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

RENILMA PEREIRA DA SILVA

Distribuições de probabilidade generalizadas: propriedades e refinamento de inferências

Recife
2018

RENILMA PEREIRA DA SILVA

DISTRIBUIÇÕES DE PROBABILIDADE GENERALIZADAS: PROPRIEDADES E REFINAMENTO DE INFERÊNCIAS

Tese apresentada ao Programa de Pós-graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco como requisito para obtenção do grau de Doutora em Estatística.

Orientadora: Profa. Dra. Audrey Helen Mariz de Aquino Cysneiros

Recife
2018

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

S586d Silva, Renilma Pereira da
Distribuições de probabilidade generalizadas: propriedades e refinamento de inferências / Renilma Pereira da Silva. – 2018.
105 f.: il., fig., tab.

Orientadora: Audrey Helen Mariz de Aquino Cysneiros.
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2018.
Inclui referências e apêndices.

1. Estatística matemática. 2. Teoria assintótica. I. Cysneiros, Audrey Helen Mariz de Aquino (orientadora). II. Título.

519.5 CDD (23. ed.) UFPE- MEI 2018-042

RENILMA PEREIRA DA SILVA

DISTRIBUIÇÕES DE PROBABILIDADE GENERALIZADAS: PROPRIEDADES E REFINAMENTO DE INFERÊNCIAS

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 2018.

BANCA EXAMINADORA

Prof.ª Audrey Helen Mariz de Aquino Cysneiros
UFPE

Prof. Aldo William Medina Garay
UFPE

Prof. Roberto Ferreira Manghi
UFPE

Prof.ª Mariana Correia de Araújo
UFRN

Prof.ª Juliana Freitas Pires
UFPB

Ao meu esposo Claudio Tablada, dedico. Pelo seu amor, paciência, companheirismo e, principalmente, por tornar os meus dias mais felizes.

Agradecimentos

Ao escrever esses agradecimentos, faço uma reflexão sobre todas as dificuldades e batalhas travadas durante toda a minha trajetória acadêmica, principalmente nesses quatro anos de doutorado. Sem dúvida, esta etapa da minha vida foi um grande desafio, mas graças às contribuições de pessoas mais que especiais consegui concluir este trabalho.

Gostaria de expressar meus agradecimentos, primeiramente, à Deus por ter me ouvido, nos momentos de desânimo, quando orei pedindo-lhe forças, por me fazer uma pessoa persistente e assim alcançar os meus objetivos. Ao meu esposo, Claudio Tablada, por ser paciente comigo nos momentos difíceis, por ser o meu melhor amigo e pelas suas sugestões para esta tese.

Agradeço, aos meus queridos pais, Renivaldo Pereira e Maria Ramos, que mesmo com poucos recursos financeiros e pouca instrução acadêmica batalharam muito para que meus irmãos e eu pudéssemos estudar. Aos meus irmãos, Renivaldo Júnior, Rumenick e Reniery pelo apoio incondicional e companheirismo. Aos meus sobrinhos, por alegrar a minha vida. A minha avó, Francisca Cardoso, por sempre me incluir em suas orações.

Sou grata a minha orientadora, Profa. Audrey Cysneiros, pela orientação durante o mestrado e doutorado, e sobretudo, por me compreender, aconselhar e compartilhar comigo os seus conhecimentos.

Não poderia deixar de agradecer, a todos os colegas de doutorado, especialmente a Enai, Pedro Rafael, Daniele Brito, Jéssica Rivas, Giannini e Fernando Arturo, pela boa convivência e parceria nos estudos.

Gostaria de expressar a minha gratidão, a todos os professores que fizeram parte da minha formação acadêmica, especialmente aos professores: Gauss Cordeiro, Francisco Cribari, Francisco Cyneiros e Getúlio Amorim. Agradeço também a Valéria Bittencourt, pela sua atenção, gentileza e por fazer seu trabalho com tanta eficiência.

Finalmente agradeço aos membros da banca examinadora, pelas suas valiosas contribuições e à CAPES, pelo apoio financeiro.

Resumo

As distribuições de probabilidade são aplicadas na modelagem de dados em diversas áreas. No entanto, em algumas situações as distribuições clássicas podem não ser as mais apropriadas para explicar o fenômeno estudado. Nestes casos, as distribuições generalizadas ou estendidas podem ser utilizadas, uma vez que são mais flexíveis. No contexto de distribuições de probabilidade generalizadas, esta tese, primeiramente, apresenta um estudo sobre refinamento de inferências nas distribuições Lomax exponencializada e BurrX escalonada. Com base na função de verossimilhança perfilada e em algumas de suas versões modificadas, faz-se estimação pontual e teste de hipóteses para uma parte do vetor paramétrico. As versões modificadas da função de verossimilhança perfilada são derivadas com o objetivo de atenuar o impacto dos parâmetros de perturbação sobre as inferências. Para a distribuição Lomax exponencializada, obteve-se diferentes ajustes à função de verossimilhança perfilada ao considerar um e dois parâmetros de interesse e amostra completa (sem censura). Enquanto que, para a distribuição BurrX escalonada, obteve-se ajustes à função de verossimilhança perfilada para o parâmetro de forma em dois enfoques: amostra completa e amostra censurada do tipo II. Resultados de simulação de Monte Carlo são apresentados com o objetivo de avaliar e comparar os desempenhos dos estimadores de máxima verossimilhança, assim como do teste da razão de verossimilhanças baseado nas diferentes funções. Além disso, incluímos na comparação um teste baseado na estatística da razão de verossimilhanças usual que utiliza o quantil da distribuição empírica da estatística da razão de verossimilhanças sob hipótese nula, em que esse quantil é obtido por meio da técnica bootstrap. Considerou-se ainda, a estatística da razão de verossimilhanças Bartlett bootstrap, a qual utiliza um fator de correção obtido numericamente através da técnica bootstrap. A teoria desenvolvida é ilustrada através de aplicações utilizando dados reais. Finalmente, neste trabalho é proposto um novo modelo de quatro parâmetros, denominado de distribuição transmutada Marshall-Olkin estendida Lomax (TMOEL). Mostra-se que a função densidade de probabilidade da TMOEL pode ser escrita como uma combinação linear de densidades de Lomax exponencializada e com isso, derivam-se algumas propriedades matemáticas e estatísticas. Realiza-se um estudo de simulação de Monte Carlo para avaliar o comportamento dos estimadores de máxima verossimilhanaça dos parâmetros do modelo. Uma aplicação a um conjunto de dados reais ilustra o uso da nova distribuição.

Palavras-chave: Bootstrap paramétrico. Distribuições generalizadas. Refinamento de inferências. Verossimilhança perfilada. Verossimilhança perfilada modificada.

Abstract

Probability distributions are utilized for modeling a wide variety of applications in several areas. However, in some situations the classic distributions can not be appropriated for explain the phenomenon studied. In these cases, generalized or extended distributions can be used, once they are more flexible. In the context of generalized probability distributions, this thesis first presents a study on profile likelihood adjustments in the exponentiated Lomax (EL) and scaled Burr type X distributions. Based on the profile likelihood function and some of its modified versions, we consider point estimation and hypothesis test for a part of the parameter vector. The objective of profile likelihood adjustments are reduce the impact of the nuisance parameters on the likelihood-based inference. In the case of the EL distribution, we obtain different adjustments for the profile likelihood to considering one and two parameters of interest and complete sample (uncensored data). Whereas, for the scaled Burr type X distribution, we obtain different adjustments for the profile likelihood on shape parameter to considering complete samples and type II censoring. Numerical results on estimation and hypotheses test are provided with the aim of avaliate and compare the performance of the maximum likelihood estimators, as the tests based on the different adjustments. In addition, we include in the comparison a test based on the usual likelihood ratio that uses the quantile of the empirical distribution of the likelihood ratio under null hypothesis, in which this quantile is obtained through the bootstrap method. Empirical applications are presented and discussed. Finally, we propose a four-parameter generalized model called the transmuted Marshall-Olkin extended Lomax (TMOEL). We prove that the probability density function of the TMOEL can be written as a linear combination of exponential densities of Lomax and so we derive several mathematical and statistical properties of the new distribution. Numerical results on estimation are provided. We prove empirically the applicability of the new model to real data.

Keywords: Generalized distributions. Improved inference. Modified profile likelihood. Parametric bootstrap. Profile likelihood.

Lista de Figuras

Lista de Tabelas

Sumário

1 Introdução

Distribuições de probabilidade são comumente utilizadas para descrever fenômenos do mundo real. Por exemplo, um pesquisador pode estar interessado em modelar os tempos de falhas de determinado equipamento, a proporção de pessoas infectadas com alguma doença, o tempo de ação de um medicamento, dentre outros. Porém, as distribuições clássicas podem não ser as mais adequadas para explicar tais fenômenos. Por essa razão, tem crescido o interesse em obter novas distribuições a partir de um modelo de base ("*baseline*").

Em geral, as distribuições generalizadas são mais flexíveis que o modelo de base, pois as generalizações, além de incluírem a distribuição de base, são mais tratáveis para estudar as propriedades de cauda. Observou-se ainda que, na maioria das vezes, a inserção de um parâmetro de forma à distribuição de base implica em uma redução nos valores de alguns critérios de bondade de ajuste, como por exemplo, o critério de informação de Akaike (TAHIR; NADARAJAH, 2015).

Na literatura, encontramos diversas técnicas para construir famílias de distribuições generalizadas. Uma família de distribuições bastante difundida é a alternativa Lehmann tipo I, também conhecida como família exponencializada-G (Exp-G). Essa família é obtida elevando $G(x)$ (função de distribuição acumulada de uma variável aleatória X) a uma potência $a > 0$. Outra família de distribuições que tem recebido bastante atenção nos últimos anos é a família transmutada generalizada (T-G), proposta por Shaw e Buckley (2009), que introduz um parâmetro à distribuição de base usando um mapa de transmutação quadrático (QRTM) visando obter um modelo mais flexível.

A distribuição Lomax (LOMAX, 1954) tem sido útil para modelar dados em diferentes áreas, inclusive dados de tempo de vida. Assim, várias generalizações dessa distribuição foram propostas (ASHOUR; ELTEHIWY, 2013; LEMONTE; CORDEIRO, 2013; RADY *et al.*, 2016). Dentre essas generalizações, podemos destacar as distribuições de três parâmetros: Marshall-Olkin estendida Lomax (MOEL), proposta por Ghitany *et al.* (2007) e a distribuição Lomax exponencializada (LE), introduzida por Abdul-Moniem e Abdel-Hameed (2012).

Outra distribuição bastante conhecida é a Burr tipo X, proposta por Burr (1942). Por ter suporte nos reais positivos, assim como a distribuição Lomax e suas generalizações, também pode ser usada para explicar dados de tempo de vida. Surles e Padgett (2001) adicionaram um parâmetro de escala à distribuição Burr tipo X e obtiveram a distribuição BurrX escalonada, a qual foi utilizada para fazer inferência sobre a confiabilidade entre dois tipos de fibras de carbono. É importante salientar que a BurrX escalonada generaliza a distribuição Rayleigh quando

o parâmetro de forma da primeira é igual a um. Por essa razão, a distribuição BurrX escalonada é denominada por alguns autores como distribuição Rayleigh generalizada ou Rayleigh exponencializada (RAQAB; MADI, 2011; ABD-ELFATTAH, 2011; AHMAD *et al.*, 2017).

Ao estudar um fenômeno, supomos um modelo e muitas vezes estamos interessados em fazer inferências apenas sobre alguns parâmetros que indexam esse modelo. Assim, denominamos esses parâmetros de *parâmetros de interesse* e os demais, de *parâmetros de perturbação*. Neste caso, as inferências podem ser feitas com base na função de verossimilhança perfilada. A ideia da verossimilhança perfilada consiste em substituir o parâmetro de perturbação por uma estimativa consistente na verossimilhança original, que pode ser a estimativa de máxima verossimilhança para valores fixos do parâmetro de interesse. Logo, essa função resultante depende apenas desse parâmetro.

A função de verossimilhança perfilada não é uma função de verossimilhança genuína, assim algumas propriedades básicas de uma função de verossimilhança não são válidas. Por exemplo, a função escore não necessariamente tem média zero e a informação pode apresentar vício. Além disso, se o tamanho da amostra não for suficientemente grande ou houver muitos parâmetros de perturbação, a aproximação da distribuição da estatística da razão de verossimilhanças perfilada, sob a hipótese nula, pela distribuição qui-quadrado pode ser bastante pobre.

Visando atenuar as limitações da função de verossimilhança perfilada, diversos trabalhos foram propostos. Barndorff-Nielsen (1983) propôs uma versão modificada da função de verossimilhança perfilada que usa a fórmula p^* (BARNDORFF-NIELSEN, 1980), a qual aproxima a função densidade de probabilidade do estimador de máxima verossimilhança condicionado a uma estatística ancilar. Porém, existem certas dificuldades no cálculo dessa função de verossimilhança perfilada modificada, como por exemplo, a especificação da estatística ancilar, que dependendo da situação pode ser inviável a sua obtenção.

Fraser e Reid (1995), Fraser *et al.* (1999) e Severini (1999) propuseram aproximações para a função de verossimilhança perfilada modificada introduzida por Barndorff-Nielsen (1983). A aproximação dada por Fraser e Reid (1995) e Fraser *et al.* (1999) é baseada em uma estatística aproximadamente ancilar, enquanto que a aproximação proposta por Severini (1999) é baseada em uma matriz de covariâncias empíricas. Uma descrição sobre função de verossimilhança perfilada modificada e suas aproximações pode ser encontrada em Severini (2000).

Cox e Reid (1987) propuseram uma função de verossimilhança perfilada modificada que não é invariante sob reparametrizações e requer ortogonalidade dos parâmetros. Na maioria das distribuições generalizadas não há ortogonalidade dos parâmetros e a obtenção de uma parametrização ortogonal pode ser bastante complicada. Logo, nesses casos, a função de verossimilhança perfilada modificada proposta por Cox e Reid (1993) é a mais adequada, pois para o seu cálculo os parâmetros do modelo não precisam ser ortogonais.

Este trabalho está organizado em seis capítulos e três apêndices. No Capítulo 2 é definida a função de verossimilhança perfilada e algumas de suas versões modificadas que serão abordadas ao longo da tese. Portanto, esse capítulo deve ser lido antes dos Capítulos 3 e 4. No entanto, os Capítulos 3, 4 e 5 são independentes e estão apresentados no formato de artigo, isso significa que eles podem ser lidos separadamente, dessa forma alguns resultados e notações são apresentados mais de uma vez. Apesar do nosso estudo estar relacionado ao tema de distribuições generalizadas, as ideias e os objetivos apresentados nos Capítulos 3 e 4 são distintos daqueles apresentados no Capítulo 5. No Capítulo 3 e 4, fizemos refinamento de inferências em duas distribuições generalizadas já propostas na literatura, enquanto que no capítulo 5, propomos um novo modelo e estudamos algumas de suas propriedades.

Nos Capítulos 3 e 4 estudamos aspectos inferenciais da distribuição LE e BurrX escalonada, respectivamente. Com base nas funções de verossimilhança propostas por Cox e Reid (1993), Fraser e Reid (1995), Fraser *et al.* (1999) e Severini (1999), objetivamos obter três versões modificadas da função de verossimilhança perfilada que forneçam inferências mais confiáveis para dois parâmetros da distribuição LE e para o parâmetro de forma da distribuição BurrX escalonada. Especificamente, no Capítulo 3 consideramos dois cenários: no primeiro, fizemos inferências para um dos parâmetros de forma da distribuição LE supondo o parâmetro de escala conhecido; no segundo, além do parâmetro de forma, tratamos também o parâmetro de escala como parâmetro de interesse. No Capítulo 4, utilizamos os diferentes ajustes para a função de verossimilhança perfilada para fazer inferências sobre o parâmetro de forma da distribuição BurrX escalonada ao considerar amostra completa e amostra censurada do tipo II.

Tanto no Capítulo 3 quanto no Capítulo 4, utilizando os diferentes ajustes para a função de verossimilhança perfilada, realizamos estimação por máxima verossimilhança e testes de hipóteses baseados na estatística da razão de verossimilhanças. Consideramos também o teste baseado na estatística da razão de verossimilhanças bootstrap Bartlett (ROCKE, 1989) e o teste da razão de verossimilhanças bootstrap usual que utiliza o quantil da distribuição empírica nula, obtido através do método bootstrap (EFRON; TIBSHIRANI, 1994). Apresentamos um estudo de simulação para avaliar e comparar os desempenhos da função de verossimilhança perfilada e suas versões modificadas no processo de estimação pontual e teste de hipóteses. Mostramos através de aplicações à conjuntos de dados reais a utilidade dos diferentes ajustes.

No Capítulo 5, propomos uma nova distribuição chamada *transmutada Marshall-Olkin estendida Lomax* (TMOEL), a qual é obtida tomando a distribuição Marshall-Olkin estendida Lomax (GHITANY *et al.*, 2007) como modelo de base na família T-G de distribuições. Discutimos várias propriedades da nova distribuição, incluindo a função taxa de falha, os momentos ordinários e incompletos, a função característica, os desvios médios e as curvas de Bonferroni e Lorenz. Apresentamos um estudo de simulação para avaliar o desempenho dos estimadores de

máxima verossimilhança dos parâmetros que indexam a distribuição proposta. Mostramos empiricamente a flexibilidade do novo modelo através de uma aplicação à um conjunto de dados reais.

O Capítulo 6 é destinado às considerações finais deste trabalho. O Apêndice 6 contém as segundas derivadas do logaritmo da função de verossimilhança da distribuição LE. No Apêndice 6, discutimos alguns aspectos computacionais referentes aos capítulos 3 e 4. Finalmente, no Apêndice 2, encontram-se as funções de densidade de probabilidade das distribuições utilizadas para a comparação com a distribuição TMOEL no Capítulo 5.

2 Ajustes para a função de verossimilhança perfilada

Seja $\boldsymbol{\theta} = (\mu, \nu)^\top$ um vetor de parâmetros que indexa uma distribuição de probabilidade e $L(\boldsymbol{\theta})$ a função de verossimilhança associada a essa distribuição, em que μ e ν podem ser vetores ou escalares.

No que segue, μ será parâmetro de interesse e ν, de perturbação. Dessa forma, a função de verossimilhança perfilada é dada por

$$L_p(\mu) = L(\mu, \hat{\nu}_\mu),$$

em que $\hat{\nu}_\mu$ é a estimativa de máxima verossimilhança de ν para μ fixo. Em outras palavras, $\hat{\nu}_\mu$ é a raiz da equação $\partial\ell(\boldsymbol{\theta})/\partial\nu = 0$, tal que $\ell(\boldsymbol{\theta}) = \log[L(\boldsymbol{\theta})]$. Temos ainda que L_p é invariante sob reparametrizações da forma $(\mu, \nu) \to (\xi, \rho)$, em que $\xi = \xi(\mu)$ e $\rho = \rho(\mu, \nu)$.

Seja $\ell_p(\mu) = \log[L_p(\mu)]$. Então, o estimador de máxima verossimilhança perfilado de μ, $\hat{\mu}_p$, é obtido maximizando $\ell_p(\mu)$. Assim, a estatística da razão de verossimilhanças para testar $\mathcal{H}_0 : \ \mu = \mu_0$ *versus* $\mathcal{H}_1 : \ \mu \neq \mu_0$ é dada por

$$LR_p = 2\left\{\ell_p(\hat{\mu}_p) - \ell_p(\mu_0)\right\}.$$

Sob as condições gerais de regularidade (CORDEIRO, 1999) e sob a hipótese nula, a estatística LR_p tem distribuição assintótica qui-quadrado com q graus de liberdade (χ^2_q), em que q é a dimensão do vetor μ_0.

É importante ressaltar que, quando utilizamos a função de verossimilhança perfilada para realizar inferências sobre um modelo estatístico qualquer, estamos tratando ν como se fosse um parâmetro conhecido e igual a $\hat{\nu}_\mu$. Porém, esse procedimento pode não ser o mais coerente caso os dados não tenham informação suficiente sobre o parâmetro de perturbação.

Outro ponto a ser considerado é o fato de que a função de verossimilhança perfilada não deve ser tratada como uma função de verossimilhança genuína, uma vez que foi derivada a partir da substituição de ν por $\hat{\nu}_\mu$ na verossimilhança usual. Assim, algumas propriedades básicas da verossimilhança usual, tais como função escore e informação não viesadas não são válidas (FERRARI *et al.*, 2005). Entretanto, a função de verossimilhança perfilada tem algumas propriedades interessantes (PACE; SALVAN, 1997), dentre elas podemos citar:

(i) $\hat{\mu}_p = \hat{\mu}$, em que $\hat{\mu}$ é a estimativa de máxima verossimilhança de μ;

(ii) ao testar hipóteses sobre μ, a estatística da razão de verossimilhanças baseada em $\ell_p(\mu)$ é igual à baseada na $\ell(\mu, \nu)$ (logaritmo da função de verossimilhança usual).

Conforme foi mencionado, $L_p(\mu)$ apresenta alguns problemas provenientes do fato desta função não ser uma função de verossimilhança genuína, então com o objetivo de contornar possíveis prejuízos na qualidade das aproximações envolvidas nas inferências, foram propostas na literatura diversas versões modificadas da função de verossimilhança perfilada. Essas versões foram construídas visando a redução dos vieses da função escore perfilada e da correspondente informação. Várias dessas versões são discutidas por Pace e Salvan (1997), Severini (2000) e algumas delas serão abordadas neste capítulo.

2.1 Aproximações para ℓ_{BN}

Barndorff-Nielsen (1983) propôs uma função de verossimilhança perfilada modificada que pode ser deduzida como uma aproximação de uma verossimilhança marginal ou condicional para μ, se uma destas duas funções existir. Em qualquer dos casos, a dedução é feita com base na fórmula p^* (BARNDORFF-NIELSEN, 1980; BARNDORFF-NIELSEN, 1983) que é uma aproximação para a função de densidade condicional do estimador de máxima verossimilhança do vetor de parâmetros do modelo, dada uma estatística ancilar. Assim, a função de verossimilhança perfilada modificada proposta por Barndorff-Nielsen (1983) é dada por

$$L_{BN}(\mu) = \left| \frac{\partial \hat{\nu}_\mu}{\partial \hat{\nu}} \right|^{-1} |j_{\nu\nu}(\mu, \hat{\nu}_\mu)|^{-1/2} L_p(\mu), \tag{2.1}$$

em que $j_{\nu\nu}(\mu, \nu) = -\partial^2 \ell(\mu, \nu)/\partial\nu\partial\nu^\top$ é a informação observada e $\partial\hat{\nu}_\mu/\partial\hat{\nu}$ é a matriz de derivadas parciais de $\hat{\nu}_\mu$ em relação a $\hat{\nu}$.

Tomando o logaritmo em (2.1), obtém-se

$$\ell_{BN}(\mu) = \ell_p(\mu) - \log \left| \frac{\partial \hat{\nu}_\mu}{\partial \hat{\nu}} \right| - \frac{1}{2} \log |j_{\nu\nu}(\mu, \hat{\nu}_\mu)|. \tag{2.2}$$

É possível expressar $\partial\hat{\nu}_\mu/\partial\hat{\nu}$ em termos de uma derivada do espaço amostral do logaritmo da função de verossimilhança da seguinte forma:

$$\partial\hat{\nu}_\mu/\partial\hat{\nu} = j_{\nu\nu}(\mu, \hat{\nu}_\mu)^{-1} \ell_{\nu;\hat{\nu}}(\mu, \hat{\nu}_\mu; \hat{\mu}, \hat{\nu}, a),$$

em que a é uma estatística ancilar e $\ell_{\nu;\hat{\nu}}(\mu, \hat{\nu}_\mu; \hat{\mu}, \hat{\nu}, a) = \partial^2 \ell(\mu, \hat{\nu}_\mu; \hat{\mu}, \hat{\nu}, a)/\partial\nu\partial\hat{\nu}$. Dessa

forma, podemos reescrever (2.2) como

$$\ell_{BN}(\mu) = \ell_p(\mu) + \frac{1}{2}\log|j_{\nu\nu}(\mu, \hat{\nu}_\mu)| - \log|\ell_{\nu;\hat{\nu}}(\mu, \hat{\nu}_\mu; \hat{\mu}, \hat{\nu}, a)|. \tag{2.3}$$

Assim como $L_p(\cdot)$, a função dada por (2.3) também é invariante sob reparametrizações da forma $(\mu, \nu) \to (\xi, \rho)$, em que $\xi = \xi(\mu)$ e $\rho = \rho(\mu, \nu)$. Entretanto, o seu cálculo requer a especificação de uma estatística ancilar, a, a fim de que $(\hat{\mu}, \hat{\nu}, a)$ seja uma estatística suficiente minimal do modelo. Porém, muitas vezes a obtenção da estatística a é bastante complicada. Por essa razão, diversas aproximações para $\ell_{BN}(\mu)$ foram propostas. Duas delas serão apresentadas a seguir.

Uma aproximação para $\ell_{BN}(\mu)$ baseada em uma estatística aproximadamente ancilar foi proposta por Fraser *et al.* (1999). O logaritmo da função de verossimilhança, neste caso, é dado por

$$\tilde{\ell}_{BN}(\mu) = \ell_p(\mu) + \frac{1}{2}\log|j_{\nu\nu}(\mu, \hat{\nu}_\mu)| - \log|\ell_{\nu;X}(\mu, \hat{\nu}_\mu)\widehat{V}_\nu|, \tag{2.4}$$

em que $\ell_{\nu;X}(\mu, \hat{\nu}_\mu) = \partial\ell_\nu(\mu, \nu)/\partial\boldsymbol{x}^\top$, $\ell_\nu(\mu, \nu)$ é a função escore para ν, $\boldsymbol{x}^\top = (x_1, \ldots, x_n)$ é a amostra observada,

$$\widehat{V}_\nu = \left(-\frac{\partial F(x_1; \hat{\mu}, \hat{\nu})/\partial\hat{\nu}}{f(x_1; \hat{\mu}, \hat{\nu})}, \ldots, -\frac{\partial F(x_n; \hat{\mu}, \hat{\nu})/\partial\hat{\nu}}{f(x_n; \hat{\mu}, \hat{\nu})}\right)^\top,$$

tal que $f(x_j; \mu, \nu)$ é a função densidade de probabilidade avaliada em x_j e $F(x_j; \mu, \nu)$ é a função de distribuição acumulada. O estimador de máxima verossimilhança referente à (2.4) será denotado por $\widehat{\tilde{\mu}}_{BN}$.

Uma outra aproximação foi proposta por Severini (1999). Nesta versão, $\ell_{\nu;\hat{\nu}(\mu,\hat{\nu}_\mu;\hat{\mu},\hat{\nu},a)}$ é aproximada por meio das covariâncias empíricas. Esta pode ser facilmente calculada e é conveniente em situações em que há dificuldade para calcular a esperança do produto de derivadas do logaritmo da verossimilhança. Assim, o logaritmo da função de verossimilhança baseada nesta proposta é dada por

$$\breve{\ell}_{BN}(\mu) = \ell_p(\mu) + \frac{1}{2}\log|j_{\nu\nu}(\mu, \hat{\nu}_\mu)| - \log|\breve{I}_\nu(\mu, \hat{\nu}_\mu; \hat{\mu}, \hat{\nu})|, \tag{2.5}$$

em que

$$\breve{I}_\nu(\mu, \nu; \hat{\mu}, \hat{\nu}) = \sum_{j=1}^{n} l_\nu^{(j)}(\mu, \nu) l_\nu^{(j)}(\hat{\mu}, \hat{\nu})^\top. \tag{2.6}$$

Em (2.6), $l_\nu^{(j)}(\cdot)$ é a função escore para ν baseada na j-ésima observação. O estimador de máxima verossimilhança correspondente à (2.5) será denotado por $\widehat{\widetilde{\mu}}_{BN}$.

2.2 Verossimilhança perfilada modificada de Cox e Reid

O logaritmo da função de verossimilhança perfilada modificada proposta por Cox e Reid (1987) é dada por

$$\ell_{CR}(\mu) = \ell_p(\mu) - \frac{1}{2}\log|j_{\nu\nu}(\mu, \hat{\nu}_\mu)|. \tag{2.7}$$

Existem duas desvantagens em relação a esse ajuste. Uma delas é que ele requer uma parametrização ortogonal e outra, é que ele não é invariante sob reparametrizações.

Quando o parâmetro de interesse é escalar, está assegurada a existência de uma parametrização ortogonal utilizando o procedimento dado por Cox e Reid (1987). A obtenção dessa parametrização está associado à resolução de equações diferenciais que envolvem elementos da matriz de informação de Fisher. Vale salientar que em algumas distribuições é difícil a obtenção de tal matriz, o que ocorre com a maioria das distribuições generalizadas. A obtenção de uma parametrização ortogonal pode ser bastante custosa mesmo nos casos em que a matriz de informação pode ser calculada, pois a equação diferencial resultante pode ser de difícil resolução.

Visando contornar a limitação do ajuste (2.7), Cox e Reid (1993) propuseram uma função de verossimilhança perfilada modificada que não exige ortogonolização dos parâmetros, cujo o logaritmo da função de verossimilhança é dada por

$$\ell_{CR_a}(\mu) = \ell_p(\mu) - \frac{1}{2}\log|j_{\nu\nu}(\mu, \hat{\nu}_\mu)| - (\mu - \hat{\mu})m_\nu(\hat{\mu}, \hat{\nu}), \tag{2.8}$$

em que $m_\nu(\hat{\mu}, \hat{\nu})$ é o traço da matriz $\partial m/\partial\nu$. Sendo que $m = i^{\nu\nu}i_{\mu\nu}$, $i^{\nu\nu}$ denota a inversa da matriz $i_{\nu\nu} = -\mathrm{E}[\partial^2\ell(\mu,\nu)/\partial\nu^2]$ e $i_{\mu\nu} = -\mathrm{E}[\partial^2\ell(\mu,\nu)/\partial\mu\partial\nu]$. Porém, essa aproximação permanece não invariante sob reparametrizações. Denotaremos o estimador de máxima verossimilhança correspondente à (2.8) por $\hat{\mu}_{CR_a}$

2.3 Teste de Hipóteses

As estatísticas da razão de verossimilhanças baseadas nas propostas de Fraser *et al.* (1999), Severini (1999) e Cox e Reid (1993) para testar

$$\mathcal{H}_0 : \mu = \mu_0 \textit{ versus } \mathcal{H}_1 : \mu \neq \mu_0$$

são dadas, respectivamente, por

$$\begin{aligned} LR_{BN_1} &= 2\left\{\tilde{\ell}_{BN}(\widehat{\tilde{\mu}}_{BN}) - \tilde{\ell}_{BN}(\mu_0)\right\}, \\ LR_{BN_2} &= 2\left\{\breve{\ell}_{BN}(\widehat{\breve{\mu}}_{BN}) - \breve{\ell}_{BN}(\mu_0)\right\} \text{ e} \\ LR_{CR_a} &= 2\left\{\ell_{CR_a}(\hat{\mu}_{CR_a}) - \ell_{CR_a}(\mu_0)\right\}. \end{aligned}$$

Sob as condições de regularidade e sob a hipótese nula, as estatísticas definidas acima têm distribuição assintótica χ^2_q.

2.3.1 *Teste da razão de verossimilhanças bootstrap usual*

A técnica bootstrap, descrita em Efron e Tibshirani (1994), é empregada aqui para encontrar a distribuição empírica da estatística LR, sob hipótese nula, diretamente da amostra observada $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$. Essa metodologia consiste em gerar um número grande de B amostras bootstrap $(x_1^{*1}, \ldots, x_n^{*B})$ a partir da amostra original $\boldsymbol{x}$, sob hipótese nula. Para isso, calcula-se o valor da estatística LR para cada amostra, obtendo LR^{*b}, $b = 1, \ldots, B$. Finalmente, esses valores são colocados em ordem crescente.

Fixando o nível de significância (ξ), o percentil $1 - \xi$ de LR^{*b} é estimado pelo valor $\widehat{q}_{1-\xi}$ tal que

$$\frac{\#\{LR^{*b} \leq \widehat{q}_{(1-\xi)}\}}{B} = 1 - \xi,$$

em que $\#$ indica o número de ocorrências do evento $\{LR^{*b} \leq \widehat{q}_{1-\xi}\}$. A quantidade $\widehat{q}_{1-\xi}$ corresponde ao valor de LR^{*b} que ocupa a posição $B \times (1-\xi)$, supondo que $B \times (1-\xi)$ é um número inteiro.

O teste da razão de verossimilhanças bootstrap (LR_b) rejeita a hipótese nula se $LR > \widehat{q}_{(1-\xi)}$.

2.3.2 *Correção de Bartlett bootstrap*

Rocke (1989) introduziu uma forma alternativa de calcular o fator de correção Bartlett para a estatística da razão de verossimilhanças usando o método bootstrap paramétrico (EFRON, B., 1979). A estatística da razão de verossimilhanças corrigida por esse fator é denotada por LR_{bb} e para a sua obtenção, deve-se calcular:

1. LR^{*b}, $b = 1, \ldots, B$, conforme procedimento descrito anteriormente;

2. $\overline{LR}^{*b} = B^{-1} \sum_{b=1}^{B} LR^{*b}$;

3. $LR_{bb} = qLR/\overline{LR}^{*b}$, em que LR é o valor da estatística da razão de verossimilhanças com base na amostra original.

A estatística LR_{bb}, sob $\mathcal{H}_0$, também tem distribuição assintótica χ^2_q. Por requerer um número menor de reamostragens da amostra original, a correção de Bartlett bootstrap é computacionalmente mais eficiente do que a correção bootstrap usual (ROCKE, 1989).

3 Refinamento de inferências na distribuição Lomax exponencializada

Resumo

A função de verossimilhança perfilada ao ser tratada como uma função de verossimilhança genuína pode acarretar em alguns problemas tais como inconsistência e ineficiência dos estimadores de máxima verossimilhança e a aproximação da distribuição nula da estatística da razão de verossimilhanças pela distribuição qui-quadrado pode ser bastante pobre na presença de parâmetros de perturbação. Com o objetivo de fazer inferências para dois parâmetros da distribuição Lomax exponencializada, derivamos a função de verossimilhança perfilada e três de suas versões modificadas. Essas versões modificadas foram obtidas visando atenuar o impacto do parâmetro de perturbação sobre as inferências. Realizamos um estudo de simulação de Monte Carlo para avaliar o desempenho dos estimadores de máxima verossimilhança e dos testes baseados na estatística da razão de verossimilhanças utilizando os diferentes ajustes. Além disso, consideramos o teste da razão de verossimilhanças que utiliza o quantil da sua distribuição empírica nula, obtido através do método bootstrap. Incluímos na comparação o teste baseado na estatística da razão de verossimilhanças Bartlett bootstrap. Para ilustrar a teoria desenvolvida, apresentamos dois exemplos numéricos usando conjuntos de dados reais.

Palavras-chave: Bootstrap paramétrico; Estimação pontual; Lomax exponencializada; Verossimilhança perfilada modificada

3.1 Introdução

A distribuição Lomax, que também é conhecida como distribuição Pareto do tipo II, foi proposta por Lomax (1954) para modelar dados de insucesso empresarial, a partir daí tem sido utilizada para modelar dados em diferentes áreas (HARRIS, 1968; BRYSON, 1974; ARNOLD, 2015; HOLLAND *et al.*, 2006). Devido à grande aplicabilidade da distribuição Lomax, várias generalizações dela foram propostas (ASHOUR; ELTEHIWY, 2013; LEMONTE; CORDEIRO, 2013; RADY *et al.*, 2016). Neste trabalho, será considerada uma generalização de três parâmetros proposta por Abdul-Moniem e Abdel-Hameed (2012), a distribuição Lomax exponencializada (LE). Essa extensão foi obtida utilizando a distribuição Lomax como modelo de

base na família de distribuições exponencializadas (Exp-G). Por ter suporte nos reais positivos e possuir diferentes formas (crescente, decrescente e unimodal) para a função taxa de falha, a distribuição LE é adequada para modelar dados de tempo de vida.

É comum em situações práticas que estejamos interessados em estudar alguns componentes do vetor de parâmetros que indexam um modelo, esses componentes são denominados de parâmetros de interesse e os demais, de parâmetros de perturbação. As inferências sobre os parâmetros de interesse podem ser feitas com base na função de verossimilhança perfilada. Porém, esta função não é uma verossimilhança genuína, sendo assim, os estimadores de máxima verossimilhança podem ser ineficientes e inconsistentes.

Outro inconveniente é que a aproximação da distribuição nula da estatística da razão de verossimilhanças pela distribuição qui-quadrado de referência pode não ser satisfatória em amostras finitas quando existem parâmetros de perturbação ou quando a amostra é pequena. Neste contexto, surgiram diversas propostas para reduzir o impacto dos parâmetros de perturbação nas inferências baseadas na função de verossimilhanças perfilada.

Baseando-se nas propostas de Barndorff-Nielsen (1983), Cox e Reid (1987), Cox e Reid (1993), Fraser e Reid (1995), Fraser *et al.* (1999) e Severini (1999), o nosso principal objetivo é obter três versões modificadas da função de verossimilhança perfilada que forneçam inferências mais confiáveis para um dos parâmetros de forma e o parâmetro de escala da distribuição LE. Estudos nessa linha já foram realizados por Cysneiros *et al.* (2008) e Barreto *et al.* (2013) para a distribuição Birnbaum-Saunders para dados completos e dados censurados, respectivamente. Ferrari *et al.* (2007) obtiveram ajustes da função de verossimilhança perfilada para a distribuição Weibull e recentemente, Pires (2014), para a distribuição Lomax.

No estudo inferencial apresentado nesta tese, consideramos dois cenários: no primeiro, assumimos que o parâmetro de escala que indexa a distribuição LE é conhecido. No segundo, consideramos dois parâmetros (um de forma e o outro de escala) como parâmetros de interesse. Utilizando a função de verossimilhança perfilada e suas versões modificadas, realizamos estimação por máxima verossimilhança e testes de hipóteses com base na estatística da razão de verossimilhanças. Consideramos também no estudo o teste da razão de verossimilhanças bootstrap usual e o teste baseado na estatística da razão de verossimilhanças Bartlett bootstrap.

O capítulo está organizado da seguinte forma: na Seção 3.2, faremos uma discussão sobre a distribuição LE; na Seção 3.3, obtemos os ajustes da verossimilhança perfilada para a distribuição LE. Resultados de simulação de Monte Carlo são apresentados e discutidos na Seção 3.4. Dois exemplos numéricos utilizando dados reais estão na Seção 3.5. As conclusões finais deste trabalho são dadas na Seção 3.6. No Apêndice 6, discutimos alguns aspectos computacionais utilizados neste capítulo.

3.2 Distribuição Lomax exponencializada

Uma variável aleatória Y tem distribuição Lomax (LOMAX, 1954) se sua função densidade de probabilidade (fdp) é

$$f(y) = \beta\lambda(1+\beta y)^{-(\lambda+1)} \tag{3.1}$$

e a sua função de distribuição acumulada (fda) é dada por

$$F(y) = 1-(1+\beta y)^{-\lambda}, \tag{3.2}$$

para $y>0,\ \beta>0$ e $\lambda>0$. Os parâmetros λ e β são de forma e escala, respectivamente.

A distribuição LE, proposta por Abdul-Moniem e Abdel-Hameed (2012), é uma generalização da distribuição Lomax. Essa generalização consiste em introduzir um parâmetro real positivo elevando a fda dada por (3.2) a uma potência $\alpha>0$. Dessa forma, se X é uma variável aleatória com distribuição LE, então a sua fda é definida como

$$F(x) = [1-(1+\beta x)^{-\lambda}]^{\alpha}. \tag{3.3}$$

Derivando (3.3) em relação a x, obtém-se a fdp de X, a qual é dada por

$$f(x) = \alpha\beta\lambda[1-(1+\beta x)^{-\lambda}]^{\alpha-1}(1+\beta x)^{-(\lambda+1)}, \tag{3.4}$$

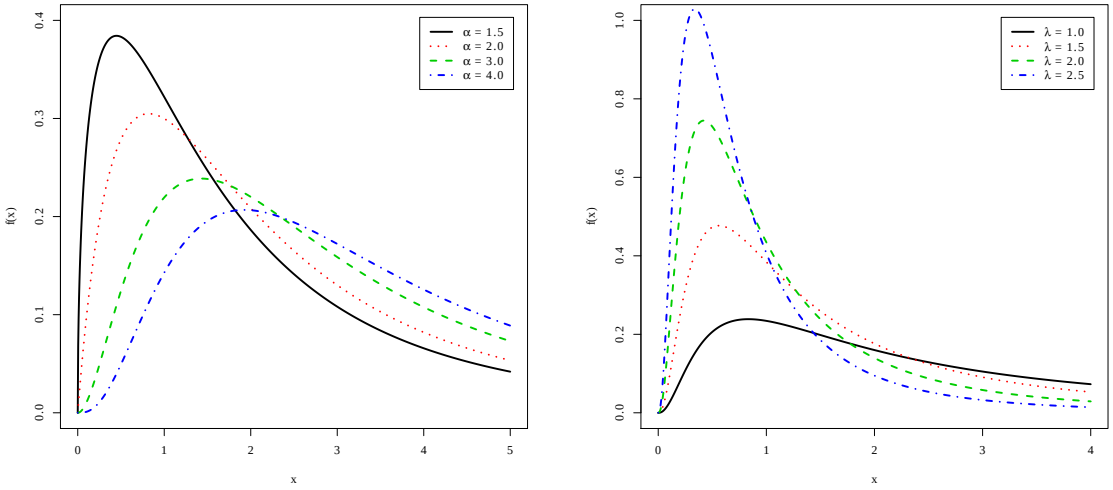
em que $x>0,\ \alpha>0,\ \beta>0$ e $\lambda>0$. Sendo α e λ parâmetros de forma, e β parâmetro de escala. A partir daqui, denotaremos uma variável aleatória X com fdp (3.4) por $X\sim \mathrm{LE}(\alpha,\beta,\lambda)$.

Podemos obter a fdp das distribuições Pareto exponencializada, Pareto e Lomax, colocando em (3.4) os valores $\beta=1$, $\beta=\alpha=1$ e $\alpha=1$, respectivamente.

Na Figura 3.1 são mostrados alguns gráficos da fdp da distribuição LE para diferentes valores de parâmetros. Avaliando esses gráficos, podemos notar que os parâmetros α e λ tem um controle sobre as caudas. Da Figura 3.1(a), concluímos que a cauda direita se torna mais pesada à medida que o parâmetro α cresce, enquanto que na Figura 3.1(b), conforme λ cresce a cauda direita se torna mais leve.

Se $X\sim \mathrm{LE}(\alpha,\beta,\lambda)$, então o r-ésimo momento de X é

$$\mathrm{E}(X^r) = \frac{\alpha}{\lambda^r}\sum_{i=0}^{r}\binom{r}{i}(-1)^i B\left(1-\frac{1}{\lambda}(r-i),\alpha\right),\ r=1,2,3,\ldots$$



(a) $\beta = 0,25$ e $\lambda = 3,0$ (b) $\beta = 1,5$ e $\alpha = 3,5$

Figura 3.1: Gráficos para a densidade LE para diferentes valores de parâmetros

e a variância é dada por

$$\mathrm{Var}(X) = \left(\frac{\alpha}{\lambda}\right)^2 \left[\frac{1}{\alpha} B\left(1 - \frac{2}{\lambda}, \alpha\right) - B^2\left(1 - \frac{1}{\lambda}, \alpha\right)\right],$$

em que $B(\cdot, \cdot)$ é a função beta (GRADSHTEYN; RYZHIK, 2007).

Seja $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ uma amostra aleatória de tamanho n proveniente da distribuição LE cuja fdp é dada em (3.4). O logaritmo da função de verossimilhança para $\boldsymbol{\theta} = (\alpha, \beta, \lambda)^\top$ é dado por

$$\begin{aligned} \ell(\boldsymbol{\theta}; \boldsymbol{x}) &= n[\log(\alpha) + \log(\beta) + \log(\lambda)] + (\alpha - 1) \sum_{i=1}^{n} \log(1 - u_i^{-\lambda}) \\ &- (\lambda + 1) \sum_{i=1}^{n} \log(u_i), \end{aligned} \tag{3.5}$$

em que $\log(a) = \log_{\mathrm{e}} a = \ln(a)$ e $u_i = 1 + \beta x_i$. O vetor escore $\boldsymbol{\ell_\theta} = (\ell_\alpha, \ell_\beta, \ell_\lambda)^\top$ é obtido pela diferenciação da função (3.5) em relação aos parâmetros desconhecidos. Dessa forma, seus componentes são dados, respectivamente, por

$$\ell_\alpha = \frac{n}{\alpha} + \sum_{i=1}^{n} \log(1 - u_i^{-\lambda}),$$

$$\ell_\beta = \frac{n}{\beta} - (1+\lambda)\sum_{i=1}^{n}\frac{x_i}{u_i} + \lambda(\alpha-1)\sum_{i=1}^{n}\frac{x_i u_i^{-(\lambda+1)}}{1-u_i^{-\lambda}}$$

e

$$\ell_\lambda = \frac{n}{\lambda} - \sum_{i=1}^{n}\log(u_i) + (\alpha-1)\sum_{i=1}^{n}\frac{\log(u_i)(u_i)^{-\lambda}}{1-(u_i)^{-\lambda}}.$$

O estimador de máxima verossimilhança de $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$, pode ser obtido resolvendo numericamente a equação não-linear $\boldsymbol{\ell_\theta} = \mathbf{0}$. Alternativamente, $\hat{\boldsymbol{\theta}}$ pode ser obtido maximizando diretamente (3.5) utilizando o método BFGS que está implementado na plataforma computacional `Ox` através do comando `MaxBFGS`.

A matriz de informação observada é dada por

$$J(\boldsymbol{\theta}) = \begin{bmatrix} -\frac{\partial^2\ell(\theta)}{\partial\alpha^2} & -\frac{\partial^2\ell(\theta)}{\partial\alpha\partial\beta} & -\frac{\partial^2\ell(\theta)}{\partial\alpha\partial\lambda} \\ -\frac{\partial^2\ell(\theta)}{\partial\beta\partial\alpha} & -\frac{\partial^2\ell(\theta)}{\partial\beta^2} & -\frac{\partial^2\ell(\theta)}{\partial\beta\partial\lambda} \\ -\frac{\partial^2\ell(\theta)}{\partial\lambda\partial\alpha} & -\frac{\partial^2\ell(\theta)}{\partial\lambda\partial\beta} & -\frac{\partial^2\ell(\theta)}{\partial\lambda^2} \end{bmatrix},$$

em que as quantidades $\partial^2\ell(\boldsymbol{\theta})/\partial r\partial s$ para $r, s = \alpha, \beta, \lambda$ são dados no Apêndice 6. Mais adiante, elementos da matriz de informação observada serão úteis para obter os diferentes ajustes para a função de verossimilhança perfilada.

3.3 Verossimilhança perfilada ajustada para a distribuição LE

Seja $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ uma amostra aleatória de tamanho n proveniente da distribuição LE(α, β, λ) cuja fdp é dada por (3.4). Com base nas funções definidas nas Seções 2.1 e 2.2, obtemos os ajustes para a função de verossimilhanças perfilada e suas versões modificadas para a distribuição LE considerando dois cenários. No primeiro cenário supomos que o parâmetro de escala β é conhecido e α é parâmetro de interesse. Enquanto que, no segundo cenário supomos que os parâmetros α e β são de interesse. Em todas as situações, λ é parâmetro de perturbação.

3.3.1 *Supondo o parâmetro β conhecido*

Inicialmente, vamos supor que o parâmetro β é conhecido, α é parâmetro de interesse e λ, de perturbação, isto é, $\mu = \alpha$ e $\nu = \lambda$. Substituindo λ por seu estimador restrito, $\hat{\lambda}_\alpha$, na

equação (3.5), obtemos o logaritmo da função de verossimilhança perfilada para o parâmetro α que é dado por

$$\begin{aligned}\ell_p(\alpha) &= n[\log(\alpha)+\log(\hat{\lambda}_\alpha)]+(\alpha-1)\sum_{i=1}^{n}\log[1-u_i^{-\hat{\lambda}_\alpha}] \\ &- (\hat{\lambda}_\alpha+1)\sum_{i=1}^{n}\log(u_i). \end{aligned} \tag{3.6}$$

$\hat{\lambda}_\alpha$ é obtido como solução da equação $\ell_\lambda = 0$ para α fixo. Como $\hat{\lambda}_\alpha$ não tem expressão em forma fechada, deve ser encontrado por meio de métodos de otimização não-linear restrita. Para maiores detalhes sobre esses métodos ver Nocedal e Wright (2006). O estimador de máxima verossimilhança perfilado de α, que maximiza $\ell_p(\alpha)$, é denotado por $\hat{\alpha}_p$.

Derivamos duas aproximações para o logaritmo da função de verossimilhança perfilada modificada de Barndorff-Nielsen (1983) , $\tilde{\ell}_{BN}(\alpha)$ e $\breve{\ell}_{BN}(\alpha)$, as quais são dadas por

$$\tilde{\ell}_{BN}(\alpha) = \ell_p(\alpha)+\frac{1}{2}\log|j_{\lambda\lambda}(\alpha,\hat{\lambda}_\alpha)| - \log|\ell_{\lambda;X}(\alpha,\hat{\lambda}_\alpha)\widehat{V}_\lambda|, \tag{3.7}$$

$$\breve{\ell}_{BN}(\alpha) = \ell_p(\alpha)+\frac{1}{2}\log|j_{\lambda\lambda}(\alpha,\hat{\lambda}_\alpha)| - \log|\breve{I}_\lambda(\alpha,\hat{\lambda}_\alpha;\hat{\alpha},\hat{\lambda})|, \tag{3.8}$$

tal que

$$\ell_{\lambda;X}(\alpha,\hat{\lambda}_\alpha)\widehat{V}_\lambda = \sum_{i=1}^{n}\frac{\log(u_i)[(\alpha-1)\hat{\lambda}_\alpha\, u_i^{\hat{\lambda}_\alpha}\log(u_i)+(u_i^{\hat{\lambda}_\alpha}-1)(u_i^{\hat{\lambda}_\alpha}-\alpha)]}{\hat{\lambda}(u_i^{\hat{\lambda}_\alpha}-1)^2},$$

$$j_{\lambda\lambda}(\alpha,\hat{\lambda}_\alpha) = \frac{n}{\hat{\lambda}_\alpha^2}+(\alpha-1)\sum_{i=1}^{n}\left[\frac{\log^2(u_i)\,u_i^{-2\hat{\lambda}_\alpha}}{(1-u_i^{-\hat{\lambda}_\alpha})^2}\right]+(\alpha-1)\sum_{i=1}^{n}\left[\frac{\log^2(u_i)\,u_i^{-\hat{\lambda}_\alpha}}{(1-u_i^{-\hat{\lambda}_\alpha})}\right],$$

$$\breve{I}_\lambda(\alpha,\hat{\lambda}_\alpha;\hat{\alpha},\hat{\lambda}) = \sum_{i=1}^{n}\left[\frac{u_i^{\hat{\lambda}_\alpha}-1+\hat{\lambda}_\alpha\log(u_i)(\alpha-u_i^{\hat{\lambda}_\alpha})}{\hat{\lambda}_\alpha\,(u_i^{\hat{\lambda}_\alpha}-1)}\right]\left[\frac{u_i^{\hat{\lambda}}-1+\hat{\lambda}\log(u_i)(\hat{\alpha}-u_i^{\hat{\lambda}})}{\hat{\lambda}\,(u_i^{\hat{\lambda}}-1)}\right],$$

em que $\hat{\alpha}$ e $\hat{\lambda}$ são, respectivamente, os estimadores de máxima verossimilhança de α e λ.

Conforme foi mencionado, o ajuste (2.7) requer uma parametrização ortogonal dos parâmetros que indexam o modelo. Porém, os parâmetros da distribuição LE não satisfazem essa condição. Além disso, a obtenção de uma parametrização ortogonal pelo procedimento des-

crito em Cox e Reid (1987) é bastante complicada. Sendo assim, aplicamos para a distribuição LE o ajuste dado pela equação (2.8), que não requer ortogonalidade explícita dos parâmetros, obtendo

$$\ell_{CR_a}(\alpha) = \ell_p(\alpha) - \frac{1}{2}\log|j_{\lambda\lambda}(\alpha, \hat{\lambda}_\alpha)| - (\alpha - \hat{\alpha})m_\lambda(\hat{\alpha}, \hat{\lambda}), \tag{3.9}$$

em que

$$m_\lambda(\hat{\alpha}, \hat{\lambda}) = \frac{(1 - \gamma - \Psi(\hat{\alpha}+1))}{(\hat{\alpha}-1)}\left\{1 + \frac{\hat{\alpha}[\pi^2 + 6(\gamma-2)\gamma + 6\Psi(\hat{\alpha})(2\gamma + \Psi(\hat{\alpha}) - 2) - 6\Psi^{(1)}(\hat{\alpha})]}{6(\hat{\alpha}-2)}\right\}^{-1},$$

tal que γ é a constante de Euler, $\Psi(\cdot)$ é a função digama e $\Psi^{(1)}(\cdot)$ é a primeira derivada da função digama. É importante notar que m_λ, neste caso, não depende de $\hat{\lambda}$ e está definido apenas para $\hat{\alpha} \neq 2$ e $\hat{\alpha} \neq 1$, portanto a função (3.9) somente pode ser ajustada sob essa condição.

Os estimadores de máxima verossimilhança perfilados ajustados de α obtidos de (3.7), (3.8), e (3.9) são denotados, respectivamente, por $\hat{\tilde{\alpha}}_{BN}$, $\hat{\tilde{\alpha}}_{BN}$ e $\hat{\alpha}_{CR_a}$. Esses estimadores e o estimador de máxima verossimilhança perfilado, $\hat{\alpha}_p$, não têm forma fechada e devem ser obtidos numericamente a partir da maximização do logaritmo das respectivas funções de verossimilhança.

3.3.2 *Supondo o parâmetro β desconhecido*

Supomos agora que α e β são parâmetros de interesse, ou seja, $\boldsymbol{\mu} = (\alpha, \beta)^\top$. Neste caso, temos que o logaritmo da função de verossimilhança perfilada para o parâmetro $\boldsymbol{\mu}$ é expresso como

$$\begin{aligned}\ell_p(\boldsymbol{\mu}) &= n[\log(\alpha) + \log(\beta) + \log(\hat{\lambda}_{\boldsymbol{\mu}})] + (\alpha - 1)\sum_{i=1}^{n}\log[1 - u_i^{-\hat{\lambda}_{\boldsymbol{\mu}}}] \\ &- (\hat{\lambda}_{\boldsymbol{\mu}} + 1)\sum_{i=1}^{n}\log(u_i),\end{aligned} \tag{3.10}$$

em que $\hat{\lambda}_{\boldsymbol{\mu}} = \hat{\lambda}_{\alpha,\beta}$ é obtido como solução da equação $\ell_\lambda = 0$ para α e β fixos. Aqui, $\hat{\lambda}_{\boldsymbol{\mu}}$ também não possui forma fechada, logo deverá ser encontrado usando um método de maximização restrita.

Como na subseção anterior, obtemos aqui as duas aproximações para o ajuste da função de verossimilhança perfilada modificada de Barndorff-Nielsen (1983) as quais são dadas por

$$\tilde{\ell}_{BN}(\boldsymbol{\mu}) = \ell_p(\alpha) + \frac{1}{2}\log|j_{\lambda\lambda}(\boldsymbol{\mu}, \hat{\lambda}_{\boldsymbol{\mu}})| - \log|\ell_{\lambda;X}(\boldsymbol{\mu}, \hat{\lambda}_{\boldsymbol{\mu}})\widehat{V}_\lambda|, \tag{3.11}$$

$$\breve{\ell}_{BN}(\boldsymbol{\mu}) = \ell_p(\alpha) + \frac{1}{2}\log|j_{\lambda\lambda}(\boldsymbol{\mu}, \hat{\lambda}_{\boldsymbol{\mu}})| - \log|\breve{I}_{\lambda}(\boldsymbol{\mu}, \hat{\lambda}_{\boldsymbol{\mu}}; \hat{\boldsymbol{\mu}}, \hat{\lambda})|, \tag{3.12}$$

em que

$$j_{\lambda\lambda}(\boldsymbol{\mu}, \hat{\lambda}_{\boldsymbol{\mu}}) = \frac{n}{\hat{\lambda}^2_{\boldsymbol{\mu}}} + (\alpha-1)\sum_{i=1}^{n}\frac{\log^2(u_i)(u_i^{-2\hat{\lambda}_{\boldsymbol{\mu}}})}{(1-u_i^{-\hat{\lambda}_{\boldsymbol{\mu}}})^2} + (\alpha-1)\sum_{i=1}^{n}\frac{\log^2(u_i)(u_i^{-\hat{\lambda}_{\boldsymbol{\mu}}})}{(1-u_i^{-\hat{\lambda}_{\boldsymbol{\mu}}})},$$

$$\ell_{\lambda;X}(\boldsymbol{\mu}, \hat{\lambda}_{\boldsymbol{\mu}})\widehat{V}_{\lambda} = \frac{\beta}{\hat{\beta}\hat{\lambda}}\sum_{i=1}^{n}\frac{\hat{u}_i\log(\hat{u}_i)(u_i^{\hat{\lambda}_{\boldsymbol{\mu}}}-1)(u_i^{\hat{\lambda}_{\boldsymbol{\mu}}}-\alpha) + (\alpha-1)\hat{\lambda}_{\boldsymbol{\mu}}u_i^{\hat{\lambda}_{\boldsymbol{\mu}}}\log(u_i)}{u_i(u_i^{\hat{\lambda}_{\boldsymbol{\mu}}}-1)},$$

$$\breve{I}_{\lambda}(\boldsymbol{\mu}, \hat{\lambda}_{\boldsymbol{\mu}}; \hat{\boldsymbol{\mu}}, \hat{\lambda}) = \sum_{i=1}^{n}\left[\frac{u_i^{\hat{\lambda}_{\boldsymbol{\mu}}}-1+\hat{\lambda}_{\boldsymbol{\mu}}\log(u_i)(\alpha-u_i^{\hat{\lambda}_{\boldsymbol{\mu}}})}{\hat{\lambda}_{\boldsymbol{\mu}}\,(u_i^{\hat{\lambda}_{\boldsymbol{\mu}}}-1)}\right]\left[\frac{\hat{u}_i^{\hat{\lambda}}-1+\hat{\lambda}\log(\hat{u}_i)(\hat{\alpha}-\hat{u}_i^{\hat{\lambda}})}{\hat{\lambda}\,(\hat{u}_i^{\hat{\lambda}}-1)}\right].$$

Sendo $\hat{u}_i = 1 + \hat{\beta}x_i$. Os estimadores de máxima verossimilhança perfilados ajustados de $\boldsymbol{\mu}$ obtidos de (3.10), (3.11) e (3.12) são denotados, respectivamente, por $\hat{\boldsymbol{\mu}}_p$, $\hat{\tilde{\boldsymbol{\mu}}}_{BN}$ e $\hat{\breve{\boldsymbol{\mu}}}_{BN}$. Esses estimadores não têm forma fechada e devem ser obtidos maximizando numericamente as respectivas funções de verossimilhança.

3.4 Estudo de simulação

Nesta seção, realizamos um estudo de simulação de Monte Carlo para avaliar e comparar os desempenhos dos estimadores de máxima verossimilhança e da estatística da razão de verossimilhanças, baseados nas funções de verossimilhanças apresentadas na Seção 3.3. Incluímos também na análise o teste baseado na estatística LR_{bb} e o teste da razão de verossimilhanças bootstrap usual (LR_b). Os resultados estão organizados conforme os dois cenários discutidos na Seção 3.3.

Os resultados são baseados em $10\,000$ réplicas de Monte Carlo. Para os testes baseados em bootstrap paramétrico, verificamos que 600 réplicas bootstrap foram suficientes para alcançar a estabilidade dos resultados, assim, assumimos $B = 600$. As simulações foram realizadas utilizando a linguagem de programação `Ox` (DOORNIK, 2009). O estimador restrito do parâmetro λ, $\hat{\lambda}_{\mu}$, foi estimado utilizando o algoritmo de otimização não-linear restrito SQP (*Sequential Quadratic Programming*) (NOCEDAL; WRIGHT, 2006). Esse algoritmo está implementando na plataforma `Ox` através da função `MaxSQP`.

3.4.1 *Cenário: supondo o parâmetro β conhecido*

Neste cenário, apresentamos os resultados numéricos para o parâmetro α da distribuição LE. Os resultados foram obtidos definindo $\beta = 1,0$. Os tamanhos amostrais considerados neste caso foram $n = 20,\ 30,\ 40,\ 50$ e os níveis nominais foram $\xi = 10\%$, 5% e 1%.

Para os resultados numéricos referentes à estimação pontual do parâmetro de forma α da distribuição LE, consideramos os estimadores de máxima verossimilhança $\hat{\alpha}_p$, $\hat{\tilde{\alpha}}_{BN}$, $\hat{\mathring{\alpha}}_{BN}$ e $\hat{\alpha}_{CR_a}$, os quais são obtidos maximizando as funções dadas por (3.6), (3.7), (3.8) e (3.9), respectivamente. A análise desses estimadores é baseada na medida do viés relativo (VR), do erro quadrático médio (EQM), assimetria e curtose.

Estamos interessados também em testar a hipótese nula $\mathcal{H}_0 : \alpha = \alpha_0$ *versus* $\mathcal{H}_1 : \alpha \neq \alpha_0$. Avaliamos os desempenhos dos testes baseados nas estatísticas LR_p, LR_{BN_1}, LR_{BN_2}, LR_{CR_a}, LR_{bb} e do teste LR_b. Para cada tamanho de amostra e cada nível nominal, calculamos as taxas de rejeição dos testes baseados nas estatísticas citadas, ou seja, são estimadas via simulação de Monte Carlo as seguintes probabilidades: $P(LR_p \geq \chi^2_{(1-\xi;1)})$, $P(LR_{BN_1} \geq \chi^2_{(1-\xi;1)})$, $P(LR_{BN_2} \geq \chi^2_{(1-\xi;1)})$, $P(LR_{CR_a} \geq \chi^2_{(1-\xi;1)})$, $P(LR_{bb} \geq \chi^2_{(1-\xi;1)})$, tal que $\chi^2_{(1-\xi;1)}$ é o quantil $(1-\xi)$ da distribuição χ^2_1. Para o teste LR_b a taxa de rejeição é obtida estimando a probabilidade $P(LR \geq \hat{q}_{(1-\xi)})$, em que $\hat{q}_{(1-\xi)}$ é estimado seguindo o procedimento descrito na Subseção 2.3.1.

Para a obtenção dos resultados apresentados nas Tabelas 3.1, 3.2 e 3.3, assumimos o valor verdadeiro de $\lambda = 5,0$ e os valores do parâmetro de interesse α variando em: $3,5$; $4,0$ e $4,5$. Podemos verificar na Tabela 3.1 que os estimadores perfilados modificados apresentam vieses relativos ligeiramente menores em relação ao viés relativo do estimador $\hat{\alpha}_p$, independente do tamanho amostral. Por exemplo, para $n = 20$ e $\alpha = 3,5$, o viés relativo de $\hat{\alpha}_p$ é aproximadamente de 29%, enquanto que os vieses relativos dos estimadores perfilados modificados não ultrapassam 22%.

Além disso, podemos notar que as medidas do VR, EQM, assimetria e curtose diminuem conforme o tamanho da amostra aumenta. Outro critério para a avaliação dos estimadores é o EQM, quanto menor essa medida, melhor é o estimador. Baseando-se nos valores do EQM de cada estimador apresentados na Tabela 3.1, concluímos que os estimadores $\hat{\tilde{\alpha}}_{BN}$ e $\hat{\mathring{\alpha}}_{BN}$ tiveram comportamentos similares e menores valores de EQM, logo são os melhores segundo esse critério. Considerando o mesmo critério, o estimador $\hat{\alpha}_p$ teve o pior desempenho.

A Tabela 3.2 é referente às taxas de rejeições nulas correspondentes aos testes baseados nas estatísticas LR_p, LR_{BN_1}, LR_{BN_2}, LR_{CR_a}, LR_{bb} e ao teste LR_b. Avaliando essa tabela, observamos que em todas as situações consideradas, o teste baseado na estatística LR_p é liberal, ou seja, as taxas de rejeições estimadas estão acima do nível nominal especificado. Por exemplo, para $\alpha = 3,5$, $n = 20$, ao nível $\xi = 10\%$, a taxa de rejeição do teste baseado em LR_p é $12,07$. Notamos ainda, que os testes baseados na estatística LR_{CR_a} apresentaram resultados

semelhantes aos testes baseados em LR_p.

Ainda analisando a Tabela 3.2, percebemos que os testes baseados nas estatísticas LR_{BN_1} e LR_{BN_2} tiveram comportamento similares. Em algumas situações se mostraram liberais e em outras, foram conservativos. No entanto, em geral, apresentaram os melhores desempenhos quando comparados aos testes baseados nas estatísticas LR_p e LR_{CR_a}, principalmente para amostras de tamanho pequeno a moderado. Por exemplo, para $\alpha = 4,5$ e $n = 20$, ao nível de $\xi = 5\%$, a taxa de rejeição dos testes baseados em LR_p e LR_{CR_a} foram $6,37$ e $6,42$, respectivamente, enquanto que as taxas de rejeição baseadas em LR_{BN_1} e LR_{BN_2} foram $5,25$ e $5,32$, respectivamente.

Concluímos também que os testes baseados na estatística LR_{bb} e o teste LR_b fornecem taxas bem próximas aos níveis nominais, porém essas taxas, em geral, não apresentam o padrão de se aproximar do nível nominal especificado conforme o tamanho da amostra aumenta. É importante frisar que em todos os cenários considerados, o teste baseado na estatística LR_{bb} apresentou resultados ligeiramente melhores que o teste baseado em LR_b. Isso significa que o fator de correção Bartlett bootstrap incorporado à estatística da razão de verossimilhanças atenuou significativamente a sua tendência liberal.

A Figura 3.2 mostra a distorção dos testes, ou seja, a diferença entre a taxa de rejeição estimada e o nível nominal correspondente. Para a construção desse gráfico, consideramos o caso em que $\alpha = 4,5$ e $\xi = 10\%$. Essa figura, evidencia toda análise feita anteriormente sobre as estatísticas de testes e mostra claramente que as distorções do testes baseado na estatística LR_{bb} e do teste LR_b são as que mais se aproximam da linha de referência (ordenada em zero), logo esses testes tiveram o melhor desempenho. Além disso, nota-se que as curvas de distorção dos testes baseados em LR_p, LR_{BN_1} e LR_{bb} apresentam o mesmo padrão das curvas de distorção dos testes baseados em LR_{BN_2}, LR_{CR_a} e LR_b, respectivamente.

A Figura 3.3 apresenta o gráfico das discrepâncias relativas de quantis (a diferença entre quantil exato e assintótico dividida pelo quantil assintótico) das quatro estatísticas de teste em relação ao quantil assintótico correspondente. Quanto mais próxima da ordenada zero a curva estiver, melhor é a aproximação assintótica utilizada no teste. Sendo assim, as estatísticas de teste LR_{BN_1}, LR_{BN_2} e LR_{bb} apresentam uma melhor aproximação da distribuição nula assintótica do que as estatísticas LR_p e LR_{CR_a}. Nota-se também que as curvas neste gráfico apresentam o mesmo padrão. No entanto, para quantis maiores, a distribuição nula das estatísticas LR_{BN_1} e LR_{BN_2} se aproximam mais da distribuição de referência quando comparada com a estatística LR_{bb}. É importante observar que o teste LR_b não foi considerado nesta comparação porque as taxas de rejeição desse teste foram estimadas utilizando a distribuição empírica da estatística da razão de verossimilhanças.

A Tabela 3.3 apresenta a média e a variância das estatísticas de teste para $\alpha = 3,5$ e $\alpha = 4,5$

Tabela 3.1: Estimação pontual de α (λ = 5,0)

		$\alpha = 3,5$				$\alpha = 4,0$				$\alpha = 4,5$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
20	$\widehat{\alpha}_p$	0,289	7,367	3,998	39,397	0,308	11,026	4,347	46,215	0,326	15,799	4,693	53,446
	$\widehat{\tilde{\alpha}}_{BN}$	0,212	5,834	3,826	36,490	0,226	8,655	4,147	42,601	0,239	12,301	4,464	49,063
	$\widehat{\breve{\alpha}}_{BN}$	0,211	5,861	3,846	36,355	0,225	8,711	4,165	42,267	0,239	12,400	4,479	48,494
	$\widehat{\alpha}_{CR_a}$	0,218	5,971	3,854	36,952	0,233	8,873	4,181	43,196	0,247	12,632	4,505	49,816
30	$\widehat{\alpha}_p$	0,170	2,729	2,074	11,837	0,180	3,936	2,176	12,735	0,189	5,447	2,271	13,612
	$\widehat{\tilde{\alpha}}_{BN}$	0,126	2,329	2,037	11,508	0,133	3,341	2,134	12,352	0,140	4,603	2,225	13,175
	$\widehat{\breve{\alpha}}_{BN}$	0,125	2,326	2,041	11,549	0,132	3,340	2,139	12,402	0,139	4,603	2,231	13,235
	$\widehat{\alpha}_{CR_a}$	0,128	2,351	2,041	11,546	0,136	3,377	2,139	12,397	0,143	4,654	2,231	13,228
40	$\widehat{\alpha}_p$	0,121	1,617	1,481	6,890	0,128	2,307	1,541	7,245	0,134	3,160	1,597	7,590
	$\widehat{\tilde{\alpha}}_{BN}$	0,090	1,430	1,464	6,796	0,095	2,032	1,523	7,136	0,100	2,774	1,578	7,466
	$\widehat{\breve{\alpha}}_{BN}$	0,089	1,429	1,468	6,821	0,094	2,032	1,527	7,168	0,099	2,774	1,582	7,504
	$\widehat{\alpha}_{CR_a}$	0,091	1,438	1,466	6,805	0,096	2,044	1,525	7,147	0,128	3,114	1,593	7,567
50	$\widehat{\alpha}_p$	0,093	1,146	1,355	6,402	0,073	1,466	1,392	6,585	0,103	2,214	1,449	6,873
	$\widehat{\tilde{\alpha}}_{BN}$	0,069	1,037	1,344	6,349	0,085	1,514	1,394	6,594	0,077	1,992	1,436	6,808
	$\widehat{\breve{\alpha}}_{BN}$	0,069	1,037	1,345	6,353	0,073	1,466	1,394	6,590	0,076	1,992	1,438	6,814
	$\widehat{\alpha}_{CR_a}$	0,070	1,041	1,344	6,353	0,074	1,472	1,393	6,589	0,101	2,791	1,579	7,478

Tabela 3.2: Taxas de rejeições nulas, inferências sobre α (λ = 5,0)

		$n = 20$						$n = 30$					
α	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}
3,5	10	12,07	10,61	10,59	12,02	10,05	10,03	10,83	9,80	9,79	10,78	10,19	10,13
	5	6,31	5,28	5,36	6,39	5,21	4,96	5,57	4,70	4,77	5,43	5,14	5,02
	1	1,39	1,00	0,97	1,36	1,06	0,92	1,24	1,03	1,02	1,20	1,07	0,90
4,0	10	12,07	10,57	10,49	12,10	10,10	10,07	10,86	9,83	9,85	10,84	10,08	10,01
	5	6,39	5,24	5,27	6,39	5,17	4,91	5,58	5,58	4,71	5,40	5,09	4,97
	1	1,39	0,97	0,95	1,38	1,05	0,90	1,22	1,02	1,00	1,22	1,09	0,91
4,5	10	12,20	10,44	10,48	12,15	10,17	10,13	10,83	9,81	9,88	10,84	10,26	10,08
	5	6,37	5,25	5,32	6,42	5,14	4,92	5,57	4,69	4,64	5,37	5,06	4,92
	1	1,39	0,93	0,94	1,37	1,06	0,89	1,23	1,01	1,02	1,23	1,08	0,92
		$n = 40$						$n = 50$					
α	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}
3,5	10	10,32	9,56	9,58	10,41	10,18	10,06	10,30	9,88	9,89	10,38	10,30	10,19
	5	5,08	4,50	4,49	4,98	4,96	4,83	5,16	4,77	4,77	5,18	5,090	4,94
	1	1,15	0,96	0,93	1,07	1,05	0,94	1,12	0,98	0,97	1,08	0,94	0,85
4,0	10	10,39	9,56	9,61	10,34	10,10	10,04	10,30	9,79	9,80	10,41	10,30	10,20
	5	5,00	4,49	4,50	4,96	4,95	4,79	5,13	4,72	4,73	5,17	5,040	4,95
	1	1,16	1,21	0,95	1,07	1,01	0,90	1,11	1,00	0,98	1,07	0,90	0,85
4,5	10	10,35	9,62	9,65	10,37	10,09	9,99	10,32	9,74	9,73	10,38	10,28	10,15
	5	5,02	4,50	4,53	5,03	4,93	4,78	5,18	4,74	4,76	5,19	5,07	4,89
	1	1,15	0,95	0,94	1,08	1,00	0,89	1,11	0,97	0,96	1,06	0,91	0,83

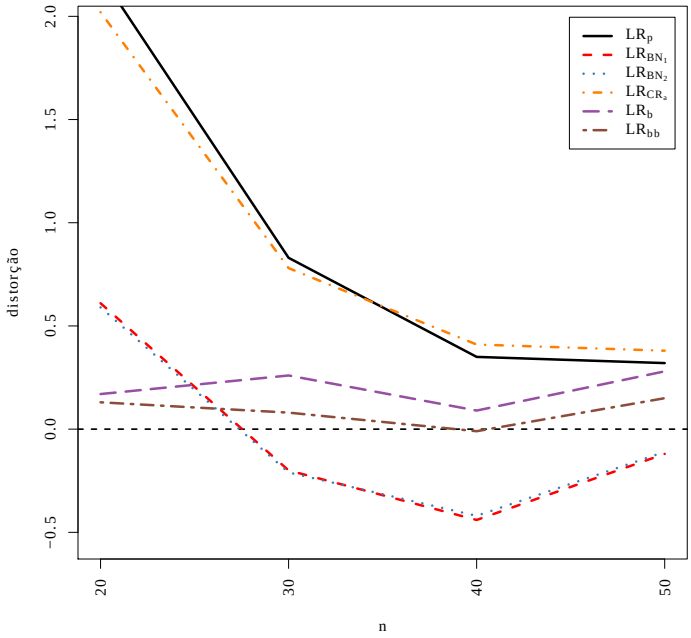


Figura 3.2: Distorções dos tamanhos dos testes: $\xi = 10\%$ e $\alpha = 4,5$

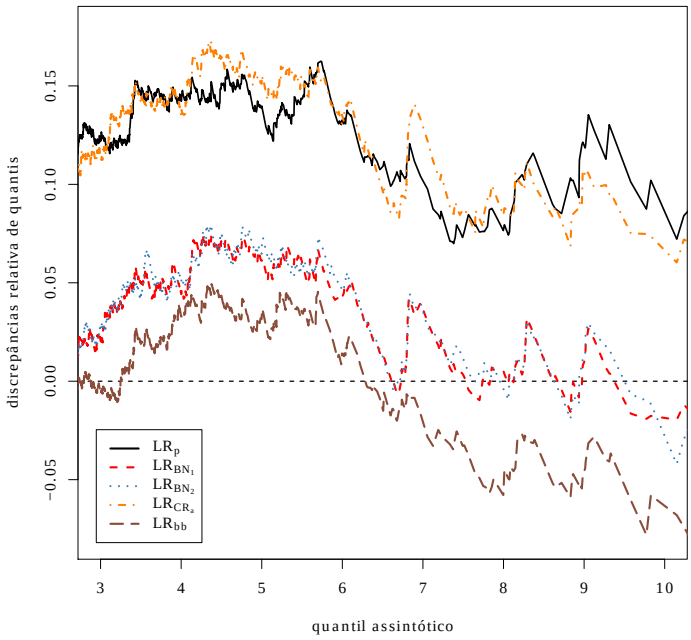


Figura 3.3: Discrepâncias relativas de quantis: $n = 20$ e $\alpha = 4,5$

Tabela 3.3: Média e variância (var.) das estatísticas de teste, β conhecido

			$\alpha = 3,5$					$\alpha = 4,5$				
n		χ^2_1	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_{bb}
20	média	1,000	1,101	1,020	1,021	1,107	1,005	1,103	1,021	1,021	1,112	1,006
	var.	2,000	2,417	2,061	2,065	2,328	1,893	2,422	2,057	2,064	2,437	1,895
30	média	1,000	1,051	0,997	0,997	1,053	1,001	1,051	0,996	0,995	1,054	0,999
	var.	2,000	2,166	1,952	1,951	2,174	1,939	2,164	1,945	1,943	2,176	1,941
40	média	1,000	1,028	0,988	0,988	1,029	0,999	1,027	0,986	0,986	1,029	1,000
	var.	2,000	2,014	1,848	1,847	2,004	1,927	2,015	1,844	1,843	2,008	1,926
50	média	1,000	1,017	0,986	0,986	1,018	1,006	1,017	0,984	0,984	1,019	1,004
	var.	2,000	2,083	1,953	1,953	2,084	1,940	2,083	1,949	1,949	2,086	1,933

e diferentes tamanhos amostrais. Avaliando esses resultados, constatamos que a média e a variância amostral das estatísticas corrigidas LR_{BN_1}, LR_{BN_2} e LR_{bb}, em geral, são as que mais se aproximam da média e da variânca da distribuição assintótica χ^2_1 sob hipótese nula. Claramente, a estatística LR_{bb} têm a melhor aproximação.

A Tabela 3.4 apresenta resultados da simulação de poder dos testes, obtidos considerando a hipótese alternativa, para $n = 30$ e diferentes valores de α, ao nível nominal $\xi = 10\%$. Observar que essas simulações correspondem à situação apresentada na Tabela 3.2 para $n = 30$ e $\alpha = 3, 5$. Como os tamanhos dos testes baseados nas estatísticas LR_p, LR_{BN_1}, LR_{BN_2}, LR_{CR_a} e LR_{bb} apresentam diferenças entre as taxas de rejeição para um nível nominal fixado, os testes foram realizados utilizando os valores críticos estimados para cada estatística, dessa forma colocamos todos os testes em um mesmo patamar. Aqui, não avaliamos o teste LR_b porque não é possível corrigi-lo em relação ao seu tamanho. Avaliando os resultados, observamos que os testes baseados nas estatísticas LR_{BN_1}, LR_{BN_2} e LR_{CR_a} sao mais poderosos que os testes baseados em LR_p e LR_{bb}. Ainda neste cenário, fizemos o mesmo estudo de simulação, considerando agora o valor verdadeiro de $\lambda = 0, 5$ e o parâmetro de interesse α variando em: $0, 5$, $1, 0$ e $1, 5$. Os resultados são mostrados na Tabela 3.5 e na Tabela 3.6. Como podemos observar, as conclusões para essas tabelas são análogas às conclusões feitas para a Tabela 3.1 e a Tabela 3.2, respectivamente.

3.4.2 *Cenário: supondo o parâmetro β desconhecido*

Nesta seção, apresentamos os resultados de simulação para o cenário em que o vetor de parâmetros de interesse é $\boldsymbol{\mu} = (\alpha, \beta)^\top$ e o parâmetro de perturbação $\nu = \lambda$. Os tamanhos amostrais considerados foram $n = 30,\ 40,\ 50$ e 80. As estatísticas LR_p, LR_{BN_1}, LR_{BN_2} são baseadas na funções de verossimilhanças dadas na Subseção 3.3.2.

O nosso objetivo é comparar as diferentes estatísticas, incluíndo a estatística LR_{bb} e o teste

Tabela 3.4: Taxas de rejeições não nulas, inferências sobre α ($n = 30$)

	$\xi = 10\%$				
α	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_{bb}
$4,5$	$17,05$	$20,30$	$20,48$	$20,24$	$17,93$
$5,5$	$34,14$	$38,88$	$39,22$	$38,84$	$35,14$
$6,5$	$52,89$	$57,93$	$58,18$	$57,91$	$54,17$
$7,5$	$67,84$	$72,59$	$72,65$	$72,58$	$67,16$
$8,5$	$78,38$	$81,95$	$82,10$	$81,98$	$77,67$
$9,5$	$84,34$	$87,20$	$87,23$	$81,98$	$84,62$
$10,5$	$89,25$	$91,29$	$91,27$	$91,33$	$89,46$
$11,5$	$92,91$	$94,40$	$94,46$	$94,44$	$92,46$
$12,5$	$94,28$	$95,65$	$95,73$	$95,73$	$94,05$
$13,5$	$96,09$	$97,03$	$97,02$	$97,09$	$95,93$
$14,5$	$96,85$	$97,55$	$97,55$	$97,57$	$96,70$
$15,5$	$97,53$	$98,16$	$98,15$	$98,17$	$97,71$

LR_b, para testar a hipótese nula $\mathcal{H}_0 : \boldsymbol{\mu} = \boldsymbol{\mu_0}$ *versus* $\mathcal{H}_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu_0}$. Para isso, Fixamos $\alpha = 0,5$, assumimos o valor verdadeiro do parâmetro $\lambda = 0,25$ e o parâmetro β variando em $0,25$ e $0,5$. Os níveis nominais considerados foram $\xi = 10\%,\ 5\%$ e 1%.

A Tabela 3.7 apresenta as taxas de rejeições nulas estimadas. Avaliando essa tabela, notamos que, de modo geral, as estatísticas LR_{BN_1}, LR_{BN_2}, LR_b e LR_{bb} tiveram os melhores desempenhos, pois os testes correspondentes forneceram taxas de rejeições mais próximas do nível nominal especificado. Por exemplo, para $n = 30$ e $\beta = 0,5$, ao nível de 10%, o teste baseado em LR_p tem taxa de rejeição igual a $12,13\%$ e os testes baseados nas demais estatísticas, as taxas de rejeições não ultrapassam $10,37\%$. Na Figura 3.4, a qual ilustra a situação dada na Tabela 3.7 quando $(\alpha;\beta) = (0,5;0,25)$ e $\xi = 10\%$, podemos notar que as curvas de distorções dos testes baseados em LR_{BN_1} e LR_{BN_2} seguem o mesmo padrão.

A Figura 3.5 apresenta o gráfico das discrepâncias relativas de quantis das estatísticas LR_p, LR_{BN_1}, LR_{BN_2} e LR_{bb}. Nota-se que as curvas da discrepância relativa de quantil possuem o mesmo padrão para todas as estatísticas. Observa-se ainda que, à medida que os quantis aumentam, a discrepância, em módulo, de LR_{bb} se torna maior que a discrepância de LR_{BN_2}. No entanto, para valores altos de quantis, a discrepância de LR_p é a que mais se aproxima da ordenada zero.

A Tabela 3.8 apresenta a média e a variância das estatísticas LR_p, LR_{BN_1}, LR_{BN_2} e LR_{bb}, estimadas através de simulação de Monte Carlo com $\alpha = 0,5$ e β variando em $0,25$ e $0,5$, para diferentes valores de n. Ainda com o intuito de avaliar a aproximação da distribuição nula das estatísticas pela distribuição χ^2_2, comparamos a média e a variância estimadas das

Tabela 3.5: Estimação pontual de α (λ = 0,5)

		$\alpha = 0,5$				$\alpha = 1,0$				$\alpha = 1,5$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
20	$\widehat{\alpha}_p$	0,131	0,034	1,625	7,981	0,167	0,209	2,109	12,033	0,197	0,634	2,529	16,589
	$\widetilde{\alpha}_{BN}$	0,092	0,030	1,605	7,844	0,120	0,177	2,065	11,622	0,143	0,526	2,460	15,823
	$\check{\alpha}_{BN}$	0,091	0,030	1,604	7,861	0,118	0,176	2,079	11,789	0,141	0,524	2,479	16,044
	$\widehat{\alpha}_{CR_a}$	0,095	0,030	1,603	7,838	0,123	0,179	2,071	11,679	0,147	0,534	2,469	15,926
30	$\widehat{\alpha}_p$	0,085	0,018	1,115	5,478	0,105	0,099	1,371	6,787	0,122	0,283	1,556	7,909
	$\widetilde{\alpha}_{BN}$	0,060	0,016	1,108	5,445	0,076	0,088	1,357	6,711	0,089	0,248	1,536	7,786
	$\check{\alpha}_{BN}$	0,059	0,016	1,107	5,438	0,074	0,088	1,361	6,712	0,088	0,247	1,538	7,790
	$\widehat{\alpha}_{CR_a}$	0,061	0,016	1,107	5,442	0,077	0,089	1,355	6,695	0,090	0,250	1,539	7,799
40	$\widehat{\alpha}_p$	0,061	0,012	0,885	4,330	0,076	0,064	1,047	4,855	0,087	0,177	1,167	5,343
	$\widetilde{\alpha}_{BN}$	0,043	0,011	0,881	4,319	0,055	0,058	1,041	4,831	0,063	0,160	1,159	5,305
	$\check{\alpha}_{BN}$	0,043	0,011	0,883	4,327	0,054	0,058	1,040	4,830	0,063	0,160	1,161	5,316
	$\widehat{\alpha}_{CR_a}$	0,044	0,011	0,887	4,342	0,056	0,059	1,041	4,833	0,065	0,161	1,156	5,281
50	$\widehat{\alpha}_p$	0,049	0,009	0,828	4,396	0,060	0,047	0,977	4,864	0,068	0,130	1,068	5,141
	$\widetilde{\alpha}_{BN}$	0,035	0,008	0,825	4,387	0,043	0,044	0,972	4,847	0,050	0,119	1,062	5,117
	$\check{\alpha}_{BN}$	0,035	0,008	0,825	4,388	0,043	0,044	0,972	4,847	0,049	0,119	1,063	5,119
	$\widehat{\alpha}_{CR_a}$	0,035	0,008	0,826	4,387	0,044	0,044	0,972	4,845	0,051	0,120	1,078	5,218

Tabela 3.6: Taxas de rejeições nulas, inferências sobre α (λ = 0,5)

		$n = 20$						$n = 30$					
α	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}
0,5	10	11,04	10,00	10,03	11,02	10,37	10,26	10,56	10,02	10,04	10,68	10,15	10,06
	5	5,89	5,25	5,19	5,88	4,99	4,85	5,51	5,01	5,06	5,51	5,20	5,00
	1	1,40	1,14	1,13	1,34	1,07	0,89	1,09	1,04	1,05	1,17	1,22	1,10
1,0	10	11,30	10,29	10,25	11,31	10,20	10,24	10,68	10,08	10,10	10,70	10,11	9,95
	5	6,19	5,37	5,30	6,20	5,08	4,96	5,60	5,06	5,01	5,57	4,97	4,91
	1	1,45	1,04	1,03	1,47	1,08	0,95	1,13	0,95	0,94	1,16	1,04	0,81
1,5	10	11,60	10,39	10,39	11,69	10,05	9,97	10,85	9,87	9,89	10,60	10,31	10,10
	5	6,25	5,41	5,36	6,27	5,17	5,00	5,59	4,96	4,93	5,56	5,07	4,94
	1	1,36	1,05	1,04	1,40	1,10	0,95	1,14	1,00	0,99	1,14	1,01	0,82
		$n = 40$						$n = 50$					
α	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}
0,5	10	10,12	9,50	9,52	10,06	10,10	9,97	10,32	9,79	9,77	10,23	10,69	10,56
	5	5,11	4,71	4,74	5,12	5,01	4,84	5,04	4,80	4,80	4,98	5,77	5,52
	1	1,09	0,98	0,99	1,11	1,16	0,94	1,15	1,07	1,07	1,13	1,06	0,93
1,0	10	10,09	9,69	9,73	10,19	9,97	9,87	10,08	9,65	9,68	10,15	10,19	0,96
	5	5,00	4,77	4,81	5,10	5,09	5,04	5,03	4,75	4,76	5,05	5,15	5,12
	1	1,12	0,92	0,93	1,12	0,99	0,96	1,21	1,06	1,06	1,15	1,06	0,94
1,5	10	9,98	9,55	9,53	10,20	10,01	9,92	10,09	9,71	9,76	10,24	10,09	10,01
	5	5,03	4,72	4,73	5,10	5,02	4,89	5,11	4,84	4,80	5,13	5,12	5,07
	1	1,10	0,97	0,97	1,08	1,02	0,94	1,19	1,05	1,03	1,16	1,05	0,93

Tabela 3.7: Taxas de rejeições nulas (%), β desconhecido

		$n = 30$					$n = 40$				
β	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_b	LR_{bb}
0,25	10	12,10	9,79	10,26	9,62	10,03	11,26	9,68	9,89	10,12	10,71
	5	5,83	4,61	4,79	4,88	4,24	5,65	4,75	4,79	5,72	5,72
	1	0,96	0,68	0,71	1,15	0,68	1,24	0,84	0,87	1,32	0,89
0,5	10	12,13	9,97	10,00	10,31	10,37	9,68	9,81	9,74	10,65	10,83
	5	5,83	4,77	4,59	5,14	4,67	5,67	4,76	4,71	5,26	5,16
	1	0,98	0,77	0,67	1,13	0,68	1,04	0,86	0,86	1,21	0,90
		$n = 50$					$n = 80$				
β	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_b	LR_{bb}
0,25	10	11,11	9,83	9,96	10,17	10,38	10,21	9,75	9,85	10,12	10,05
	5	6,03	5,21	5,34	5,41	5,50	5,25	4,82	4,92	5,28	5,18
	1	1,15	1,07	1,09	1,32	0,97	1,15	1,02	1,07	1,21	1,09
0,5	10	11,09	9,85	10,02	10,35	9,99	10,28	9,76	9,79	10,47	10,31
	5	6,02	5,33	5,46	5,44	5,32	5,32	4,82	4,88	5,48	5,34
	1	1,17	1,07	1,08	1,15	0,92	1,12	1,05	1,07	1,12	1,07

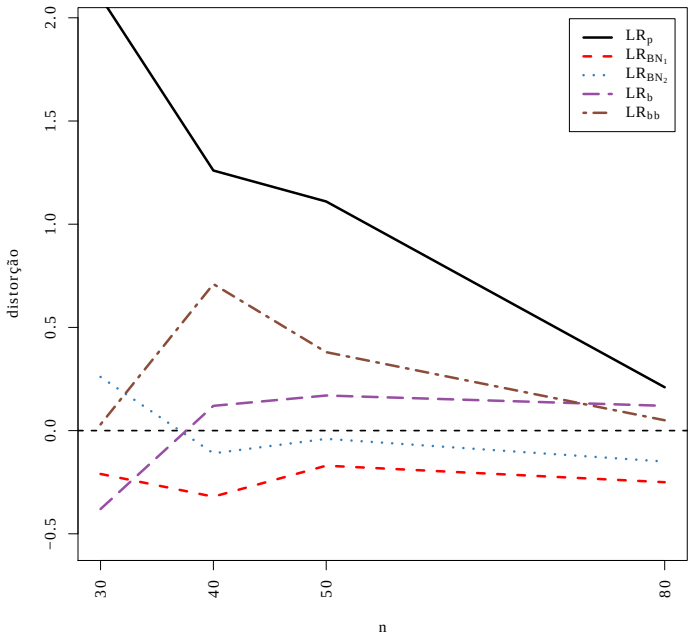


Figura 3.4: Distorções dos tamanhos dos testes: $\xi = 10\%$ e $(\alpha; \beta) = (0,5; 0,25)$

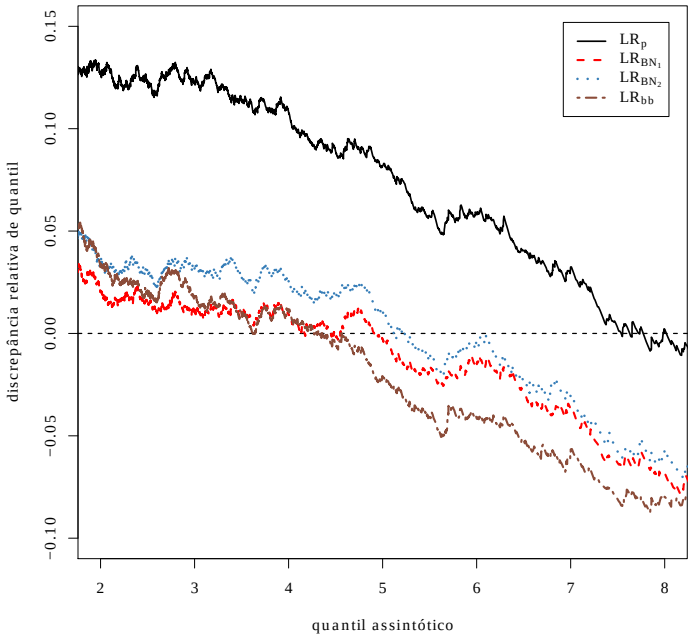


Figura 3.5: Discrepâncias relativas de quantis: $n = 30$ e $(\alpha; \beta) = (0,5; 0,25)$

Tabela 3.8: Média e variância (var.) das estatísticas de teste, β desconhecido

			$(\alpha;\beta)=(0,5;0,25)$				$(\alpha;\beta)=(0,5;0;5)$			
n		χ^2_2	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{bb}
30	média	2,00	2,09	1,97	2,00	1,98	2,15	1,98	1,98	2,00
	var.	4,00	4,07	3,64	3,75	3,57	4,27	3,71	3,66	3,70
40	média	2,00	2,16	1,98	1,99	2,02	2,10	1,98	1,99	2,02
	var.	4,00	4,25	3,73	3,79	3,99	4,16	3,74	3,78	3,92
50	média	2,00	2,09	2,00	2,01	2,00	2,09	2,00	2,01	2,00
	var.	4,00	4,37	3,98	4,06	4,04	4,37	4,02	4,07	3,96
80	média	2,00	2,02	1,98	1,98	2,00	2,03	1,97	1,98	2,04
	var.	4,00	4,13	3,99	3,99	4,09	4,15	3,94	3,97	4,13

Tabela 3.9: Taxas de rejeições não nulas (%), β desconhecido, $n = 30$

	$\xi = 10\%$			
δ	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{bb}
0,8	40,67	43,58	43,55	39,08
1,1	72,78	74,98	75,44	72,02
1,4	88,36	89,44	89,92	87,38
1,7	93,84	94,44	94,69	93,79
2,0	96,34	96,69	96,82	97,04
2,3	97,99	98,32	98,35	97,93
2,6	98,44	98,61	98,91	98,47
2,9	98,91	99,12	99,12	98,95
3,2	99,17	99,27	99,31	99,40

diferentes estatísticas com a média e a variância da distribuição χ^2_2. Notamos que as médias e as variâncias estimadas dessas estatísticas se aproximam da média e da variância da distribuição de referência.

A Tabela 3.9 apresenta os resultados de simulação obtidos levando em consideração a hipótese alternativa para $n = 30$, $\lambda = 0,25$ e diferentes valores para $\alpha = \beta = \delta$ foram considerados ao nível de 10%. Observar que essas simulações de poder correspondem à situação abordada na Tabela 3.7 para $n = 30$ e $\beta = 0,5$. Avaliando esta tabela concluímos os testes baseados nas estatísticas LR_{BN_1} e LR_{BN_2} apresentaram poderes semelhantes e foram mais poderosos que o teste baseado em LR_p e LR_{bb}.

3.5 Exemplos numéricos

Mostraremos agora dois exemplos nos quais utilizamos conjuntos de dados reais. Para o primeiro exemplo, consideramos um conjunto de dados (sem observações censuradas) referente ao tempo de sobrevivência (em anos) de um grupo de 36 pacientes tratados com quimioterapia e radioterapia, esses dados podem ser encontrados em Bekker *et al.* (2000). Nosso objetivo é fazer inferências para o parâmetro α que indexa a distribuição LE. Para isso, primeiramente, realizamos o teste de Kolmogorov-Smirnov para verificar a aderência dos dados à distribuição Pareto exponencializada de parâmetros $\hat{\alpha} = 2.182$ e $\hat{\lambda} = 2.865$, ou seja, à distribuição LE$(\hat{\alpha}, 1, \hat{\lambda})$. Para esse teste, o valor da estatística e o p-valor são, respectivamente, $0,096$ e $0,859$, o que indica que não há evidências para rejeitar a distribuição Pareto exponencializada para esses valores dos parâmetros.

Utilizando a função de verossimilhança perfilada e a suas versões modificadas dadas na Subseção 3.3.1, obtemos os seguintes resultados para estimação pontual de α: $\hat{\alpha}_p = 2,182$, $\hat{\tilde{\alpha}}_{BN} = 2,127$, $\hat{\mathring{\alpha}}_{BN} = 2,121$ e $\hat{\alpha}_{CR_a} = 2,129$. Aqui, estamos interessados em testar a hipótese nula $\mathcal{H}_0 : \alpha = 1,0$, ou seja, estamos testando o modelo Pareto. Neste caso, temos $LR_p = 8,778$, $LR_{BN_1} = 8,127$, $LR_{BN_2} = 8,073$, $LR_{CR_a} = 8,434$ e $LR_{bb} = 8,336$ e os correspondentes p-valores são: $0,003$; $0,004$; $0,004$; $0,004$ e $0,003$. O teste baseado em LR_b tem p-valor igual a $0,008$. Dessa forma, considerando os níveis de significância usuais, todos os testes rejeitam a hipótese nula, ou seja, o modelo Pareto é rejeitado. Adicionalmente, calculamos os valores de AIC (Akaike information criterion) correspondentes às distribuições Pareto exponencializada e Pareto, obtendo, respectivamente, $62,914$ e $69,693$. Logo, de acordo com este critério, para esse conjunto de dados, a distribuição Pareto exponencializada oferece um melhor ajuste (menor AIC) que o modelo Pareto.

Para o segundo exemplo, supomos que os dados seguem distribuição LE(α, β, λ) e consideramos um conjunto de dados referente ao tempo de reparação (em horas) de um transceptor, que é um aparelho usado para comunicação aérea (JØRGENSEN, 2012, p.165). Para esta análise obtemos a partir dos dados originais a seguinte subamostra aleatória: 2,0; 1,5; 5,4; 0,5; 1,5; 10,3; 1,0; 0,8; 24,5; 7,5; 9,0; 22,0; 1,3; 2,5; 8,8; 1,5; 3,0; 0,3; 0,5 e 2,7. Supomos que os dados seguem distribuição LE(α, β, λ). Neste caso, o nosso objetivo é fazer inferências para o vetor de parâmetros $\boldsymbol{\mu} = (\alpha; \beta)^\top$ usando as funções dadas na Subseção 3.3.2. Os resultados da estimação pontual para $\boldsymbol{\mu}$ são: $\hat{\boldsymbol{\mu}}_p = (3,919; 1,871)^\top$, $\hat{\tilde{\boldsymbol{\mu}}}_{BN} = (4,857; 2,854)^\top$ e $\hat{\mathring{\boldsymbol{\mu}}}_{BN} = (4,335; 2,400)^\top$. Estamos interessados em testar a hipótese nula $\mathcal{H}_0 : (\alpha; \beta) = (1; 1)$, isto é, estamos testando se os dados seguem a distribuição Pareto. As estatísticas $LR_p = 7,246$, $LR_{BN_1} = 6,504$, $LR_{BN_2} = 6,348$ e $LR_{bb} = 7.909$ e os p-valores correspondentes são: $0,02$; $0,04$; $0,04$ e $0,02$. O p-valor correspondente ao teste LR_b é igual a $0,02$. Portanto, ao nível

nominal de 5% todos os testes rejeitam a hipótese nula, ou seja, o modelo Pareto também é rejeitado para esse conjunto de dados.

3.6 Conclusões finais

Fizemos inferências (estimação pontual e teste de hipóteses) para dois parâmetros da distribuição Lomax exponencializada. Baseamo-nos em diferentes propostas e derivamos ajustes para a função de verossimilhança perfilada com o objetivo de diminuir o impacto do parâmetro de perturbação (λ), e consequentemente, obter resultados mais precisos, principalmente quando a amostra é pequena. Realizamos um estudo de simulação de Monte Carlo para avaliar os desempenhos da função de verossimilhança perfilada e suas versões modificadas no processo de estimação pontual e teste de hipóteses com base na estatística da razão de verossimilhanças. Consideramos também o teste baseado na estatística da razão de verossimilhanças bootstrap Bartlett e o teste da razão de verossimilhanças bootstrap usual. O estudo apontou que os estimadores de máxima verossimilhança perfilado ajustados tiveram os melhores desempenhos. Os testes baseados na estatística da razão de verossimilhanças que utiliza as versões modificadas da função de verossimilhança perfilada tiveram as menores distorções de tamanho. No entanto, os testes baseados em bootstrap, em geral, apresentaram resultados satisfatórios em todos os cenários considerados. Por fim, apresentamos dois exemplos numéricos utilizando conjuntos de dados reais.

4 Refinamento de inferências na distribuição BurrX escalonada

Resumo

A distribuição BurrX escalonada, além de ser bastante útil para modelar dados de tempo de vida, quando seu parâmetro de forma é igual a um, tem a distribuição Rayleigh como caso particular. Este capítulo apresenta três diferentes ajustes para a função de verossimilhança perfilada com o objetivo de obter inferências mais confiáveis sobre o parâmetro de forma da distribuição BurrX escalonada. Especificamente, derivamos a função de verossimilhança perfilada e suas versões ajustadas para fazer estimação pontual e teste de hipóteses com base na estatística da razão de verossimilhanças (LR), sob o enfoque de amostra completa (dados não censurados) e amostra contendo dados censurados do tipo II. O estudo também inclui o teste baseado na estatística da razão de verossimilhanças bootstrap Bartlett (ROCKE, 1989) e o teste da razão de verossimilhanças bootstrap usual (EFRON; TIBSHIRANI, 1994). Os resultados de simulação indicam que em pequenas amostras as inferências baseadas nas funções de verossimilhança perfilada ajustadas são mais precisas. Duas aplicações usando conjuntos de dados reais são apresentadas para ilustrar a teoria desenvolvida.

Palavras-chave: BurrX escalonada; Censura tipo II; Estimação pontual; Teste de hipóteses; Verossimilhança perfilada ajustada

4.1 Introdução

Burr (1942) introduziu doze famílias de distribuições, dentre elas destacam-se as distribuições Burr tipo III e a Burr tipo X, as quais têm recebido muita atenção de pesquisadores em diversas áreas. Surles e Padgett (2001) propuseram a distribuição Burr tipo X com um parâmetro adicional, denominada de BurrX escalonada, com o objetivo de fazer inferência para $R = P(X < Y)$ (confiabilidade) de fibras de carbono, em que X e Y são variáveis aleatórias independentes que seguem distribuição BurrX escalonada.

Por ser uma generalização da distribuição Rayleigh, a BurrX escalonada também é chamada por alguns autores de Rayleigh generalizada ou Rayleigh exponencializada (RAQAB; MADI,

2011; ABD-ELFATTAH, 2011). Nesta tese, usamos a denominação dada por Surles e Padgett (2001) e será denotada por BurrX (σ, λ), tal que σ e λ são parâmetros de escala e forma, respectivamente. Essa distribuição é bastante apropriada para analisar dados de tempo de vida assimétricos (AHMAD *et al.*, 2017).

A função de verossimilhança perfilada pode ser usada quando desejamos fazer inferências apenas para uma parte dos parâmetros que indexam um modelo. No entanto, a função de verossimilhança perfilada não é uma verossimilhança genuína, assim algumas propriedades tais como função escore e informação não viesadas podem não ser satisfeitas. Além disso, a distribuição nula da estatística da razão de verossimilhanças pode não ser bem aproximada pela distribuição qui-quadrado de referência na presença de muitos parâmetros de perturbação ou ainda quando a amostra é pequena. Sendo assim, com base em duas aproximações para a função de verossimilhança proposta por Barndorff-Nielsen (1983) e na função de verossimilhança de Cox e Reid (1993), o nosso principal objetivo é obter três versões modificadas da função de verossimilhança perfilada que forneçam inferências mais confiáveis para o parâmetro λ da distribuição BurrX escalonada. Estudos nessa linha já foram realizados por Cysneiros *et al.* (2008) e Barreto *et al.* (2013) para a distribuição Birnbaum-Saunders para dados completos e dados censurados do tipo II, respectivamente. Ferrari *et al.* (2007) obtiveram diferentes ajustes da verossimilhança perfilada para a distribuição Weibull.

Para nosso estudo inferencial consideramos dois enfoques. No primeiro enfoque, obtivemos três ajustes para a função de verossimilhança perfilada para o parâmetro λ em amostra completa (dados não censurados). No segundo, derivamos dois ajustes para a função de verossimilhança perfilada para amostra contendo observações censuradas do tipo II. Nos dois enfoques, fizemos estimação por máxima verossimilhança e teste de hipóteses usando a estatística da razão de verossimilhanças com base nos diferentes ajustes. Consideramos também o teste baseado na estatística da razão de verossimilhanças bootstrap Bartlett (ROCKE, 1989) e o teste da razão de verossimilhanças bootstrap usual (EFRON; TIBSHIRANI, 1994).

O capítulo está organizado da seguinte forma: na Seção 4.2 são discutidos alguns aspectos e propriedades da distribuição BurrX escalonada; na Seção 4.3, obtemos os ajustes para a função de log-verossimilhança perfilada para a distribuição BurrX escalonada. Resultados de simulação de Monte Carlo são mostrados e discutidos na Seção 4.4. Duas aplicações a conjunto de dados reais, uma para amostra completa e outra para amostra censurada do tipo II, são apresentadas na Seção 4.5. As conclusões finais deste capítulo são dadas na Seção 4.6. No Apêndice 6, discutimos alguns aspectos computacionais utilizados neste capítulo.

4.2 A distribuição BurrX escalonada

A distribuição BurrX escalonada tem função de distribuição acumulada (fda), função densidade de probabilidade (fdp) e função de sobrevivência dadas, respectivamente, por

$$F(x) = \left[1 - e^{-\left(\frac{x}{\sigma}\right)^2}\right]^{\lambda}, \tag{4.1}$$

$$f(x) = \frac{2\lambda x e^{-\left(\frac{x}{\sigma}\right)^2}}{\sigma^2}\left[1 - e^{-\left(\frac{x}{\sigma}\right)^2}\right]^{\lambda-1} \tag{4.2}$$

e

$$S(x) = 1 - \left[1 - e^{-\left(\frac{x}{\sigma}\right)^2}\right]^{\lambda}, \tag{4.3}$$

em que $x > 0$, $\sigma > 0$ e $\lambda > 0$.

A distribuição BurrX escalonada é denominada por alguns autores de Burr X biparamétrica, Rayleigh exponencializada ou Rayleigh generalizada (RAQAB; KUNDU, 2006). As duas últimas denominações são devido ao fato de que (4.1) pode ser obtida elevando a fda da distribuição Rayleigh a um parâmetro positivo.

A fdp da distribuição BurrX escalonada pode ser assimétrica unimodal ($\lambda > 1/2$) ou descrescente ($\lambda \leq 1/2$) como podemos observar na Figura 4.1. No que segue, uma variável X que tem fdp dada por (4.2) será denotada por $X \sim$ BurrX(σ, λ).

A função quantílica (fq) de X é dada por

$$Q(u; \sigma, \lambda) = F^{-1}(u) = \sigma\left[-\log\left(1 - u^{\frac{1}{\lambda}}\right)\right]^{\frac{1}{2}}, 0 < u < 1. \tag{4.4}$$

Podemos usar (4.4) para gerar números aleatórios da distribuição BurrX escalonada. Portanto, se $U \sim U(0, 1)$ então $Q(U; \sigma, \lambda) \sim$ BurrX(σ, λ).

4.2.1 *Resultados gerais para inferência*

Seja $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ uma amostra aleatória de tamanho n proveniente da distribuição BurrX escalonada cuja a fdp é dada em (4.2), então a função de log-verossimilhança para $\boldsymbol{\theta} = (\sigma, \lambda)^\top$ é dada por

$$\ell(\boldsymbol{\theta}) = n\log\left(\frac{2\lambda}{\sigma}\right) + (\lambda - 1)\sum_{i=1}^{n}\log[1 - e^{-u_i^2}] + \sum_{i=1}^{n}[\log(u_i) - u_i^2]. \tag{4.5}$$

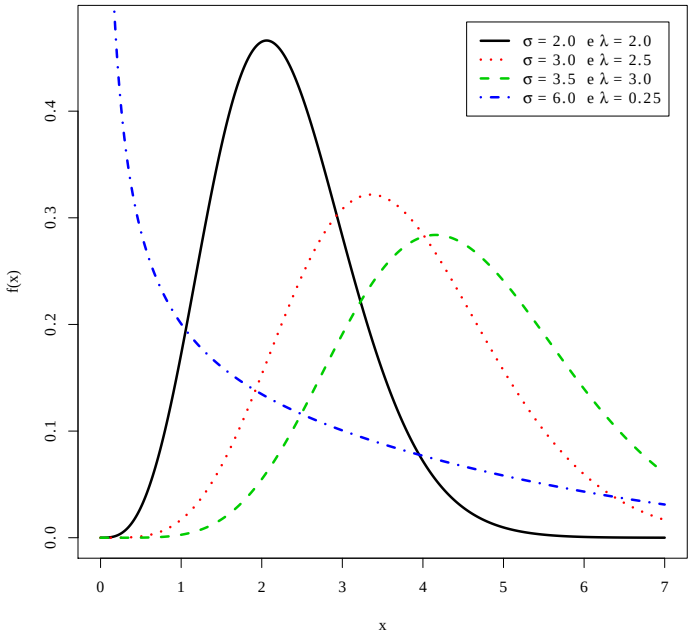


Figura 4.1: Fdp da distribuição BurrX escalonada para diferentes valores de parâmetros

Derivando a função log-verossimilhança dada por (4.5), obtemos as componentes do vetor escore $\ell_{\boldsymbol{\theta}} = (\ell_\sigma, \ell_\lambda)^\top$, as quais são dadas por

$$\ell_\sigma = -\frac{2n}{\sigma} + \frac{2}{\sigma}\sum_{i=1}^{n} u_i^2 - \frac{2(\lambda-1)}{\sigma}\sum_{i=1}^{n}\frac{u_i^2 e^{-u_i^2}}{G(u_i)},$$

$$\ell_\lambda = \frac{n}{\lambda} + \sum_{i=1}^{n}\log[G(u_i)],$$

em que $u_i = \frac{x_i}{\sigma}$ e $G(t) = 1 - e^{-t^2},\ t > 0$.

O estimador de máxima verossimilhança (EMV) de $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}} = (\hat{\sigma}, \hat{\lambda})^\top$, pode ser obtido resolvendo a equação não-linear $\ell_\sigma = \ell_\lambda = 0$. Neste caso, esta equação não pode ser resolvida analiticamente. Contudo, pode ser avaliada numericamente (NOCEDAL; WRIGHT, 2006). Alternativamente, $\hat{\boldsymbol{\theta}}$ pode ser obtido maximizando diretamente (4.5) usando as plataformas computacionais `SAS` (`PROC NLMIXED`), `R` (funções `optim` e `MaxLik`) ou plataforma `Ox` (função `MaxBFGS`).

A matriz de informação observada é dada por

$$J(\boldsymbol{\theta}) = \begin{bmatrix} j_{\sigma\sigma} & j_{\sigma\lambda} \\ \\ j_{\lambda\sigma} & j_{\lambda\lambda} \end{bmatrix},$$

em que

$$j_{\sigma\sigma} = -\frac{\partial^2 \ell(\theta)}{\partial \sigma^2} = -\frac{2n}{\sigma^2} + \frac{6}{\sigma^2}\sum_{i=1}^{n} u_i^2 - \frac{2(\lambda-1)}{\sigma^2}\sum_{i=1}^{n} \frac{u_i^2 e^{-u_i^2}}{G(u_i)}\left[3 - 2u_i^2 - \frac{2u_i^2 e^{-u_i^2}}{G(u_i)}\right],$$

$$j_{\sigma\lambda} = j_{\lambda\sigma} = -\frac{\partial^2 \ell(\theta)}{\partial \sigma \partial \lambda} = \frac{2}{\sigma}\sum_{i=1}^{n} \frac{u_i^2 e^{-u_i^2}}{G(u_i)}$$

e

$$j_{\lambda\lambda} = -\frac{\partial^2 \ell(\theta)}{\partial \lambda^2} = \frac{n}{\lambda^2}.$$

Resultado 1. *Se* $X \sim \mathrm{BurrX}(\sigma, \lambda)$*, então para* $\lambda \neq 1$ *e* $\lambda \neq 2$

$$\begin{aligned} -\mathrm{E}\left\{\frac{\partial^2 \log[f(X;\boldsymbol{\theta})]}{\partial \sigma^2}\right\} &= a_{\sigma,\lambda} + \frac{2}{3\sigma^2 b_\lambda}\{12 - 24\gamma + c\lambda + d\lambda^2 + f\lambda^3 + 6(\lambda-1)^2 \\ &\times \left[4\gamma\Psi(\lambda-1) + \lambda(\lambda-1)\Psi^2(\lambda-1)\right] + 6b_\lambda\Psi(\lambda)[2\gamma \\ &- \lambda\Psi(\lambda)] + 12g_\lambda\Psi^{(1)}(\lambda-1) + 6h_\lambda\Psi^{(1)}\}, \end{aligned}$$

em que

$$a_{\sigma,\lambda} = a(\sigma,\lambda) = \frac{6H_\lambda - 2 + 6\sigma^2[1 - \gamma - \Psi(\lambda)]}{\sigma^2}, \quad b_\lambda = b(\lambda) = (\lambda-2)(\lambda-1)^2,$$

$$g_\lambda = g(\lambda) = -(\lambda-1)^2, \quad h_\lambda = h(\lambda) = 2 - 5\lambda + 4\lambda^2 - 6\lambda^3,$$

$$c = 6 + 12\gamma + 6\gamma^2 + \pi^2, \quad d = -2c + 6, \quad f = -12\gamma + 6\gamma^2 + \pi^2,$$

tal que $H_z = H(z) = \Psi(z+1) + \gamma$ *é o número harmônico,* $\Psi(z) = \dfrac{d}{dz}\log[\Gamma(z)]$ *é a função digama,* $\gamma \simeq 0.577216$ *é a constante de Euler e* $\Psi^{(n)}(z) = \dfrac{d^n\Psi(z)}{dz^n}$.

Resultado 2. *Se* $X \sim \mathrm{BurrX}(\sigma, \lambda)$*, então para* $\lambda \neq 1$

$$-\mathrm{E}\left\{\frac{\partial^2 \log[f(X;\boldsymbol{\theta})]}{\partial\sigma\lambda}\right\} = \frac{2}{\lambda\sigma}\left\{\frac{\lambda[\gamma + \Psi(\lambda)] - (\lambda - 1)}{\lambda - 1}\right\}$$

Resultado 3. *Se* $X \sim \mathrm{BurrX}(\sigma, \lambda)$*, então*

$$-\mathrm{E}\left\{\frac{\partial^2 \log[f(X;\boldsymbol{\theta})]}{\partial\lambda^2}\right\} = \frac{n}{\lambda^2}.$$

Os Resultados 1-3 serão úteis para a obtenção da função de verossimilhança perfilada de Cox e Reid (1993) na próxima seção.

4.3 Verossimilhanças perfiladas ajustadas para a BurrX escalonada

Obteremos ajustes para a função de verossimilhança perfilada e suas versões modificadas para o parâmetro λ da distribuição BurrX escalonada. Neste caso, o parâmetro σ é um parâmetro de perturbação. Consideramos dados completos e dados sob censura do tipo II.

4.3.1 *Amostra completa (dados não censurados)*

Seja $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ uma amostra aleatória da distribuição BurrX (σ, λ). A função de log-verossimilhança para o parâmetro λ é obtido substituindo σ por seu estimador restrito, $\hat{\sigma}_\lambda$, na equação (4.5). Dessa forma, obtemos

$$\ell_p(\lambda) = n\log\left(\frac{2\lambda}{\hat{\sigma}_\lambda}\right) + (\lambda - 1)\sum_{i=1}^{n}\log\left[1 - e^{-\tilde{u}_i^2}\right] + \sum_{i=1}^{n}[\log(\tilde{u}_i) - \tilde{u}_i^2], \tag{4.6}$$

em que $\tilde{u}_i = x_i/\hat{\sigma}_\lambda$ e $\hat{\sigma}_\lambda$ é obtido como solução da equação $\ell_\sigma = 0$ para λ fixo. Como $\hat{\sigma}_\lambda$ não tem forma fechada deve ser encontrado utilizando um método de otimização não-linear restrita. Denotaremos o estimador de máxima verossimilhança perfilado de λ, que maximiza $\ell_p(\lambda)$, por $\hat{\lambda}_p$.

Derivamos duas aproximações para o logaritmo da função de verossimilhança perfilada modificada de Barndorff-Nielsen (1983), $\tilde{\ell}_{BN}(\lambda)$ e $\breve{\ell}_{BN}(\lambda)$, as quais são dadas por

$$\tilde{\ell}_{BN}(\lambda) = \ell_p(\lambda) + \frac{1}{2}\log|j_{\sigma\sigma}(\hat{\sigma}_\lambda, \lambda)| - \log|\ell_{\sigma;X}(\hat{\sigma}_\lambda, \lambda)\widehat{V}_\sigma|, \tag{4.7}$$

$$\breve{\ell}_{BN}(\lambda) = \ell_p(\lambda) + \frac{1}{2}\log|j_{\sigma\sigma}(\hat{\sigma}_\lambda, \lambda)| - \log|\breve{I}_\sigma(\hat{\sigma}_\lambda, \lambda; \hat{\sigma}, \hat{\lambda})|. \tag{4.8}$$

Sendo que

$$j_{\sigma\sigma}(\hat{\sigma}_\lambda, \lambda) = -\frac{2n}{\hat{\sigma}_\lambda^2} + \frac{6}{\hat{\sigma}_\lambda^2}\sum_{i=1}^{n}\tilde{u}_i^2 - \frac{2(\lambda-1)}{\hat{\sigma}_\lambda^2}\sum_{i=1}^{n}\frac{\tilde{u}_i^2 e^{-\tilde{u}_i^2}}{G(\tilde{u}_i)}\left[3 - 2\tilde{u}_i^2 - \frac{2\tilde{u}_i^2 e^{-\tilde{u}_i^2}}{G(\tilde{u}_i)}\right]$$

$$\ell_{\sigma;X}(\hat{\sigma}_\lambda, \lambda)\widehat{V}_\sigma = \sum_{i=1}^{n}\frac{4\hat{u}_i\tilde{u}_i\{e^{2\tilde{u}_i^2} + \lambda + e^{\tilde{u}_i^2}[-1 - \tilde{u}_i^2 + \lambda(\tilde{u}_i^2 - 1)]\}}{\hat{\sigma}_\lambda^2(\tilde{u}_i^2 - 1)^2},$$

$$\breve{I}_\sigma(\hat{\sigma}_\lambda, \lambda; \hat{\sigma}, \hat{\lambda}) = \sum_{i=1}^{n}\frac{4[1 - \lambda\tilde{u}_i^2 + e^{\tilde{u}_i^2}(\tilde{u}_i^2 - 1)][1 - \hat{\lambda}\hat{u}_i^2 + e^{\hat{u}_i^2}(\hat{u}_i^2 - 1)]}{\hat{\sigma}_\lambda\hat{\sigma}(e^{\tilde{u}_i^2} - 1)(e^{\hat{u}_i^2} - 1)},$$

em que $\hat{u}_i = x_i/\hat{\sigma}$, $\hat{\sigma}$ e $\hat{\lambda}$ são os EMVs dos parâmetros σ e λ, respectivamente.

Os parâmetros σ e λ que indexam a distribuição BurrX escalonada não são ortogonais. Assim, não foi possível obter o ajuste dado em (2.7). Alternativamente, usando (2.8) obtemos a função de log-verossimilhança perfilada modificada dada por

$$\ell_{CR_a}(\lambda) = \ell_p(\lambda) - \frac{1}{2}\log|j_{\sigma\sigma}(\hat{\sigma}_\lambda, \lambda)| - (\lambda - \hat{\lambda})m_\sigma(\hat{\sigma}, \hat{\lambda}), \tag{4.9}$$

em que

$$m_\sigma(\hat{\sigma}, \hat{\lambda}) = K_{\hat{\sigma},\hat{\lambda}}\frac{A_{\hat{\sigma},\hat{\lambda}}}{B_{\hat{\sigma},\hat{\lambda}}},$$

tal que

$$K_{\hat{\sigma},\hat{\lambda}} = \frac{4\hat{\lambda}[\gamma + \Psi(\hat{\lambda})] - 4(\hat{\lambda} - 1)}{3\hat{\lambda}^2\hat{\sigma}^4 b_{\hat{\lambda}}(\hat{\lambda} - 1)},$$

$$\begin{aligned}
A_{\hat{\sigma},\hat{\lambda}} &= \hat{\lambda}[9(2 + 7\hat{\sigma}^2 - 2\gamma\hat{\sigma}^2 + b_{\hat{\lambda}}H_{\hat{\lambda}}) - 24\gamma] - \hat{\lambda}^2(15 - c + 81\hat{\sigma}^2 - 45\gamma\hat{\sigma}^2) \\
&+ \hat{\lambda}^3(6 - 2f + 45\hat{\sigma}^2 - 36\gamma\hat{\sigma}^2) + \hat{\lambda}^4[f - 3 - 9\hat{\sigma}^2(1 - \gamma)] - 6\hat{\lambda}g_{\hat{\lambda}}[4\gamma \\
&+ \hat{\lambda}(\hat{\lambda} - 1)\Psi(\hat{\lambda} - 1)]\Psi(\hat{\lambda} - 1) - 3\hat{\lambda}(h_{\hat{\lambda}} + 5\hat{\lambda}^3)[4\gamma + 3\hat{\sigma}^2 - 2\hat{\lambda}\Psi(\hat{\lambda})]\Psi(\hat{\lambda}) \\
&+ 6\hat{\lambda}[2g_{\hat{\lambda}}\Psi^{(1)}(\hat{\lambda} - 1) + (h_{\hat{\lambda}} + 5\hat{\lambda}^3)\Psi^{(1)}(\hat{\lambda}) - 18\hat{\sigma}^2],
\end{aligned}$$

e

$$\begin{aligned} B_{\hat{\sigma},\hat{\lambda}} &= \{\hat{\lambda}^{-1}(6-\hat{\lambda}a_{\hat{\sigma},\hat{\lambda}}) - \frac{2}{3\hat{\sigma}^2 b_{\hat{\lambda}}}[12(1-2\gamma)+c\hat{\lambda}-2(f+3)\hat{\lambda}^2+f\hat{\lambda}^3 \\ &- 6g_{\hat{\lambda}}[4\gamma+\hat{\lambda}(\hat{\lambda}-1)]\Psi(\hat{\lambda}-1)+6b_{\hat{\lambda}}(2\gamma-\hat{\lambda}\Psi(\hat{\lambda}))\Psi(\hat{\lambda})+2g_{\hat{\lambda}}\Psi^{(1)}(\hat{\lambda}-1) \\ &+ 6(h_{\hat{\lambda}}+5\hat{\lambda}^3)\Psi^{(1)}(\hat{\lambda})]\}^2. \end{aligned}$$

Os termos $a_{\hat{\sigma},\hat{\lambda}}$, $b_{\hat{\lambda}}$, $H_{\hat{\lambda}}$, $h_{\hat{\lambda}}$ e $g_{\hat{\lambda}}$ são obtidos avaliando, respectivamente, as funções $a_{\sigma,\lambda}$, b_λ, H_λ, h_λ e g_λ (definidas na Subseção 4.2.1) em $\hat{\lambda}$ e $\hat{\sigma}$. É importante ressaltar que a log-verossimilhança dada por (4.9) somente está definida quando $\hat{\lambda} \neq 1$ e $\hat{\lambda} \neq 2$.

Os estimadores de máxima verossimilhança perfilados ajustados de λ obtidos a partir de (4.7), (4.8), e (4.9) são denotados, respectivamente, por $\hat{\tilde{\lambda}}_{BN}$, $\hat{\tilde{\lambda}}_{BN}$ e $\hat{\lambda}_{CR_a}$. Esses estimadores e o estimador de máxima verossimilhança perfilado, $\hat{\lambda}_p$, não têm forma fechada e devem ser obtidos maximizando numericamente as respectivas funções de log-verossimilhança.

4.3.2 *Dados censurados do tipo II*

Vamos considerar agora dados sob censura do tipo II, em que a observação termina na r-ésima ($r < n$) falha. Sejam $x_{(1)}, \ldots, x_{(r)}$ as r menores estatísticas de ordem, obtidas a partir de uma amostra de tamanho n da distribuição BurrX (σ, λ). Neste caso, a função de verossimilhança é dada por

$$L(\sigma,\lambda) = [S(x_{(r)};\sigma,\lambda)]^{n-r}\prod_{i=1}^{r} f(x_{(i)};\sigma,\lambda).$$

em que as funções $f(\cdot)$ e $S(\cdot)$ são dadas por (4.2) e (4.3), respectivamente. Consequentemente,

$$\begin{aligned} \ell(\sigma,\lambda) &= r\log\left(\frac{2\lambda}{\sigma}\right) + (n-r)\log[1-G^\lambda(u_{(r)})] + \sum_{i=1}^{r}\log[G(u_{(i)})] \\ &+ \sum_{i=1}^{r}[\log(u_{(i)}) - u_{(i)}^2]. \end{aligned} \tag{4.10}$$

Assim,

$$\begin{aligned} \ell_p(\lambda) &= r\log\left(\frac{2\lambda}{\hat{\sigma}_\lambda}\right) + (n-r)\log[1-G^\lambda(\tilde{u}_{(r)})] + \sum_{i=1}^{r}\log[G(\tilde{u}_{(i)})] \\ &+ \sum_{i=1}^{r}[\log(\tilde{u}_{(i)}) - \tilde{u}_{(i)}^2]. \end{aligned} \tag{4.11}$$

Aqui, $G^a(t) = [G(t)]^a$, $\tilde{u}_{(i)} = x_{(i)}/\hat{\sigma}_\lambda$ e $\hat{\sigma}_\lambda$ é solução da equação

$$-\frac{2r}{\sigma} + \frac{2}{\sigma}\sum_{i=1}^{r}\left[u_{(i)}^2 - \frac{(\lambda-1)e^{-u_{(i)}^2}u_{(i)}^2}{G(u_{(i)})}\right] + \frac{2\lambda(n-r)u_{(r)}^2 e^{-u_{(r)}^2}G^{\lambda-1}(u_{(r)})}{\sigma[1-G^\lambda(u_{(r)})]} = 0$$

para λ fixo.

A partir da Equação (2.4), derivamos

$$\tilde{\ell}_{BN}(\lambda) = \ell_p(\lambda) + \frac{1}{2}\log|j_{\sigma\sigma}(\hat{\sigma}_\lambda,\lambda)| - \log|\ell_{\sigma;X}(\hat{\sigma}_\lambda,\lambda)\widehat{V}_\sigma|, \tag{4.12}$$

em que

$$\begin{aligned} j_{\sigma\sigma}(\hat{\sigma}_\lambda,\lambda) &= -\frac{2r}{\hat{\sigma}_\lambda^2} + \frac{6}{\hat{\sigma}_\lambda^2}\sum_{i=1}^{r}\tilde{u}_{(i)}^2 - \frac{(\lambda-1)}{\hat{\sigma}_\lambda^2}\sum_{i=1}^{r}\frac{\tilde{u}_{(i)}^2 e^{-\tilde{u}_{(i)}^2}}{G(\tilde{u}_{(i)})}\left[6 - 4\tilde{u}_{(i)}^2 - \frac{4\tilde{u}_{(i)}^2 e^{-\tilde{u}_{(i)}^2}}{G(\tilde{u}_{(i)})}\right] \\ &+ \frac{6\lambda(n-r)e^{-\tilde{u}_{(r)}^2}G^{\lambda-1}(\tilde{u}_{(r)})}{\hat{\sigma}_\lambda^2[1-G^\lambda(\tilde{u}_{(r)})]} - \frac{4\lambda(n-r)\tilde{u}_{(r)}^4 e^{-\tilde{u}_{(r)}^2}G^{\lambda-1}(\tilde{u}_{(r)})}{\hat{\sigma}_\lambda^2[1-G^\lambda(\tilde{u}_{(r)})]} \\ &\times \left[1 - (\lambda-1)e^{-\tilde{u}_{(r)}^2}G^{-1}(\tilde{u}_{(r)}) - \frac{\lambda e^{-\tilde{u}_{(r)}^2}G^{\lambda-1}(\tilde{u}_{(r)})}{1-G^\lambda(\tilde{u}_{(r)})}\right], \end{aligned}$$

$\widehat{V}_\sigma = (\hat{u}_{(1)},\ldots,\hat{u}_{(r-1)},\overbrace{\hat{u}_{(r)},\ldots,\hat{u}_{(r)}}^{(n-r+1)\text{ elementos}})^\top$ e $\ell_{\sigma;\boldsymbol{X}}(\lambda,\hat{\sigma}_\lambda) = (\ell_{\sigma;X_1},\ldots,\ell_{\sigma;X_{r-1}},\overbrace{\ell_{\sigma;X_r},\ldots,\ell_{\sigma;X_r}}^{(n-r+1)\text{ elementos}})$,

tal que

$$\ell_{\sigma;X_i} = \frac{\partial\ell_\sigma}{\partial x_{(i)}} = \frac{4\tilde{u}_{(i)}}{\hat{\sigma}_\lambda^2}\left\{1 + \frac{(\lambda-1)e^{-\tilde{u}_{(i)}^2}}{G(\tilde{u}_{(i)})}\left[-1 + \tilde{u}_{(i)}^2 + \frac{\tilde{u}_{(i)}^2 e^{-\tilde{u}_{(i)}^2}}{G(\tilde{u}_{(i)})}\right]\right\},\ i = 1,\ldots,r-1$$

e

$$\begin{aligned} \ell_{\sigma;X_r} &= \frac{4\lambda(n-r)\tilde{u}_{(r)}e^{-\tilde{u}_{(r)}^2}G(\tilde{u}_{(r)})^{\lambda-1}}{\hat{\sigma}_\lambda^2[1-G^\lambda(\tilde{u}_{(r)})]}\Big\{1 - \tilde{u}_{(r)}^2 + (\lambda-1)\tilde{u}_{(r)}^2 e^{-\tilde{u}_{(r)}^2}G^{-1}(\tilde{u}_{(r)}) \\ &+ \frac{\lambda\tilde{u}_{(r)}^2 e^{-\tilde{u}_{(r)}^2}G(\tilde{u}_{(r)})^{\lambda-1}}{[1-G(\tilde{u}_{(r)})^\lambda]}\Big\} + \frac{4\tilde{u}_{(r)}}{\hat{\sigma}_\lambda^2}\left\{1 + \frac{(\lambda-1)e^{-\tilde{u}_{(r)}^2}}{G(\tilde{u}_{(r)})}\left[-1 + \tilde{u}_{(r)}^2 + \frac{\tilde{u}_{(r)}^2 e^{-\tilde{u}_{(r)}^2}}{G(\tilde{u}_{(r)})}\right]\right\}. \end{aligned}$$

Assim, temos

$$\begin{aligned}\ell_{\sigma;X}(\hat{\sigma}_\lambda,\lambda)\widehat{V}_\sigma &= 4\sum_{i=1}^{r-1}\frac{\tilde{u}_{(i)}\hat{u}_{(i)}}{\hat{\sigma}_\lambda^2}\left[1+\frac{(\lambda-1)e^{-\tilde{u}_{(i)}^2}}{G(\tilde{u}_{(i)})}\left(-1+\tilde{u}_{(i)}^2+\frac{\tilde{u}_{(i)}^2e^{-\tilde{u}_{(i)}^2}}{G(\tilde{u}_{(i)})}\right)\right]\\ &+ \frac{4\lambda(n-r)(n-r+1)\tilde{u}_{(r)}\hat{u}_{(r)}e^{-\tilde{u}_{(r)}^2}G^{\lambda-1}(\tilde{u}_{(r)})}{\hat{\sigma}_\lambda^2[1-G^\lambda(\tilde{u}_{(r)})]}\\ &\times \left[1-\tilde{u}_{(r)}^2+(\lambda-1)\tilde{u}_{(r)}^2e^{-\tilde{u}_{(r)}^2}G^{-1}(\tilde{u}_{(r)})+\frac{\lambda\tilde{u}_{(r)}^2e^{-\tilde{u}_{(r)}^2}G^{\lambda-1}(\tilde{u}_{(r)})}{1-G^\lambda(\tilde{u}_{(r)})}\right]\\ &+ \frac{4(n-r+1)\hat{u}_{(r)}\tilde{u}_{(r)}}{\hat{\sigma}_\lambda^2}\left[1+\frac{(\lambda-1)e^{-\tilde{u}_{(r)}^2}}{G(\tilde{u}_{(r)})}\left(-1+\tilde{u}_{(r)}^2+\frac{\tilde{u}_{(r)}^2e^{-\tilde{u}_{(r)}^2}}{G(\tilde{u}_{(r)})}\right)\right].\end{aligned}$$

Agora, usando o ajuste (2.5), obtemos

$$\breve{\ell}_{BN}(\lambda)=\ell_p(\lambda)+\frac{1}{2}\log|j_{\sigma\sigma}(\hat{\sigma}_\lambda,\lambda)|-\log|\breve{I}_\sigma(\hat{\sigma}_\lambda,\lambda;\hat{\sigma},\hat{\lambda})|, \tag{4.13}$$

em que

$$\begin{aligned}\breve{I}_\sigma(\hat{\sigma}_\lambda,\lambda;\hat{\sigma},\hat{\lambda}) &= \frac{4}{\hat{\sigma}_\lambda\hat{\sigma}}\sum_{i=1}^{r-1}\frac{[1-\lambda\tilde{u}_{(i)}^2+e^{\tilde{u}_{(i)}^2}(\tilde{u}_{(i)}^2-1)][1-\hat{\lambda}\hat{u}_{(i)}^2+e^{\hat{u}_{(i)}^2}(\hat{u}_{(i)}^2-1)]}{[e^{\tilde{u}_{(i)}^2}-1][e^{\hat{u}_{(i)}^2}-1]}\\ &+ \frac{4(n-r+1)[1-\lambda\tilde{u}_{(r)}^2+e^{\tilde{u}_{(r)}^2}(\tilde{u}_{(r)}^2-1)][1-\hat{\lambda}\hat{u}_{(r)}^2+e^{\hat{u}_{(r)}^2}(\hat{u}_{(r)}^2-1)]}{\hat{\sigma}_\lambda\hat{\sigma}(e^{\tilde{u}_{(r)}^2}-1)(e^{\hat{u}_{(r)}^2}-1)}.\end{aligned}$$

Aqui $\hat{u}_r = x_r/\hat{\sigma}$, $\hat{\sigma}$ e $\hat{\lambda}$ são, respectivamente, EMVs de σ e λ sob censura do tipo II e podem ser obtidos maximizando a função de log-verossimilhança dada por (4.10).

4.4 Resultados numéricos

Nesta seção apresentamos os resultados das simulações de Monte Carlo, obtidos para fazer inferências sobre o parâmetro de forma λ da distribuição BurrX escalonada e assim, comparar o desempenho das funções de verossimilhança perfilada e suas versões modificadas. Para isso, consideramos amostras completas (dados não censurados) e amostras com dados censurados (censura do tipo II).

Todas as simulações foram realizadas usando a linguagem de programação `Ox` (DOORNIK, 2009) e os resultados são baseados em 10 000 réplicas de Monte Carlo. Os tamanhos amostrais considerados neste estudo foram $n = 20$, $n = 30$, $n = 50$ e $n = 100$. Os níveis nominais foram $\xi = 10\%$, $\xi = 5\%$ e $\xi = 1\%$. Para os testes baseados em bootstrap paramétrico, utilizamos

$B = 1\,000$ réplicas.

4.4.1 *Amostra completa*

Neste enfoque, para os parâmetros da distribuição BurrX escalonada, supomos o valor verdadeiro de σ igual a 50 e λ variando em: $3,0$; $3,5$ e $4,0$. Apresentamos os resultados numéricos referentes à estimação pontual e teste de hipóteses utilizando amostras completas. Para isso, consideramos os estimadores de máxima verossimilhança $\hat{\lambda}_p$, $\hat{\tilde{\lambda}}_{BN}$, $\hat{\breve{\lambda}}_{BN}$ e $\hat{\lambda}_{CR_a}$ obtidos maximizando as funções ℓ_p, $\tilde{\ell}_{BN}$, $\breve{\ell}_{BN}$ e ℓ_{CR_a}, respectivamente. A comparação desses estimadores é baseada na medida do viés relativo (VR), do erro quadrático médio (EQM), assimetria e curtose.

Em relação, ao teste de hipóteses, avaliamos os desempenhos dos testes baseados nas estatísticas LR_p, LR_{BN_1}, LR_{BN_2}, LR_{CR_a}, LR_{bb} e do teste LR_b. As estatísticas supracitadas foram definidas na Seção 2.3.

O estimador restrito do parâmetro σ, $\hat{\sigma}_\lambda$, foi estimado utilizando o algoritmo de otimização não-linear restrito SQP (*Sequential Quadratic Programming*) (NOCEDAL; WRIGHT, 2006). Esse algoritmo está implementando na plataforma `Ox` através da função `MaxSQP`.

Na Tabela 4.1, apresentamos os resultados de simulação para os estimadores de máxima verossimilhança e suas versões modificadas. Podemos verificar nesta tabela que os estimadores perfilados modificados possuem vieses relativos menores quando comparado ao viés relativo do estimador $\hat{\lambda}_p$. Por exemplo, para $n = 20$ e $\lambda = 3,0$, o viés relativo de $\hat{\lambda}_p$ é aproximadamente 27%, enquanto que o viés relativo dos estimadores corrigidos não ultrapassam 20%.

Podemos notar também que os valores associados ao VR, EQM, assimetria e curtose dos estimadores avaliados, diminuem à medida que o tamanho da amostra aumenta. Por exemplo, considerando $\lambda = 3.0$, para $n = 20$, as medidas do VR, EQM, assimetria e curtose dos estimadores de λ não ultrapassam, respectivamente, 0,269; 4,652; 3,645; 33,012. No entanto, para $n = 100$, temos que as medidas do VR, EQM, assimetria e curtose não ultrapassam, respectivamente, 0,039; 0,295; 0,866; 4,764. Da Tabela 4.1, pode-se conluir ainda que $\hat{\lambda}_{CR_a}$ apresentou menor valor em todas as medidas em todos os cenários.

Na Tabela 4.2, mostramos as taxas de rejeição nulas estimadas referentes ao teste da hipótese nula $\mathcal{H}_0 : \lambda = \lambda_0$ *versus* $\mathcal{H}_1 : \lambda \neq \lambda_0$ usando as estatísticas de teste LR_p, LR_{BN_1}, LR_{BN_2} e LR_{CR_a}, LR_b e LR_{bb}. Avaliando essa tabela, observamos que, de modo geral, os testes baseados nas estatísticas LR_{BN_1}, LR_{BN_2}, LR_{CR_a}, LR_b e LR_{bb} possuem taxas de rejeições estimadas mais próximas do nível nominal especificado. Por exemplo, para $\lambda = 3,0$, $n = 20$ e $\xi = 10\%$, as taxas de rejeições estimadas não ultrapassam $10,61\%$, enquanto que o teste baseado em LR_p possue taxa de rejeição estimada de $11,93\%$.

A Figura 4.2 mostra a distorção dos testes, ou seja, a diferença entre a taxa de rejeição estimada e o nível nominal correspondente. O gráfico ilustra a situação da Tabela 4.2 quando

Tabela 4.1: Estimação pontual de λ (σ = 50)

		$\lambda = 3,0$				$\lambda = 3,5$				$\lambda = 4,0$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
20	$\widehat{\lambda}_p$	0,269	4,652	3,645	33,012	0,289	7,367	3,998	39,397	0,308	11,026	4,347	46,215
	$\widehat{\tilde{\lambda}}_{BN}$	0,197	3,720	3,499	30,746	0,212	5,835	3,826	36,489	0,226	8,656	4,147	42,600
	$\widehat{\breve{\lambda}}_{BN}$	0,195	3,730	3,522	30,770	0,211	5,861	3,846	36,355	0,225	8,711	4,165	42,267
	$\widehat{\lambda}_{CR_a}$	0,153	3,263	3,439	29,838	0,166	5,098	3,754	35,318	0,178	7,537	4,063	41,133
30	$\widehat{\lambda}_p$	0,156	1,745	1,861	9,708	0,166	2,650	1,953	10,371	0,175	3,816	2,037	11,012
	$\widehat{\tilde{\lambda}}_{BN}$	0,114	1,498	1,830	9,470	0,122	2,262	1,918	10,087	0,129	3,240	1,998	10,681
	$\widehat{\breve{\lambda}}_{BN}$	0,113	1,496	1,836	9,523	0,121	2,261	1,925	10,162	0,128	3,240	2,007	10,781
	$\widehat{\lambda}_{CR_a}$	0,088	1,370	1,817	9,371	0,094	2,065	1,903	9,969	0,100	2,951	1,982	10,544
50	$\widehat{\lambda}_p$	0,092	0,768	1,389	7,508	0,097	1,148	1,450	7,916	0,102	1,629	1,506	8,309
	$\widehat{\tilde{\lambda}}_{BN}$	0,069	0,697	1,377	7,429	0,073	1,038	1,436	7,824	0,077	1,467	1,491	8,204
	$\widehat{\breve{\lambda}}_{BN}$	0,068	0,696	1,380	7,442	0,073	1,037	1,439	7,837	0,077	1,467	1,494	8,218
	$\widehat{\lambda}_{CR_a}$	0,054	0,658	1,372	7,397	0,058	0,979	1,431	7,786	0,061	1,383	1,485	8,161
100	$\widehat{\lambda}_p$	0,039	0,295	0,866	4,764	0,042	0,436	0,906	4,945	0,044	0,612	0,941	5,115
	$\widehat{\tilde{\lambda}}_{BN}$	0,028	0,281	0,863	4,749	0,030	0,414	0,902	4,927	0,032	0,581	0,937	5,095
	$\widehat{\breve{\lambda}}_{BN}$	0,028	0,281	0,864	4,753	0,030	0,414	0,903	4,932	0,032	0,581	0,938	5,101
	$\widehat{\lambda}_{CR_a}$	0,021	0,273	0,861	4,743	0,023	0,403	0,900	4,919	0,025	0,564	0,935	5,085

$\lambda = 4,0$ e $\xi = 10\%$. Como podemos observar, os testes baseados nas estatísticas LR_{CR_a}, LR_b e LR_{bb} são os que mais se aproximam da linha de referência (ordenada zero) quando $n \leq 50$, portanto nestes casos, essas estatísticas tiveram o melhor desempenho. Porém, os testes baseados em LR_{CR_a} têm um comportamento mais estável (distorção com pouca variação à medida que n cresce).

É importante destacar também que as taxas de rejeições estimadas, dadas na Tabela 4.2, associadas aos testes baseados nas estatísticas LR_b e LR_{bb} apresentam uma diferença mínima. Isso indica que essas estatísticas tiveram um desempenho semelhante.

A Figura 4.3 apresenta o gráfico das discrepâncias relativas de quantis (diferença entre o quantil exato e quantil assintótico dividida por este último) das estatísticas LR_p, LR_{BN_1}, LR_{BN_2}, LR_{CR_a} e LR_{bb}. Quanto mais próxima da ordenada zero a curva estiver, melhor será a aproximação assintótica utilizada no teste. Então, de acordo com o gráfico, as distribuições das estatísticas LR_{CR_a} e LR_{bb} são as que mais se aproximam da distribuição nula assintótica (χ^2_1). A estatística LR_b não foi considerada nesta comparação porque, ao contrário das demais estatísticas, as taxas de rejeições dos testes baseados em LR_b foram estimadas utilizando o quantil da distribuição empírica da estatística LR_p.

Uma avaliação do impacto do número de réplicas bootstrap no desemenho de LR_b e LR_{bb} é apresentada na Tabela 4.3. Neste caso, obtemos as taxas de rejeições estimadas considerando $n = 20$, $\lambda = 4,0$. Variando o nível nominal ξ em: 10%, 5% e 1% e o número de réplicas bootstrap B em: 100, 300, 600 e 1 000. Como podemos notar para $B = 100$ o teste baseado em LR_{bb} teve o melhor desempenho, pois forneceram taxas de rejeições mais próximas do nível nominal especificado. Esse fato é esperado, pois o teste baseado em LR_{bb} necessita de menos réplicas bootstrap que o teste LR_b para ser preciso. Nos demais casos, as duas estatísticas tiveram desempenho bastante similar.

A Tabela 4.4 apresenta a média e a variância das estatísticas LR_p, LR_{BN_1}, LR_{BN_2}, LR_{CR_a} e LR_{bb}, estimadas através de simulação de Monte Carlo quando $\lambda = 3,0$ e $\lambda = 3,5$, para diferentes valores de n. Ainda com o intuito de avaliar a aproximação da distribuição nula das estatísticas pela distribuição χ^2_1, comparamos a média e a variância estimadas das diferentes estatísticas com a média e a variância da distribuição χ^2_1. Notamos que as médias e as variâncias estimadas das estatísticas LR_{BN_1}, LR_{BN_2}, LR_{CR_a} e LR_{bb} são as que mais aproximam da média e variância da distribuição de referência, principalmente quando a amostra é pequena. Por exemplo, para $n = 20$ e $\lambda = 3,0$, a média e a variância das estatísticas LR_{BN_1}, LR_{BN_2}, LR_{CR_a} e LR_{bb} não ultrapassam, respectivamente, 1,020 e 2,067. Porém a média e a variância da estatística LR_p são, respectivamente, 1,099 e 2,414. No entanto, para amostras de tamanho grande, a média e a variância de LR_p são semelhantes às demais estatísticas. Isso ocorre porque assintoticamente essas estatísticas têm o mesmo comportamento.

Tabela 4.2: Taxas de rejeições nulas, inferências sobre λ (σ = 50)

		$n = 20$						$n = 30$					
λ	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}
3,0	10	11,93	10,61	10,55	10,01	9,95	9,80	10,87	10,14	10,12	9,88	10,35	10,30
	5	6,30	5,36	5,35	5,01	5,11	5,18	5,33	4,78	4,82	4,66	4,76	4,73
	1	1,35	1,01	0,98	0,94	1,04	0,87	1,14	0,97	0,95	0,88	0,94	0,87
3,5	10	12,07	10,61	10,59	9,89	10,50	10,46	10,92	10,08	10,12	9,82	9,68	9,63
	5	6,31	5,28	5,36	4,99	5,28	5,34	5,36	4,82	4,83	4,68	4,87	4,92
	1	1,39	1,00	0,97	0,93	1,23	1,14	1,19	0,95	0,92	0,84	1,04	0,97
4,0	10	12,07	10,57	10,49	9,87	9,63	9,57	10,95	10,06	10,08	9,88	10,28	10,19
	5	6,39	5,24	5,27	4,96	4,75	4,76	5,44	4,79	4,73	4,64	5,18	5,11
	1	1,39	0,97	0,95	0,94	1,05	1,00	1,23	0,94	0,94	0,83	1,00	0,92
		$n = 50$						$n = 100$					
λ	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_b	LR_{bb}
3,0	10	9,86	9,46	9,52	9,47	10,46	10,40	9,93	9,79	9,77	9,73	10,18	10,02
	5	5,19	4,75	4,76	4,63	5,48	5,43	4,92	4,69	4,66	4,56	4,72	4,64
	1	1,06	0,87	0,87	0,78	1,05	0,95	0,83	0,83	0,82	0,84	0,99	0,90
3,5	10	9,85	9,51	9,48	9,40	9,89	9,82	9,89	9,83	9,84	9,73	10,73	10,68
	5	5,22	4,78	4,78	4,70	4,62	4,57	4,90	4,67	4,66	4,61	5,48	5,44
	1	1,02	0,88	0,87	0,78	1,08	1,01	0,84	0,81	0,82	0,82	1,29	1,25
4,0	10	9,96	9,49	9,57	9,36	10,06	9,99	9,91	9,86	9,85	9,80	9,85	9,80
	5	5,24	4,78	4,76	4,67	5,07	5,01	4,86	4,77	4,79	4,60	4,93	4,96
	1	0,99	0,88	0,88	0,78	1,12	1,01	0,85	0,80	0,80	0,81	1,24	1,14

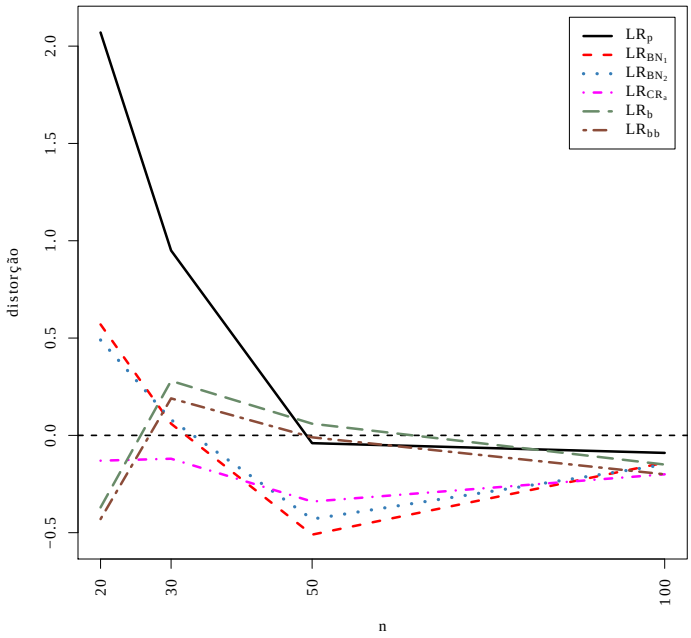


Figura 4.2: Distorção dos testes: $n = 20$, $\xi = 10\%$ e $\lambda = 4,0$ ($\sigma = 50$)

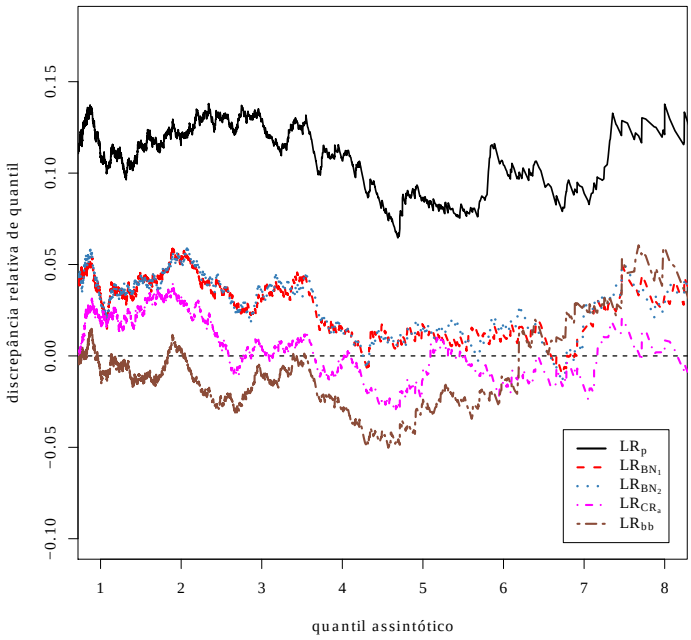


Figura 4.3: Discrepância relativa de quantil para $n = 20$ e $\lambda = 4,0$ ($\sigma = 50$)

Tabela 4.3: Taxas de rejeições nulas (%), $(\sigma;\lambda) = (50;4)$ e $n = 20$

	Nível nominal					
	$\xi = 10\%$		$\xi = 5\%$		$\xi = 1\%$	
B	LR_b	LR_{bb}	LR_b	LR_{bb}	LR_b	LR_{bb}
100	10,63	10,28	5,78	5,34	1,90	1,24
300	10,22	10,04	5,08	4,86	1,19	0,93
600	9,72	9,62	5,18	5,09	1,02	0,89
1000	10,50	10,33	5,20	5,02	1,12	1,00

Tabela 4.4: Média e variância (var.) das estatísticas de teste ($\sigma = 50$)

			$\lambda = 3,0$					$\lambda = 3,5$				
n		χ^2_1	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_{bb}
20	média	1,000	1,099	1,020	1,020	0,998	0,999	1,101	1,020	1,021	0,999	1,021
	var.	2,000	2,414	2,064	2,067	1,964	2,007	2,418	2,061	2,066	1,961	2,029
30	média	1,000	1,045	0,994	0,994	0,980	0,995	1,045	0,993	0,993	0,980	0,984
	var.	2,000	2,128	1,932	1,931	1,882	1,939	2,130	1,931	1,931	1,882	1,933
50	média	1,000	1,007	0,974	0,974	0,963	1,000	1,008	0,974	0,973	0,963	1,006
	var.	2,000	2,032	1,900	1,899	1,859	2,039	2,031	1,897	1,896	1,857	1,980
100	média	1,000	0,992	0,979	0,979	0,976	0,998	0,993	0,979	0,979	0,976	1,037
	var.	2,000	1,972	1,917	1,917	1,904	1,941	1,977	1,921	1,921	1,908	2,143

A Tabela 4.5 apresenta os resultados da simulação de poder, ou seja, as taxas de rejeições ao testar a hipótese nula $\mathcal{H}_0 : \lambda = 3,0$, quando essa hipótese é falsa. Neste caso, as taxas de rejeições foram obtidas considerando $n = 50$, o nível nominal $\xi = 10\%$ e diferentes valores de λ na hipótese alternativa $\mathcal{H}_1 : \lambda \neq 3,0$. Como os tamanhos dos testes baseados nas estatísticas LR_p, LR_{BN_1}, LR_{BN_2}, LR_{CR_a} e LR_{bb} apresentam diferenças entre as taxas de rejeição para um nível nominal fixado, os testes foram realizados utilizando os valores críticos estimados para cada estatística, dessa forma colocamos todos os testes em um mesmo patamar. Aqui, não avaliamos os testes baseados na estatística LR_b porque não é possível corrigi-lo em relação ao seu tamanho. Avaliando os resultados do poder dos testes, observamos que o teste baseado e LR_{CR_a} apresentou maior poder.

Tabela 4.5: Taxas de rejeições não nulas (%), $n = 50$

	$\xi = 10\%$				
λ	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{CR_a}	LR_{bb}
$3,5$	$14,91$	$16,89$	$16,91$	$18,01$	$14,75$
$4,0$	$30,33$	$33,85$	$33,88$	$35,98$	$29,35$
$4,5$	$47,98$	$51,80$	$51,85$	$54,14$	$47,12$
$5,0$	$63,45$	$67,12$	$67,10$	$68,85$	$62,41$
$5,5$	$75,23$	$77,95$	$77,95$	$79,34$	$75,65$
$6,0$	$83,68$	$85,76$	$85,79$	$86,99$	$83,08$
$6,5$	$89,16$	$90,94$	$90,97$	$91,75$	$89,11$
$7,0$	$92,99$	$94,08$	$94,05$	$94,52$	$92,49$
$7,5$	$95,99$	$96,57$	$96,57$	$96,94$	$95,97$
$8,0$	$96,81$	$97,50$	$97,48$	$97,84$	$96,77$
$8,5$	$98,17$	$98,53$	$98,54$	$98,71$	$97,97$

4.4.2 *Dados censurados do tipo II*

Aqui, apresentamos os resultados numéricos referentes às simulações de Monte Carlo sob o enfoque de censura do tipo II. Neste cenário, os valores verdadeiros do parâmetro de interesse λ são: 0,75 e 1,5; o valor de σ é fixado em 30 e as porcentagens de observações censuradas consideradas para todos os tamanhos amostrais foram $p = 10\%, 30\%, \ 50\%$.

A Tabela 4.6 apresenta os resultados da estimação pontual de $\hat{\lambda}_p$, $\hat{\hat{\lambda}}_{BN}$ e $\hat{\overset{\circ}{\lambda}}_{BN}$, obtidos maximizando as funções (4.11), (4.12) e (4.13), respectivamente. Nesta tabela, podemos observar que $\hat{\hat{\lambda}}_{BN}$ e $\hat{\overset{\circ}{\lambda}}_{BN}$, têm desempenho superior com relação ao estimador $\hat{\lambda}_p$ em todos os casos. Por exemplo, para $n = 20$ e $n = 30$, à medida que o percentual de observações censuradas aumenta, $\hat{\hat{\lambda}}_{BN}$ e $\hat{\overset{\circ}{\lambda}}_{BN}$ apresentam melhores medidas para o VR, EQM, assimetria e curtose quando comparadas com as medidas referentes à λ_p. Por exemplo, para $\lambda = 0,75$, $n = 20$ e $p = 50\%$, o VR de $\hat{\hat{\lambda}}_{BN}$ e $\hat{\overset{\circ}{\lambda}}_{BN}$ são, respectivamente, 0,037 e 0,125, enquanto que o VR e o EQM de λ_p são, respectivamente, 0,339 e 0,405.

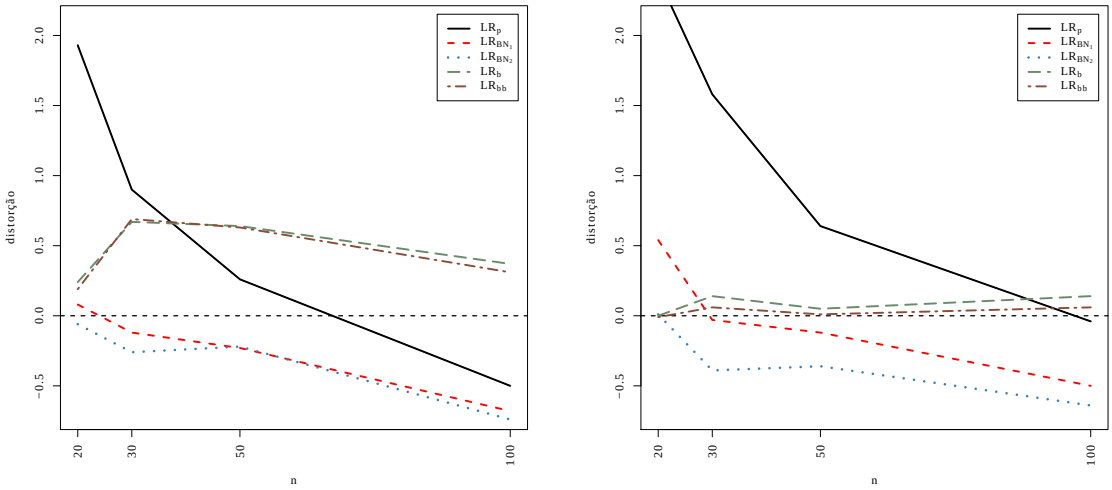
Os resultados sobre o teste de hipóteses baseados nas estatísticas LR_p, LR_{BN_1}, LR_{BN_2}, LR_b e LR_{bb} são dados na Tabela 4.7. Ao analisar essa tabela notamos que os testes baseados em LR_p, na maioria dos casos, foram os que apresentaram as maiores taxas de rejeições. Por exemplo, para $n = 20$, $\lambda = 1,5$ e $p = 30\%$, ao nível nominal de 10%, a taxa de rejeição de LR_p é aproximadamente $12,41\%$, enquanto que as taxas de rejeições dos testes baseados nas demais estatísticas não ultrapassam $10,54\%$. A Figura 4.4 mostra a distorção dos testes supracitados e ilustra a situação apresentada na Tabela 4.7 quando $\lambda = 1,5$; $\xi = 10\%$ para as diferentes valores de n e porcentagens de observações censuradas. Nesta figura, notamos clara-

Tabela 4.6: Estimação pontual de λ sob censura tipo II ($\sigma = 30$)

			$\lambda = 0,75$				$\lambda = 1,5$			
n	$p(\%)$	EMV	VR	EQM	Assimetria	Curtose	VR	EQM	Assimetria	Curtose
20	10	$\hat{\lambda}_p$	0,168	0,115	2,186	13,158	0,220	0,774	3,086	23,857
		$\hat{\hat{\lambda}}_{BN}$	0,065	0,086	2,154	12,805	0,112	0,569	2,983	22,429
		$\hat{\check{\lambda}}_{BN}$	0,054	0,082	2,140	12,707	0,094	0,537	2,985	22,518
	30	$\hat{\lambda}_p$	0,229	0,193	4,246	52,975	0,306	1,598	8,953	200,488
		$\hat{\hat{\lambda}}_{BN}$	0,054	0,129	4,179	51,062	0,137	1,092	8,367	178,350
		$\hat{\check{\lambda}}_{BN}$	0,069	0,124	3,984	46,968	0,127	0,981	7,902	162,429
	50	$\hat{\lambda}_p$	0,339	0,405	6,168	89,946	0,478	4,234	10,545	210,973
		$\hat{\hat{\lambda}}_{BN}$	0,037	0,234	6,118	85,714	0,171	2,479	9,784	185,109
		$\hat{\check{\lambda}}_{BN}$	0,125	0,247	5,065	66,495	0,261	2,567	9,091	164,039
30	10	$\hat{\lambda}_p$	0,105	0,054	1,287	6,282	0,133	0,316	1,596	8,307
		$\hat{\hat{\lambda}}_{BN}$	0,041	0,044	1,283	6,256	0,068	0,254	1,581	8,193
		$\hat{\check{\lambda}}_{BN}$	0,034	0,042	1,276	6,220	0,056	0,243	1,573	8,144
	30	$\hat{\lambda}_p$	0,139	0,076	1,528	7,790	0,177	0,470	1,925	10,506
		$\hat{\hat{\lambda}}_{BN}$	0,032	0,057	1,536	7,828	0,077	0,355	1,914	10,399
		$\hat{\check{\lambda}}_{BN}$	0,042	0,056	1,512	7,673	0,072	0,336	1,878	10,116
	50	$\hat{\lambda}_p$	0,201	0,130	1,998	11,294	0,260	0,879	2,652	16,892
		$\hat{\hat{\lambda}}_{BN}$	0,007	0,084	2,016	11,470	0,087	0,589	2,628	16,542
		$\hat{\check{\lambda}}_{BN}$	0,075	0,095	1,969	11,021	0,142	0,641	2,541	15,664
50	10	$\hat{\lambda}_p$	0,061	0,026	0,980	4,914	0,075	0,146	1,167	5,631
		$\hat{\hat{\lambda}}_{BN}$	0,024	0,023	0,980	4,911	0,039	0,128	1,163	5,613
		$\hat{\check{\lambda}}_{BN}$	0,020	0,023	0,976	4,898	0,032	0,124	1,159	5,596
	30	$\hat{\lambda}_p$	0,076	0,034	1,077	5,087	0,095	0,196	1,307	6,104
		$\hat{\hat{\lambda}}_{BN}$	0,016	0,029	1,081	5,098	0,040	0,164	1,304	6,089
		$\hat{\check{\lambda}}_{BN}$	0,021	0,028	1,073	5,067	0,037	0,159	1,292	6,030
	50	$\hat{\lambda}_p$	0,107	0,053	1,561	9,954	0,134	0,320	2,237	19,245
		$\hat{\hat{\lambda}}_{BN}$	0,001	0,040	1,562	10,007	0,041	0,247	2,226	18,957
		$\hat{\check{\lambda}}_{BN}$	0,037	0,043	1,550	9,803	0,072	0,262	2,173	18,125
100	10	$\hat{\lambda}_p$	0,031	0,011	0,698	4,408	0,038	0,058	0,801	4,711
		$\hat{\hat{\lambda}}_{BN}$	0,014	0,010	0,698	4,407	0,022	0,054	0,800	4,708
		$\hat{\check{\lambda}}_{BN}$	0,011	0,010	0,697	4,402	0,017	0,053	0,798	4,700
	30	$\hat{\lambda}_p$	0,040	0,014	0,014	4,293	0,049	0,075	0,865	4,684
		$\hat{\hat{\lambda}}_{BN}$	0,011	0,013	0,739	4,298	0,023	0,068	0,865	4,683
		$\hat{\check{\lambda}}_{BN}$	0,014	0,012	0,736	4,288	0,021	0,067	0,861	4,668
	50	$\hat{\lambda}_p$	0,053	0,019	0,836	4,428	0,065	0,106	1,012	5,090
		$\hat{\hat{\lambda}}_{BN}$	0,003	0,016	0,838	4,436	0,021	0,092	1,012	5,089
		$\hat{\check{\lambda}}_{BN}$	0,020	0,017	0,834	4,423	0,036	0,095	1,006	5,063

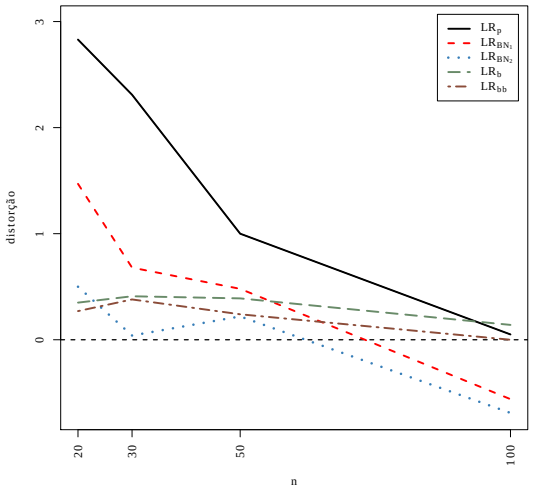
Tabela 4.7: Taxas de rejeições nulas (%), inferência sobre λ sob censura tipo II ($\sigma = 30$)

			$\lambda = 0,75$					$\lambda = 1,5$				
n	$p(\%)$	$\xi(\%)$	LR_p	LR_{BN_1}	LR_{BN_2}	LR_b	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_b	LR_{bb}
20	10	10	11,64	10,06	9,97	10,28	10,17	11,93	10,08	9,94	10,24	10,19
		5	6,14	5,18	5,06	5,08	5,00	6,18	5,20	5,08	4,96	4,89
		1	1,41	1,10	1,12	1,18	1,11	1,36	1,14	1,08	1,07	0,92
	30	10	12,27	10,83	9,94	9,90	9,90	12,41	10,54	10,01	10,00	9,99
		5	6,63	5,65	5,00	4,83	4,72	6,59	5,33	5,10	5,39	5,28
		1	1,53	1,21	1,02	0,88	0,78	1,65	1,14	1,04	1,11	0,89
	50	10	12,70	9,85	10,58	9,90	9,89	12,83	11,47	10,50	10,35	10,27
		5	7,20	4,10	5,42	5,11	4,97	7,38	5,87	5,12	10,35	4,91
		1	1,68	0,60	1,26	1,04	0,97	1,81	1,43	1,15	1,10	0,95
30	10	10	10,91	9,86	9,75	10,26	10,23	10,90	9,88	9,74	10,67	10,69
		5	5,39	4,87	4,77	5,16	5,06	5,44	4,83	4,87	5,55	5,50
		1	1,15	1,01	0,98	1,01	0,95	1,17	0,91	0,91	1,36	1,18
	30	10	11,53	10,25	9,64	9,95	9,78	11,58	9,97	9,61	10,14	10,06
		5	5,92	5,03	4,71	4,86	4,86	5,90	4,91	4,77	5,07	4,98
		1	1,30	1,11	1,04	1,19	1,14	1,35	1,06	1,03	1,00	0,98
	50	10	12,36	11,43	10,30	10,59	10,54	12,31	10,68	10,04	10,41	10,38
		5	6,44	6,12	5,13	5,27	5,22	6,44	5,45	5,01	5,52	5,54
		1	1,44	1,20	1,03	1,44	1,31	1,45	1,16	0,95	1,27	1,14
50	10	10	10,21	9,67	9,67	10,03	10,12	10,26	9,77	9,78	10,64	10,63
		5	5,15	5,02	5,04	5,12	4,99	5,30	4,81	4,79	5,38	5,35
		1	1,16	0,98	0,97	1,14	1,06	1,18	0,94	0,98	1,17	1,12
	30	10	10,65	10,18	9,90	9,62	9,50	10,64	9,88	9,64	10,05	10,01
		5	5,46	5,17	4,95	4,89	4,79	5,56	4,89	4,76	5,14	5,02
		1	1,28	1,06	1,02	1,04	0,94	1,28	0,99	0,96	1,28	1,18
	50	10	11,01	10,86	10,16	10,67	10,55	11,00	10,48	10,22	10,39	10,24
		5	5,54	5,41	4,69	5,18	5,18	5,58	4,80	4,68	5,13	5,08
		1	1,19	1,15	1,02	1,08	0,99	1,20	1,10	1,05	1,01	0,95
100	10	10	9,51	9,17	9,08	10,30	10,19	9,50	9,32	9,26	10,37	10,31
		5	5,01	4,66	4,59	5,30	5,27	4,90	4,63	4,52	5,17	5,15
		1	0,91	0,88	0,91	1,19	1,13	0,90	0,86	0,85	1,10	1,03
	30	10	10,23	9,58	9,43	10,13	9,96	9,96	9,50	9,36	10,14	10,06
		5	5,07	4,75	4,67	5,19	5,12	5,07	4,77	4,68	4,91	4,86
		1	1,10	0,95	0,93	1,09	1,04	1,06	0,94	0,92	1,08	0,97
	50	10	10,02	9,66	9,41	10,68	10,55	10,05	9,44	9,31	10,14	10,00
		5	5,21	5,13	4,96	5,48	5,31	5,23	4,83	4,84	4,93	4,92
		1	1,07	0,97	0,93	1,15	1,04	1,07	0,98	0,91	1,01	0,93



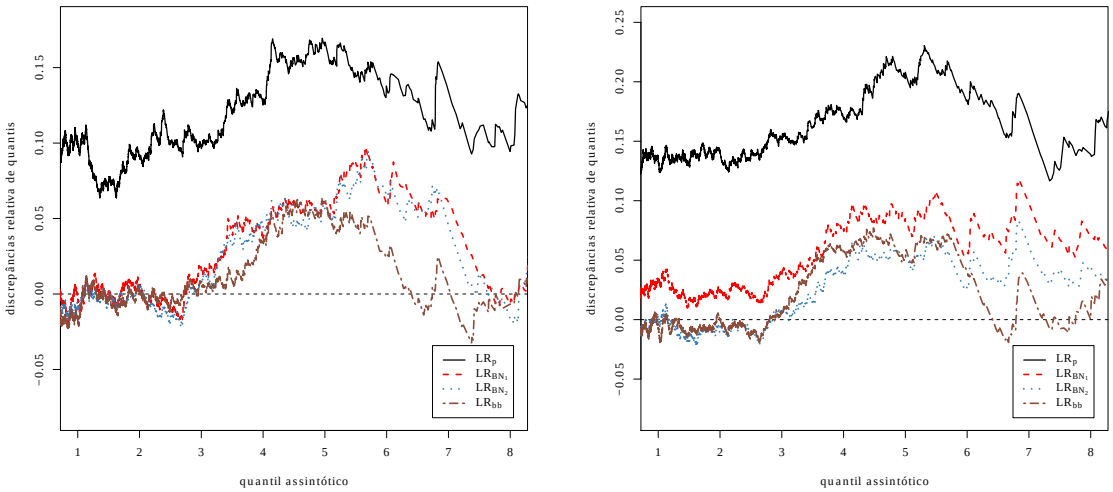
(a) 10% de censura

(b) 30% de censura



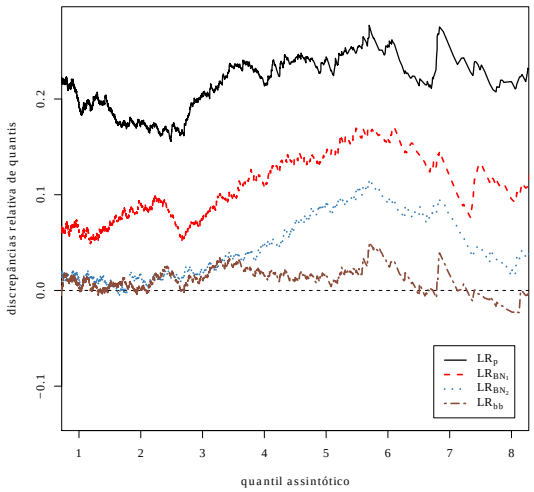
(c) 50% de censura

Figura 4.4: Distorção dos testes para $\xi = 10\%$ e $\lambda = 1,5$ sob censura do tipo II



(a) 10% de censura

(b) 30% de censura



(c) 50% de censura

Figura 4.5: Discrepância relativa de quantil para $n = 20$ e $\lambda = 1,5$ sob censura do tipo II

mente que os testes baseados nas estatísticas LR_{BN_1}, LR_{BN_2}, LR_b e LR_{bb} tiveram os melhores desempenhos, principalmente para amostras de tamanho pequeno a moderado. Além disso, à medida que a porcentagem de censura aumenta, os testes baseados em bootstrap apresentaram a menores distorções, de forma geral.

O gráfico de discrepância relativa de quantis sob censura do tipo II para $n = 20$ e diferentes porcentagens de censura é dado pela Figura 4.5. A inspeção desse gráfico sugere que as distribuições das estatísticas LR_{BN_2} e LR_{bb} são as que mais se aproximam da distribuição de referência. Essas estatísticas também foram as que sofreram um menor impacto com o aumento de observações censuradas na amostra.

A Tabela 4.8 mostra os resultados de simulação de Monte Carlo para a média e a variância das estatísticas de teste. Observamos nesta tabela que a média e a variância estimadas das estatísticas LR_{BN_1}, LR_{BN_2} e LR_{bb} são as que mais aproximam à média e a variância da distribuição χ^2_1. No entanto, para $n = 100$ os resultados sugerem que a estatística LR_p e suas versões modificadas são equivalentes.

As taxas de rejeições não nulas (poder) são dadas na Tabela 4.9. Os resultados foram obtidos ao considerar $n = 50$, os valores de λ que variam de $1,7$ a $4,5$, os níveis nominais de 10% e 5%. A hipótese nula a ser testada é $\mathcal{H}_0 : \lambda = 1,5$. Os resultados apresentados nessa tabela mostram que o teste baseado em LR_p é menos poderoso que os testes baseados em LR_{BN_1} e LR_{BN_2}. Porém, o teste corrigido bootstrap apresenta poder ligeiramente menor que LR_p. Observamos também que todos os testes se tornam menos poderosos à medida que a quantidade de observações censuradas aumenta. Por exemplo, para $\lambda = 3,0$ e o nível nominal de 10%, quando 10% da amostra é censurada, os poderes de LR_p, LR_{BN_1}, LR_{BN_2} e LR_{bb} são, respectivamente: $90,20$; $92,25$; $92,63$ e $89,36$. No entanto, quando metade das observações são censuradas as taxas de rejeições passam a ser: $72,95$; $79,34$; $78,11$ e $72,46$.

4.5 Aplicações

Nesta seção, mostraremos dois exemplos numéricos, um para amostra completa e outro, para amostra com dados censurados do tipo II. Para as duas aplicações assumimos que as observações são independentes e provenientes da distribuição BurrX (σ, λ). O interesse é testar a hipótese nula $\mathcal{H}_0 : \lambda = 1,0$ *versus* $\mathcal{H}_1 : \lambda \neq 1,0$, ou seja, estamos testando o modelo Rayleigh.

Primeiramente, vamos considerar o seguinte conjunto de dados não censurados: 38,4; 32,9; 48,5; 20,9; 11,6; 22,3; 30,2; 33,4; 26,7; 39,0; 12,8; 14,6; 12,2; 23,1; 29,4; 16,0; 20,1; 23,3; 22,9; 22,5; 15,1; 31,0; 16,9; 16,1; 10,8, o qual consiste em 25 observações referentes às taxas de fecundidade de um tipo de mosca. Essa taxa é calculada com base no número de ovos colocados por fêmea por dia pelos primeiros 14 dias de vida (HAND *et al.*, 1993, p. 17).

Tabela 4.8: Média e variância (var.) das estatísticas de teste sob censura tipo II ($\sigma = 30$)

				$\lambda = 1,5$			
n	$p(\%)$		χ^2_1	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{bb}
20	10	média	1,0	1,100	1,011	1,007	0,999
		var.	2,0	2,460	2,070	2,051	1,983
	30	média	1,0	1,148	1,038	1,007	1,006
		var.	2,0	2,680	2,223	2,093	2,071
	50	média	1,0	1,203	1,083	1,024	1,002
		var.	2,0	2,863	2,374	2,121	1,930
30	10	média	1,0	1,046	0,985	0,981	1,030
		var.	2,0	2,135	1,919	1,907	2,180
	30	média	1,0	1,084	1,009	0,989	1,001
		var.	2,0	2,317	2,055	1,974	1,986
	50	média	1,0	1,127	1,033	1,003	1,014
		var.	2,0	2,437	2,061	1,941	2,056
50	10	média	1,0	1,022	0,988	0,984	1,025
		var.	2,0	2,114	1,961	1,945	2,119
	30	média	1,0	1,040	1,001	0,988	2,119
		var.	2,0	2,167	1,989	1,939	2,039
	50	média	1,0	1,064	1,017	0,998	1,001
		var.	2,0	2,274	2,038	1,975	1,931
100	10	média	1,0	0,985	0,965	0,961	1,011
		var.	2,0	2,011	1,926	1,911	1,968
	30	média	1,0	1,002	0,975	0,968	1,004
		var.	2,0	2,037	1,927	1,900	2,003
	50	média	1,0	1,009	0,977	0,971	0,998
		var.	2,0	2,075	1,946	1,922	1,961

Tabela 4.9: Taxas de rejeições não nulas sob censura tipo II $(n = 50)$

		$\xi = 10\%$				$\xi = 5\%$			
$p(\%)$	λ	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{bb}	LR_p	LR_{BN_1}	LR_{BN_2}	LR_{bb}
10	1,7	14,12	17,24	17,68	12,97	7,61	10,10	10,61	7,22
	2,0	35,52	41,47	42,62	33,68	25,35	31,12	32,05	23,56
	2,5	73,03	77,47	78,34	70,79	63,37	69,00	70,17	60,37
	3,0	90,20	92,25	92,63	89,36	84,74	87,90	88,59	83,86
	3,5	96,58	97,39	97,52	96,62	94,61	95,83	96,01	94,25
	4,0	98,90	99,20	99,28	98,60	98,04	98,52	98,62	97,56
	4,5	99,57	99,68	99,71	99,61	99,27	99,44	99,46	99,07
30	1,7	12,77	16,22	16,54	12,40	6,38	9,72	10,07	6,92
	2,0	30,79	38,28	39,12	29,41	20,94	27,94	28,62	20,12
	2,5	64,66	71,43	72,42	63,26	53,73	61,71	62,90	53,07
	3,0	83,51	87,47	88,11	84,04	76,93	81,64	82,45	76,75
	3,5	93,31	95,21	95,54	92,99	89,61	92,24	92,75	89,25
	4,0	97,35	98,10	98,27	97,01	95,19	96,70	97,04	94,84
	4,5	98,66	99,07	99,17	98,68	97,69	98,35	98,54	97,55
50	1,7	11,20	15,06	13,70	11,07	5,85	9,44	8,45	6,22
	2,0	23,98	33,36	30,70	22,93	16,26	24,20	22,23	14,72
	2,5	53,14	62,49	59,98	52,12	43,19	53,44	51,10	41,42
	3,0	72,95	79,34	78,11	72,46	64,69	72,78	71,41	63,61
	3,5	85,48	89,63	88,83	85,73	79,78	85,20	84,20	79,43
	4,0	92,20	94,56	94,16	92,22	88,73	91,89	91,50	88,23
	4,5	95,39	96,82	96,54	95,85	93,12	95,07	94,83	92,96

As estatísticas de teste são $LR_p = 4,15$, $LR_{BN_1} = 3,62$; $LR_{BN_2} = 3,62$; $LR_{CR_a} = 3,27$ e $LR_{bb} = 4,07$ e os correspondentes p-valores são: $0,042$; $0,057$; $0,057$; $0,070$ e $0,044$. O teste baseado em LR_b tem p-valor igual a $0,041$. Assim, temos que os testes baseados nas estatísticas LR_p, LR_b e LR_{bb} rejeitam a hipótese nula ao nível de 5%. As estimativas de máxima verossimilhança, obtidas maximizando as funções (4.6), (4.7), (4.8) e (4.9) são: $\hat{\lambda}_p = 1,888$; $\hat{\tilde{\lambda}}_{BN} = 1,818$; $\hat{\mathring{\lambda}}_{BN} = 1,818$ e $\hat{\lambda}_{CR_a} = 1,770$, respectivamente.

Consideramos agora, um exemplo numérico usando um conjunto de dados apresentado por Mann e Fertig (1973), os quais correspondem a tempos de falhas de componentes de uma aeronave. A amostra contém observações censuradas do tipo II e foi obtida colocando 13 componentes sob teste de vida até ocorrer a décima falha. Dessa forma, os tempos de falhas observados (em horas) são: 0,22; 0,50; 0,88; 1,00; 1,32; 1,33; 1,54; 1,76; 2,50; 3,00. Esse mesmo conjunto de dados foi analisado por Ferrari *et al.* (2007) para fazer inferências sobre um dos parâmetros da distribuição Weibull. Aqui, obtemos $LR_p = 1,66$, $LR_{BN_1} = 2,98$ e $LR_{BN_2} = 2,96$ e os p-valores correspondentes iguais a: $0,198$; $0,084$ e $0,085$. Logo, ao nível nominal de 10%, o teste baseado em LR_p é o único que não rejeita o modelo Rayleigh. As estimativas de máxima verossimilhança para o parâmetro λ, obtidas maximizando as funções (4.11), (4.12) e (4.13), são respectivamente: $\hat{\lambda}_p = 0,633$; $\hat{\tilde{\lambda}}_{BN} = 0,506$ e $\hat{\mathring{\lambda}}_{BN} = 0,513$.

4.6 Conclusão

Neste capítulo, fizemos inferências (estimação pontual e teste de hipóteses) baseadas na função de verossimilhança perfilada para o parâmetro de forma da distribuição BurrX escalonada ao considerar dois enfoques: amostras completas (dados não censurados) e amostras com dados sob censura do tipo II. Derivamos três ajustes para a função de verossimilhança perfilada com base nas propostas de Fraser *et al.* (1999), Severini (1999) e Cox e Reid (1993) com o objetivo de atenuar o impacto do parâmetro de perturbação nas inferências. Obtemos e comparamos os estimadores de máxima verossimilhança perfilado e perfilado ajustados. Tanto para dados completos quanto para dados censurados, os resultados de simulação indicaram um melhor desempenho para os estimadores perfilado ajustados, principalmente para amostras de tamanho pequeno a moderado. Também comparamos os testes baseados na estatística da razão de verossimilhanças usando os diferentes ajustes. Incluímos na comparação o teste baseado na estatística da razão de verossimilhanças bootstrap Bartlett e o teste da razão de verossimilhanças bootstrap usual. De modo geral, a estatística da razão de verossimilhanças baseada na função de verossimilhança perfilada teve um desempenho inferior as demais em pequenas amostras. No entanto, em grandes amostras todas as estatísticas de teste se mostraram equivalentes. Concluímos ainda que no caso de amostras contendo censurados, em geral, as inferências tornam-se

menos precisas à medida que a quantidade de observações censuradas aumenta. No entanto, os testes baseados em bootstrap, sofreram menor impacto com o aumento da quantidade de observações censuradas. Finalmente, através de dois exemplos numéricos mostramos que a diferentes funções de verossimilhanças podem levar à conclusões distintas.

5 Distribuição transmutada Marshall-Olkin estendida Lomax

Resumo

A família transmutada de distribuições tem recebido bastante atenção nos últimos anos. Neste capítulo generalizamos a distribuição Marshall-Olkin estendida Lomax utilizando o mapeamento de transmutação de posto quadrático para obter a distribuição transmutada Marshall-Olkin estendida Lomax. Discutimos várias propriedades da nova distribuição, incluindo a função taxa de falha, os momentos ordinário e incompleto e função característica. Apresentamos um estudo de simulação para avaliar o desempenho dos estimadores de máxima verossimilhança dos parâmetros que indexam a distribuição proposta. Por fim, provamos empiricamente a flexibilidade do novo modelo através de uma aplicação para um conjundo de dados reais.

Palavras-chave: Distribuições generalizadas. Distribuição Lomax. Família Marshall-Olkin estendida. Família transmutada.

5.1 Introdução

As distribuições de probabilidade são utilizadas para modelagem de dados em diferentes áreas, tais como análise de sobrevivência, demografia, confiabilidade, atuária e outras. Com isso, tem crescido o interesse em construir novas distribuições que satisfaçam certas propriedades e que sejam adequadas para modelar dados de qualquer natureza.

Um dos métodos mais usados para construir novas distribuições é por meio da composição de distribuições já existentes, geralmente chamadas de classes de distribuições generalizada G (TAHIR; NADARAJAH, 2015). A principal razão para isso é que, em geral, as distribuições generalizadas são mais flexíveis que a distribuição de base G e, portanto, podem fornecer melhores ajustes para os dados (PESCIM *et al.*, 2010).

Algumas das famílias generalizadas mais conhecidas são: a Marshall-Olkin estendida (MOE) (MARSHALL; OLKIN, 1997), a exponencializada generalizada (exp-G) (CORDEIRO *et al.*, 2013; GUPTA *et al.*, 1998), a beta generalizada (beta-G) (EUGENE *et al.*, 2002), Kumaraswamy generalizada (Kw-G) (CORDEIRO; CASTRO, 2011), gama generalizada (gama-G)

(NADARAJAH *et al.*, 2015; RISTIĆ; BALAKRISHNAN, 2012; ZOGRAFOS; BALAKRISHNAN, 2009) e a McDonald generalizada (Mc-G) (ALEXANDER *et al.*, 2012). Uma descrição detalhada dessas famílias pode ser encontrada em Tahir e Nadarajah (2015).

Shaw e Buckley (2009) foram os pioneiros em usar o mapa de transmutação quadrático (QRTM) para adicionar um parâmetro a uma distribuição existente para obter mais flexibilidade. A classe gerada, denominada família transmutada, inclui a distribuição de base como caso especial e tem sido bastante utilizada para modelar dados de tempo de vida. Os resultados gerais para esta família e novos modelos são discutidos por Bourguignon *et al.* (2016).

Neste trabalho, consideramos a família transmutada generalizada (T-G) de distribuições para construir um novo modelo denominado de *Transmutada Marshall-Olkin estendida Lomax* (TMOEL), o qual é obtido tomando a distribuição Marshall-Olkin estendida Lomax (GHITANY *et al.*, 2007) como modelo de base na família T-G. Dado que a nova distribuição tem suporte nos reais positivos, o nosso objetivo é definir uma distribuição flexível para aplicações em tempo de vida.

O capítulo está organizado da seguinte forma: na Seção 5.2, definimos a distribuição TMOEL. Na Seção 5.3, discutimos as formas da função de densidade e função taxa de falha. Na Seção 5.4, obtemos uma expansão para a função densidade da TMOEL como uma combinação linear de densidades da distribuição Lomax exponencializada (exp-Lomax). Nas Seções 5.5-5.8, apresentamos as expressões explícitas para a função quantílica (fq), momentos completo e incompleto, função característica, desvios médios, curvas de Bonferroni e Lorenz, e estatísticas de ordem. Na Seção 5.9, descrevemos como estimar os parâmetros da distribuição TMOEL por máxima verossimilhança e, na Seção 5.10 realizamos um estudo de simulação para avaliar o comportamento desses estimadores. Na Seção 5.11, consideramos uma aplicação da distribuição TMOEL e comparamos com outras distribuições concorrentes, baseando-se nas estatísticas de bondade de ajuste. Finalmente, na Seção 5.12 sumarizamos as conclusões deste trabalho.

5.2 A nova distribuição

Uma variável aleatória absolutamente contínua segue *distribuição Lomax* (também conhecida como distribuição Pareto do tipo II) se sua função de distribuição acumulada (fda) é dada por

$$R(x) = 1 - (1 + \beta x)^{-\gamma}, \quad x, \beta, \gamma > 0, \tag{5.1}$$

em que γ e β são, respectivamente, parâmetros de forma e escala.

A distribuição *Marshall-Olkin estendida Lomax* (MOEL), proposta por Ghitany *et al.* (2007), foi obtida considerando a função de distribuição acumulada dada pela equação (5.1) como mo-

delo de base na família Marshall-Olkin estendida, cuja fda é dada por

$$G(x) = 1 - \frac{\alpha}{(1+\beta x)^{\gamma} - \bar{\alpha}}, \quad x, \alpha, \beta, \gamma > 0,\ \bar{\alpha} = 1 - \alpha. \tag{5.2}$$

Uma variável aleatória tem distribuição pertencente à família T-G se sua fda e a função densidade de probabilidade (fdp) são dadas, respectivamente, por

$$T(x) = T(x; \boldsymbol{\tau}, \lambda) = (1+\lambda)G(x; \boldsymbol{\tau}) - \lambda G(x; \boldsymbol{\tau})^2, \quad \lambda \in [-1, 1] \tag{5.3}$$

e

$$t(x) = t(x; \boldsymbol{\tau}, \lambda) = [1 + \lambda - 2\lambda G(x; \boldsymbol{\tau})] g(x; \boldsymbol{\tau}),\ x \in \mathcal{D} \subseteq \mathbb{R}, \tag{5.4}$$

em que $G(x; \boldsymbol{\tau})$, $g(x; \boldsymbol{\tau})$ e $\boldsymbol{\tau}$ são, respectivamente, a fda, a fdp e o vetor de parâmetros da distribuição de base. Observar que, para $\lambda = 0$, obtemos o modelo de base. Bourguignon *et al.* (2016) mostraram que a fdp (5.4) pode ser expressa como uma combinação linear de densidades de exp-G.

Baseando-se na família T-G e na distribuição MOEL, construímos uma nova distribuição indexada por quatro parâmetros, a distribuição TMOEL. Considerando (5.2) como modelo de base na equação (5.3), obtemos a fda da distribuição TMOEL (para $x > 0$), dada por

$$F(x) = \frac{[(1+\beta x)^{\gamma} - 1][(1+\beta x)^{\gamma} + \alpha\lambda - \bar{\alpha}]}{[(1+\beta x)^{\gamma} - \bar{\alpha}]^2}. \tag{5.5}$$

Sua fdp é expressa como

$$f(x) = \frac{\alpha\,\beta\,\gamma\,(1+\beta x)^{\gamma-1}\{[1-(1+\beta x)^{\gamma}](\lambda - 1) + \alpha(\lambda + 1)\}}{[(1+\beta x)^{\gamma} - \bar{\alpha}]^3}, \tag{5.6}$$

em que $\alpha > 0,\ \beta > 0,\ \gamma > 0$ e $\lambda \in [-1, 1]$. No que segue, uma variável aleatória X cuja fdp é dada por (5.6) será denotada por $X \sim \text{TMOEL}(\alpha, \beta, \gamma, \lambda)$.

Em análise de tempo de vida, uma função bastante útil é a função taxa de falha (hrf). A hrf de X é dada por

$$h(x) = \frac{\beta\gamma(1+\beta x)^{\gamma-1}\{(\lambda-1)[(1+\beta x)^{\gamma} - 1] - \alpha(1+\lambda)\}}{[(1+\beta x)^{\gamma} - \bar{\alpha}]\{(\lambda-1)[(1+\beta x)^{\gamma} - 1] - \alpha\}}. \tag{5.7}$$

Para valores selecionados dos parâmetros α, β, γ e λ, alguns sub-modelos da distribuição TMOEL publicados na literatura são listados na Tabela 5.1.

5.3 Formas da densidade e da função taxa de falha

A forma da fdp (5.6) pode ser descrita analiticamente examinando as raízes da equação $f'(x) = 0$ e analisando os seus limites quando $x \to 0$ ou $x \to \infty$. Como $f(x)$ é contínua, então $\lim_{x\to\infty} f(x) = 0$. Além disso, temos

$$\lim_{x\to 0} f(x) = \frac{\beta\,\gamma\,(\lambda+1)}{\alpha}$$

e, portanto, $\lim_{x\to 0} f(x) = 0$ se e somente se $\lambda = -1$. A Figura 5.1 mostra alguns gráficos da fdp da distribuição TMOEL para diferentes valores de parâmetros. Esses gráficos mostram que a fdp de X pode ser estritamente decrescente ou unimodal com o modo em

$$x_0 = \frac{1}{\beta}\left(\frac{c_{\alpha,\lambda} + 2\alpha\gamma\lambda - \{\alpha^2\lambda^2 + 4\alpha\gamma\lambda\, c_{\alpha,\lambda} + \gamma^2[2\alpha\,(\lambda-1) + (\lambda-1)^2 + \alpha^2(3\lambda^2+1)]\}^{1/2}}{(1+\gamma)(\lambda-1)}\right)^{\frac{1}{\gamma}} - \frac{1}{\beta},$$

tal que $c_{\alpha,\lambda} = -1 + \alpha + \lambda$.

A seguir, obtemos condições, em termos dos parâmetros, para o comportamento da fdp da distribuição TMOEL. A partir de (5.4), temos que a densidade $t(x)$ é decrescente quando $t'(x) = g'(x)[1 + \lambda - 2\,\lambda\,G(x)] - 2\,\lambda\,g^2(x) < 0$ para todo $x > 0$, tal que $G(x)$ e $g(x)$ são, respectivamente, a fda e a fdp da distribuição MOEL. Assim, $t(x)$ é decrescente quando

$$\frac{g'(x)}{g(x)}[1 + \lambda - 2\,\lambda\,G(x)] < 2\,\lambda\,g(x), \quad \forall x > 0.$$

Depois de algumas manipulações algébricas e considerando que $x \to 0^+$ (x se aproxima de zero pela direita), a inequação acima é equivalente a

$$\alpha + 2\,\gamma - \alpha\gamma + \alpha\lambda + 4\,\gamma\lambda - \alpha\gamma\lambda > 0.$$

Portanto, $t(x)$ é unimodal se e somente se

$$\alpha + 2\,\gamma - \alpha\gamma + \alpha\lambda + 4\,\gamma\lambda - \alpha\gamma\lambda < 0.$$

Logo, os parametros α, γ e λ controlam a forma da fdp, enquanto que β é um parâmetro de escala. Os parâmetros de forma permitem um controle extensivo na cauda direita, fornecendo caudas mais pesadas (leves) quando o parâmetro α decresce (cresce) e os parâmetros γ e λ crescem (decrescem).

A correspondente hrf pode ter formas tais como decrescente e unimodal, conforme é mostrado na Figura 5.2. Assim, a nova distribuição pode ser apropriada para diferentes aplicações de análise de tempo de vida. Para as condições sobre o comportamento da função $h(x)$, notamos

que $h'(x) = \frac{t'(x)[1-T(x)]+t^2(x)}{[1-T(x)]^2}$ e, dessa forma, $h(x)$ é decrescente se e somente se

$$t'(x)[1 - T(x)] + t^2(x) < 0, \quad \forall x > 0.$$

A partir de (5.4) e considerando que $x \to 0^+$, esta última inequação é equivalente a

$$\alpha(\gamma - 1)(\lambda + 1) + \gamma(\lambda^2 - 2\,\lambda - 1) < 0.$$

Assim, $h(x)$ é unimodal se e somente se

$$\alpha(\gamma - 1)(\lambda + 1) + \gamma(\lambda^2 - 2\,\lambda - 1) > 0.$$

Portanto, os parâmetros α, γ e λ também controlam a forma da função hrf de X.

Tabela 5.1: Sub-modelos da distribuição TMOEL: MOEL, TL (transmutada Lomax) e Lomax

α	β	γ	λ	Modelo	Referências
α	β	γ	0	$\mathrm{MOEL}(\alpha, \beta, \gamma)$	Ghitany *et al.* (2007)
1	β	γ	λ	$\mathrm{TL}(\beta, \gamma, \lambda)$	Ashour e Eltehiwy (2013)
1	β	γ	0	$\mathrm{Lomax}(\beta, \gamma)$	Lomax (1954)

5.4 Expansões úteis

A fda $F(x)$ de X pode ser escrita em termos da fda da distribuição Lomax, $R(x)$, como

$$F(x) = \frac{(1+\lambda)\,R(x)}{\alpha\left[1 + \left(\frac{1-\alpha}{\alpha}\right)R(x)\right]} - \frac{\lambda\,R^2(x)}{\alpha^2\left[1 + \left(\frac{1-\alpha}{\alpha}\right)R(x)\right]^2}. \tag{5.8}$$

Usando a expansão binomial generalizada, temos, para $\alpha > 1/2$,

$$\left[1 + \left(\frac{1-\alpha}{\alpha}\right)R(x)\right]^{-1} = \sum_{k=0}^{\infty}\left(\frac{\alpha-1}{\alpha}\right)^k R^k(x) \tag{5.9}$$

e

$$\left[1 + \left(\frac{1-\alpha}{\alpha}\right)R(x)\right]^{-2} = \sum_{k=0}^{\infty}(k+1)\left(\frac{\alpha-1}{\alpha}\right)^k R^k(x). \tag{5.10}$$

Substituindo (5.9) e (5.10) na equação (5.8) obtemos

$$F(x) = \sum_{k=0}^{\infty} a_k \, R^{k+1}(x), \tag{5.11}$$

em que $a_k = \frac{(\alpha-1)(1+\lambda)-k\lambda}{\alpha^2} \left(\frac{\alpha-1}{\alpha}\right)^{k-1}$.

Para obter uma expansão para $0 < \alpha \leq 1/2$, reescrevemos a equação (5.2) como

$$G(x) = \frac{R(x)}{1 - \bar{\alpha}\bar{R}(x)},$$

em que $\bar{R}(x) = 1 - R(x)$. Considerando a expansão

$$[1 - \bar{\alpha}\,\bar{R}(x)]^{-1} = \sum_{j=0}^{\infty} \bar{\alpha}^j \bar{R}^j(x)$$

e expandindo $\bar{R}^j(x)$ temos

$$G(x) = \sum_{j=0}^{\infty} \sum_{k=0}^{j} (-1)^k \, \bar{\alpha}^j \binom{j}{k} R^{k+1}(x).$$

Trocando $\sum_{j=0}^{\infty} \sum_{k=0}^{j}$ por $\sum_{k=0}^{\infty} \sum_{j=k}^{\infty}$ na última equação, obtemos

$$G(x) = \sum_{k=0}^{\infty} b_k \, R^{k+1}(x),$$

em que $b_k = \sum_{j=k}^{\infty} (-1)^k \, \bar{\alpha}^j \binom{j}{k}$.

Substituindo a última equação em (5.3), temos

$$F(x) = \sum_{k=0}^{\infty} [(1+\lambda)\, b_k - \lambda \, c_k] R^{k+1}(x), \tag{5.12}$$

tal que $c_0 = b_0^2$ e $c_m = \frac{1}{m\, b_0} \sum_{i=1}^{m} (3i - m)\, c_{m-i} b_i$ para $m \geq 1$.

Finalmente, a partir das equações (5.11) e (5.12), podemos escrever, para todo $\alpha > 0$,

$$F(x) = \sum_{k=0}^{\infty} p_k \, H_{k+1}(x), \tag{5.13}$$

em que $H_{k+1}(x)$ é a fda da exp-Lomax com parâmetro de potência $k+1$ e

$$p_k = \begin{cases} (1+\lambda)\, b_k - \lambda\, c_k, & \text{se } 0 < \alpha \leq 1/2, \\ a_k, & \text{se } \alpha > 1/2. \end{cases}$$

Diferenciando (5.13) em relação a x, obtemos

$$f(x) = \sum_{k=0}^{\infty} p_k\, h_{k+1}(x), \tag{5.14}$$

em que $h_{k+1}(x)$ é a densidade da exp-Lomax com parâmetro de potência $k+1$.

A equação (5.14) mostra que a fdp de X pode ser expressa como uma combinação linear de densidades de exp-Lomax. Assim, algumas propriedades estruturais da distribuição TMOEL podem ser obtidas a partir das propriedades da distribuição exp-Lomax (SALEM, 2014).

5.5 Função quantílica

Visto que a fda $F(x)$ dada em (5.5) é contínua e estritamente crescente, a fq de X é $Q(u) = F^{-1}(u)$, para $0 < u < 1$. A partir dos resultados dados em Bourguignon *et al.* (2016), obtemos a fq da distribuição TMOEL como

$$Q(u;\boldsymbol{\tau},\lambda) = \begin{cases} Q_G\left(\frac{1+\lambda-\sqrt{(1+\lambda)^2-4\lambda u}}{2\lambda};\boldsymbol{\tau}\right), & \text{se } \lambda \neq 0, \\ Q_G(u;\boldsymbol{\tau}), & \text{se } \lambda = 0, \end{cases} \tag{5.15}$$

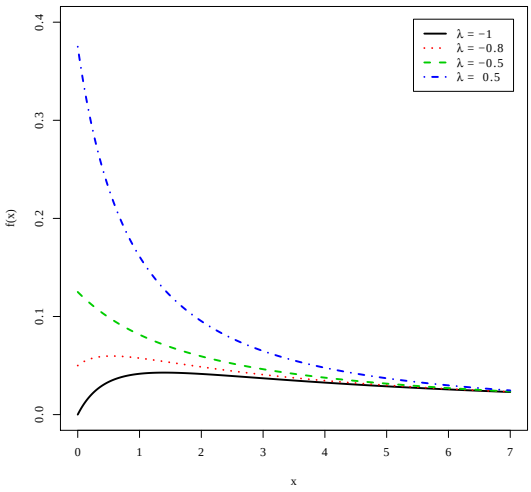
em que $\boldsymbol{\tau} = (\alpha,\beta,\gamma)^\top$ é o vetor de parâmetros e $Q_G(u;\boldsymbol{\tau})$ é a fq do modelo MOEL, a qual é dada por

$$Q_G(u;\boldsymbol{\tau}) = \beta^{-1}\left[\left(1 - \frac{\alpha u}{u-1}\right)^{1/\gamma} - 1\right].$$

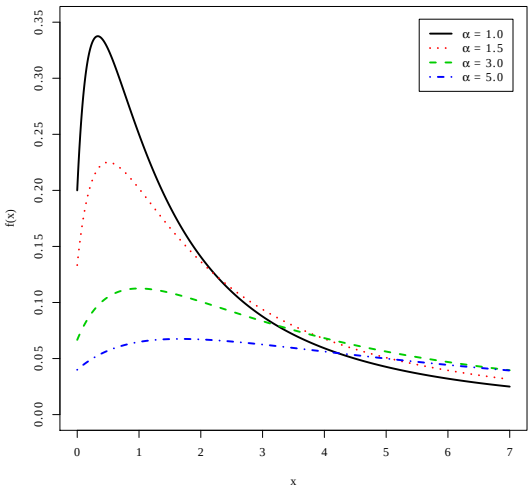
Usando (5.15), podemos gerar números aleatórios da distribuição TMOEL da seguinte forma. Se $U \sim \mathcal{U}(0,1)$, então

$$Q(U;\alpha,\beta,\gamma,\lambda) \sim \text{TMOEL}\,(\alpha,\beta,\gamma,\lambda).$$

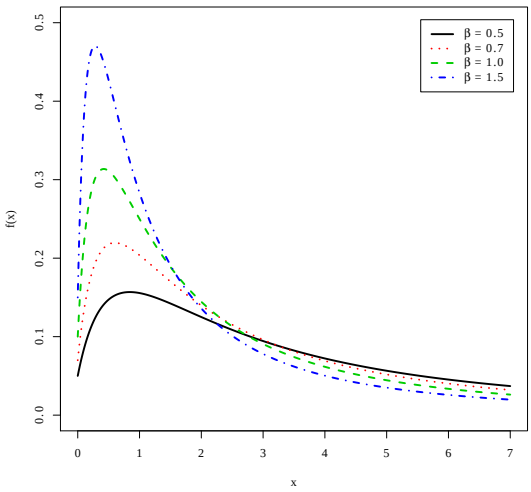
Kozubowski e Podgórski (2016) apresentaram uma forma alternativa de gerar números aleatórios de distribuições transmutadas utilizando extremos aleatórios. Seguindo as idéias desse trabalho, definindo X_1 e X_2 variáveis aleatórias que seguem a distribuição MOEL e N_p uma



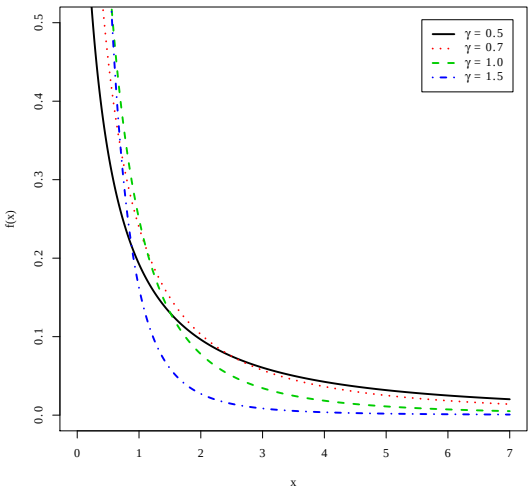
(a) $\alpha = 2,0$; $\beta = 1,0$ e $\gamma = 0,5$



(b) $\beta = \gamma = 1,0$ e $\lambda = -0,8$

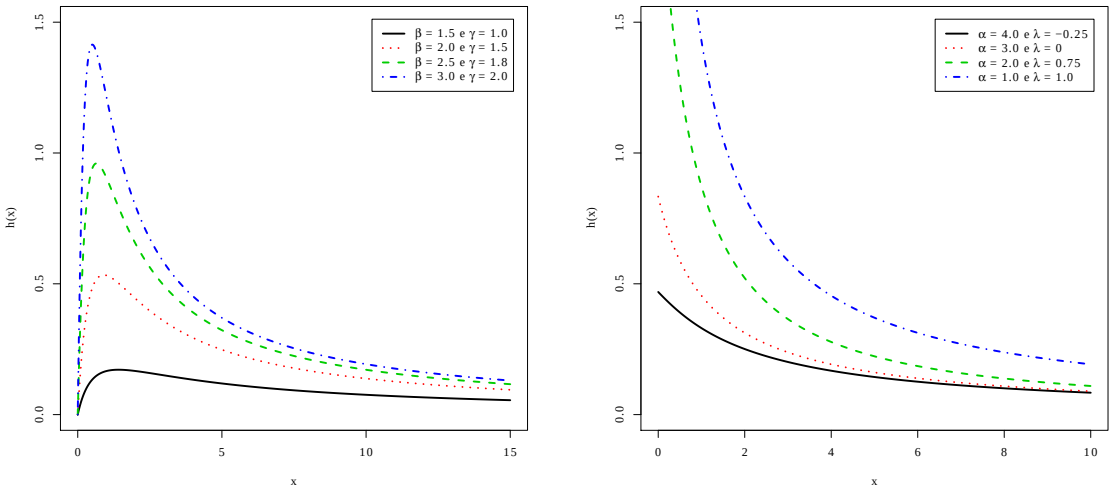


(c) $\alpha = \gamma = 1,0$ e $\lambda = -0,9$

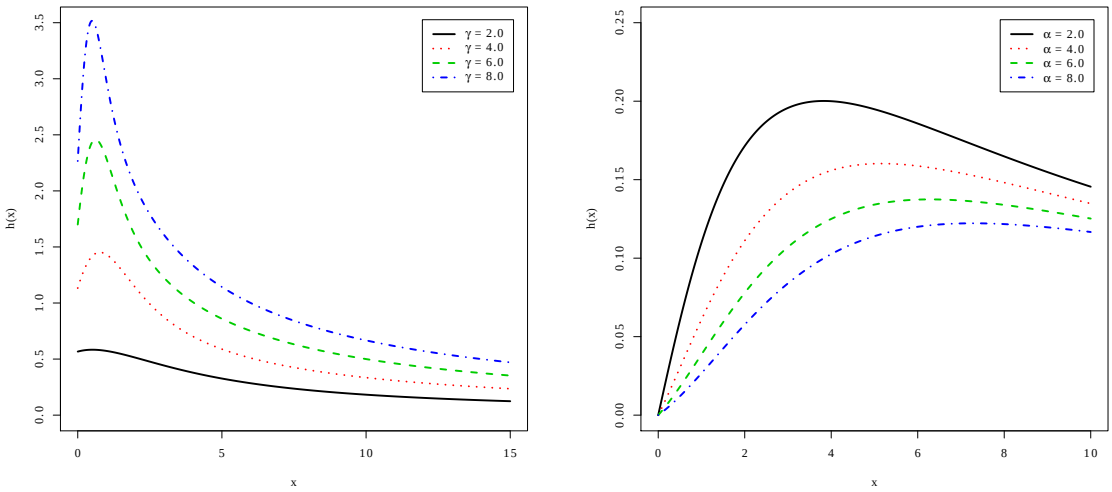


(d) $\alpha = \beta = 4,0$ e $\lambda = 0,9$

Figura 5.1: Fdp da distribuição TMOEL para alguns valores dos parâmetros



(a) $\alpha = 3,0$ e $\lambda = -1,0$ (b) $\beta = 2,5$ e $\gamma = 1,0$



(c) $\alpha = 3,0$; $\beta = 0,5$ e $\lambda = 0,7$ (d) $\beta = 0,5$; $\gamma = 2,0$ e $\lambda = -1,0$

Figura 5.2: Hrf da distribuição TMOEL para alguns valores dos parâmetros

variável aleatória que assume valores inteiros tal que $N_p - 1 \sim \text{Bernoulli}(p)$, $0 \leq p \leq 1$. Agora, supondo que N_p e X_i, $i = 1, 2$, são variáveis aleatórias independentes, então uma observação y que segue distribuição TMOEL pode ser gerada da seguinte forma:

Gerando variáveis aleatórias $y_1, \ldots, y_n$ da distribuição TMOEL:

1 Gerar u_1 e u_2 independentes a partir da distribuição $\mathcal{U}(0, 1)$.

2 Calcular $x_i = Q_G(u_i)$, $i = 1, 2$.

3 I) Se $\lambda \in [-1, 0]$, defina $p = -\lambda \in [0, 1]$ e gerar $n_p - 1$ a partir da distribuição de $\text{Bernoulli}(p)$.

4 Finalmente, obtenha $y = \vee_{i=1}^{n_p} x_i$, em que $\vee$ denota o valor mínimo.

5 II) Se $\lambda \in [0, 1]$, defina $p = \lambda$ e gerar $n_p - 1$ a partir da distribuição de $\text{Bernoulli}(p)$.

6 Finalmente, obtenha $y = \wedge_{i=1}^{n_p} x_i$, em que $\wedge$ denota o valor máximo.

É importante notar que quando $\lambda = 0$ na expressão (5.3), temos que $n_p = 1$ quase certamente (q.c.) e os passos 4 e 6 são satisfeitos simultaneamente e os lados direitos dessas igualdades se reduzem a x_1. Enquanto que nos casos extremos ($\lambda = \pm 1$), temos $n_p = 2$ q.c. e os passos 4 e 6 são reduzidos a $y = \max(x_1, x_2)$ e $y = \min(x_1, x_2)$, respectivamente.

Na Figura 5.3, comparamos a densidade exata da TMOEL e os histogramas construídos a partir de dois conjuntos de dados simulados utilizando o algoritmo descrito acima para valores de parâmetros especificados. Para simular os dados, usamos a plataforma computacional `R` (versão 3.2.3). Com isso, ilustramos a consistência dos valores simulados a partir do algoritmo.

5.6 Assimetria e curtose

As medidas de assimetria e curtose são determinadas por $\alpha_3 = \mu_3/\sigma^3$ e $\alpha_4 = \mu_4/\sigma^4$, respectivamente, em que μ_j é o j-ésimo momento central e σ é o desvio padrão. Para algumas distribuições da família T-G, a obtenção do terceiro e quarto momentos pode ser inviável. Medidas alternativas para a assimetria e curtose baseadas nos quantis às vezes são mais apropriadas. A medida de assimetria S de Bowley e a medida de curtose de Moors K são, respectivamente, definidas por

$$S = \frac{Q(6/8) + Q(2/8) - 2\,Q(4/8)}{Q(6/8) - Q(2/8)}$$

e

$$K = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)},$$

em que $Q(\cdot)$ é a função quantílica dada por (5.15). É importante ressaltar que as medidas S e K existem mesmo para as distribuições sem momentos.

Na Figura 5.4 mostramos a assimetria e a curtose como função do parâmetro λ para alguns valores de parâmetros. Analisando esses gráficos, notamos que a assimetria e a curtose de X decrescem rapidamente quando λ converge para um.

5.7 Momentos e função característica

Momentos são importantes em qualquer análise estatística. Por exemplo, algumas características de uma distribuição podem ser descritas usando medidas tais como média, variância, assimetria e curtose, as quais são determinadas utilizando os quatro primeiros momentos ordinários.

Para $r \in \mathbb{N}$, seja $\mu'_r = E(X^r)$ o r-ésimo momento ordinário de X. A partir da equação (5.14), podemos expressar μ'_r como uma combinação linear do r-ésimo momento ordinário de variáveis aleatórias exp-Lomax. De fato, para $r < \gamma$, obtemos

$$\mu'_r = \sum_{k=0}^{\infty} p_k \, E(Y_k^r), \tag{5.16}$$

em que $Y_k \sim \exp-\mathrm{Lomax}(k+1, \beta, \gamma)$.

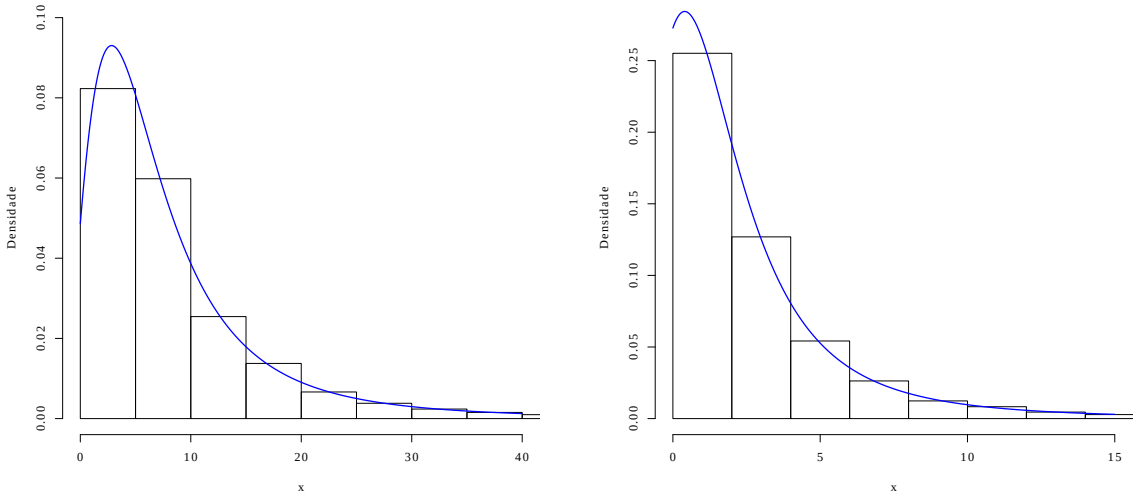
O r-ésimo momento ordinário de Y_k é dado em Salem (2014) (para $r < \gamma$) como

$$E(Y_k^r) = \frac{(k+1)}{\beta^r} \sum_{j=0}^{r} (-1)^j \binom{r}{j} B\left(1 - \frac{1}{\gamma}(r-j), k+1\right), \tag{5.17}$$

tal que $B(a,b) = \Gamma(a)\,\Gamma(b)/\Gamma(a+b)$ é a função beta e $\Gamma(\cdot)$ é a função gama. Substituindo (5.17) na equação (5.16), temos (para $r < \gamma$)

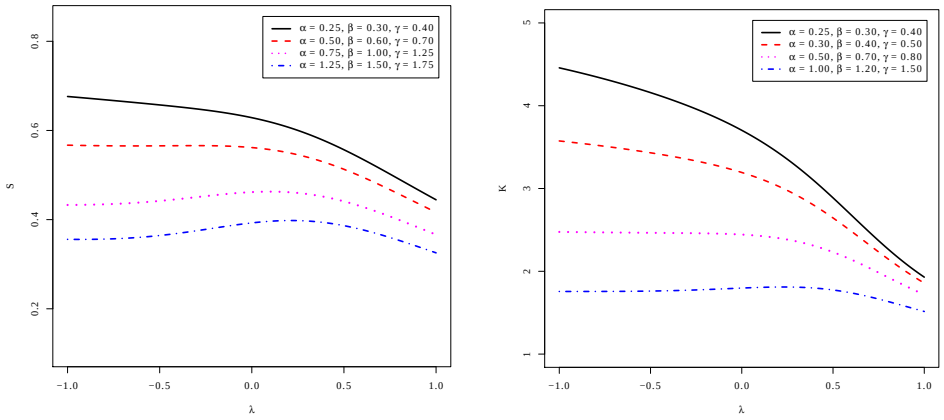
$$\mu'_r = \sum_{k=0}^{\infty} \sum_{j=0}^{r} b_r(k,j) \, B\left(1 - \frac{1}{\gamma}(r-j), k+1\right), \tag{5.18}$$

em que $b_r(k,j) = (-1)^j \binom{r}{j} \frac{(k+1)}{\beta^r} \, p_k$. A partir da equação (5.18), notamos que $\mu'_r < \infty$ para $r < \gamma$, uma condição que também se aplica à distribuição Lomax.



(a) $\alpha = 2,5$; $\beta = 0,15$; $\gamma = 2,7$ e $\lambda = -0,7$ (b) $\alpha = 5,5$; $\beta = 0,5$; $\gamma = 2,5$ e $\lambda = 0,2$

Figura 5.3: Gráficos das densidades exatas da distribuição TMOEL e histogramas dos dados simulados.



(a) Assimetria (b) Curtose

Figura 5.4: Assimetria e curtose para alguns valores dos parâmetros

5.7.1 *Resultados numéricos*

Nesta seção, apresentamos e comparamos os resultados numéricos para média, variância, assimetria e curtose de uma variável aleatória $X \sim$ TMOEL $(\alpha, \beta, \gamma, \lambda)$, obtidos a partir de três métodos: truncamento, integração numérica e Monte Carlo (50 000 réplicas). Para o método do truncamento utilizamos o r-ésimo momento ordinário truncado de X, o qual é obtido a partir da equação (5.18) como

$$\mu'_{r,N} = \sum_{k=0}^{N} \sum_{j=0}^{r} b_r(k,j)\, B\left(1 - \frac{1}{\gamma}(r-j), k+1\right), \quad r < \gamma,\ N \in \mathbb{N}.$$

Utilizamos a plataforma computacional Ox para a obtenção dos resultados referentes aos métodos do truncamento e Monte Carlo. Enquanto que, para os resultados obtidos por integração numérica, usamos os algoritmos do "software" Mathematica, que dividem recursivamente a região de integração.

Para o método do truncamento, consideramos os valores de $N = 5,\ 10$ e 20. Os resultados deste estudo são dados na Tabela 5.2. Ao utilizar os resultados referentes à integração numérica como parâmetro de precisão, notamos que os valores correspondentes ao truncamento são mais precisos quando N cresce, o que é esperado devido ao fato de que $\mu'_r = \lim_{N\to\infty} \mu'_{r,N}$. Para $N = 20$ os resultados são suficientemente precisos quando comparados com os resultados obtidos através de integração numérica. Além disso, observar que para o método de Monte Carlo, os resultados são menos precisos para as medidas de assimetria e curtose.

5.7.2 *Momento incompleto*

O r-ésimo momento incompleto de X é determinado a partir da expressão (5.14) como

$$m_X^{(r)}(z) = \int_0^z x^r f(x)\, dx = \sum_{k=0}^{\infty} p_k\, m_{Y_k}^{(r)}(z). \tag{5.19}$$

Além disso, temos

$$m_{Y_k}^{(r)}(z) = \int_0^z x^r\, h_{k+1}(x)\, dx = \beta\,\gamma\,(k+1) \int_0^z x^r \left[1 - (1+\beta x)^{-\gamma}\right]^k (1+\beta x)^{-\gamma-1}\, dx.$$

Considerando a seguinte expansão binomial

$$\left[1 - (1+\beta x)^{-\gamma}\right]^k = \sum_{j=0}^{k} (-1)^j \binom{k}{j} (1+\beta x)^{-j\gamma}. \tag{5.20}$$

e substituindo-a na equação (5.19), obtemos

$$m_X^{(r)}(z) = \sum_{j=0}^{\infty} c(k,j)\, m_{V_j}^{(r)}(z), \tag{5.21}$$

em que $V_j \sim \text{Lomax}(\beta, (j+1)\gamma)$, $c(k,j) = \sum_{k=j}^{\infty} (-1)^j \binom{k}{j} \frac{(k+1)}{(j+1)}\, p_k$,

$$m_{V_j}^{(r)}(z) = \frac{(j+1)\beta\,\gamma\, z^{r+1}\, {}_2F_1(r+1, 1+(j+1)\gamma; r+2; -\beta z)}{r+1},$$

e ${}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; x)$ é a função hipergeométrica (GRADSHTEYN; RYZHIK, 2007).

5.7.3 *Desvios médios e curvas de Bonferroni e Lorenz*

Os desvios médios são usados frequentemente como medidas de dispersão. Os desvios médios de X sobre a média (δ_1) e a mediana (δ_2), são dados, respectivamente, por

$$\delta_1 = 2\,\mu_1' \,F(\mu_1') - 2\, m_X^{(1)}(\mu_1'), \quad \delta_2 = \mu_1' - 2\, m_X^{(1)}(M),$$

em que $M = Q(0,5)$, a mediana de X, pode ser obtida a partir da equação (5.15) avaliada em $u = 0,5$ e as quantidades μ_1' e $m_X^{(1)}(z)$ podem ser determinadas a partir de (5.18) e (5.21), respectivamente.

As curvas de Bonferroni e Lorenz, usadas em várias áreas tais como economia, confiabilidade, demografia e medicina, são definidas por $B(\pi) = m_X^{(1)}(q)/(\pi\mu_1')$ e $L(\pi) = m_X^{(1)}(q)/\mu_1'$, respectivamente, em que $q = Q(\pi)$ é calculada a partir de (5.15) para $0 < \pi < 1$.

5.7.4 *Função característica*

As funções geradora de momentos e característica são ferramentas muito úteis. Para uma variável aleatória que segue a distribuição Lomax, a função geradora de momentos é definida somente para $t \leq 0$. Consequentemente, a função geradora de $X \sim \text{TMOEL}(\alpha, \beta, \gamma, \lambda)$ também é definida somente nestes casos. Contudo, a função característica de uma distribuição existe para todo $t \in \mathbb{R}$. Sendo assim, a função característica de X é dada por

$$\phi_X(t) = E(\mathrm{e}^{itX}) = \int_0^{\infty} \mathrm{e}^{itx} f(x)\, dx,$$

em que $i = \sqrt{-1}$ e $t \in \mathbb{R}$.

Tabela 5.2: Resultados para a média, variância, assimetria e curtose

Parâmetros	**Método**	N	**média**	**variância**	**assimetria**	**curtose**
$\alpha = 0,75$		5	$0,7356$	$0,7407$	$5,1369$	$134,01$
$\beta = 0,5$	Truncamento	10	$0,6960$	$0,6916$	$5,2047$	$138,12$
$\gamma = 4,5$		20	$0,6963$	$0,6923$	$5,2040$	$138,06$
$\lambda = -0,75$	Int. Num.		$0,6963$	$0,6923$	$5,2040$	$138,07$
	Monte Carlo		$0,6949$	$0,7069$	$5,7817$	$99,360$
$\alpha = 1,5$		5	$0,3861$	$0,1688$	$3,8940$	$55,140$
$\beta = 1,0$	Truncamento	10	$0,3933$	$0,1715$	$3,8878$	$54,934$
$\gamma = 5,0$		20	$0,3934$	$0,1715$	$3,8877$	$54,930$
$\lambda = -0,5$	Int. Num.		$0,3934$	$0,1715$	$3,8877$	$54,930$
	Monte Carlo		$0,3927$	$0,1735$	$4,3059$	$56,020$
$\alpha = 3,0$		5	$0,1621$	$0,0344$	$3,1759$	$32,507$
$\beta = 2,0$	Truncamento	10	$0,2060$	$0,0402$	$3,0934$	$31,455$
$\gamma = 5,5$		20	$0,2174$	$0,0420$	$3,0889$	$31,254$
$\lambda = -0,25$	Int. Num.		$0,2178$	$0,0421$	$3,0887$	$31,238$
	Monte Carlo		$0,2175$	$0,0424$	$3,3848$	$35,354$
$\alpha = 5,0$		5	$0,0956$	$0,0117$	$2,8857$	$24,900$
$\beta = 3,0$	Truncamento	10	$0,1239$	$0,0137$	$2,7908$	$23,949$
$\gamma = 6,0$		20	$0,1314$	$0,0143$	$2,7898$	$23,838$
$\lambda = 0,25$	Int. Num.		$0,1304$	$0,0142$	$2,7891$	$23,886$
	Monte Carlo		$0,1302$	$0,0142$	$3,0287$	$27,900$
$\alpha = 7,0$		5	$0,0587$	$0,0043$	$2,4756$	$17,285$
$\beta = 4,0$	Truncamento	10	$0,0820$	$0,0051$	$2,3795$	$16,737$
$\gamma = 7,0$		20	$0,0888$	$0,0053$	$2,4155$	$16,961$
$\lambda = 0,5$	Int. Num.		$0,0828$	$0,0048$	$2,3729$	$17,031$
	Monte Carlo		$0,0827$	$0,0048$	$2,5454$	$20,037$

Usando a equação (5.14), temos

$$\phi_X(t) = \sum_{k=0}^{\infty} p_k \, \phi_{Y_k}(t),$$

em que ϕ_{Y_k} é a função característica de Y_k. Agora, usando a equação (5.20), podemos escrever

$$\phi_X(t) = \sum_{j=0}^{\infty} c(k,j) \, \phi_{V_j}(t), \tag{5.22}$$

em que

$$\phi_{V_j}(t) = (j+1)\,\gamma\,\mathrm{e}^{-it/\beta}(-it)^{(j+1)\gamma}\beta^{-(j+1)\gamma}\,\Gamma\left(-(j+1)\gamma, -\frac{it}{\beta}\right) \tag{5.23}$$

e $\Gamma(a,z) = \int_z^{\infty} t^{a-1}\mathrm{e}^{-t}dt$ é a função gama superior incompleta. Substituindo (5.23) na equação (5.22) e definindo

$$d(k,j;t) = c(k,j)\,(j+1)\,\gamma\,\mathrm{e}^{-it/\beta}(-it)^{(j+1)\gamma}\beta^{-(j+1)\gamma},$$

obtemos

$$\phi_X(t) = \sum_{j=0}^{\infty} d(k,j;t)\,\Gamma\left(-(j+1)\,\gamma, -\frac{it}{\beta}\right). \tag{5.24}$$

5.8 Estatísticas de ordem

Seja $X_1, \ldots, X_n$ uma amostra aleatória de tamanho n da distribuição $F(x)$. Então, para $1 \leq m \leq n$, a fdp da m-ésima estatística de ordem, $X_{(m)}$, pode ser expressa como (SEVERINI, 2005, p. 218)

$$f_{(m)}(x) = M\,F(x)^{m-1}(1-F(x))^{n-m}f(x) = M\,f(x)\sum_{j=0}^{n-m}(-1)^j\binom{n-m}{j}F(x)^{m+j-1},$$

em que $M = n!/[(m-1)!\,(n-m)!]$.

Baseando-se na equação (5.13) e usando a expansão para séries de potência definida para um número inteiro positivo (GRADSHTEYN; RYZHIK, 2007), temos

$$F(x)^{m+j-1} = \left(\sum_{k=0}^{\infty} p_k\,H_{k+1}(x)\right)^{m+j-1} = \sum_{k=0}^{\infty} c_{m+j-1,k}\,H_{k+1}(x),$$

em que $c_{m+j-1,0} = p_0^{m+j-1}$ e, para $i \geq 1$,

$$c_{m+j-1,i} = \frac{1}{i\,p_0} \sum_{s=1}^{i} [(m+j)s - i]\, p_s\, c_{m+j-1,i-s}.$$

Assim, obtemos

$$f_{(m)}(x) = Mf(x) \sum_{j=0}^{n-m} \sum_{k=0}^{\infty} (-1)^j \binom{n-m}{j} c_{m+j-1,k}\, H_{k+1}(x).$$

Substituindo $f(x)$ na equação acima pela expansão (5.14), depois de algumas manipulações algébricas, escrevemos

$$f_{(m)}(x) = M \sum_{j=0}^{n-m} \sum_{k,\ell=0}^{\infty} (-1)^j \binom{n-m}{j} \frac{(\ell+1)}{(\ell+k+2)}\, c_{m+j-1,k}\, p_\ell\, h_{\ell+k+2}(x).$$

Dessa forma, nota-se que $f_{(m)}(x)$ é uma combinação linear de densidades exp-Lomax.

5.9 Estimação

Nesta seção, discutiremos sobre a estimação dos parâmetros da distribuição TMOEL pelo método de máxima verossimilhança. Para tanto, seja $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ uma amostra aleatória de tamanho n de $X \sim \text{TMOEL}(\alpha, \beta, \gamma, \lambda)$ e $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \lambda)^\top$ o vetor de parâmetros. O logaritmo da função de verossimilhança de $\boldsymbol{\theta}$, denotado por $\ell(\boldsymbol{\theta})$, é dado por

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= n[\log(\alpha) + \log(\beta) + \log(\gamma)] + (\gamma - 1) \sum_{i=1}^{n} \log(1 + \beta x_i) \\ &+ \sum_{i=1}^{n} \log[u_i(\lambda - 1) + \alpha(\lambda + 1)] - 3 \sum_{i=1}^{n} \log(\alpha - u_i), \end{aligned} \tag{5.25}$$

em que $\log(a) = \log_{\text{e}} a = \ln(a)$ e $u_i = 1 - (1 + \beta x_i)^\gamma$.

O estimador de máxima verossimilhança (EMV) de $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$, pode ser obtido maximizando diretamente (5.25) usando as plataformas computacionais `SAS` (`PROC NLMIXED`), `R` (funções `optim` ou `MaxLik`) e `Ox` (sub-rotina `MaxBFGS`).

Os componentes do vetor escore $\ell_{\boldsymbol{\theta}} = (\ell_\alpha, \ell_\beta, \ell_\gamma, \ell_\lambda)^\top$ são dados por

$$\ell_\alpha = \frac{\partial\, \ell(\boldsymbol{\theta})}{\partial \alpha} = \frac{n}{\alpha} + 3 \sum_{i=1}^{n} (u_i - \alpha)^{-1} + (\lambda + 1) \sum_{i=1}^{n} [u_i(\lambda - 1) + \alpha(\lambda + 1)]^{-1},$$

$$\ell_\beta = \frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\beta} = \frac{n}{\beta} + (\gamma-1)\sum_{i=1}^{n}\frac{x_i}{1+\beta x_i} + 3\sum_{i=1}^{n}\frac{\gamma x_i(1+\beta x_i)^{\gamma-1}}{u_i-\alpha} - \sum_{i=1}^{n}\frac{\gamma x_i(\lambda-1)(1+\beta x_i)^{\gamma-1}}{u_i(\lambda-1)+\alpha(\lambda+1)},$$

$$\ell_\gamma = \frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\gamma} = \frac{n}{\alpha} + \sum_{i=1}^{n}\log(1+\beta x_i) + 3\sum_{i=1}^{n}\frac{(1+\beta x_i)^{\gamma}\log(1+\beta x_i)}{u_i-\alpha} + (\lambda-1)\sum_{i=1}^{n}\frac{(1+\beta x_i)^{\gamma}\log(1+\beta x_i)}{u_i(1-\lambda)-\alpha(1+\lambda)},$$

$$\ell_\lambda = \frac{\partial\,\ell(\boldsymbol{\theta})}{\partial\lambda} = \sum_{i=1}^{n}\frac{\alpha-u_i}{u_i(\lambda-1)+\alpha(1+\lambda)}.$$

Alternativamente, o EMV $\hat{\boldsymbol{\theta}}$ pode ser determinado resolvendo numericamente a equação não-linear $\ell_{\boldsymbol{\theta}} = \mathbf{0}$ (NOCEDAL; WRIGHT, 2006). Neste caso, essa equação pode ser avaliada numericamente usando o algoritmo de Newton-Raphson.

Sob as condições gerais de regularidade, temos $(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) \overset{a}{\sim} N_4(0, K(\boldsymbol{\theta})^{-1})$, em que $K(\boldsymbol{\theta})$ é a matriz de informação esperada de dimensão 4×4 e a notação $\overset{a}{\sim}$ denota distribuição assintótica. Para n grande, $K(\boldsymbol{\theta})$ pode ser aproximado pela matriz de informação observada. A aproximação normal para o EMV $\hat{\boldsymbol{\theta}}$ pode ser usada para construir intervalos de confiança aproximados e testar hipóteses sobre os parâmetros α, β, γ e λ.

Em algumas situações é de interesse fazer inferência sobre uma parte do vetor de parâmetros que indexam o modelo assumindo que os demais parâmetros são conhecidos. Considere o EMV do parâmetro λ assumindo que os parâmetros α, β e γ são conhecidos, temos

$$\ell_{\lambda\lambda} = -\frac{n}{\lambda^2} - \sum_{i=1}^{n}\frac{(u_i+\alpha)^2}{[u_i(\lambda-1)+\alpha(\lambda+1)]^2} < 0.$$

Além disso,

$$\lim_{\lambda\to-1}\ell_\lambda = \sum_{i=1}^{n}\frac{-\alpha-3\,u_i}{2\,u_i} < 0 \quad \text{e} \quad \lim_{\lambda\to1}\ell_\lambda = \sum_{i=1}^{n}\frac{u_i+3\,\alpha}{2\,\alpha} > 0.$$

Com isso, mostramos que neste caso especial o EMV $\hat{\lambda}$ existe e é único.

5.10 Estudo de simulação

Nesta seção, realizamos simulações de Monte Carlo para avaliar o comportamento do EMV de $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\lambda})$, estimando as medidas de viés relativo (VR) e erro quadrático médio (EQM) para amostras de tamanhos $n = \{100, 200, 250\}$.

Para esse estudo, consideramos $10\,000$ réplicas de Monte Carlo e usamos o método BFGS implementado na plataforma computacional `Ox` (versão 7.10, função `MaxBFGS`) para maximizar o logaritmo da função de verossimilhança (5.25). Fixamos os valores dos parâmetros β e γ em $0,25$ e $0,3$, respectivamente e variamos os parâmetros α e λ.

Os resultados, dados na Tabela 5.3, mostram que, geralmente, os valores dos vieses relativos e os EQMs diminuem quando n cresce. Os menores valores absolutos para o viés e o EQM são iguais a $0,001$ e $0,003$, respectivamente. Enquanto que o máximo valor absoluto para o viés e o EQM são $1,632$ e $4,467$, respectivamente. Além disso, notamos na Tabela 5.3 que o parâmetro λ foi subestimado na maioria dos casos (possui vieses relativos negativos).

Tabela 5.3: Valores para o viés e EQM do EMV de $\boldsymbol{\theta}$ ($\beta = 0,25$ e $\gamma = 0,3$)

			VR				EQM			
λ	α	n	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$
$-0,5$	$0,25$	100	$0,954$	$0,975$	$0,057$	$-0,086$	$0,687$	$0,932$	$0,012$	$0,142$
		200	$0,549$	$0,696$	$0,045$	$-0,049$	$0,144$	$0,337$	$0,007$	$0,124$
		250	$0,471$	$0,632$	$0,040$	$-0,049$	$0,111$	$0,301$	$0,006$	$0,117$
	$0,5$	100	$0,742$	$0,884$	$0,021$	$-0,158$	$2,155$	$1,025$	$0,007$	$0,153$
		200	$0,403$	$0,695$	$0,020$	$-0,073$	$0,521$	$0,290$	$0,004$	$0,139$
		250	$0,348$	$0,671$	$0,020$	$-0,059$	$0,225$	$0,279$	$0,003$	$0,133$
$0,25$	$0,25$	100	$1,245$	$1,632$	$0,098$	$-0,349$	$2,260$	$3,752$	$0,021$	$0,151$
		200	$0,495$	$0,751$	$0,063$	$-0,312$	$0,215$	$0,432$	$0,013$	$0,126$
		250	$0,382$	$0,570$	$0,049$	$-0,272$	$0,117$	$0,238$	$0,010$	$0,109$
	$0,5$	100	$0,802$	$1,122$	$0,043$	$-0,110$	$3,145$	$1,638$	$0,013$	$0,122$
		200	$0,376$	$0,528$	$0,001$	$0,062$	$1,097$	$0,452$	$0,008$	$0,119$
		250	$0,277$	$0,429$	$-0,011$	$0,076$	$0,301$	$0,144$	$0,007$	$0,122$
$0,5$	$0,25$	100	$0,931$	$1,131$	$0,144$	$-0,292$	$1,436$	$1,867$	$0,028$	$0,136$
		200	$0,379$	$0,462$	$0,087$	$-0,202$	$0,227$	$0,297$	$0,018$	$0,105$
		250	$0,308$	$0,385$	$0,062$	$-0,154$	$0,153$	$0,221$	$0,015$	$0,096$
	$0,5$	100	$0,772$	$1,057$	$0,101$	$-0,324$	$4,467$	$2,560$	$0,019$	$0,121$
		200	$0,237$	$0,348$	$0,040$	$-0,222$	$0,609$	$0,288$	$0,011$	$0,101$
		250	$0,144$	$0,249$	$0,022$	$-0,196$	$0,379$	$0,122$	$0,009$	$0,098$

Tabela 5.4: EMVs e erros padrões

	EMV			
Distribution	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$
TMOEL	$0,524\ (0,229)$	$0,014\ (0,006)$	$0,456\ (0,060)$	$-1,000\ (0,037)$
MOEL	$1,097\ (0,471)$	$0,008\ (0,003)$	$0,508\ (0,065)$	-
TL	-	$0,025\ (0,004)$	$0,519\ (0,035)$	$-1,000\ (0,038)$
Lomax	-	$0,008\ (0,001)$	$0,495\ (0,042)$	-

5.11 Aplicação

Nesta seção, apresentamos uma aplicação da distribuição TMOEL usando dados reais e comparamos com os seus sub-modelos, os quais são dados na Tabela 5.1. Para tanto, usamos um conjunto de dados que correspondem aos tamanhos de 269 arquivos de computador. Esse conjunto de dados foi analisado por Giles *et al.* (2013). Todos os resultados foram obtidos usando o *software* `R` versão 3.2.3, pacote `AdequacyModel`. Empregamos o método BFGS para maximizar o logaritmo da função de verossimilhança. Os EMVs dos parâmetros dos modelos e dos sub-modelos são dados na Tabela 5.4 (com os respectivos erros padrões entre parênteses).

Para fins de comparação, calculamos algumas estatísticas de bondade de ajuste: critério de informação de Akaike (AIC) (AKAIKE, 1974), critério de informação Bayesiano (BIC) (SCHWARZ, 1978), critério de informação de Hannan-Quinn (HQIC) (HANNAN; QUINN, 1979), critério Cramér-von Mises (W*) e critério Anderson-Darling (A*) (CHEN; BALAKRISHNAN, 1995). Em geral, menores valores dessas estatísticas sugerem melhor ajuste.

Também incluímos na comparação os seguintes modelos não encaixados: beta Weibull (BW) (LEE *et al.*, 2007), transmutada Marshall-Olkin Fréchet (TMOFr) (AFIFY *et al.*, 2015) e Kumaraswamy Weibull (KW) (CORDEIRO *et al.*, 2010). As funções densidade de probabilidade dessas distribuições são dadas no Apêndice 2. Os valores das estatísticas de bondade de ajuste são listadas na Tabela 5.5.

A Tabela 5.5 mostra que os valores das estatíticas AIC, BIC e HQIC correspondentes à distribuição TL são ligeiramente menores que os correspondentes à distribuição TMOEL. Porém, a distribuição TMOEL apresenta os menores valores das estatísticas W* e A* entre todos os modelos ajustados. Avaliando as estimativas e os erros padrões dados na Tabela 5.4, concluimos que os parâmetros de todos os modelos são significativos. Na Figura 5.5 mostramos os gráficos das densidades estimadas das distribuições KW, TL e TMOEL.

Tabela 5.5: Estatísticas de bondade de ajuste

	Estatística				
Distribuição	AIC	BIC	HQIC	W*	A*
TMOEL	$4590,497$	$4604,876$	$4596,272$	$\mathbf{1,234}$	$\mathbf{7,205}$
MOEL	$4611,764$	$4622,549$	$4616,095$	$1,579$	$8,832$
BW	$4602,505$	$4616,884$	$4608,280$	$1,248$	$7,278$
TMOFr	$4640,043$	$4654,422$	$4645,818$	$1,929$	$10,543$
KW	$4603,025$	$4617,404$	$4608,799$	$1,245$	$7,266$
Lomax	$4609,768$	$4616,957$	$4612,655$	$1,576$	$8,827$
TL	$4590,340$	$4601,124$	$4594,671$	$1,283$	$7,388$

5.12 Conclusões finais

Neste trabalho, estudamos um novo modelo de quatro parâmetros, a distribuição transmutada Marshall–Olkin estendida Lomax (TMOEL), como caso especial da família de distribuições transmutada-G (T-G) (SHAW; BUCKLEY, 2009) quando o modelo de base é a distribuição Marshall-Olkin Lomax (MOEL) (GHITANY *et al.*, 2007). Alguns sub-modelos da distribuição TMOEL são apresentados. Obtemos expressões simples para as funções densidade e acumulada. Estudamos algumas de suas propriedades estatísticas e matemáticas. Demonstramos que a função densidade da TMOEL pode ser expressa como uma combinação linear de funções densidade da Lomax exponencializada e assim algumas de suas propriedades podem ser determinadas a partir desse modelo. Obtemos expressões explícitas para a função quantílica e momentos ordinário e incompleto, função característica, desvios médios e estatísticas de ordem. Através de simulações de Monte Carlo avaliamos os estimadores de máxima verossimilhança dos parâmetros que indexam a distribuição TMOEL em amostras finitas. Comparamos o novo modelo com outras distribuições usando as estatísticas de bondade de ajuste clássicas. Mostramos a utilidade do modelo proposto através de uma aplicação a um conjunto de dados reais.

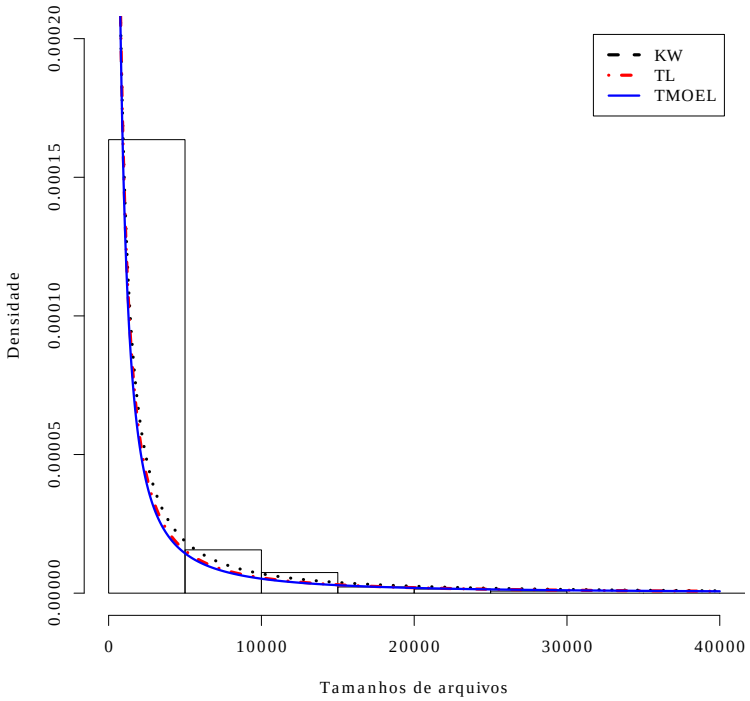


Figura 5.5: Comparação das densidades estimadas de TMOEL, KW e TL

6 Considerações finais

Neste trabalho, fizemos inferências (estimação pontual e teste de hipóteses) para dois parâmetros (um de forma e outro de escala) da distribuição Lomax exponencializada (LE) e para o parâmetro de forma da distribuição BurrX escalonada, baseadas na função de verossimilhança perfilada e suas versões modificadas (SEVERINI, 2000; COX; REID, 1993). Para a distribuição BurrX escalonada derivamos os ajustes para a função de verossimilhança perfilada considerando amostras completas e amostras censuradas do tipo II. Em todos os casos, estimamos pontualmente através do método de máxima verossimilhança e utilizamos a estatística da razão de verossimilhanças para realizar teste de hipóteses. Consideramos também o teste baseado na estatística da razão de verossimilhanças bootstrap Bartlett e o teste da razão de verossimilhanças bootstrap usual. Realizamos um estudo de simulação para avaliar e comparar os estimadores de máxima verossimilhança perfilado e perfilados ajustados. O estudo mostrou que, em geral, os estimadores de máxima verossimilhança perfilados ajustados tiveram os melhores desempenhos, pois foram os menos viesados e também apresentaram os menores valores para o erro quadrático médio.

Quanto ao teste de hipóteses, o teste baseado na estatística da razão de verossimilhanças bootstrap Bartlett, o teste da razão de verossimilhanças bootstrap usual e os testes que utilizam as funções de verossimilhança perfilada modificada apresentaram os melhores desempenhos (com exceção da estatística de teste baseada na proposta de Cox e Reid (1993) para a distribuição LE), principalmente em amostras de tamanho pequeno. Mostramos por meio de aplicações a conjuntos de dados reais a importância dos diferentes ajustes da função de verossimilhança perfilada para fazer inferências mais precisas sobre os parâmetros da distribuição LE e sobre o parâmetro de forma da distribuição BurrX escalonada.

Além disso, propomos uma distribuição de quatro parâmetros denominada *transmutada Marshall-Olkin estendida Lomax* (TMOEL). Mostramos que as distribuições Marshall-Olkin estendida Lomax, transmutada Lomax e Lomax são casos especiais do modelo TMOEL. Demonstramos que a função densidade de probabilidade do modelo proposto pode ser expressa como uma combinação linear de funções densidade da distribuição Lomax exponencializada. Obtemos expressões explícitas para a função quantílica e momentos ordinário e incompleto, função característica, desvios médios e estatísticas de ordem. Avaliamos, através de Monte Carlo, o comportamento dos estimadores de máxima verossimilhança dos parâmetros que indexam a distribuição TMOEL em amostras finitas. Ilustramos a utilidade do novo modelo por meio de uma aplicação a um conjunto de dados reais.

Referências

ABD-ELFATTAH, A. M. Goodness of fit test for the generalized Rayleigh distribution with unknown parameters. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 81, p. 357–366, 2011.

ABDUL-MONIEM, I.; ABDEL-HAMEED, H. On exponentiated Lomax distribution. *International Journal of Mathematical Archive*, v. 3, p. 2144–2150, 2012.

AFIFY, A. Z.; HAMEDANI, G.; GHOSH, I.; MEAD, M. The transmuted Marshall-Olkin Fréchet distribution: Properties and applications. *International Journal of Statistics and Probability*, v. 4, p. 132–148, 2015.

AHMAD, M. A.; RAQAB, M. Z.; KUNDU, D. Discriminating between the generalized Rayleigh and Weibull distributions: Some comparative studies. *Communications in Statistics - Simulation and Computation*, Taylor & Francis, v. 46, p. 4880–4895, 2017.

AKAIKE, H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions*, v. 19, p. 716–723, 1974.

ALEXANDER, C.; CORDEIRO, G. M.; ORTEGA, E. M. M.; SARABIA, J. M. Generalized beta-generated distributions. *Computational Statistics and Data Analysis*, v. 56, p. 1880–1897, 2012.

ARNOLD, B. *Pareto Distributions*. 2. ed. Boca Raton: CRC Press, 2015.

ASHOUR, S. K.; ELTEHIWY, M. A. Transmuted Lomax distribution. *American Journal of Applied Mathematics and Statistics*, Science and Education Publishing, v. 1, p. 121–127, 2013.

BARNDORFF-NIELSEN, O. Conditionality resolutions. *Biometrika*, v. 67, p. 293–310, 1980.

____. On a formula for the distribution of the maximum likelihood estimator. *Biometrika*, v. 70, p. 343–365, 1983.

BARRETO, L. S.; CYSNEIROS, A. H. M. A.; CRIBARI-NETO, F. Improved Birnbaum–Saunders inference under type II censoring. *Computational Statistics & Data Analysis*, v. 57, p. 68–81, 2013.

BEKKER, A.; ROUX, J.; MOSTEIT, P. A generalization of the compound Rayleigh distribution: using a bayesian method on cancer survival times. *Communications in Statistics - Theory and Methods*, v. 29, n. 7, p. 1419–1433, 2000.

BOURGUIGNON, M.; GHOSH, I.; CORDEIRO, G. M. General results for the transmuted family of distributions and new models. *Journal of Probability and Statistics*, Hindawi Publishing Corporation, 2016.

BRYSON, M. C. Heavy-tailed distributions: properties and tests. *Technometrics*, v. 16, p. 61–68, 1974.

BURR, I. W. Cumulative frequency functions. *The Annals of mathematical statistics*, JSTOR, v. 13, p. 215–232, 1942.

CHEN, G.; BALAKRISHNAN, N. A general purpose approximate goodness-of-fit test. *Journal of Quality Technology*, American Society for Quality Control, v. 27, p. 154–161, 1995.

CORDEIRO, G. *Introdução à teoria assintótica*. Rio de Janeiro: Conselho Nacional de Desenvolvimento Scientifico e Tecnológico, Institutode Matemática Pura e Aplicada, 1999.

CORDEIRO, G. M.; CASTRO, M. A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, v. 81, p. 883–898, 2011.

CORDEIRO, G. M.; ORTEGA, E. M.; CUNHA, D. C. The exponentiated generalized class of distributions. *Journal of Data Science*, v. 11, p. 1–27, 2013.

CORDEIRO, G. M.; ORTEGA, E. M.; NADARAJAH, S. The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, v. 347, p. 1399 – 1429, 2010.

COX, D. R.; REID, N. Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 1–39, 1987.

____. A note on the calculation of adjusted profile likelihood. *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 467–471, 1993.

CYSNEIROS, A. H. M. A.; CRIBARI-NETO, F.; ARAÚJO, C. A. On Birnbaum–Saunders inference. *Computational Statistics & Data Analysis*, v. 52, p. 4939–4950, 2008.

DOORNIK, J. A. An object-oriented matrix programming language Ox 6. Timberlake Consultants Ltd, 2009.

EFRON, B. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, v. 7, p. 1–26, 1979.

EFRON, B.; TIBSHIRANI, R. J. *An introduction to the bootstrap*. Boca Raton: CRC press, 1994.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*, v. 31, p. 497–512, 2002.

FERRARI, S. L.; LUCAMBIO, F.; CRIBARI-NETO, F. Improved profile likelihood inference. *Journal of Statistical Planning and Inference*, v. 134, p. 373 – 391, 2005.

FERRARI, S. L.; SILVA, M. F. D.; CRIBARI-NETO, F. Adjusted profile likelihoods for the Weibull shape parameter. *Journal of Statistical Computation and Simulation*, v. 77, p. 531–548, 2007.

FRASER, D.; REID, N. Ancillaries and third order significance. *Utilitas Mathematica*, v. 7, p. 33–53, 1995.

FRASER, D. A. S.; REID, N.; WU, J. A simple general formula for tail probabilities for frequentist and bayesian inference. *Biometrika*, v. 86, p. 249–264, 1999.

FRERY, A.; CRIBARI-NETO, F. *Elementos de estatística computacional usando plataformas de software livre/gratuito*. Rio de Janeiro: IMPA, 2005.

GHITANY, M. E.; AL-AWADHI, F. A.; ALKHALFAN, L. A. Marshall-Olkin extended Lomax distribution and its application to Censored Data. *Communications in Statistics - Theory and Methods*, v. 36, p. 1855–1866, 2007.

GILES, D. E.; FENG, H.; GODWIN, R. T. On the bias of the maximum likelihood estimator for the two-parameter Lomax distribution. *Communications in Statistics - Theory and Methods*, v. 42, p. 1934–1950, 2013.

GRADSHTEYN, I. S.; RYZHIK, I. M. *Table of integrals, series, and products*. 7. ed. Amsterdam: Elsevier/Academic Press, 2007.

GUPTA, R. C.; GUPTA, P. L.; GUPTA, R. D. Modeling failure time data by Lehmann alternatives. *Communications in Statistics - Theory and Methods*, v. 27, p. 887–904, 1998.

HAND, D.; DALY, F.; MCCONWAY, K.; LUNN, D.; OSTROWSKI, E. *A Handbook of Small Data Sets*. Boca Raton: CRC Press, 1993.

HANNAN, E. J.; QUINN, B. G. The Determination of the Order of an Autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 41, p. 190–195, 1979.

HARRIS, C. M. The Pareto distribution as a queue service discipline. *Operations Research*, v. 16, p. 307–313, 1968.

HOLLAND, O.; GOLAUP, A.; AGHVAMI, A. Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration. *IEEE Proceedings-Communications*, v. 153, p. 683–690, 2006.

JØRGENSEN, B. *Statistical properties of the generalized inverse gaussian distribution*. New York: Springer Science & Business Media, 2012.

KOZUBOWSKI, T. J.; PODGÓRSKI, K. Transmuted distributions and random extrema. *Statistics & Probability Letters*, Elsevier, v. 116, p. 6–8, 2016.

LEE, C.; FAMOYE, F.; OLLUMOLADE, O. Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern Applied Statistical Methods*, v. 6, p. 173–186, 2007.

LEMONTE, A. J.; CORDEIRO, G. M. An extended Lomax distribution. *Statistics*, v. 47, p. 800–816, 2013.

LOMAX, K. S. Business failures: another example of the analysis of failure data. *Journal of the American Statistical Association*, v. 49, p. 847–852, 1954.

MANN, N. R.; FERTIG, K. W. Tables for obtaining Weibull confidence bounds and tolerance bounds based on best linear invariant estimates of parameters of the extreme-value distribution. *Technometrics*, Taylor & Francis, v. 15, p. 87–101, 1973.

MARSHALL, A. W.; OLKIN, I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, v. 84, p. 641–652, 1997.

NADARAJAH, S.; CORDEIRO, G. M.; ORTEGA, E. M. M. The Zografos–Balakrishnan-G family of distributions: Mathematical properties and applications. *Communications in Statistics - Theory and Methods*, v. 44, p. 186–215, 2015.

NOCEDAL, J.; WRIGHT, S. *Numerical optimization*. 2: Springer Science & Business Media, 2006.

PACE, L.; SALVAN, A. *Principles of statistical inference: from a Neo-Fisherian perspective*. London: World scientific, 1997.

PESCIM, R. R.; DEMÉTRIO, C. G. B.; CORDEIRO, G. M.; ORTEGA, E. M. M.; URBANO, M. R. The beta generalized half-normal distribution. *Computational Statistics and Data Analysis*, v. 54, p. 945–957, 2010.

PIRES, J. F. *Refinamento de inferências nas distribuições gaussiana inversa triparamétrica, Pareto generalizada e Lomax*. 89 p. Tese (Doutorado) — Matemática computacional, CCEN, Universidade Federal de Pernambuco, Recife, 2014.

RADY, E.-H. A.; HASSANEIN, W.; ELHADDAD, T. The power Lomax distribution with an application to bladder cancer data. *SpringerPlus*, v. 5, p. 1838, 2016.

RAQAB, M. Z.; KUNDU, D. Burr type X distribution: revisited. *Journal of Probability and Statistical Sciences*, v. 4, p. 179–193, 2006.

RAQAB, M. Z.; MADI, M. T. Inference for the generalized Rayleigh distribution based on progressively censored data. *Journal of Statistical Planning and Inference*, v. 141, p. 3313 – 3322, 2011.

RISTIĆ, M. M.; BALAKRISHNAN, N. The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, v. 82, p. 1191–1206, 2012.

ROCKE, D. Bootstrap Bartlett adjustment in seemingly unrelated regression. v. 84, p. 598–601, 1989.

SALEM, H. M. The exponentiated Lomax distribution: different estimation methods. *American Journal of Applied Mathematics and Statistics*, Science and Education Publishing, v. 2, p. 364–368, 2014.

SCHWARZ, G. Estimating the dimension of a model. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 6, p. 461–464, 1978.

SEVERINI, T. A. An empirical adjustment to the likelihood ratio statistic. *Biometrika*, v. 86, p. 235–247, 1999.

____. *Likelihood methods in statistics*. New York: Oxford University Press, 2000.

____. *Elements of distribution theory*. Cambridge: Cambridge University Press, 2005.

SHAW, W. T.; BUCKLEY, I. R. C. The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *ArXiv*, 2009.

SURLES, J.; PADGETT, W. Inference for reliability and stress-strength for a scaled Burr type X distribution. *Lifetime Data Analysis*, Springer, v. 7, p. 187–200, 2001.

TAHIR, M. H.; NADARAJAH, S. Parameter induction in continuous univariate distributions: Well-established G families. *Anais da Academia Brasileira de Ciências*, v. 87, p. 539–568, 2015.

ZOGRAFOS, K.; BALAKRISHNAN, N. On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, v. 6, p. 344 – 362, 2009.

Apêndice A - Matriz de informação observada

Segundas derivadas do logaritmo da função de verossimilhança da distribuição Lomax exponencializada para a obtenção da matriz de informação observada.

$$\frac{\partial^2 \ell(\theta)}{\partial \alpha^2} = -\frac{n}{\alpha^2},$$

$$\frac{\partial^2 \ell(\theta)}{\partial \alpha \partial \beta} = \lambda \sum_{i=1}^{n} \frac{x_i u_i^{-(1+\lambda)}}{1-u_i^{-\lambda}},$$

$$\frac{\partial^2 \ell(\theta)}{\partial \alpha \partial \lambda} = \sum_{i=1}^{n} \frac{\log(u_i) u_i^{-\lambda}}{1-u_i^{\lambda}},$$

$$\begin{aligned}\frac{\partial^2 \ell(\theta)}{\partial \beta^2} &= -\frac{n}{\beta^2} + (\lambda+1)\sum_{i=1}^{n}\left(\frac{x_i}{u_i}\right)^2 - (\alpha-1)\lambda^2 \sum_{i=1}^{n} \frac{x_i^2 u_i^{-2(\lambda+1)}}{(1-u_i^{-\lambda})^2} \\ &- \lambda(\alpha-1)(\lambda+1)\sum_{i=1}^{n}\frac{x_i^2 u_i^{-(2+\lambda)}}{1-u_i^{-\lambda}},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \ell(\theta)}{\partial \beta \partial \lambda} &= -\sum_{i=1}^{n}\frac{x_i}{u_i} - \lambda(\alpha-1)\sum_{i=1}^{n}\frac{x_i \log(u_i) u_i^{-(1+2\lambda)}}{(1-u_i^{-\lambda})^2} + (\alpha-1)\sum_{i=1}^{n}\frac{x_i u_i^{-(1+\lambda)}}{1-u_i^{\lambda}} \\ &- \lambda(\alpha-1)\sum_{i=1}^{n}\frac{x_i \log(u_i) u_i^{-(1+\lambda)}}{1-u_i^{-\lambda}},\end{aligned}$$

$$\frac{\partial^2 \ell(\theta)}{\partial \lambda^2} = -\frac{n}{\lambda^2} - (\alpha-1)\sum_{i=1}^{n}\frac{\log(u_i^2) u_i^{-2\lambda}}{(1-u_i^{-\lambda})^2} - (\alpha-1)\sum_{i=1}^{n}\frac{\log^2(u_i) u_i^{-\lambda}}{1-u_i^{-\lambda}},$$

em que $u_i = 1 + \beta x_i$.

Apêndice B - Aspectos Computacionais

Neste apêndice, detalhamos alguns aspectos computacionais referentes aos Capítulos 3 e 4. Todos os resultados das simulações de Monte Carlo e aplicações apresentados foram obtidos utilizando programas implementados na linguagem de programação `Ox`.

`Ox` é uma linguagem de programação matricial orientada a objetos que, utilizando uma sintaxe semelhante com as de `C` e `C++`, oferece uma grande quantidade de recursos matemáticos e estatísticos. Do ponto de vista numérico, `Ox` é uma das mais confiáveis plataformas para computação científica (DOORNIK, 2009; FRERY; CRIBARI-NETO, 2005). Uma versão gratuita do `Ox` está disponível em: <https://www.doornik.com>.

Geradores de números aleatórios

Para gerar números aleatórios das distribuições Lomax exponencializada e BurrX escalonada estudadas, respectivamente, nos Capítulos 3 e 4, utilizamos o gerador de números aleatórios de George Marsaglia (GM). Este gerador com período de aproximadamente 2^{60} e que usa duas sementes passa por testes rigorosos de aleatoriedade. Além disso, possui uma eficiente implementação em `Ox`.

No Capítulo 3, o método da inversão foi empregado para gerar números aleatórios da distribuição LE, respectivamente, da seguinte forma:

Gerando número aleatório x da distribuição LE:

1 Gerar u a partir da distribuição $\mathcal{U}(0,1)$.

2 Calcular $x = \beta^{-1}\left[\left(1-u^{\frac{1}{\alpha}}\right)^{-\frac{1}{\lambda}}-1\right]$.

Assim, a função para gerar uma amostra de tamanho n da distribuição Lomax exponencalizada é dada pela seguinte rotina:

```
//função para gerar uma amostra de tamanho
n da distribuição exp_lomax(alpha, beta, lambda)
amostraEL(const N, const vP)
{
 decl u         = ranu(N, 1);
```

```
 decl amostra  = (1/vP[1])* (((1 - ( u .^(1/vP[0])))
                .^(-1/vP[2])) - 1);
 return amostra;
}
```

Também usando o método da inversão, no Capítulo 4, geramos números aleatórios da distribuição BurrX escalonada, seguindo o algoritmo:

Gerando número aleatório x da distribuição BurrX escalonada:

1 Gerar u a partir da distribuição $\mathcal{U}(0,1)$.

2 Calcular $x = \sigma\left[-\log\left(1-u^{\frac{1}{\lambda}}\right)\right]^{\frac{1}{2}}$.

Maximização das funções de verossimilhança

Os estimadores de máxima verossimilhança, dados nos Capítulos 3 e 4, obtidos maximizando as funções de verossimilhança perfilada e suas versões modificadas, não possuem forma fechada. Além disso, em todos os casos considerados, não foi possível obter explicitamente as funções de verossimilhança perfilada e suas versões modificadas devido ao fato de $\hat{\nu}_\mu$ também não ter forma fechada. Dessa forma, para maximizar essas funções e assim, obter os estimadores de máxima verossimilhança perfilado e perfilado ajustados, utilizamos a rotina `MAXSQP` implementada na plataforma `OX`. Essa rotina maximiza funções com restrições nas variáveis (parâmetros) e está implementada no `OX` (Versão 7.10, para Windows) da seguinte forma:

```
#import <maxsqp>
MaxSQP(const func, const avP, const adFunc, const amHessian,
 const fNumDer, const cfunc_gt0, const cfunc_eq0, vLo, vHi);
```

- `func`

  Entrada: a função que deve ser maximizada com p parâmetros, neste caso, a função de log-verossimilhança.

- `avP`

  Entrada: uma matriz de dimensão $p \times 1$ com os valores iniciais.

  Saída: matriz de dimensão $p \times 1$ com os valores que maximizam `func`.

- `adFunc`

  Entrada: endereçamento de um objeto.

  Saída: valor máximo de `func`.

- `amHessian`

  Entrada: 0 ou a matriz hessiana endereçada (usamos 0).

- `fNumDer`

  Entrada: 0, usa primeiras derivadas analíticas ou 1, usa primeiras derivadas numéricas (utilizamos 1).

- `vLo`

  Entrada: matriz de dimensão $p \times 1$ com os limites inferiores dos parâmetros, ou $<>$ se a função é ilimitada inferiormente.

- `vHi`

  Entrada: matriz de dimensão $p \times 1$ com os limites superiores, ou $<>$ se a função é ilimitada superiormente

- `cfunc_gt0`

  Entrada: 0

- `cfunc_eq0` é uma matriz cujos os elementos são as funções de restrições não-lineares (a qual deve ser restrita em zero) com o seguinte formato:

  ```
  cfunc_eq0(const avF, const vP);
  ```

  - `avF`

    Entrada: matriz de restrições para `vP`.

  - `vP`

    Entrada: matriz de dimensão $p \times 1$ de parâmetros.

Por exemplo, no Capítulo 4, quando fizemos inferência para o parâmetro da λ da distribuição BurrX escalonada ao considerar amostra completa (dados não censurados), temos:

$$\ell_p(\lambda) = n \log\left(\frac{2\lambda}{\hat{\sigma}_\lambda}\right) + (\lambda - 1)\sum_{i=1}^{n} \log\left[1 - e^{-\tilde{u}_i^2}\right] + \sum_{i=1}^{n}[\log(\tilde{u}_i) - \tilde{u}_i^2], \tag{6.1}$$

em que $\tilde{u}_i = x_i/\hat{\sigma}_\lambda$ e $\hat{\sigma}_\lambda$ satisfaz a equação

$$-\frac{2n}{\hat{\sigma}_\lambda} + \frac{2}{\hat{\sigma}_\lambda}\sum_{i=1}^{n}\tilde{u}_i^2 - \frac{2(\lambda-1)}{\hat{\sigma}_\lambda}\sum_{i=1}^{n}\frac{\tilde{u}_i^2 e^{-\tilde{u}_i^2}}{G(\tilde{u}_i)} = 0,$$

para λ fixo.

Neste caso, quando λ é parâmetro de interesse, o logaritmo da função de verossimilhança perfilada é expresso em função da estimativa de máxima verossimilhança restrita $\hat{\sigma}_\lambda$, que não possui forma fechada. Devido ao desconhecimento de $\hat{\sigma}_\lambda$, $\ell_p(\lambda)$ não está explicitamente definida em função do parâmetro λ. Dessa forma, a estimativa de máxima verossimilhança perfilada $\hat{\lambda}_p$ é obtida da maximização da função

$$\ell(\sigma,\lambda) = n\log\left(\frac{2\lambda}{\sigma}\right) + (\lambda-1)\sum_{i=1}^{n}\log[1-e^{-u_i^2}] + \sum_{i=1}^{n}[\log(u_i) - u_i^2], \tag{6.2}$$

sujeita à restrição de igualdade

$$-\frac{2n}{\sigma} + \frac{2}{\sigma}\sum_{i=1}^{n}u_i^2 - \frac{2(\lambda-1)}{\sigma}\sum_{i=1}^{n}\frac{u_i^2 e^{-u_i^2}}{G(u_i)} = 0, \tag{6.3}$$

em que $u_i = x_i/\sigma$. As equações (6.2) e (6.3) são implementadas, usando linguagem de programação Ox, da seguinte forma:

```
//função de log-verossimilhança perfilada
func(const vP, const adFunc, const avScore, const amHessian)
{
 decl u = (x ./vP[0]);
 decl g = 1 - exp(-(u .^2));

 adFunc[0] = n*log((2*vP[1])/vP[0]) + (vP[1]-1)*sumc(log(g))
           - sumc(u .^2) + sumc(log(u);
  if(avScore)
{
    (avScore[0])[0] = -n*2/vP[0] + (2/vP[0])*sumc(u .^2)
                    -((2*(vP[1]-1))/vP[0])*sumc((u .^2
                    .* exp(-(u.^2))) ./ g);
(avScore[0])[1] = n/vP[1] + sumc(log(g));
}
  if(isnan(adFunc[0]) || isdotinf(adFunc[0]))
```

```
 return 0;
  else
     return 1;
}

cfunc_eq0(const avF, const vP) // restrição
{
 decl u = (x ./vP[0]);
 decl g = 1-exp(-(u .^2));
 avF[0] = -n*2/vP[0] + ( 2/vP[0])*sumc(u .^2)
       - ((2*(vP[1]-1))/vP[0])* sumc((u .^2
       .* exp(-(u .^2))) ./ g);
 return 1;
}
```

Finalmente, a função (6.1) é maximizada utilizando a seguinte rotina:

```
 MaxSQP(func, &vp, &f, 0, TRUE, 0, cfunc_eq0, vlo, vhi);
```

Algumas observações:

- salientamos que as funções de verossimilhança perfilada modificadas por também não apresentarem uma forma explícita em função do parâmetro de interesse, devem ser maximizadas utilizando o método SQP, seguindo um procedimento semelhante ao descrito anteriormente;
- os estimadores de máxima verossimilhança $\hat{\mu}$ e $\hat{\nu}$ nos ajustes (2.4), (2.5) e (2.8) podem ser encontrados maximizando diretamente a função de verossimilhança usual utilizando o método BFGS (implementado na plataforma `Ox` através do comando `MaxBFGS`);
- nos estudos de simulação ao maximizar as funções de verossimilhança, o número de não convergências em cada cenário, não ultrapassaram 10% do total de réplicas de Monte Carlo considerado;
- os valores iniciais utilizados para maximizar as diferentes funções de verossimilhança foram atribuídos de forma arbitrária repetidas vezes até conseguir estabilidade na convergência das estimativas;
- os códigos que implementam bootstrap paramétrico, em geral, foram os mais custosos computacionalmente. Por exemplo, para a distribuição BurrX escalonada, sob o enfoque

de amostras completas, considerando: $n = 20$, $1\,000$ réplicas bootstrap e $10\,000$ réplicas de Monte, o tempo de execução do código foi de aproximadamente um minuto. Ao aumentar o tamanho amostral n, fazendo $n = 100$, o tempo de execução passou a ser de aproximadamente oito minutos.

Apêndice C - Funções densidade de probabilidade

As densidades dos modelos usados para comparar com a distribuição TMOEL são dados a seguir.

A densidade da transmutada Marshall-Olkin Fréchet (TMOFr) é dada por

$$f(x) = \frac{\alpha\beta\sigma^{\beta}x^{-(\beta+1)}\mathrm{e}^{-\left(\frac{\sigma}{\mathrm{x}}\right)^{\beta}}}{\left[\alpha+(1-\alpha)\mathrm{e}^{-\left(\frac{\sigma}{\mathrm{x}}\right)^{\beta}}\right]^2}\left[1+\lambda-\frac{2\lambda\mathrm{e}^{-\left(\frac{\sigma}{\mathrm{x}}\right)^{\beta}}}{\alpha+(1-\alpha)\mathrm{e}^{-\left(\frac{\sigma}{\mathrm{x}}\right)^{\beta}}}\right],\ x>0,$$

em que β, α, σ são parâmetros positivos e $\mid \lambda \mid \leq 1$.

A densidade beta Weibull (BW) é dada por

$$f(x) = \frac{\alpha\theta}{B(a,b)}x^{\theta-1}(1-\mathrm{e}^{-\alpha x^{\theta}})^{a-1}\mathrm{e}^{-\alpha b x^{\theta}},\ x>0,$$

em que a, b, α e θ são parâmetros positivos e $B(\cdot,\cdot)$ é a função beta.

A densidade Kumaraswamy Weibull (KW) é dada por

$$f(x) = ab\alpha x^{\theta-1}\mathrm{e}^{-\alpha x^{\theta}}(1-\mathrm{e}^{-\alpha x^{\theta}})^{a-1}[1-(1-\mathrm{e}^{-\alpha x^{\theta}})^{a}]^{b-1},\ x>0,$$

em que a, b, α e θ são parâmetros positivos.