa φ^{-1} . Consequentemente, se denotarmos por G a função de distribuição de X , teremos

$$G(x) = P\left(\frac{1}{\alpha} \varphi\left(\frac{T}{\beta}\right) \leq x\right) = P(T \leq \beta \varphi^{-1}(\alpha x)) = F_T(\beta \varphi^{-1}(\alpha x)), \quad x > 0.$$

Utilizando a equação (5.9), obtemos $G(x) = 1 - (1 + \gamma x)^{-1/\gamma}$. Logo, $X = \frac{1}{\alpha} \varphi\left(\frac{T}{\beta}\right) \sim \text{GPD}(0, 1, \gamma)$. A justificativa dos casos $\gamma = 0$ ou $\gamma < 0$ é análoga.

Reciprocamente, admitamos que $X = \frac{1}{\alpha} \varphi\left(\frac{T}{\beta}\right) \sim \text{GPD}(0, 1, \gamma)$, com $\alpha > 0$, $\beta > 0$ e $\gamma \in \mathbb{R}$. Logo, dado que φ é crescente, teremos

$$P(T \leq t) = P\left(\frac{1}{\alpha} \varphi\left(\frac{T}{\beta}\right) \leq \frac{1}{\alpha} \varphi(t/\beta)\right) = G\left(\frac{1}{\alpha} \varphi(t/\beta)\right);$$

Por meio da relação (5.5), concluímos que $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$.

Propriedade 7. Se $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ e $\gamma < 1$, então

1.
$$\mathbb{E} \left(\left[1 + \frac{\gamma}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right]^{-r} \right) = \frac{1}{1 + \gamma r}, \quad r > -\frac{1}{\gamma}, \quad (5.10)$$

2.
$$\mathbb{E} \left(\log \left[1 + \frac{\gamma}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right] \right)^k = \gamma^k k!, \quad \text{para todo } k \in \mathbb{N}. \quad (5.11)$$

3.
$$\mathbb{E} \left[\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \left(\overline{G} \left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right) \right)^r \right] = \frac{1}{(r + 1 - \gamma)(r + 1)}, \quad (r + 1)/|\gamma| > 0. \quad (5.12)$$

4. Se $r \in \mathbb{N}$ e $\gamma < 1/r$, então

$$\mathbb{E} \left[\left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right)^r \right] = \frac{\Gamma(\gamma^{-1} - r)}{\gamma^{r+1} \Gamma(1 + \gamma^{-1})} r!. \quad (5.13)$$

Prova: os itens 1, 2, 3 e 4 são consequências imediatas da Propriedade 6 combinada com propriedades do valor esperado de uma variável que segue uma distribuição GPD, veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997). A título de ilustração, no Apêndice D exibimos algumas provas para os casos em que $\gamma > 0$ ou $\gamma = 0$.

Observação: embora os itens da Propriedade 7 expressem resultados do valor esperado de uma função de T , eles serão úteis para assegurar que a esperança das entradas da função score, associadas com as derivadas parciais em relação a α ou a γ , são nulas.

Propriedade 8. O quantil de ordem q , com $0 < q < 1$, de uma variável aleatória $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ é dado por:

$$t_q = \beta \left(\frac{\alpha x_q}{2} + \sqrt{\left(\frac{\alpha x_q}{2} \right)^2 + 1} \right)^2, \quad (5.14)$$

onde x_q é o quantil de ordem q da variável aleatória que segue a distribuição GPD padrão.

Prova: a justificativa é imediata utilizando a equação (5.5).

A Propriedade 8 leva-nos ao seguinte algoritmo para geração de números aleatórios de uma variável GPD-BS :

Algoritmo 2. Gerador de números aleatórios de uma variável

$$T \sim \text{GPD-BS}(\alpha, \beta, \gamma).$$

Passo 1. Gere um número aleatório, u , de uma variável $U \sim (0, 1)$.

Passo 2. Fixe os valores dos parâmetros α , β e γ .

Passo 3. Se $\gamma = 0$, considere $x_u = -\log(1 - u)$ e calcule:

$$t = \beta \left(\frac{\alpha x_u}{2} + \sqrt{\left(\frac{\alpha x_u}{2} \right)^2 + 1} \right)^2 ;$$

Se $\gamma \neq 0$, considere $x_u = \frac{1}{\gamma}[(1 - u)^{-\gamma} - 1]$ e calcule:

$$t = \beta \left(\frac{\alpha x_u}{2} + \sqrt{\left(\frac{\alpha x_u}{2} \right)^2 + 1} \right)^2 .$$

Passo 4. Repita os passos 1, 2 e 3, n vezes, a fim de obter uma amostra de n números aleatórios.

5.3 ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA

O propósito desta seção é obter equações, a serem utilizadas em rotinas computacionais, que nos levem a encontrar numericamente as estimativas de MV dos parâmetros do modelo GPD-BS. Logo, não faremos uma investigação profunda acerca da temática de estimação na família GPD-BS, pois ela revela as mesmas dificuldades encontradas na família EVBS. Concentraremos-nos no caso em $\gamma \geq -1$, pois se $\gamma < -1$, então não existe EMV que assuma valores no interior do espaço paramétrico, conforme justificaremos adiante.

Sejam T_1, \dots, T_n variáveis aleatórias independentes e identicamente distribuídas a variável $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$. Para fins de estimação, a função de log-verossimilhança de $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$, dada uma amostra de observações $\mathbf{t} = (t_1, t_2, \dots, t_n)^\top$ pode ser expressa como

$$\ell(\boldsymbol{\theta}; \mathbf{t}) = \begin{cases} -n \log(\alpha\beta) + \sum_{i=1}^n \log(\varphi'(t_i/\beta)) - \frac{1}{\alpha} \sum_{i=1}^n \varphi(t_i/\beta); & \gamma = 0. \\ -n \log(\alpha\beta) + \sum_{i=1}^n \log(\varphi'(t_i/\beta)) - \frac{1+\gamma}{\gamma} \sum_{i=1}^n \log\left(1 + \frac{\gamma}{\alpha} \varphi(t_i/\beta)\right); & \gamma \neq 0. \end{cases} \quad (5.15)$$

No caso em que $\gamma = 0$, a primeira sentença do sistema (5.15) possui derivadas parciais

dadas por:

$$\begin{cases} \frac{\partial}{\partial \alpha} \ell(\boldsymbol{\theta}; \mathbf{t}) = -\frac{n}{\alpha} + \frac{1}{\alpha^2} \sum_{i=1}^n \varphi(t_i/\beta). \\ \frac{\partial}{\partial \beta} \ell(\boldsymbol{\theta}; \mathbf{t}) = -\frac{n}{\beta} + \frac{1}{2\beta} \sum_{i=1}^n \frac{(t_i/\beta) + 3}{(t_i/\beta) + 1} + \frac{1}{\alpha\beta} \sum_{i=1}^n (t_i/\beta) \varphi'(t_i/\beta). \end{cases} \quad (5.16)$$

Notemos que primeira sentença do sistema (5.16) possibilita expressarmos a estimativa $\hat{\alpha}$ em função de $\hat{\beta}$, mais precisamente temos

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^n \left\{ \left(\frac{t_i}{\hat{\beta}} \right)^{\frac{1}{2}} - \left(\frac{t_i}{\hat{\beta}} \right)^{-\frac{1}{2}} \right\} = \frac{1}{n} \sum_{i=1}^n \varphi(t_i/\hat{\beta}).$$

Agora, ao fazermos uso do Item 3 da Propriedade 7 temos que se β é conhecido, então $\mathbb{E}(\hat{\alpha}) = \alpha$, ou seja, $\hat{\alpha}$ é não-viesado para α . Além disso, aplicando o mesmo item sobre a primeira sentença do sistema (5.16) implica que $\mathbb{E} \left(\frac{\partial}{\partial \alpha} \ell(\boldsymbol{\theta}; \mathbf{T}) \right) = 0$.

A seguir, passemos a tratar do caso $\gamma \neq 0$. A segunda sentença da equação (5.15) possui derivadas parciais dadas por:

$$\begin{cases} \frac{\partial}{\partial \alpha} \ell(\boldsymbol{\theta}; \mathbf{t}) = -\frac{n}{\alpha} + \frac{1+\gamma}{\alpha} \sum_{i=1}^n \frac{1}{\alpha} \varphi(t_i/\beta) \frac{1}{(1 + \frac{\gamma}{\alpha} \varphi(t_i/\beta))}. \\ \frac{\partial}{\partial \beta} \ell(\boldsymbol{\theta}; \mathbf{t}) = -\frac{n}{\beta} + \frac{1}{2\beta} \sum_{i=1}^n \frac{(t_i/\beta) + 3}{(t_i/\beta) + 1} + \frac{1+\gamma}{\alpha\beta} \sum_{i=1}^n \frac{(t_i/\beta) \varphi'(t_i/\beta)}{(1 + \frac{\gamma}{\alpha} \varphi(t_i/\beta))}. \\ \frac{\partial}{\partial \gamma} \ell(\boldsymbol{\theta}; \mathbf{t}) = \frac{1}{\gamma^2} \sum_{i=1}^n \log \left(1 + \frac{\gamma}{\alpha} \varphi(t_i/\beta) \right) - \frac{1+\gamma}{\gamma} \sum_{i=1}^n \frac{1}{\alpha} \varphi(t_i/\beta) \frac{1}{(1 + \frac{\gamma}{\alpha} \varphi(t_i/\beta))}. \end{cases} \quad (5.17)$$

Sendo $\varphi(t_i) = t_i^{1/2} - t_i^{-1/2}$, com $i = 1, 2, \dots, n$. Neste ponto, observemos que $t_i > \beta$ implica $\varphi(t_i/\beta) > 0$. Logo, se $\gamma < -1$, ou seja $1 + \gamma < 0$, o membro direito da primeira equação do sistema (5.17) é negativo. Isso acarreta que o gradiente de ℓ não se anula no interior do espaço paramétrico $\Theta = (0, \infty) \times (0, \infty) \times (-\infty, -1)$. Portanto, no caso em que $\gamma < -1$, não existe EMV no interior de Θ .

A função escore de $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$, dado $\mathbf{t} = (t_1, \dots, t_n)^\top$, é

$$\mathbf{U}(\boldsymbol{\theta}, \mathbf{t}) = \frac{\partial \ell(\boldsymbol{\theta}, \mathbf{t})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial}{\partial \alpha} \ell(\boldsymbol{\theta}, \mathbf{t}) \\ \frac{\partial}{\partial \beta} \ell(\boldsymbol{\theta}, \mathbf{t}) \\ \frac{\partial}{\partial \gamma} \ell(\boldsymbol{\theta}, \mathbf{t}) \end{pmatrix}_{3 \times 1}. \quad (5.18)$$

onde as entradas da matriz acima são as equações que compõem o sistema (5.17). Logo um EMV de $\boldsymbol{\theta}$, denotado por $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})^\top$ deve satisfazer, como condição necessária, ao sistema de equações $\mathbf{U}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$. Porém, uma vez que o sistema de equações $\mathbf{U}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$, em geral, possui solução definida implicitamente, que não possibilita a determinação de uma solução

analítica, é natural o uso de procedimentos numéricos de otimização não-linear a fim de obter as estimativas de máxima verossimilhança. Mais uma vez, observemos que os Itens 2 e 3 da Proposição 7 acarretam $\mathbb{E}\left(\frac{\partial}{\partial\alpha}\ell(\boldsymbol{\theta}; \mathbf{T})\right) = 0$ e $\mathbb{E}\left(\frac{\partial}{\partial\gamma}\ell(\boldsymbol{\theta}; \mathbf{T})\right) = 0$, supondo $\gamma < 1$ para que se possa assegurar a existência dos momentos. Já em relação ao valor esperado $\mathbb{E}\left(\frac{\partial}{\partial\beta}\ell(\boldsymbol{\theta}; \mathbf{T})\right)$, talvez os itens da Propriedades 7 não possibilitam obtê-lo analiticamente, sendo necessário integração numérica.

5.4 APLICAÇÃO

5.4.1 Introdução

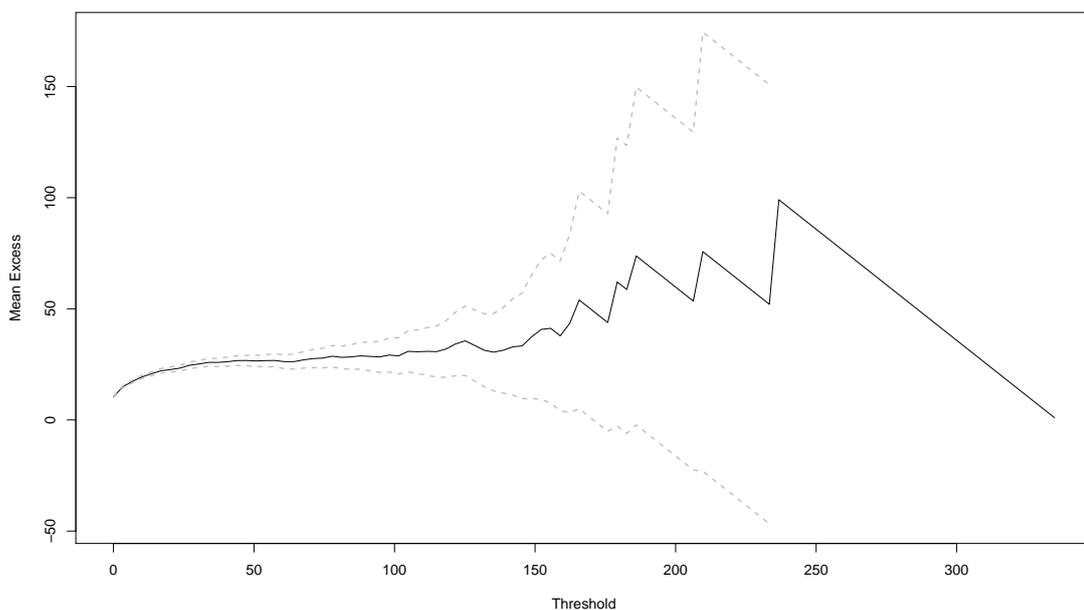
Na aplicação desenvolvida no Capítulo 2, Seção 2.6, empregamos o método BM para investigar o comportamento da precipitação pluviométrica máxima anual (mm) na estação Recife-Curado. Ao considerarmos a partição dos dados em blocos anuais e tomarmos o máximo em cada bloco, é evidente a possibilidade de nos depararmos com observações na amostra de máximos que são inferiores a valores ocorridos em outros anos. Isso motiva o uso da metodologia POT que considera como extremos as observações que superam um limiar elevado u , as quais denominamos de excedentes. Se a variável X sob investigação é tal que $F_X \in \text{MDA}(H_\gamma)$, de acordo com (BALKEMA; HANN, 1974), o excesso de X acima do limiar u possui a GPD como distribuição limite, quando u é suficientemente elevado. Nesta aplicação usamos a metodologia POT e a empregamos sobre os dados de precipitações pluviométricas registrados na estação Recife-Curado de 01/01/1962 até 31/12/2019.

A variável aleatória de interesse é o excesso de precipitação pluviométrica acima de um limiar elevado u . Iniciamos com uma análise exploratória dos dados com o propósito de selecionarmos o *threshold*, u , e obter informações preliminares sobre o comportamento da cauda direita da distribuição subjacente aos dados. Em seguida, lidamos com estimação pontual dos parâmetros de interesse em Estatísticas de Extremos. Obtemos estimativas provenientes de dois modelos paramétricos: o GPD, que é o candidato natural para modelagem, e o GPD-BS, aquele cuja potencialidade deseja-se inspecionar.

5.4.2 Análise exploratória dos dados

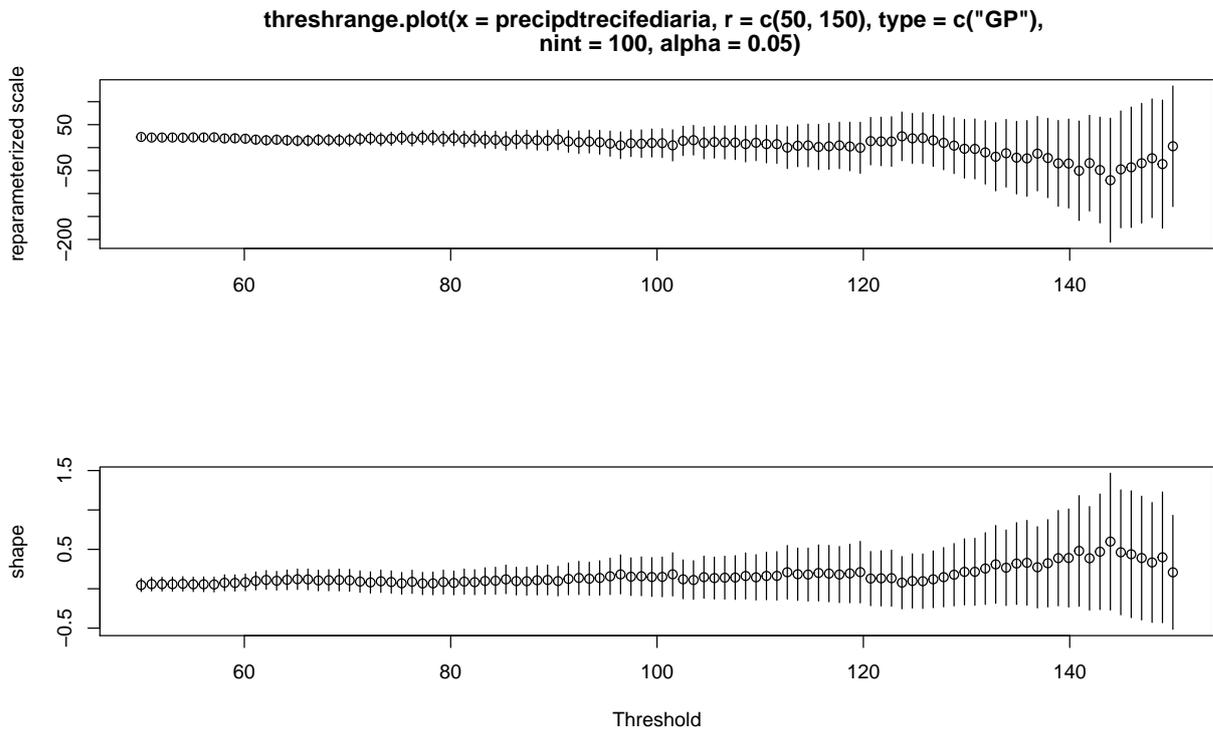
Um dos instrumentos gráficos que analisamos para auxiliar na escolha do *threshold* e na identificação do tipo de cauda da distribuição é ME-plot, Figura 21, que é o gráfico da função empírica correspondente à função de excesso médio, detalhes acerca da definição e propriedades dessa função podem ser encontrados em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997), págs. 294-303. Dado que o gráfico apresenta comportamento predominantemente crescente, a expectativa é de que a cauda da distribuição subjacente aos dados pertença ao máx-domínio Fréchet da família GPD, ou seja, tenhamos uma distribuição de cauda pesada ($\gamma > 0$). Visto que o gráfico da função excesso médio de uma GPD é uma função afim, podemos inspecionar o ME-plot e identificar um intervalo onde tal comportamento seja verificado, possibilitando uma escolha de *threshold* nesse intervalo. Outra ferramenta que nos guiará na escolha do *threshold* é o gráfico gerado pela função `thresrange.plot()`, disponível no pacote *extRemes*, exibido na Figura 22, que exprime o comportamento das estimativas dos parâmetros de escala e forma do modelo GDP versus o *threshold*. Na Figura 22, nota-se certa estabilidade da estimativas desses parâmetros em torno do valor real 96. Além disso, este número pertence ao intervalo onde o ME-plot possui comportamento linear. Diante das ponderações feitas, fixamos o número real $u = 96$ como sendo o *threshold*, resultando em uma amostra de 90 excessos.

Figura 21 – ME-plot das precipitações pluviométricas registradas na estação Recife-Curado (1962:2019).



Fonte: elaborada pelo autor (2022)

Figura 22 – Estimativas dos parâmetros de escala e forma versus *threshold*.



Fonte: elaborada pelo autor (2022)

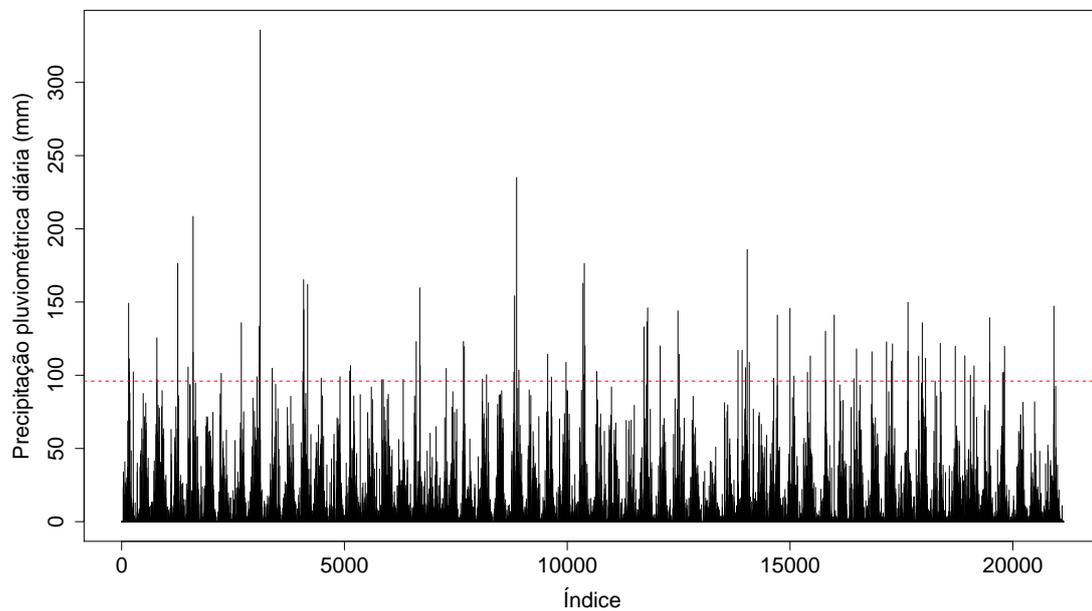
Uma vez fixado o *threshold* em $u = 96$, determina-se o excesso como sendo a diferença $Y_i = X_i - u | X_i > u$. Ou equivalentemente, os excedentes, isto é, os próprios valores de X_i tal que $X_i > u$. Os gráficos de dispersão da Figura 23 exibem a amostra original de dados de precipitações pluviométricas diárias registradas na estação Recife-Curado (1962:2019) e os respectivos excessos além do limiar $u = 96$ mm.

O histograma dos excessos, acima do limiar $u = 96$ mm, registrados na estação Recife-Curado, Figura 24 (a), revela que a distribuição pertence ao máx-domínio que agrega as distribuições de cauda pesada na família GDP. Já o Box-plot, exibido na Figura 24 (b) dá indícios de que a distribuição é assimétrica.

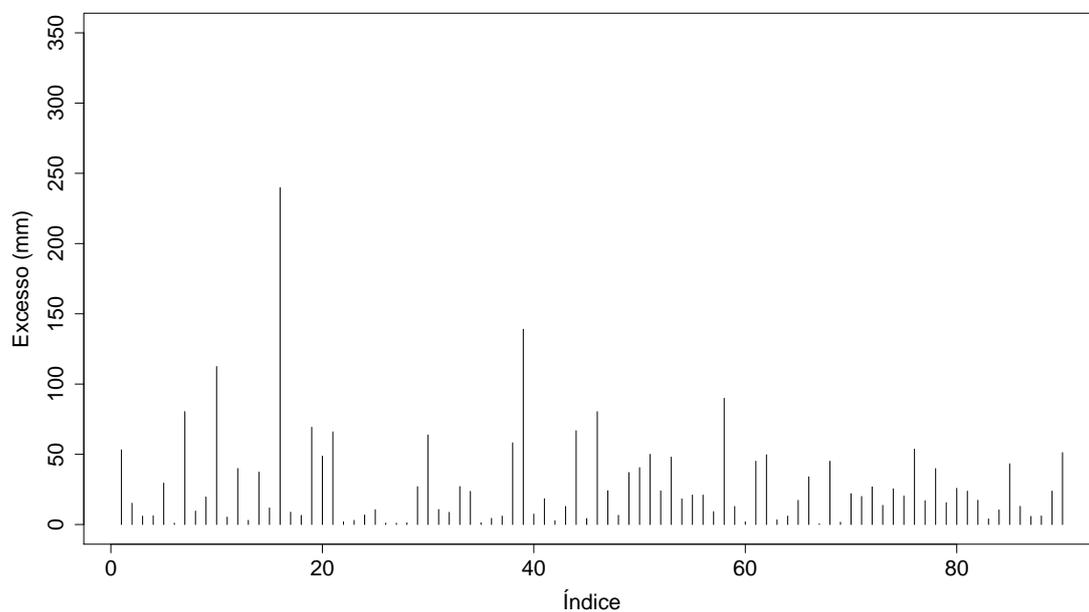
Em seguida, observemos que os gráficos das funções de autocorrelação amostral e parcial, exibidos nas Figuras 25 (a) e 25 (b), respectivamente, indicam que os excessos de precipitações pluviométricas registradas na estação Recife-Curado (1962:2019) possuem uma leve correlação entre si, revelando uma dependência linear não significativa. Sendo assim, admitiremos a hipótese de independência entre os mesmos.

Por fim, notemos que o QQ-plot exponencial ilustrado na Figura 26 ratifica que a distribuição subjacente aos dados possui cauda pesada dentre as que integram a família GDP, pois

Figura 23 – Precipitações pluviométricas diárias registradas na estação Recife-Curado (1962:2019) em (a) e os respectivos excessos além do limiar $u = 96$ mm em (b).



(a)

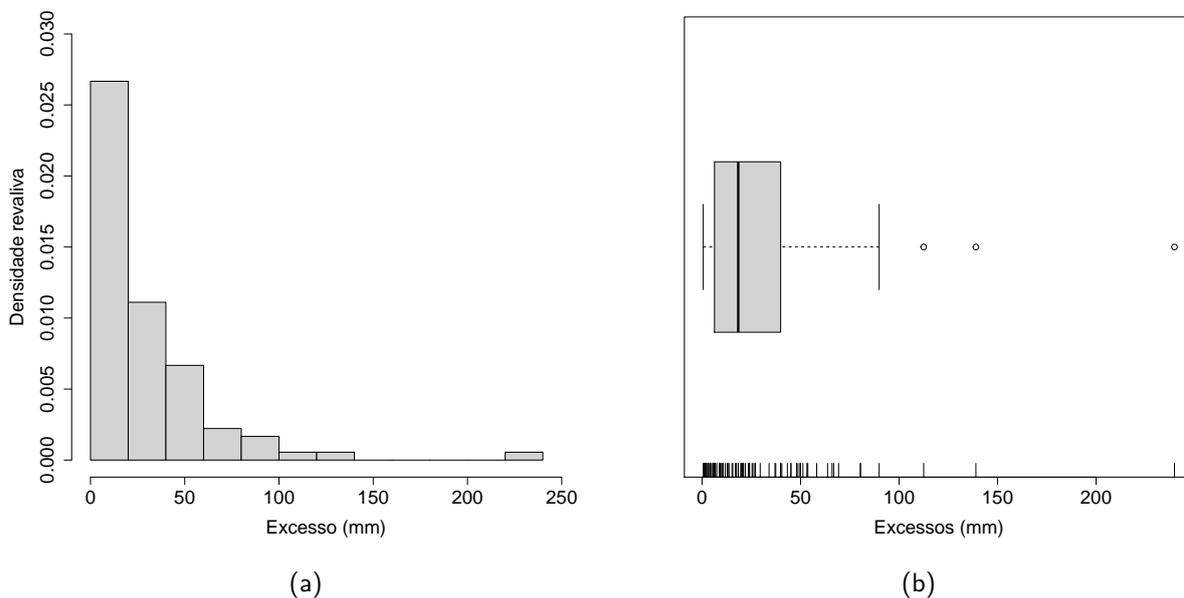


(b)

Fonte: elaborada pelo autor (2022)

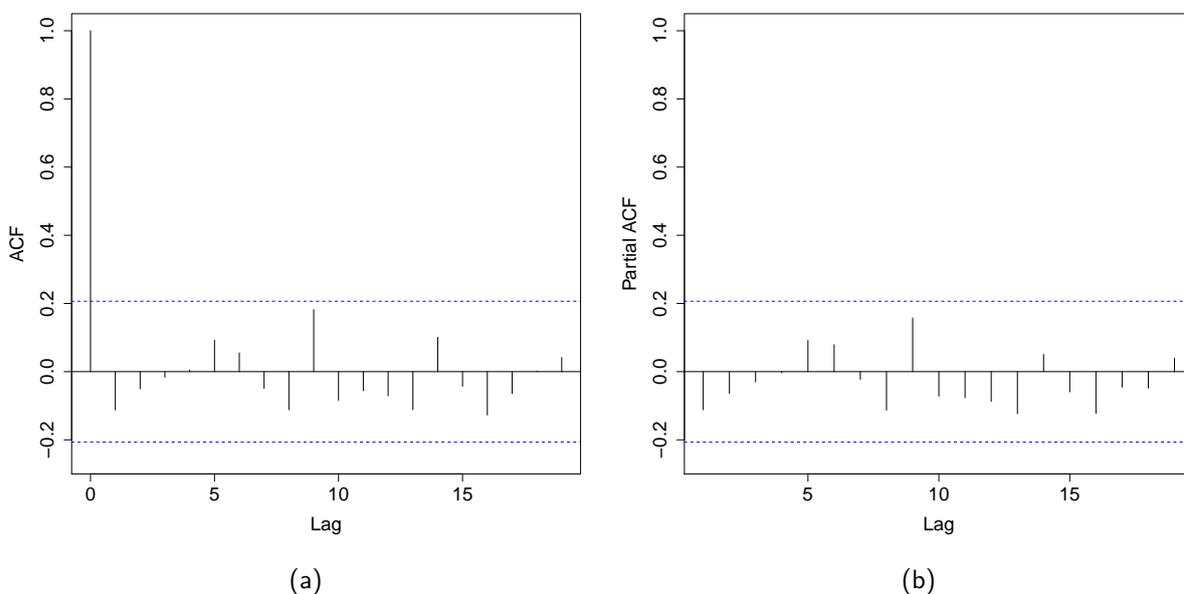
a nuvem de pontos possui padrão similar aos pontos de uma função convexa, afastando a possibilidade de modelagem por meio da distribuição exponencial (caso em que $\gamma = 0$).

Figura 24 – Histograma e Box-plot dos excessos, acima do limiar $u = 96$ mm, das precipitações pluviométricas na estação Recife-Curado (1962:2019).



Fonte: elaborada pelo autor (2022)

Figura 25 – Gráficos das funções de autocorrelação amostral em (a) e parcial em (b) dos excessos, acima do limiar $u = 96$ mm, das precipitações pluviométricas na estação Recife-Curado (1962:2019).

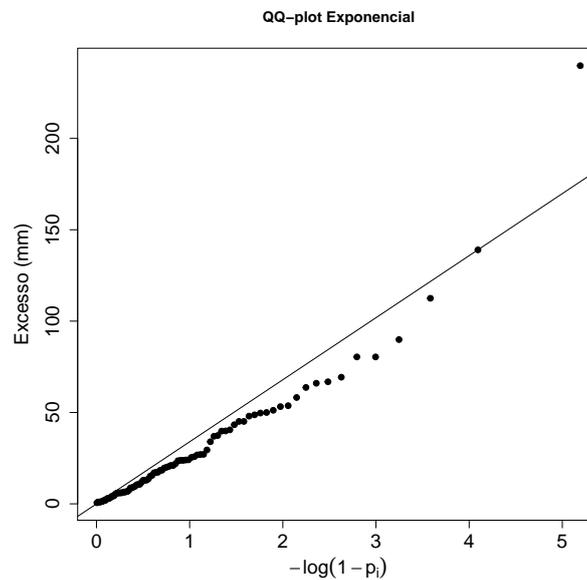


Fonte: elaborada pelo autor (2022)

5.4.3 Estimação por meio do modelo GPD

Sob o ajuste GPD, as estimativas de MV para os parâmetros de escala (σ) e forma (γ) são exibidas na Tabela 29. Nota-se que o sinal da estimativa do parâmetro γ é positivo. Logo,

Figura 26 – QQ-plot exponencial dos excessos, acima do limiar $u = 96$ mm, das precipitações pluviométricas na estação Recife-Curado (1962:2019).



Fonte: elaborada pelo autor (2022)

Tabela 28 – Estatísticas descritivas para os excessos de precipitação pluviométrica (mm) acima do limiar $u = 96$ mm, na estação Recife-Curado (1962:2019).

n	Mínimo	Q_1	Mediana	Média	Q_3	Máximo	Desvio	Assimetria	Curtose
90	0.60	6.35	18.35	28.27	39.88	239.80	34.53	3.24	18.01

Fonte: elaborada pelo autor (2022)

somos levados a classificar a distribuição dos excessos de precipitação pluviométrica na estação Recife-Curado como sendo detentora de cauda pesada, compondo o máx-domínio de atração Fréchet. Porém, a estimativa do desvio padrão do EMV do parâmetro γ é relativamente grande quando comparada com a estimativa do próprio parâmetro. Isso já era esperado pela nossa escolha de limiar, pois, de acordo com (MENDES, 2004), a escolha de um valor muito alto para o *threshold* implicará em um número pequeno de observações da cauda direita, podendo ocasionar maior variabilidade dos estimadores; por outro lado, se o limiar não for suficientemente alto, as suposições teóricas necessárias para o uso da GPD como distribuição limite dos excessos não estão satisfeitas.

Passemos a estimação de quantis elevados para os excedentes. No caso da distribuição GPD, os mesmos são obtidos por $q_p = \mu + \sigma \frac{((1-p)^{-\gamma}-1)}{\gamma}$ para $0 < p < 1$. Ao usarmos a propriedade de invariância do EMV, obtemos as estimativas expressas na Tabela 30.

A seguir, obtemos estimativas para os níveis de retorno. O nível de retorno T -anos, denotado por $U(T)$, é o quantil correspondente à probabilidade $1 - \frac{1}{T}$. Aqui, ele pode ser

Tabela 29 – Estimativas de MV dos parâmetros σ e γ do modelo GPD para o excesso acima do limiar $u = 96$, acompanhadas do respectivo erro padrão.

	$\hat{\sigma}$	$\hat{\gamma}$
emv	23.758	0.1591
Estimativas		
se	3.7669	0.1196

Fonte: elaborada pelo autor (2022)

Tabela 30 – Estimativas de MV dos quantis elevados no modelo GPD, excedentes acima do limiar $u = 96$.

	$\hat{q}_{0.95}$	$\hat{q}_{0.99}$	$\hat{q}_{0.999}$	$\hat{q}_{0.9999}$
Estimativas	187.1761	257.3524	394.7833	592.9991

Fonte: elaborada pelo autor (2022)

interpretado como o excedente de precipitação pluviométrica (mm) que é excedido, em média, uma vez a cada T anos. Mais uma vez, ressaltamos que a função `is.fixedfevd()`, quando aplicada ao modelo ajustado, sugere estacionariedade do mesmo. A Tabela 31 exibe as estimativas de MV para os níveis de retorno para 10, 20, 50, 100 e 200 anos. De onde se infere que aproximadamente 187 mm é o nível de retorno que é excedido, em média uma vez, a cada período de 20 anos. Observemos que este nível quase que foi superado no ano de 2000, mais precisamente no 01/08/2000 foi registrada uma precipitação pluviométrica diária igual 185.9 mm. Se o período é alongado para um século, a estimativa para o nível de retorno é aproximadamente 257 mm.

Tabela 31 – Estimativas de MV dos níveis de retorno, excedentes, no modelo GPD.

	$\hat{U}(10)$	$\hat{U}(20)$	$\hat{U}(50)$	$\hat{U}(100)$	$\hat{U}(200)$
Estimativas	162.0659	187.1761	224.9167	257.3524	293.5686

Fonte: elaborada pelo autor (2022)

Agora, calculemos as estimativas para a probabilidade de excedência de um determinado nível e e para o correspondente período de retorno, denotado por $\mathbb{I}EL(e)$ e definido como o número médio de anos de espera até que o nível e seja ultrapassado. Na Tabela 32 encontram-se as estimativas de MV das probabilidades de excedência associadas aos quatro maiores níveis pluviométricos registrados na estação Recife-Curado no período 1962:2019, paralelamente são exibidas as estimativas dos períodos de retorno associados aos respectivos níveis. Destaquemos que a estimativa para o período de retorno correspondente ao nível pluviométrico de 235.8 mm é aproximadamente 62.5 anos, sendo 1986 o último ano em que este nível foi registrado.

Já a estimativa do período de retorno associada a maior precipitação pluviométrica registrada na estação, que foi de 335.8 mm, é aproximadamente 412 anos.

Tabela 32 – Estimativas de MV da probabilidade de excedência do nível e e do período de retorno correspondente, em anos, no modelo GPD.

Níveis	Probabilidade de excedência	Período de retorno
$e = 185.9$	0.0517	19.3422
$e = 208.5$	0.0293	34.1151
$e = 235.0$	0.0160	62.5396
$e = 335.8$	0.0024	411.7906

Fonte: elaborada pelo autor (2022)

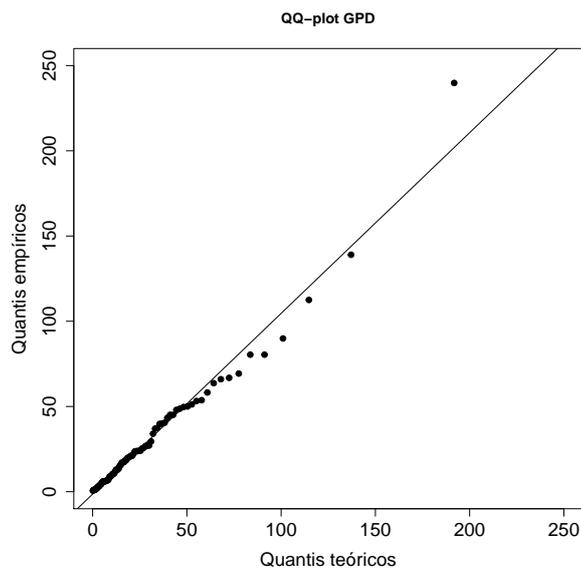
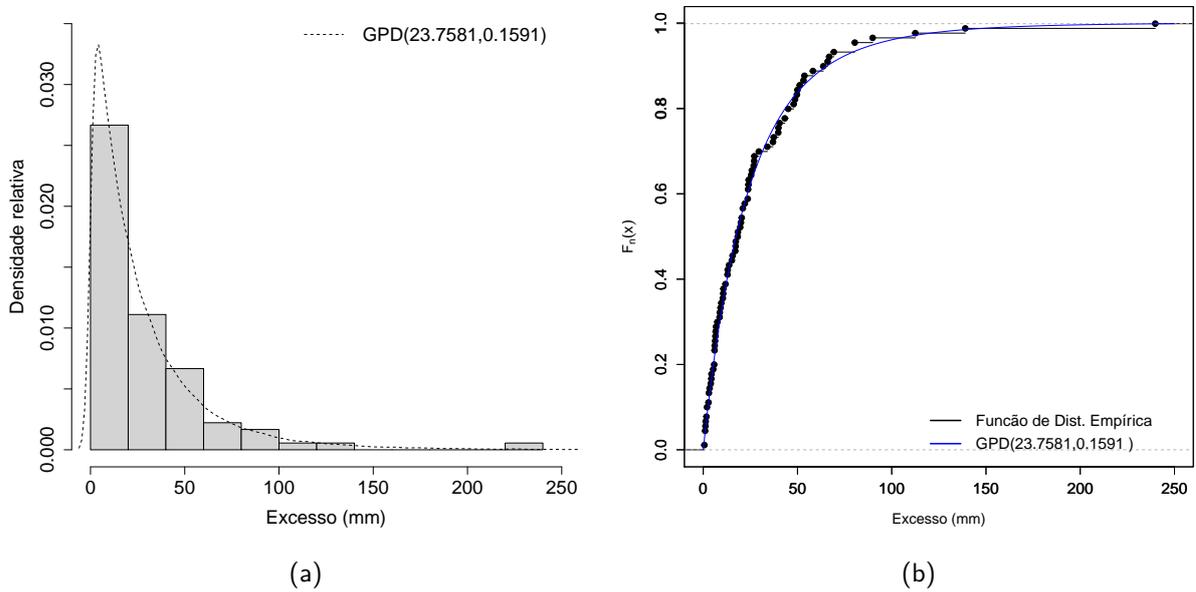
Por fim, apresentamos na Figura 27 os gráficos que refletem a qualidade do ajuste obtido por meio do modelo GPD. Particularmente, o QQ-plot-GPD, Figura 27 (c), indica que não há evidências contrárias ao ajuste dos excessos de precipitação pluviométrica, acima do limiar $u = 96$, registrados na estação Recife-Curado por meio do modelo GPD.

5.4.4 Estimação via o modelo GPD-BS

A fim de implementarmos o ajuste por meio do modelo GPD-BS, consideraremos os excedentes de precipitação pluviométrica, acima do limiar $u = 96$ mm. Dessa forma, a variável que modela tais excedentes deve, obrigatoriamente, assumir valores maiores do que u , isso nos leva a fixar o valor do parâmetro β do modelo GPD-BS em 96. Um procedimento similar foi adotado no ajuste GPD quando admitiu-se que parâmetro de locação $\mu = 96$. A Tabela 33 exhibe as estimativas de MV dos parâmetros do modelo GPD-BS. Destaque-se que o sinal da estimativa de γ é negativo, levando-nos a classificar a distribuição dos excedentes de precipitação pluviométrica como tendo cauda curta dentre aquelas que compõem a família GPD-BS. Em teoria, isso possibilita a estimação do limite superior do suporte da distribuição subjacente aos dados. Mas, dado que a estimativa de γ está próxima de zero, o valor numérico a ser estimado será, provavelmente, muito elevado, inviabilizando seu uso prático no caso sob estudo.

A Tabela 34 exhibe as estimativas de MV de alguns dos quantis elevados. As estimativas dos quantis $q_{0.95}$ e $q_{0.99}$ são superiores àquelas obtidas sob o ajuste GPD, mas relativamente próximas. Por outro lado, as estimativas dos quantis $q_{0.999}$ e $q_{0.9999}$ são inferiores e relativamente

Figura 27 – Gráficos da qualidade do ajuste por meio do modelo GPD.



Fonte: elaborada pelo autor (2022)

Tabela 33 – Estimativas de MV dos parâmetros α e γ no modelo GPD-BS, excedentes acima do limiar $u = 96$, acompanhadas do respectivo erro padrão.

	$\hat{\alpha}$	$\hat{\gamma}$
emv	0.2411	-0.0330
se	0.0346	0.0975

Fonte: elaborada pelo autor (2022)

distantes.

Tabela 34 – Estimativas de MV dos quantis elevados no modelo GPD-BS, excedentes acima do limiar $u = 96$.

	$\hat{q}_{0.95}$	$\hat{q}_{0.99}$	$\hat{q}_{0.999}$	$\hat{q}_{0.9999}$
Estimativas	188.5375	258.1959	380.7798	526.6031

Fonte: elaborada pelo autor (2022)

A Tabela 35 apresenta as estimativas dos níveis de retorno para os períodos de 10, 20, 50, 100 e 200 anos obtidas pelo modelo GPD-BS. Nota-se que elas estão muito próximas daquelas provenientes do ajuste GPD.

Tabela 35 – Estimativas de MV dos níveis de retorno no modelo GPD-BS.

	$\hat{U}(10)$	$\hat{U}(20)$	$\hat{U}(50)$	$\hat{U}(100)$	$\hat{U}(200)$
Estimativas	162.854	188.5375	226.4988	258.1959	292.3802

Fonte: elaborada pelo autor (2022)

Agora, notamos que as estimativas das probabilidades de excedência associadas aos 4 maiores níveis de precipitação pluviométrica, provenientes do ajuste GPD-BS e exibidas na Tabela 36, são diferentes daquelas advindas do modelo GPD. Em particular, as probabilidades de excedência associadas a maior observação amostral, 335.8 mm, estão relativamente afastadas uma da outra, fazendo com que as emv para os períodos de retorno sejam distintas. De modo que a diferença absoluta entre as emv dos períodos de retorno associado ao nível 335.8 mm é de aproximadamente 41.5 anos.

Tabela 36 – Estimativas de MV da probabilidade de excedência do nível e e do período de retorno correspondente, em anos, no modelo GPD-BS.

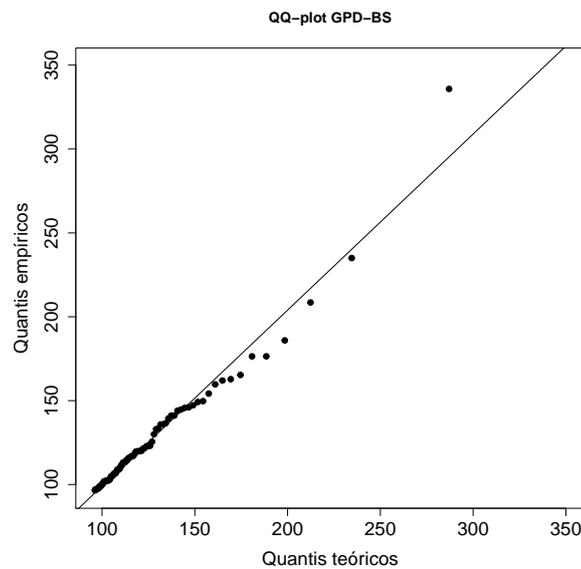
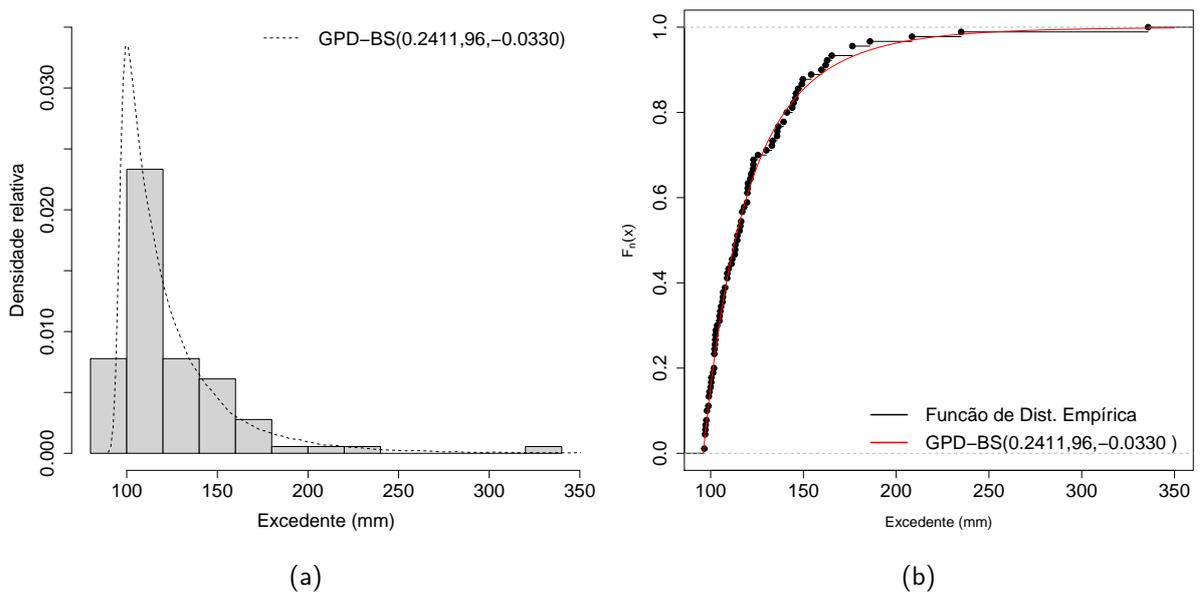
Níveis	Probabilidade de excedência	Período de retorno
$e = 185.9$	0.0535	18.6843
$e = 208.5$	0.0305	32.8246
$e = 235.0$	0.0165	60.5498
$e = 335.8$	0.0022	453.2847

Fonte: elaborada pelo autor (2022)

Neste ponto, destaquemos que, para o caso sob estudo, o ajuste por meio do modelo GPD-BS possibilita a estimação do limite superior do suporte da distribuição; que é igual $\hat{x}^* = 5312.403$ mm. Porém, nota-se claramente que este valor é muito elevado, inviabilizando seu uso prático.

Por fim, os gráficos da Figura 28 atestam que o ajuste por meio do modelo GPD-BS é tão satisfatório quando aquele em que se utilizou a distribuição GPD. Em particular, o QQ-plot-GPD-BS dos excedentes, Figura 28 (c), sugere que não há evidências contrárias ao ajuste por meio da GPD-BS, pois o mesmo acomoda satisfatoriamente a nuvem de pontos em torno da reta bissetriz.

Figura 28 – Gráficos da qualidade do ajuste por meio do modelo GPD-BS.



(c)

Fonte: elaborada pelo autor (2022)

6 CONCLUSÕES E TRABALHOS FUTUROS

Ao desenvolvermos o presente trabalho obtemos ferramentas que viabilizam a implementação de uma análise de influência local sobre o modelo de regressão EVBS, pois derivamos expressões que possibilitam o cálculo das curvaturas normal e conforme para três diferentes esquemas de perturbação: ponderação de casos, perturbação na resposta e perturbação em uma variável explicativa, permitindo a identificação criteriosa de observações influentes. Paralelamente, obtemos uma condição necessária e suficiente para existência dos dois primeiros momentos da distribuição EVBS e deduzimos algumas características relativas à monotonicidade e à existência de máximos locais para a função densidade de probabilidade da distribuição log-EVBS. Além disso, abordamos o problema de estimação no modelo de regressão EVBS e acentuamos a problemática decorrente do fato de estarmos diante de um caso não-regular. Obtivemos, por meio de simulações de Monte Carlo, suporte numérico que corrobora com a hipótese de consistência assintótica do EMV dos parâmetros do modelo de regressão EVBS.

Nossas aplicações se concentraram em dados reais registrados nas estações meteorológicas Recife-Curado-PE e Itajaí-SC. Em uma delas nosso interesse incidiu sobre as temperaturas máximas anuais, na outra foi a velocidade máxima mensal de ventos. Em ambas, confirmamos a eficácia da técnica de influência local em detectar observações influentes. Ainda nesse contexto, uma terceira aplicação possui grande relevo social, pois investigamos os níveis da precipitação pluviométrica máxima anual registrados na estação Recife-Curado, utilizando os modelos GEV e EVBS, combinados com o método de Gumbel, onde obtivemos estimativas de máxima verossimilhança para os parâmetros tais como: nível de retorno e período de retorno. Agregasse a isso, a obtenção de uma estimativa, proveniente do ajuste EVBS, para o limite superior do suporte da distribuição subjacente dos dados, ou seja, uma estimativa para o supremo dos níveis de precipitação máxima anual na estação Recife-Curado.

Concluimos o desenvolvimento do trabalho com a propositura de uma nova família de distribuições BS derivada da GPD para modelagem de dados extremos obtidos segundo a metodologia POT. Deduzimos algumas propriedades desta distribuição e tratamos de estimação por máxima verossimilhança. O potencial desta nova família na modelagem de dados extremos foi evidenciado a partir de uma aplicação na qual modelamos os excedentes anuais de precipitação pluviométrica, acima de um limiar elevado, registrados na estação Recife-Curado, utilizando a metodologia POT, resultando em um ajuste tão satisfatório quanto aquele em

que se utilizou o modelo GPD.

Por fim, indicaremos alguns tópicos que podem ser aprofundados ou desenvolvidos em pesquisas futuras. No que tange à estimação dos parâmetros do modelo EVBS e do seu correspondente modelo de regressão, deve-se investigar minuciosamente as condições de regularidades que são violadas pela função de log-verossimilhança e usar os resultados obtidos por (DOMBRY, 2015), (BUCHER; SEGERS, 2017) (DOMBRY; FERREIRA, 2019) e (ZHANG; SHABY, 2020) para derivar condições suficientes que assegurem a consistência e a normalidade assintótica do EMV. Em relação à análise de influência local sobre o modelo de regressão EVBS, o pressuposto de suporte comum, fixado por (AMARI, 1985) para o desenvolvimento do arcabouço geométrico-diferencial, limitou nossas ações ao caso regular. Neste sentido, é frutífero investigarmos quais as consequências do relaxamento desse pressuposto na estrutura geométrica e estabelecermos o alicerce para lidarmos com situações em que tal condição não seja atendida.

Acerca da aplicação que trata dos níveis de precipitação pluviométrica registrados na estação Recife-Curado, é natural investigar o comportamento da precipitação pluviométrica máxima mensal, sendo necessário, talvez, incluir uma estrutura que admita, por exemplo, a hipótese de dependência linear entre os dados. Dada a natureza dessa variável, uma outra alternativa consiste em empregar a metodologia na qual se volta a atenção para os conglomerados de excessos, usando os resultados obtidos por (ANDERSON, 1994), na tentativa de compreender as cheias ocorridas em meados da década de 70 no Recife-PE, possivelmente causadas por soma de excessos de precipitação.

Em relação ao modelo GPD-BS, há algumas direções que podem ser examinadas. De início, provavelmente as expressões obtidas para os valores esperados de funções da variável GPD-BS possibilitem a construção de rotinas que nos levem aos estimadores de L-momentos dos parâmetros. Em seguida, cabe uma investigação sobre a função de distribuição da GPD-BS acerca das condições citadas por (MENDES, 2004), no tocante à monotonicidade côncava e a dominação, ao comentar sobre a forma funcional da distribuição assintótica do excesso agregado.

REFERÊNCIAS

- ABRAMOWITZ, M.; STEGUN, I. A. *Handbook of Mathematical Functions*. [S.l.]: Dover Press New York, 1970.
- ACHCAR, J. A. Inference for the birnbaum-saunders distribution fatigue life model using bayesian methods. *Comput. Stat. Data Analysis*, v. 15, n. 4, p. 367–380, 1993.
- ACHCAR, J. A.; MOALA, F. A. Use of mcmc methods to obtain bayesian inferences for the birnbaum-saunders distribution in the presence of censored data and covariates. *Adv. Appl. Stat.*, v. 17, p. 1–27, 2010.
- AMARI, S. *Differential-Geometrical Methods in Statistics*. [S.l.]: Springer-Verlag Berlin, 1985.
- AMARI, S.; NAGOAKA, H. *Methods of Information Geometry*. [S.l.]: American Mathematical Society, 2000.
- ANDERSON, C. W. The aggregate excess measure of severity of extremes events. *Journal of Research of the National Institute of Standards and Technology*, v. 99, p. 555–561, 1994.
- ASTM-INTERNATIONAL, . Standard terminology relatın to fatigue and fracture testing. *ASTM International, West Conshohocken, US*. [http://: www.astm.org/Sity_tandart/E1823.](http://www.astm.org/Sity_tandart/E1823.), 2013.
- ATHAYDE, E.; AZEVEDO, A.; BARROS, M.; LEIVA, V. Failure rate of birnbaum-saunders distributions: Shape, change-point, estimation and robustness. *Brazilian Journal of Probability and Statistics*, v. 33, n. 2, p. 301–328, 2019.
- BALAKRISHNAN, N.; KUNDU, D. Birnbaum-saunders distribution: A review of models, analysis and applications. *Appl Stochastic Models Bus Ind*, John Willey & Sons , Ltd, v. 35, p. 4–49, 2019.
- BALAKRISHNAN, N.; ZHU, X. Inference for the birnbaum-saunders lifetime regression model with applications. *Communications in Statistics - simulation and computation*, v. 44, n. 8, p. 2073–2100, 2015.
- BALAKRISHNAN, N.; ZHU, X. Inference for the bivariate birnbaum-saunders lifetime regression model and associated inference. *Metrika*, v. 78, p. 853–872, 2015.
- BALKEMA, A. A.; HANN, L. Resid life time at great age. *Ann. Probability*, v. 2, p. 792–804, 1974.
- BARROS, M.; PAULA, G. A.; LEIVA, V. A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis*, v. 14, p. 316–332, 2008.
- BARROS, M.; PAULA, G. A.; LEIVA, V. An r implementation for generalized birnbaum-saunders distributions. *Computational Statistics and Data Analysis*, v. 53, p. 1511–1028, 2009.
- BEBBINGTON, M.; C-D, L.; ZITIKIS, A. A proof of the shape of the birnbaum-saunders hazard rate function. *Math. Sci.*, v. 33, p. 49–56, 2008.

- BHATTACHARYYA, G. K.; FRIES, A. Fatigue failure models birnbaum-saunders vs. inverse gaussian. *IEEE Trans Reliab*, v. 31, n. 5, p. 439–440, 1982.
- BIRNBAUM, Z. W.; SAUNDERS, S. C. A new family of the distributions. *J. Appl. Prob.*, v. 6, p. 319–327, 1969a.
- BIRNBAUM, Z. W.; SAUNDERS, S. C. Estimation for a family of life distributions with applications to fatigue. *J. Appl. Prob.*, v. 6, p. 328–347, 1969b.
- BOURGUINGON, M.; LEÃO, J.; LEIVA, V.; SANTOS-NETO, M. The transmuted birnbaum-saunders distribution. *REVSTAT- Statistical Journal*, v. 15, n. 4, p. 601–628, 2017.
- BUCHER, A.; SEGERS, J. On the maximum likelihood estimator for the generalized extreme-value distribution. *Extremes*, v. 20, p. 839–872, 2017.
- CAM, L. L. *Asymptotic Methods in Statistical Decision Theory*. [S.l.]: Springer Series in Statistics- Springer Verlag, 1986.
- CANCHO, V. G.; ORTEGA, E. M. M.; PAULA, G. A. Os estimation and influence diagnostics for log-birnbaum-saunders student-t regression models: full bayesian analysis. *Journal of Statistical Planning and Inference*, v. 140, n. 9, p. 2486–2496, 2010.
- CARMO, M. P. D. *Geometria riemanniana*. [S.l.]: Projeto Euclides - IMPA, 1998.
- CARMO, M. P. D. *Geometria Diferencial de Curvas e Superfícies*. [S.l.]: Sociedade Brasileira de Matemática - SBM, 2010.
- CASTILLO, N.; GÓMEZ, H. W. O.; BOLFARINE, H. Epsilon birnbaum-saunders distribution family: properties and inference. *Statistical Papers*, v. 52, n. 4, p. 871–883, 2011.
- CHANG, D. S.; TANG, L. C. Reability bounds and critical time for the birnbaum-saunders distribution. *IEEE Trans. Device Mater. Reliab.*, v. 24, p. 464–469, 1993.
- CHANG, D. S.; TANG, L. C. Random number generator for the birnbaum-saunders distribution. *Computers and Industrial Engineering*, v. 27, n. 1-4, p. 345–348, 1994.
- COOK, R. D. Assessment of local influence. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 133–169, 1986.
- CRAMÉR, H. *Mathematical Methods of Statistics*. [S.l.]: Princeton University Press, 1946.
- CRIBARI-NETO, F.; FONSECA, R. V. A new log-linear bimodal birnbaum-saunders regression model with applications to survival data. *Brazilian Journal of Probability and Statistics*, v. 33, n. 2, p. 329–355, 2019.
- DEKKERS, A. L.; EINMAHL, J. H.; HAAN, L. D. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, JSTOR, p. 1833–1855, 1989.
- DESMOND, A. Stochastic models of failure in random environments. *The Canadian Journal of Statistics*, v. 13, n. 2, p. 171–183, 1985.
- DESMOND, A. On the relationship between two fatigue-life models. *IEEE Trans Reliab*, v. 35, n. 2, p. 167–169, 1986.

- DESMOND, A.; RODRIGUEZ-YAM, G. A.; LU, X. Estimation of parameters for a birnbaum-saunders regression model with censored data. *Journal of Statistical Computation and Simulation*, v. 78, n. 11, p. 983–997, 2008.
- DESOUSA, M.; SAULO, H.; LEIVA, V.; SCALCO, P. On a tobit-birnbaum-saunders model with an application to medical data. *Journal Appl. Stat.*, v. 45, n. 5, p. 932–955, 2017.
- DOMBRY, C. Existence and consistency of the maximum likelihood estimators for the extreme value index within the block maxima framework. *Bernoulli*, v. 21, 2015.
- DOMBRY, C.; FERREIRA, A. Maximum likelihood estimators based on the block maxima method. *Bernoulli*, v. 25, 2019.
- DRESS H., e. a. On maximum likelihood estimation of the extreme value index. *Annals of Applied Probability*, v. 14, n. 3, p. 1179–1201, 2004.
- DUPUIS, D. J.; MILLS, J. E. Robust estimation of the birnbaum-saunders distribution. *IEEE Transactions on Reliability*, v. 47, n. 1, p. 88–95, 1998.
- DÍAZ-GARCÍA, J. A.; LEIVA, V. A new family of life distributions based on elliptically contoured distributions. *J. Statistic. Plan. Infer.*, v. 128, p. 445–457, 2005.
- EMBRECHTS, R.; KLUPPELBERG, C.; MIKOSCH, T. *Modelling extremal events for insurance and finance*. [S.l.]: Springer, 1997.
- ENGELHARDT, M.; BAIN, L. J.; WRIGHT, F. T. Inferences on the parameters of the birnbaum-saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics*, v. 23, n. 3, p. 251–256, 1981.
- FARIAS, R. B. A.; LEMONTE, A. J. Bayesian inference for the birnbaum-saunders nonlinear regression model. *Statistical Methods & Applications*, v. 20, p. 423–438, 2011.
- FERREIRA, M.; GOMES, M. I.; LEIVA, V. On an extreme value version of the birnbaum-saunders distribution. *REVISTAT Statistical Journal*, v. 10, n. 2, p. 181–210, 2012.
- FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In: CAMBRIDGE UNIVERSITY PRESS. *Mathematical Proceedings of the Cambridge Philosophical Society*. [S.l.], 1928. v. 24, n. 2, p. 180–190.
- FLETCHER, H. A. A verification of the theory of brownian movements and a direct determination of the value of n_e for gases ionization. *Phys Rev*, v. 33, p. 81–110, 1911.
- FREUDENTHAL, A.; M., S. Structural safety under conditions of ultimate load failure and fatigue. WADD-TR-61-77. Wright Air Development Division, 1961.
- FROM, S. G.; LI, L. Estimation of the parameters of the birnbaum-saunders distribution. *Commun. Stat. Theory Methods*, v. 35, n. 12, p. 2157–2169, 2006.
- FUNG, W. K.; KWAN, C. A note on local influence based on normal curvature. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 59, n. 4, p. 839–843, 1997.
- GALEA, M.; LEIVA, V.; A, P. G. Influence diagnostic in log-birnbaum-saunders regression models. *Journal of Applied Statistics*, v. 13, n. 9, p. 1049–1064, 2004.

- GARCIA-PAPANI, F.; URIBE-OPAZO, M. A.; V., L.; AUKROYD, R. G. Birnbaum–saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stochastic Environmental Research and Risk Assessment*, v. 31, n. 105–124, 2017.
- GLASER, R. Bathub and related failure rate characterization. *J. Amer. Statist. Assoc.*, v. 75, p. 667–672, 1980.
- GNEDENKO, B. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, JSTOR, p. 423–453, 1943.
- GOMES, M.; GUILLOU, A. Extreme value theory and statistical of univariate extremes: A review. *International Statistical Review*, v. 83, n. 2, p. 263–292, 2015.
- GOMES, M. I.; FERREIRA, M.; LEIVA, V. The extreme value birnbaum-saunders model, its moments and an application in biometry. *Biometrical Letters*, v. 49, n. 2, p. 81–94, 2012.
- GUIRAUD, P.; LEIVA, V.; FIERRO, R. A non-central version of the birnbaum-saunders distribution for reliability analysis. *IEEE Trans. Rel.*, v. 58, p. 152–160, 2009.
- GÓMEZ, H. W.; OLIVARES, J.; BOLFARINE, H. An extension of the generalized birnbaum-saunders distribution. *Statistic Probab. Letters*, v. 79, p. 331–338, 2009.
- HAAN, L. de; FERREIRA, A. *Extreme Value Theory: An introduction*. [S.l.]: Springer, 2006.
- HAAN, L. F. M. On regular variation and its application to the weak convergence of sample extremes. 1970.
- HILL, B. M. A simple general approach to inference about the tail of a distribution. *Ann. Statistics*, v. 3, p. 1163–1174, 1975.
- HOSKING, J. R. M.; WALIIS, J. R.; WOOD, E. F. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Tecnometrics*, v. 27, p. 251–261, 1985.
- KONSTANTINOWSKY, D. Elektrische ladungen und brownscher bewegung schrkleiner metallteilchen gasen. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften*, v. 123, p. 1697–1752, 1914.
- KUNDU, D.; BALAKRISHNAN, N.; JAMALIZADEH, A. Bivariate birnbaum-saunders distribution and associated inference. *J. Multivar Analysis*, v. 102, n. 1, p. 113–125, 2010.
- KUNDU, D.; BALAKRISHNAN, N.; JAMALIZADEH, A. Generalized multivariate birnbaum-saunders distributions and related inferential issues. *J. Multivar Analysis*, v. 116, p. 230–244, 2013.
- KUNDU, D.; KANNAN, N.; BALAKRISHNAN, N. On the hazard function of birnbaum-saunders distibution and associated inference. *Computational Statistics & Data Analysis*, v. 52, p. 2692–2702, 2008.
- LAWLESS, J. F. *Statistical Models and Methods for Lifetime Data, 2nd*. [S.l.]: John Wiley & Sons, 2003.
- LEE, J. M. *Introduction to smooth manifolds*. [S.l.]: Springer, 2013.

- LEHMANN, E. L.; CASELLA, G. *Theory of Point Estimations, 2nd*. [S.I.]: Springer-Verlag, 1998.
- LEIVA, V. *The Birnbaum-Saunders Distribution*. [S.I.]: The Netherlands: Elsevier-Academic Press, 2016.
- LEIVA, V.; BARROS M. GILBERTO, A. P.; GALEA, M. Influence diagnostic in log-birnbaum-saunders regression models with censored data. *Computational Statistics & Data Analysis*, v. 51, p. 5694–4707, 2007.
- LEIVA, V.; FERREIRA, M.; GOMES, M. I.; LILLO, C. Extreme value birnbaum-saunders regression models applied to environmental data. *Stochastic environmental research and risk assessment*, Springer, v. 30, n. 3, p. 1045–1058, 2016.
- LEIVA, V.; HERNÁNDEZ, H.; RIQUELME, M. A new package for the birnbaum-saunders distribution. *Teh newsletter of the R project*, v. 6, p. 35–40, 2006.
- LEIVA, V.; SANHUEZA, A.; ANGULO, J. M. A length-biased version of the birnbaum-saunders distribution with application in water quality. *Stoch. Env. Res. Risk A*, v. 23, n. 3, p. 299–307, 2009.
- LEIVA, V.; SANHUEZA, A.; SEN, P. K.; PAULA, G. A. Random number generators for the generalized birnbaum-saunders distribution. *Journal Stat. Comput. and Simulation*, v. 78, n. 11, p. 299–307, 2008.
- LEIVA, V.; SANTOS-NETO, M.; CYSNEIROS, F. J. A.; BARROS, M. Birnbaum-saunders statistical modelling: a new approach. *Statistical Modelling*, v. 14, n. 1, p. 21–48, 2014.
- LEMONTE, A. J. Covariance matrix formula for birnbaum-saunders regression models. *Journal of Statistical Computational and Simulation*, v. 81, n. 7, p. 899–908, 2011.
- LEMONTE, A. J. A log-birnbaum-saunders regression model with asymmetric errors. *Journal of Statistical Computational and Simulation*, v. 82, n. 12, p. 1775–1787, 2012.
- LEMONTE, A. J. Multivariate birnbaum-saunders regression model. *Journal of Statistical Computational and Simulation*, v. 83, n. 12, p. 2244–2257, 2013.
- LEMONTE, A. J. A new extended birnbaum-saunders regression model for lifetime modeling. *Computational Statistics & Data Analysis*, v. 64, p. 34–50, 2013.
- LEMONTE, A. J.; CORDEIRO, G. M. Birnbaum-saunders nonlinear regression models. *Computational. Statistics and Data Analysis*, v. 53, n. 12, p. 4441–4452, 2009.
- LEMONTE, A. J.; CORDEIRO, G. M. Asymptotic skewness in birnbaum-saunders nonlinear regression models. *Statistics & Probability Letters*, v. 80, n. 9-10, p. 892–898, 2010.
- LEMONTE, A. J.; CORDEIRO, G. M.; MORENO, G. Bartlett corrections in birnbaum-saunders nonlinear regression models. *Journal of Statistical Computational and Simulation*, v. 82, n. 6, p. 927–935, 2012.
- LEMONTE, A. J.; CRIBARI-NETO, F.; VASCONCELOS, K. L. P. Improved statistical inference for the two-parameter birnbaum-saunders distribution. *Comput. Stat. Data Analysis*, v. 51, n. 9, p. 4656–4691, 2007.

- LEMONTE, A. J.; FERRARI, S. L. P. Signed likelihood ratio test in the birnbaum-saunders regression model. *Journal of Statistical Planning and Inference*, v. 141, n. 2, p. 1031–1040, 2011.
- LEMONTE, A. J.; PATRIOTA, A. G. Influence diagnostic in birnbaum-saunders regression models. *Journal of Applied Statistics*, v. 38, n. 5, p. 871–884, 2011.
- LI, A.-P.; CHEN, Z.-X.; F-C., X. Diagnostic analysis for heterogeneous log-birnbaum-saunders regression models. *Statistics & Probability Letters*, v. 82, n. 9, p. 1690–1698, 2012.
- LI, A.-P.; F-C., X. Diagnostic for a class of survival regression models with heavy-tailed errors. *Computational Statistics & Data Analysis*, v. 56, n. 12, p. 4204–4214, 2012.
- LILLO, C.; LEIVA, V.; NICOLIS, O.; AYKROYD, R. L-moments of the birnbaum-saunders distribution and its extreme value version: estimation, goodness of fit and application to earthquake data. *Journal of Applied Statistics*, v. 45, n. 2, p. 187–209, 2018.
- LIMA, E. L. *Variedades Diferenciáveis*. [S.l.]: Publicações Matemáticas- IMPA, 2007.
- MARCHANT, C.; LEIVA, V.; CYSNEIROS, F. J. A. A multivariate log-linear for birnbaum-saunders distribution. *IEEE Transactions on Reliability*, v. 65, n. 2, p. 816–827, 2015.
- MARCHANT, C.; LEIVA, V.; CYSNEIROS, F. J. A. Diagnostic in multivariate generalized birnbaum-saunders regression models. *Journal of Applied Statistics*, v. 43, n. 15, p. 2829–2849, 2016.
- MARTÍNEZ-FLÓREZ, G.; BOLFARINE, H.; GÓMEZ, H. W. The log-linear birnbaum-saunders power model. *Methodology and Computing in Applied Probability*, v. 19, p. 913–933, 2017.
- MARTÍNEZ, S.; GIRALDO, R.; LEIVA, V. Birnbaum-saunders functional regression models for spatial data. *Stochastic Environmental Research and Risk Assessment*, v. 33, p. 1765–1780, 2019.
- MENDES, B. V. d. M. *Introdução à Análise de Eventos Extremos*. [S.l.]: E-papers Serviços Editoriais, 2004.
- MURRAY, M. K.; RICE, J. W. *Differential Geometry and Statistics*. [S.l.]: Chapman & Hall, 1993.
- NG, H. K. T.; KUNDU, D.; BALAKRISHNAN, N. Modified moment estimation for the two-parameter birnbaum-saunders distribution. *Comput. Stat. Data Analysis*, v. 43, n. 3, p. 283–298, 2003.
- NG, H. K. T.; KUNDU, D.; BALAKRISHNAN, N. Point and interval estimation for the birnbaum-saunders distribution based on type-ii censored samples. *Comput. Stat. Data Analysis*, v. 50, n. 11, p. 3222–3242, 2006.
- ORTEGA, E. M.; CORDEIRO, G. M.; LEMONTE, A. J. A log-linear regression model for the beta-birnbaum-saunders distribution with censored data. *Computational Statistics & Data Analysis*, v. 56, n. 3, p. 698–718, 2013.
- OWEN, W. J. A new three-parameters extension to the birnbaum-saunders distribution. *IEEE Transactions on Reliability*, v. 55, n. 3, p. 475–479, 2006.

- OWEN, W. J.; PADGETT, W. J. A birnbaum-saunders accelerated life model. *IEEE Transactions on Reliability*, v. 49, n. 2, p. 224–229, 2000.
- PADGETT, W. J. On bayes estimation of reliability for the birnbaum-saunders fatigue life model. *IEEE Trans. Reliab.*, v. 31, n. 5, p. 436–438, 1982.
- PAULA, G. A.; LEIVA, V.; BARROS, M.; LUI, S. Robust statistical modeling using the birnbaum-saunders-t distribution applied to insure. *Applied Stochastic Models in Business and Industry*, v. 28, n. 1, p. 16–34, 2012.
- PICKANDS, J. Statistical inference using extreme order statistics. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 3, n. 1, p. 119–131, 1975.
- POON, W.-Y.; POON, Y. S. Conformal normal curvature and assessment of local influence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 61, n. 1, p. 51–61, 1999.
- POON, Y.-S.; POON, W.-Y. *Application of elementary differential geometry to influence analysis*. [S.l.]: Higher education Press-, 2010.
- RESNICK, S. I. *Extreme Values, Regular Variation and Point Processes*. [S.l.]: Srpinger, 1987.
- RIECK, J. R. *Statistical Analysis for the Birnbaum-Saunders Fatigue Distribution*. Tese (Doutorado) — Graduate School of Clemson University, Dept. of Mathematical Sciences, 1989.
- RIECK, J. R. Parametric estimation for the birnbaum-saunders distribution based on symmetrically censored samples. *Comput. Stat. Theory Methods*, v. 24, n. 7, p. 1721–1736, 1995.
- RIECK, J. R. A moment-generating function with application to the birnbaum-saunders distribution. *Comput. Stat. Theory Methods*, v. 28, n. 9, p. 2213–2222, 1999.
- RIECK, J. R. A comparison of two random number generators for the birnbaum-saunders distribution. *Comput. Stat. Theory Methods*, v. 32, p. 929–934, 2003.
- RIECK, J. R.; NEDELMAN, J. R. Maximum likelihood estimation for log-linear models for the birnbaum-saunders distribution. *Technical Report 1*, State University of New York at Albany, Dept. of Biometry and Statistics, 1990.
- RIECK, J. R.; NEDELMAN, J. R. A log-linear model for the birnbaum-saunders distribution. *Technometrics*, v. 33, n. 1, p. 51–60, 1991.
- SANTANNA, L.; VILCA, F.; LEIVA, V. Influence analysis in skew-birnbaum-saunders regression models and applications. *Journal of Applied Statistics*, v. 38, n. 8, p. 1633–1649, 2011.
- SANTOS-NETO, M.; CYSNEIROS, F. J. A.; LEIVA, V.; AHMED, S. E. On new parameterizations of the birnbaum-saunders distributions. *Pak. J. Statist.*, v. 28, n. 1, p. 1–26, 2012.
- SHRÖEDINGER, E. Zur theorie der fall-und steigversuche an teilchenn mit brownscher bewegung. *Physikalische Zeitschrift.*, v. 16, p. 289–295, 1915.

- SMITH, R. L. Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, v. 72, p. 67–90, 1985.
- THORPE, J. A. *Elementary Topics in Differential Geometry*. [S.l.]: Springer-Verlag, 1979.
- TOMAZELLA, V.; NOBRE, J. S.; PEREIRA, G. H. A.; SANTOS-NETO, M. Zero-adjusted brinbaum-saunders regression model. <https://arxiv.org/abs/1802.00517v1>, Cornell University, 2018.
- TOMAZELLA, V.; NOBRE, J. S.; PEREIRA, G. H. A.; SANTOS-NETO, M. Zero-adjusted reparameterized brinbaum-saunders regression model. *Statistics & Probability Letters*, v. 149, p. 142–145, 2019.
- TSIOMAS, E. G. Bayesian inference birnbaum-saunders regression. *Comput. Stat. Theory Methods*, v. 30, n. 1, p. 179–193, 2001.
- TSUYUGUCHI, A.; PAULA, G. A.; BARROS, M. Analysis of correlated birnbaum-saunders data based on estimating equations. *Test*, Springer, 2019.
- VAART, A. van der. *Asymptotic Statistics*. [S.l.]: Cambridge University Press, 1998.
- VILCA, F.; LEIVA, V. A new fatigue life model based on the family of skew elliptic distributions. *Comm. Statist. Theor. Meth.*, v. 35, p. 229–244, 2006.
- VILCA, F.; SANHUEZA, A.; LEIVA, V.; CHISTAKOS, G. An extended birnbaum-saunders model and its applications in the study of environmental quality in santiago-chile. *Stoch. Env. Res Riks Analysis*, v. 24, n. 5, p. 782–771, 2010.
- VILLEGAS, C.; PAULA, G.; LEIVA, V. Birnbaum-saunders mixed models for censored reliability data analysis. *IEEE Transactions on Reliability*, v. 60, n. 4, p. 748–758, 2011.
- WANG, M.; SUN, X.; PARK, C. Bayesian analysis of birnbaum-saunders distribution via the generalized ratio-of-uniforms method. *Comput. Stat.*, v. 31, n. 1, p. 207–225, 2016.
- WANG, Z.; DESMOND, A. F.; LU, X. Modified censored moment estimation for the two-parameter birnbaum-saunders distribution. *Comput. Stat. Data Analysis*, v. 50, n. 4, p. 1033–1051, 2006.
- XIE, F.-C.; WEI, B.-C. Diagnostic analysis for log-birnbaum-saunders regression models. *Comput. Stat. Data Analysis*, v. 51, n. 9, p. 4692–4706, 2007.
- ZHANG, L.; SHABY, B. Uniqueness and global optimality of the maximum likelihood estimator for the generalized value distribution. *arXiv:2008.06400v1*, p. 1–37, 2020.
- ZHOU, C. Existence and consistency of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis*, v. 100, p. 794–815, 2009.
- ZHOU, C. The extent of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis*, v. 101, p. 971–983, 2010.
- ZHU, H.; IBRAHIM, J. G.; LEE, S.; ZHANG, H. et al. Perturbation selection and influence measures in local influence analysis. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 35, n. 6, p. 2565–2588, 2007.

APÊNDICE A – A ESTATÍSTICA DE EXTREMOS UNIVARIADOS

A.1 MODELAGEM PROBABILÍSTICA EM ESTATÍSTICA DE EXTREMOS

Neste apêndice, apresentamos uma breve introdução ao arcabouço probabilístico contido na Teoria de Valores Extremos e que fundamenta a Estatística de Extremos Univariados. O material exibido nas seções seguintes é uma síntese do Capítulo 3 do livro (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997).

A.1.1 Distribuições limites para o máximo amostral

Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias independentes, identicamente distribuídas (iid), não-degeneradas e com função de distribuição comum denotada por F . Definimos o máximo amostral, e indicamos M_n , por

$$M_1 = X_1 \text{ e } M_n = \max\{X_1, X_2, \dots, X_n\}, \quad n \geq 2. \quad (\text{A.1})$$

Interessa-nos investigar o comportamento da estatística M_n . Ressaltamos que os resultados a serem obtidos também são válidos, com as devidas adaptações, para a estatística mínimo amostral. Isso é uma consequência da identidade $\min\{X_1, \dots, X_n\} = -\max\{-X_1, \dots, -X_n\}$.

Antes de apresentarmos alguns resultados básicos sobre comportamento assintótico de M_n , passemos a definição de limite superior do suporte da variável cuja distribuição é F , denotado por x_F , e definido por:

$$x_F = \sup\{x \in \mathbb{R}; F(x) < 1\}. \quad (\text{A.2})$$

Intuitivamente, valores extremos ocorrem próximo ao limite superior do suporte de F . Dessa forma, o comportamento assintótico do máximo amostral deve ser similar ao da cauda direita de F próximo de x_F . Isso de fato se verifica e a justificativa está contida na combinação dos seguintes resultados básicos:

Resultado 1. Para todo $x < x_F$, tem-se $P(M_n \leq x) \rightarrow 0$, quando $n \rightarrow \infty$. E, no caso em que $x_F < \infty$, se $x \geq x_F$, então $P(M_n \leq x) \rightarrow 1$.

Resultado 2. Dada $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias iid, o máximo amostral M_n converge em probabilidade para x_F . Em símbolos, $M_n \xrightarrow{P} x_F$, quando $n \rightarrow \infty$.

Resultado 3. Dado que a sequência de variáveis aleatórias $\{M_n\}_{n \geq 1}$ é não-decrescente em n e $M_n \xrightarrow{P} x_F$, segue que M_n converge quase certamente (a.s.) para x_F . Em símbolos, $M_n \xrightarrow{a.s.} x_F$, quando $n \rightarrow \infty$.

O fato de que a distribuição assintótica de M_n é uma função degenerada em x_F não agrega informação satisfatória se almejamos calcular probabilidades da forma $P(M_n \leq x)$ para valores moderados de n . Um resultado fundamental na Teoria de Valores Extremos, conhecido como Teorema de Fisher-Tippett, assegura que se existem constantes $c_n > 0$ e $d_n \in \mathbb{R}$ tais que

$$c_n^{-1}(M_n - d_n) \xrightarrow{D} H, \quad n \rightarrow \infty \quad (\text{A.3})$$

para alguma distribuição não-degenerada H , então H é do tipo valor extremo. Porém, nem sempre é possível padronizarmos o máximo amostral de modo que a distribuição assintótica seja não-degenerada. Um exemplo disso é o máximo de variáveis aleatórias iid que seguem uma distribuição de Poisson, para uma justificativa desse fato veja (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997).

Neste ponto, interessa-nos determinar condições que a função de distribuição subjacente F deve satisfazer a fim de que se possa assegurar a existência de limite não-degerado para $P(M_n \leq u_n)$ quando $n \rightarrow \infty$ e $u_n = c_n x + d_n$. Os resultados a seguir expressam condições que garante a existência e a natureza do limite em termos do comportamento da cauda da função F , ou seja, $\bar{F} = 1 - F$.

Teorema A.1 (*Aproximação Poisson*)

Dados $\tau \in [0, \infty]$ e uma sequência $(u_n)_{n \geq 1}$, as afirmações a seguir são equivalentes:

- i) $n\bar{F}(u_n) \rightarrow \tau, \quad n \rightarrow \infty.$
- ii) $P(M_n \leq u_n) \rightarrow e^{-\tau}, \quad n \rightarrow \infty.$

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) p. 116.

Corolário A.1.1

Seja F uma função de distribuição com limite superior do suporte x_F finito e $\bar{F}(x_F^-) = F(x_F) - F(x_F^-) > 0$. Se $(u_n)_{n \geq 1}$ é uma sequência de números reais tal que $P(M_n \leq u_n) \rightarrow \rho, \quad n \rightarrow \infty$, então ou $\rho = 0$ ou $\rho = 1$.

Teorema A.2

Seja F uma função de distribuição com limite superior do suporte $x_F \leq \infty$ e $\tau \in (0, \infty)$. Existe uma sequência $(u_n)_{n \geq 1}$ satisfazendo $n\bar{F}(u_n) \rightarrow \tau$, $n \rightarrow \infty$ se, e somente se, $\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{F}(x^-)} = 1$ e $F(x_F) = \lim_{x \uparrow x_F} F(x) = 1$.

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) p. 117.

A.1.2 Convergência fraca do máximo amostral sob transformações afins

Nesta seção, temos por objetivo caracterizar as possíveis distribuições limite para o máximo amostral, M_n , quando o mesmo for submetido a uma transformação afim positiva.

Definição A.1 (Distribuições Max-estáveis)

Uma variável aleatória X , não-degenerada, ou sua correspondente função de distribuição, é dita *max-estável* se, dada uma sequência de variáveis, X_1, \dots, X_n , independentes e identicamente distribuídas a X , é válida a equivalência

$$M_n \stackrel{d}{=} c_n X + d_n$$

para todo $n \geq 2$ e constantes apropriadas $c_n > 0$ e $d_n \in \mathbb{R}$. A notação $Y \stackrel{d}{=} X$ significa que Y e X seguem a mesma distribuição.

Exemplo. Seja X uma variável aleatória cuja função de distribuição é $\Lambda(x) = \exp(-\exp(-x))$, $x \in \mathbb{R}$. Mostraremos que X é max-estável com $c_n = 1$ e $d_n = \log(n)$. De fato, seja X_1, \dots, X_n uma sequência de variáveis aleatórias iid a X . Se $Y = X + \log(n)$, então a fda de Y é dada por

$$F_Y(y) = P(X + \log(n) \leq y) = \exp(-\exp(-(y - \log(n)))) = \exp(-n \exp(-y)), \quad (\text{A.4})$$

usando o fato de que $\Lambda(x) = \exp(-\exp(-x))$, $x \in \mathbb{R}$, temos

$$F_Y(y) = [\Lambda(y)]^n = F_{M_n}(y). \quad (\text{A.5})$$

Ou seja, $M_n \stackrel{d}{=} X + \log(n)$.

Teorema A.3 (Propriedade limite das leis max-estáveis)

A classe das distribuições max-estáveis coincide com a classe de todas as distribuições limites possíveis, não-degeneradas, para o máximo amostral, apropriadamente padronizado, de variáveis aleatórias independentes e identicamente distribuídas.

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997), p. 121.

A seguir, enunciaremos um resultado que é fundamental na Teoria dos Valores Extremos, pois apresenta as distribuições limites possíveis para o máximo amostral.

Teorema A.4 (FISHER; TIPPETT, 1928)

Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas. Se existem constantes padronizadoras $c_n > 0$, $d_n \in \mathbb{R}$ e uma distribuição não-degenerada H tal que

$$c_n^{-1}(M_n - d_n) \xrightarrow{D} H, \quad (\text{A.6})$$

então H é do mesmo tipo que uma das três seguintes distribuições:

$$\begin{aligned} \text{Fréchet: } \Phi_\alpha(x) &= \begin{cases} 0; & \text{se } x \leq 0. \\ \exp(-x^{-\alpha}); & \text{se } x > 0, \quad \alpha > 0. \end{cases} \\ \text{Weibull: } \Psi_\alpha(x) &= \begin{cases} \exp(-(-x)^\alpha); & \text{se } x \leq 0. \\ 1; & \text{se } x > 0, \quad \alpha > 0. \end{cases} \\ \text{Gumbel: } \Lambda(x) &= \exp(-\exp(-x)), \quad x \in \mathbb{R}. \end{aligned}$$

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) pp. 122 e 123.

Observações:

- 1ª) A distribuição limite, H , apresentada em (A.6) é única sob transformações afins positivas. Em outras palavras, se o limite figurar na forma $H(cx + d)$, com $c > 0$ e $d \in \mathbb{R}$, ou seja,

$$\lim_{n \rightarrow \infty} P(c_n^{-1}(M_n - d_n) \leq x) = H(cx + d),$$

ao considerarmos $\tilde{c}_n = \frac{c_n}{c}$ e $\tilde{d}_n = d_n - d \cdot \frac{c_n}{c}$, obtemos

$$\lim_{n \rightarrow \infty} P(\tilde{c}_n^{-1}(M_n - \tilde{d}_n) \leq x) = H(x).$$

- 2ª) Para fins de modelagem, as funções de distribuição Φ_α , Ψ_α e Λ são muito diferentes. Porém, do ponto de vista matemático, elas estão relacionadas segundo a seguinte equivalência: se $X > 0$, então

$$X \sim \Phi_\alpha \iff \log(X^\alpha) \sim \Lambda \iff -X^{-1} \sim \Psi_\alpha.$$

Definição A.2 *Distribuição valor extremo*

As distribuições Φ_α , Ψ_α e Λ , que figuram no teorema (A.4), são chamadas distribuições padrão de valores extremos. A mesma denominação recai sobre as variáveis aleatórias que seguem tais distribuições.

Observação: As distribuições de valores extremos são max-estáveis. Em particular,

1. Se $X \sim \Phi_\alpha$, então $M_n \stackrel{d}{=} n^{1/\alpha} X$;
2. Se $X \sim \Psi_\alpha$, então $M_n \stackrel{d}{=} n^{-1/\alpha} X$;
3. Se $X \sim \Lambda$, então $M_n \stackrel{d}{=} X + \log(n)$.

A.1.3 Max-Domínios de atração e constantes padronizadoras

Nesta subsecção, investigaremos as condições que a função de distribuição, F , da variável subjacente X deve satisfazer a fim de que o máximo amostral, devidamente padronizado, convirja fracamente para uma dada distribuição valor extremo padrão H . Uma questão relacionada com o propóstio citado consiste em determinar as constantes de normalização $c_n > 0$ e $d_n \in \mathbb{R}$ tais que se verifica a convergência $c_n^{-1}(M_n - d_n) \xrightarrow{D} H$. Destaquemos que diferentes constantes de padronização para as quais essa convergência é válida não acarretam distribuições limites distintas.

Definição A.3 (*Máximo domínio de atração*)

Dizemos que uma variável aleatória X , ou sua respectiva função de distribuição F , pertence ao máximo domínio de atração da distribuição valor extremo H , e denotamos $F \in \text{MDA}(H)$, se existem constantes reais $c_n > 0$ e $d_n \in \mathbb{R}$ tais que

$$c_n^{-1}(M_n - d_n) \xrightarrow{D} H. \quad (\text{A.7})$$

Teorema A.5 (*Caracterização do max-domínio H*)

A função de distribuição $F \in \text{MDA}(H)$, com constantes de padronização $c_n > 0$ e $d_n \in \mathbb{R}$, se, e somente se,

$$\lim_{n \rightarrow \infty} n\bar{F}(c_n x + d_n) = -\log(H(x)); \quad \forall x \in \mathbb{R}. \quad (\text{A.8})$$

Quando $H(x) = 0$, o limite é interpretado como sendo ∞ .

Observação: Usando a teoria das funções de variação regular é possível caracterizarmos cada um dos domínios de atração das distribuições Φ_α , Ψ_α e Λ . Como veremos adiante, a caracterização dos domínios de atração Fréchet (Φ_α) e Weibull (Ψ_α) revela-se mais simples do que a do domínio Gumbel (Λ), este último consiste de distribuições cuja cauda direita decresce mais rápido do que uma função potência. Antes de passarmos as caracterizações, apresentemos um conceito útil quando se compara caudas de distribuições.

Definição A.4

Dizemos que duas distribuições, F e G , possuem caudas equivalentes, e denotamos $\bar{F} \sim \bar{G}$, se elas têm o mesmo limite superior do suporte, isto é, $x_F = x_G \leq \infty$, e para alguma constante real c , com $0 < c < \infty$, tem-se $\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{G}(x)} = c$.

Observações:

- 1ª) A relação de equivalência de caudas (\sim) é reflexiva, simétrica e transitiva.
- 2ª) Cada um dos domínios de atração, Φ_α , Ψ_α e Λ , é fechado com respeito à equivalência de caudas, no seguinte sentido:

$$F \in \text{MDA}(H) \iff G \in \text{MDA}(H).$$

- 3ª) Dadas duas distribuições de caudas equivalentes, as sequências de constantes apropriadas para padronização do máximo amostral podem ser as mesmas.

A.1.3.1 Caracterização do max-domínio Fréchet

Uma caracterização para classe das distribuições que compõem o max-domínio Fréchet, (Φ_α), com $\alpha > 0$, é dada em termos do conceito de função de variação regular. Mostraremos que o max-domínio de atração Fréchet é constituído por todas as funções de distribuição cuja cauda direita possui variação regular com índice $-\alpha$. Para $F \in \text{MDA}(\Phi_\alpha)$ as constantes c_n e d_n podem ser escolhidas como $c_n = F^{\leftarrow}(1 - n^{-1}) = \left(1/\bar{F}\right)^{\leftarrow}(n)$ e $d_n = 0$.

Teorema A.6 (Caracterização do max-domínio Fréchet)

Uma função de distribuição F pertence ao máximo-domínio de atração Fréchet, e denotamos $F \in \text{MDA}(\Phi_\alpha)$, com $\alpha > 0$, se, e somente se, $\bar{F}(x) = x^{-\alpha}L(x)$, para alguma função L de variação lenta.

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) pp 131 e 132.

Observações: O Teorema A.6 tem as seguintes implicações:

- 1ª) Se $F \in \text{MDA}(\Phi_\alpha)$, então $c_n^{-1}M_n \xrightarrow{D} \Phi_\alpha$, com $c_n = F^{\leftarrow}(1 - n^{-1})$.
- 2ª) Se $F \in \text{MDA}(\Phi_\alpha)$, então o limite superior do suporte da variável associada a F é infinito, isto é, $x_F = \infty$.
- 3ª) A classe de distribuições do max-domínio Fréchet contém as distribuições de cauda pesada, no sentido de que $\mathbb{E}(X^+)^\delta = \infty$, para $\delta > \alpha$.

Corolário A.6.1 (*Condição de suficiência de Von Mises*)

Seja F uma função de distribuição absolutamente contínua e f sua correspondente função densidade de probabilidade. Se

$$\lim_{x \rightarrow \infty} \frac{xf(x)}{\bar{F}(x)} = \alpha > 0, \quad (\text{A.9})$$

então $F \in \text{MDA}(\Phi_\alpha)$.

O resultado a seguir assegura que a classe de funções de distribuição cuja cauda direita, \bar{F} , possui variação regular é fechada com respeito à equivalência de cauda dada pela Definição A.4.

Teorema A.7 (*Propriedade de fechamento do $\text{MDA}(\Phi_\alpha)$*)

Sejam F e G funções de distribuição. Suponhamos que $F \in \text{MDA}(\Phi_\alpha)$ com constantes normalizadoras $c_n > 0$, ou seja

$$\lim_{n \rightarrow \infty} F^n(c_n x) = \Phi_\alpha(x), \quad x > 0. \quad (\text{A.10})$$

Nessas condições,

$$\lim_{n \rightarrow \infty} G^n(c_n x) = \Phi_\alpha(cx), \quad x > 0. \quad (\text{A.11})$$

para algum $c > 0$ se, e somente se, F e G possuem caudas equivalentes com

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x)}{\bar{G}(x)} = c^\alpha. \quad (\text{A.12})$$

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) p. 133.

A.1.3.2 Caracterização do max-domínio Weibull

Assim como o max-domínio Fréchet, a caracterização do max-domínio Weibull, Ψ_α , para $\alpha > 0$, também recorre ao conceito de função de variação regular.

Teorema A.8 (Caracterização do max-domínio Weibull)

Uma função de distribuição F pertence ao máximo-domínio de atração Weibull, e denotamos $F \in \text{MDA}(\Psi_\alpha)$, com $\alpha > 0$, se, e somente se, $x_F < \infty$ e $\overline{F}(x_F - x^{-1}) = x^{-\alpha}L(x)$, para alguma função L que possui variação lenta.

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) p. 135.

Observações:

1ª) Se $F \in \text{MDA}(\Psi_\alpha)$, então $c_n^{-1}(M_n - d_n) \xrightarrow{D} \Psi_\alpha$, com $c_n = x_F - F^{\leftarrow}(1 - n^{-1})$ e $d_n = x_F$.

2ª) Se $F \in \text{MDA}(\Psi_\alpha)$, então o limite superior do suporte da variável associada a F é finito.

3ª) Os max-domínios Ψ_α e Φ_α se relacionam por meio da seguinte expressão:

$$\Psi_\alpha(-x^{-1}) = \Phi_\alpha(x), \quad x > 0. \quad (\text{A.13})$$

Corolário A.8.1 (Condição de suficiência de Von Mises)

Seja F uma função de distribuição absolutamente contínua cuja função densidade de probabilidade correspondente f é positiva em um intervalo limitado (z, x_F) . Se

$$\lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{\overline{F}(x)} = \alpha > 0, \quad (\text{A.14})$$

então $F \in \text{MDA}(\Psi_\alpha)$.

Teorema A.9 (Propriedade de fechamento do $\text{MDA}(\Psi_\alpha)$)

Sejam F e G funções de distribuição tais que $x_F = x_G < \infty$. Suponhamos que $F \in \text{MDA}(\Psi_\alpha)$ com constantes normalizadoras $c_n > 0$, ou seja,

$$\lim_{n \rightarrow \infty} F^n(c_n x + x_F) = \Psi_\alpha(x), \quad x < 0. \quad (\text{A.15})$$

Nessas condições,

$$\lim_{n \rightarrow \infty} G^n(c_n x + x_F) = \Psi_\alpha(cx), \quad x < 0. \quad (\text{A.16})$$

para algum $c > 0$ se, e somente se, F e G possuem caudas equivalentes com

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{G}(x)} = c^{-\alpha}. \quad (\text{A.17})$$

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) p. 136.

A.1.3.3 Caracterização do max-domínio Gumbel

O máximo domínio de atração Gumbel, $\text{MDA}(\Lambda)$, agrega uma classe de distribuições com caudas muito diferentes, algumas possuindo caudas leves, por exemplo a normal, e outras tendo caudas moderadamente pesadas, por exemplo a log-normal. Ele é formado por distribuições que estão estreitamente relacionadas com as funções que admitem uma representação na forma de função de Von Mises.

A função distribuição Gumbel é definida por

$$\Lambda(x) = \exp(-\exp(-x)), \quad x \in \mathbb{R}. \quad (\text{A.18})$$

De imediato, podemos identificar, usando expansão em série de Taylor, que $\Lambda(x)$ possui cauda equivalente a $\exp(-x)$, ou seja, $\bar{\Lambda}(x) = 1 - \Lambda(x) \sim e^{-x}$. De fato, de início observemos que

$$\bar{\Lambda}(x) = \sum_{n=1}^{\infty} (-1)^n \frac{\exp(-nx)}{n!}, \quad x \in \mathbb{R}.$$

Daí,

$$\lim_{x \rightarrow \infty} \frac{\bar{\Lambda}(x)}{e^{-x}} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} = c < \infty$$

A convergência da série $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n!}$ é assegurada pelo Teste de Leibniz, pois os termos $u_n = \frac{1}{n!}$ são positivos, decrescentes e $u_n \rightarrow 0$. Do exposto, segue que a cauda de Λ decresce para zero na razão exponencial.

Definição A.5 (Funções von Mises)

Sejam F uma função de distribuição e $x_F \leq \infty$ o limite superior do suporte. Se existe algum $z < x_F$ tal que $\bar{F} = 1 - F$ admite uma representação na forma

$$\bar{F}(x) = c \exp\left(-\int_z^x \frac{1}{a(t)} dt\right), \quad z < x < x_F, \quad (\text{A.19})$$

onde c é uma constante real positiva, $a(\cdot)$ é uma função positiva absolutamente contínua (com respeito a medida de Lebesgue), com densidade $a'(\cdot)$ e $\lim_{x \uparrow x_F} a'(x) = 0$, então diz-se que F é uma função von Mises com função auxiliar $a(\cdot)$.

Exemplo. Seja F uma fd com $x_F \leq \infty$ e assuma que existe algum $z < x_F$ tal que F é duas vezes derivável no intervalo (z, x_F) com densidade positiva $f = F'$ e $F''(x) < 0$ para todo $x \in (z, x_F)$. Nesses termos, F é uma função von Mises com função auxiliar $a = \frac{\bar{F}}{f}$ se, e somente se, $\lim_{x \uparrow x_F} \frac{\bar{F}(x)F''(x)}{f^2(x)} = -1$.

Teorema A.10 (*Propriedades das funções von Mises*)

Se uma distribuição, F , é uma função von Mises, então F é absolutamente contínua no intervalo (z, x_F) , para algum $z < x_F \leq \infty$, com função densidade f . A função auxiliar pode ser tomada como $a(x) = \frac{\bar{F}(x)}{f(x)}$. Além disso, são válidas as seguintes afirmativas:

- i) Se $x_F = \infty$, então $\bar{F} \in \mathcal{R}_{-\infty}$ e $\lim_{x \rightarrow \infty} \frac{xf(x)}{\bar{F}(x)} = \infty$.
- ii) Se $x_F < \infty$, então $\bar{F}(x_F - x^{-1}) \in \mathcal{R}_{-\infty}$ e $\lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{\bar{F}(x)} = \infty$.

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) p. 141.

A notação $\bar{F} \in \mathcal{R}_{-\infty}$ indica que \bar{F} possui variação rápida com índice $-\infty$. Mais detalhes podem ser encontrados na última seção deste apêndice.

A proposição adiante assegura que a classe de distribuições von Mises pertence ao máximo domínio de atração Gumbel e sugere valores para as sequências de constantes para padronização.

Teorema A.11 (*Funções von Mises e o MDA(Λ)*)

Se uma distribuição F é uma função von Mises, então $F \in \text{MDA}(\Lambda)$. Além disso, uma escolha possível para as sequências de constantes para padronização do máximo é dada por

$$d_n = F^{\leftarrow}(1 - n^{-1}) \quad \text{e} \quad c_n = a(d_n),$$

onde $a(\cdot)$ é a função auxiliar de F .

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997), pp. 141 e 142.

Teorema A.12 (*Primeira caracterização do MDA(Λ)*)

Uma função de distribuição $F \in \text{MDA}(\Lambda)$ se, e somente se, existe $z < x_F$ tal que F admite uma representação na forma

$$\bar{F}(x) = c(x) \exp\left(-\int_z^x \frac{g(t)}{a(t)} dt\right), \quad z < x < x_F, \quad (\text{A.20})$$

onde $c(\cdot)$ e $g(\cdot)$ são funções mensuráveis tais que $\lim_{x \rightarrow x_F} c(x) = 0$ e $\lim_{x \rightarrow x_F} g(x) = 1$. Além disso, $a(\cdot)$ é positiva absolutamente contínua, com respeito a medida de Lebesgue, e sua densidade $a'(\cdot)$ é tal que $\lim_{x \rightarrow x_F} a'(x) = 0$. Para F com representação dada por A.20 podemos considerar as seguintes sequências de constantes para padronização do máximo

$$d_n = F^{\leftarrow}(1 - n^{-1}) \quad \text{e} \quad c_n = a(d_n),$$

onde $a(\cdot)$ é uma função auxiliar de F que pode ser tomada como

$$a(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt, \quad x < x_F. \quad (\text{A.21})$$

Prova: veja em (RESNICK, 1987).

Teorema A.13 (*Segunda caracterização do MDA(Λ)*)

Uma função de distribuição $F \in \text{MDA}(\Lambda)$ se, e somente se, existe uma função positiva $\tilde{a}(\cdot)$ tal que

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x + t\tilde{a}(x))}{\bar{F}(x)} = e^{-t}, \quad t \in \mathbb{R}. \quad (\text{A.22})$$

Uma escolha possível para a função \tilde{a} é $\tilde{a}(x) = a(x)$, com $a(x)$ dada pela expressão A.21.

Prova: consta em (HAAN, 1970).

Teorema A.14 (*Propriedade de fechamento do max-domínio Gumbel*)

Sejam F e G funções de distribuição tais que $x_F = x_G$. Suponhamos que $F \in \text{MDA}(\Lambda)$ com constantes normalizadoras $c_n > 0$ e $d_n \in \mathbb{R}$, isto é,

$$\lim_{n \rightarrow \infty} F^n(c_n x + d_n) = \Lambda(x), \quad x > 0. \quad (\text{A.23})$$

Nessas condições,

$$\lim_{n \rightarrow \infty} G^n(c_n x + d_n) = \Lambda(x + b), \quad x > 0. \quad (\text{A.24})$$

para algum $b \in \mathbb{R}$ se, e somente se, F e G possuem caudas equivalentes com

$$\lim_{x \uparrow x_F} \frac{\overline{F}(x)}{\overline{G}(x)} = e^b. \quad (\text{A.25})$$

Prova: veja em (EMBRECHTS; KLUPPELBERG; MIKOSCH, 1997) p. 144.

A.1.4 Inversas de Funções Monótonas.

Seja $H : \mathbb{R} \rightarrow (a, b)$ uma função não-decrescente com imagem (a, b) , onde $-\infty \leq a < b \leq \infty$. Adotada a convenção de que o ínfimo do conjunto vazio é $+\infty$, definimos a inversa generalizada de H , e denotamos H^\leftarrow , por

$$H^\leftarrow : (a, b) \rightarrow \mathbb{R}, \quad H^\leftarrow(y) = \inf\{s; H(s) \geq y\}. \quad (\text{A.26})$$

Observações:

1ª) A inversa generalizada também é chamada inversa contínua à esquerda.

2ª) A função H^\leftarrow é contínua à esquerda em $x \in (a, b)$.

3ª) Sob a hipótese adicional de que H é contínua à direita, temos:

3a) $A(y) := \{s; H(s) \geq y\}$ é um conjunto fechado .

3b) $H(H^\leftarrow(y)) \geq y$.

3c) $H^\leftarrow(y) \leq t \iff y \leq H(t)$ e $t < H^\leftarrow(y) \iff H(t)$.

Agora, apresentamos dois resultados importantes relativos à convergência de sequências de funções monótonas. Dada uma função monótona $H : \mathbb{R} \rightarrow \mathbb{R}$, definamos o conjunto $\mathcal{C}(H) = \{x \in \mathbb{R}; H \text{ é finita e contínua em } x\}$. Dizemos que uma sequência $\{H_n\}_{n \in \mathbb{N}}$ de funções não-decrescentes definidas sobre \mathbb{R} converge fracamente para a função H_0 se $H_n(x) \rightarrow H_0(x)$, para todo $x \in \mathcal{C}(H_0)$, quando $n \rightarrow \infty$. E, denotamos tal convergência por $H_n \rightarrow H_0$. Se $\{X_n\}_{n \in \mathbb{N}}$ é uma sequência de variáveis aleatórias que converge para X_0 e cujas correspondentes funções de distribuição formam a sequência $\{F_n\}_{n \in \mathbb{N}}$, usamos a notação $X_n \Rightarrow X_0$ para denotar que $F_n \rightarrow F_0$.

Teorema A.15

Se $\{H_n : \mathbb{R} \rightarrow (a, b)\}_{n \in \mathbb{N}}$ é uma sequência de funções não-decrescentes com imagem (a, b) tal que $H_n \rightarrow H_0$, então $H_n^{\leftarrow} \rightarrow H_0^{\leftarrow}$ no seguinte sentido: para todo $t \in (a, b) \cap \mathcal{C}(H_0^{\leftarrow})$, tem-se $H_n^{\leftarrow}(t) \rightarrow H_0^{\leftarrow}(t)$

Teorema A.16

Se , para cada $n \geq 0$, X_n é uma variável aleatória real no espaço de probabilidade $(\Omega_n, \mathcal{B}_n, P_n)$ tal que $X_n \Rightarrow X_0$, então existe uma sequência de variáveis aleatórias $\{\tilde{X}_n, n \geq 0\}$ definidas no espaço de probabilidade de Lebesgue $([0, 1], \mathcal{B}[0, 1], \mu)$ tal que

- (i) $\tilde{X}_n \stackrel{d}{=} X_n$ para cada $n \geq 0$,
- (ii) $\tilde{X}_n \rightarrow \tilde{X}_0$, quase certamente, em relação a medida μ .

A.1.5 Variação Regular

Definição.

- (a) Uma função positiva e mensurável, $L : (0, \infty) \rightarrow (0, \infty)$, possui variação lenta no infinito, e denotamos $L \in \mathcal{R}_0$, se

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1, \quad t > 0. \quad (\text{A.27})$$

- (b) Uma função positiva e mensurável, $h : (0, \infty) \rightarrow (0, \infty)$, diz-se possuir variação regular no infinito com índice $\alpha \in \mathbb{R}$, e denotamos $h \in \mathcal{R}_\alpha$, se

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = t^\alpha, \quad t > 0. \quad (\text{A.28})$$

- (c) Uma função positiva e mensurável, $h : (0, \infty) \rightarrow (0, \infty)$, diz-se possuir variação rápida com índice $-\infty$, e denotamos $h \in \mathcal{R}_{-\infty}$, se

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = \begin{cases} 0; & \text{se, } t > 1, \\ \infty; & \text{se, } 0 < t < 1. \end{cases} \quad (\text{A.29})$$

APÊNDICE B – GEOMETRIA EM SUPERFÍCIES QUE SÃO GRÁFICOS DE FUNÇÕES DIFERENCIÁVEIS

Há muitas formas de apresentarmos um estudo sobre a geometria do gráfico de uma função diferenciável. Uma abordagem mais geral, usa o fato de que o gráfico de uma função suave é uma variedade diferenciável, daí as propriedades geométricas desta são válidas para aquela, isso é feito, por exemplo, em (LEE, 2013). Outra forma, identifica o gráfico como sendo uma superfície (ou hiperfície) contida no espaço euclidiano, veja em (THORPE, 1979). Uma terceira, que atende plenamente aos nossos propósitos, é apresentada por (POON; POON, 2010) que desenvolve um estudo específico para superfícies obtidas a partir do gráfico de uma função diferenciável. Nas subseções seguintes, apresentamos uma síntese da exposição que consta em (POON; POON, 2010). O exposto a seguir fundamenta substancialmente a análise de influência local por meio da função afastamento pela verossimilhança e contribui para uma compreensão mais profunda das medidas de influência.

B.0.1 Nocões básicas

Sejam $U \subset \mathbb{R}^q$ um aberto e $f : U \rightarrow \mathbb{R}$ uma função diferenciável que a cada ponto $\mathbf{x} = (x_1, \dots, x_q)^\top \in \mathbb{R}^q$ associa um único número real $f(\mathbf{x})$. O gráfico de f é o conjunto de pontos em \mathbb{R}^{q+1} definido por

$$G = \{\mathbf{y} \in \mathbb{R}^{q+1} : y_j = x_j \text{ para todo } 1 \leq j \leq q \text{ e } y_{q+1} = f(\mathbf{x}), \mathbf{x} \in U\}. \quad (\text{B.1})$$

Notemos que o gráfico de f pode ser identificado como a imagem da aplicação $\phi : U \rightarrow \mathbb{R}^{q+1}$ tal que

$$\phi(\mathbf{x}) = \begin{pmatrix} \mathbf{x} \\ f(\mathbf{x}) \end{pmatrix}, \quad \mathbf{x} \in U. \quad (\text{B.2})$$

Nesses termos, segue que G é uma superfície q -dimensional em \mathbb{R}^{q+1} , ou seja, uma hiperfície, sendo ϕ uma parametrização de todo o conjunto G .

A diferencial da aplicação ϕ é a transformação linear injetora $d\phi : \mathbb{R}^q \rightarrow \mathbb{R}^{q+1}$ cuja

matriz de representação na base canônica é

$$\mathbf{d}\phi = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial x_3} & \cdots & \frac{\partial f}{\partial x_q} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_q \\ \mathbf{d}f \end{bmatrix}_{(q+1) \times q}. \quad (\text{B.3})$$

onde \mathbf{I}_q representa a matriz identidade de ordem q e $\mathbf{d}f$ a diferencial da função f . Logo, a imagem de um vetor $\mathbf{v} = (v_1, \dots, v_q)^\top \in \mathbb{R}^q$ através de $\mathbf{d}\phi$ é dada por

$$\mathbf{d}\phi(\mathbf{v}) = \begin{pmatrix} \mathbf{I}_q \\ \mathbf{d}f \end{pmatrix} \cdot \mathbf{v} = \begin{pmatrix} \mathbf{v} \\ \sum_{k=1}^q \frac{\partial f}{\partial x_k} \cdot v_k \end{pmatrix}. \quad (\text{B.4})$$

Ou equivalentemente,

$$\mathbf{d}\phi(\mathbf{v}) = \begin{pmatrix} \mathbf{v} \\ \langle \nabla f, \mathbf{v} \rangle \end{pmatrix}. \quad (\text{B.5})$$

Considere a superfície G dada pela equação (B.1) e uma parametrização $\phi : U \subset \mathbb{R}^q \rightarrow V \subset \mathbb{R}^{q+1}$ com $\mathbf{x} \in U$ e $\phi(\mathbf{x}) \in G$. O espaço tangente a G no ponto $\phi(\mathbf{x})$ é o espaço vetorial de dimensão q cujos vetores são os elementos do conjunto $T_{\phi(\mathbf{x})}G$ dado por

$$T_{\phi(\mathbf{x})}G = \{\mathbf{w} = \mathbf{d}\phi(\mathbf{v}) : \mathbf{v} \in \mathbb{R}^q\}. \quad (\text{B.6})$$

Geometricamente, o espaço tangente $T_{\phi(\mathbf{x})}G$ consiste de todos os vetores velocidade em $\phi(\mathbf{x})$ dos caminhos diferenciáveis contidos em G que passa por $\phi(\mathbf{x})$. Além disso, a definição de espaço tangente em um ponto $\mathbf{p} \in G$ independe da escolha da parametrização ϕ . Por fim, observemos que a coleção de vetores $\mathcal{B} = \{\mathbf{d}\phi(\mathbf{e}_1), \dots, \mathbf{d}\phi(\mathbf{e}_q)\}$ é uma base $T_{\phi(\mathbf{x})}G$.

Diz-se que $\mathbf{n} \in \mathbb{R}^{q+1}$ é um vetor normal à superfície G no ponto $\mathbf{p} \in G$ se $\langle \mathbf{n}, \mathbf{w} \rangle = 0$ para todo $\mathbf{w} \in T_{\mathbf{p}}G$. O conjunto dos vetores normais à G no ponto \mathbf{p} é um subespaço vetorial de \mathbb{R}^{q+1} e o denotamos por $T_{\mathbf{p}}G^\perp$. Em símbolos,

$$T_{\mathbf{p}}G^\perp = \{\mathbf{n} : \langle \mathbf{n}, \mathbf{w} \rangle = 0 \text{ para todo } \mathbf{w} \in T_{\mathbf{p}}G\}. \quad (\text{B.7})$$

Observemos que se \mathbf{e}_j é o j -ésimo vetor canônico de \mathbb{R}^q e $\mathbf{x}_0 \in U$, então a diferencial de $\phi(\mathbf{x}) = (\mathbf{x}, f(\mathbf{x}))^\top$ no ponto \mathbf{x}_0 avaliada no vetor canônico \mathbf{e}_j é dada por

$$\mathbf{d}\phi(\mathbf{x}_0) \cdot \mathbf{e}_j = (\mathbf{e}_j, \mathbf{d}f(\mathbf{x}_0) \cdot \mathbf{e}_j)^\top = \left(\mathbf{e}_j, \sum_{k=1}^q \frac{\partial f}{\partial x_k}(\mathbf{x}_0) \cdot e_{j,k} \right)^\top = \left(\mathbf{e}_j, \frac{\partial f}{\partial x_j}(\mathbf{x}_0) \right)^\top. \quad (\text{B.8})$$

Agora, consideremos o vetor $\mathbf{n} = (-\nabla_f(\mathbf{x}_0), 1)^\top$ e observemos que

$$\langle \mathbf{n}, \mathbf{d}\phi(\mathbf{e}_j) \rangle = \langle (-\nabla_f(\mathbf{x}_0), 1)^\top, \left(\mathbf{e}_j, \frac{\partial f}{\partial x_j}(\mathbf{x}_0) \right)^\top \rangle = - \sum_{k=1}^q \frac{\partial f}{\partial x_k}(\mathbf{x}_0) e_{j,k} + \frac{\partial f}{\partial x_j}(\mathbf{x}_0) = 0. \quad (\text{B.9})$$

Daí, se $\mathbf{w} \in T_{\phi(\mathbf{x}_0)}G$, então

$$\langle \mathbf{n}, \mathbf{w} \rangle = \langle \mathbf{n}, \sum_{k=1}^q \mathbf{d}\phi(\mathbf{e}_k) \rangle = \sum_{k=1}^q \langle \mathbf{n}, \mathbf{d}\phi(\mathbf{e}_k) \rangle = 0. \quad (\text{B.10})$$

Ou seja $\mathbf{n} = (-\nabla_f(\mathbf{x}_0), 1)^\top$ é um vetor normal à superfície G no ponto $\phi(\mathbf{x}_0) \in G$. Em particular, se \mathbf{x}_0 é um ponto crítico de f , então \mathbf{n} é igual ao vetor canônico \mathbf{e}_{q+1} .

Proposição Seja $f : U \rightarrow \mathbb{R}$ uma função diferenciável no aberto $U \subset \mathbb{R}^q$. Se \mathbf{x}_0 é um ponto crítico de f , então a diferencial $\mathbf{d}\phi$ identifica o espaço tangente de \mathbb{R}^q no ponto \mathbf{x}_0 com o subespaço vetorial

$$\{(w_1, \dots, w_n, w_{q+1})^\top \in \mathbb{R}^{q+1} : w_{q+1} = 0\}. \quad (\text{B.11})$$

Em outras palavras, $\mathbf{d}\phi$ é a imersão canônica de \mathbb{R}^q em \mathbb{R}^{q+1} .

Prova. Veja em (POON; POON, 2010).

B.0.2 A seção normal

Existem diferentes formas de analisarmos o comportamento geométrico do gráfico, G , de uma função diferenciável $f : U \subset \mathbb{R}^q \rightarrow \mathbb{R}$. Uma delas, consiste em investigarmos propriedades de $f(\mathbf{x})$ em relação a cada uma das coordenadas do vetor $\mathbf{x} = (x_1, \dots, x_q)^\top \in \mathbb{R}^q$. Nesse contexto, o exame do comportamento de f próximo ao ponto \mathbf{x} , em relação à coordenada x_j , para algum $j = 1, \dots, q$, pode ser feito considerando-se a função a valores reais, f_j , definida no intervalo aberto $(-\varepsilon, \varepsilon)$, para algum $\varepsilon > 0$, e cuja imagem de t é dada por

$$f_j(t) = f(x_1, \dots, x_{j-1}, t + x_j, x_{j+1}, \dots, x_q). \quad (\text{B.12})$$

Assim, o gráfico de f_j está contido em \mathbb{R}^2 e pode visto como sendo o traço do caminho diferenciável $\phi_j : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^2$, $\phi_j(t) = (t, f_j(t))^\top$. De um modo geral, se $\mathbf{v} = (v_1, \dots, v_q)^\top$ é um vetor não-nulo em \mathbb{R}^q , podemos considerar a função de uma variável, $t \mapsto f(\mathbf{x} + t\mathbf{v})$, cujo gráfico, contido em \mathbb{R}^2 , é o traço de um caminho diferenciável.

De forma alternativa, dado um vetor não-nulo $\mathbf{v} \in \mathbb{R}^q$ consideremos o espaço vetorial gerados pelos vetores $\mathbf{d}\phi(\mathbf{v})$ e o vetor normal \mathbf{n} . Esse espaço vetorial possui dimensão 2 e

corta transversalmente o gráfico de $f : U \rightarrow \mathbb{R}$, determinando sobre G uma curva suave. Mais precisamente, sejam $\mathbf{w} \in \mathbb{R}^{q+1}$ um vetor tangente a G no ponto $\phi(\mathbf{x})$ e $N(\mathbf{w})$ o conjunto definido por

$$N(\mathbf{w}) = \{\mathbf{y} \in \mathbb{R}^{q+1} : \mathbf{y} = \phi(\mathbf{x}) + \mu\mathbf{w} + \nu\mathbf{n}, \text{ onde } \mu, \nu \in \mathbb{R}\}, \quad (\text{B.13})$$

onde \mathbf{n} denota o vetor normal a \mathbf{w} . A seção normal na direção do vetor \mathbf{w} no ponto $\phi(\mathbf{x})$, denotada por $G(\mathbf{w})$, é definida como sendo a seguinte intersecção:

$$G(\mathbf{w}) = \phi(U) \cap N(\mathbf{w}) = G \cap N(\mathbf{w}). \quad (\text{B.14})$$

Em termos geométricos, $N(\mathbf{w})$ é o plano em \mathbb{R}^{q+1} gerado pelos vetores \mathbf{w} e \mathbf{n} e que passa pelo ponto $\phi(\mathbf{x})$. Por outro lado, $G(\mathbf{w})$ é uma curva contida no plano $N(\mathbf{w})$ e que passa pelo ponto $\phi(\mathbf{x})$. Além disso, dado que \mathbf{n} é perpendicular a G no ponto $\phi(\mathbf{x})$, segue que o vetor tangente a curva $G(\mathbf{w})$, no ponto $\phi(\mathbf{x})$, pertence ao espaço vetorial gerado \mathbf{w} .

Por fim, sejam \mathbf{v} é um vetor unitário em \mathbb{R}^q e $\mathbf{x}_0 + t\mathbf{v}$ o segmento de reta que está contido no aberto U para $t \in (-\varepsilon, \varepsilon)$. Agora, consideremos a função de uma variável real $\nu(t) = f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)$, onde \mathbf{x}_0 é um ponto crítico de f . Isso nos possibilita obter uma seção normal ao gráfico de f , no ponto $\phi(\mathbf{x}_0)$ e na direção do vetor tangente $d\phi(\mathbf{v})$. De fato, observemos que

$$\phi(\mathbf{x}_0 + t\mathbf{v}) = \begin{pmatrix} \mathbf{x}_0 + t\mathbf{v} \\ f(\mathbf{x}_0) + \nu \end{pmatrix} = \phi(\mathbf{x}_0) + t \begin{pmatrix} \mathbf{v} \\ 0 \end{pmatrix} + \nu \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = \phi(\mathbf{x}_0) + td\phi(\mathbf{v}) + \nu\mathbf{n}. \quad (\text{B.15})$$

Logo, $\phi(\mathbf{x}_0 + t\mathbf{v}) \in G(d\phi(\mathbf{v}))$ para $t \in (-\varepsilon, \varepsilon)$.

B.0.3 As seções transversais em dimensão superior

O conceito de seção normal pode ser estendido para dimensão superior. De fato, sejam $\mathbf{w}_1, \dots, \mathbf{w}_k$ vetores linearmente independentes em \mathbb{R}^{q+1} e tangentes ao gráfico G no ponto $\phi(\mathbf{x})$. Define-se o conjunto $N(\mathbf{w}_1, \dots, \mathbf{w}_k)$ do seguinte modo:

$$N(\mathbf{w}_1, \dots, \mathbf{w}_k) = \{\mathbf{y} \in \mathbb{R}^{q+1} : \mathbf{y} = \phi(\mathbf{x}) + \sum_{j=1}^k \mu_j \mathbf{w}_j + \nu\mathbf{n}, \text{ onde } \nu, \mu_1, \dots, \mu_k \in \mathbb{R}\}. \quad (\text{B.16})$$

Dessa forma, o conjunto $G(\mathbf{w}_1, \dots, \mathbf{w}_k) = \phi(U) \cap N(\mathbf{w}_1, \dots, \mathbf{w}_k)$ é chamado de seção transversal k -dimensional no ponto $\phi(\mathbf{x})$. Quando \mathbf{w} é uma combinação linear de \mathbf{w}_1 e \mathbf{w}_2 , tem-se $G(\mathbf{w}) \subset G(\mathbf{w}_1, \mathbf{w}_2)$. Além disso, se $\{\mathbf{w}_1, \dots, \mathbf{w}_q\}$ forma uma base de $T_{\phi(\mathbf{x})}G$, então $G = G(\mathbf{w}_1, \dots, \mathbf{w}_q)$. Por fim, tem-se que são válidas as seguintes inclusões:

$$G(\mathbf{w}_i) \subset G(\mathbf{w}_i, \mathbf{w}_j) \subset G(\mathbf{w}_i, \mathbf{w}_j, \mathbf{w}_k) \subset \dots \subset G(\mathbf{w}_i, \dots, \mathbf{w}_q). \quad (\text{B.17})$$

B.0.4 As formas fundamentais

A partir do produto interno usual $\langle \cdot, \cdot \rangle$ definido sobre \mathbb{R}^q , podemos calcular o comprimento de um vetor $\mathbf{v} \in \mathbb{R}^q$ que seja tangente a uma curva diferenciável $\alpha : (-\varepsilon, \varepsilon) \rightarrow U \subset \mathbb{R}^q$ em $t = 0$. Com efeito, ao produto interno $\langle \cdot, \cdot \rangle$ associamos a norma $\|\cdot\| : \mathbb{R}^q \rightarrow \mathbb{R}$, $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$. Dada a superfície G , obtida a partir do gráfico da função diferenciável $f : U \rightarrow \mathbb{R}$, interessa-nos calcular o comprimento de vetores que são tangentes a G em um ponto $\mathbf{p} = \phi(\mathbf{x}) \in G$, onde $\phi : U \rightarrow G$ é uma parametrização de G .

Dada uma direção arbitrária $\mathbf{v} \in \mathbb{R}^q$, sendo $\mathbf{v} \neq \mathbf{0}$, consideremos o caminho retilíneo diferenciável $\alpha : (-\varepsilon, \varepsilon) \rightarrow U$, $\alpha(t) = \mathbf{x} + t\mathbf{v}$, com $\mathbf{x} \in U$, onde U é um aberto em \mathbb{R}^q . Seja G a superfície gerada pelo gráfico de $f : U \rightarrow \mathbb{R}$ e $\phi : U \rightarrow G$, $\phi(\mathbf{x}) = (\mathbf{x}, f(\mathbf{x}))$ uma parametrização de G . Notemos que a curva $\beta = \phi \circ \alpha : (-\varepsilon, \varepsilon) \rightarrow G$, $\beta(t) = \phi \circ \alpha(t)$ é tal que $\beta(0) = \phi(\mathbf{x}) \in G$ e $\beta'(0) = d\phi(\mathbf{v}) \in T_{\phi(\mathbf{x})}G$. Sabemos que se $\mathbf{w} \in T_{\phi(\mathbf{x})}G$, então $\mathbf{w} = d\phi(\mathbf{v})$ para algum $\mathbf{v} \in \mathbb{R}^q$. Consequentemente,

$$\|\mathbf{w}\| = \|d\phi(\mathbf{v})\| = \sqrt{\langle d\phi(\mathbf{v}), d\phi(\mathbf{v}) \rangle} = \sqrt{\|\mathbf{v}\|^2 + \langle \nabla_f, \mathbf{v} \rangle^2}. \quad (\text{B.18})$$

Em geral, se $\mathbf{u}, \mathbf{v} \in \mathbb{R}^q$, então

$$\langle d\phi(\mathbf{u}), d\phi(\mathbf{v}) \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \nabla_f, \mathbf{u} \rangle \langle \nabla_f, \mathbf{v} \rangle. \quad (\text{B.19})$$

Neste caso, a primeira forma fundamental associada à superfície G , obtida a partir do gráfico de f , é a transformação bilinear simétrica \mathbb{I} que associa a cada ponto $\mathbf{p} \in G$ um produto interno em $T_{\mathbf{p}}G$. Assim,

$$\mathbb{I}(\mathbf{u}, \mathbf{v}) = \mathbf{u}^\top (\mathbf{I}_q + \nabla_f^\top \nabla_f) \mathbf{v} = \mathbf{u}^\top \mathbf{v} + \langle \nabla_f, \mathbf{u} \rangle \langle \nabla_f, \mathbf{v} \rangle, \quad (\text{B.20})$$

onde \mathbf{I}_q denota a matriz identidade de ordem q . Em particular, se \mathbf{x}_0 é um ponto crítico de f , temos $\nabla_f = \mathbf{0}$ em \mathbf{x}_0 . Daí, a primeira forma fundamental associada a G no ponto $\mathbf{p} = \phi(\mathbf{x}_0)$ é tal que $\mathbb{I}(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle$. Além disso, a restrição de \mathbb{I} ao $T_{\phi(\mathbf{x}_0)}G$ acarreta que $\mathbb{I}(\mathbf{w}, \mathbf{w}) = \langle \mathbf{w}, \mathbf{w} \rangle = \|\mathbf{w}\|^2$.

Sejam \mathbf{v} e \mathbf{w} vetores tangentes a G em um mesmo ponto $\mathbf{p} = \phi(\mathbf{x}_0) \in G$ e \mathbf{n} o vetor normal. A segunda forma fundamental avaliada no par de vetores \mathbf{v} e \mathbf{w} é dada por

$$\mathbb{II}(\mathbf{v}, \mathbf{w}) = -\frac{1}{\|\mathbf{n}\|} \langle \nabla_{\mathbf{v}} \mathbf{n}, \mathbf{w} \rangle. \quad (\text{B.21})$$

Em particular, se $\mathbf{v} = d\phi(\mathbf{e}_j)$ e $\mathbf{w} = d\phi(\mathbf{e}_k)$, então

$$\langle \nabla_{d\phi(\mathbf{e}_j)} \mathbf{n}, d\phi(\mathbf{e}_k) \rangle = -\frac{\partial^2 f(\mathbf{x}_0)}{\partial x_j \partial x_k}.$$

Consequentemente, dado que $\mathcal{B} = \{\mathbf{d}\phi(\mathbf{e}_1), \dots, \mathbf{d}\phi(\mathbf{e}_q)\}$ é uma base $T_{\phi(\mathbf{x})}G$ e o produto interno é bilinear, segue que

$$\Pi(\mathbf{v}, \mathbf{w}) = -\frac{1}{\|\mathbf{n}\|} \langle \nabla_{\mathbf{v}} \mathbf{n}, \mathbf{w} \rangle = \frac{\mathbf{v}^\top \ddot{\mathbf{F}} \mathbf{w}}{\sqrt{1 + \|\nabla_f\|^2}}. \quad (\text{B.22})$$

B.0.5 A curvatura normal

Dada a superfície G , obtida a partir do gráfico de uma função diferenciável $f : U \subset \mathbb{R}^q \rightarrow \mathbb{R}$, consideremos a parametrização $\phi : U \rightarrow G$, $\phi(\mathbf{x}) = (\mathbf{x}, f(\mathbf{x}))^\top$. A curvatura normal de G em $\phi(\mathbf{x}) \in G$, na direção de um vetor não-nulo $\mathbf{v} \in \mathbb{R}^q$, é dada por

$$C_{\mathbf{v}} = \frac{\Pi(\mathbf{v}, \mathbf{v})}{\mathbb{I}(\mathbf{v}, \mathbf{v})} = \frac{\mathbf{v}^\top \ddot{\mathbf{F}} \mathbf{v}}{(\|\mathbf{v}\|^2 + \langle \nabla_f, \mathbf{v} \rangle^2) \sqrt{1 + \|\nabla_f\|^2}}. \quad (\text{B.23})$$

onde ∇_f denota o gradiente da função f e $\ddot{\mathbf{F}}$ a matriz Hessiana de f no ponto \mathbf{x} .

Restringindo-nos ao cálculo da curvatura ao longo de vetores não-nulos pertencentes ao espaço tangente a G no ponto $\phi(\mathbf{x})$, obtemos a seguinte aplicação

$$C : T_{\phi(\mathbf{x})}G \rightarrow \mathbb{R}, \quad C_{\mathbf{w}} = \frac{\Pi(\mathbf{w}, \mathbf{w})}{\mathbb{I}(\mathbf{w}, \mathbf{w})}. \quad (\text{B.24})$$

Neste ponto, observemos que se a é um número real não-nulo, então aplicando a definição (B.23) facilmente constata-se que $C_{a\mathbf{v}} = C_{\mathbf{v}}$ para todo vetor não-nulo $\mathbf{v} \in T_{\phi(\mathbf{x})}G$, ou seja, a curvatura normal não muda ao longo de uma dada direção \mathbf{v} . Isso permite restringirmos o cálculo da curvatura a vetores unitários pertencentes ao $T_{\phi(\mathbf{x})}G$. Dessa forma, consideramos a aplicação

$$C : \bar{T}_{\phi(\mathbf{x})}G \rightarrow \mathbb{R}, \quad C_{\mathbf{v}} = \frac{\Pi(\mathbf{v}, \mathbf{v})}{\mathbb{I}(\mathbf{v}, \mathbf{v})}, \quad (\text{B.25})$$

onde $\bar{T}_{\phi(\mathbf{x})}G$ denota o conjunto de vetores unitários tangentes a G no ponto $\phi(\mathbf{x})$.

Por fim, tendo em vista a análise de influência local, a ser considerada posteriormente na Seção 3.3, podemos particularizar o cálculo da curvatura normal do gráfico de f em pontos críticos da função f . Dessa forma, se \mathbf{x}_0 é um ponto crítico de f e \mathbf{v} é um vetor unitário não-nulo pertencente a $\bar{T}_{\phi(\mathbf{x})}G$, então $\nabla_f(\mathbf{x}_0)$ é o funcional identicamente nulo. Consequentemente,

$$C_{\mathbf{v}} = \left. \frac{\mathbf{v}^\top \ddot{\mathbf{F}} \mathbf{v}}{(\|\mathbf{v}\|^2 + \langle \nabla_f, \mathbf{v} \rangle^2) \sqrt{1 + \|\nabla_f\|^2}} \right|_{\mathbf{x}=\mathbf{x}_0} = \mathbf{v}^\top \ddot{\mathbf{F}}(\mathbf{x}_0) \mathbf{v}. \quad (\text{B.26})$$

Agora, fazemos uma conexão entre a curvatura normal de G em $\phi(\mathbf{x}_0)$, na direção do vetor não-nulo $\mathbf{v} \in \overline{T}_{\phi(\mathbf{x}_0)}G$, e a curvatura de uma curva plana, contida no plano que determina a seção normal $G(\mathbf{v})$ no ponto $\phi(\mathbf{x}_0)$. Dados um vetor unitário \mathbf{v} e um ponto crítico \mathbf{x}_0 de f , a seção normal $G(\mathbf{v})$, definida pela equação (B.14), pode ser vista como uma curva cujo gráfico é dado por

$$\phi(\mathbf{x}_0 + t\mathbf{v}) = \phi(\mathbf{x}_0) + t\mathbf{d}\phi(\mathbf{v}) + \nu(t)\mathbf{n}, \quad (\text{B.27})$$

onde $\nu(t) = f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)$, conforme atesta a equação (B.15).

Consideremos a curva contida no plano gerado pelos vetores $\mathbf{d}\phi(\mathbf{x}_0)$ e \mathbf{n} e que associa a cada $t \in (-\varepsilon, \varepsilon)$ o ponto $(t, \nu(t))$, onde $\nu(t) = f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)$. É imediato verificarmos que

$$\begin{aligned} \nu(0) &= 0 \text{ e } \nu'(0) = 0, \\ \frac{d\nu(t)}{dt} &= \sum_{j=1}^q \frac{\partial f}{\partial x_j}(\mathbf{x}_0 + t\mathbf{v})v_j, \end{aligned}$$

e

$$\frac{d^2\nu(t)}{dt^2} = \sum_{j,k=1}^q \frac{\partial^2 f}{\partial x_k \partial x_j}(\mathbf{x}_0 + t\mathbf{v})v_k v_j = \mathbf{v}^\top \ddot{\mathbf{F}}(\mathbf{x}_0 + t\mathbf{v})\mathbf{v}.$$

Logo,

$$\left. \frac{d^2\nu(t)}{dt^2} \right|_{t=0} = \mathbf{v}^\top \ddot{\mathbf{F}}(\mathbf{x}_0)\mathbf{v} = C_{\mathbf{v}}.$$

Portanto, a curvatura normal de G no ponto $\phi(\mathbf{x}_0)$ e na direção do vetor \mathbf{v} está estreitamente relacionada com a concavidade da seção normal $G(\mathbf{d}\phi(\mathbf{y}))$ no ponto $\phi(\mathbf{x}_0)$.

Vimos que se \mathbf{x}_0 é um ponto crítico de f então, a curvatura normal de G no ponto $\phi(\mathbf{x}_0)$ e na direção do vetor unitário \mathbf{v} é dada por

$$C_{\mathbf{v}} = \mathbf{v}^\top \ddot{\mathbf{F}}(\mathbf{x}_0)\mathbf{v}, \quad (\text{B.28})$$

onde $\ddot{\mathbf{F}}(\mathbf{x}_0)$ é a matriz da forma Hessiana de f no ponto \mathbf{x}_0 .

Nessas condições, a curvatura normal pode ser vista como uma forma quadrática definida em $\overline{T}_{\phi(\mathbf{x}_0)}G$ e cuja matriz simétrica associada é $\ddot{\mathbf{F}}(\mathbf{x}_0)$. Usando a Proposição 1.2.3, que consta em (POON; POON, 2010), segue a aplicação $\mathbf{v} \mapsto C_{\mathbf{v}}$ possui q pontos críticos, $\mathbf{v}_1, \dots, \mathbf{v}_q$, e existem escalares $\lambda_1, \dots, \lambda_q$ tais que

$$\|\mathbf{v}_j\| = 1; \quad \langle \mathbf{v}_j, \mathbf{v}_k \rangle = 0, \text{ se } j \neq k; \quad \ddot{\mathbf{F}}\mathbf{v}_j = \lambda_j \mathbf{v}_j. \quad (\text{B.29})$$

Consequentemente,

$$C_{\mathbf{v}_j} = \mathbf{v}_j^\top \ddot{\mathbf{F}}\mathbf{v}_j = \lambda_j \mathbf{v}_j^\top \mathbf{v}_j = \lambda_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle = \lambda_j. \quad (\text{B.30})$$

Em linguagem geométrica, os escalares, $\lambda_1, \dots, \lambda_q$, são chamados de curvaturas principais e os vetores, $\mathbf{v}_1, \dots, \mathbf{v}_q$, são denominados direções principais de G no ponto $\phi(\mathbf{x}_0)$. Para os nossos propósitos, interessa-nos identificar $\lambda_{\max} = \max_{1 \leq i \leq q} \lambda_i$, ou equivalentemente, C_{\max} .

Por fim, destaquemos que a curvatura média (\bar{C}) de G no ponto $\phi(\mathbf{x}_0)$ é dada por

$$\bar{C} = \frac{1}{q}(\lambda_1 + \dots + \lambda_q) = \frac{1}{q} \text{tr} \ddot{\mathbf{F}}(\mathbf{x}_0). \quad (\text{B.31})$$

APÊNDICE C – MÉTODOS DE GEOMETRIA DA INFORMAÇÃO

C.1 NOÇÕES BÁSICAS

C.1.1 Variedades topológicas

Intuitivamente, uma variedade diferenciável é um conceito que generaliza as noções de curva suave e superfície regular em espaços euclidianos. A idéia é considerar um conjunto arbitrário, denotado genericamente por \mathcal{S} , que admite um sistema de coordenadas. Por um sistema de coordenadas locais em um ponto $p \in \mathcal{S}$ entende-se uma aplicação um-a-um $\varphi : U \subset \mathcal{S} \mapsto \mathbb{R}^n$, definida em um aberto U , contendo o ponto, e que associa cada $p \in U$ a um vetor $\varphi(p) = \xi \in \mathbb{R}^n$. Um sistema de coordenadas cujo domínio é toda variedade \mathcal{S} é dito um sistema global. A síntese aqui exposta é baseada em (CARMO, 1998) e (AMARI; NAGOAKA, 2000).

Uma topologia em um conjunto \mathcal{S} é uma coleção \mathcal{T} de subconjuntos de \mathcal{S} , chamados de abertos segundo a topologia \mathcal{T} , que cumpre as seguintes condições:

1. $\emptyset \in \mathcal{T}$ e $\mathcal{S} \in \mathcal{T}$;
2. Se $\{A_\alpha\}_{\alpha \in \Gamma}$ é uma família qualquer de abertos pertencentes a topologia \mathcal{T} , então a reunião $\bigcup_{\alpha \in \Gamma} A_\alpha \in \mathcal{T}$;
3. Se A_1, \dots, A_n é uma família finita de abertos em \mathcal{T} , então a interseção $\bigcap_{i=1}^n A_i \in \mathcal{T}$.

Satisfeitas as condições 1, 2 e 3, o par $(\mathcal{S}, \mathcal{T})$ é dito um espaço topológico. Dado um ponto $\mathbf{p} \in \mathcal{S}$, qualquer aberto $V \in \mathcal{T}$ que contem \mathbf{p} é denominado vizinhança de \mathbf{p} . O espaço topológico \mathcal{S} é dito espaço de *Hausdorff* se dados dois pontos arbitrários $\mathbf{p} \neq \mathbf{q}$ em \mathcal{S} , existem abertos $A, B \subset \mathcal{S}$ tais que $\mathbf{p} \in A$, $\mathbf{q} \in B$ e $A \cap B = \emptyset$. Uma base de um espaço topológico $(\mathcal{S}, \mathcal{T})$ é uma coleção \mathcal{B} de abertos pertencentes a \mathcal{T} com a seguinte propriedade: todo aberto $A \subset \mathcal{S}$ se exprime como uma reunião $A = \bigcup B_\lambda$ de abertos $B_\lambda \in \mathcal{B}$.

Uma aplicação $\varphi : \mathcal{S}_1 \mapsto \mathcal{S}_2$, de um espaço topológico $(\mathcal{S}_1, \mathcal{T}_1)$ em um espaço topológico $(\mathcal{S}_2, \mathcal{T}_2)$, diz-se contínua em um ponto $\mathbf{p} \in \mathcal{S}_1$ quando, para cada aberto $B \subset \mathcal{S}_2$, com $\varphi(\mathbf{p}) \in B$, existe um aberto $A \subset \mathcal{S}_1$, com $\mathbf{p} \in A$, tal que $\varphi(A) \subset B$. Diz-se que $\varphi : \mathcal{S}_1 \mapsto \mathcal{S}_2$ é contínua quando a mesma é contínua em todos os pontos de \mathcal{S}_1 , ou equivalentemente, quando a imagem inversa $\varphi^{-1}(B)$ de todo aberto $B \subset \mathcal{S}_2$ é um aberto em \mathcal{S}_1 . No caso em

que $\varphi : \mathcal{S}_1 \mapsto \mathcal{S}_2$ é contínua, biunívoca e possui inversa $\varphi^{-1} : \mathcal{S}_2 \mapsto \mathcal{S}_1$ também contínua, diz-se que φ é um homeomorfismo de \mathcal{S}_1 em \mathcal{S}_2 .

Um espaço topológico \mathcal{S} diz-se localmente euclidiano de dimensão n se cada ponto $\mathbf{p} \in \mathcal{S}$ possui uma vizinhança U tal que existe um homeomorfismo φ de U em um subconjunto aberto de \mathbb{R}^n . Neste caso, dizemos que $(U, \varphi : U \mapsto \mathbb{R}^n)$ é uma carta local em \mathcal{S} e $\varphi : U \mapsto \mathbb{R}^n$ é um sistema de coordenadas locais em \mathcal{S} .

Definição. Uma variedade topológica n -dimensional é um espaço topológico \mathcal{S} que cumpre às seguintes condições:

- i) \mathcal{S} é um espaço de *Hausdorff*;
- ii) \mathcal{S} possui uma base \mathcal{B} enumerável;
- iii) \mathcal{S} é um espaço localmente euclidiano de dimensão n .

Suponha que $(U, \varphi : U \mapsto \mathbb{R}^n)$ e $(V, \phi : V \mapsto \mathbb{R}^n)$ são duas cartas em uma variedade topológica \mathcal{S} . Dado que $U \cap V = W$ é um aberto em U e φ é um homeomorfismo, a imagem $\varphi(W)$ é um aberto em \mathbb{R}^n . De modo similar, conclui-se que $\phi(W)$ também é um aberto em \mathbb{R}^n . Logo, faz sentido as seguintes definições:

Definição Duas cartas $(U, \varphi : U \mapsto \mathbb{R}^n)$ e $(V, \phi : V \mapsto \mathbb{R}^n)$ em uma variedade topológica \mathcal{S} são ditas cartas compatíveis se as aplicações $\phi \circ \varphi^{-1} : \varphi(W) \mapsto \phi(W)$ e $\varphi \circ \phi^{-1} : \phi(W) \mapsto \varphi(W)$ são aplicações de classe C^∞ .

Definição Um atlas, em um espaço localmente euclidiano \mathcal{S} , é uma coleção $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}$ de cartas compatíveis tais que $\bigcup_\alpha U_\alpha = \mathcal{S}$.

C.1.2 Variedades diferenciáveis

Seja \mathcal{S} uma variedade topológica n -dimensional e $\varphi : \mathcal{S} \mapsto \mathbb{R}^n$ um sistema de coordenadas global para \mathcal{S} . Logo, φ associa a cada ponto $\mathbf{p} \in \mathcal{S}$ o vetor n -dimensional dado por $\varphi(\mathbf{p}) = (\xi^1(\mathbf{p}), \dots, \xi^n(\mathbf{p})) = (\xi^1, \dots, \xi^n) = [\xi^i] \in \mathbb{R}^n$, chamado de coordenadas do ponto \mathbf{p} . Cada ξ^i pode ser visto como uma função a valores reais, $\mathbf{p} \mapsto \xi^i(\mathbf{p})$, que associa a cada ponto $\mathbf{p} \in \mathcal{S}$ a i -ésima coordenada de $\varphi(\mathbf{p})$. As n funções, $\xi^i : \mathcal{S} \mapsto \mathbb{R}$, $i = 1, 2, \dots, n$, são denominadas funções coordenadas.

Seja $\psi : \mathcal{S} \mapsto \mathbb{R}^n$ um outro sistema de coordenadas em $\mathbf{p} \in \mathcal{S}$ tal que $\psi(\mathbf{p}) = (\rho^1, \dots, \rho^n)$. Neste caso, existe uma aplicação um-a-um entre as coordenadas $\xi = (\xi^1, \dots, \xi^n)$ e $\rho = (\rho^1, \dots, \rho^n)$ dada por $\psi \circ \varphi^{-1} : \mathbb{R}^n \mapsto \mathbb{R}^n$ tal que $\psi \circ \varphi^{-1}(\xi) = \rho$.

Agora, apresentaremos a definição de variedade diferenciável para qual existe um sistema de coordenadas global. Seja \mathcal{S} uma variedade topológica e \mathcal{A} um conjunto de sistemas de coordenadas globais para \mathcal{S} , isto é, um atlas sobre \mathcal{S} que cumpre as seguintes condições

- i) Cada elemento $\varphi \in \mathcal{A}$ é uma aplicação um-a-um de \mathcal{S} em um subconjunto aberto $U \subset \mathbb{R}^n$.
- ii) Para todo $\varphi \in \mathcal{A}$ e qualquer aplicação um-a-um $\psi : \mathcal{S} \mapsto \mathbb{R}^n$ tem-se que $\psi \in \mathcal{A}$ se, e somente se $\psi \circ \varphi^{-1}$ é um difeomorfismo de classe C^∞ .

Diz-se que $(\mathcal{S}, \mathcal{A})$ é uma variedade diferenciável n -dimensional de classe C^∞ .

C.1.3 Aplicações diferenciáveis

Seja \mathcal{S} uma variedade diferenciável n -dimensional. Uma função $f : \mathcal{S} \mapsto \mathbb{R}$ é dita diferenciável em $\mathbf{p} \in \mathcal{S}$ se existe um sistema de coordenadas $\varphi : U \mapsto \mathbb{R}^n$ em \mathbf{p} tal que $f \circ \varphi^{-1} : U \subset \mathbb{R}^n \mapsto \mathbb{R}$ é diferenciável em \mathbf{p} . Diz-se que $f : \mathcal{S} \mapsto \mathbb{R}$ é diferenciável se f for diferenciável em todos os pontos de \mathcal{S} .

Observações:

- 1ª. A função $\bar{f} = f \circ \varphi^{-1}$ é denominada representação local de f ;
- 2ª. Se $\varphi : \mathcal{S} \mapsto \mathbb{R}^n$ é uma carta global para \mathcal{S} tal que $\varphi(\mathbf{p}) = (\xi^1, \dots, \xi^n)$, então $f(\mathbf{p}) = \bar{f}(\xi^1, \dots, \xi^n)$;
- 3ª. Quando $f \circ \varphi^{-1} = \bar{f}$ é uma função de classe C^∞ em \mathbb{R}^n , ou seja, possui derivadas parciais de todas as ordens e as mesmas forem funções contínuas, diz-se que f é de classe C^∞ em \mathcal{S} ;
- 4ª. O conjunto $\mathcal{F}(\mathcal{S}) = \{f : \mathcal{S} \mapsto \mathbb{R}; f \text{ é diferenciável de classe } C^\infty\}$, munido das operações usuais de adição de funções e multiplicação de função por um escalar, é um espaço vetorial de dimensão infinita.

Sejam \mathcal{S} e \mathcal{Q} variedades diferenciáveis cujas dimensões são m e n , respectivamente. Diz-se que uma aplicação $f : \mathcal{S} \mapsto \mathcal{Q}$ é diferenciável em um ponto $\mathbf{p} \in \mathcal{S}$ se existem sistemas de coordenadas $\varphi : U \mapsto \mathbb{R}^m$ em \mathcal{S} e $\psi : V \mapsto \mathbb{R}^n$ em \mathcal{Q} , com $\mathbf{p} \in U$ e $f(U) \subset V$ tais que $\psi \circ f \circ \varphi^{-1} : \varphi(U) \mapsto \psi(V)$ é diferenciável no ponto $\varphi(\mathbf{p})$.

Observações:

- 1ª. Se $f : \mathcal{S} \mapsto \mathcal{Q}$ é diferenciável em \mathbf{p} , então f é contínua em \mathbf{p} ;
- 2ª. A definição de diferenciabilidade independe da escolha do sistema de coordenadas;
- 3ª. Diz-se que $f : \mathcal{S} \mapsto \mathcal{Q}$ é diferenciável em \mathcal{S} quando for diferenciável em todos os pontos de \mathcal{S} . Além disso, diz-se que $f : \mathcal{S} \mapsto \mathcal{Q}$ é de classe C^∞ se $\psi \circ f \circ \varphi^{-1} : \varphi(U) \mapsto \psi(V)$ é de classe C^∞ ;
- 4ª. A composta de duas aplicações $f : \mathcal{S} \mapsto \mathcal{N}$ e $g : \mathcal{N} \mapsto \mathcal{Q}$ de classe C^∞ é também uma aplicação de classe C^∞ ;
- 5ª. Um difeomorfismo $f : \mathcal{S} \mapsto \mathcal{Q}$ é uma bijeção diferenciável cuja inversa $f^{-1} : \mathcal{Q} \mapsto \mathcal{S}$ é também diferenciável.

C.1.4 O espaço tangente

Sejam \mathcal{S} uma variedade diferenciável e $\lambda : (-\varepsilon, \varepsilon) \mapsto \mathcal{S}$ uma curva diferenciável em \mathcal{S} tal que $\lambda(0) = \mathbf{p}$. Denotemos por $\mathcal{F}(\mathcal{S})$ o conjunto das funções $f : \mathcal{S} \mapsto \mathbb{R}$ de classe C^∞ diferenciáveis em \mathbf{p} . O vetor tangente à curva λ em $t = 0$ é o funcional linear $\lambda'(0) : \mathcal{F} \mapsto \mathbb{R}$ definido por

$$\lambda'(0)f = \left. \frac{d(f \circ \lambda)}{dt} \right|_{t=0}, \quad f \in \mathcal{F}. \quad (\text{C.1})$$

Um vetor tangente à variedade \mathcal{S} em um ponto $\mathbf{p} \in \mathcal{S}$ é um vetor tangente a alguma curva $\lambda : (-\varepsilon, \varepsilon) \mapsto \mathcal{S}$ com $\lambda(0) = \mathbf{p}$. O conjunto dos vetores tangentes à variedade \mathcal{S} em \mathbf{p} é denominado espaço tangente de \mathcal{S} em \mathbf{p} e denotado por $T_{\mathbf{p}}(\mathcal{S})$.

De acordo com (LIMA, 2007), é possível dar a $T_{\mathbf{p}}(\mathcal{S})$ uma estrutura de espaço vetorial. Isso é feito da seguinte forma: cada sistema de coordenadas $\varphi : \mathcal{S} \mapsto \mathbb{R}^n$ em \mathcal{S} , com $\mathbf{p} \in \mathcal{S}$, dá origem a uma bijeção $\bar{\varphi} : T_{\mathbf{p}}(\mathcal{S}) \mapsto \mathbb{R}^n$, definida por $\bar{\varphi}(\dot{\lambda}) = (\varphi \circ \lambda)'(0)$. Daí, exige-se que a bijeção seja $\bar{\varphi} : T_{\mathbf{p}}(\mathcal{S}) \mapsto \mathbb{R}^n$ seja um isomorfismo. Com isso, as operações de adição e multiplicação de um vetor por um escalar são definidas por

$$\dot{\lambda} + \dot{\gamma} = (\bar{\varphi})^{-1}(\bar{\varphi}(\dot{\lambda}) + \bar{\varphi}(\dot{\gamma})), \quad \forall \dot{\lambda}, \dot{\gamma} \in T_{\mathbf{p}}(\mathcal{S})$$

e

$$c\dot{\lambda} = (\bar{\varphi})^{-1}(c\bar{\varphi}(\dot{\lambda})), \quad \forall c \in \mathbb{R}, \forall \dot{\lambda} \in T_{\mathbf{p}}(\mathcal{S}).$$

Dados um sistema de coordenadas locais $\varphi : U \mapsto \mathbb{R}^n$ em \mathcal{S} e um ponto $\mathbf{p} \in U \subset \mathcal{S}$ tal que $\varphi(\mathbf{p}) = (\xi^1, \dots, \xi^n)$, indica-se por $\left\{ \frac{\partial}{\partial \xi^1}(\mathbf{p}), \dots, \frac{\partial}{\partial \xi^n}(\mathbf{p}) \right\}$ a base de $T_{\mathbf{p}}(\mathcal{S})$ que é levada pelo isomorfismo $\bar{\varphi}$ sobre a base canônica $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. Daí segue que

$$T_{\mathbf{p}}(\mathcal{S}) = \left\{ \sum_{i=1}^n c^i \frac{\partial}{\partial \xi^i}(\mathbf{p}); c^i \in \mathbb{R} \right\}. \quad (\text{C.2})$$

C.1.5 Métricas riemannianas

De acordo com (CARMO, 2010), uma métrica riemanniana em uma variedade diferenciável \mathcal{S} é uma correspondência g que associa a cada ponto $\mathbf{p} \in \mathcal{S}$ um produto interno no espaço tangente $T_{\mathbf{p}}(\mathcal{S})$, que varia diferenciavelmente no seguinte sentido: se $x : U \subset \mathbb{R}^n \mapsto \mathcal{S}$ é um sistema de coordenadas em torno de \mathbf{p} , então para cada (i, j) , a função $g_{ij} : U \mapsto \mathbb{R}$ definida por

$$g_{ij}(x^1, \dots, x^n) = \left\langle \frac{\partial}{\partial x^i}(\mathbf{q}), \frac{\partial}{\partial x^j}(\mathbf{q}) \right\rangle_{\mathbf{q}}, \quad (\text{C.3})$$

onde $\mathbf{q} = x(x^1, \dots, x^n)$, é diferenciável em U .

Observações:

- 1ª. As funções g_{ij} são denominadas expressões da métrica riemanniana no sistema de coordenadas (x, U) . Uma variedade diferenciável munida com uma métrica riemanniana, (\mathcal{S}, g) , é dita variedade riemanniana;
- 2ª. A qualquer variedade diferenciável, \mathcal{S} , é possível associar uma métrica riemanniana.

C.1.6 Conexões afins e riemannianas

Um campo de vetores X em uma variedade diferenciável \mathcal{S} é uma correspondência que a cada ponto $\mathbf{p} \in \mathcal{S}$ associa um vetor $X(\mathbf{p}) \in T_{\mathbf{p}}(\mathcal{S})$. O campo X diz-se diferenciável se $X : \mathcal{S} \mapsto T_{\mathbf{p}}(\mathcal{S})$ é diferenciável. O conjunto dos campos de vetores em \mathcal{S} que são diferenciáveis é denotado por $\mathcal{T}(\mathcal{S})$.

Uma conexão afim em uma variedade diferenciável \mathcal{S} é uma aplicação $\nabla : \mathcal{T}(\mathcal{S}) \times \mathcal{T}(\mathcal{S}) \rightarrow \mathcal{T}(\mathcal{S})$, indicada por $(X, Y) \mapsto \nabla_X Y$, que cumpre as seguintes condições:

$$(i) \quad \nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z,$$

$$(ii) \quad \nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z,$$

$$(iii) \quad \nabla_X(fY) = f\nabla_X Y + X(f)Y,$$

onde $X, Y, Z \in \mathcal{T}(\mathcal{S})$ e $f, g \in \mathcal{F}(\mathcal{S})$.

Observações:

- 1ª. Seja X e Y campos vetoriais em uma variedade diferenciável \mathcal{S} . Existe um campo vetorial Z em \mathcal{S} tal que $Zf = (XY - YX)f$ para todo $f \in \mathcal{F}(\mathcal{S})$. Em geral, o campo vetorial Z é denominado colchete de *Lie* dos campos X e Y e denotado por $[X, Y]$. Dessa forma, escreve-se $[X, Y] = XY - YX$.
- 2ª. Uma conexão afim ∇ em uma variedade riemanniana $(\mathcal{S}, g = \langle \cdot, \cdot \rangle)$ é dita compátivel com a métrica g se $X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle$, para todo $X, Y, Z \in \mathcal{T}(\mathcal{S})$.
- 3ª. Uma conexão afim ∇ em uma variedade diferenciável \mathcal{S} diz-se simétrica se para todo $X, Y, Z \in \mathcal{T}(\mathcal{S})$ tem-se $\nabla_X Y - \nabla_Y X = XY - YX = [X, Y]$.
- 4ª. Seja (\mathcal{S}, g) uma variedade riemanniana. Se ∇ é uma conexão afim em \mathcal{S} que é simétrica e compátivel com a métrica g , diz-se que ∇ é conexão riemanniana.

C.2 A ESTRUTURA GEOMÉTRICA DE MODELOS ESTATÍSTICOS

O estudo geométrico-diferencial voltado para a Estatística considera o conjunto de distribuições de probabilidade que constituem o modelo estatístico como sendo uma variedade diferenciável e analisa as relações entre a estrutura geométrica desta variedade e a estimação usando tal modelo.

C.2.1 Modelos estatísticos

Consideremos uma família \mathcal{S} de distribuições de probabilidades definidas sobre \mathcal{X} . Admitamos que cada elemento de \mathcal{S} possa ser parametrizado por meio de um vetor n -dimensional

$\theta = (\theta^1, \theta^2, \dots, \theta^n)$. Assim,

$$\mathcal{S} = \{p_\theta = p(x; \theta); \theta \in \Theta\}, \quad (\text{C.4})$$

onde Θ é um subconjunto de \mathbb{R}^n e a aplicação $\theta \mapsto p_\theta$ é injetiva. Neste caso, diz-se que \mathcal{S} é um modelo estatístico paramétrico n -dimensional. Se p é uma distribuição em \mathcal{S} , denotemos o suporte de p_θ por $\text{supp}(p)$, logo $\text{supp}(p) = \{x \in \mathcal{X}; p(x) > 0\}$.

Em conformidade com (AMARI, 1985), suponhamos que o modelo estatístico $\mathcal{S}\{p(x; \theta)\}$ satisfaz as seguintes hipóteses:

- H1. O suporte das funções $p(x; \theta) \in \mathcal{S}$ é o mesmo, não dependendo do parâmetro θ ;
- H2. Para todo θ fixado, as funções $\frac{\partial}{\partial \theta^i} \log p(x; \theta)$, $i = 1, \dots, n$, são linearmente independentes;
- H3. Os momentos das variáveis aleatórias $\frac{\partial}{\partial \theta^i} \log p(X, \theta)$ existem até uma ordem suficientemente necessária;
- H4.

$$\int \partial_i p(x; \theta) dx = \partial_i \int p(x; \theta) dx$$

, ou seja, a ordem das operações de integração e derivação sobre os elementos $p(x; \theta)$ possa ser alterada.

Dado um modelo estatístico $\mathcal{S} = \{p_\theta; \theta \in \Theta\}$, a aplicação $\varphi : \mathcal{S} \mapsto \mathbb{R}^n$ tal que $\varphi(p_\theta) = \theta$ é um sistema de coordenadas para \mathcal{S} . Além disso, se $\psi : \Theta \mapsto \psi(\Theta)$ é uma aplicação um-a-um em que ψ e ψ^{-1} são de classe C^∞ , então pode-se usar $\rho = \psi(\theta)$ em vez de θ como sendo o vetor de parâmetros do modelo que passa a ser dado por $\mathcal{S} = \{p_{\psi^{-1}(\rho)}; \rho \in \psi(\Theta)\}$. Assim, o modelo \mathcal{S} pode ser visto como sendo uma variedade diferenciável, a qual denominamos de variedade estatística.

C.2.2 A métrica Fisher

Seja $\mathcal{S} = \{p_\theta; \theta \in \Theta\}$ um modelo estatístico n -dimensional. Dado um ponto $\theta \in \Theta$, a matriz de informação de Fisher de \mathcal{S} em θ é a matriz quadrada de ordem n , denotada por $\mathbf{G}(\theta) = (g_{ij}(\theta))$, cujo termo (i, j) é definido por

$$g_{ij}(\theta) = \mathbb{E}_\theta[\partial_i \ell_\theta \partial_j \ell_\theta] = \int \partial_i \ell(x; \theta) \partial_j \ell(x; \theta) p(x; \theta) dx, \quad (\text{C.5})$$

em que $\partial_i = \frac{\partial}{\partial \theta^i}$ e $\ell_\theta(x) = \ell(x; \theta) = \log p(x; \theta)$.

Observações:

- 1ª. Embora existam modelos para os quais a integral dada pela equação (C.5) diverge, nas discussões desta seção admite-se que $g_{ij}(\theta)$ é finito para todo $\theta \in \Theta$ e para quaisquer $i, j = 1, 2, \dots, n$. Além disso, assume-se que a aplicação $g_{ij} : \Theta \mapsto \mathbb{R}$ é de classe C^∞ .
- 2ª. Sob as hipóteses fixadas na seção anterior, mostra-se que

$$g_{ij}(\theta) = -\mathbb{E}_\theta (\partial_i \partial_j \ell_\theta)$$

e

$$g_{ij}(\theta) = 4 \int \partial_i \sqrt{p(x; \theta)} \partial_j \sqrt{p(x; \theta)} dx.$$

- 3ª. A matriz $\mathbf{G}(\theta) = (g_{ij}(\theta))$ é simétrica e positiva semi-definida. Porém, ao longo desta seção, assume-se que a mesma é positiva definida. Isso é equivalente a admitir que os elementos em $\{\partial_1 \ell_\theta, \dots, \partial_n \ell_\theta\}$, quando vistos como funções sobre \mathcal{X} , são linearmente independentes.

C.2.3 A α -conexão

Seja $\mathcal{S} = \{p_\theta; \theta \in \Theta\}$ um modelo estatístico n -dimensional. Consideremos a correspondência que associa a cada $\theta \in \Theta$ a função definida por

$$\left(\Gamma_{ij,k}^{(\alpha)} \right)_\theta = \mathbb{E}_\theta \left[\left(\partial_i \partial_j \ell_\theta + \frac{1-\alpha}{2} \partial_i \ell_\theta \partial_j \ell_\theta \right) (\partial_k \ell_\theta) \right], \quad (\text{C.6})$$

sendo α um número real arbitrário e $i, j, k \in \{1, 2, \dots, n\}$. Consequentemente, defini-se uma conexão afim, $\nabla^{(\alpha)}$, sobre \mathcal{S} , denominada α -conexão, por

$$\langle \nabla_{\partial_i}^{(\alpha)} \partial_j, \partial_k \rangle = \Gamma_{ij,k}^{(\alpha)}, \quad (\text{C.7})$$

onde $g = \langle \cdot, \cdot \rangle$ é a métrica Fisher.

Observações:

- 1ª. A α -conexão é uma conexão simétrica.
- 2ª. Para quaisquer $\alpha, \beta \in \mathbb{R}$, tem-se a seguinte relação:

$$\Gamma_{ij,k}^{(\beta)} = \Gamma_{ij,k}^{(\alpha)} + \frac{\alpha - \beta}{2} \mathbf{T}_{ijk}, \quad (\text{C.8})$$

onde \mathbf{T}_{ijk} é um tensor simétrico covariante de grau 3 definido por

$$(\mathbf{T}_{ijk})_{\theta} = \mathbb{E} [\partial_i \ell_{\theta} \partial_j \ell_{\theta} \partial_k \ell_{\theta}].$$

3ª. A α -conexão admite as seguintes combinações lineares convexas:

$$\nabla^{(\alpha)} = (1 - \alpha)\nabla^{(0)} + \alpha\nabla^{(1)} = \frac{1 + \alpha}{2}\nabla^{(1)} + \frac{1 - \alpha}{2}\nabla^{(-1)}. \quad (\text{C.9})$$

Teorema C.1 (AMARI; NAGOAKA, 2000) *A 0-conexão é uma conexão riemanniana com respeito à métrica Fisher.*

APÊNDICE D – PROVAS

D.1 JUSTIFICATIVA DAS PROPRIEDADES DO CAPÍTULO 2

Prova da Propriedade 1 (Seção 2.7.2) A identificação de máximos locais de f pode ser feita por meio do estudo do sinal das derivadas da função $g = \log(f)$. Se $\gamma = 0$, então para todo $y \in \mathbb{R}$, $g(y)$ é dada por

$$g(y) = -\log(\alpha) + \log\left(\cosh\left(\frac{y-\eta}{2}\right)\right) - \frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right) - \exp\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)\right). \quad (\text{D.1})$$

Notemos que g é derivável em toda reta real e sua derivada primeira, g' , pode ser escrita na forma

$$g'(y) = \frac{1}{2} \cdot \frac{\sinh\left(\frac{y-\eta}{2}\right)}{\cosh\left(\frac{y-\eta}{2}\right)} - \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right) \left[1 - \exp\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)\right)\right]. \quad (\text{D.2})$$

Neste ponto, observemos que, independente do valor de α , tem-se $g'(\eta) = 0$. No que tange à derivada segunda de g , obtemos

$$g''(y) = \frac{1}{4} - \frac{1}{4} \tanh^2\left(\frac{y-\eta}{2}\right) - \frac{1}{2\alpha} \sinh\left(\frac{y-\eta}{2}\right) + \frac{1}{\alpha} \exp\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)\right) \left[\frac{1}{2} \sinh\left(\frac{y-\eta}{2}\right) - \frac{1}{\alpha} \cosh^2\left(\frac{y-\eta}{2}\right)\right].$$

Logo, $g''(\eta) = \frac{1}{4} - \frac{1}{\alpha^2}$. Levando-nos as seguintes conclusões:

- i) Se $0 < \alpha < 2$, então $g''(\eta) < 0$.
- ii) Se $\alpha = 2$, então $g''(\eta) = 0$.
- iii) Se $\alpha > 2$, então $g''(\eta) > 0$.

Portanto, η é um máximo local de f quando $0 < \alpha < 2$. Por outro lado, se $\alpha > 2$, então η é um mínimo local de f .

A seguir, mostraremos que se $0 < \alpha < 2$, então g é estritamente crescente no intervalo $(-\infty, \eta)$. Com efeito, para todo número real z tem-se $\cosh(z) \geq 1$. Consequentemente,

$$\frac{1}{2} \geq \frac{1}{2} \frac{1}{\cosh^2\left(\frac{y-\eta}{2}\right)}. \quad (\text{D.3})$$

A hipótese $0 < \alpha < 2$ implica

$$\frac{1}{2} < \frac{1}{\alpha}. \quad (\text{D.4})$$

A partir das desigualdades (D.3) e (D.4), agregando o fato de que a função $\cosh(\cdot)$ é estritamente positiva, obtemos

$$\frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right) > \frac{1}{2 \cosh\left(\frac{y-\eta}{2}\right)}. \quad (\text{D.5})$$

Agora, observemos que quando $y < \eta$, temos $\sinh\left(\frac{y-\eta}{2}\right) < 0$, ou equivalentemente, $-\sinh\left(\frac{y-\eta}{2}\right) > 0$. Daí, a partir da desigualdade (D.5) concluímos que

$$-\frac{1}{\alpha} \sinh\left(\frac{y-\eta}{2}\right) \cosh\left(\frac{y-\eta}{2}\right) > -\frac{1}{2} \frac{\sinh\left(\frac{y-\eta}{2}\right)}{\cosh\left(\frac{y-\eta}{2}\right)}. \quad (\text{D.6})$$

Por outro lado, os pressupostos $y < \eta$ e $0 < \alpha < 2$ acarretam $\sinh\left(\frac{y-\eta}{2}\right) < 0$ e $-\frac{2}{\alpha} < -1$, respectivamente. Daí,

$$-\frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right) > -\sinh\left(\frac{y-\eta}{2}\right). \quad (\text{D.7})$$

Uma vez que a função exponencial é crescente, temos

$$\exp\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)\right) > \exp\left(-\sinh\left(\frac{y-\eta}{2}\right)\right). \quad (\text{D.8})$$

Neste ponto, ao usarmos o fato de que $e^x > 1 + x$ para todo $x > 0$ sobre o membro do lado direito da desigualdade (D.8) estabelecemos que

$$\exp\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)\right) > 1 - \sinh\left(\frac{y-\eta}{2}\right). \quad (\text{D.9})$$

Ao multiplicarmos ambos os membros da desigualdade (D.9) por $\frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right)$, obtemos

$$\frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right) \exp\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)\right) > \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right) \left(1 - \sinh\left(\frac{y-\eta}{2}\right)\right). \quad (\text{D.10})$$

Assim, a partir de (D.6) e (D.10), concluímos que é válida a seguinte desigualdade:

$$\frac{1}{2} \cdot \frac{\sinh\left(\frac{y-\eta}{2}\right)}{\cosh\left(\frac{y-\eta}{2}\right)} - \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right) \left[1 - \exp\left(-\frac{2}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)\right)\right] > 0$$

Do exposto acima, fica assegurado que g , e portanto f , é estritamente crescente em $(-\infty, \eta)$.

Prova da Propriedade 2 (Seção 2.7.2). Investiguemos o comportamento da função $g(y) = \log(f(y))$ cuja derivada primeira, quando $\gamma \neq 0$, admite ser escrita na forma

$$g'(y) = \frac{1}{2} \tanh\left(\frac{y-\eta}{2}\right) + \left[-(1+\gamma) + \left(1 + \frac{2\gamma}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)\right)^{-1/\gamma} \right] \left(\frac{\frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right)}{1 + \frac{2\gamma}{\alpha} \sinh\left(\frac{y-\eta}{2}\right)} \right), \quad (\text{D.11})$$

Se usarmos a hipótese adicional que $\gamma < -1$, podemos justificar que g é estritamente crescente no intervalo aberto $(\eta, \eta + 2\operatorname{arcsinh}(-\alpha/2\gamma))$. De fato, observemos que se $\gamma < -1$, então

$$\left[-(1 + \gamma) + \left(1 + \frac{2\gamma}{\alpha} \operatorname{senh} \left(\frac{y - \eta}{2} \right) \right)^{-1/\gamma} \right] \left(\frac{\frac{1}{\alpha} \cosh \left(\frac{y - \eta}{2} \right)}{1 + \frac{2\gamma}{\alpha} \operatorname{senh} \left(\frac{y - \eta}{2} \right)} \right) > 0, \quad (\text{D.12})$$

pois $\cosh(\frac{y-\eta}{2}) > 0$ para todo $y \in \mathbb{R}$, $1 + \frac{2\gamma}{\alpha} \operatorname{senh}(\frac{y-\eta}{2})$ e $-(1 + \gamma)$ são positivos por hipótese. Agora, usando o fato de que $\tanh(\frac{y-\eta}{2})$ é positiva para todo $y > \eta$, tem-se $g'(y) > 0$ para todo $y \in (\eta, \eta + 2\operatorname{arcsinh}(-\alpha/2\gamma))$, isto implica que g é estritamente crescente nesse intervalo. Visto que a função logarítmica é estritamente crescente, tal propriedade pode ser estendida de g para f .

Seja $y_0 = \eta + 2 \operatorname{arcsinh}(-\alpha/2\gamma)$, notemos que

$$\lim_{y \rightarrow y_0^-} 1 + \frac{2\gamma}{\alpha} \operatorname{senh} \left(\frac{y - \eta}{2} \right) = 0.$$

Logo,

$$\lim_{y \rightarrow y_0^-} f(y) = +\infty.$$

Assegurando-nos que f é ilimitada superiormente quando $\gamma < -1$.

Prova da Propriedade 3 (Seção 2.7.2). Sob a hipótese de que $\gamma = -1$, a função densidade de Y assume a forma

$$f(y) = \frac{1}{\alpha} \cosh \left(\frac{y - \eta}{2} \right) \exp \left(- \left(1 - \frac{2}{\alpha} \operatorname{senh} \left(\frac{y - \eta}{2} \right) \right) \right), \quad y < \eta + 2 \operatorname{arcsinh}(\alpha/2).$$

Consequentemente,

$$g(y) = \log(f(y)) = -\log(\alpha) + \log \left(\cosh \left(\frac{y - \eta}{2} \right) \right) - 1 + \frac{2}{\alpha} \operatorname{senh} \left(\frac{y - \eta}{2} \right).$$

Notemos que g é derivável e sua derivada primeira é

$$g'(y) = \frac{1}{2} \frac{\operatorname{senh} \left(\frac{y - \eta}{2} \right)}{\cosh \left(\frac{y - \eta}{2} \right)} + \frac{1}{\alpha} \cosh \left(\frac{y - \eta}{2} \right). \quad (\text{D.13})$$

Assim, ao usarmos a identidade $\cosh^2(z) - \operatorname{senh}^2(z) = 1$, válida para todo $z \in \mathbb{R}$, obtemos

$$g'(y) = 0 \iff \frac{1}{\alpha} \operatorname{senh}^2 \left(\frac{y - \eta}{2} \right) + \frac{1}{2} \operatorname{senh} \left(\frac{y - \eta}{2} \right) + \frac{1}{\alpha} = 0 \quad (\text{D.14})$$

A fim de que a equação $g'(y) = 0$ possua solução, faz-se necessário que $\alpha \geq 4$. Nessas condições, temos as raízes

$$y_1 = \eta + 2\operatorname{arcsenh} \left(-\frac{\alpha}{4} - \frac{1}{4} \sqrt{\alpha^2 - 16} \right)$$

e

$$y_2 = \eta + 2\operatorname{arcsinh}\left(-\frac{\alpha}{4} + \frac{1}{4}\sqrt{\alpha^2 - 16}\right)$$

Em resumo, sob a hipótese de que $\alpha \geq 4$, temos:

- i) $g'(y) < 0$ quando $y_1 < y < y_2$,
- ii) $g'(y) = 0$ quando $y = y_1$ ou $y = y_2$,
- iii) $g'(y) > 0$ quando $y < y_1$ ou $y > y_2$.

Em outros termos, f é crescente nos intervalos $(-\infty, \eta + 2\operatorname{arcsinh}\left(-\frac{\alpha}{4} - \frac{1}{4}\sqrt{\alpha^2 - 16}\right))$ e $(\eta + 2\operatorname{arcsinh}\left(-\frac{\alpha}{4} + \frac{1}{4}\sqrt{\alpha^2 - 16}\right), \eta + 2\operatorname{arcsinh}(-\alpha/2\gamma))$. E, decrescente no intervalo $(\eta + 2\operatorname{arcsinh}\left(-\frac{\alpha}{4} - \frac{1}{4}\sqrt{\alpha^2 - 16}\right), \eta + 2\operatorname{arcsinh}\left(-\frac{\alpha}{4} + \frac{1}{4}\sqrt{\alpha^2 - 16}\right))$, isso quando $\alpha \geq 4$ e $\gamma = -1$. Além disso, $\lim_{y \rightarrow y_0^-} f(y) = \frac{1}{\alpha} \sqrt{1 + \frac{\alpha^2}{4}}$.

Agora, passemos a lidar com o caso $0 < \alpha < 2$. Para todo $y \in \mathbb{R}$, tem-se

$$-1 < \frac{\sinh\left(\frac{y-\eta}{2}\right)}{\cosh\left(\frac{y-\eta}{2}\right)} < 1.$$

Consequentemente,

$$-\frac{1}{2} + \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right) < \frac{1}{2} \frac{\sinh\left(\frac{y-\eta}{2}\right)}{\cosh\left(\frac{y-\eta}{2}\right)} + \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right) < \frac{1}{2} + \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right). \quad (\text{D.15})$$

A partir de (D.15) e (D.13), obtemos

$$-\frac{1}{2} + \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right) < g'(y) < \frac{1}{2} + \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right). \quad (\text{D.16})$$

Por outro lado, se $0 < \alpha < 2$, então $0 < -\frac{1}{2} + \frac{1}{\alpha}$. Além disso, usando o fato de que $\cosh\left(\frac{y-\eta}{2}\right) \geq 1$ para todo $y \in \mathbb{R}$, obtemos a seguinte desigualdade:

$$0 < -\frac{1}{2} + \frac{1}{\alpha} \leq -\frac{1}{2} + \frac{1}{\alpha} \cosh\left(\frac{y-\eta}{2}\right). \quad (\text{D.17})$$

Logo, $g'(y) > 0$. Isso implica que f é estritamente crescente quando $0 < \alpha < 2$ e $\gamma = -1$.

Por fim, tratemos do caso $2 \leq \alpha < 4$. Aqui vale destacar que g' não muda de sinal quando $\alpha \in (0, 4)$. De fato, suponhamos por contradição que existem t_1 e t_2 pertencentes ao intervalo $(-\infty, \eta_i + 2\operatorname{arcsinh}(-\alpha/2\gamma))$ tais que $g'(t_1) < 0$ e $g'(t_2) > 0$. Dado que g' é contínua, o

teorema do valor intermediário nos assegura a existência de $t_0 \in (-\infty, \eta_i + 2\text{arcsinh}(-\alpha/2\gamma))$ tal que $g'(t_0) = 0$. Isso não é possível, pois g' só admite pontos críticos quando $\alpha \geq 4$.

Prova da Propriedade 4 (Seção 2.7.2) Se $X \sim \text{GEV}(0, 1, \gamma)$ e $\gamma \neq 0$, então a fda de X é

$$G_\gamma(x) = \exp\left(-(1 + \gamma x)^{-\frac{1}{\gamma}}\right), \quad 1 + \gamma x > 0. \quad (\text{D.18})$$

Observemos que a função de distribuição, F_Y , da variável $Y = \eta + 2\text{arcsinh}(\frac{\alpha X}{2})$ está relacionada com a função G_γ por meio da expressão

$$F_Y(y) = G_\gamma\left(\frac{2}{\alpha}\text{sinh}\left(\frac{y - \eta}{2}\right)\right),$$

de modo que $1 + \gamma\frac{2}{\alpha}\text{sinh}\left(\frac{y - \eta}{2}\right) > 0$. Consequentemente, a função densidade de Y , f_Y , é dada por

$$f_Y(y) = \frac{d}{dy}G_\gamma\left(\frac{2}{\alpha}\text{sinh}\left(\frac{y - \eta}{2}\right)\right) = \frac{1}{2}\xi_1(y)(1 + \gamma\xi_2(y))^{-1-\frac{1}{\gamma}}\exp(-(1 + \gamma\xi_2(y))^{-\frac{1}{\gamma}}).$$

Portanto, $Y = \eta + 2\text{arcsinh}(\frac{\alpha X}{2}) \sim \text{log-EVBS}(\alpha, \eta, \gamma)$. A justificativa do caso em que $\gamma = 0$ é simples e inteiramente análoga.

D.2 JUSTIFICATIVAS DAS PROPRIEDADES DO CAPÍTULO 5

A partir deste ponto, exprimimos as provas dos itens 1, 2, 3 e 4 da propriedade 7, exibida na Seção 5.2.1.

1. Suponhamos que $T \sim \text{GPD-BS}(\alpha, \beta, \gamma)$ com $\gamma > 0$. Logo,

$$\mathbb{E}\left[1 + \frac{\gamma}{\alpha}\varphi\left(\frac{T}{\beta}\right)\right]^{-r} = \int_\beta^\infty \frac{1}{\alpha\beta}\varphi'(t/\beta)\left(1 + \frac{\gamma}{\alpha}\varphi(t/\beta)\right)^{-1-(1/\gamma)-r} dt. \quad (\text{D.19})$$

Ao fazermos a mudança de variável $u = 1 + \frac{\gamma}{\alpha}\varphi(t/\beta)$, obtemos

$$\mathbb{E}\left[1 + \frac{\gamma}{\alpha}\varphi\left(\frac{T}{\beta}\right)\right]^{-r} = \int_1^\infty \frac{1}{\gamma}u^{-1-(1/\gamma)-r} du = \lim_{x \rightarrow \infty} \left(-\frac{1}{1 + \gamma r}\right) \left(\frac{1}{x^{(1/\gamma)+r}} - 1\right) = \frac{1}{1 + \gamma r}, \quad (\text{D.20})$$

quando $(1/\gamma) + r > 0$.

2. Se $\gamma > 0$, então

$$\mathbb{E}\left(\log\left[1 + \frac{\gamma}{\alpha}\varphi\left(\frac{T}{\beta}\right)\right]\right)^k = \int_\beta^\infty \left(\log\left[1 + \frac{\gamma}{\alpha}\varphi(t/\beta)\right]\right)^k \frac{1}{\alpha\beta}\varphi'(t/\beta)\left(1 + \frac{\gamma}{\alpha}\varphi(t/\beta)\right)^{-1-(1/\gamma)} dt. \quad (\text{D.21})$$

As mudanças de variáveis $u = \log \left(1 + \frac{\gamma}{\alpha} \varphi(t/\beta) \right)$ e, em seguida, $z = \frac{u}{\gamma}$ acarretam

$$\mathbb{E} \left(\log \left[1 + \frac{\gamma}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right] \right)^k = \int_0^\infty u^k \frac{1}{\gamma} e^{-\frac{u}{\gamma}} du = \int_0^\infty \gamma^k z^{(k+1)-1} e^{-z} dz = \gamma^k \Gamma(k+1). \quad (\text{D.22})$$

Usando o Princípio da Indução Matemática, mostra-se que $\Gamma(k+1) = k!$ para todo $k \in \mathbb{N}$. Portanto,

$$\mathbb{E} \left(\log \left[1 + \frac{\gamma}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right] \right)^k = \gamma^k k!. \quad (\text{D.23})$$

3. Observemos que

$$\mathbb{E} \left[\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \left(\overline{G} \left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right) \right)^r \right] = \int_\beta^\infty \frac{1}{\alpha^2 \beta} \varphi(t/\beta) \varphi'(t/\beta) \left[1 + \frac{\gamma}{\alpha} \varphi(t/\beta) \right]^{-1 - \left(\frac{1+r}{\gamma} \right)} dt. \quad (\text{D.24})$$

Ao fazermos a mudança de variável $u = 1 + \frac{\gamma}{\alpha} \varphi(t/\beta)$, obtemos

$$\mathbb{E} \left[\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \left(\overline{G} \left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right) \right)^r \right] = \int_1^\infty (u-1) \frac{1}{\gamma^2} u^{-1 - \left(\frac{1+r}{\gamma} \right)} du. \quad (\text{D.25})$$

Se $\frac{1+r}{\gamma} > 0$, tem-se assegurada a convergência da integral imprópria e, conseqüentemente,

$$\begin{aligned} \mathbb{E} \left[\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \left(\overline{G} \left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right) \right)^r \right] &= \frac{1}{\gamma^2} \lim_{x \rightarrow \infty} \left[\frac{\gamma}{\gamma - (1+r)} \left(\frac{1}{x^{\frac{1+r}{\gamma} + 1}} - 1 \right) + \right. \\ &\quad \left. + \frac{\gamma}{(1+r)} \left(\frac{1}{x^{\frac{1+r}{\gamma}}} - 1 \right) \right]. \end{aligned}$$

Portanto,

$$\mathbb{E} \left[\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \left(\overline{G} \left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right) \right)^r \right] = \frac{1}{(r+1-\gamma)(r+1)} \quad (\text{D.26})$$

4. No caso em que $\gamma = 0$, temos

$$\mathbb{E} \left[\left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right)^r \right] = \int_\beta^\infty \left(\frac{1}{\alpha} \varphi(t/\beta) \right)^r \frac{1}{\alpha \beta} \varphi'(t/\beta) \exp \left(-\frac{1}{\alpha} \varphi(t/\beta) \right) dt. \quad (\text{D.27})$$

A mudança de variável $u = \frac{1}{\alpha} \varphi(t/\beta)$ implica

$$\mathbb{E} \left[\left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right)^r \right] = \int_0^\infty u^r \exp(-u) du = \Gamma(r+1). \quad (\text{D.28})$$

Se $r \in \mathbb{N}$, então $\Gamma(r+1) = r!$. Logo,

$$\mathbb{E} \left[\left(\frac{1}{\alpha} \varphi \left(\frac{T}{\beta} \right) \right)^r \right] = r!. \quad (\text{D.29})$$