

**Jaine de Moura Carvalho**

**Distribuição Slashed Lomax: Medidas de Qualidade de Ajuste com Base na Transformada de Mellin**

**Recife**

**28 de fevereiro de 2023**



**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

**Distribuição Slashed Lomax: Medidas de Qualidade de Ajuste com Base na  
Transformada de Mellin**

Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 28/02/2023 pela comissão examinadora.

**Área de concentração: Biometria e  
Estatística Aplicada**

**Orientador: Dr. Frank Sinatra Gomes  
da Silva**

**Coorientador: Dr. Josimar Mendes de  
Vasconcelos**

**Recife**

**28 de fevereiro de 2023**

Dados Internacionais de Catalogação na Publicação  
Universidade Federal Rural de Pernambuco  
Sistema Integrado de Bibliotecas  
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

---

C331d

Carvalho, Jaine de Moura

Distribuição Slashed Lomax: Medidas de Qualidade de Ajuste com Base na Transformada de Mellin / Jaine de Moura Carvalho. - 2023.

43 f. : il.

Orientador: Frank Sinatra Gomes da Silva.

Coorientador: Josimar Mendes de Vasconcelos.

Inclui referências e apêndice(s).

Dissertação (Mestrado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Biometria e Estatística Aplicada, Recife, 2023.

1. Classe Slashed. 2. Log-Cumulantes. 3. Medidas de bondade de ajuste. 4. Estatística  $T^2$  de Hotelling. 5. Bootstrap. I. Silva, Frank Sinatra Gomes da, orient. II. Vasconcelos, Josimar Mendes de, coorient. III. Título

CDD 519.5

---

Jaine de Moura Carvalho

## **Distribuição Slashed Lomax: Medidas de Qualidade de Ajuste com Base na Transformada de Mellin**

Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 28/02/2023 pela comissão examinadora.

Banca examinadora:

---

**Dr. Frank Sinatra Gomes da Silva**  
Universidade Federal Rural de Pernambuco  
Orientador

---

**Dr. Josimar Mendes de Vasconcelos**  
Universidade Federal Rural de Pernambuco  
Coorientador

---

**Dr. Jhonnata Bezerra de Carvalho**  
Universidade Federal do Amazonas  
Examinador externo

Recife

28 de fevereiro de 2023

*Este trabalho é dedicado às crianças  
adultas que, quando pequenas, sonharam em  
se tornar cientistas.*

# Agradecimentos

Gratidão primeiramente a Deus, que mais uma vez me fortaleceu e guiou meus caminhos para que eu pudesse concluir esse desafio.

Agradeço aos meus pais, João e Teresinha, por me concederem à vida e pelos ensinamentos transmitidos.

Aos meus amigos, os velhos e os novos, pelo apoio incessante nesta etapa. Ao meu companheiro de vida, Acyllon Wilk, por todo amor e compreensão demonstrada, sobretudo quando estive ausente.

A todos os meus professores da Universidade Federal do Piauí - CSHNB e Universidade Federal Rural de Pernambuco, em especial, Dr. Josimar Mendes de Vasconcelos e Dr. Alex Sandro Lopes Santos, por nunca medirem esforços para me ajudar desde a graduação até o presente momento.

Aos meus estimados orientadores, Dr. Frank Sinatra Gomes da Silva e Dr. Josimar Mendes de Vasconcelos por tamanha competência, dedicação, esforço e pelos ensinamentos repassados durante esse período.

Gratidão ao Programa de Pós-Graduação em Biometria e Estatística Aplicada, a Universidade Federal Rural de Pernambuco e todos os colegas de departamento, em especial, Isaac, Jonnathan, Taís, Marília e Natália pela troca de conhecimento e companheirismo durante a jornada do mestrado.

*“Não vos amoldeis às estruturas deste mundo,  
mas transformai-vos pela renovação da mente,  
a fim de distinguir qual é a vontade de Deus:  
o que é bom, o que Lhe é agradável, o que é perfeito.  
(Bíblia Sagrada, Romanos 12, 2)*

# Resumo

Recentemente, várias distribuições de probabilidade foram propostas buscando alcançar resultados satisfatórios, em especial, modelos com maior flexibilidade e capazes de modelar dados de tempo de duração de componentes ou tempo de vida de indivíduos. Dentre estes, os modelos da classe Slashed, em particular, a distribuição Slashed Lomax, que é um modelo assimétrico, definido para valores reais positivos, destacando-se por possuir representação estocástica e por se ajustar a conjunto de dados de cauda pesada. Apesar desse crescente número de novos modelos probabilísticos, que correspondam as especificidades das amostras, poucas ferramentas estatísticas para avaliar a qualidade do ajuste têm sido introduzidas para as distribuições recém-desenvolvidas. Diante desse déficit, utilizamos a metodologia apresentada em Nicolas (2002) e, empregamos as estatísticas do segundo tipo (log-cumulantes) que são definidas a partir da Transformada de Mellin (TM), para fornecer novas medidas de qualidade de ajuste considerando aspectos qualitativos e quantitativos, para a distribuição Slashed Lomax. Derivamos a expressão da TM para a distribuição Slashed Lomax, calculamos os log-cumulantes (LCs) e construímos o diagrama de LCs ( $\tilde{k}_3, \tilde{k}_2$ ). Posteriormente, através de uma combinação entre a estatística  $T^2$  de Hotelling e o método delta multivariado, foi proposta uma estatística de teste para testar hipóteses sobre os LCs. Por fim, dois bancos de dados reais no contexto da análise de sobrevivência foram submetidos a metodologia proposta, afim de demonstrar a eficiência dos critérios de ajuste desenvolvidos. Experimentos de *bootstrap* foram realizados para avaliar o poder do teste da medida de bondade de ajuste quantitativa, como também verificar o desempenho dos estimadores do método dos log-cumulantes (MLC), método dos momentos (MM) e método de máxima verossimilhança (MV) em relação a essa ferramenta. As análises evidenciaram um bom desempenho das ferramentas de ajuste e a efetividade do MLC enquanto critério de estimação.

**Palavras-chave:** Classe Slashed, Log-Cumulantes, Medidas de bondade de ajuste, Estatística  $T^2$  de Hotelling, *Bootstrap*.



# Abstract

Recently, various probability distributions have been proposed to achieve satisfactory results, specifically models with increased flexibility that can model data on the duration of components or the lifetime of individuals. Among these, the Slashed class models, particularly the Slashed Lomax distribution, have gained attention. This asymmetric model is defined for positive real values, and is notable for its stochastic representation and ability to fit heavy-tailed data sets. Despite the increasing number of new probabilistic models that cater to specific samples, there have been few statistical tools introduced to evaluate their goodness of fit. To address this deficit, we employed the methodology outlined in Nicolas (2002) and utilized second-type statistics (log-cumulative) derived from the Mellin Transform (MT) to provide new measures of goodness of fit for the Slashed Lomax distribution. These measures consider both qualitative and quantitative aspects. We derived the MT expression for the Slashed Lomax distribution, calculated the log-cumulants (LCs) and created the log-cumulative diagram  $(\tilde{k}_3, \tilde{k}_2)$ . Then, we proposed a test statistic using a combination of Hotelling's  $T^2$  statistic and the multivariate Delta method to test hypotheses about the LCs. Finally, we applied the proposed methodology to two real databases in the context of survival analysis to show its effectiveness in evaluating the fit criteria. We conducted bootstrap experiments to assess the power of the proposed test and to evaluate the performance of the estimators using the log-cumulative method (MLC), method of moments (MM), and maximum likelihood method (ML). The results revealed that the adjustment tools performed well, and that the MLC proved to be an effective estimation criterion.

**Keywords:** Slashed class, Log-Cumulants, Goodness-of-fit measures, Hotelling's  $T^2$  statistic, Bootstrap.

# Lista de figuras

Figura 1 – Fluxograma ilustrando o estudo desenvolvido . . . . .	5
Figura 2 – Gráficos da fdp Slomax $(\alpha, \beta, \lambda)$ . . . . .	8
Figura 3 – Diagrama dos LCs $(\tilde{k}_3, \tilde{k}_2)$ para o modelo Slashed Lomax . . . . .	20
Figura 4 – Histogramas e densidade estimada . . . . .	25
Figura 5 – Diagrama $(\tilde{k}_3, \tilde{k}_2)$ e os LCs amostrais estimados . . . . .	27
Figura 6 – Elipses de confiança para os conjunto de dados . . . . .	29

## Lista de tabelas

Tabela 1 – Estatísticas descritivas dos dados . . . . .	24
Tabela 2 – Estimativas, estatísticas $T^2$ de Hotelling e p-valores para os conjuntos de dados . . . . .	28
Tabela 3 – Poder de teste da estatística $T^2$ de Hotelling . . . . .	31

# Lista de abreviaturas e siglas

AIC	Critério de Informação de Akaike
BIC	Critério de Informação Bayesiano
fdp	Função densidade de probabilidade
LCs	Log-Cumulantes
LMs	Log-Momentos
MLC	Método dos Log-Cumulantes
MM	Método dos Momentos
MV	Método de Máxima Verossimilhança
PolSAR	Polarimetric Synthetic Aperture Radar
SAR	Synthetic Aperture Radar
TM	Transformada de Mellin

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação	1
1.2	Objetivos	3
1.2.1	Objetivo geral	3
1.2.2	Objetivos específicos	3
1.3	Contribuições	4
1.4	Plataforma computacional	4
1.5	Organização da dissertação	4
<b>2</b>	<b>Revisão de literatura</b>	<b>6</b>
2.1	A distribuição Slashed Lomax	6
2.2	Método de máxima verossimilhança	9
2.3	Método dos momentos	9
2.4	Transformada de Mellin	10
2.5	Método dos log-cumulantes	11
2.6	Estatística $T^2$ de Hotelling	12
2.7	Método <i>bootstrap</i>	13
2.8	Poder do teste	13
<b>3</b>	<b>Distribuição Slashed Lomax: contribuições</b>	<b>15</b>
3.1	Momento de ordem $r$	15
3.2	Estimação pelo método dos momentos	15
3.3	Estimação de máxima verossimilhança	16
3.4	Transformada de Mellin	17
3.5	Diagrama de log-cumulantes	17
3.6	Teste de bondade de ajuste para a distribuição Slashed Lomax	21
<b>4</b>	<b>Aplicação</b>	<b>24</b>
4.1	Conjuntos de dados e estatísticas descritivas	24
4.2	Diagrama $(\tilde{k}_3, \tilde{k}_2)$	25
4.3	Estatística $T^2$ de Hotelling	28
4.4	Elipses de confiança	28
4.5	<i>Bootstrap</i> da estatística de teste	29
<b>5</b>	<b>Considerações finais</b>	<b>32</b>

Referências bibliográficas . . . . .	33
APÊNDICE A Primeiro apêndice . . . . .	38
A.1 Convergência da integral . . . . .	38
APÊNDICE B Segundo apêndice . . . . .	39
B.1 Log-verossimilhança . . . . .	39
B.2 Vetor escore . . . . .	39
APÊNDICE C Terceiro apêndice . . . . .	41
C.1 Estatística $T^2$ de Hotelling . . . . .	41
C.2 Cálculo das estatísticas . . . . .	43

# 1 Introdução

## 1.1 Motivação

Recentemente, várias distribuições da classe Slashed têm sido consideradas na literatura estatística, a exemplo os modelos: Slashed Rayleigh ([IRIARTE et al., 2015](#)), Slashed Rayleigh Exponenciada ([SALINAS; IRIARTE; BOLFARINE, 2015](#)), Slashed Exponencial Generalizada ([ASTORGA; GÓMEZ; BOLFARINE, 2017](#)), Slashed Exponencial de Momento ([IRIARTE et al., 2017](#)), Slashed Gumbel ([GÓMEZ; BOLFARINE; GÓMEZ, 2018](#)), Slashed Exponencial Generalizada Modificada ([ASTORGA et al., 2020](#)), Slashed Lomax ([LI; TIAN, 2020](#)) e Slashed Power Alfa-normal ([BARRIOS et al., 2022](#)). Temos também a família de distribuições Rayleigh Generalizada com base no modelo Slashed Modificado, que reúne o modelo Slashed Rayleigh Generalizado Modificado com seus casos particulares: Slashed Maxwell Modificado e Slashed Qui-quadrado Modificado ([BARRANCO-CHAMORRO et al., 2021](#)), além das distribuições Slashed Rayleigh Modificada ([IRIARTE et al., 2018](#)) e Slashed Alfa-normal Modificada ([OLMOS et al., 2019](#)). Tal classe de distribuições é formada por modelos probabilísticos não encaixados, determinados por representações estocásticas. Característica que torna esses modelos mais flexíveis do que as distribuições-base das quais foram obtidos.

Além disso, as distribuições do tipo Slashed fornecem um bom ajuste quando aplicadas para descrever fenômenos e analisar dados no contexto da análise de sobrevivência. Em [Olmos et al. \(2019\)](#), dados de falha do sistema de ar condicionado de um avião são modelados pela distribuição Slashed Alfa-normal Modificada. [Barranco-Chamorro et al. \(2021\)](#) empregaram o modelo Slashed Rayleigh Generalizado Modificado para modelar dados que descrevem os tempos de remissão (em meses) de pacientes com câncer de bexiga. Enquanto [Barrios et al. \(2022\)](#) trabalharam com dados de vida útil de tensão-ruptura do Kevlar 49/epóxi, que está sujeito a pressão constante a 90% de nível de tensão até a falha e dados de tempos de sobrevivência de cobaias injetadas com diferentes doses de bacilos da tuberculose, em que são ajustados sob a distribuição Slashed Power Alfa-normal.

A distribuição Slashed Lomax é caso especial dentre os modelos probabilísticos da classe Slashed. É um modelo assimétrico, definido para valores reais positivos, destacando-se por se adequar a conjunto de dados de cauda pesada. [Li e Tian \(2020\)](#) desenvolveram várias propriedades dessa distribuição. Por exemplo, as funções de densidade e taxa de risco,

a entropia de Rényi, os momentos de ordem  $r$  e os coeficientes de assimetria e curtose. Além disso, os autores estudaram também o modelo de regressão Slashed Lomax e apresentaram como obter os estimadores de máxima verossimilhança para os parâmetros da distribuição. Por fim, concluíram que, quando aplicado a dados de falha sucessivas dos sistemas de ar condicionado de aviões a jato Boeing 720 e dados de tempo de vida de pacientes com insuficiência cardíaca, o modelo Slashed Lomax forneceu o melhor ajuste comparado com as distribuições beta Lomax (LEMONTE; CORDEIRO, 2013), Exponencial Lomax Padrão (LEMONTE; CORDEIRO, 2013) e Gumbel Lomax (TAHIR et al., 2016).

A escolha do modelo é um fator importante da análise de dados, uma vez que o sucesso da modelagem de dados está diretamente relacionado a qualidade do ajuste da distribuição em relação a um determinado grupo de observações. Desse modo, é essencial ter um amplo conjunto de ferramentas que qualifiquem e/ou quantifiquem o ajuste, indicando assim, a adequação ou discrepância do modelo em teste. São muitos os estudos concentrados na geração de novas distribuições que se ajustam às particularidades das amostras. Apesar disso, poucas medidas de bondade de ajuste têm sido propostas para essas novas distribuições, com o propósito de criar novos critérios de seleção de modelo.

Nesse contexto, Chu, Dickin e Nadarajah (2019) fornecem um estudo com mais de vinte testes conhecidos de bondade de ajuste para as distribuições generalizadas de Pareto, Pareto tipo I e Pareto tipo II, dentre estes, testes como Kolmogorov Smirnov, Cramér-von Mises e Anderson-Darling. Os critérios de ajuste supracitados verificam a adequação do modelo com base nas medidas de distância entre as funções de distribuição acumulada e empírica, que nem sempre possui solução analítica.

Outros indicadores de seleção de modelos amplamente utilizados são os Critério de Informação de Akaike (AIC) (AKAIKE, 1974) e Critério de Informação Bayesiano (BIC) (SCHWARZ, 1978). Em Khalil et al. (2021) uma nova distribuição da classe Weibull é introduzida, a distribuição de Weibull Aditiva Flexível. Para validação desse modelo, a qualidade do seu ajuste quando aplicado a dados reais é comparada com a distribuição Weibull, distribuição Weibull Aditiva e a distribuição Weibull Aditiva Transmutada utilizando critérios de bondade de ajuste, incluindo os critérios AIC e BIC. A formulação original desses critérios não é necessariamente de fácil implementação devido ao cálculo da verossimilhança dos modelos que podem apresentar expressões analiticamente intratáveis (ROSSI et al., 2020).

Um método clássico para seleção de modelos e avaliação de bondade de ajuste é o diagrama de Pearson (PEARSON, 1895) que se baseia nas medidas de assimetria e curtose, discutido em Delignon, Garello e Hillion (1997) e Tison et al. (2004). Todavia, Nicolas (2002) verificou que o diagrama de Pearson pode ser mal adaptado para variáveis aleatórias



positivas e propôs um método alternativo, empregando as estatísticas do segundo tipo (log-cumulantes), que são definidas a partir da TM. Nessa nova abordagem introduziu o diagrama de LCs, uma alternativa gráfica para seleção de modelos. A nova ferramenta de ajuste, apesar de similar, se sobressai em relação o diagrama de Pearson por se adequar a variáveis definidas para valores reais positivos, possuir implementação computacional prática e capta a flexibilidade da distribuição no sentido de assimetria e curtose (LI et al., 2011). Uma descrição detalhada dessa metodologia baseada na TM é dada em Nicolas (2002). Em seus estudos, Bombrun, Anfinson e Harant (2011), Khan e Guida (2013), Deng e López-Martínez (2016) e Deng et al. (2018) empregaram o diagrama de LCs como ferramenta para classificar distribuições para dados de imagens *Polarimetric Synthetic Aperture Radar* (PolSAR). Em Anfinson, Doulgeris e Eltoft (2011) os LCs do tipo Mellin são utilizados para fornecer testes de qualidade de ajuste com hipóteses simples e compostas, aplicados a dados de imagens *Synthetic Aperture Radar* (SAR). Enquanto em Vasconcelos, Cintra e Nascimento (2021), tais estatísticas são aplicadas para obter medidas de bondade de ajuste qualitativas e quantitativas para seleção de modelos da classe beta-G.

## 1.2 Objetivos

### 1.2.1 Objetivo geral

Apresentar uma técnica de estimação independente da função de verossimilhança e propor medidas de bondade de ajuste considerando aspectos qualitativos e quantitativos para a distribuição Slashed Lomax.

### 1.2.2 Objetivos específicos

- Desenvolver a TM;
- Construir os LCs do segundo tipo;
- Elaborar o diagrama de LCs enquanto medida de bondade de ajuste qualitativa;
- Formular uma estatística de teste com base na estatística  $T^2$  de Hotelling e nos LCs como medida de bondade de ajuste quantitativa;
- Ilustrar as ferramentas desenvolvidas em aplicações com dados reais no contexto da análise de sobrevivência;
- Quantificar, por meio de simulação, o poder do teste da medida de bondade de ajuste quantitativa.
- Comparar o efeito do uso das estimativas pelo MLC, MM e MV para a estatística de teste.

### 1.3 Contribuições

Neste trabalho, aplicamos as estatísticas do tipo Mellin para fornecer novas medidas de bondade de ajuste para a distribuição Slashed Lomax, nas quais consideramos aspectos qualitativos e quantitativos. Expressões de forma fechada para a TM e para os LCs do segundo tipo são introduzidas para esse modelo. Apresentamos o diagrama de LCs enquanto medida de bondade de ajuste qualitativa. Para mais, é gerada uma estatística de teste a partir da inserção dos cumulantes do segundo tipo na estatística  $T^2$  de Hotelling, fornecendo assim uma medida de bondade de ajuste quantitativa. Adicionalmente, uma junção entre a estatística  $T^2$  de Hotelling e método delta é apresentada para determinar as elipses de confiança para os conjuntos de dados. Dois conjuntos de dados reais no contexto da análise de sobrevivência são submetidos à metodologia proposta. Para quantificar a eficiência da ferramenta de ajuste quantitativa, estudos de *bootstrap* são realizados.

### 1.4 Plataforma computacional

Os resultados que apresentamos nesta dissertação foram obtidos utilizando o *software* estatístico R, versão 4.1.0 (R Core Team, 2021). As estimativas dos parâmetros do modelo de probabilidade Slashed Lomax foram obtidas aplicando o pacote `BB` e a função `BBsolve` para os métodos de log-cumulantes e dos momentos; para o MV utilizou-se a função `maxLik` do pacote `maxLik`.

Empregamos o *software* WolframAlpha<sup>1</sup> para auxiliar na obtenção de algumas expressões. O texto foi escrito usando o editor *on-line* de texto em LaTeX, Overleaf<sup>2</sup>.

### 1.5 Organização da dissertação

O restante desta dissertação está organizado da seguinte forma:

- No Capítulo 2, apresentamos um resumo com as principais ferramentas estatísticas da distribuição Slashed Lomax. Revisamos também os métodos de estimação de máxima verossimilhança e dos momentos, a TM, o MLC e a estatística  $T^2$  de Hotelling, além do poder de teste de uma estatística de teste. No capítulo 3 a TM e suas propriedades são desenvolvidas. Apresentamos também novas medidas de qualidade de ajuste para o modelo Slashed Lomax. No Capítulo 4 é feita uma aplicação das ferramentas desenvolvidas em dois conjuntos de dados reais no contexto da análise de sobrevivência. Resultados numéricos

<sup>1</sup> Para obter mais informações sobre esse *software*, visite <https://www.wolframalpha.com>.

<sup>2</sup> Para mais informações e detalhes sobre esse editor de texto, acesse <https://pt.overleaf.com>.

são exibidos. Por fim, as conclusões e trabalhos futuros são apresentadas no Capítulo 5. A Figura 1 ilustra todo o arcabouço do nosso estudo nessa dissertação.

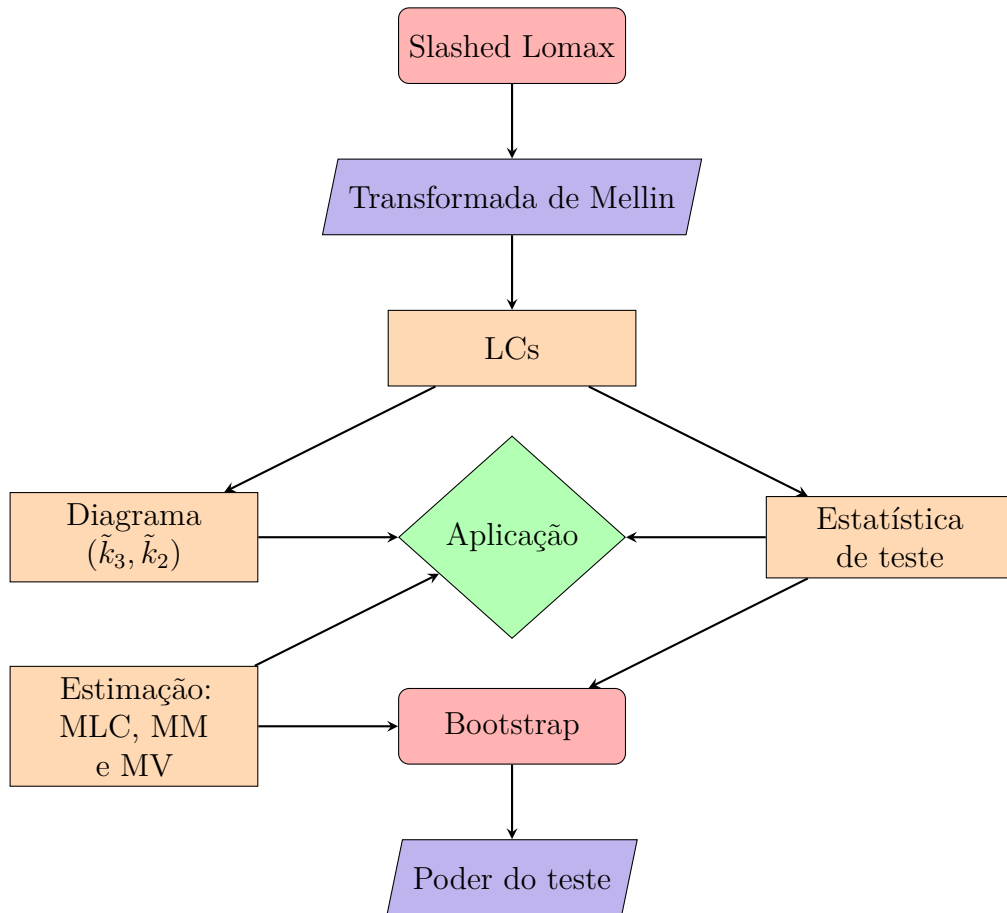


Figura 1 – Fluxograma ilustrando o estudo desenvolvido

## 2 Revisão de literatura

### 2.1 A distribuição Slashed Lomax

Nesta seção, descrevemos as principais propriedades da distribuição Slashed Lomax ilustrando-a graficamente para observarmos sua flexibilidade. Sendo assim, temos a primeira definição da construção do modelo de distribuição de probabilidade.

**Definição 1.** Uma variável aleatória  $Y$  segue uma distribuição Slashed Lomax, denotada como  $Y \sim \text{Slomax}(\alpha, \beta, \lambda)$ , se ela dispõe de uma representação estocástica dada por (LI; TIAN, 2020):

$$Y = XV^{-\frac{1}{2}},$$

em que  $X \sim \text{Lomax}(\alpha, \beta)$  e  $V \sim \text{beta}(\lambda, 1)$  são variáveis aleatórias independentes com  $\alpha, \beta, \lambda > 0$ .

**Definição 2.** Sendo  $Y \sim \text{Slomax}(\alpha, \beta, \lambda)$ , sua função densidade de probabilidade (fdp) é definida por (LI; TIAN, 2020):

$$f_Y(y; \alpha, \beta, \lambda) = \frac{2\lambda\alpha\beta^{2\lambda}}{y^{2\lambda+1}} \int_0^{\frac{y}{\beta}} \frac{w^{2\lambda}}{(1+w)^{\alpha+1}} dw, \quad \forall y \in \mathbb{R}^+.$$

**Corolário 1.** Seja  $Y \sim \text{Slomax}(\alpha, \beta, \lambda)$ , então:

i) Para  $\frac{y}{\beta} < 1$ ,

$$f_Y(y; \alpha, \beta, \lambda) = \frac{2\lambda\alpha}{(2\lambda+1)\beta} F\left(\alpha+1, 2\lambda+1, 2\lambda+2; -\frac{y}{\beta}\right),$$

em que

$$F(\alpha, \beta, \gamma; x) = \frac{1}{B(\beta, \gamma-\beta)} \int_0^1 (1-t)^{\gamma-\beta-1} t^{\beta-1} (1-xt)^{-\alpha} dt,$$

denota a função hipergeométrica (LI; TIAN, 2020).

ii) Sendo  $2\lambda$  e  $\alpha - 2\lambda$  números inteiros positivos e,  $0 < \frac{y}{\beta} < 1$ ,

$$f_Y(y; \alpha, \beta, \lambda) = \frac{2\lambda\alpha}{(2\lambda + 1)\beta} B\left(\frac{y}{\beta}; 2\lambda + 1, \alpha - 2\lambda\right),$$

em que

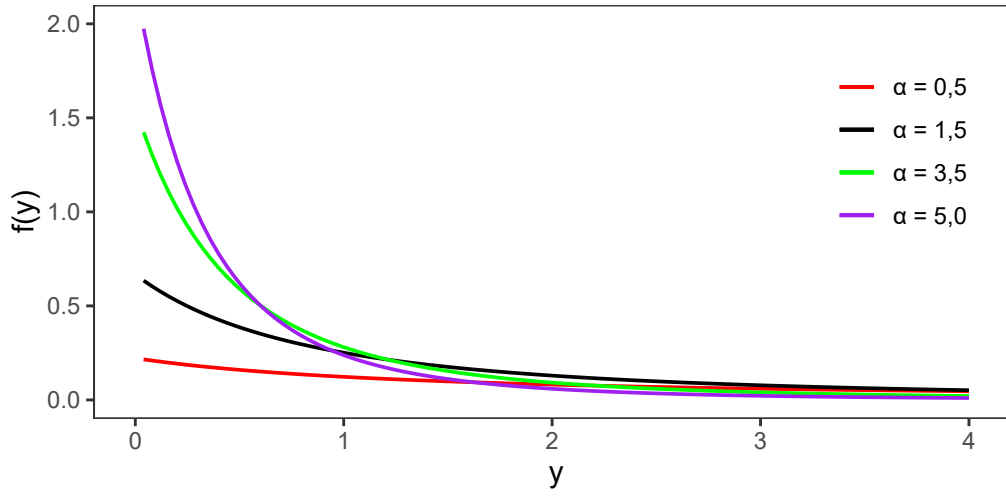
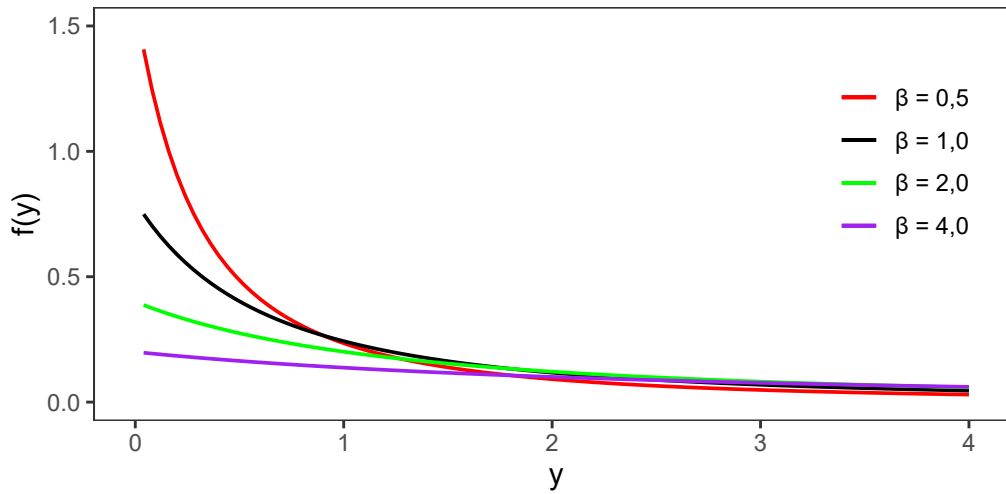
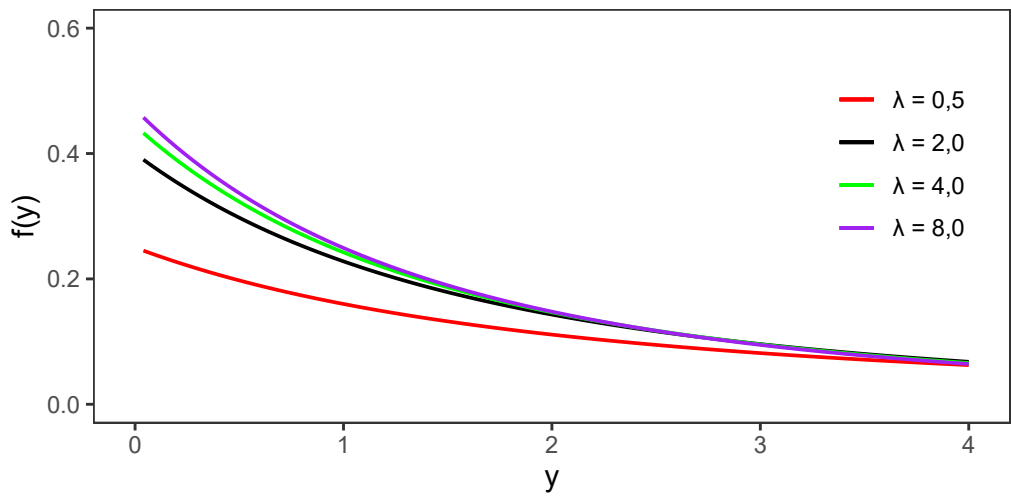
$$B(x; a, b) = \int_0^x t^{a-1} (1-t)^{b-1} dt,$$

é a função beta incompleta (LI; TIAN, 2020).

iii) Para  $\lambda = 1$ , a distribuição de  $Y$  se reduz a distribuição Slashed Lomax canônica, com fdp dada por (LI; TIAN, 2020):

$$f_Y(y; \alpha, \beta, \lambda) = \int_0^1 \frac{\alpha v^{\frac{1}{2}}}{\beta \left(1 + \frac{v^{\frac{1}{2}} y}{\beta}\right)^{\alpha+1}} dv.$$

A Figura 2 ilustra a curva da fdp da distribuição Slashed Lomax com diferentes valores de  $\alpha$ ,  $\beta$  e  $\lambda$ . Ao fixarmos dois desses parâmetros, identificamos o comportamento de cauda mais pesada à medida que  $\beta$  ou  $\lambda$  aumenta, enquanto que o oposto acontece para  $\alpha$ .

(a) Slomax ( $\alpha, 1,5, 1$ )(b) Slomax ( $1, \beta, 2$ )(c) Slomax ( $2, 4, \lambda$ )Figura 2 – Gráficos da fdp Slomax ( $\alpha, \beta, \lambda$ )

## 2.2 Método de máxima verossimilhança

Sejam  $X_1, \dots, X_n$  uma amostra aleatória da variável aleatória  $X$  com fdp ou função de probabilidade  $f(x; \theta_1, \dots, \theta_k)$ , de modo que a função de verossimilhança correspondente à amostra aleatória observada  $\mathbf{x} = (x_1, \dots, x_n)$  é dada por  $L(\theta_1, \dots, \theta_k; \mathbf{x})$ .

O estimador de máxima verossimilhança de  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$  é uma solução das equações (CASELLA; BERGER, 2001):

$$\frac{\partial L(\theta_1, \dots, \theta_k; \mathbf{x})}{\partial \theta_1} = 0,$$

$$\frac{\partial L(\theta_1, \dots, \theta_k; \mathbf{x})}{\partial \theta_2} = 0,$$

$$\vdots$$

$$\frac{\partial L(\theta_1, \dots, \theta_k; \mathbf{x})}{\partial \theta_k} = 0.$$

Como visto, o estimador de máxima verossimilhança de  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$  é o valor  $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_k)$  no espaço paramétrico  $\Theta$  que maximiza a função de verossimilhança. Porém, essa maximização pode acontecer utilizando o log  $L(\boldsymbol{\theta}; \mathbf{x})$  (BOLFARINE; SANDOVAL, 2010). Note que a função logarítmica é monótona crescente. Então, qualquer valor de  $\boldsymbol{\theta}$  que maximiza a função de log-verossimilhança dada por  $l(\boldsymbol{\theta}; \mathbf{x}) = \log L(\boldsymbol{\theta}; \mathbf{x})$  também maximiza  $L(\boldsymbol{\theta}; \mathbf{x})$ .

Consequentemente, temos a função escore definida da seguinte forma:

$$U(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta}; \mathbf{x})}{\partial \boldsymbol{\theta}},$$

no qual obtêm-se o vetor escore dado por

$$U(\theta_1, \dots, \theta_k) = \left( \frac{\partial l(\boldsymbol{\theta}; \mathbf{x})}{\partial \theta_1}, \dots, \frac{\partial l(\boldsymbol{\theta}; \mathbf{x})}{\partial \theta_k} \right)^\top.$$

Dessa forma, o estimador de  $\hat{\boldsymbol{\theta}}$  é determinado igualando os elementos do vetor escore a zero.

## 2.3 Método dos momentos

O MM é um dos métodos de estimação mais simples e um dos mais antigos que existe (BOLFARINE; SANDOVAL, 2010).

Sejam  $X_1, \dots, X_n$  uma amostra aleatória da variável aleatória  $X$  com fdp ou função de probabilidade  $f(x; \theta_1, \dots, \theta_k)$  e seu  $r$ -ésimo momento amostral definido por:

$$m_r = \frac{1}{n} \sum_{i=1}^n X_i^r, \quad r \geq 1.$$

Considere ainda,

$$\mu_r = E[X^r],$$

$r \geq 1$ , o  $r$ -ésimo momento populacional. Em geral,  $\mu_r$  é uma função conhecida dos parâmetros  $\theta_1, \dots, \theta_k$ , denotada por  $\mu_r(\theta_1, \dots, \theta_k)$ . O MM consiste na obtenção dos estimadores para  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$  resolvendo-se o seguinte sistemas de equações (CASSELLA; BERGER, 2001):

$$m_r = \mu_r(\theta_1, \dots, \theta_k), \quad (2.1)$$

$r = 1, \dots, k$ .

## 2.4 Transformada de Mellin

Proposta por Haljar Mellin (1854-1933), a TM emprega, na sua formulação, a mesma axiomática da estatística tradicional (baseada na transformada de Fourier) para definir as funções características, momentos e cumulantes (NICOLAS, 2002). Dessa forma, produz expressões que permitem uma análise mais prática das distribuições definidas em  $\mathbb{R}_+$  (NICOLAS, 2002). Formalmente, Nicolas (2006) expressa a TM pela relação entre uma função  $f(u)$  definida sobre o eixo real positivo e uma função  $\phi(s)$ , com  $s \in \mathbb{C}$ , que é dada por:

$$\phi(s) = \mathcal{M}[f(u)](s) = \int_0^{\infty} u^{s-1} f(u) du = E(U^{s-1}), \quad (2.2)$$

denominada de primeira função característica do segundo tipo.

A TM é determinada de forma análoga a derivada da expressão para a transformada de Fourier (função característica ordinária) na escala logaritmica (KRYLOV et al., 2013). Isto é, dada uma variável aleatória  $Y$  com fdp definida para valores reais positivos, expressa por  $f_Y(u)$  e, considerando a transformada de Fourier de  $f_Y(u)$  denotada como  $\Phi_{\log X}(v)$ , então (NICOLAS, 2002; KRYLOV et al., 2013):

$$\Phi_{\log X}(v) = \phi_X(s)|_{s=1+iv}, \quad v \in \mathbb{R}. \quad (2.3)$$

A relação apresentada em (2.3) é a TM de  $f_Y(u)$  com  $s = 1 + iv$ .



A segunda função característica do segundo tipo, referida como  $\psi(s)$ , é determinada empregando o logaritmo natural na primeira função característica do segundo tipo (NICOLAS, 2006), assim:

$$\psi(s) = \log(\phi(s)).$$

## 2.5 Método dos log-cumulantes

Por analogia às estatísticas clássicas, os log-momentos (LMs) e os LCs de ordem  $\nu$  são construídos por meio das derivadas da primeira e segunda função característica do segundo tipo, no ponto  $s = 1$  (NICOLAS, 2002). Assim,  $\forall \nu \in \mathbb{N}$ , suas respectivas expressões são dadas por:

$$\tilde{m}_\nu = \left. \frac{d^\nu \phi(s)}{ds^\nu} \right|_{s=1} = \int_0^\infty (\log u)^\nu f(u) du$$

e

$$\tilde{k}_\nu = \left. \frac{d^\nu \psi(s)}{ds^\nu} \right|_{s=1}. \quad (2.4)$$

Existe uma relação na qual os LCs podem ser determinados a partir de “polinômios” de LMs e vice-versa (NICOLAS, 2002; ANFINSEN; ELTOFT, 2011; ANFINSEN; DOULGERIS; ELTOFT, 2011). Dessa forma, os três primeiros LCs podem ser descritos como:

$$\begin{aligned} \tilde{k}_1 &= \tilde{m}_1 \\ \tilde{k}_2 &= \tilde{m}_2 - \tilde{m}_1^2 \\ \tilde{k}_3 &= \tilde{m}_3 - 3\tilde{m}_1\tilde{m}_2 + 2\tilde{m}_1^3. \end{aligned}$$

Resultando dessa relação, uma expressão geral para os LCs de ordem  $\nu$  é fornecida (ANFINSEN; DOULGERIS; ELTOFT, 2011):

$$\tilde{k}_\nu = \tilde{m}_\nu - \sum_{i=1}^{\nu-1} \binom{\nu-1}{i-1} \tilde{k}_i \tilde{m}_{\nu-i}. \quad (2.5)$$

Para obter os LCs amostrais, substituímos  $\tilde{m}_\nu$  pelos LMs da amostra dados por

$$\widehat{\tilde{m}}_\nu = \frac{1}{n} \sum_{i=1}^n (\log x_i)^\nu,$$

com  $x_i$  sendo a  $i$ -ésima observação amostral e  $n$  o tamanho da amostra.

A estimação pelo MLC ocorre de forma similar ao MM. No MLC os estimadores dos parâmetros do modelo probabilístico são obtidos resolvendo o sistema de equações formado pela igualdade entre as expressões dos LCs amostrais e populacionais.

## 2.6 Estatística $T^2$ de Hotelling

A estatística  $T^2$  de Hotelling é uma generalização multivariada da estatística (univariada)  $t$  de Student (ANDERSON, 2003). Essa ferramenta tem a aplicabilidade para realizar testes de hipóteses sobre a média (JOHNSON; WICHERN, 2007). Neste caso o teste de hipóteses será sobre os LCs.

Definido um vetor de variáveis aleatórias  $\mathbf{X} = (X_1, \dots, X_n)$  tal que  $\mathbf{X}$  é um vetor Normal  $\nu$ -dimensional  $\mathbf{X} \sim N_\nu(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , em que  $\boldsymbol{\mu} = E(\mathbf{X})$  e  $\boldsymbol{\Sigma} = E(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^\top$  denotam o vetor de médias populacional e a matriz de covariância, respectivamente. A estatística  $T^2$  de Hotelling é dada por (ANDERSON, 2003):

$$T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu})^\top \mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}), \quad (2.6)$$

sendo  $n$  é o tamanho da amostra,  $\bar{\mathbf{X}} = \frac{1}{n} \sum_{j=1}^n \mathbf{X}_j$  o vetor de médias amostrais e  $\mathbf{S} = \frac{1}{n} \sum_{j=1}^n \mathbf{X}_j \mathbf{X}_j^\top - n \bar{\mathbf{X}} \bar{\mathbf{X}}^\top$  a matriz de covariância amostral. A estatística  $T^2$  têm distribuição  $\frac{(n-1)\nu}{(n-\nu)} F_{\nu, n-\nu}$ . Neste caso,  $F_{\nu, n-\nu}$  descreve uma variável aleatória com distribuição  $F$  de Snedecor com  $\nu$  e  $n - \nu$  graus de liberdade (JOHNSON; WICHERN, 2007). Portanto, com o nível de significância  $\eta$  e sob a hipótese  $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ ,

$$\begin{aligned} \eta &= \mathcal{P} \left[ T^2 > \frac{(n-1)\nu}{(n-\nu)} F_{\nu, n-\nu}(\eta) \right], \\ &= \mathcal{P} \left[ n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^\top \mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > \frac{(n-1)\nu}{(n-\nu)} F_{\nu, n-\nu}(\eta) \right], \end{aligned}$$

em que  $F_{\nu, n-\nu}(\eta)$  denota o percentil superior da distribuição  $F_{\nu, n-\nu}$ .

Essa afirmação leva imediatamente a rejeitar  $H_0$  se

$$T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^\top \mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > \frac{(n-1)\nu}{(n-\nu)} F_{\nu, n-\nu}(\eta)$$

(ANDERSON, 2003; JOHNSON; WICHERN, 2007).

De forma análoga, o teste da razão de verossimilhança rejeita  $H_0$  para um nível de significância  $\eta$ , se  $T^2 \geq \frac{(n-1)\nu}{(n-\nu)} F_{\nu, n-\nu}(\eta)$  (ANDERSON, 2003). Adicionalmente, podemos

esboçar elipsóides de confiança  $\nu$ -dimensional definidas por (ANDERSON, 2003)

$$n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^\top \mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \leq \frac{(n-1)\nu}{(n-\nu)} F_{\nu, n-\nu}(\eta).$$

Para amostras grandes, a distribuição de  $T^2$  é dada por sua distribuição limite, que é a distribuição Qui-quadrado com  $\nu$  graus de liberdade ( $\chi_\nu^2$ ) (ANDERSON, 2003; ANFINSEN; ELTOFT, 2011).

## 2.7 Método *bootstrap*

O método *bootstrap* introduzido por Efron (1979) com o estudo “*Bootstrap methods: another look at the jackknife*” é uma técnica estatística de reamostragem. Consiste basicamente na reamostragem com reposição de um conjunto de dados, que pode ocorrer de forma direta ou através de um modelo probabilístico ajustado. Com isso, produz réplicas (repetições) dos dados, possibilitando avaliar a variabilidade de estatísticas, parâmetros ou quantidade de interesse.

Esse mecanismo de sorteio aleatório pode ser de natureza paramétrica ou não paramétrica. A sua formulação consiste em aproximar a distribuição de uma função das observações pela distribuição empírica obtida de uma amostra finita (EFRON, 1979). A amostragem é feita com reposição da distribuição da qual os dados são obtidos. Para situações em que se têm alguma suposição da distribuição que originou os dados utiliza-se o termo *bootstrap* paramétrico, já para os casos que não envolve suposições paramétricas denomina-se *bootstrap* não paramétrico (EFRON, 1979; CORDEIRO; CRIBARI-NETO, 2014).

O *bootstrap* paramétrico gera as amostras *bootstrap* com base na distribuição de probabilidade dos dados. No caso não paramétrico, a amostra é considerada como uma população (objeto de estudo) (DAVISON; HINKLEY, 1997). Nele utiliza-se a amostra original para obter as amostras *bootstrap*. Por exemplo, seja  $\mathbf{x} = (x_1, \dots, x_n)$  uma amostra finita contendo  $n$  observações. Constrói-se então  $B$  amostras (número de amostras *bootstrap*)  $x^{*(1)}, \dots, x^{*(B)}$  independentes, obtidas a partir da reamostragem da amostra inicial  $\mathbf{x}$ .

## 2.8 Poder do teste

Um teste de hipóteses é, de forma geral, formulado sob duas hipóteses: a hipótese nula e hipótese alternativa. Mood, Graybill e Boes (1974) definem como hipótese nula

( $H_0$ ) a afirmação ou suposição que está sendo testada, enquanto a afirmação que nega à alegação de  $H_0$  é denominada de hipótese alternativa ( $H_1$ ). O interesse é testar  $H_0$  versus  $H_1$ .

Se  $\theta$  denota um parâmetro populacional e  $\Theta_0$  é algum subconjunto do espaço paramétrico, Casella e Berger (2001) determinam, sem pormenorizar, as hipóteses nula e alternativa como sendo:  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_0^c$ . Ainda de acordo com Casella e Berger (2001) a finalidade de um teste de hipóteses é decidir, com base em uma amostra da população, qual das duas hipóteses complementares é verdadeira. A concepção é que se a hipótese de nulidade é falsa, então a hipótese alternativa é verdadeira, e vice-versa.

No entanto, ao tomar essa decisão é possível que se cometa erros. Esses erros são denominados erro do tipo I e erro do tipo II. Definidas a hipótese nula e a hipótese alternativa, dadas respectivamente por  $H_0$  e  $H_1$ , o erro tipo I consiste em rejeitar  $H_0$  quando esta é verdadeira e, o erro do tipo II é cometido ao não rejeitar  $H_0$  quando esta é falsa. Associados a esses erros temos os riscos  $\alpha$  (também denominado de nível de significância) e  $\beta$ , que são definidos pelas probabilidades dos erros dos tipos I e II, respectivamente (CASELLA; BERGER, 2001; ROUSSAS, 2014).

Erros do tipo I e II não podem ser totalmente evitados, mas é indispensável que se busque mantê-los com probabilidades de ocorrência relativamente pequenas (MEYER, 1983). Controlar os erros de modo que seja garantida essa condição é um fenômeno de difícil alcance, uma vez que a redução da probabilidade de um tipo de erro ocasiona o aumento da probabilidade do outro para um determinado tamanho de amostra. No entanto, a diminuição simultânea desses erros pode ser atingida pelo aumento do tamanho da amostra (CARMER, 1976).

O poder de um teste é descrito como sendo  $1 - \beta$ . Isto é,  $1 - P(\text{Erro tipo II}) = 1 - P(\text{Aceitar } H_0 | H_0 \text{ é falsa}) = P(\text{Rejeitar } H_0 | H_0 \text{ não é verdadeira})$  (CASELLA; BERGER, 2001; ROUSSAS, 2014). Formalmente, considerando  $f(\cdot, \theta)$  a função de distribuição da amostra parametrizada por  $\theta$  pertencente ao espaço paramétrico  $\Theta$  e,  $\Upsilon$  um teste de hipótese nula  $H_0$ . Então, a função poder de teste  $\Upsilon$ , denotada por  $\pi_\Upsilon(\theta)$ , é a probabilidade que  $H_0$  seja rejeitada quando a distribuição da qual a amostra foi obtida é parametrizada por  $\theta$  (MOOD; GRAYBILL; BOES, 1974).

Nesse entendimento, considera-se um bom teste ou um teste poderoso, aquele que apresenta alta probabilidade de aceitação para a hipótese  $H_0$  quando o parâmetro populacional em estudo pertence ao espaço paramétrico estabelecido pela hipótese de nulidade. Além disso, tem uma probabilidade elevada de rejeição para a mesma hipótese quando o parâmetro não pertence ao espaço paramétrico pré-fixado.

### 3 Distribuição Slashed Lomax: contribuições

A seguir descrevemos as contribuições para a distribuição Slashed Lomax.

#### 3.1 Momento de ordem $r$

**Proposição 1.** O  $r$ -ésimo momento da distribuição de  $Y$ , definido  $\forall r \in \mathbb{N}^*$  e  $r < 2\lambda$ , é determinado por:

$$\mu_r = E(Y^r) = \frac{2\lambda\beta^r\Gamma(\alpha-r)\Gamma(r+1)}{(2\lambda-r)\Gamma(\alpha)}, \quad (3.1)$$

em que  $\Gamma(\cdot)$  é a função gama.

*Demonstração.* Como  $X$  e  $V$  são variáveis aleatórias independentes, temos:

$$\mu_r = E(Y^r) = E(X^r) E\left(V^{-\frac{r}{2}}\right).$$

Apoiado nas propriedades da distribuição beta e da distribuição Lomax, temos que  $E(X^r) = \frac{\beta^r\Gamma(\alpha-r)\Gamma(r+1)}{\Gamma(\alpha)}$  e  $E\left(V^{-\frac{r}{2}}\right) = \frac{2\lambda}{2\lambda-r}$ , para  $r < 2\lambda$ . Por conseguinte, o resultado é obtido.  $\square$

**Observação:** O  $r$ -ésimo momento do artigo de [Li e Tian \(2020\)](#) está errado!

#### 3.2 Estimação pelo método dos momentos

Seja o  $j$ -ésimo momento amostral de uma amostra aleatória com distribuição Slomax  $(\alpha, \beta, \lambda)$ , com  $j \geq 1$ :

$$m_j = \frac{1}{n} \sum_{i=1}^n Y_i^j.$$

Seja também, o  $j$ -ésimo momento populacional

$$\mu_j = E[Y^j], \quad j \geq 1.$$

De acordo com (2.1) os estimadores de  $\boldsymbol{\theta} = (\alpha, \beta, \lambda)$  são obtidos resolvendo-se o sistema de equações a seguir:

$$\begin{cases} \frac{1}{n} \sum_{i=1}^n Y_i = \frac{2\lambda\beta\Gamma(\alpha-1)\Gamma(2)}{(2\lambda-1)\Gamma(\alpha)}, \\ \frac{1}{n} \sum_{i=1}^n Y_i^2 = \frac{2\lambda\beta^2\Gamma(\alpha-2)\Gamma(3)}{(2\lambda-2)\Gamma(\alpha)}, \\ \frac{1}{n} \sum_{i=1}^n Y_i^3 = \frac{2\lambda\beta^3\Gamma(\alpha-3)\Gamma(4)}{(2\lambda-3)\Gamma(\alpha)}. \end{cases}$$

### 3.3 Estimação de máxima verossimilhança

Sejam  $Y_1, \dots, Y_n$  uma amostra aleatória de uma variável aleatória  $Y \sim \text{Slomax}(\alpha, \beta, \lambda)$ , a função de log-verossimilhança associada é dada por:

$$\begin{aligned} l(\boldsymbol{\theta}) &= n \log(2) + n \log(\lambda) + n \log(\alpha) + 2\lambda n \log(\beta) - (2\lambda + 1) \sum_{i=1}^n \log(y_i) \\ &+ \sum_{i=1}^n \log \left[ \int_0^{\frac{y_i}{\beta}} \frac{w^{2\lambda}}{(1+w)^{\alpha+1}} dw \right]. \end{aligned}$$

Sendo  $\boldsymbol{\theta}$  o vetor de parâmetros da distribuição, isto é,  $\boldsymbol{\theta} = (\alpha, \beta, \lambda)$ , os componentes do vetor escore são dados por:

$$U(\boldsymbol{\theta}) = (U_\alpha, U_\beta, U_\lambda) = \left( \frac{\partial l(\boldsymbol{\theta})}{\partial \alpha}, \frac{\partial l(\boldsymbol{\theta})}{\partial \beta}, \frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} \right)^\top, \quad (3.2)$$

em que as equações de (3.2) são:

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \left\{ \frac{\frac{\partial}{\partial \alpha} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw \right]}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw} \right\}, \quad (3.3)$$

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta} = \frac{2\lambda n}{\beta} + \sum_{i=1}^n \left\{ \frac{\frac{\partial}{\partial \beta} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw \right]}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw} \right\} \quad (3.4)$$

e

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} = \frac{n}{\lambda} + 2n \log \beta - 2 \sum_{i=1}^n \log(y_i) + \sum_{i=1}^n \left\{ \frac{\frac{\partial}{\partial \lambda} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw \right]}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw} \right\}. \quad (3.5)$$

As expressões para o vetor escore aqui determinadas não foram obtidas anteriormente. A prova de (3.3), (3.4) e (3.5) é apresentada no Apêndice B.

Neste caso, é notório que os estimadores de máxima verossimilhança para o modelo especificado não possuem expressões em forma fechada, fazendo necessário o uso de métodos iterativos para realizar a estimação. Assim, utilizaremos o algoritmo BFGS (Broyden Fletcher-Goldfarb-Shanno) (NOCEDAL; WRIGHT, 1999) para determinar as estimativas dos parâmetros.

### 3.4 Transformada de Mellin

**Teorema 1.** Seja  $Y \in \mathbb{R}^+$  uma variável aleatória seguindo o modelo Slashed Lomax com fdp dada por  $f_Y(y)$ . A primeira função característica do segundo tipo de  $Y$ , expressa pela TM de  $Y$ , é dada por:

$$\phi_Y(s) = \frac{2\lambda\beta^{s-1}\Gamma(s)\Gamma(\alpha - s + 1)}{(2\lambda - s + 1)\Gamma(\alpha)}.$$

*Demonstração.* De acordo com (2.2),

$$\phi_Y(s) = \int_0^\infty y^{s-1} f_Y(y) dy = E(Y^{s-1}).$$

Considerando o resultado dado em (3.1) e fazendo  $r = s - 1$ , temos que:

$$E(Y^{s-1}) = \frac{2\lambda\beta^{s-1}\Gamma(s)\Gamma(\alpha - s + 1)}{(2\lambda - s + 1)\Gamma(\alpha)}.$$

□

**Proposição 2.** A segunda função característica do segundo tipo de  $Y \sim \text{Slomax}(\alpha, \beta, \lambda)$  é definida como:

$$\begin{aligned} \psi_Y(s) &= \log(2) + \log(\lambda) + s \log(\beta) - \log(\beta) + \log[\Gamma(s)] + \log[\Gamma(\alpha - s + 1)] - \log[\Gamma(\alpha)] \\ &\quad - \log(2\lambda - s + 1). \end{aligned}$$

A demonstração é imediata do  $\log(\phi_Y(s))$ .

### 3.5 Diagrama de log-cumulantes

O diagrama de Pearson, abordado em Tison et al. (2004) consiste em escolher o modelo que melhor se adequa a amostra em análise com base nos coeficientes de assimetria

( $\beta_1$ ) e curtose ( $\beta_2$ ). Essas medidas são definidas em termos dos momentos centrais de segunda ( $M_2$ ), terceira ( $M_3$ ) e quarta ordem ( $M_4$ ):

$$\beta_1 = \frac{M_3^2}{M_2^3} \text{ e } \beta_2 = \frac{M_4}{M_2^2},$$

em que  $M_k = E[(\mathbf{X} - E[\mathbf{X}])^k]$ .

No entanto, a quadratura do momento de terceira ordem em  $\beta_1$  não permite diferir as distribuições que tem assimetria de mesma magnitude, mas com sinal oposto, impossibilitando separar modelos com normalidade e modelos com cauda pesada (NICOLAS, 2002).

À vista disso, Nicolas (2002) classificou esse diagrama como mal-adaptado às distribuições definidas em  $\mathbb{R}^+$  e como solução propôs o diagrama  $(\tilde{k}_3, \tilde{k}_2)$  (diagrama de log-cumulantes). Semelhante ao diagrama de Pearson, este diagrama, em vez de utilizar os coeficientes de assimetria e curtose, faz uso dos LCs, ou seja, as estatísticas  $\tilde{k}_3$  versus  $\tilde{k}_2$ . A representação dos modelos no diagrama  $(\tilde{k}_3, \tilde{k}_2)$  é dada por subespaços e suas dimensões determinadas pelo número de parâmetros contidos nas expressões dos LCs (ANFINSSEN; ELTOFT, 2011; ANFINSSEN; DOULGERIS; ELTOFT, 2011; VASCONCELOS; CINTRA; NASCIMENTO, 2021).

Posteriormente, Anfinsen e Eltoft (2011) introduziram o diagrama da matriz de LCs, visto como uma extensão do diagrama de LCs para a Transformada de Mellin univariada discutido em Nicolas (2002). A nova ferramenta propõe uma forma de visualizar o espaço multidimensional, cada dimensão é representada por uma matriz log-cumulante de ordem  $n$ . Em Anfinsen, Doulgeris e Eltoft (2011), Anfinsen e Eltoft (2011), Li et al. (2011), Bombrun, Anfinsen e Harant (2011), Khan e Guida (2013), Deng e López-Martínez (2016) e Deng et al. (2018) o diagrama de LCs é apresentado como critério para classificar modelos probabilísticos para dados de imagens SAR e PolSAR. Por outro lado, em Vasconcelos, Cintra e Nascimento (2021) o diagrama  $(\tilde{k}_3, \tilde{k}_2)$  é utilizado para dados reais de sobrevivência através da classificação das distribuições beta-G.

Neste trabalho, empregamos o digrama  $(\tilde{k}_3, \tilde{k}_2)$  como critério para avaliar o ajuste da distribuição Slashed Lomax aplicada a dados de sobrevivência.

Os LCs do modelo Slashed Lomax, obtidos a partir de (2.4), estão assim definidos:

$$\tilde{k}_\nu = \begin{cases} \log \beta + \Psi(1) - \Psi(\alpha) + \frac{1}{2\lambda}, & \text{se } \nu = 1; \\ \Psi^{(\nu-1)}(1) + (-1)^\nu \Psi^{(\nu-1)}(\alpha) + \frac{(\nu-1)!}{(2\lambda)^\nu}, & \text{se } \nu \geq 2. \end{cases} \quad (3.6)$$



e, encontram-se em termos das funções digama e poligama, dadas por  $\Psi(x) = \frac{d}{dx} \log \Gamma(x)$  e  $\Psi^m(x) = \frac{d^m}{dx^m} \Psi(x)$  (ABRAMOWITZ; STEGUN, 1948), respectivamente.

**Proposição 3.** Seja  $Y \sim \text{Slomax}(\alpha, \beta, \lambda)$  e  $\psi_Y(s)$  a segunda função característica do segundo tipo de  $Y$ . Então,  $\forall n \in \mathbb{N}, n > 1$ , temos:

$$\tilde{k}_n = \left. \frac{d^n \psi_Y(s)}{ds^n} \right|_{s=1} = \Psi^{(n-1)}(1) + (-1)^n \Psi^{(n-1)}(\alpha) + \frac{(n-1)!}{(2\lambda)^n}, \quad (3.7)$$

onde

$$\frac{d^n \psi_Y(s)}{ds^n} = \Psi^{(n-1)}(s) + (-1)^n \Psi^{(n-1)}(\alpha + 1 - s) + \frac{(n-1)!}{(2\lambda - s + 1)^n}.$$

*Demonstração.* Por se tratar de uma propriedade definida para números naturais, utilizaremos o Princípio de Indução (LIMA, 2014) para demonstrá-la. De acordo com o método de indução (LIMA, 2017), se a propriedade  $\tilde{k}_n$  é válida para  $n = 2$  e ao supor sua validade para  $n = \nu$  resultar  $\tilde{k}_n$  válida para  $\nu + 1$ , então a propriedade é válida para todos os números naturais  $n > 1$ . De fato:

i) Para  $n = 2$ , temos:

$$\frac{d^2 \psi_Y(s)}{ds^2} = \Psi^{(1)}(s) + \Psi^{(1)}(\alpha + 1 - s) + \frac{1}{(2\lambda - s + 1)^2},$$

e quando  $s = 1$ ,

$$\left. \frac{d^2 \psi_Y(s)}{ds^2} \right|_{s=1} = \Psi^{(1)}(1) + \Psi^{(1)}(\alpha) + \frac{1}{(2\lambda)^2}.$$

Então,

$$\tilde{k}_2 = \left. \frac{d^2 \psi_Y(s)}{ds^2} \right|_{s=1} = \Psi^{(2-1)} + (-1)^2 \Psi^{(2-1)}(\alpha) + \frac{(2-1)!}{(2\lambda)^2}.$$

ii) Supondo válida a igualdade (3.7) para  $n = \nu$ , temos:

$$\frac{d^\nu \psi_Y(s)}{ds^\nu} = \Psi^{(\nu-1)}(s) + (-1)^\nu \Psi^{(\nu-1)}(\alpha + 1 - s) + \frac{(\nu-1)!}{(2\lambda - s + 1)^\nu}, \quad (3.8)$$

e para  $s = 1$ ,

$$\tilde{k}_\nu = \left. \frac{d^\nu \psi_Y(s)}{ds^\nu} \right|_{s=1} = \Psi^{(\nu-1)}(1) + (-1)^\nu \Psi^{(\nu-1)}(\alpha) + \frac{(\nu-1)!}{(2\lambda)^\nu}.$$

iii) Derivando (3.8), obtemos:

$$\frac{d^{\nu+1}\psi_Y(s)}{ds^{\nu+1}} = \Psi^{(\nu)}(s) + (-1)^{\nu+1}\Psi^{(\nu)}(\alpha + 1 - s) + \frac{\nu!}{(2\lambda - s + 1)^{\nu+1}}.$$

Assim, para  $s = 1$ ,

$$\left. \frac{d^{\nu+1}\psi_Y(s)}{ds^{\nu+1}} \right|_{s=1} = \Psi^{(\nu)}(1) + (-1)^{\nu+1}\Psi^{(\nu)}(\alpha) + \frac{\nu!}{(2\lambda)^{\nu+1}}.$$

Logo,

$$\tilde{k}_{\nu+1} = \left. \frac{d^{\nu+1}\psi_Y(s)}{ds^{\nu+1}} \right|_{s=1} = \Psi^{[(\nu+1)-1]}(1) + (-1)^{\nu+1}\Psi^{[(\nu+1)-1]}(\alpha) + \frac{[(\nu+1)-1]!}{(2\lambda)^{\nu+1}}.$$

□

A Figura 3 exibe a região do diagrama  $(\tilde{k}_3, \tilde{k}_2)$  vinculada a distribuição Slashed Lomax. A área é simbolizada por um subespaço bidimensional, cuja dimensão foi fixada pelos dois parâmetros dos quais os LCs apresentados em (3.6) estão associados.

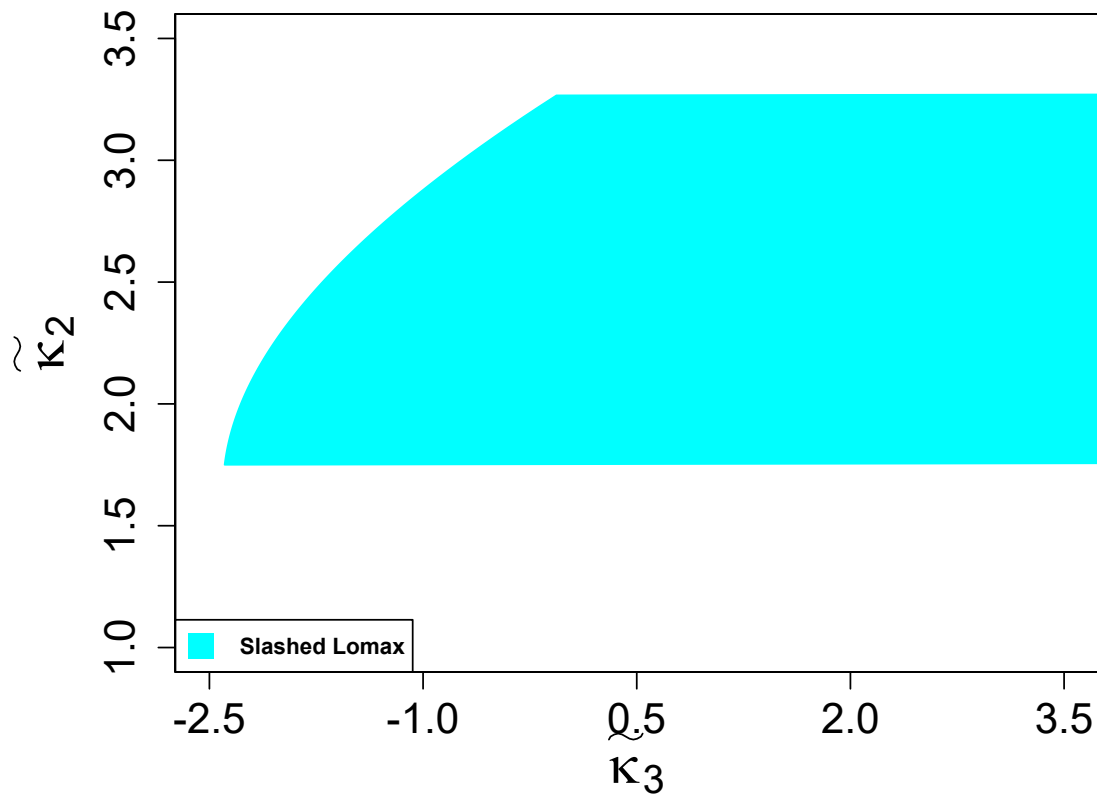


Figura 3 – Diagrama dos LCs  $(\tilde{k}_3, \tilde{k}_2)$  para o modelo Slashed Lomax

### 3.6 Teste de bondade de ajuste para a distribuição Slashed Lomax

Aqui, propomos uma medida de bondade de ajuste quantitativa através da inserção dos LCs na estatística  $T^2$  de Hotelling. Com essa ferramenta, pretende-se classificar o ajuste do modelo aos dados de acordo com a localização dos LCs estimados  $\begin{bmatrix} \hat{k}_2 & \hat{k}_3 \end{bmatrix}^\top$  sobre o diagrama  $(\tilde{k}_3, \tilde{k}_2)$ .

Para isso, carecemos de uma estatística de teste para a hipótese de nulidade  $H_0 : E \left( \begin{bmatrix} \hat{k}_2 & \hat{k}_3 \end{bmatrix} \right) = \begin{bmatrix} \tilde{k}_2 & \tilde{k}_3 \end{bmatrix}$ . Este teste deve aceitar ou não as estimativas dos LCs para a região teórica do diagrama  $(\tilde{k}_2, \tilde{k}_3)$ , indicando assim a adequação ou discrepância do modelo em teste. Os LCs tendem a ser quantidades analiticamente bem definidas e, para a distribuição Slashed Lomax, admitem expressões em forma fechada (Equação 3.6). Tal relação entre parâmetros do modelo e LCs pode ser usada para derivar estimadores para os LCs. À vista disso, consideraremos as seguintes medidas:

$$\tilde{m} = (\tilde{m}_1, \dots, \tilde{m}_p)^\top \text{ e } \hat{m} = (\hat{m}_1, \dots, \hat{m}_p)^\top.$$

Pelo teorema do limite central de Cramér-Wold (BILODEAU; BRENNER, 2008), obtemos a relação a seguir:

$$\sqrt{n}(\hat{m} - \tilde{m}) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \mathcal{N}_v(0, \mathbf{M}_v),$$

em que a amostra segue a distribuição normal multivariada de média  $\boldsymbol{\mu}$  e matriz de covariância  $\boldsymbol{\Sigma}$ , expressa por  $\mathbf{X} \sim \mathcal{N}_v(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  e  $\xrightarrow[n \rightarrow \infty]{\mathcal{D}}$  denota convergência em distribuição.

Além do que,  $\mathbf{M}_v$  é definido por:

$$nE \left\{ (\hat{m} - \tilde{m})(\hat{m} - \tilde{m})^\top \right\} = \left\{ \tilde{m}_{r+i} - \tilde{m}_r \tilde{m}_i \right\} \Bigg|_{r,i=1}^v.$$

Ainda, levando em conta as funções compostas que transforma os LCs em polinômios dos LMs (DENG et al., 2018), temos que:

$$\tilde{k}_v = g_v(\tilde{m}_1, \dots, \tilde{m}_v) \text{ e } \hat{k}_v = g_v(\hat{m}_1, \dots, \hat{m}_v).$$

Dessa forma, pelo teorema do limite central de Cramér-Wold, temos

$$\sqrt{n}(\hat{k} - \tilde{k}) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} z \sim \mathcal{N}_v(0, \mathbf{K}_v),$$

com  $\xrightarrow[n \rightarrow \infty]{\mathcal{D}}$  denotando convergência em distribuição e  $\mathbf{K}_v$  determinado por:

$$\mathbf{K}_v = \mathbf{J}_v^\top \cdot \mathbf{M}_v \cdot \mathbf{J}_v, \quad (3.9)$$

$$\mathbf{J}_v = [\nabla g_1 | \cdots | \nabla g_v]$$

e

$$\nabla g_v = \left[ \frac{\partial g_v(\tilde{m})}{\partial \tilde{m}_1}, \dots, \frac{\partial g_v(\tilde{m})}{\partial \tilde{m}_v} \right].$$

Para mais, de acordo com o método delta (CASELLA; BERGER, 2001; BILODEAU; BRENNER, 2008), se considerarmos grandes amostras, é conhecida a distribuição do vetor estimado  $\left[ \hat{k}_2 \quad \hat{k}_3 \right]^\top$ , que têm distribuição gaussiana bivariada com média  $\left[ \tilde{k}_2 \quad \tilde{k}_3 \right]^\top$  e uma matriz de covariância assintótica  $\mathbf{K}$ . Isto é:

$$\left( \left[ \begin{array}{c} \hat{k}_2 \\ \hat{k}_3 \end{array} \right] - \left[ \begin{array}{c} \tilde{k}_2 \\ \tilde{k}_3 \end{array} \right] \right) \stackrel{a}{\sim} \mathcal{N}_2(0, \mathbf{K}), \quad (3.10)$$

em que  $\stackrel{a}{\sim}$  indica distribuição assintótica e  $\mathbf{K} = Cov(\hat{m}_r, \hat{m}_i) \Big|_{r,i=1}^v = \frac{1}{n}(\tilde{m}_{r+i} - \tilde{m}_r \tilde{m}_i)$ .

Neste caso, comparando o resultado obtido na Equação (3.10) com a estatística  $T^2$  de Hotelling expressa por (2.6), temos que  $\boldsymbol{\mu} = \left[ \tilde{k}_2 \quad \tilde{k}_3 \right]^\top$  e  $\bar{\mathbf{X}} = \left[ \hat{k}_2 \quad \hat{k}_3 \right]^\top$ . A matriz  $\mathbf{S}$  em (2.6) será substituída pela matriz de covariância assintótica estimada  $\widehat{\mathbf{K}}$ . Assim sendo, para um nível de significância  $\eta$  fixado, temos a seguinte estatística de teste:

$$T^2 = n \left( \left[ \begin{array}{c} \hat{k}_2 \\ \hat{k}_3 \end{array} \right] - \left[ \begin{array}{c} \tilde{k}_2 \\ \tilde{k}_3 \end{array} \right] \right)^\top \widehat{\mathbf{K}}^{-1} \left( \left[ \begin{array}{c} \hat{k}_2 \\ \hat{k}_3 \end{array} \right] - \left[ \begin{array}{c} \tilde{k}_2 \\ \tilde{k}_3 \end{array} \right] \right) \quad (3.11)$$

em que  $\widehat{\mathbf{K}}^{-1}$  é a inversa de  $\widehat{\mathbf{K}}$  se a matriz é não singular, do contrário, utiliza-se a inversa generalizada de Moore-Pearson (RAO et al., 2008; SEBER, 2008).

Conforme mencionado em Anfinen e Eltoft (2011) e Anfinen, Doulergis e Eltoft (2011), para grandes amostras, a variável aleatória  $T^2$  têm distribuição Qui-quadrado com  $\nu$  graus de liberdade ( $\chi_\nu^2$ ). Dessa forma é possível adotar em (3.11) a aproximação  $F_{\nu, n-\nu}(\cdot) \approx \chi_\nu^2(\cdot)$ , em que  $\chi_\nu^2(\cdot)$  é o percentil superior da distribuição  $\chi_\nu^2$ . Além disso, podemos obter elipses de confiança com um nível de significância especificado  $\eta$ , que são dadas por:

$$\left( \left[ \begin{array}{c} \hat{k}_2 \\ \hat{k}_3 \end{array} \right] - \left[ \begin{array}{c} \tilde{k}_2 \\ \tilde{k}_3 \end{array} \right] \right)^\top \widehat{\mathbf{K}}^{-1} \left( \left[ \begin{array}{c} \hat{k}_2 \\ \hat{k}_3 \end{array} \right] - \left[ \begin{array}{c} \tilde{k}_2 \\ \tilde{k}_3 \end{array} \right] \right) \leq \frac{2(n-1)}{n(n-2)} \chi_2^2(\eta), \quad (3.12)$$

sendo  $\chi_2^2(\eta)$  o percentil superior para  $\chi_2^2$ . A elipse descrita em (3.12) possui centro em  $(\tilde{k}_2, \tilde{k}_3)$  e eixos direcionados de acordo com os autovetores da matriz  $\widehat{\mathbf{K}}$  (JOHNSON; WICHERN, 2007; HÄRDLE; SIMAR, 2019).

A proposição a seguir expressa a estatística de teste  $T^2$  de Hotelling para a distribuição Slashed Lomax.

**Proposição 4.** Seja  $Y$  uma variável aleatória com distribuição Slomax  $(\alpha, \beta, \lambda)$ . Então, a estatística  $T^2$  de Hotelling de  $Y$ , baseada nos LCs, é dada por:

$$T_{SL}^2 = \frac{n}{\hat{\delta}_{22}\hat{\delta}_{33} - \hat{\delta}_{23}^2} \left[ \hat{\delta}_{33} \left( \hat{k}_2 - \tilde{k}_2 \right)^2 + \hat{\delta}_{22} \left( \hat{k}_3 - \tilde{k}_3 \right)^2 - 2\hat{\delta}_{23} \left( \hat{k}_2 - \tilde{k}_2 \right) \left( \hat{k}_3 - \tilde{k}_3 \right) \right].$$

A prova desta proposição juntamente com os algoritmos para gerar a estatística de teste encontram-se no Apêndice C.

## 4 Aplicação

### 4.1 Conjuntos de dados e estatísticas descritivas

Selecionamos dois conjuntos de dados reais a serem submetidos à metodologia abordada, determinando, de acordo com o critério de bondade de ajuste introduzido, se a distribuição Slashed Lomax é um modelo adequado para estudar esses dados. A seguir, descreveremos brevemente as bases de dados.

- i) Sobrevivência (SO): 33 observações sobre os tempos de sobrevivência, em semanas, de pacientes que sofrem de leucemia mielóide (SELIM, 2018). Os dados são: 65, 156, 100, 134, 16, 108, 121, 4, 39, 143, 56, 26, 22, 1, 1, 5, 65, 56, 65, 17, 7, 16, 22, 3, 4, 2, 3, 8, 4, 3, 30, 4, 43;
- ii) Falha (FA): 18 observações que fornecem o tempo até a falha de dispositivos eletrônicos (WANG, 2000). Os dados são: 5, 11, 21, 31, 46, 75, 98, 122, 145, 165, 196, 224, 245, 293, 321, 330, 350, 420.

Dados	Média	Mediana	DP	Assimetria	Curtose	CV (%)	$\hat{k}_2$	$\hat{k}_3$
SO	40.88	22	46.7	1.11	-0.06	114.25	2.29	-0.73
FA	172.11	155	131.54	0.29	-1.35	76.43	1.57	-1.94

Tabela 1 – Estatísticas descritivas dos dados

A Tabela 1 contém as estatísticas descritivas para cada conjunto de dados. Os dados SO e FA estão bem distribuídos em torno da média, apresentando desvio-padrão (DP) de 46.7 e 131.54, respectivamente. Além disso, possuem média maior que a mediana, resultando em uma assimetria positiva. Com relação a curtose, ambos os dados têm medidas de curtose menor do que 3, ou seja, tem curva de inclinação negativa com distribuição platicúrtica. Os dois conjuntos de dados são heterogêneos apresentando coeficiente de variação (CV) de 114.25% e 76.43%. As quantidades  $\hat{k}_2$  e  $\hat{k}_3$  representam os LCs estimados da amostra.

A Figura 4 apresenta os histogramas e a densidade estimada para os conjuntos de dados SO e FA. As estimativas para os parâmetros foram obtidas através dos seguintes

métodos de estimação: MLC, MM e MV. A estimação pelo MV mostrou-se mais adequada para ambos os conjuntos de dados.

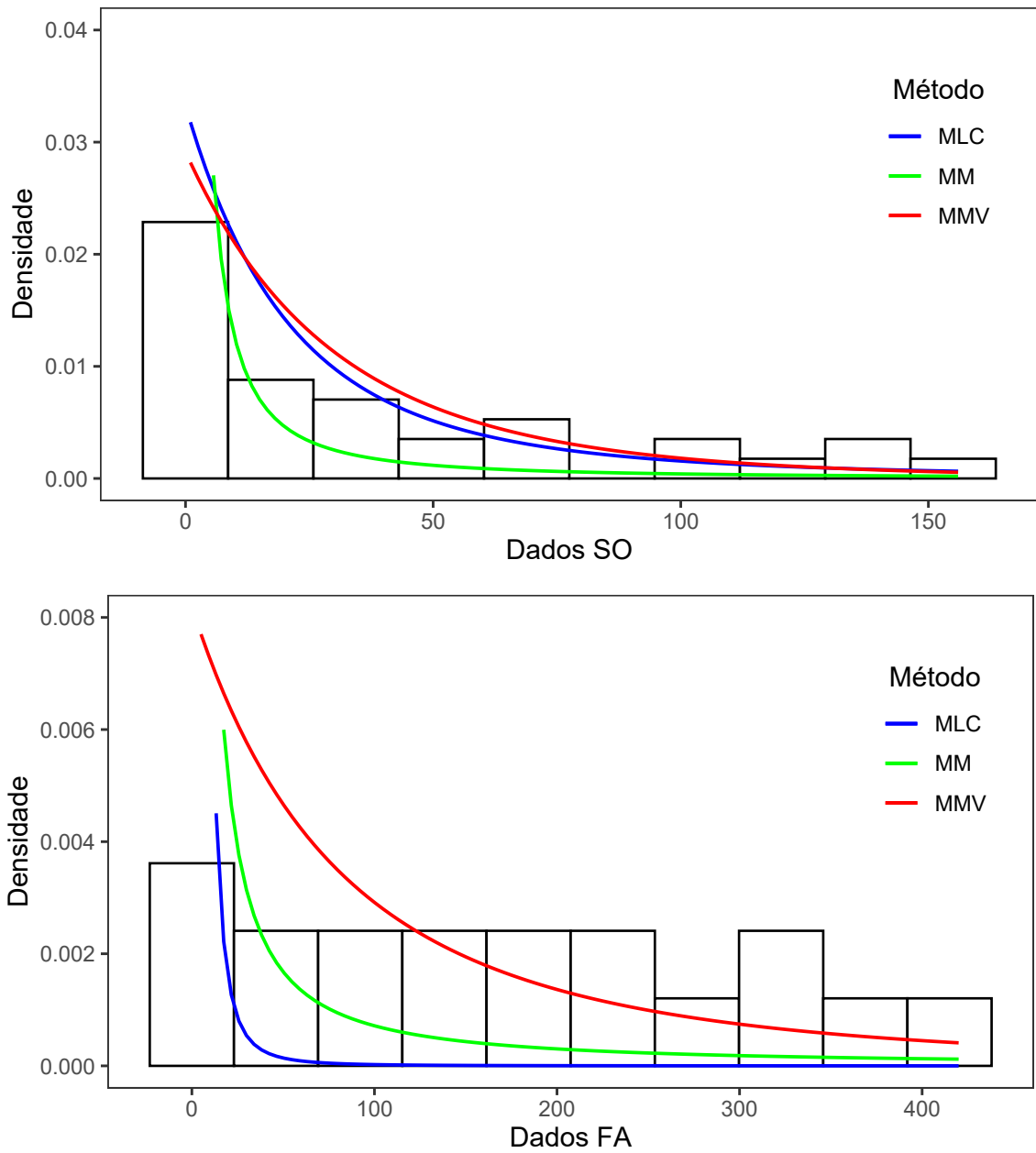
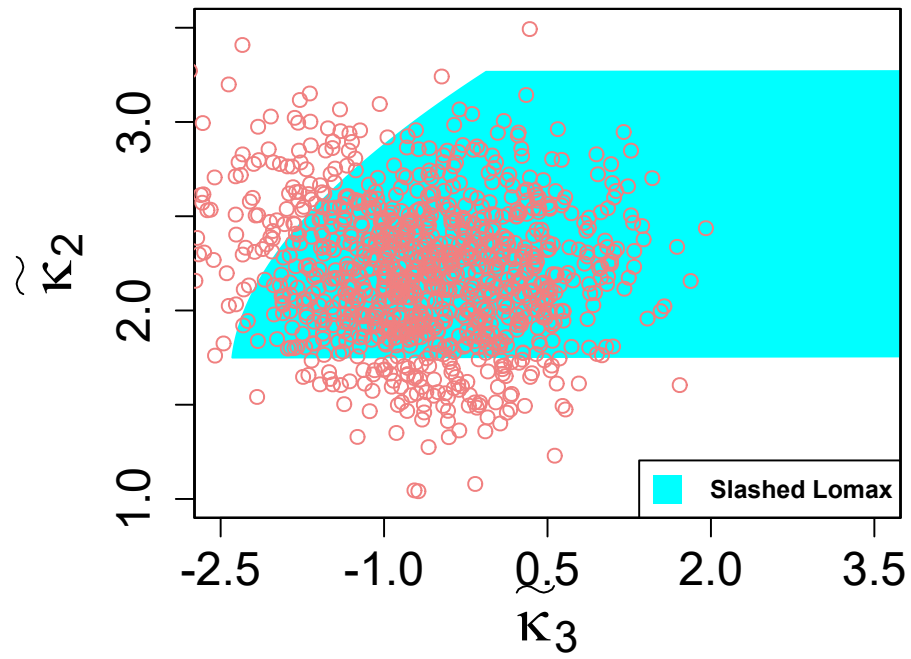


Figura 4 – Histogramas e densidade estimada

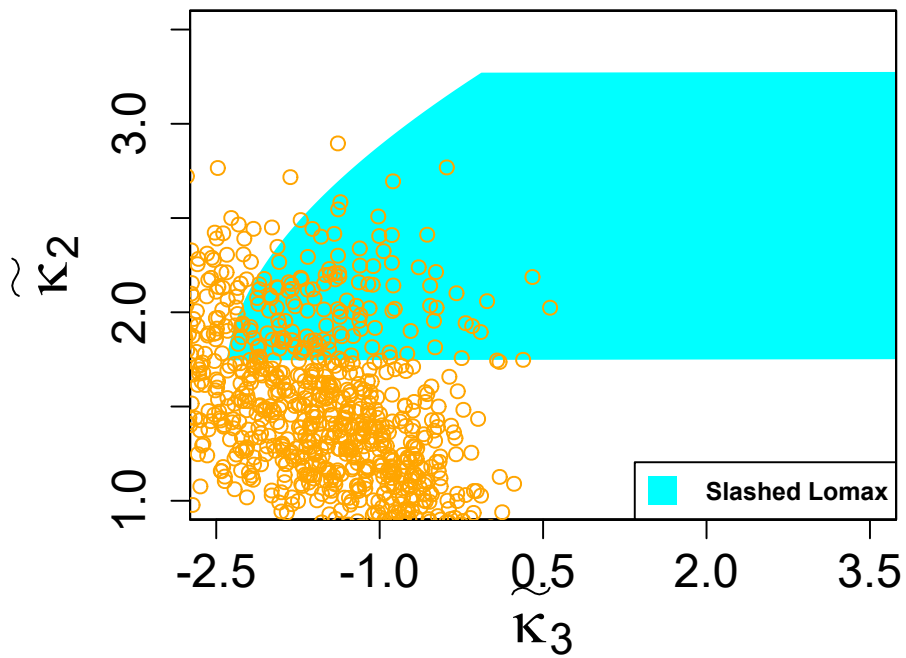
## 4.2 Diagrama $(\tilde{k}_3, \tilde{k}_2)$

A Figura 5 exibe os diagramas de LCs  $(\tilde{k}_3, \tilde{k}_2)$  enquanto medida de qualidade de ajuste. Para ilustrar cada conjunto de dados, calculamos os LCs amostrais considerando uma amostragem via *bootstrap*, na qual geramos 1.000 réplicas com tamanhos amostrais de 95% da amostra, ou seja, para o conjunto de dados SO tivemos uma amostra de 31 e para o segundo banco de dados, uma amostra de 17. A utilização da técnica de *bootstrap* se justifica uma vez que cada conjunto de dados produz um log-cumulante que é associado a um único ponto no Diagrama  $(\tilde{k}_3, \tilde{k}_2)$ . Observamos que para o conjunto de dados SO, a maioria dos pontos estão localizados na região vinculada ao modelo Slashed Lomax. Já para o conjunto de observações FA apenas parte dos pontos estão sob a área teórica do diagrama. Como em ambos os conjuntos de dados as nuvens de pontos não estão sob a região do modelo de probabilidade, vamos nos debruçar sobre a estatística (de teste)  $T^2$  de Hotelling para verificar se de fato o modelo de probabilidade Slashed Lomax é ajustável aos bancos de dados.





(a) Dados SO



(b) Dados FA

Figura 5 – Diagrama  $(\tilde{k}_3, \tilde{k}_2)$  e os LCs amostrais estimados

### 4.3 Estatística $T^2$ de Hotelling

Na Tabela 2 temos as estimativas dos parâmetros  $\alpha$ ,  $\beta$  e  $\lambda$  obtidos pelo MLC e a estatística  $T^2$  de Hotelling para cada conjunto de dados, bem como seus respectivos p-valores. Essas estatísticas podem ser vistas como uma medida de distância entre o modelo

Dados	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$T^2$ de Hotelling	p-valor
SO	16.49	255.05	0.53	0.75	0.70
FA	117.09	160.14	0.81	1.25	0.57

Tabela 2 – Estimativas, estatísticas  $T^2$  de Hotelling e p-valores para os conjuntos de dados

em teste e os conjuntos de dados. Valores mais baixos para a estatística  $T^2$  evidenciam uma proximidade entre o modelo e os dados, sugerindo, portanto, um bom ajuste.

### 4.4 Elipses de confiança

Para o complemento da aplicação anterior, exploramos novamente o campo da análise visual esboçando as elipses de confiança para cada conjunto de dados. Além disso, quantificamos o grau de cobertura (GC) para cada uma dessas amostras. A construção das elipses é dada a partir de (3.12). Para isto, empregamos o menor valor da estatística  $T^2$  de Hotelling apresentado na Tabela 2, a distribuição Slashed Lomax utilizando os LCs  $(\tilde{k}_3, \tilde{k}_2)$  e a matriz de covariância assintótica  $\mathbf{K}$ . Para os dados SO, a elipse apresentou uma cobertura de 82.58% dos dados, enquanto para os dados FA o grau de cobertura é de 91.73%. A Figura 6 contém as elipses obtidas e seus respectivos graus de cobertura.

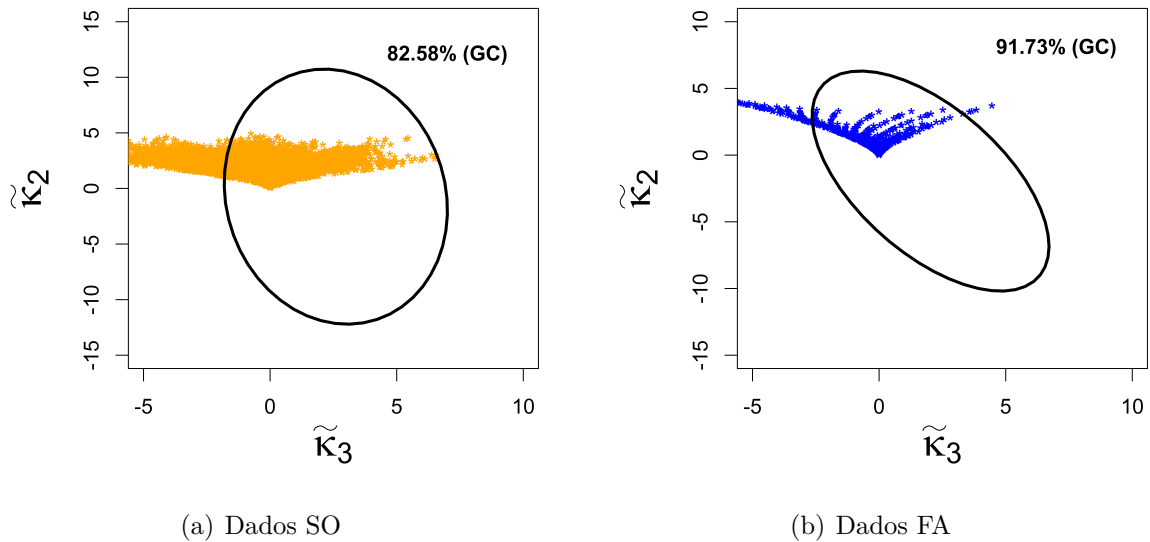


Figura 6 – Elipses de confiança para os conjunto de dados

## 4.5 *Bootstrap* da estatística de teste

Nesta seção, conduzimos estudo de *bootstrap* para avaliar a eficiência da medida de bondade de ajuste quantitativa aplicada a dados reais. Tal “eficácia” será mensurada pelo poder do teste da estatística  $T^2$  de Hotelling. Consideramos os procedimentos MLC, MM e MV para o cálculo da estatística de teste, objetivando analisar o efeito do uso das estimativas para a nova ferramenta de ajuste.

Através dos conjuntos de dados SO e FA, trabalhamos a técnica *bootstrap* para verificar a eficácia da estatística de teste  $T^2$  com a seguinte configuração: com 1000 réplicas de *bootstrap* tomamos os tamanhos amostrais  $n_1 \in \{100, 200, 300, 400, 500\}$  e  $n_2 \in \{100, 200, 300, 400, 500, 600\}$ , em que  $n_1$  e  $n_2$  correspondem ao primeiro e segundo banco de dados, respectivamente.

No que tange ao poder do teste, fixamos níveis de significância  $\eta \in \{1\%, 5\%, 10\%\}$  e calculamos a probabilidade de decisão correta, isto é, a probabilidade de rejeitar  $H_0$  quando esta não é verdadeira. A seguir, no Algoritmo 1, apresentamos um resumo contextual do *bootstrap* não paramétrico utilizado para ilustrar eficiência da estatística de teste dada na **Proposição 4**.

A Tabela 3 exhibe o poder do teste da estatística  $T^2$  de Hotelling aplicada as bases

---

**Algoritmo 1:** *Bootstrap* da estatística de teste  $T^2$  de Hotelling
 

---

**Passo 1.** Determine o nível de significância  $\eta$ ;

**Passo 2.** Obtenha as estimativas de  $\alpha, \beta$  e  $\lambda$  da distribuição de probabilidade Slashed Lomax empregando o MLC, MM e MV e determine os LCs deste modelo;

**Passo 3.** Gere a matriz de covariância assintótica;

**Passo 4.** Calcule os LCs da amostra por meio da transformação de momentos para cumulantes;

**Passo 5.** Obtenha a estatística  $T^2$  de Hotelling dada na **Proposição 4**, aqui denominada  $T_p^2$ ;

**Passo 6.** Gere aleatoriamente, com reposição, 1000 amostras de tamanho  $n$  sob a base de dados estudada. Para cada amostra, repita os **Passos 2-5** e armazene as estatísticas de teste simuladas como  $\{T_s^2(i)\}_{i=1}^m$ ;

**Passo 7.** Some as estatísticas de teste simuladas que são maiores do que a estatística de teste  $T_p^2$  obtida no **Passo 5** e calcule a fração em relação ao número de réplicas. Isso fornece o p-valor, ou seja,

$$P = \frac{1}{1000} \sum_{i=1}^m \mathbb{I}(T_s^2(i) > T_p^2),$$

em que  $\mathbb{I}(\cdot)$  é a função indicadora. O p-valor simulado  $P$  é então comparado com o nível de significância  $\eta$  determinado no teste, rejeitando  $H_0$  se  $P < \eta$ ;

**Passo 8.** Execute o teste de hipóteses comparando o p-valor com o nível de significância adotado;

**Passo 9.** Calcule o poder do teste.

---

de dados SO e FA. Observa-se que para 10% de significância têm-se um teste mais flexível, resultando em um elevado poder de teste, enquanto que 1% e 5% de significância produzem testes mais rigorosos, reduzindo o poder do teste. No geral, para ambos os valores de significância assumidos, o poder de teste da estatística  $T^2$  de Hotelling forneceu altas probabilidades de tomada de decisão correta, conforme aumenta o tamanho da amostra, demonstrando assim a eficiência da medida de bondade de ajuste proposta.

Além disso, é possível constatar um teste mais poderoso quando as estimativas dos parâmetros são produzidas pelo MLC, evidenciando assim um melhor desempenho do MLC enquanto critério de estimação.

	MLC			MV			MM		
				Poder-SO					
$n_1$	1%	5%	10%	1%	5%	10%	1%	5%	10%
100	0	0	0,019	0	0	0	0	0	0,003
200	0,015	0,563	0,981	0	0	0	0	0,006	0,26
300	0,346	0,995	1	0	0,052	0,592	0	0,175	0,851
400	0,78	1	1	0,049	0,862	1	0,015	0,562	0,989
500	0,947	1	1	0,348	0,995	1	0,069	0,845	1
				Poder-FA					
$n_2$	1%	5%	10%	1%	5%	10%	1%	5%	10%
100	0	0	0,021	0	0	0	0	0	0,005
200	0	0,115	0,765	0	0	0	0	0,018	0,392
300	0,034	0,718	0,998	0	0	0	0,002	0,265	0,922
400	0,188	0,97	1	0	0	0	0,024	0,7	0,999
500	0,451	0,998	1	0	0,079	0,656	0,106	0,951	1
600	0,682	1	1	0,478	0,998	1	0,264	0,993	1

Tabela 3 – Poder de teste da estatística  $T^2$  de Hotelling

## 5 Considerações finais

Neste estudo, fizemos uma revisão de literatura com respeito ao modelo de probabilidade Slashed Lomax. Nessa revisão identificamos divergências no valor esperado. Neste caso, mostramos detalhadamente uma nova expressão para o valor esperado. Determinamos também os elementos do vetor escore, ainda não obtidos anteriormente. Após esse desenrolar focamos na parte principal dos nossos objetivos que foi desenvolver medidas de qualidade de ajuste através dos LCs do segundo tipo e da estatística  $T^2$  de Hotelling para o modelo de probabilidade Slashed Lomax. Proporcionamos análises qualitativas e quantitativas para as ferramentas de ajuste desenvolvidas. Encontramos expressões de forma fechada para as funções características do segundo tipo, LCs e estatística  $T^2$  de Hotelling do modelo estudado. Derivamos elipses de confiança para os LCs da distribuição Slashed Lomax. Aplicações com dois bancos de dados reais no contexto da análise de sobrevivência foram feitas afim de demonstrar a eficiência dos critérios de ajuste desenvolvidos. Experimentos de *bootstrap* foram realizados para avaliar o poder do teste da medida de bondade de ajuste quantitativa. Essas análises indicaram um bom desempenho das ferramentas de ajuste e a efetividade do MLC enquanto técnica de estimação.

Como trabalhos futuros, pretendemos avaliar, utilizando simulação de Monte Carlo, o desempenho da medida de bondade de ajuste quantitativa, comparando-a com testes já estabelecidos (por exemplo: Kolmogorov-Smirnov, Cramér-von Mises e Anderson-Darling). Nesse estudo vamos inicialmente em busca de um método de otimização razoável porque os estimadores não tem forma fechada. Depois escolher um bom estimador para enfim comparar a estatística de teste  $T^2$  com esses testes já existentes na literatura. Um segundo artigo que pretendemos trabalhar é correção de viés de Cox-Snell.

## Referências bibliográficas

- ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of mathematical functions with formulas, graphs, and mathematical tables**. [S.l.]: US Government printing office, 1948. v. 55.
- AKAIKE, H. A new look at the statistical model identification. **IEEE transactions on automatic control**, Ieee, v. 19, n. 6, p. 716–723, 1974.
- ANDERSON, T. W. **An introduction to multivariate statistical analysis**. [S.l.]: Wiley-Interscience, 2003. v. 3.
- ANFENSEN, S. N.; DOULGERIS, A. P.; ELTOFT, T. Goodness-of-fit tests for multilook polarimetric radar data based on the Mellin transform. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 49, n. 7, p. 2764–2781, 2011.
- ANFENSEN, S. N.; ELTOFT, T. Application of the matrix-variate Mellin transform to analysis of polarimetric radar images. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 49, n. 6, p. 2281–2295, 2011.
- ASTORGA, J. M.; GÓMEZ, H. W.; BOLFARINE, H. Slashed generalized exponential distribution. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 46, n. 5, p. 2091–2102, 2017.
- ASTORGA, J. M. et al. Modified slashed generalized exponential distribution. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 49, n. 19, p. 4603–4617, 2020.
- BARRANCO-CHAMORRO, I. et al. A generalized Rayleigh family of distributions based on the modified slash model. **Symmetry**, MDPI, v. 13, n. 7, p. 1226, 2021.
- BARRIOS, L. et al. The slashed power half-normal distribution with applications. **Mathematics**, MDPI, v. 10, n. 9, p. 1528, 2022.
- BILODEAU, M.; BRENNER, D. **Theory of multivariate statistics**. [S.l.]: Springer Science & Business Media, 2008.
- BOLFARINE, H.; SANDOVAL, M. C. **Introdução à inferência estatística**. [S.l.]: SBM, 2010. v. 2.
- BOMBRUN, L.; ANFENSEN, S. N.; HARANT, O. A complete coverage of log-cumulant space in terms of distributions for polarimetric sar data. In: **POLinSAR 2011-5th International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry**. [S.l.: s.n.], 2011. p. 1–8.

- CARMER, S. Optimal significance levels for application of the least significant difference in crop performance trials 1. **Crop Science**, Wiley Online Library, v. 16, n. 1, p. 95–99, 1976.
- CASELLA, G.; BERGER, R. L. **Statistical inference**. [S.l.]: Cengage Learning, 2001.
- CHU, J.; DICKIN, O.; NADARAJAH, S. A review of goodness of fit tests for Pareto distributions. **Journal of Computational and Applied Mathematics**, Elsevier, v. 361, p. 13–41, 2019.
- CORDEIRO, G. M.; CRIBARI-NETO, F. **An Introduction to Bartlett Correction and Bias Reduction**. [S.l.]: Springer Science & Business Media, 2014.
- DAVISON, A. C.; HINKLEY, D. V. **Bootstrap methods and their application**. [S.l.]: Cambridge University Press, 1997.
- DELIGNON, Y.; GARELLO, R.; HILLION, A. Statistical modelling of ocean sar images. **IEE Proceedings-Radar, Sonar and Navigation**, IET, v. 144, n. 6, p. 348–354, 1997.
- DENG, X. et al. Log-cumulants of the finite mixture model and their application to statistical analysis of fully polarimetric uavsar data. **Geo-spatial Information Science**, Taylor & Francis, v. 21, n. 1, p. 45–55, 2018.
- DENG, X.; LÓPEZ-MARTÍNEZ, C. Higher order statistics for texture analysis and physical interpretation of polarimetric sar data. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 13, n. 7, p. 912–916, 2016.
- EFRON, B. Bootstrap methods: Another look at the jackknife. **Annals of Statistics**, v. 7, p. 1–26, 1979.
- GÓMEZ, Y. M.; BOLFARINE, H.; GÓMEZ, H. W. Gumbel distribution with heavy tails and applications to environmental data. **Mathematics and Computers in Simulation**, 2018.
- GUIDORIZZI, H. L. Um curso de cálculo. vol 2. - 5.ed. [reimpr.]. **Rio de Janeiro: LTC**, 2014.
- HÄRDLE, W. K.; SIMAR, L. **Applied multivariate statistical analysis**. [S.l.]: Springer Nature, 2019.
- IRIARTE, Y. A. et al. Slashed moment exponential distribution. **J. Stat. Theory Appl.**, v. 16, n. 3, p. 354–365, 2017.
- IRIARTE, Y. A. et al. Modified slashed-Rayleigh distribution. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 47, n. 13, p. 3220–3233, 2018.
- IRIARTE, Y. A. et al. Slashed Rayleigh distribution. **Revista Colombiana de Estadística**, Universidad Nacional de Colombia., v. 38, n. 1, p. 31–44, 2015.
- JOHNSON, R. A.; WICHERN, D. W. Applied multivariate statistical analysis. 6th. **New Jersey, US: Pearson Prentice Hall**, 2007.



- KHALIL, A. et al. A novel flexible additive Weibull distribution with real-life applications. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 50, n. 7, p. 1557–1572, 2021.
- KHAN, S.; GUIDA, R. Application of Mellin-kind statistics to polarimetric  $\mathcal{G}$  distribution for sar data. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 52, n. 6, p. 3513–3528, 2013.
- KRYLOV, V. A. et al. On the method of logarithmic cumulants for parametric probability density function estimation. **IEEE Transactions on Image Processing**, IEEE, v. 22, n. 10, p. 3791–3806, 2013.
- LEMONTE, A. J.; CORDEIRO, G. M. An extended Lomax distribution. **Statistics**, Taylor & Francis, v. 47, n. 4, p. 800–816, 2013.
- LI, H.; TIAN, W. Slashed Lomax distribution and regression model. **Symmetry**, Multidisciplinary Digital Publishing Institute, v. 12, n. 11, p. 1877, 2020.
- LI, H.-C. et al. On the empirical-statistical modeling of sar images with generalized gamma distribution. **IEEE Journal of selected topics in signal processing**, IEEE, v. 5, n. 3, p. 386–397, 2011.
- LIMA, E. L. **Curso de análise, vol 1**. [S.l.]: Rio de Janeiro: IMPA, 2014.
- LIMA, E. L. **Análise Real Volume 1: Funções de Uma Variável Real**. [S.l.]: Coleção Matemática Universitária. IMPA, Rio de Janeiro, 2017.
- MEYER, P. L. **Probabilidade: aplicações à estatística**. [S.l.]: Livros Técnicos e Científicos Rio de Janeiro, 1983.
- MOOD, A.; GRAYBILL, F. A.; BOES, D. **Introduction to the theory of statistics, McGraw-Hill series in probability and statistics**. [S.l.]: McGraw-Hill New York, NY, 1974.
- NICOLAS, J.-M. Introduction aux statistiques de deuxième espèce: Applications des logs-moments et des logs-cumulants à l'analyse des lois d'images radar. **TS. Traitement du signal**, Lavoisier, v. 19, n. 3, p. 139–167, 2002.
- NICOLAS, J.-M. Application de la transformée de Mellin: étude des lois statistiques de l'imagerie cohérente. **Rapport de recherche, 2006D010**, 2006.
- NOCEDAL, J.; WRIGHT, S. J. **Numerical optimization**. [S.l.]: Springer, 1999.
- OLMOS, N. M. et al. An asymmetric distribution with heavy tails and its expectation-maximization (em) algorithm implementation. **Symmetry**, MDPI, v. 11, n. 9, p. 1150, 2019.
- PEARSON, K. X. contributions to the mathematical theory of evolution.—ii. skew variation in homogeneous material. **Philosophical Transactions of the Royal Society of London.(A.)**, The Royal Society London, n. 186, p. 343–414, 1895.

- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>.
- RAO, C. R. et al. Linear models and generalizations. **Least Squares and Alternatives (3rd edition)** Springer, Berlin Heidelberg New York, 2008.
- ROSSI, R. et al. Upgrading model selection criteria with goodness of fit tests for practical applications. **Entropy**, MDPI, v. 22, n. 4, p. 447, 2020.
- ROUSSAS, G. G. **An Introduction to Probability and Statistical Inference**. [S.l.]: Academic Press, 2014.
- SALINAS, H. S.; IRIARTE, Y. A.; BOLFARINE, H. Slashed exponentiated Rayleigh distribution. **Revista Colombiana de Estadística**, Universidad Nacional de Colombia., v. 38, n. 2, p. 453–466, 2015.
- SCHWARZ, G. Estimating the dimension of a model. **The annals of statistics**, JSTOR, p. 461–464, 1978.
- SEBER, G. A. **A matrix handbook for statisticians**. [S.l.]: John Wiley & Sons, 2008.
- SELIM, M. A. The generalized power generalized Weibull distribution: Properties and applications. **arXiv preprint arXiv:1807.10763**, 2018.
- TAHIR, M. H. et al. The gumbel-lomax distribution: properties and applications. **Journal of Statistical Theory and Applications**, Atlantis Press, v. 15, n. 1, p. 61–79, 2016.
- TISON, C. et al. A new statistical model for markovian classification of urban areas in high-resolution sar images. **IEEE transactions on geoscience and remote sensing**, IEEE, v. 42, n. 10, p. 2046–2057, 2004.
- VASCONCELOS, J. M.; CINTRA, R. J.; NASCIMENTO, A. D. Goodness-of-fit measures based on the Mellin transform for beta generalized lifetime data. **Mathematical Methods in the Applied Sciences**, Wiley Online Library, v. 44, n. 18, p. 14823–14848, 2021.
- WANG, F. A new model with bathtub-shaped failure rate using an additive Burr XII distribution. **Reliability Engineering & System Safety**, Elsevier, v. 70, n. 3, p. 305–312, 2000.

## Apêndices

# APÊNDICE A – Primeiro apêndice

## A.1 Convergência da integral

Um método clássico para verificar a convergência de integrais é o critério da comparação (GUIDORIZZI, 2014), em que dado duas funções  $f$  e  $g$  integráveis em  $[a, s]$ ,  $\forall s > a$ , e que, para todo  $x \geq a$ ,  $0 \leq f(x) \leq g(x)$  é válido:  $\int_0^\infty g(x)dx$  convergente  $\implies \int_0^\infty f(x)dx$  convergente. Temos para  $\alpha, \beta, \lambda, y_i > 0$  e  $w \geq 0$ ,

$$0 \leq \frac{w^{2\lambda}}{(1+w)^{\alpha+1}} \leq w^{2\lambda}.$$

Por outro lado,

$$\int_0^{\frac{y_i}{\beta}} w^{2\lambda} dw = \frac{\left(\frac{y_i}{\beta}\right)^{2\lambda+1}}{2\lambda+1}. \quad (\text{A.1})$$

Então, visto a convergência da integral dada em (A.1), conclui-se, pelo critério da comparação, que

$$\int_0^{\frac{y_i}{\beta}} \frac{w^{2\lambda}}{(1+w)^{\alpha+1}} dw \quad (\text{A.2})$$

é convergente.

## APÊNDICE B – Segundo apêndice

### B.1 Log-verossimilhança

As funções de verossimilhança e de log-verossimilhança para o modelo Slashed Lomax são dadas por:

$$L(\mathbf{y}, \boldsymbol{\theta}) = 2^n \lambda^n \alpha^n \beta^{2\lambda n} \prod_{i=1}^n \left[ y_i^{-(2\lambda+1)} \int_0^{\frac{y_i}{\beta}} \frac{w^{2\lambda}}{(1+w)^{\alpha+1}} dw \right]$$

e

$$\begin{aligned} l(\boldsymbol{\theta}) &= n \log(2) + n \log(\lambda) + n \log(\alpha) + 2\lambda n \log(\beta) - (2\lambda + 1) \sum_{i=1}^n \log(y_i) \\ &+ \sum_{i=1}^n \log \left[ \int_0^{\frac{y_i}{\beta}} \frac{w^{2\lambda}}{(1+w)^{\alpha+1}} dw \right]. \end{aligned}$$

### B.2 Vetor escore

As derivadas parciais da função de log-verossimilhança na direção de  $\alpha$ ,  $\beta$  e  $\lambda$  são dadas por i), ii) e iii), respectivamente:

i)

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \left\{ \frac{\frac{\partial}{\partial \alpha} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw \right]}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw} \right\}.$$

Pela convergência da integral dada em (A.2), temos que:

$$\frac{\partial}{\partial \alpha} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw \right] = \int_0^{\frac{y_i}{\beta}} \frac{\partial}{\partial \alpha} \left[ w^{2\lambda} (1+w)^{-(\alpha+1)} \right] dw.$$

Assim,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^n \left[ \frac{\int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} \log(1+w) dw}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw} \right].$$

ii)

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta} = \frac{2\lambda n}{\beta} + \sum_{i=1}^n \left\{ \frac{\frac{\partial}{\partial \beta} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw \right]}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda} (1+w)^{-(\alpha+1)} dw} \right\}.$$

Utilizando o Teorema Fundamental do Cálculo (GUIDORIZZI, 2014), temos que

$$\frac{\partial}{\partial \beta} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda}(1+w)^{-(\alpha+1)} dw \right] = -y_i^{2\lambda+1} \beta^{-(2\lambda+2)} \left( 1 + \frac{y_i}{\beta} \right)^{-(\alpha+1)}.$$

Desse modo,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \beta} = \frac{2\lambda n}{\beta} - \sum_{i=1}^n \left[ \frac{y_i^{2\lambda+1} \beta^{-(2\lambda+2)} \left( 1 + \frac{y_i}{\beta} \right)^{-(\alpha+1)}}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda}(1+w)^{-(\alpha+1)} dw} \right].$$

iii)

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} = \frac{n}{\lambda} + 2n \log \beta - 2 \sum_{i=1}^n \log(y_i) + \sum_{i=1}^n \left\{ \frac{\frac{\partial}{\partial \lambda} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda}(1+w)^{-(\alpha+1)} dw \right]}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda}(1+w)^{-(\alpha+1)} dw} \right\}.$$

O cálculo de  $\frac{\partial}{\partial \lambda} \left[ \int_0^{\frac{y_i}{\beta}} w^{2\lambda}(1+w)^{-(\alpha+1)} dw \right]$ , é dado por:

$$\int_0^{\frac{y_i}{\beta}} \frac{\partial}{\partial \lambda} \left[ w^{2\lambda}(1+w)^{-(\alpha+1)} \right] dw = 2 \int_0^{\frac{y_i}{\beta}} w^{2\lambda}(1+w)^{-(\alpha+1)} \log(w) dw.$$

Assim sendo,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \lambda} = \frac{n}{\lambda} + 2n \log \beta - 2 \sum_{i=1}^n \log(y_i) + \sum_{i=1}^n \left[ 2 \frac{\int_0^{\frac{y_i}{\beta}} w^{2\lambda}(1+w)^{-(\alpha+1)} \log(w) dw}{\int_0^{\frac{y_i}{\beta}} w^{2\lambda}(1+w)^{-(\alpha+1)} dw} \right].$$

# APÊNDICE C – Terceiro apêndice

## C.1 Estatística $T^2$ de Hotelling

Neste apêndice, fornecemos as expressões para os elementos da matriz assintótica dos LCs estimados. Considere a relação analítica entre LMs e LCs dada em (2.5),  $\mathbf{M}_1$  em (3.6) e  $\mathbf{K}_1$  em (3.9). Então, temos:

$$\mathbf{K}_1 = \mathbf{J}_1^\top \cdot \mathbf{M}_1 \cdot \mathbf{J}_1,$$

com a matriz jacobiana das transformações de momentos para cumulantes dada por (ANFINSEN; DOULGERIS; ELTOFT, 2011)

$$\mathbf{J}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2m & 1 & 0 \\ -3(m_2 - 2m_1^2) & -3m_1 & 1 \end{bmatrix} \quad (\text{C.1})$$

e

$$\mathbf{M}_1 = \begin{bmatrix} \tilde{k}_2 & \tilde{k}_3 + 2\tilde{k}_1\tilde{k}_2 & M_{13} \\ \tilde{k}_3 + 2\tilde{k}_1\tilde{k}_2 & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}, \quad (\text{C.2})$$

em que

$$M_{13} = \tilde{k}_4 + 3\tilde{k}_1\tilde{k}_3 + 3\tilde{k}_2^2 + 3\tilde{k}_1^2\tilde{k}_2,$$

$$M_{22} = \tilde{k}_4 + 4\tilde{k}_1\tilde{k}_3 + 2\tilde{k}_2^2 + 4\tilde{k}_1^2\tilde{k}_2$$

$$M_{23} = \tilde{k}_5 + 5\tilde{k}_1\tilde{k}_4 + 9\tilde{k}_2\tilde{k}_3 + 9\tilde{k}_1^2\tilde{k}_3 + 12\tilde{k}_1\tilde{k}_2^2 + 6\tilde{k}_1^3\tilde{k}_2$$

$$M_{33} = \tilde{k}_6 + 6\tilde{k}_1\tilde{k}_5 + 15\tilde{k}_2\tilde{k}_4 + 15\tilde{k}_1^2\tilde{k}_4 + 9\tilde{k}_3^2 + 54\tilde{k}_1\tilde{k}_2\tilde{k}_3 + 18\tilde{k}_1^3\tilde{k}_3 + 15\tilde{k}_2^3 + 36\tilde{k}_1^2\tilde{k}_2^2 + 9\tilde{k}_1^4\tilde{k}_2.$$

Portanto,

$$\mathbf{K}_1 = \begin{bmatrix} \tilde{k}_2 & \tilde{k}_3 & \tilde{k}_4 \\ \tilde{k}_3 & \tilde{k}_4 + 2\tilde{k}_2^2 & \tilde{k}_5 + 6\tilde{k}_2\tilde{k}_3 \\ \tilde{k}_4 & \tilde{k}_5 + 6\tilde{k}_2\tilde{k}_3 & \tilde{k}_6 + 9\tilde{k}_2\tilde{k}_4 + 9\tilde{k}_3^2 + 6\tilde{k}_2^3 \end{bmatrix}. \quad (\text{C.3})$$

Com base na matriz (C.3), derivamos a matriz de covariância assintótica para o modelo Slashed Lomax, isto é:

$$\mathbf{K} = \begin{bmatrix} \tilde{k}_4 + 2\tilde{k}_2^2 & \tilde{k}_5 + 6\tilde{k}_2\tilde{k}_3 \\ \tilde{k}_5 + 6\tilde{k}_2\tilde{k}_3 & \tilde{k}_6 + 9\tilde{k}_2\tilde{k}_4 + 9\tilde{k}_3^2 + 6\tilde{k}_2^3 \end{bmatrix} = \begin{bmatrix} \delta_{22} & \delta_{23} \\ \delta_{32} & \delta_{33} \end{bmatrix},$$

em que

$$\begin{aligned}\delta_{22} &= \psi(3, 1) + \psi(3, \alpha) + \frac{6}{(2\lambda)^4} + 2 \left[ \psi(1, 1) + \psi(1, \alpha) + \frac{1}{(2\lambda)^2} \right]^2, \\ \delta_{23} &= \delta_{32} = \psi(4, 1) - \psi(4, \alpha) + \frac{24}{(2\lambda)^5} + 6 \left[ \psi(1, 1) + \psi(1, \alpha) + \frac{1}{(2\lambda)^2} \right] \\ &\quad \cdot \left[ \psi(2, 1) - \psi(2, \alpha) + \frac{2}{(2\lambda)^3} \right], \\ \delta_{33} &= \psi(5, 1) + \psi(5, \alpha) + \frac{120}{(2\lambda)^6} + 9 \left[ \psi(1, 1) + \psi(1, \alpha) + \frac{1}{(2\lambda)^2} \right] \left[ \psi(3, 1) + \psi(3, \alpha) + \frac{6}{(2\lambda)^4} \right] \\ &\quad + 9 \left[ \psi(2, 1) - \psi(2, \alpha) + \frac{2}{(2\lambda)^3} \right]^2 + 6 \left[ \psi^3(1, 1) + \psi^3(1, \alpha) + \frac{1}{(2\lambda)^2} \right]^3.\end{aligned}$$

Se  $\mathbf{K}$  é uma matriz não-singular, então seu inverso é dada por:

$$\mathbf{K}^{-1} = \frac{1}{\delta_{33}\delta_{22} - \delta_{23}^2} \begin{bmatrix} \delta_{33} & -\delta_{32} \\ -\delta_{23} & \delta_{22} \end{bmatrix}.$$

Do contrário, se  $|\mathbf{K}| = 0$ , aplicamos a inversa generalizada de Moore-Pearson (RAO et al., 2008; SEBER, 2008). Por fim, a estatística  $T^2$  de Hotelling é dada por:

$$\begin{aligned}T_{SL}^2 &= n \left( \begin{bmatrix} \hat{k}_2 \\ \hat{k}_3 \end{bmatrix} - \begin{bmatrix} \tilde{k}_2 \\ \tilde{k}_3 \end{bmatrix} \right)^\top \widehat{\mathbf{K}}^{-1} \left( \begin{bmatrix} \hat{k}_2 \\ \hat{k}_3 \end{bmatrix} - \begin{bmatrix} \tilde{k}_2 \\ \tilde{k}_3 \end{bmatrix} \right) \\ &= \frac{n}{\hat{\delta}_{22}\hat{\delta}_{33} - \hat{\delta}_{23}^2} \left[ \hat{\delta}_{33} \left( \hat{k}_2 - \tilde{k}_2 \right)^2 + \hat{\delta}_{22} \left( \hat{k}_3 - \tilde{k}_3 \right)^2 - 2\hat{\delta}_{23} \left( \hat{k}_2 - \tilde{k}_2 \right) \left( \hat{k}_3 - \tilde{k}_3 \right) \right].\end{aligned}$$



## C.2 Cálculo das estatísticas

A seguir, listamos a sequência de algoritmos utilizados para determinar na prática a estatística  $T^2$  de Hotelling. De acordo com (3.11), inicialmente obtêm-se  $\hat{k}_2$  e  $\hat{k}_3$  (estimativas de  $\tilde{k}_2$  e  $\tilde{k}_3$ ), seguido da matriz  $\mathbf{K}_1^{-1}$ .

---

**Algoritmo 2:** Cômputo de  $\hat{k}_2$  e  $\hat{k}_3$

---

**Passo 1.** Considere o vetor de parâmetros  $\boldsymbol{\theta} = [\alpha \ \beta \ \lambda]^\top$ ;

**Passo 2.** Compute os três primeiros LCs e os três primeiros LMs;

**Passo 3.** Determine  $\hat{\boldsymbol{\theta}}$  utilizando a função `BBsolve` para resolver o sistema de equações;

**Passo 4.** Obtenha  $\hat{k}_2$  e  $\hat{k}_3$  de acordo com  $\hat{k}_v = g_v(\hat{m}_1, \dots, \hat{m}_v)$ .

---



---

**Algoritmo 3:** Cômputo de  $\widehat{\mathbf{K}}_1^{-1}$

---

**Passo 1.** Compute a matriz  $\mathbf{J}_1$  conforme (C.1);

**Passo 2.** Compute a matriz  $\mathbf{M}_1$  de acordo (C.2);

**Passo 3.** Compute  $\widehat{\mathbf{K}}_1^{-1} = \widehat{\mathbf{J}}_1^\top \widehat{\mathbf{M}}_1 \widehat{\mathbf{J}}_1$ ;

**Passo 4.** Se  $\widehat{\mathbf{K}}_1$  é não-singular, compute a sua inversa usual  $\widehat{\mathbf{K}}_1^{-1}$ , do contrário, compute a inversa generalizada de  $\widehat{\mathbf{K}}_1$ .

---