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

ALEXSANDRO ARRUDA FERREIRA

REFINAMENTO ASSINTÓTICO NA DISTRIBUIÇÃO BURR XII E NO MODELO DE REGRESSÃO LOG-BURR XII

Recife

2021

ALEXSANDRO ARRUDA FERREIRA

REFINAMENTO ASSINTÓTICO NA DISTRIBUIÇÃO BURR XII E NO MODELO DE REGRESSÃO LOG-BURR XII

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de mestre em Estatística.

Área de Concentração: Estatística Matemática

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

Recife

2021

Catalogação na fonte
Bibliotecário Cristiano Cosme S. dos Anjos, CRB4-2290

F383r Ferreira, Alexsandro Arruda
Refinamento assintótico na distribuição Burr XII e no modelo de regressão log-Burr XII / Alexsandro Arruda Ferreira. – 2021.
104 f.: il., fig., tab.

Orientador: Audrey Helen Mariz de Aquino Cysneiros.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatistica, Recife, 2021.
Inclui referências.

1. Distribuição Burr XII. 2. Modelo de regressão log-Burr XII. 3. Estimador de máxima verossimilhança. I. Cysneiros, Audrey Helen Mariz de Aquino (orientador). II. Título.

519.5 CDD (22. ed.) UFPE-CCEN 2021-53

ALEXSANDRO ARRUDA FERREIRA

REFINAMENTO ASSINTÓTICO NA DISTRIBUIÇÃO BURR XII E NO MODELO DE REGRESSÃO LOG-BURR XII

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

Aprovada em: 25 de fevereiro de 2021.

BANCA EXAMINADORA

Profa. Dra. Audrey Helen Mariz de Aquino Cysneiros
UFPE

Profa. Dra. Aline Barbosa Tsuyuguchi
UFPE

Prof. Dr. José Nilton da Cruz
UFMT

AGRADECIMENTOS

- Primeiramente, agradeço à minha mãe Maria e ao meu pai Waldemar, pelos valores transmitidos em todo o percurso da minha vida e apoio em todas as horas, por sempre me orientarem a continuar estudando e fazer por onde ser melhor.
- À minha orientadora Audrey Helen Mariz de Aquino Cysneiros, pela paciência e confiança. Com certeza, foi um aprendizado muito grande para minha formação, ser orientado por uma professora com tanto conhecimento e experiência. Sua orientação foi essencial.
- A todos os professores do departamento de estatística da UFPE, em especial aos professores: Alex Dias Ramos, Gauss Montinho Cordeiro, Francisco Cribari Neto, Francisco José de Azevêdo Cysneiros, Maria do Carmo Soares de Lima e Pablo Martin Rodriguez. Serei grato a todos vocês!
- Por fim, agradeço a todos que me ajudaram de maneira direta e indiretamente nessa jornada com seus conselhos, conhecimentos e paciência. Que de certa forma fizeram possível a realização desse trabalho.
- À CAPES, pelo apoio financeiro.

RESUMO

Nesta dissertação, derivamos ajustes para a função de verossimilhança perfilada, baseado nas propostas de Barndorff-Nielsen (1983) e Cox e Reid (1987) e nas aproximações propostas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998) na distribuição Burr XII e no modelo de regressão log-Burr XII. Obtivemos os estimadores de máxima verossimilhança perfilado e perfilados modificados. Além disso, mostramos como ficam os testes baseados nas estatísticas da razão de verossimilhanças perfiladas e perfiladas modificadas. Através de simulações de Monte Carlo, avaliamos numericamente o comportamento dos estimadores pontuais, sob diferentes cenários, assim como, os desempenhos dos testes da razão de verossimilhanças baseados nas verossimilhanças perfiladas modificadas, em relação ao tamanho e poder. Os resultados evidenciam que tanto os testes quanto os estimadores baseados nas versões modificadas da verossimilhança perfilada, que foram desenvolvidas neste trabalho, possuem desempenho superior em pequenas amostras quando comparados com sua contrapartida não modificada. Por fim, cinco aplicações a dados reais são apresentadas.

Palavras-chaves: Distribuição Burr XII. Modelo de regressão log-Burr XII. Estimador de máxima verossimilhança. Verossimilhança perfilada. Verossimilhança perfilada modificada. Teste da razão de verossimilhanças.

ABSTRACT

In this dissertation, We consider the adjustments to the profile likelihood function proposed by Barndorff-Nielsen (1983), Cox e Reid (1987), Fraser e Reid (1995), Fraser, Reid e Wu (1999) and Severini (1998) Burr XII distribution and in the log-Burr XII regression models. Addition, we show how the tests are based on the profile likelihood ratio statistics and profile likelihood adjusted. Monte Carlo simulation results on the finite sample performances of the usual profile maximum likelihood estimator and profile likelihood ratio test and also their modified versions is presented and discussed. The numerical evidence favors the modified profile maximum likelihood estimators and tests we propose. Finally, we consider five real datasets as illustrations.

Keywords: Burr XII distribution. log-Burr XII regression models. Maximum likelihood estimator. Profile likelihood. Modified profile likelihood. Likelihood ratio test.

LISTA DE FIGURAS

Figura 1 – Gráficos das funções de densidade (a) e distribuição acumulada (b) para alguns valores dos parâmetros da distribuição Burr XII. 27
Figura 2 – Gráfico da função de risco para alguns valores dos parâmetros da distribuição Burr XII. 28
Figura 3 – Gráfico de discrepância relativa dos quantis da distribuição Burr XII para $n = 55$, $s = 0,5$, $c = 3$ e $k = 0,25$. 39
Figura 4 – Gráfico de discrepância relativa dos quantis da distribuição Burr XII para $n = 40$, $\boldsymbol{\mu} = (0,5; 0,5)^\top$ e $c = 5$. 47
Figura 5 – Gráfico de discrepância relativa dos quantis da distribuição log-logística para $n = 20$, $c = 4$ e $s = 2$. 58
Figura 6 – Gráfico de discrepância relativa dos quantis da distribuição Burr XII com $n = 40$, $\boldsymbol{\mu} = (0,5; 1,5)$, $s = 2$. 65
Figura 7 – Gráfico de discrepância relativa dos quantis da distribuição Lomax para $n = 30$, $k = 0,25$ e $s = 3$. 73
Figura 8 – Função densidade do modelo LBXII para diferentes valores de σ para $\mu = 5$ e $k = 2$. 77
Figura 9 – Gráfico de discrepância relativa dos quantis da distribuição LBXII para $n = 55$, $\sigma = 0,5$, $\beta_0 = -1$, $\beta_1 = 1$ e $k = 0,25$. 87

LISTA DE TABELAS

Tabela 1 – Medidas descritivas de diferentes estimadores de k para $c = 3$ e $s = 0,5$. . 38

Tabela 2 – Médias e variâncias amostrais de diferentes estatísticas de teste e os valores assintóticos destas, $H_0 : k = k_0$, distribuição χ^2_1. 39

Tabela 3 – Taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : k = k_0$, para $s = 0,5$ e $c = 3$. 40

Tabela 4 – Taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 55$ e nível nominal de 10%, indexados pelos parâmetros $c = 3$ e $s = 0,5$. 41

Tabela 5 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo Burr XII indexado pelos parâmetros $c = 3$, $s = 0,5$, $k = 0,25$ e $n = 35$. . . 41

Tabela 6 – Medidas descritivas de diferentes estimadores de $\boldsymbol{\mu} = (k, s)^\top$ para $c = 5$. . 46

Tabela 7 – Médias e variâncias amostrais de diferentes estatísticas de teste e os valores assintóticos destas, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, distribuição χ^2_2. 47

Tabela 8 – taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, $\boldsymbol{\mu} = (k, s)^\top$ para $c = 5$. 48

Tabela 9 – taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 30$ e nível nominal de 10%. 49

Tabela 10 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo Burr XII indexado pelos parâmetros $c = 5$, $\boldsymbol{\mu} = (0,5; 0,5)$ e $n = 30$. . . . 49

Tabela 11 – Medidas descritivas de diferentes estimadores de s para $c = 4$. 55

Tabela 12 – Médias e variâncias amostrais de diferentes estatísticas de teste e os valores assintóticos destas, $H_0 : s = s_0$, distribuição χ^2_1. 56

Tabela 13 – taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : s = s_0$, para $c = 4$. 57

Tabela 15 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo log-logística indexado pelos parâmetros $c = 4$, $s = 2$ e $n = 20$. 58

Tabela 14 – taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 20$ e nível nominal de 10%. 59

Tabela 16 – Medidas descritivas de diferentes estimadores de $\boldsymbol{\mu} = (c, k)^\top$ para $s = 2$. . 64

Tabela 17 – Médias e variâncias amostrais das estatísticas de teste e os valores assintóticos destas, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, distribuição χ^2_2. 65

Tabela 18 – taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, $\boldsymbol{\mu} = (c, k)^\top$ para $(s = 2)$. 66
Tabela 19 – taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 30$ e nível nominal de 10%. 67
Tabela 20 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo Burr XII indexado pelos parâmetros $\boldsymbol{\mu} = (1,5; 0,5)$, $s = 2$ e $n = 30$. . . . 67
Tabela 21 – Medidas descritivas de diferentes estimadores de k para $s = 3$. 72
Tabela 22 – Médias e variâncias amostrais das estatísticas de teste e os valores assintóticos destas, $H_0 : k = k_0$, distribuição χ^2_1. 73
Tabela 23 – taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : k = k_0$, para $s = 3$. 74
Tabela 24 – taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 30$ e nível nominal de 10%. 75
Tabela 25 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo Lomax indexado pelos parâmetros $k = 0,25$, $s = 3$ e $n = 30$. 75
Tabela 26 – Medidas descritivas de diferentes estimadores de k para $\sigma = 0,5$, $\beta_0 = -1$ e $\beta_1 = 1$. 86
Tabela 27 – Médias e variâncias amostrais das estatísticas de teste e os valores assintóticos destas, $H_0 : k = k_0$, distribuição χ^2_1. 87
Tabela 28 – Taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : k = k_0$, para $\sigma = 0,5$, $\beta_0 = -1$ e $\beta_1 = \beta_2 = \beta_3 = 1$. . . . 88
Tabela 29 – Taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 55$ e nível nominal de 10%, indexados pelos parâmetros $\sigma = 0,5$, $\beta_0 = -1$ e $\beta_1 = 1$. 89
Tabela 30 – velocidade mensal do vento (m/s) em Cameron Highland nos anos de 2004 a 2006. 91
Tabela 31 – Medidas descritivas. 91
Tabela 32 – Tempos de remissão (em meses) de pacientes com câncer de bexiga. . . . 92
Tabela 33 – Medidas descritivas. 93
Tabela 34 – Intervalos de tempo (em dias) dos sucessivos terremotos. 94
Tabela 35 – Medidas descritivas. 94
Tabela 36 – Medidas descritivas. 95
Tabela 37 – Concentração de ferritina plasmática (ng/dl) em 202 atletas. 96

Tabela 38 – Medidas descritivas. 97
Tabela 39 – Tempo de durabilidade (t) e leituras de odômetro (v) de 98 carros (por 1000 km). 98

SUMÁRIO

1 INTRODUÇÃO 13

2 AJUSTES À VEROSSSIMILHANÇA PERFILADA 18

2.1 DEFINIÇÕES PRELIMINARES 18

2.2 VEROSSIMILHANÇA PERFILADA 20

2.3 AJUSTES PARA VEROSSIMILHANÇA PERFILADA PROPOSTA POR BARNDORFF-NIELSEN 22

2.4 VEROSSIMILHANÇA PERFILADA MODIFICADA PROPOSTA POR COX E REID 24

2.5 TESTE DE HIPÓTESES 25

3 DISTRIBUIÇÃO BURR XII 26

3.1 INTRODUÇÃO 26

3.2 DEFINIÇÕES BÁSICAS 26

3.3 ESTIMAÇÃO DOS PARÂMETROS 29

3.4 VEROSSIMILHANÇA PERFILADA AJUSTADA PARA A DISTRIBUIÇÃO BURR XII 30

3.5 CASO (i) 31

3.5.1 Resultados numéricos 34

3.6 CASO (ii) 41

3.6.1 Resultados numéricos 43

3.6.2 Caso particular: distribuição log-logística 49

3.6.3 Resultados numéricos 52

3.7 CASO (iii) 59

3.7.1 Resultados numéricos 61

3.7.2 Caso particular: distribuição Lomax 67

3.7.3 Resultados numéricos 69

4 MODELO DE REGRESSÃO LOG-BURR XII 76

4.1 DEFINIÇÕES BÁSICAS 76

4.2 ESTIMAÇÃO DOS PARÂMETROS 78

4.3 VEROSSIMILHANÇA PERFILADA MODIFICADA PARA O MODELO DE REGRESSÃO LBXII 80

4.4 RESULTADOS NUMÉRICOS 83
5 APLICAÇÕES 90
5.1 EXEMPLO 1 90
5.2 EXEMPLO 2 92
5.3 EXEMPLO 3 94
5.4 EXEMPLO 4 95
5.5 EXEMPLO 5 96
6 CONCLUSÃO 99
REFERÊNCIAS 101

1 INTRODUÇÃO

A distribuição Burr XII faz parte de um conjunto de doze distribuições contínuas introduzidas por Burr (1942), é uma distribuição muito usada em análise de sobrevivência para modelar fenômenos com funções de risco monótonas e unimodais, possui como submodelos as distribuições Weibull, log-logística e pareto tipo II (Lomax). Além disso, produz uma ampla gama de valores de assimetria e curtose e é amplamente utilizada na análise de confiabilidade como uma alternativa mais flexível à distribuição Weibull.

Como exemplos de trabalhos publicados sobre a distribuição Burr XII, podemos mencionar: Singh e Moddola (1976), Rodriguez (1977), Papadopoulos (1978), Lewis (1981), Evans e Ragab (1983), Wang e Keats (1983), Khan e Khan (1987), Zimmer, Keats e Wang (1998), Soliman (2002), Silva et al. (2008) e Jalali e Watkins (2009).

Singh e Moddola (1976) mostraram que a distribuição Burr XII possui excelentes propriedades para modelar o tamanho da distribuição de renda. Rodriguez (1977) mostrou que a região de cobertura Burr XII em um plano específico é ocupado por algumas distribuições bem conhecidas como, a distribuição normal, log-normal, gama, logística e de valor extremo tipo I também conhecida como distribuição Gumbel. Papadopoulos (1978) utilizou uma abordagem bayesiana para fazer inferências sobre os parâmetros da distribuição Burr XII. Lewis (1981) declarou que muitas distribuições teóricas padrão, incluindo às duas distribuições de tempo de falha comuns, Weibull e a Exponencial são casos especiais ou casos limitantes da distribuição Burr XII. Evans e Ragab (1983) utilizaram uma abordagem bayesiana para estimação do parâmetro de forma, baseada em amostras censuradas tipo II. Wang e Keats (1983) utilizaram o método de máxima verossimilhança para a obtenção de estimativas pontuais e intervalares dos parâmetros da distribuição na presença, ou não de censura. Khan e Khan (1987) forneceram relações de recorrência para estatísticas de ordem na distribuição Burr XII. Zimmer, Keats e Wang (1998) apresentaram algumas propriedades estatísticas, como, a média, a variância, a assimetria e a curtose. Além disso, introduziram um terceiro parâmetro na distribuição Burr XII. Soliman (2002) usou a distribuição Burr XII para comparar os estimadores da função de confiabilidade GLM (generalized life model). Silva et al. (2008) propuseram o modelo de regressão log-Burr XII como uma alternativa ao modelo de regressão log-logística. Jalali e Watkins (2009) consideram três aspectos relacionados aos estimadores de máxima verossimilhança dos dois parâmetros da distribuição Burr XII.

Em algumas situações, é de interesse fazer inferências apenas em um subconjunto do vetor de parâmetros do modelo. Dessa maneira, denominamos esses parâmetros de interesse e os demais de parâmetros de incômodo. Quando o modelo envolve parâmetros de incômodo, neste caso, as inferências podem ser baseadas na função de verossimilhança perfilada. Essa verossimilhança é obtida substituindo o parâmetro de incômodo por uma estimativa consistente na verossimilhança original. Embora, a verossimilhança perfilada possui algumas propriedades da verossimilhança original, como, por exemplo, o estimador de máxima verossimilhança (EMV) e o EMV perfilado com relação ao parâmetro de interesse são iguais. O mesmo ocorre com o teste da razão de verossimilhanças, isto é, ao se testar hipóteses sobre o parâmetro de interesse, a estatística da razão de verossimilhanças perfiladas é igual à estatística da razão de verossimilhanças original. Infelizmente, ela não usufrui de todas as propriedades de uma verossimilhança genuína. Por exemplo, a função escore e a informação de Fisher podem apresentar viés. Assim, tratar a verossimilhança perfilada como uma verossimilhança genuína pode levar a alguns problemas, como, inconsistência e ineficiência dos estimadores. Quando há um número considerável de parâmetros de incômodo, há uma deterioração na qualidade das aproximações envolvidas nas inferências que se baseiam em resultados assintóticos.

Na literatura, várias modificações foram propostas por diversos pesquisadores com objetivo de melhorar as propriedades inferenciais da verossimilhança perfilada, a saber: Barndorff-Nielsen (1983), Cox e Reid (1987), McCullagh e Tibishirani (1990), Cox e Reid (1993), Barndorff-Nielsen (1994), Fraser e Reid (1995), Stern (1997), Severini (1998), Severini (1999), Fraser, Reid e Wu (1999) e Severini (2000). A presente dissertação dará ênfase às propostas de Cox e Reid (1987) e Barndorff-Nielsen (1983). Neste último, usaremos as aproximações sugeridas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998).

Barndorff-Nielsen (1983) propôs uma versão modificada da função de verossimilhança perfilada baseada em uma aproximação da função densidade de probabilidade do EMV do vetor de parâmetros do modelo condicionado a uma estatística ancilar. Entretanto, é frequentemente difícil a obtenção dessa verossimilhança perfilada modificada em problemas particulares, uma vez que esta depende da obtenção de uma estatística ancilar e envolve derivadas com respeito ao espaço amostral. Vale salientar aqui, que a versão modificada da função de verossimilhança perfilada proposta, é invariante sob reparametrizações da forma $(\boldsymbol{\mu}, \boldsymbol{\phi}) \rightarrow (\boldsymbol{\mu}, \zeta(\boldsymbol{\mu}, \boldsymbol{\phi}))$, sendo $\boldsymbol{\mu}$ o vetor de parâmetros de interesse, $\boldsymbol{\phi}$ o vetor de parâmetros de incômodo e ζ uma função de $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$.

Cox e Reid (1987) apresentou uma verossimilhança perfilada modificada que não é invariante

sob reparametrizações e requer que os parâmetros sejam ortogonais. Entretanto, em casos particulares em que não há ortogonalidade, a obtenção de uma reparametrização ortogonal pode ser muito complicada. Assim, Cox e Reid (1993) obtiveram uma aproximação para o ajuste de Cox e Reid (1987) que não exige que os parâmetros de interesse e os de incômodo sejam ortogonais, para o caso em que o parâmetro de interesse é escalar.

Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998) apresentaram aproximações para a função de verossimilhança perfilada modificada proposta por Barndorff-Nielsen (1983). Fraser e Reid (1995), Fraser, Reid e Wu (1999) derivaram aproximações para a derivada do espaço amostral do logaritmo da função de verossimilhança perfilada modificada de Barndorff-Nielsen (1983) que não requer a determinação de uma estatística ancilar. Severini (1998) derivou uma aproximação que é conveniente em situações em que há dificuldade para calcular a esperança do produto de derivadas do logaritmo da função de verossimilhança. A aproximação baseia-se no cálculo de uma matriz de covariâncias empíricas.

Em geral, as verossimilhanças perfiladas modificadas possuem propriedades mais próximas à de uma verossimilhança original. Ferguson, Reid e Cox (1991) apresentaram resultados em que a função escore baseada na proposta de Barndorff-Nielsen (1983) e Barndorff-Nielsen (1994) tem aproximadamente média zero.

Como exemplos de trabalhos publicados sobre ajustes, neste sentido, podemos mencionar: Ferrari, Silva e Cribari-Neto (2005), Cysneiros e Ferrari (2006), Ferrari, Silva e Cribari–neto (2007), Cysneiros, Cribari-Neto e Júnior (2008), Melo, Ferrari e Cribari-Neto (2009), Barreto, Cysneiros e Cribari-Neto (2013), Pires, Cysneiros e Cribari-Neto (2018) e Araújo, Cysneiros e Montenegro (2020).

Ferrari, Silva e Cribari-Neto (2005) obtiveram diferentes ajustes para a verossimilhança perfilada na família exponencial de dois parâmetros. Eles consideraram os ajustes baseados na proposta de Barndorff-Nielsen (1983), Cox e Reid (1987), Cox e Reid (1993), Severini (1999) e Severini (2000). Esses ajustes visam reduzir o impacto do parâmetro de incômodo na inferência em relação ao parâmetro de interesse.

Cysneiros e Ferrari (2006) desenvolveram uma correção de Bartlett para o teste de heteroscedasticidade baseado na estatística da razão de verossimilhanças perfiladas modificadas em modelos não-lineares da família exponencial, mais especificamente nos modelos normal e normal inverso. Estendendo, portanto, Ferrari e Cribari-Neto (2002) que é restrito ao modelo de regressão normal linear. Vale ressaltar aqui que o modelo é heteroscedástico se os parâmetros de dispersão são iguais para todas as observações.

Ferrari, Silva e Cribari–neto (2007) obtem os ajustes propostos por Barndorff-Nielsen (1983), Cox e Reid (1987), Fraser e Reid (1995), Fraser, Reid e Wu (1999), Severini (1998) e Severini (1999) para o parâmetro de forma da distribuição Weibull para dados completos, assim como, sob censuras do tipo I e tipo II.

Cysneiros, Cribari-Neto e Júnior (2008) desenvolveram ajustes para a função de verossimilhança perfilada na distribuição Birnbaum-Saunders. Eles consideraram os ajustes baseados nas propostas de Cox e Reid (1987) e Barndorff-Nielsen (1983). Obtiveram os estimadores de máxima verossimilhança perfilados modificados e os testes da razão de verossimilhanças perfilados modificados.

Melo, Ferrari e Cribari-Neto (2009) obtiveram fatores de correção de Bartlett para as estatísticas da razão de verossimilhanças perfiladas original e modificada baseada na proposta de Cox e Reid (1987), para o caso em que o interesse consiste em testar um vetor de parâmetros de efeitos fixos.

Barreto, Cysneiros e Cribari-Neto (2013) utilizou a verossimilhança perfilada e seus ajustes propostos por Barndorff-Nielsen (1983) e Cox e Reid (1987), Fraser e Reid (1995), Fraser, Reid e Wu (1999), Severini (1998) e Severini (1999) no aperfeiçoamento de inferências para a distribuição Birnbaum-Saunders sob censura tipo II.

Pires, Cysneiros e Cribari-Neto (2018) derivaram ajustes para a função de verossimilhança perfilada da distribuição Pareto generalizada, considerando as propostas de Barndorff-Nielsen (1983) e Cox e Reid (1993), Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1999). Obtiveram os estimadores de máxima verossimilhanças perfilados modificados e os testes da razão de verossimilhanças perfilados modificados e boostrap paramétrico.

Araújo, Cysneiros e Montenegro (2020) propuseram teste de heteroscedasticidade na classe dos modelos não-lineares simétricos heteroscedásticos. Desenvolveram fatores de correção de Bartlett para os testes da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas baseado na proposta por Cox e Reid (1987), a qual exige ortogonalidade entre parâmetros de interesse e incômodo.

O principal objetivo desse trabalho consiste na obtenção, mais especificamente, do ajuste proposto por Barndorff-Nielsen (1983) através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998) para a distribuição Burr XII de três parâmetros proposta por Zimmer, Keats e Wang (1998) e no modelo de regressão log-Burr XII proposto por Silva et al. (2008), considerando diferentes cenários em relação aos parâmetros de interesse e de incômodo.

Esta dissertação está dividida em seis capítulos. No segundo capítulo, resumimos a teoria dos diferentes ajustes para a função de verossimilhança perfilada. Mais especificamente, a ênfase será dada aos ajustes propostos por Barndorff-Nielsen (1983), Cox e Reid (1987), Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998).

No terceiro capítulo, apresentamos uma pequena introdução sobre a distribuição Burr XII, bem como, algumas características, propriedades e aspectos inferenciais. Derivamos os ajustes para a função de verossimilhanca perfilada modificada, baseados nas propostas de Barndorff-Nielsen (1983), Cox e Reid (1987). Para a verossimilhança perfilada modificada de Barndorff-Nielsen (1983) utilizamos as aproximações propostas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998) à distribuição Burr XII. Obtivemos os estimadores de verossimilhança perfilado modificados, assim como, os testes da razão de verossimilhanças perfiladas modificados.

No quarto capítulo, derivamos os ajustes para a função de verossimilhança perfilada modificada, baseado na proposta de Barndorff-Nielsen (1983) através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998) ao modelo de regressão log-Burr XII. Obtivemos os estimadores de verossimilhança perfilado modificados, assim como, os testes da razão de verossimilhanças perfiladas modificados. Avaliamos numericamente os desempenhos dos estimadores pontuais e os testes da razão de verossimilhanças perfiladas modificados em amostras finitas.

No quinto capítulo, aplicamos alguns dos resultados obtidos neste trabalho a cinco conjuntos de dados reais e comparamos os resultados obtidos com relação à estimação dos parâmetros e a testes de hipóteses. Por fim, no sexto capítulo, apresentamos as conclusões finais referente aos resultados obtidos neste trabalho.

Na literatura, não há nada em relação a verossimilhança perfilada na distribuição Burr XII, assim como, no modelo de regressão log-Burr XII. Nosso trabalho veio a preencher parcialmente esta lacuna. Vale salientar que, os capítulos três e quatro são independentes e podem ser lidos separadamente. Dessa maneira, algumas expressões e notações são apresentadas mais de uma vez.

2 AJUSTES À VEROSSSIMILHANÇA PERFILADA

Neste capítulo, faremos uma revisão sobre a função de verossimilhança perfilada e seus ajustes. Os ajustes abordados aqui foram propostos por Barndorff-Nielsen (1983) e Cox e Reid (1987). Para a verossimihança perfilada modificada por Barndorff-Nielsen (1983), usaremos as aproximações sugeridas por Fraser e Reid (1995), Severini (1998) e Fraser, Reid e Wu (1999). Daremos ênfase ao teste da razão de verossimilhanças.

2.1 DEFINIÇÕES PRELIMINARES

Seja o vetor $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$, o valor observado da variável aleatória $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$ de um experimento aleatório representado por um espaço de probabilidade $(\Omega, \mathscr{F}, P)$, em que Ω é o conjunto de todos os resultados possíveis, $\mathscr{F}$ é uma σ-álgebra de subconjunto de Ω e P uma medida de probabilidade definida em $\mathscr{F}$. A variável $\boldsymbol{Y}$ é caracterizada por uma função de densidade ou de probabilidade na forma analítica $f(\boldsymbol{y}; \boldsymbol{\theta})$ conhecida, de parâmetros desconhecidos $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)^\top$. O espaço paramétrico $\Theta \subseteq I\!R^p$ representa o conjunto de possíveis valores do vetor $\boldsymbol{\theta}$. A função $f(\boldsymbol{y}; \boldsymbol{\theta})$ denominada de função modelo estatístico, que define alguma família $\mathscr{F}$ de distribuição de probabilidade.

A definição da função de verossimilhança é igual à função modelo, mas interpretada de maneira diferente, como função de $\boldsymbol{\theta}$ para $\boldsymbol{y}$ conhecido. Desta maneira

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}; \boldsymbol{y}) = f(\boldsymbol{y}; \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \Theta.$$

Seja $Y_1, \ldots, Y_n$ variáveis aleatórias independentes com função de distribuição acumulada (f.d.a.) $F(y_1, \ldots, y_n|\boldsymbol{\theta})$, em que, $\boldsymbol{\theta} \in \Theta$ é desconhecido. A função de verossimilhança é $L(\boldsymbol{\theta}) = f(y_1, \ldots, y_n|\boldsymbol{\theta})$. Observe que, se $Y_1, \ldots, Y_n$ são variáveis aleatórias independentes e identicamente distribuídos, então $L(\boldsymbol{\theta})$ pode ser definido como

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}; \boldsymbol{y}) = \prod_{i=1}^{n} f(y_i; \boldsymbol{\theta}).$$

O estimador de máxima verossimilhança $\widehat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$ é o valor em Θ que maximiza a função de verossimilhança $L(\boldsymbol{\theta}; \boldsymbol{y})$. Assim, em muitas situações, é mais conveniente trabalhar com o logaritmo da função de verossimilhança de $\boldsymbol{\theta}$, denotada por

$$\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} \log\{f(y_i; \boldsymbol{\theta})\}.$$

Como a função logaritmo é monótona, o valor de $\boldsymbol{\theta}$ que maximiza a função de verossimilhança $L(\boldsymbol{\theta}; \boldsymbol{y})$ é o mesmo que maximiza $\ell(\boldsymbol{\theta}; \boldsymbol{y})$. Assim, o estimador de máxima verossimilhança é, em muitos casos, raiz da equação de verossimilhança $\ell'(\widehat{\boldsymbol{\theta}}) = 0$. Em casos mais simples, a solução desta equação possui forma analítica. Contudo, em casos mais complexos, tal solução dever ser obtida numericamente. A função escore $u(\boldsymbol{\theta}) = (u_1(\boldsymbol{\theta}), \ldots, u_p(\boldsymbol{\theta}))^\top$ é definida como

$$u_i(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_i}, \quad i = 1, \ldots, p.$$

Para garantir que o estimador de máxima verossimilhança de $\boldsymbol{\theta}$ seja consistente e possua distribuição limite conhecida, assumimos que a função de verossimilhança satisfaz às condições de regularidade, isto implica que $E[u(\boldsymbol{\theta})] = 0$, isto é, o valor esperado da função escore é sempre igual a zero. Outro resultado importante mostra que

$$E\left[\frac{\partial \ell}{\partial \boldsymbol{\theta}} \frac{\partial \ell}{\partial \boldsymbol{\theta}^\top}\right] = -E\left[\frac{\partial^2 \ell}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\right],$$

sendo esta expressão conhecida como igualdade da informação.

Seja $j(\boldsymbol{\theta})$, a matriz de informação observada, dada por

$$j = j(\boldsymbol{\theta}) = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top},$$

e $i(\boldsymbol{\theta})$ é a matriz de informação esperada ou matriz de informação de Fisher, definida por

$$i = i(\boldsymbol{\theta}) = E[j(\boldsymbol{\theta})].$$

Suponha que o vetor $\boldsymbol{\theta}$ pode ser decomposto como $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\phi})^\top$, em que $\boldsymbol{\mu}$ tem dimensão r e $\boldsymbol{\phi}$ tem dimensão $p - r$. Deste modo, a matriz de informação observada particionada é dada por

$$j = j(\boldsymbol{\theta}) = \begin{pmatrix} j_{\boldsymbol{\mu}\boldsymbol{\mu}} & j_{\boldsymbol{\mu}\boldsymbol{\phi}} \\ \\ j_{\boldsymbol{\phi}\boldsymbol{\mu}} & j_{\boldsymbol{\phi}\boldsymbol{\phi}} \end{pmatrix},$$

em que

$$j_{\boldsymbol{\mu}\boldsymbol{\mu}} = -\frac{\partial^2 \ell(\boldsymbol{\mu}, \boldsymbol{\phi})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^\top}, \quad j_{\boldsymbol{\mu}\boldsymbol{\phi}} = j_{\boldsymbol{\phi}\boldsymbol{\mu}}^\top = -\frac{\partial^2 \ell(\boldsymbol{\mu}, \boldsymbol{\phi})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\phi}^\top} \quad \text{e} \quad j_{\boldsymbol{\phi}\boldsymbol{\phi}} = -\frac{\partial^2 \ell(\boldsymbol{\mu}, \boldsymbol{\phi})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^\top}.$$

De maneira semelhante, a matriz de informação esperada é dada por

$$i = i(\boldsymbol{\theta}) = \begin{pmatrix} i_{\boldsymbol{\mu}\boldsymbol{\mu}} & i_{\boldsymbol{\mu}\boldsymbol{\phi}} \\ i_{\boldsymbol{\phi}\boldsymbol{\mu}} & i_{\boldsymbol{\phi}\boldsymbol{\phi}} \end{pmatrix},$$

em que

$$i_{\boldsymbol{\mu}\boldsymbol{\mu}} = -E\left[\frac{\partial^2 \ell(\boldsymbol{\mu},\boldsymbol{\phi})}{\partial \boldsymbol{\mu}\partial \boldsymbol{\mu}^\top}\right], \quad i_{\boldsymbol{\mu}\boldsymbol{\phi}} = i_{\boldsymbol{\phi}\boldsymbol{\mu}}^\top = -E\left[\frac{\partial^2 \ell(\boldsymbol{\mu},\boldsymbol{\phi})}{\partial \boldsymbol{\mu}\partial \boldsymbol{\phi}^\top}\right] \quad \text{e} \quad i_{\boldsymbol{\phi}\boldsymbol{\phi}} = -E\left[\frac{\partial^2 \ell(\boldsymbol{\mu},\boldsymbol{\phi})}{\partial \boldsymbol{\phi}\partial \boldsymbol{\phi}^\top}\right].$$

Muitas vezes é de interesse testar hipóteses sobre uma parte do vetor de parâmetros $\boldsymbol{\theta}$, isto é, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ contra $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$. Neste caso, considere $L(\boldsymbol{\mu}, \boldsymbol{\phi})$ a função de verossimilhança e $\ell(\boldsymbol{\mu}, \boldsymbol{\phi})$ o logaritmo da função de verossimilhança, a estatística da razão de verossimilhanças para testar a hipótese nula H_0 contra H_1 é expressa por

$$\lambda = \frac{L(\boldsymbol{\mu}_0, \tilde{\boldsymbol{\phi}})}{L(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\phi}})},$$

em que, $\widehat{\boldsymbol{\mu}}$ e $\widehat{\boldsymbol{\phi}}$ são os estimadores de máxima verossimilhança irrestritos de $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ e $\tilde{\boldsymbol{\phi}}$ é o estimador de máxima verossimilhança restrito de $\boldsymbol{\phi}$. Quando não conseguimos explicitar a região crítica através de uma estatística com distribuição conhecida, sob H_0, utiliza-se uma conveniente transformação de λ definida por

$$LR = 2\left\{\ell(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\phi}}) - \ell(\boldsymbol{\mu}_0, \tilde{\boldsymbol{\phi}})\right\},$$

que têm, assintoticamente e sob certas condições de regularidade, distribuição χ^2 com r graus de liberdade igual à dimensão do vetor $\boldsymbol{\mu}$. Assim, rejeita-se a hipótese nula quando LR for maior que o quantil de nível $1-\alpha$ da distribuição χ^2_r, sendo α o nível de significância adotado no teste. A aproximação pela distribuição χ^2_r tem erro de ordem $O(n^{-1})$. [1]

2.2 VEROSSIMILHANÇA PERFILADA

No caso de modelos com parâmetros de incômodo, costuma-se fazer inferência usando a verossimilhança perfilada. Adotando a decomposição $\boldsymbol{\theta} = (\boldsymbol{\mu}^\top, \boldsymbol{\phi}^\top)^\top$, o vetor de parâmetros

[1] Se $\{a_n\}_{n\geq 1}$ e $\{b_n\}_{n\geq 1}$ são duas sequências de números reais, dizemos que a_n é de ordem menor que b_n e escrevemos $a_n = o(b_n)$ se $\lim_{n\to\infty} a_n/b_n = 0$. Dizemos que a_n é de ordem no máximo igual a b_n, denotado por $a_n = O(b_n)$, se existe um número real $M > 0$ tal que $|a_n/b_n| \leq M$, isto é, a razão $|a_n/b_n|$ é limitada. Observamos que se $a_n = o(b_n)$, então $a_n = O(b_n)$ e que quando $b_n \to 0$ a ordem fornece noção da taxa de convergência de a_n para zero.

particionado nos vetores $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ são comumente denominados de parâmetros de interesse e de incômodo, respectivamente.

A função de verossimilhança perfilada, em geral, é obtida substituindo o seu vetor de parâmetros de incômodo $\boldsymbol{\phi}$ pelo seu respectivo estimador de máxima verossimilhança para valores fixados do parâmetro de interesse $\boldsymbol{\mu}$ denotado por $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}$ em sua função de verossimilhança genuína.

Desta maneira, podemos escrever $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\mu}} = (\boldsymbol{\mu}, \widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}})$, em que, $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}$ é a solução em $\boldsymbol{\phi}$ de $\frac{\partial \ell(\boldsymbol{\mu},\boldsymbol{\phi})}{\partial \boldsymbol{\phi}} = 0$, de maneira que, a função de verossimilhança perfilada é definida por

$$L_p(\boldsymbol{\mu}) = L(\boldsymbol{\mu}, \widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}).$$

Sendo a estimativa $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}$ única, o logaritmo da função de verossimilhança perfilada para $\boldsymbol{\mu}$ é denotada como

$$\ell_p(\boldsymbol{\mu}) = \ell(\boldsymbol{\mu}, \widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}; y) = \sup_{\boldsymbol{\phi}} \ell(\boldsymbol{\mu}, \boldsymbol{\phi}; y).$$

O processo de maximização descrito acima, sugere ser realizado em duas etapas. A primeira consiste em encontrar o valor único de $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}$ que maximiza $\ell(\boldsymbol{\mu}, \boldsymbol{\phi})$ com relação a $\boldsymbol{\phi}$ com $\boldsymbol{\mu}$ fixo. A segunda etapa é encontrar o valor de $\boldsymbol{\mu}$ que maximiza $\ell_p(\boldsymbol{\mu}) = \log L_p(\boldsymbol{\mu})$.

Desta maneira, o estimador de máxima verossimilhança perfilado $\widehat{\boldsymbol{\mu}}_p$ é obtido como solução da equação

$$\frac{\partial \ell_p(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} = 0.$$

A matriz de informação observada perfilada em relação ao parâmetro de interesse $\boldsymbol{\mu}$ é definida de maneira análoga à matriz de informação observada $j(\boldsymbol{\theta})$, ou seja,

$$j_p = j_p(\boldsymbol{\mu}) = -\frac{\partial^2 \ell_p(\boldsymbol{\mu})}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^\top}.$$

Desta maneira, a função de máxima verossimilhança $L_p(\boldsymbol{\mu})$ não deve ser tratada como uma verossimilhança genuína. Em particular, algumas propriedades válidas para a função de verossimilhança não são mais satisfeitas para a verossimilhança perfilada, como as propriedades que garantem que o valor esperado da função escore seja sempre nula e que a identidade da informação não apresente viés. Ou seja, as igualdades

$$E\left[u_p(\boldsymbol{\mu})\right] = 0 \quad \text{e} \quad E\left[u_p(\boldsymbol{\mu}) u_p^\top(\boldsymbol{\mu})\right] + E\left[\frac{\partial u_p(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}\right] = 0,$$

em que, $u_p(\boldsymbol{\mu}) = \partial \ell_p(\boldsymbol{\mu})/\partial \boldsymbol{\mu}$ é a função escore perfilada relacionada a $L_p(\boldsymbol{\mu})$, não são sempre satisfeitas para a função de verossimilhança perfilada.

A estatística da razão de verossimilhanças perfiladas para testar hipóteses sobre $\boldsymbol{\mu}$, isto é, testar $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ contra $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, é

$$LR_p = 2\{\ell_p(\widehat{\boldsymbol{\mu}}) - \ell_p(\boldsymbol{\mu}_0)\} = 2\{\ell(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}) - \ell(\boldsymbol{\mu}_0, \widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}})\}.$$

Quando as condições de regularidade são satisfeitas, a distribuição assintótica de LR_p sob a hipótese nula é uma distribuição χ^2_r, em que, o número de graus de liberdade r é a dimensão do vetor $\boldsymbol{\mu}$. É importante observar que a estatística da razão de verossimilhanças LR é proporcional à diferença entre os logaritmos de duas funções de verossimilhanças perfiladas maximizadas.

Devido a sua irrestrita aplicabilidade, a verossimilhança perfilada desempenha um papel importante na inferência estatística, relacionados a problemas com parâmetros de incômodo. Sendo uma ferramenta atraente em análises inferenciais sobre parâmetros de interesse $\boldsymbol{\mu}$ pelo fato de possuir algumas semelhanças com a função de verossimilhança verdadeira. Muitas vezes, usar $L_p(\boldsymbol{\mu})$ e como se $\boldsymbol{\phi}$ fosse conhecido e igual a $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}$ não é adequado em especial quando os dados não contêm muita informação sobre $\boldsymbol{\phi}$, fato que frequentemente ocorre quando a dimensão do parâmetro $\boldsymbol{\phi}$ é grande. A presença de um número considerável de parâmetros de incômodo também deteriora a qualidade das aproximações envolvidas nas inferências que se baseiam em resultados assintóticos. Por esta razão existe uma necessidade de ajustar a função de verossimilhança perfilada de alguma maneira.

2.3 AJUSTES PARA VEROSSIMILHANÇA PERFILADA PROPOSTA POR BARNDORFF-NIELSEN

A verossimilhança perfilada modificada proposta por Barndorff-Nielsen (1983) pode ser derivada como uma aproximação de uma verossimilhança marginal ou de uma verossimilhança condicional para $\boldsymbol{\mu}$, se algumas dessas funções existir. Em ambos os casos, utiliza-se uma aproximação para função densidade do estimador de máxima verossimilhança condicionado a uma estatística ancilar. O logaritmo da função de verossimilhança perfilada modificada pode ser expressa como

$$\ell_{BN}(\boldsymbol{\mu}) = \ell_p(\boldsymbol{\mu}) - \log\left|\frac{\partial\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}}{\partial\widehat{\boldsymbol{\phi}}}\right| - \frac{1}{2}\log|j_{\phi\phi}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}};\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}},a)|, \tag{2.1}$$

em que $j_{\phi\phi}(\boldsymbol{\mu},\boldsymbol{\phi};a) = -\partial^2\ell/\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}^\top$ é a matriz de informação observada com relação a $\boldsymbol{\phi}$, $\ell(\cdot)$ é o logaritmo da função de verossimilhança e $\partial\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}/\partial\widehat{\boldsymbol{\phi}}$ corresponde a matriz de derivadas parciais de $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}$ com relação a $\boldsymbol{\phi}$.

Uma expressão alternativa para o logaritmo da função de verossimilhança em (2.1) que não envolve o termo $|\partial\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}/\partial\widehat{\boldsymbol{\phi}}|$ é expressa por

$$\ell_{BN}(\boldsymbol{\mu}) = \ell_p(\boldsymbol{\mu}) + \frac{1}{2}\log|j_{\phi\phi}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}};\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}},a)| - \log|\ell_{\boldsymbol{\phi};\widehat{\boldsymbol{\phi}}}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}};\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}},a)|,$$

em que $\ell_{\boldsymbol{\phi};\widehat{\boldsymbol{\phi}}}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}};\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}},a) = \partial^2\ell(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}};\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}},a)/\partial\boldsymbol{\phi}\partial\widehat{\boldsymbol{\phi}}^\top$. Observe, que a última expressão envolve derivadas com respeito ao espaço amostral além de uma especificação de uma estatística ancilar a de modo que $(\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}},a)$ seja uma estatística suficiente minimal. Entretanto, em casos particulares, a obtenção dessa quantidade pode ser bastante complicada.

Algumas aproximações foram propostas para o logaritmo da função de verossimilhança de Barndorff-Nielsen (1983).

Fraser e Reid (1995) e Fraser, Reid e Wu (1999) propuseram uma aproximação para a função de verossimilhança perfilada modificada de Barndorff-Nielsen (1983) que não envolve uma estatística ancilar. Neste caso, o logaritmo da função de verossimilhança é dada por

$$\breve{\ell}_{BN}(\boldsymbol{\mu}) = \ell_p(\boldsymbol{\mu}) + \frac{1}{2}\log|j_{\phi\phi}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}})| - \log|\ell_{\boldsymbol{\phi};Y}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}})\widehat{V}_{\boldsymbol{\phi}}|,$$

em que $\ell_{\boldsymbol{\phi};Y}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}) = \partial\ell_{\boldsymbol{\phi}}(\boldsymbol{\mu},\boldsymbol{\phi})/\partial\boldsymbol{Y}^\top$, $\ell_{\boldsymbol{\phi}}(\boldsymbol{\mu},\boldsymbol{\phi})$ é a função escore avaliada em $\boldsymbol{\phi}$, $\boldsymbol{Y}^\top = (y_i,\ldots,y_n)$ é a amostra observada,

$$\widehat{V}_{\boldsymbol{\phi}} = \Big(-\frac{\partial F(y_1;\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}})/\partial\widehat{\boldsymbol{\phi}}}{f(y_1;\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}})},\ldots,-\frac{\partial F(y_n;\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}})/\partial\widehat{\boldsymbol{\phi}}}{f(y_n;\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}})}\Big)^\top,$$

em que $f(y_j;\boldsymbol{\mu},\boldsymbol{\phi})$ é a função densidade de probabilidade avaliada em y_j e $F(y_j;\boldsymbol{\mu},\boldsymbol{\phi})$ é a função de distribuição acumulada. O estimador de máxima verossimilhança será denotado por $\widehat{\breve{\boldsymbol{\mu}}}_{BN}$

Outro método de aproximação é baseado na proposta de Severini (1998). Logo, uma aproximação para a verossimilhança proposta por Barndorff-Nielsen é expressa por

$$\bar{\ell}_{BN}(\boldsymbol{\mu}) = \ell_p(\boldsymbol{\mu}) + \frac{1}{2}\log|j_{\phi\phi}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}})| - \log|I_{\boldsymbol{\phi}}(\boldsymbol{\mu},\widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}};\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\phi}})|.$$

em que

$$I_{\phi}(\boldsymbol{\mu}, \widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}; \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\phi}}) = \sum_{i=1}^{n} \ell_{\phi}^{(i)}(\boldsymbol{\mu}, \widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}}) \ell_{\phi}^{(i)}(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\phi}})^{\top},$$

sendo $\ell_{\phi}^{(i)}(.)$ a função escore avaliada em $\boldsymbol{\phi}$ baseada na i-ésima observação. O estimador de máxima verossimilhança será denotado por $\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$.

2.4 VEROSSIMILHANÇA PERFILADA MODIFICADA PROPOSTA POR COX E REID

O logaritmo da função de verossimilhança proposta por Cox e Reid (1987) é dado por

$$\ell_{CR}(\boldsymbol{\mu}) = \ell_{p}(\boldsymbol{\mu}) - \frac{1}{2}|j_{\phi\phi}(\boldsymbol{\mu}, \widehat{\boldsymbol{\phi}}_{\boldsymbol{\mu}})|,$$

em que $j_{\phi\phi}(\boldsymbol{\mu}, \widehat{\boldsymbol{\phi}})$ é a matriz de informação observada relacionada a $\boldsymbol{\phi}$ quando $\boldsymbol{\mu}$ é fixo. A utilização dessa função necessita da ortogonalidade dos parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ e não é invariante sob reparametrizações. Quando os parâmetros não são ortogonais é necessário obter uma reparametrização, para que, $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ sejam ortogonais. Entretanto, em muitos casos não é possível obter tal reparametrização, exceto nos casos em que o parâmetro de interesse é escalar. Vale ressaltar, que essa reparametrização está associada a solução de equações diferenciais que envolvem elementos da matriz de informação de Fisher, que em casos particulares, se torna bem difícil tal reparametrização.

Devido a essas dificuldades, Cox e Reid (1993) propuseram uma aproximação que não exige ortogonalidade para o caso em que o parâmetro de interesse é escalar. É possível escrever o logaritmo da função de verossimilhança proposta por Cox e Reid (1993) como

$$\ell_{CRa}(\mu) = \ell_{p}(\mu) - \frac{1}{2}\log|j_{\phi\phi}(\mu, \widehat{\boldsymbol{\phi}}_{\mu})| - (\mu - \widehat{\mu}) m_{\boldsymbol{\phi}}(\widehat{\mu}, \widehat{\boldsymbol{\phi}}),$$

em que $m_{\boldsymbol{\phi}}(\widehat{\mu}, \widehat{\boldsymbol{\phi}})$ é o traço da matriz $\partial m / \partial \boldsymbol{\phi}$. Temos que, $m = i^{\phi\phi} i_{\mu\boldsymbol{\phi}}$, $i^{\phi\phi}$ denota a inversa da matriz $i_{\boldsymbol{\phi}\boldsymbol{\phi}} = -E[\partial^2 \ell(\mu, \boldsymbol{\phi}) / \partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^{\top}]$ e $i_{\mu\boldsymbol{\phi}} = -E[\partial^2 \ell(\mu, \boldsymbol{\phi}) / \partial \mu \partial \boldsymbol{\phi}]$. Entretanto, essa aproximação permanece não invariante sob reparametrizações. O estimador de máxima verossimilhança será denotado por $\widehat{\mu}_{CRa}$.

Para uma análise mais detalhada sobre as aproximações da função de verossimilhança perfilada proposta por Barndorff-Nielsen (1983), consultar Pace e Savan (1997) e Severini (2000).

2.5 TESTE DE HIPÓTESES

As estatísticas da razão de verossimilhanças baseadas nas propostas de Fraser e Reid (1995), Fraser, Reid e Wu (1999), Severini (1998), e Cox e Reid (1987) para testar

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \quad \text{contra} \quad H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$$

são dadas, respectivamente, por

$$\begin{aligned} LR_{BN1} &= 2\{\breve{\ell}_{BN}(\widehat{\breve{\boldsymbol{\mu}}}) - \breve{\ell}_{BN}(\boldsymbol{\mu}_0)\}, \\ LR_{BN2} &= 2\{\bar{\ell}_{BN}(\widehat{\bar{\boldsymbol{\mu}}}) - \bar{\ell}_{BN}(\boldsymbol{\mu}_0)\} \quad \text{e} \\ LR_{CR} &= 2\{\ell_{CR}(\widehat{\boldsymbol{\mu}}) - \ell_{CR}(\boldsymbol{\mu}_0)\}. \end{aligned}$$

Sob condições de regularidade e sob a hipótese nula, as estatísticas definidas acima têm distribuição χ^2_r. Mesmo com a modificação a aproximação pela distribuição χ^2_r tem erro de ordem $O(n^{-1})$.

3 DISTRIBUIÇÃO BURR XII

3.1 INTRODUÇÃO

A distribuição Burr XII com três parâmetros proposta por Zimmer, Keats e Wang (1998) vem sendo aplicada em diversas áreas do conhecimento, a citar: Shao, Wong e Xia (2004) aplicaram a distribuição Burr XII em dados relativos a enchente. Tejeda e Goodwin (2008) aplicaram a distribuição Burr XII em dados agrícolas dos Estados Unidos. Okasha e Matter (2015) utilizaram a distribuição Burr XII para modelar dados de pacientes com câncer de mama.

Neste capítulo apresentaremos a distribuição Burr XII de três parâmetros, incluindo a sua função densidade, função de distribuição acumulada, o valor esperado, a variância, a assimetria, a curtose e alguns aspectos inferenciais. Desenvolvemos ajustes para função de verossimilhança perfilada para os parâmetros de forma e escala da distribuição. Enfoque tem sido dado também nas aproximações da log-verossimilhança de Barndorff-Nielsen (1983) baseadas nas propostas de Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Vale ressaltar, que não existem trabalhos publicados sobre refinamentos assintóticos na distribuição Burr XII. Neste sentido, nosso trabalho contribui para preencher essa lacuna. Além disso, apresentaremos os estimadores de máxima verossimilhança perfilado e perfilados modificados, mostraremos como ficam os testes baseados nas estatísticas da razão de verossimilhanças provenientes das verossimilhanças perfilada e perfiladas modificadas. Avaliaremos o desempenho dos estimadores de máxima verossimilhança perfilado e perfilados modificados, assim como, os testes baseados na razão de verossimilhanças perfiladas e perfiladas modificadas em relação ao tamanho e poder.

3.2 DEFINIÇÕES BÁSICAS

Seja Y uma variável aleatória da distribuição Burr XII, denotada por $Y \sim \mathrm{Burr}(c, k, s)$, então a função densidade de probabilidade (f.d.p.) é dada por

$$f(y; c, k, s) = cks^{-c}y^{c-1}\left[1 + \left(\frac{y}{s}\right)^{c}\right]^{-(k+1)}, \quad 0 < y < \infty, \tag{3.1}$$

em que, $k, c > 0$ são parâmetros de forma e $s > 0$ é parâmetro de escala. As correspondentes

funções de distribuição acumulada (f.d.a.) e de risco (f.r.) são, respectivamente, dadas por

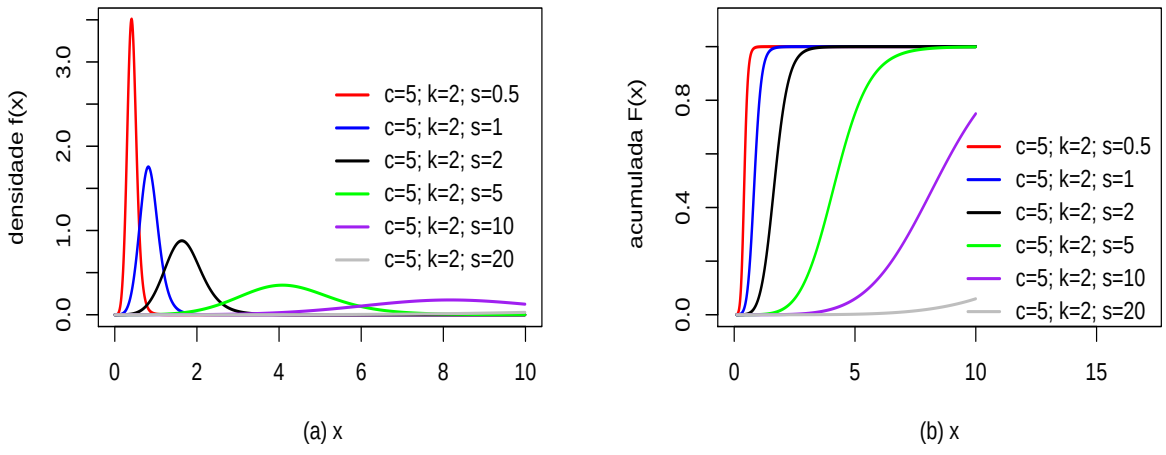
$$F(y;c,k,s) = 1 - \left[1 + \left(\frac{y}{s}\right)^{c}\right]^{-k}, \quad 0 < y < \infty, \quad c,k,s > 0,$$

e

$$h(y;c,k,s) = \frac{ck\left(\frac{y}{s}\right)^{c-1}}{s\left[1 + \left(\frac{y}{s}\right)^{c}\right]} \quad 0 < y < \infty, \quad c,k,s > 0.$$

Para entender o comportamento dessa distribuição de probabilidade, alguns gráficos da função densidade, função de distribuição acumulada e da função de risco foram obtidos e encontram-se nas Figuras 1 e 2. Pode-se observar que esse modelo pode comportar dados do tipo assimétricos à direita. Nota-se que, para diferentes valores do parâmetro de escala s com $c = 5$ e $k = 2$ fixos, a função densidade fica mais "achatada". Observamos que a distribuição acumulada comporta-se de diferentes formas para valores diferentes de s. A função de risco é decrescente quando $c \leq 1$ e unimodal quando $c > 1$. Além disso, para valores de c entre 1 e 2, a função de risco pode ser, praticamente, constante sob quase toda amplitude da função dependendo do valor de s.

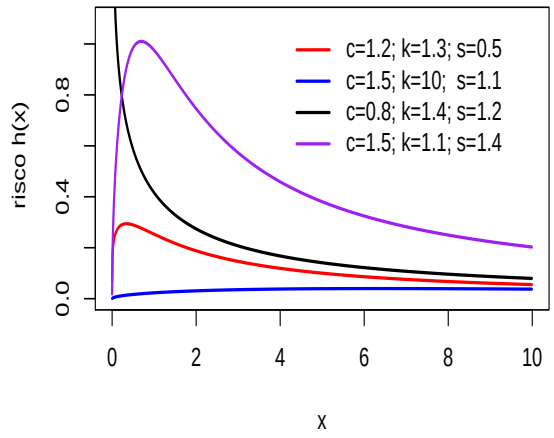
Figura 1 – Gráficos das funções de densidade (a) e distribuição acumulada (b) para alguns valores dos parâmetros da distribuição Burr XII.



Fonte: Próprio autor (2021).

O r-ésimo momento ordinário da distribuição Burr XII é obtido pela expressão $E(Y^r) = \int_0^\infty y^r f(y)dy$, em que $f(y)$ é a f.d.p. da distribuição Burr XII.

Figura 2 – Gráfico da função de risco para alguns valores dos parâmetros da distribuição Burr XII.



Fonte: Próprio autor (2021).

$$E(Y^r) = \mu^r = ks^r B\left(\frac{r}{c} + 1, k - \frac{r}{c}\right), \quad s, c > r,$$

em que $B(\cdot)$ é a função beta completa (LAWLESS, 2003).

Usando a expressão acima, o valor esperado, a variância, a assimetria e a curtose são dadas por

$$E(Y) = \mu^1 = \frac{s}{c} B\left(\frac{1}{c}, k - \frac{1}{c}\right), \quad c, s > 1,$$

$$Var(Y) = \mu^2 - (\mu^1)^2 = \frac{2s^2}{c} B\left(\frac{2}{c}, k - \frac{2}{c}\right) - \frac{s^2}{c^2} B\left(\frac{1}{c}, k - \frac{1}{c}\right)^2, \quad c, s > 2,$$

$$\text{assimetria} = 2(\mu^1)^3 - 3\mu^1\mu^2 + \mu^3,$$

$$\text{curtose} = -3(\mu^1)^4 + 6(\mu^1)^2\mu^2 - 4\mu^1\mu^3 + \mu^4.$$

A geração de números pseudo-aleatório da distribuição Burr XII é obtido por meio do método da inversão, desta maneira obtemos a expressão $Q(y; c, k, s) = s[((1-u)^{-1/k} - 1)^{1/c}]$. Para mais detalhes sobre o método da inversão ver (GENTLE, 2004).

3.3 ESTIMAÇÃO DOS PARÂMETROS

Seja $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$ uma amostra aleatória de tamanho n proveniente de uma distribuição Burr XII com f.d.p. dada em (3.1). O logaritmo da função de verossimilhança (log-verossimilhança) para $\boldsymbol{\theta} = (c, k, s)^\top$, pode ser escrito como

$$\ell(\boldsymbol{\theta}) = n[\log(c) + \log(k) - c\log(s)] + (c-1)\sum_{i=1}^{n}\log(y_i) - (k+1)\sum_{i=1}^{n}\log[1+(y_i/s)^c]. \quad (3.2)$$

O vetor escore $u(\boldsymbol{\theta}) = (u_c, u_k, u_s)^\top$ é obtido pela diferenciação da função log-verossimilhança dada em (3.2) em relação aos parâmetros desconhecidos. Dessa maneira, seus componentes são dados, respectivamente, por

$$u_c = \frac{\partial \ell(\boldsymbol{\theta})}{\partial c} = \frac{n}{c} - n\ log(s) - (k+1)\sum_{i=1}^{n}\left[\frac{(y_i/s)^c \log(y_i/s)}{1+(y_i/s)^c}\right] + \sum_{i=1}^{n}\log(y_i),$$

$$u_k = \frac{\partial \ell(\boldsymbol{\theta})}{\partial k} = \frac{n}{k} - \sum_{i=1}^{n}\log[1+(y_i/s)^c]$$

e

$$u_s = \frac{\partial \ell(\boldsymbol{\theta})}{\partial s} = -\frac{nc}{s} + \frac{c(k+1)}{s}\sum_{i=1}^{n}\left[\frac{(y_i/s)^c}{1+(y_i/s)^c}\right].$$

Os EMV são obtidos a partir das soluções das equações $\partial\ell(\boldsymbol{\theta})/\partial c = 0$, $\partial\ell(\boldsymbol{\theta})/\partial k = 0$ e $\partial\ell(\boldsymbol{\theta})/\partial s = 0$. Contudo, os EMV $\widehat{c}$, $\widehat{k}$ e $\widehat{s}$ não possuem solução analítica e devem ser obtidos numericamente. Alternativamente, $\widehat{\boldsymbol{\theta}}$ pode ser obtido maximizando a expressão (3.2) através do método BFGS implementado na plataforma computacional Ox por meio do comando MaxBFGS.

A matriz de informação observada é dada por

$$j(\boldsymbol{\theta}) = \begin{pmatrix} j_{ss} & j_{sc} & j_{sk} \\ j_{cs} & j_{cc} & j_{ck} \\ j_{ks} & j_{kc} & j_{kk} \end{pmatrix},$$

em que

$$j_{ss} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial s \partial s^\top} = -\frac{nc}{s^2} + (k+1)\sum_{i=1}^{n}\left[\frac{c(y_i/s)^c(1+c+(y_i/s)^c)}{(s(1+(y_i/s)^c))^2}\right],$$

$$j_{sc} = j_{cs} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial s \partial c} = \frac{n}{s} - (k+1)\sum_{i=1}^{n}\left[\frac{(y_i/s)^c(1+(y_i/s)^c + c\log(y_i/s))}{s(1+(y_i/s)^c)^2}\right],$$

$$j_{ks} = j_{sk} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial k \partial s} = -\sum_{i=1}^{n}\left[\frac{c(y_i/s)^c}{s(1+(y_i/s)^c)}\right],$$

$$j_{ck} = j_{kc} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial c \partial k} = \sum_{i=1}^{n}\left[\frac{(y_i/s)^c(\log(y_i/s))}{(1+(y_i/s)^c)}\right],$$

$$j_{cc} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial c \partial c^\top} = \frac{n}{c^2} + (k+1)\sum_{i=1}^{n}\left[\frac{(y_i/s)^c \log^2(y_i/s)}{(1+(y_i/s)^c)^2}\right]$$

e

$$j_{kk} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial k \partial k^\top} = \frac{n}{k^2}.$$

A matriz de informação esperada, também conhecida como matriz de informação de Fisher, é dada por

$$i(\boldsymbol{\theta}) = \begin{pmatrix} \frac{k}{(cs)^2(k+2)} & \frac{k[\psi(k)-\psi(1)-1]+1}{s(k+2)} & -\frac{1}{cs(k+1)} \\ \frac{k[\psi(k)-\psi(1)-1]+1}{s(k+2)} & \frac{c^2[k[(\psi(k)-\psi(1)-1)^2+\psi'(k)+\psi'(1)]+2(\psi(k)-\psi(1))]}{k+2} & \frac{c[\psi(1)-\psi(k)+1]}{k+1} \\ -\frac{1}{cs(k+1)} & \frac{c[\psi(1)-\psi(k)+1]}{k+1} & \frac{1}{k^2} \end{pmatrix},$$

em que $\Psi(\cdot)$ é a função digama e $\Psi'(\cdot)$ é a primeira derivada da função digama.

Mais adiante, os elementos da matriz de informação observada e da matriz de informação de Fisher serão úteis para obter os diferentes ajustes para a função de verossimilhança perfilada.

Intervalos de confiança e testes de hipóteses podem ser obtidos usando o fato de que a distribuição assintótica para os EMV é a distribuição normal com média $\boldsymbol{\theta}$ e matriz de covariância dada pelo inverso da matriz de informação de Fisher (SEN; SINGER, 1993). A matriz de covariância assintótica é dada por $\boldsymbol{I}^{-1}(\boldsymbol{\theta})$ com $\boldsymbol{I}(\boldsymbol{\theta}) = E[\ddot{\boldsymbol{L}}(\boldsymbol{\theta})]$, sendo $\ddot{\boldsymbol{L}}(\boldsymbol{\theta}) = -\left[\partial^2 \ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top\right]$. Entretanto, em muitos casos é complicado obter a matriz de informação de Fisher $\boldsymbol{I}(\boldsymbol{\theta})$. Assim, pode-se usar a matriz hessiana $-\ddot{\boldsymbol{L}}(\boldsymbol{\theta})$, avalianda em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ que é um estimador consistente. Desta maneira, a distribuição assintótica para $\widehat{\boldsymbol{\theta}}$ é dada por $\widehat{\boldsymbol{\theta}}^\top \sim N_{(3)}\left\{\widehat{\boldsymbol{\theta}}^\top; \ddot{\boldsymbol{L}}(\boldsymbol{\theta})^{-1}\right\}$, em que $\ddot{\boldsymbol{L}}(\boldsymbol{\theta})$ é a matriz de informação observada.

3.4 VEROSSIMILHANÇA PERFILADA AJUSTADA PARA A DISTRIBUIÇÃO BURR XII

Nesta seção, consideraremos o problema em realizar inferências sobre os parâmetros da distribuição Burr XII por meio da função de verossimilhança perfilada. Também obteremos as funções de verossimilhanças perfiladas modificadas de Cox e Reid (1987) e Barndorff-Nielsen

(1983). Para este último, usaremos as aproximações propostas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Apresentaremos os EMV perfilados modificados e como ficam os testes da razão de verossimilhanças baseados nestas funções. Iremos considerar três situações, a saber: caso (i): k sendo o parâmetro de interesse e $(c, s)^\top$ sendo os parâmetros de incômodo; caso (ii): sendo $(c, k)^\top$ parâmetros de interesse e s o parâmetro de incômodo e caso (iii): sendo $(k, s)^\top$ parâmetros de interesse e c o parâmetro de incômodo.

3.5 CASO (i)

No primeiro caso, estamos considerando k como o parâmetro de interesse e como parâmetros de incômodo, s e c. Isto é, $\mu = k$ e $\boldsymbol{\phi} = (s, c)^\top$.

Fixando k e resolvendo a equação $u(\boldsymbol{\phi}) = 0$, obtemos os EMV restritos de $\boldsymbol{\phi}$ $(\widehat{\boldsymbol{\phi}}_k)$. Substituindo $\boldsymbol{\phi}$ por $\widehat{\boldsymbol{\phi}}$ na log-verossimilhança original, isto é, em (3.2), obtemos a log-verossimilhança perfilada para k. A log-verossimilhança perfilada é dada por

$$\ell_p(k) = n[\log(\widehat{c}_k) + \log(k) - \widehat{c}_k \log(\widehat{s}_k)] - (k+1)\sum_{i=1}^{n}\left[1 + \left(\frac{y_i}{\widehat{s}_k}\right)^{\widehat{c}_k}\right] + (\widehat{c}_k - 1)\sum_{i=1}^{n}\log(y_i).$$

O EMV perfilado de $\widehat{k}_p$ é obtido como solução de $u_p(k) = 0$ e não tem solução em forma fechada. Logo, para encontrarmos o valor de $\widehat{k}_p$ que maximiza $\ell_p(k)$ é necessário a utilização de um procedimento de otimização não-linear, como, por exemplo, o algoritmo de Newton (Newton-Raphson, escore de Fisher) ou um algoritmo quasi-Newton (BFGS). Para maiores detalhes sobre tais algoritmos ver Nocedal e Wright (2006).

A log-verossimilhança perfilada modificada proposta por Barndorff-Nielsen (1983) para o parâmetro k é dada por

$$\ell_{BN}(k) = \ell_p(k) + \frac{1}{2}\log|j_{\phi\phi}(k, \widehat{\boldsymbol{\phi}}_k; \widehat{k}, \widehat{\boldsymbol{\phi}}, a) - \log|\ell_{\boldsymbol{\phi};\widehat{\boldsymbol{\phi}}}(k, \widehat{\boldsymbol{\phi}}_k; \widehat{k}, \widehat{\boldsymbol{\phi}}, a)|,$$

em que a quantidade $j_{\phi\phi}(k, \widehat{\boldsymbol{\phi}}_k; \widehat{k}, \widehat{\boldsymbol{\phi}}, a)$ é o bloco da matriz de informação observada para os parâmetros de incômodo $(s, c)^\top$ e $\ell_{\boldsymbol{\phi};\widehat{\boldsymbol{\phi}}}(k, \widehat{\boldsymbol{\phi}}_k; \widehat{k}, \widehat{\boldsymbol{\phi}}, a)$ envolve derivadas sobre o espaço amostral de forma que sua obtenção é tipicamente complicada. Assim, é útil adotar uma aproximação para obtenção desta quantidade. Neste trabalho, derivamos para a distribuição Burr XII, as aproximações para a log-verossimilhança perfilada de Barndorff-Nielsen (1983) sugeridas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). A log-verossimilhança

perfilada ajustada proposta por Fraser e Reid (1995) e Fraser, Reid e Wu (1999) é dada por

$$\breve{\ell}_{BN}(k) = \ell_p(k) + \frac{1}{2}\log|j_{\phi\phi}(k,\widehat{\boldsymbol{\phi}}_k)| - \log|\ell_{\phi;Y}(k,\widehat{\boldsymbol{\phi}}_k)\widehat{V}_{\boldsymbol{\phi}}|, \tag{3.3}$$

em que $j_{\phi\phi}(k,\widehat{\boldsymbol{\phi}}_k)$ é o bloco da matriz de informação observada para os parâmetros de incômodo $(s,c)^\top$, $\ell_{\phi;Y}(k,\widehat{\boldsymbol{\phi}}_k) = \partial\ell_{\boldsymbol{\phi}}/\partial Y^\top$, $\ell_{\boldsymbol{\phi}}$ é a função escore avaliada em $\boldsymbol{\phi} = (s,c)^\top$, $Y^\top = (y_1,\ldots,y_n)$ e

$$\widehat{V}_{\boldsymbol{\phi}} = \left(-\frac{F(y_1;\widehat{k},\widehat{\boldsymbol{\phi}})/\partial\widehat{\boldsymbol{\phi}}}{f(y_1;\widehat{k},\widehat{\boldsymbol{\phi}})},\ldots,-\frac{F(y_n;\widehat{k},\widehat{\boldsymbol{\phi}})/\partial\widehat{\boldsymbol{\phi}}}{f(y_n;\widehat{k},\widehat{\boldsymbol{\phi}})}\right)^\top,$$

em que $F(y_j;k,\phi)$ é a distribuição acumulada de y_j e $f(y_j;k,\phi)$ é a função densidade de probabilidade de y_j, $j = 1,\ldots,n$.

Na distribuição Burr XII, as quantidades acima são dadas por

$$j_{\phi\phi}(k,\widehat{\boldsymbol{\phi}}_k) = \begin{pmatrix} j_{ss}(k,\widehat{\boldsymbol{\phi}}_k) & j_{sc}(k,\widehat{\boldsymbol{\phi}}_k) \\ j_{cs}(k,\widehat{\boldsymbol{\phi}}_k) & j_{cc}(k,\widehat{\boldsymbol{\phi}}_k) \end{pmatrix},$$

sendo

$$j_{cc}(k,\widehat{\boldsymbol{\phi}}_k) = \frac{n}{\widehat{c}_k^2} + (k+1)\sum_{i=1}^{n}\left[\frac{(y_i/\widehat{s}_k)^c\log^2(y_i/\widehat{s}_k)}{(1+(y_i/\widehat{s}_k)^{\widehat{c}_k})^2}\right],$$

$$j_{sc}(k,\widehat{\boldsymbol{\phi}}_k) = j_{cs}(k,\widehat{\boldsymbol{\phi}}_k) = \frac{n}{\widehat{s}_k} - (k+1)\sum_{i=1}^{n}\left[\frac{(y_i/\widehat{s}_k)^{\widehat{c}_k}(1+(y_i/\widehat{s}_k)^{\widehat{c}_k}+\widehat{c}_k\log(y_i/\widehat{s}_k))}{\widehat{s}_k(1+(y_i/\widehat{s}_k)^{\widehat{c}_k})^2}\right]$$

e

$$j_{ss}(k,\widehat{\boldsymbol{\phi}}_k) = -\frac{n\widehat{c}_k}{\widehat{s}^2} + (k+1)\sum_{i=1}^{n}\left[\frac{\widehat{c}_k(y_i/\widehat{s}_k)^{\widehat{c}_k}(1+\widehat{c}_k+(y_i/\widehat{s}_k)^{\widehat{c}_k})}{(\widehat{s}_k(1+(y_i/\widehat{s}_k)^{\widehat{c}_k}))^2}\right].$$

A quantidade $\ell_{\phi;Y}(k,\widehat{\boldsymbol{\phi}}_k)\widehat{V}_{\boldsymbol{\phi}}$ é dada por

$$\ell_{\boldsymbol{\phi};Y}(k,\widehat{\boldsymbol{\phi}}_k)\widehat{V}_{\boldsymbol{\phi}} = \begin{pmatrix} \ell_{s;Y}(k,\widehat{s}_k)\widehat{V}_s & \ell_{s;Y}(k,\widehat{s}_k)\widehat{V}_c \\ \ell_{c;Y}(k,\widehat{c}_k)\widehat{V}_s & \ell_{c;Y}(k,\widehat{c}_k)\widehat{V}_c \end{pmatrix},$$

em que

$$\ell_{s;Y}(k,\widehat{s}_k)\widehat{V}_s = \sum_{i=1}^{n}\left[\frac{c^2(k+1)(y_i/\widehat{s}_k)^{\widehat{c}_k-1}}{(\widehat{s}_k(1+(y_i/\widehat{s}_k)^{\widehat{c}_k}))^2}\times\frac{y_i}{\widehat{s}}\right],$$

$$\ell_{s;Y}(k,\widehat{s}_k)\widehat{V}_c = \sum_{i=1}^{n}\left[\frac{c^2(k+1)(y_i/\widehat{s}_k)^{\widehat{c}_k-1}}{(\widehat{s}_k(1+(y_i/\widehat{s}_k)^{\widehat{c}_k}))^2}\times-\frac{y_i\log(y_i/\widehat{s})}{\widehat{c}}\right],$$

$$\ell_{c;Y}(k,\widehat{c}_k)\widehat{V}_s = \sum_{i=1}^{n}\left[-\frac{(1+(y_i/\widehat{s}_k)^{\widehat{c}_k})(k(y_i/\widehat{s}_k)^{\widehat{c}_k}-1)+\widehat{c}_k(k+1)(y_i/\widehat{s}_k)^{\widehat{c}_k}\log(y_i/\widehat{s}_k)}{y_i(1+(y_i/\widehat{s}_k)^{\widehat{c}_k})^2}\times\frac{y_i}{\widehat{s}}\right]$$

e

$$\ell_{c;Y}(k,\widehat{c}_k)\widehat{V}_c = \sum_{i=1}^{n}\left[-\frac{(1+(y_i/\widehat{s}_k)^{\widehat{c}_k})(k(y_i/\widehat{s}_k)^{\widehat{c}_k}-1)+\widehat{c}_k(k+1)(y_i/\widehat{s}_k)^{\widehat{c}_k}\log(y_i/\widehat{s}_k)}{y_i(1+(y_i/\widehat{s}_k)^{\widehat{c}_k})^2}\right.$$
$$\left.\times -\frac{y_i\log(y_i/\widehat{s})}{\widehat{c}}\right].$$

A função de log-verossimilhança perfilada modificada proposta por Severini (1998) é dada por

$$\bar{\ell}_{BN}(k) = \ell_p(k) + \frac{1}{2}\log|j_{\phi\phi}(k,\widehat{\boldsymbol{\phi}}_k)| - \log|I_{\boldsymbol{\phi}}(k,\widehat{\boldsymbol{\phi}}_k;\widehat{k},\widehat{\boldsymbol{\phi}})|, \tag{3.4}$$

em que, $I_{\boldsymbol{\phi}}(k,\widehat{\boldsymbol{\phi}}_k;\widehat{k},\widehat{\boldsymbol{\phi}}) = \sum_{i=1}^{n}\ell_{\boldsymbol{\phi}}^{(i)}(k,\widehat{\boldsymbol{\phi}}_k)\ell_{\boldsymbol{\phi}}^{(i)}(\widehat{k},\widehat{\boldsymbol{\phi}})^{\top}$, sendo $\ell_{\boldsymbol{\phi}}^{(i)}$ a função escore baseada na i-ésima observação.

Na distribuiçao Burr XII, temos que

$$I_{\boldsymbol{\phi}}(k,\widehat{\boldsymbol{\phi}}_k;\widehat{k},\widehat{\boldsymbol{\phi}}) = \begin{pmatrix} I_{ss} & I_{sc} \\ I_{cs} & I_{cc} \end{pmatrix},$$

em que

$$I_{ss} = \sum_{i=1}^{n}\left[\left(\frac{\widehat{c}_k(k(y_i/\widehat{s}_k)^{\widehat{c}_k}-1)}{\widehat{s}_k(1+(y_i/\widehat{s}_k)^{\widehat{c}_k})}\right)\left(\frac{\widehat{c}(\widehat{k}(y_i/\widehat{s})^{\widehat{c}}-1)}{\widehat{s}(1+(y_i/\widehat{s})^{\widehat{c}})}\right)\right],$$

$$I_{sc} = \sum_{i=1}^{n}\left[\left(\frac{\widehat{c}_k(k(y_i/\widehat{s}_k)^{\widehat{c}_k}-1)}{\widehat{s}_k(1+(y_i/\widehat{s}_k)^{\widehat{c}_k})}\right)\left(\frac{1}{\widehat{c}}-\log(\widehat{s})-\frac{(\widehat{k}+1)(y/\widehat{s})^{\widehat{c}}\log(y/\widehat{s})}{1+(y/\widehat{s})^{\widehat{c}}}+\log(y_i)\right)\right],$$

$$I_{cs} = \sum_{i=1}^{n}\left[\left(\frac{1}{\widehat{c}_k}-\log(\widehat{s}_k)-\frac{(k+1)(y/\widehat{s}_k)^{\widehat{c}_k}\log(y/\widehat{s}_k)}{1+(y/\widehat{s}_k)^{\widehat{c}_k}}+\log(y_i)\right)\left(\frac{\widehat{c}(\widehat{k}(y_i/\widehat{s})^{\widehat{c}}-1)}{\widehat{s}(1+(y_i/\widehat{s})^{\widehat{c}})}\right)\right]$$

e

$$I_{cc} = \sum_{i=1}^{n}\left[\left(\frac{1}{\widehat{c}_k}-\log(\widehat{s}_k)-\frac{(k+1)(y/\widehat{s}_k)^{\widehat{c}_k}\log(y/\widehat{s}_k)}{1+(y/\widehat{s}_k)^{\widehat{c}_k}}+\log(y_i)\right)\right.$$
$$\left.\times\left(\frac{1}{\widehat{c}}-\log(\widehat{s})-\frac{(\widehat{k}+1)(y/\widehat{s})^{\widehat{c}}\log(y/\widehat{s})}{1+(y/\widehat{s})^{\widehat{c}}}+\log(y_i)\right)\right].$$

Em que $\widehat{s}$, $\widehat{k}$ e $\widehat{c}$ são, respectivamente, os EMV de s, k e c. Os EMV perfilados modificados de k obtidos em (3.3) e (3.4) são denominados, respectivamente, por $\widehat{\breve{k}}_{BN}$ e $\widehat{\bar{k}}_{BN}$. Novamente, os estimadores não possuem solução analítica e devem ser obtidos numericamente por meio da maximização do logaritmo das suas respectivas funções de verossimilhanças.

As estatísticas da razão de verossimilhanças baseadas na função de verossimilhança perfilada $\ell_p(k)$ e nas funções de verossimilhanças perfiladas modificadas $\breve{\ell}_{BN}(k)$ e $\bar{\ell}_{BN}(k)$ para testar $H_0 : k = k_0$ contra $H_1 : k \neq k_0$ são dadas, respectivamente, por

$$LR_p = 2\{\ell_p(\widehat{k}) - \ell_p(k_0)\},$$

$$LR_{BN1} = 2\{\breve{\ell}_{BN}(\widehat{\breve{k}}) - \breve{\ell}_{BN}(k_0)\}$$

e

$$LR_{BN2} = 2\{\bar{\ell}_{BN}(\widehat{\bar{k}}) - \bar{\ell}_{BN}(k_0)\}.$$

Satisfeitas as condições de regularidade e sob a hipótese nula, as estatísticas da razão de verossimilhanças LR_p, LR_{BN1} e LR_{BN2} são distribuídas assintoticamente como χ^2_1. Mesmo com a modificação, a aproximação pela distribuição χ^2_1 tem erro de ordem $O(n^{-1})$.

3.5.1 Resultados numéricos

Avaliaremos numericamente a qualidade das inferências realizadas para o parâmetro de interesse k da distribuição Burr XII baseadas nas funções de verossimilhança perfilada e perfiladas modificadas com base na proposta de Barndorff-Nielsen (1983) através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Os principais objetivos dos estudos de simulação são: (i) comparar os desempenhos dos EMV perfilado e perfilados modificados, considerando as seguintes medidas descritivas dos diferentes estimadores: viés relativo, erro quadrático médio (EQM), assimetria e curtose (Tabela 1), (ii) comparar a média e a variância das estatísticas da razão de verossimilhanças perfiladas (LR_p) e razão de verossimilhanças perfiladas modificadas (LR_{BN1} e LR_{BN2}) com a média e a variância da distribuição χ^2_1 (Tabela 2), (iii) verificar a qualidade das aproximações das estatísticas de teste em relação à distribuição χ^2_1 (Figura 3) e (iv) comparar os desempenhos dos testes da razão de verossimilhanças baseados nas estatísticas LR_p, LR_{BN1} e LR_{BN2}, através dos tamanhos e poderes (Tabelas 3 e 4). As entradas das Tabelas 3 e 4 apresentadas são porcentagem. Para estimação pontual e os testes de hipóteses, o número de réplicas de Monte Carlo utilizadas

em cada simulação foi de 10.000, considerando os seguintes tamanhos amostrais: $n = 35$, 45, 55, 65 e 100. Os valores considerados para o parâmetro de interesse k foram de $0,25$ e $0,5$. Para os parâmetros de incômodo foram considerados $c = 3$ e $s = 0,5$. Para cada tamanho de amostra e cada nível nominal considerado, calculamos as taxas de rejeição de cada teste, isto é, estimamos, via simulação $P(LR_p > x_\alpha)$, $P(LR_{BN1} > x_\alpha)$ e $P(LR_{BN2} > x_\alpha)$ em que x_α é o quantil $(1 - \alpha)$ apropriado da distribuição χ^2_1. As simulações foram realizadas usando a linguagem de programação Ox (DOORNIK, 2009). Para as maximizações dos logaritmos das funções de verossimilhanças, utilizamos os métodos numéricos BFGS e SQP. O gráfico foi gerado usando a linguagem de programação R (R Core Team, 2020).

A Tabela 1, contém medidas descritivas de diferentes estimadores. Observamos que os estimadores perfilados modificados $\widehat{\tilde{k}}_{BN}$ e $\widehat{\bar{k}}_{BN}$ apresentam desempenhos semelhantes e superiores quando comparado com o EMV perfilado $\widehat{k}_p$. Por exemplo, para $n = 35$ e $k = 0,25$, o erro quadrático médio do estimador $\widehat{k}_p$ é de 0,134, enquanto os erros quadráticos médios dos estimadores perfilados ajustados $\widehat{\tilde{k}}_{BN}$ e $\widehat{\bar{k}}_{BN}$ foram de 0,057 e 0,044, respectivamente. Satisfeita as condições de regularidade, os EMV têm distribuição assintótica normal. Nota-se que a curtose de todos os estimadores são maiores que três, isto implica que a distribuição desses estimadores têm curva leptocúrtica. A assimetria é positiva, isto é, a distribuição dos estimadores apresentam valores concentrados à esquerda. À medida que o valor de k aumenta, os vieses relativos, os erros quadráticos médios, as assimetrias e as curtoses dos estimadores ficam mais elevados. Observamos também que, de modo geral, a assimetria e a curtose dos estimadores perfilados ajustados são menores quando comparado com a assimetria e a curtose do estimador perfilado. Como esperado, à medida que o tamanho da amostra aumenta, as medidas dos estimadores $\widehat{k}_p$, $\widehat{\tilde{k}}_{BN}$ e $\widehat{\bar{k}}_{BN}$ passam a ter desempenhos semelhantes. Por exemplo, para $n = 100$ e $k = 0,25$, a assimetria e a curtose dos estimadores $\widehat{k}_p$, $\widehat{\tilde{k}}_{BN}$ e $\widehat{\bar{k}}_{BN}$ são (0,965 e 5,378), (0,972 e 5,320) e (0,965 e 5,339), respectivamente.

A Tabela 2 apresenta a média e a variância das estatísticas de teste LR_p, LR_{BN1} e LR_{BN2}, obtidas via simulação de Monte Carlo para o caso em que $k = 0,25$ e $k = 0,50$, para diferentes valores de n. Aqui, comparamos a média e a variância estimadas das diferentes estatísticas de teste com a média e a variância da distribuição χ^2_1. Note que para $n = 35$, a média e a variância da estatística LR_p está mais próxima da média e variância da distribuição χ^2_1 quando comparado com a média e a variância das estatísticas LR_{BN1} e LR_{BN2}. Porém, a partir de $n = 45$ a média e variância das estatísticas LR_{BN1} e LR_{BN2} passam a ser mais próximas das medidas da distribuição de referência χ^2_1 quando comparado com a estatística LR_p. Por

exemplo, para $n = 35$ e $k = 0,50$, a média e a variância da estatística LR_p é de 1,033 e 1,856, enquanto as mesmas medidas das estatísticas LR_{BN1} e LR_{BN2} são de (0,906 e 1,422) e (0,877 e 1,314), respectivamente. Para $n = 55$ e $k = 0,25$ a média e a variância da estatística LR_p é de 1,091 e 2,322, enquanto a média e a variância das estatísticas LR_{BN1} e LR_{BN2} são de (1,005 e 1,989) e (0,997 e 1,969), respectivamente. Note também que as medidas das estatísticas LR_{BN1} e LR_{BN2} são similares em todos os tamanhos amostrais e com o aumento da amostra são as que mais se aproximam da média e variância da distribuição χ^2_1.

Na Tabela 3 comparamos os desempenhos dos testes da razão de verossimilhanças baseados nas estatísticas LR_p, LR_{BN1} e LR_{BN2}. Foram considerados os mesmos tamanhos amostrais da estimação pontual e para cada tamanho de amostra foram calculadas as taxas de rejeição dos testes considerando as hipóteses $H_0 : k = k_0$ contra a hipótese alternativa $H_1 : k \neq k_0$ com níveis nominais de 1%, 5% e 10%. Os desempenhos são avaliados em função da proximidade da probabilidade de rejeição da hipótese nula, sendo esta verdadeira, aos seus respectivos níveis nominais. Observamos que os testes baseados nas estatísticas LR_{BN1} e LR_{BN2} têm desempenhos superiores ao teste LR_p. Por exemplo, para $n = 45$, $k = 0,25$ e nível nominal de 10%, o teste LR_p apresenta taxa de rejeição igual a $11,60\%$, enquanto os testes LR_{BN1} e LR_{BN2} apresentam taxas iguais a $9,78\%$ e $9,54\%$, respectivamente, mais próximas do nível nominal. Observamos que para qualquer tamanho amostral, o teste LR_p é liberal apresentando taxas de rejeição acima dos níveis nominais. Para $n = 35$, os testes LR_{BN1} e LR_{BN2} tornam-se mais conservativos, mas à medida que o tamanho amostral aumenta, as taxas de rejeição ficam mais próximas dos níveis nominais considerados. Por exemplo, para $n = 55$, $k = 0,5$ e nível nominal de 10%, os testes LR_p, LR_{BN1} e LR_{BN2} apresentam taxas de rejeição iguais a $11,30\%$, $10,09\%$ e $9,90\%$, respectivamente. Assim, como esperado, o teste LR_p passa a apresentar desempenho similar aos testes LR_{BN1} e LR_{BN2} em amostras maiores. Por exemplo, para $n = 100$ e $k = 0,5$ e nível nominal de 5%, os testes LR_p, LR_{BN1} e LR_{BN2} apresentaram taxas de rejeição iguais a $5,55\%$, $5,31\%$ e $5,31\%$.

Avaliamos a aproximação assintótica das distribuições nula das estatísticas LR_p, LR_{BN1} e LR_{BN2} pela distribuição χ^2_1. A Figura 3 apresenta o gráfico das discrepâncias relativas de quantis (diferenças absolutas entre quantis exatos e assintóticos dividido pelo quantil assintótico), para $n = 55$ e $k = 0,25$. Mais precisamente, denotando por ST a estatística de teste (LR_p, LR_{BN1}, LR_{BN2}, conforme o caso), o quantil amostral de ordem $(1 - \alpha)$ do conjunto de valores simulados da estatística do teste sendo denotado por $ST(1 - \alpha)$ e o correspondente quantil da distribuição χ^2_1 sendo denotado por $\chi^2_1(1-\alpha)$, a discrepância relativa

é definida como

$$\frac{ST(1-\alpha)-\chi^2_1(1-\alpha)}{\chi^2_1(1-\alpha)}.$$

Quanto mais próxima à curva estiver da ordenada zero, melhor é a aproximação da distribuição nula da estatística de teste pela distribuição assintótica χ^2_1. Observe que as curvas relacionadas às estatísticas LR_{BN1} e LR_{BN2} são similares e estão mais próximas da ordenada zero quando comparada com a curva da estatística LR_p. Portanto, as distribuições nula das estatísticas da razão de verossimilhanças obtidas através das funções de verossimilhanças perfiladas modificadas são melhores aproximada pela distribuição de referência χ^2_1.

A Tabela 4 apresenta resultados de simulação obtidos sob a hipótese alternativa para $n = 55$, k variando de $0,30$ até $0,95$ e nível nominal de 10%. Uma vez que os testes de tamanho baseados nas estatísticas LR_p, LR_{BN1} e LR_{BN2} apresentam diferenças entre as taxas de rejeição para um nível nominal fixado, os testes foram obtidos usando valores críticos estimados para cada estatística, de maneira que, todos os testes tivessem o mesmo patamar. De modo geral, observamos que os desempenhos dos testes baseados nas três estatísticas consideradas foram bastante similares. Notamos que os poderes dos testes LR_{BN1} e LR_{BN2} não perdem desempenho com relação ao poder do teste LR_p. Além disso, os testes tornam-se mais poderosos à medida que ocorre um afastamento do valor verdadeiro do parâmetro especificado na hipótese nula, conforme esperado.

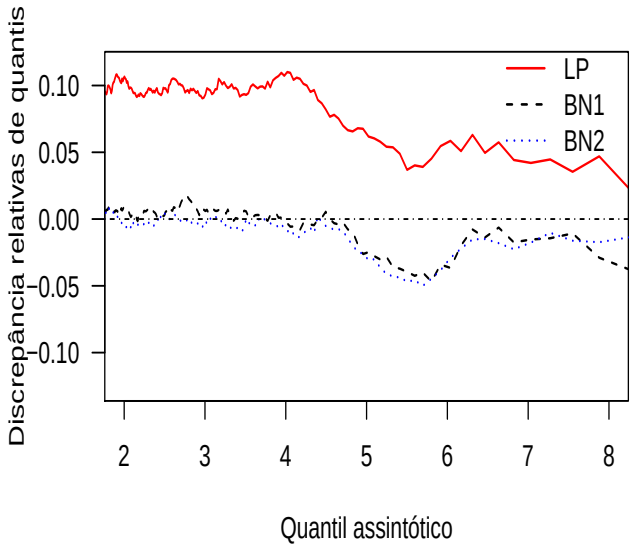
Da Tabela 5, nota-se que no extremo da cauda da distribuição χ^2_1 (que corresponde a tomar níveis nominais bem pequenos), as diferenças relativas entre as taxas de rejeição empírica do teste da razão de verossimilhanças e os correspondentes níveis nominais, são mais acentuadas, diferente do que ocorre na região mediana da distribuição. Por exemplo, para o nível nominal de 50%, a estatística LR_p apresenta uma diferença relativa de $3,75\%$ $(((51877-50000)/50000)\times 100)$, enquanto, para nível nominal de $0,05\%$ esta diferença é de 58%. Tal diferença é menos acentuadas para as estatísticas da razão de verossimilhanças perfiladas modificadas LR_{BN1} e LR_{BN2}. Por exemplo, para às duas situações citadas acima, as diferenças relativas das estatísticas LR_{BN1} e LR_{BN2} para nível nominal de 50% são de $0,26\%$ e $-0,12\%$, respectivamente. Já para nível nominal de $0,05\%$ as diferenças relativas são de -8% e -16%, respectivamente.

Tabela 1 – Medidas descritivas de diferentes estimadores de k para $c = 3$ e $s = 0,5$.

		$k = 0,25$				$k = 0,5$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
35	$\widehat{k}_p$	0,202	0,134	18,104	529,093	0,378	1,056	13,552	300,648
	$\widehat{\widehat{k}}_{BN}$	0,205	0,057	9,084	217,949	0,273	0,309	5,044	46,802
	$\widehat{\widehat{k}}_{BN}$	0,188	0,044	4,367	47,005	0,261	0,282	4,889	47,407
45	$\widehat{k}_p$	0,094	0,034	6,077	95,210	0,223	0,403	12,098	262,849
	$\widehat{\widehat{k}}_{BN}$	0,117	0,029	3,915	40,453	0,194	0,202	5,403	58,223
	$\widehat{\widehat{k}}_{BN}$	0,115	0,029	4,032	49,113	0,196	0,212	5,500	57,788
55	$\widehat{k}_p$	0,060	0,022	5,525	99,077	0,161	0,240	15,672	519,064
	$\widehat{\widehat{k}}_{BN}$	0,082	0,020	3,858	54,949	0,145	0,132	4,661	47,133
	$\widehat{\widehat{k}}_{BN}$	0,079	0,020	3,636	47,428	0,147	0,140	5,507	68,342
65	$\widehat{k}_p$	0,041	0,014	1,935	14,348	0,122	0,136	10,418	252,624
	$\widehat{\widehat{k}}_{BN}$	0,061	0,014	1,961	14,130	0,120	0,108	7,221	137,547
	$\widehat{\widehat{k}}_{BN}$	0,057	0,014	1,977	14,396	0,124	0,122	8,495	164,032
100	$\widehat{k}_p$	0,018	0,007	0,965	5,378	0,059	0,038	2,065	18,151
	$\widehat{\widehat{k}}_{BN}$	0,029	0,007	0,972	5,320	0,060	0,036	1,936	15,672
	$\widehat{\widehat{k}}_{BN}$	0,028	0,007	0,965	5,339	0,060	0,037	2,001	16,826

Fonte: Próprio autor (2021).

Figura 3 – Gráfico de discrepância relativa dos quantis da distribuição Burr XII para $n = 55$, $s = 0,5$, $c = 3$ e $k = 0,25$.



Fonte: Próprio autor (2021).

Tabela 2 – Médias e variâncias amostrais de diferentes estatísticas de teste e os valores assintóticos destas, $H_0 : k = k_0$, distribuição χ^2_1.

			$k = 0,25$			$k = 0,50$		
n		χ^2_2	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
35	média	1,00	1,049	0,912	0,897	1,033	0,906	0,877
	var	2,00	2,056	1,575	1,509	1,856	1,422	1,314
45	média	1,00	1,073	0,971	0,959	1,059	0,959	0,953
	var	2,00	2,161	1,784	1,734	2,141	1,683	1,666
55	média	1,00	1,091	1,005	0,997	1,077	1,003	0,990
	var	2,00	2,322	1,989	1,969	2,227	1,895	1,884
65	média	1,00	1,093	1,011	1,015	1,082	1,019	1,016
	var	2,00	2,332	1,950	1,985	2,358	2,035	2,027
100	média	1,00	1,051	1,007	1,005	1,031	1,004	1,002
	var	2,00	2,235	2,038	2,028	2,155	2,039	2,018

Fonte: Próprio autor (2021).

Tabela 3 – Taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : k = k_0$, para $s = 0,5$ e $c = 3$.

		$n = 35$			$n = 45$			$n = 55$			$n = 65$			$n = 100$		
k	$\alpha(\%)$	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
0,25	10	10,91	8,52	8,24	11,60	9,78	9,54	11,60	10,15	9,97	11,27	9,87	9,90	10,57	9,92	9,91
	5	5,43	3,73	3,53	5,71	4,48	4,34	6,08	5,05	4,94	6,16	5,21	5,19	5,77	5,25	5,26
	1	1,06	0,63	0,53	1,09	0,77	0,72	1,24	0,97	0,93	1,32	1,27	1,00	1,27	1,07	1,08
0,5	10	10,64	8,28	7,57	11,09	9,47	9,36	11,30	10,09	9,90	11,36	10,35	10,26	10,55	10,15	10,13
	5	5,08	3,47	3,08	5,65	4,42	4,39	5,78	4,85	4,65	5,89	5,17	5,06	5,55	5,31	5,31
	1	0,84	0,36	0,32	1,07	0,64	0,57	1,22	0,89	0,83	1,29	1,05	1,03	1,17	1,05	1,04

Fonte: Próprio autor (2021).

Tabela 4 – Taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 55$ e nível nominal de 10%, indexados pelos parâmetros $c = 3$ e $s = 0,5$.

k	LR_p	LR_{BN1}	LR_{BN2}
0,30	13,22	12,73	12,84
0,35	20,29	19,81	20,10
0,40	29,57	28,90	28,99
0,45	39,16	38,67	39,01
0,50	47,70	47,50	47,66
0,55	57,53	56,90	57,44
0,60	63,37	63,56	64,08
0,65	70,11	70,14	70,58
0,70	75,77	76,11	76,25
0,75	79,22	79,38	79,67
0,80	83,41	83,74	83,64
0,95	89,78	90,31	90,49

Fonte: Próprio autor (2021).

Tabela 5 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo Burr XII indexado pelos parâmetros $c = 3$, $s = 0,5$, $k = 0,25$ e $n = 35$.

$\alpha(\%)$	exato	LR_p	LR_{BN1}	LR_{BN2}
50	50000	51877	50130	49940
15	15000	17010	15201	15100
10	10000	11649	10166	10085
5	5000	6115	5046	4986
3	3000	3805	2998	2924
1	1000	1303	942	923
0,5	500	665	447	445
0,1	100	144	83	82
0,05	50	79	46	42
0,01	10	16	10	8

Fonte: Próprio autor (2021).

3.6 CASO (ii)

No segundo caso, estamos considerando que s e k são parâmetros de interesse, isto é, $\boldsymbol{\mu} = (k, s)^\top$ e $\phi = c$ é o parâmetro de incômodo. Fixando (k e s) e resolvendo a equação

$u_c = 0$, obtemos o EMV restrito de c ($\widehat{c}_{\boldsymbol{\mu}}$). Substituindo c por $\widehat{c}$ na log-verossimilhança original, isto é, em (3.2), obtemos a log-verossimilhança perfilada para $\boldsymbol{\mu}$. Neste caso, temos que a log-verossimilhança perfilada para o parâmetro $\boldsymbol{\mu}$ é dada como

$$\ell_p(\boldsymbol{\mu}) = n[\log(\widehat{c}_{\boldsymbol{\mu}}) + \log(k) - (\widehat{c}_{\boldsymbol{\mu}})\log(s)] + (\widehat{c}_{\boldsymbol{\mu}} - 1)\sum_{i=1}^{n}\log(y_i) - (k+1)\sum_{i=1}^{n}\log[1 + (y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}}].$$

O EMV perfilado $\widehat{\boldsymbol{\mu}}_p$ é obtido como solução de $u_p(\boldsymbol{\mu}) = 0$ e não tem solução em forma fechada. Logo, para encontrarmos o valor de $\widehat{\boldsymbol{\mu}}_p$ que maximiza $\ell_p(\boldsymbol{\mu})$ é necessário a utilização de um procedimento de otimização não-linear.

Derivamos duas aproximações para a verossimilhança perfilada modificada proposta por Barndorff-Nielsen (1983). O primeiro ajuste foi proposto por Fraser e Reid (1995) e Fraser, Reid e Wu (1999) e a log-verossimilhança modificada pode ser expressa como

$$\breve{\ell}_{BN}(\boldsymbol{\mu}) = \ell_p(\boldsymbol{\mu}) + \frac{1}{2}\log|j_{cc}(\boldsymbol{\mu}, \widehat{c}_{\boldsymbol{\mu}})| - \log|\ell_{c;Y}(\boldsymbol{\mu}, \widehat{c}_{\boldsymbol{\mu}})\widehat{V}_c|, \tag{3.5}$$

em que

$$\ell_{c;Y}(\boldsymbol{\mu}, \widehat{c}_{\boldsymbol{\mu}})\widehat{V}_c = \sum_{i=1}^{n}\left[\frac{\log(y_i/\widehat{s})[(1 + (y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}})(k(y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}} - 1) + \widehat{c}_{\boldsymbol{\mu}}(k+1)(y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}}\log(y_i/s)]}{\widehat{c}(1 + (y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}})^2}\right]$$

e

$$j_{cc}(\boldsymbol{\mu}, \widehat{c}_{\boldsymbol{\mu}}) = \frac{n}{\widehat{c}^2_{\boldsymbol{\mu}}} + (k+1)\sum_{i=1}^{n}\left[\frac{(y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}}\log^2(y_i/s)}{(1 + (y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}})^2}\right].$$

O EMV perfilado modificado obtido em (3.5) é denotado $\widehat{\breve{\boldsymbol{\mu}}}_{BN}$ e não possui solução em forma fechada devendo ser obtido maximizando numericamente a função de log-verossimilhança.

A segunda aproximação foi introduzida por Severini (1998). A log-verossimilhança perfilada modificada é dada por

$$\bar{\ell}_{BN}(\boldsymbol{\mu}) = \ell_p(\boldsymbol{\mu}) + \frac{1}{2}\log|j_{cc}(\boldsymbol{\mu}, \widehat{c}_{\boldsymbol{\mu}})| - \log|I_c(\boldsymbol{\mu}, \widehat{c}_{\boldsymbol{\mu}}; \widehat{\boldsymbol{\mu}}, \widehat{c})|, \tag{3.6}$$

em que

$$I_c(\boldsymbol{\mu}, \widehat{c}_{\boldsymbol{\mu}}; \widehat{\boldsymbol{\mu}}, \widehat{c}) = \sum_{i=1}^{n}\left[\left(\frac{1}{\widehat{c}_{\boldsymbol{\mu}}} - \log(s) - \frac{(k+1)(y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}}\log(y_i/s)}{1 + (y_i/s)^{\widehat{c}_{\boldsymbol{\mu}}}} + \log(y_i)\right)\right.$$
$$\left.\times\left(\frac{1}{\widehat{c}} - \log(\widehat{s}) - \frac{(\widehat{k}+1)(y_i/\widehat{s})^{\widehat{c}}\log(y_i/\widehat{s})}{1 + (y_i/\widehat{s})^{\widehat{c}}} + \log(y_i)\right)\right].$$

O EMV perfilado modificado $\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$ não possui solução em forma fechada e deve ser obtido por métodos numéricos de maximização.

Os testes baseados nas estatísticas da razão de verossimilhanças perfiladas e perfiladas modificadas para testar $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ contra $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$ são, respectivamente, dadas por

$$\begin{aligned} LR_p &= 2\Big\{\ell_p(\widehat{\boldsymbol{\mu}}) - \ell_p(\boldsymbol{\mu}_0)\Big\}, \\ LR_{BN1} &= 2\Big\{\breve{\ell}_{BN}(\widehat{\breve{\boldsymbol{\mu}}}) - \breve{\ell}_{BN}(\boldsymbol{\mu}_0)\Big\} \quad \text{e} \\ LR_{BN2} &= 2\Big\{\bar{\ell}_{BN}(\widehat{\bar{\boldsymbol{\mu}}}) - \bar{\ell}_{BN}(\boldsymbol{\mu}_0)\Big\}. \end{aligned}$$

Sob condições de regularidade e sob a hipótese nula, as estatísticas LR_p, LR_{BN1} e LR_{BN2} são assintoticamente distribuídas com distribuição χ^2_2. Mesmo com a modificação, a aproximação pela distribuição χ^2_2 tem erro de ordem $O(n^{-1})$.

3.6.1 Resultados numéricos

Avaliaremos numericamente a qualidade das inferências realizadas para os parâmetros de interesses $\boldsymbol{\mu} = (k, s)^\top$ da distribuição Burr XII baseadas nas funções de verossimilhança perfilada e perfiladas modificadas com base na proposta de Barndorff-Nielsen (1983) através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Os principais objetivos dos estudos de simulação são: (i) comparar os desempenhos dos EMV perfilado e perfilados modificados, considerando as seguintes medidas descritivas dos diferentes estimadores: viés relativo, erro quadrático médio (EQM), assimetria e curtose (Tabela 6), (ii) comparar a média e a variância das estatísticas LR_p, LR_{BN1} e LR_{BN2} com a média e variância da distribuição χ^2_2 (Tabela 7), (iii) verificar a qualidade das aproximações das estatísticas de teste em relação à distribuição χ^2_2 (Figura 4) e (iv) comparar os desempenhos dos testes LR_p, LR_{BN1} e LR_{BN2}, através dos tamanhos e poderes (Tabelas 8 e 9). As entradas das Tabelas 8 e 9 apresentadas são porcentagem. Para estimação pontual e os testes de hipóteses, o número de réplicas de Monte Carlo utilizadas em cada simulação foi 10.000, considerando os seguintes tamanhos amostrais: $n = 30$, 40, 50, 60 e 100. Os níveis nominais utilizados foram $\alpha = 1\%$, 5% e 10%. Os valores considerados para o parâmetro de interesse $\boldsymbol{\mu}$ foram $s = 0,5$ e variamos k em $0,25$ e $0,5$. Para o parâmetro de incômodo consideramos $c = 5$. As simulações foram realizadas usando a linguagem de programação Ox (DOORNIK, 2009). Para as maximizações das log-verossimilhanças, utilizamos os métodos numéricos BFGS e SQP. O gráfico foi gerado usando a linguagem de programação R (R Core Team, 2020).

Na Tabela 6 apresentamos os desempenhos dos EMV perfilado e perfilados modificados, a saber: $\widehat{\boldsymbol{\mu}}_p$, $\widehat{\overset{\circ}{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$. Nota-se que, os EMV perfilados modificados apresentaram vieses relativos menores quando comparado com o EMV perfilado, principalmente em amostras menores. Por exemplo, para $n = 30$ e $k = 0,25$, os vieses relativos dos estimadores $\widehat{\boldsymbol{\mu}}_p$, $\widehat{\overset{\circ}{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$ são: (26,4%; 9,8%), (15,6%; 6,3%) e (12,3%; 5,3%), respectivamente. De modo semelhante, ocorrem com os erros quadráticos médios, assimetrias e curtoses. Essas medidas para os estimadores perfilados modificados apresentam desempenhos superiores ao estimador perfilado. Como esperado, com o aumento do tamanho da amostra, o desempenho do estimador $\widehat{\boldsymbol{\mu}}_p$ se aproxima dos desempenhos dos estimadores $\widehat{\overset{\circ}{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$. A assimetria de todos os estimadores são positivas, indicando que a distribuição dos estimadores considerados apresentam valores concentrados à esquerda. À medida que o valor de k aumenta os desempenhos dos estimadores se deterioram, principalmente o estimador $\widehat{\boldsymbol{\mu}}_p$.

Apresentamos na Tabela 7 a média e a variância das estatísticas LR_p, LR_{BN1} e LR_{BN2} estimadas com $s = 0,5$ e k variando de 0,25 a 0,50 para diferentes valores de n. Note que a média e a variância da estatística LR_p é mais acentuada quando comparada com a média e a variância da distribuição χ^2_2. Os resultados mostram que, em geral, a média e a variância amostral das estatísticas corrigidas LR_{BN1} e LR_{BN2} são as que mais se aproximam da média e da variância da distribuição de referência χ^2_2.

A Tabela 8 apresenta as taxas de rejeição dos testes baseados nas estatísticas de teste, a saber: LR_p, LR_{BN1} e LR_{BN2}, para diferentes tamanhos de amostra e níveis nominais. Note que o teste LR_p é liberal, apresentando valores acima dos níveis nominais. Já os testes LR_{BN1} e LR_{BN2} em amostras menores são conservativos e conforme o tamanho da amostra aumenta os testes ficam mais próximos dos níveis nominais. Por exemplo, para $n = 30$ e $k = 0,25$ os testes LR_p, LR_{BN1} e LR_{BN2} apresentaram taxas de rejeição de $10,75\%$, $8,15\%$ e $8,06\%$, respectivamente. Para o tamanho amostral $n = 40$ e $k = 0,25$ os testes LR_p, LR_{BN1} e LR_{BN2} apresentaram taxas de rejeição de $11,59\%$, $9,74\%$ e $9,36\%$, respectivamente. Quando o tamanho amostral aumenta, os testes passam a apresentar desempenhos semelhantes. Por exemplo, quando $k = 0,5$, $s = 0,5$, $n = 100$ e nível nominal $\alpha = 10\%$, os testes apresentam as seguintes taxas de rejeição: $LR_p = 10,25\%$, $LR_{BN1} = 9,81\%$ e $LR_{BN2} = 9,71\%$. Avaliando, de modo geral, os testes baseados nas estatísticas LR_{BN1} e LR_{BN2} tiveram os melhores desempenhos, pois, os testes correspondentes forneceram taxas de rejeição mais próximas do nível nominal especificado. Por exemplo, para $n = 50$ e $k = 0,50$ ao nível de 10% os testes LR_{BN1} e LR_{BN2} apresentaram taxas de rejeição de $9,83\%$ e $9,61\%$, enquanto o teste LR_p

apresentou taxa de rejeição de $11,22\%$.

O gráfico de discrepância relativa de quantis das estatísticas LR_p, LR_{BN1} e LR_{BN2} são apresentados na Figura 4. Nota-se que o comportamento das curvas de discrepância relativa do quantil das estatísticas LR_{BN1} e LR_{BN2} são semelhantes e posicionadas abaixo da ordenada zero em toda sua extensão, indicando que os quantis amostrais das estatísticas de teste subestimaram os respectivos quantis assintóticos, como mostra a Tabela 8. As curvas relacionadas as estatísticas LR_{BN1} e LR_{BN2} estão mais próximas da ordenada zero do que a curva relacionada a estatística LR_p. Desta forma, as distribuições nula das estatísticas de teste obtidas através das funções de verossimilhanças perfiladas modificadas aproxima-se mais da distribuição χ^2_2.

A Tabela 9 apresenta os resultados das simulações das taxas de rejeição não nula. Aqui, consideramos o cenário em que $n = 30$, o parâmetro de incômodo $c = 5$ e diferentes valores para os parâmetros de interesse $k = s = \delta$, considerando um nível nominal de $\alpha = 10\%$. Novamente, as simulações de poder foram realizadas utilizando valores críticos exatos estimados e não tabulados. Adotamos este procedimento porque em determinadas situações alguns testes são conservativos. Desta maneira, as simulações ficam ajustadas para que os testes tenham o mesmo patamar. Observe que estas simulações de poder corresponde a situação abordada na Tabela 8 para $n = 30$ e $k = 0,5$. Avaliando os resultados, observamos que os desempenhos dos testes considerados são similares, com isso, podemos concluir que as modificações realizadas na verossimilhança perfilada não acarretaram perda de poder.

A Tabela 10 apresenta os resultados de simulação de 100.000 amostras de Monte Carlo. Novamente, com o interesse em avaliar os tamanhos empíricos dos testes para valores pequenos de nível nominal, para o cenário em que $n = 30$, $c = 5$ e $k = s = 0,5$. Observe que as diferenças relativas entre as taxas de rejeição empíricas e os níveis nominais dos testes LR_{BN1} e LR_{BN2} são negativas e mais acentuadas no extremo da cauda da distribuição χ^2_2. Por exemplo, para o teste LR_{BN2} ao nível nominal de 50% esta diferença é de $-8,8\%$ $(((45600 - 50000)/50000) \times 100)$, enquanto para o nível nominal de $0,05\%$ esta diferença é de -62% $(((19 - 50)/50) \times 100)$. já a diferença relativa entre a taxa de rejeição empírica e o nível nominal do teste LR_p é positiva até o nível nominal de $0,05\%$ a partir de $0,05\%$ esta diferença relativa passa a ser negativa. Novamente, essa diferença é mais acentuada no extremo da cauda da distribuição de referência. Por exemplo, ao nível nominal de 50% a diferença é de $5,43\%$ $(((52718 - 50000)/50000) \times 100)$, enquanto para o nível nominal de $0,05\%$ essa diferença é de -18% $(((41 - 50)/50) \times 100)$.

Tabela 6 – Medidas descritivas de diferentes estimadores de $\boldsymbol{\mu} = (k, s)^\top$ para $c = 5$.

		$k = 0,25$				$k = 0,5$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
30	$\widehat{\boldsymbol{\mu}}_p$	0,264; 0,098	0,118; 0,037	8,527; 5,194	135,964; 52,333	0,364; 0,083	0,600; 0,027	5,838; 2,949	60,532; 17,714
	$\widehat{\breve{\boldsymbol{\mu}}}_{BN}$	0,156; 0,063	0,042; 0,018	2,771; 2,345	17,309; 13,483	0,110; 0,028	0,126; 0,011	1,957; 1,329	9,482; 6,252
	$\widehat{\bar{\boldsymbol{\mu}}}_{BN}$	0,123; 0,053	0,035; 0,016	2,616; 2,293	17,302; 16,005	0,080; 0,019	0,148; 0,011	7,368; 2,286	219,440; 21,283
40	$\widehat{\boldsymbol{\mu}}_p$	0,127; 0,055	0,071; 0,019	27,690; 10,038	1575,320; 314,964	0,241; 0,059	0,327; 0,017	6,038; 3,224	66,900; 23,617
	$\widehat{\breve{\boldsymbol{\mu}}}_{BN}$	0,087; 0,040	0,028; 0,012	3,151; 2,732	26,728; 21,144	0,118; 0,030	0,124; 0,010	2,839; 1,574	20,573; 8,019
	$\widehat{\bar{\boldsymbol{\mu}}}_{BN}$	0,075; 0,036	0,025; 0,011	2,171; 1,921	12,668; 10,786	0,112; 0,028	0,115; 0,009	2,481; 1,527	15,831; 7,881
50	$\widehat{\boldsymbol{\mu}}_p$	0,067; 0,035	0,026; 0,011	5,800; 4,637	118,450; 83,833	0,159; 0,041	0,188; 0,011	6,924; 3,618	100,542; 41,889
	$\widehat{\breve{\boldsymbol{\mu}}}_{BN}$	0,047; 0,026	0,020; 0,009	2,760; 2,298	21,729; 16,471	0,103; 0,026	0,105; 0,008	2,987; 1,727	22,482; 9,354
	$\widehat{\bar{\boldsymbol{\mu}}}_{BN}$	0,041; 0,025	0,019; 0,009	2,832; 2,656	30,722; 27,390	0,108; 0,027	0,103; 0,008	2,605; 1,632	16,294; 8,335
60	$\widehat{\boldsymbol{\mu}}_p$	0,043; 0,025	0,017; 0,007	2,424; 2,822	26,036; 42,274	0,128; 0,033	0,131; 0,008	7,523; 2,388	157,072; 17,555
	$\widehat{\breve{\boldsymbol{\mu}}}_{BN}$	0,032; 0,020	0,015; 0,007	2,140; 1,893	18,382; 15,849	0,099; 0,025	0,092; 0,007	3,113; 1,867	23,157; 10,443
	$\widehat{\bar{\boldsymbol{\mu}}}_{BN}$	0,032; 0,019	0,015; 0,007	1,806; 1,586	10,933; 8,944	0,105; 0,027	0,094; 0,007	3,246; 1,899	26,793; 11,171
100	$\widehat{\boldsymbol{\mu}}_p$	0,017; 0,012	0,007; 0,003	0,959; 0,895	5,375; 5,103	0,057; 0,016	0,037; 0,003	2,079; 1,163	18,376; 6,689
	$\widehat{\breve{\boldsymbol{\mu}}}_{BN}$	0,007; 0,007	0,007; 0,003	0,977; 0,903	5,385; 5,163	0,049; 0,013	0,037; 0,003	2,328; 1,552	24,960; 13,225
	$\widehat{\bar{\boldsymbol{\mu}}}_{BN}$	0,006; 0,007	0,007; 0,003	0,976; 0,884	5,398; 5,008	0,050; 0,013	0,036; 0,003	1,599; 1,094	8,729; 5,824

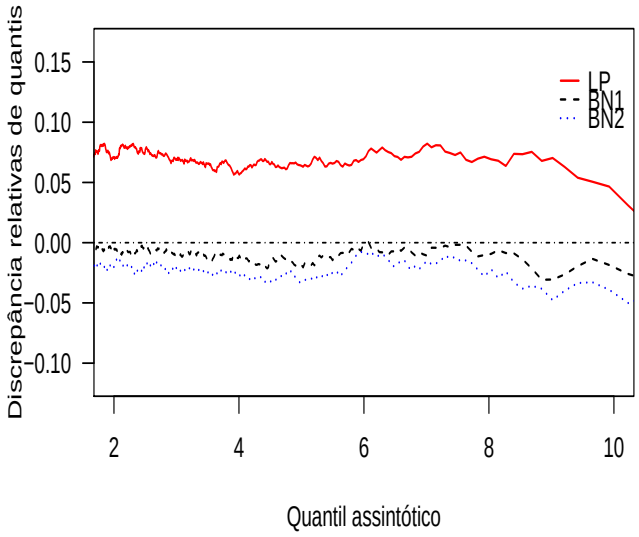
Fonte: Próprio autor (2021).

Tabela 7 – Médias e variâncias amostrais de diferentes estatísticas de teste e os valores assintóticos destas, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, distribuição χ^2_2.

			$(k;s) = (0,25;0,25)$			$(k;s) = (0,50;0,50)$		
n		χ^2_2	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
30	média	2,00	2,091	1,872	1,848	2,091	1,779	1,822
	var	4,00	4,152	3,289	3,276	4,106	3,083	3,487
40	média	2,00	2,131	1,974	1,945	2,105	1,900	1,860
	var	4,00	4,438	3,818	3,783	4,355	3,598	3,520
50	média	2,00	2,139	2,031	2,022	2,106	1,976	1,951
	var	4,00	4,463	4,004	3,947	4,367	3,781	3,710
60	média	2,00	2,144	2,050	2,040	2,117	2,014	2,001
	var	4,00	4,486	4,109	4,029	4,524	4,024	4,032
100	média	2,00	2,081	2,032	2,031	2,043	1,974	1,989
	var	4,00	4,335	4,295	4,167	4,176	4,013	3,982

Fonte: Próprio autor (2021).

Figura 4 – Gráfico de discrepância relativa dos quantis da distribuição Burr XII para $n = 40$, $\boldsymbol{\mu} = (0,5;0,5)^\top$ e $c = 5$.



Fonte: Próprio autor (2021).

Tabela 8 – taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, $\boldsymbol{\mu} = (k, s)^\top$ para $c = 5$.

		$n = 30$			$n = 40$			$n = 50$			$n = 60$			$n = 100$		
k	$\alpha(\%)$	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
0,25	10	10,75	8,15	8,06	11,59	9,74	9,36	11,64	10,43	10,35	11,90	10,87	10,70	11,14	10,48	10,45
	5	5,39	3,55	3,69	5,97	4,99	4,88	6,03	5,16	5,05	6,10	5,37	5,21	5,81	5,44	5,47
	1	1,09	0,63	0,62	1,39	0,91	0,87	1,17	0,93	0,93	1,21	0,92	0,90	1,23	1,15	1,20
0,50	10	11,09	7,30	8,06	11,33	9,06	8,72	11,22	9,83	9,61	11,39	10,25	10,02	10,25	9,81	9,71
	5	5,41	3,20	3,84	6,06	4,34	4,13	5,78	4,73	4,71	6,06	5,14	5,12	5,57	5,22	5,28
	1	1,04	0,52	0,70	1,09	0,72	0,70	1,18	0,87	0,87	1,23	0,87	0,90	1,16	1,05	1,04

Fonte: Próprio autor (2021).

Tabela 9 – taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 30$ e nível nominal de 10%.

δ	LR_p	LR_{BN1}	LR_{BN2}
0,6	38,02	40,75	41,11
0,7	78,23	79,57	78,65
0,8	94,46	94,05	92,48
0,9	98,40	97,66	96,38
1,0	99,44	98,65	98,31
1,1	99,89	99,24	99,23
1,2	99,92	99,42	99,37

Fonte: Próprio autor (2021).

Tabela 10 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo Burr XII indexado pelos parâmetros $c = 5$, $\boldsymbol{\mu} = (0,5; 0,5)$ e $n = 30$.

$\alpha(\%)$	exato	LR_p	LR_{BN1}	LR_{BN2}
50	50000	52718	46767	45600
15	15000	16624	12115	11954
10	10000	11052	7681	7603
5	5000	5525	3502	3552
3	3000	3277	2003	2005
1	1000	1051	584	593
0,5	500	513	264	279
0,1	100	84	42	46
0,05	50	41	15	19
0,01	10	10	5	4

Fonte: Próprio autor (2021).

3.6.2 Caso particular: distribuição log-logística

Quando o parâmetro $k = 1$, obtemos a distribuição log-logística (c, s) como um caso particular da distribuição Burr XII. Neste caso, suponha que s é parâmetro de interesse e c de incômodo, ou seja, $\mu = s$ e $\phi = c$. Resolvendo a expressão $u_c = 0$, obtemos o EMV restrito de c $(\widehat{c}_s)$. O estimador $\widehat{c}_s$ é o mesmo EMV original $\widehat{c}$ que pode ser obtido como solução da equação não-linear dada em (3.2) quando $k = 1$. Substituindo c por $\widehat{c}_s$ na função de verossimilhança

original obtemos a função de verossimilhança perfilada para s. A log-verossimilhança perfilada é dada por

$$\ell_p(s) = n\log(\widehat{c}_s) - n(\widehat{c}_s)\log(s) + (\widehat{c}_s - 1)\sum_{i=1}^{n}\log(y_i) - 2\sum_{i=1}^{n}\log[1 + (y_i/s)^{\widehat{c}_s}]. \tag{3.7}$$

Calculando $\partial\ell_p(s)/\partial s$, obtemos a função escore perfilada para s.

$$u_p(s) = -\frac{n\widehat{c}_s}{s} + \frac{2\widehat{c}_s}{s}\sum_{i=1}^{n}\left[\frac{(y_i/s)^{\widehat{c}_s}}{1 + (y_i/s)^{\widehat{c}_s}}\right].$$

Resolvendo a expressão $u_p(s) = 0$ obtemos o EMV perfilado $\widehat{s}_p$, como $\widehat{s}_p$ não tem solução em forma fechada, deve ser obtido por métodos de otimização não-linear restrita. Para maiores detalhes sobre esses métodos ver (NOCEDAL; WRIGHT, 2006).

Os parâmetros da distribuição log-logística são ortogonais. Neste caso, o ajuste proposto por Cox e Reid (1987) é mais apropriado. Assim, a log-verossimilhança perfilada modificada proposta por Cox e Reid (1987) é dada por

$$\ell_{CR}(s) = \ell_p(s) - \frac{1}{2}\log|j_{cc}(s, \widehat{c}_s)|, \tag{3.8}$$

em que

$$j_{cc}(s, \widehat{c}_s) = \frac{n}{\widehat{c}_s^2} + 2\sum_{i=1}^{n}\left[\frac{(y_i/s)^{\widehat{c}_s}\log^2(y_i/s)}{(1 + (y_i/s)^{\widehat{c}_s})^2}\right].$$

O EMV perfilado ajustado $\widehat{s}_{CR}$ não tem solução em forma fechada, sendo necessário a utilização de um procedimento de otimização não-linear para encontrar o valor de s que maximiza a função $\ell_{CR}(s)$, como, por exemplo, o algoritmo de Newton-Raphson, escore de Fisher ou quase-Newton(BFGS).

A log-verossimilhança perfilada proposta por Barndorff-Nielsen (1983) para o parâmetro s é dada por

$$\ell_{BN}(s) = \ell_p(s) + \frac{1}{2}\log|j_{cc}(s, \widehat{c}_s; \widehat{s}, \widehat{c}, a) - \log|\ell_{c;\widehat{c}}(s, \widehat{c}_s; \widehat{s}, \widehat{c}, a)|,$$

em que a quantidade $\ell_{c;\widehat{c}}(s, \widehat{c}_s; \widehat{s}, \widehat{c}, a) = \partial^2\ell(s, \widehat{c}_s; \widehat{s}, \widehat{c}, a)/\partial c\partial\widehat{c}$. Note que envolve derivadas sobre o espaço amostral de maneira que sua obtenção é geralmente complicada. Assim, é útil adotar uma aproximação para obtenção dessa quantidade.

A aproximação proposta por Fraser e Reid (1995) e Fraser, Reid e Wu (1999), para a distribuição log-logística é dada por

$$\breve{\ell}_{BN}(s) = \ell_p(s) + \frac{1}{2}\log|j_{cc}(s,\widehat{c}_s)| - \log|\ell_{c;Y}(s,\widehat{c}_s)\widehat{V}_c|, \tag{3.9}$$

em que

$$\ell_{c;Y}(s,\widehat{c}_s)\widehat{V}_c = \sum_{i=1}^{n}\left[\frac{\log(y_i/\widehat{s})[(1+(y_i/s)^{\widehat{c}_s})((y_i/s)^{\widehat{c}_s}-1)+2\widehat{c}_s(y_i/s)^{\widehat{c}_s}\log(y_i/s)]}{\widehat{c}(1+(y_i/s)^{\widehat{c}_s})^2}\right].$$

A função escore perfilada modificada é obtida a partir de $\breve{u}_{BN}(s) = \partial\breve{\ell}_{BN}(s)/\partial s$. O EMV perfilado $\widehat{\breve{s}}_{BN}$ novamente não apresenta solução em forma fechada, devendo ser obtido numericamente.

A aproximação proposta por Severini (1998) é baseada nas covariâncias empíricas. De modo que, tal aproximação é dada por

$$\bar{\ell}_{BN}(s) = \ell_p(s) + \frac{1}{2}\ \log\ |j_{cc}(s,\widehat{c}_s)| - \log\ |I_c(s,\widehat{c}_s;\widehat{s},\widehat{c})|, \tag{3.10}$$

em que

$$I_c(s,\widehat{c}_s;\widehat{s},\widehat{c}) = \sum_{i=1}^{n}\left[\left(\frac{1}{\widehat{c}_s} - \log(s) - \frac{2(y_i/s)^{\widehat{c}_s}\log(y_i/s)}{1+(y_i/s)^{\widehat{c}_s}} + \log(y_i)\right)\right.$$
$$\left.\times\left(\frac{1}{\widehat{c}} - \log(\widehat{s}) - \frac{2(y_i/\widehat{s})^{\widehat{c}}\log(y_i/\widehat{s})}{1+(y_i/\widehat{s})^{\widehat{c}}} + \log(y_i)\right)\right].$$

A função escore perfilada modificada é obtida a partir de $\bar{u}_{BN}(s) = \partial\bar{\ell}_{BN}(s)/\partial s$. O EMV perfilado $\widehat{\bar{s}}_{BN}$ não apresenta solução em forma fechada, devendo ser obtido numericamente.

Os testes baseados nas estatísticas da razão de verossimilhanças obtidas a partir das log-verossimilhanças perfilada e perfiladas modificadas para testar $H_0 : s = s_0$ contra $H_1 : s \neq s_0$ são, respectivamente, dadas por

$$LR_p = 2\Big\{\ell_p(\widehat{s}_p) - \ell_p(s_0)\Big\},$$
$$LR_{BN1} = 2\Big\{\breve{\ell}_{BN}(\widehat{\breve{s}}) - \breve{\ell}_{BN}(s_0)\Big\},$$
$$LR_{BN2} = 2\Big\{\bar{\ell}_{BN}(\widehat{\bar{s}}) - \bar{\ell}_{BN}(s_0)\Big\} \quad \text{e}$$
$$LR_{CR} = 2\Big\{\ell_{CR}(\widehat{s}) - \ell_{CR}(s_0)\Big\}.$$

Sob condições de regularidade e sob a hipótese nula, as estatísticas LR_p, LR_{BN1}, LR_{BN2} e LR_{CR} são assintoticamente distribuídas com distribuição χ^2_1. Mesmo com a modificação, a aproximação pela distribuição χ^2_1 tem erro de ordem $O(n^{-1})$.

3.6.3 Resultados numéricos

Avaliaremos numericamente a qualidade das inferências realizadas para o parâmetro de interesse $\mu = s$ da distribuição log-logística baseadas nas funções de verossimilhança perfilada e perfiladas modificadas com base nas propostas de Barndorff-Nielsen (1983) e Cox e Reid (1987). A verossimihança perfilada modificada de Barndorff-Nielsen (1983) será obtida através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Os principais objetivos dos estudos de simulação são: (i) comparar os desempenhos dos EMV perfilado e perfilados modificados, considerando as seguintes medidas descritivas dos diferentes estimadores: viés relativo, erro quadrático médio (EQM), assimetria e curtose (Tabela 11), (ii) comparar a média e a variância das estatísticas LR_p, LR_{BN1}, LR_{BN2} e LR_{CR} com a média e a variância da distribuição χ^2_1 (Tabela 12), (iii) verificar a qualidade das aproximações das estatísticas de teste em relação à distribuição χ^2_1 (Figura 5) e (iv) comparar os desempenhos dos testes LR_p, LR_{CR}, LR_{BN1} e LR_{BN2}, através dos tamanhos e poderes (Tabelas 13 e 14). As entradas das Tabelas 13 e 14 são porcentagem. Para estimação pontual e os testes de hipóteses, o número de réplicas de Monte Carlo utilizadas em cada simulação foi 10.000, considerando os seguintes tamanhos amostrais: $n = 20$, 30, 40, 60 e 100. Os níveis nominais utilizados foram $\alpha = 1\%$, 5% e 10%. Os valores considerados para o parâmetro de incômodo foi $c = 4$ e variamos o parâmetro de interesse s em $2; 2, 5$ e 3. As simulações foram realizadas usando a linguagem de programação Ox (DOORNIK, 2009). Para as maximizações das log-verossimilhanças, utilizamos os métodos numéricos BFGS e SQP. O gráfico foi gerado usando a linguagem de programação R (R Core Team, 2020).

Na Tabela 11 apresentamos os desempenhos dos EMV perfilado e perfilados modificados a saber: $\widehat{s}_p$, $\widehat{\widetilde{s}}_{BN}$, $\widehat{\widetilde{s}}_{BN}$ e $\widehat{s}_{CR}$. Sabemos que, sob condições de regularidade os EMV têm distribuição assintótica normal, como apresentam curtose maiores que três, implica que a distribuição desses estimadores têm uma curva leptocúrtica e com o aumento da amostra, as curtoses de todos os estimadores se aproximam de três que é a curtose da distribuição normal. Com relação à assimetria, note que, é sempre positiva, isto é, a distribuição dos estimadores apresentam valores concentrados à esquerda. Além disso, nota-se que os erros quadráticos médios de todos os estimadores são semelhantes. Como esperado, todos os vieses, os erros quadráticos médios, as assimetria e as curtoses diminuem com o aumento do tamanho da amostra. É importante observar que os desempenhos dos estimadores $\widehat{s}_p$, $\widehat{\widetilde{s}}_{BN}$, $\widehat{\widetilde{s}}_{BN}$ e $\widehat{s}_{CR}$ são semelhantes mesmo em amostras menores.

A Tabela 12 apresenta a média e a variância das estatísticas de teste LR_p, LR_{BN1}, LR_{BN2}, LR_{CR} e as mesmas medidas para a distribuição χ^2_1 para o cenário em que o parâmetro de interesse s varia de 2; 2,5 e 3 para tamanhos amostrais de $n = 20$, 30, 40, 60 e 100. Note que, em amostras pequenas a média e a variância das estatísticas LR_p e LR_{CR} são mais elevadas quando comparadas com a média e a variância da distribuição χ^2_1. Por exemplo, para $n = 20$ a média e variância da estatística LR_p é de (1,055 e 2,312). Já para a estatística LR_{CR} é de (1,112 e 2,571). O valor da média e da variância das estatísticas LR_{BN1} e LR_{BN2} ficaram mais próximos das medidas correspondentes da distribuição de referência χ^2_1. Observe que, com o aumento da amostra, a média e a variância das estatísticas estudadas se aproximam da média e da variância da distribuição assintótica χ^2_1 sob a hipótese nula.

Apresentamos na Tabela 13 as taxas de rejeição nula do teste da razão de verossimilhanças perfiladas LR_p, dos testes da razão de verossimilhanças perfilados modificados LR_{BN1}, LR_{BN2} e LR_{CR} para testar a hipótese nula $H_0 : s = s_0$ contra $H_1 : s \neq s_0$, em que $s_0 = 2; 2,5$ e 3 e considerando os seguintes níveis nominais 1%, 5% e 10%. De início, nota-se que os testes das razões de verossimilhanças perfilados ajustados LR_{BN1} e LR_{BN2} são conservativos, apresentando taxas de rejeição abaixo dos valores nominais em todos os tamanhos amostrais e diferentes valores de s. Por exemplo, para $s = 2$, $n = 30$ e nível $\alpha = 10\%$, os testes LR_{BN1} e LR_{BN2} apresentaram taxas de rejeição de $9,79\%$ e $9,83\%$, respectivamente. Já os testes LR_p e LR_{CR} são liberais, isto é, apresentam taxas de rejeição acima dos valores nominais em quase todos os tamanhos amostrais e diferentes valores de s. Por exemplo, para $s = 2$, $n = 30$ e nível $\alpha = 10\%$, os testes LR_p e LR_{CR} apresentaram taxas de rejeição de $10,95\%$ e $11,56\%$, respectivamente. Note também que o teste da razão de verossimilhanças perfiladas modificadas LR_{CR} apresentou desempenho inferior aos testes da razão de verossimilhanças LR_p, LR_{BN1} e LR_{BN2}. Por exemplo, para um teste com nível nominal de 10%, tamanho amostral $n = 20$ e para $s = 2$, o teste LR_{CR} apresentou taxa de rejeição de $11,95\%$, enquanto os testes LR_p, LR_{BN1} e LR_{BN2} apresentaram taxas de $10,99\%$, $9,02\%$ e $8,85\%$, respectivamente. Quando o tamanho amostral aumenta, os testes passam apresentar taxas de rejeição mais próximas de seus respectivos níveis nominais. Por exemplo, para $s = 2,5$, $n = 100$ e $\alpha = 5\%$, os testes LR_p, LR_{BN1}, LR_{BN2} e LR_{CR} apresentaram taxas de rejeição de $4,95\%$, $4,64\%$, $4,61\%$ e $5,05\%$, respectivamente.

Para avaliar a aproximação assintótica das distribuições nula das estatísticas LR_p, LR_{BN1}, LR_{BN2} e LR_{CR}, utilizamos o gráfico de discrepâncias relativas. Quanto mais próxima a curva estiver da ordenada zero, melhor é a aproximação da estatística do teste por sua distribuição

assintótica χ^2_1. A Figura 5 apresenta os gráficos das discrepâncias relativas das quatro estatísticas de teste contra os correspondentes quantis assintóticos para um tamanho de amostra $n = 20$, parâmetro de interesse $s = 2$ e parâmetro de incômodo $c = 4$. As curvas de discrepância para as estatísticas LR_{BN1} e LR_{BN2} estão mais próximas da ordenada zero. Já as curvas de discrepância para as estatísticas LR_p e LR_{CR} estão mais distantes da ordenada zero, o que confirma que os testes baseados nessas estatísticas são liberais. Dessa maneira, concluímos que as estatísticas de testes baseadas nas verossimilhanças perfiladas modificadas LR_{BN1} e LR_{BN2} apresentam aproximações nula melhores aproximadas pela distribuição χ^2_1.

A Tabela 14 apresenta o estudo de poder dos testes considerando a hipótese alternativa para $n = 20$, $c = 4$ e o parâmetro de interesse s variando de 2,1 a 3,1, ao nível nominal de $\alpha = 10\%$. Note que, as simulações correspondem ao cenário apresentado na Tabela 13 para $n = 20$ e $s = 2$. Utilizamos valores críticos exatos estimados para cada estatística ao invés dos valores tabulados, com isso, deixamos todos os testes em um mesmo patamar. Observamos que os poderes dos testes LR_p, LR_{BN1}, LR_{BN2} e LR_{CR} tiveram desempenhos similares. Com isso, concluímos que as verossimilhanças perfiladas modificadas não apresentam perda de poder quando comparada com a verossimilhança perfilada.

A Tabela 15 apresenta os resultados de simulação de 100.000 amostras de Monte Carlo com o objetivo de avaliar os tamanhos empíricos dos testes LR_p, LR_{BN1}, LR_{BN2} e LR_{CR} em pequenos valores para o nível nominal de 0,01% a 50%. Consideramos o cenário em que $n = 20$, $s = 2$ e $c = 4$. Avaliando os resultados, observamos que as diferenças relativas entre as taxas de rejeição empíricas e os níveis nominais dos testes LR_p e LR_{CR} são positivas e mais acentuadas no extremo da cauda da distribuição χ^2_1. Por exemplo, para o teste LR_p, ao nível nominal de 50% esta diferença é de 2,46% $((51229 - 50000)/50000) \times 100$, enquanto para o nível nominal de 0,05% é de 70% $((85 - 50)/50) \times 100$. Para o teste LR_{CR} essa diferença é mais acentuada para as situações citadas acima, as diferenças relativas são de 4,69% $((52344 - 50000)/50000) \times 100$ e 118% $((109 - 50)/50) \times 100$, respectivamente. Já as diferenças relativas entre as taxas de rejeição empíricas e os níveis nominais dos testes LR_{BN1} e LR_{BN2} são negativas e menos acentuadas em relação as estatísticas LR_p e LR_{CR} considerando as duas situações anteriores. Por exemplo, para o teste LR_{BN1} a diferença relativa é de -2,04% $((48981 - 50000)/50000) \times 100$ e -12% $((44 - 50)/50) \times 100$. Para o teste LR_{BN2} a diferença relativa é de -2,55% $((48724 - 50000)/50000) \times 100$ e -24% $((38 - 50)/50) \times 100$, respectivamente.

Tabela 11 – Medidas descritivas de diferentes estimadores de s para $c = 4$.

		$s = 2$				$s = 2,5$				$s = 3,0$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
20	$\widehat{s}_p$	0,004	0,038	0,263	3,275	0,004	0,059	0,263	3,275	0,004	0,086	0,263	3,275
	$\widehat{\breve{s}}_{BN}$	0,003	0,038	0,262	3,275	0,004	0,059	0,261	3,275	0,004	0,086	0,262	3,274
	$\widehat{\bar{s}}_{BN}$	0,004	0,039	0,330	3,940	0,004	0,060	0,303	3,642	0,004	0,086	0,255	3,277
	$\widehat{s}_{CR}$	0,004	0,038	0,264	3,276	0,004	0,060	0,264	3,276	0,003	0,086	0,264	3,276
30	$\widehat{s}_p$	0,003	0,026	0,228	3,184	0,003	0,040	0,228	3,184	0,003	0,058	0,228	3,184
	$\widehat{\breve{s}}_{BN}$	0,003	0,026	0,226	3,181	0,003	0,040	0,228	3,184	0,003	0,058	0,228	3,184
	$\widehat{\bar{s}}_{BN}$	0,003	0,026	0,225	3,186	0,003	0,040	0,227	3,176	0,003	0,058	0,226	3,176
	$\widehat{s}_{CR}$	0,003	0,026	0,228	3,186	0,003	0,040	0,228	3,186	0,003	0,058	0,228	3,186
40	$\widehat{s}_p$	0,002	0,019	0,175	3,162	0,002	0,029	0,175	3,162	0,002	0,042	0,175	3,162
	$\widehat{\breve{s}}_{BN}$	0,002	0,019	0,175	3,161	0,002	0,029	0,175	3,161	0,002	0,042	0,175	3,161
	$\widehat{\bar{s}}_{BN}$	0,002	0,019	0,166	3,147	0,002	0,029	0,169	3,154	0,002	0,042	0,171	3,159
	$\widehat{s}_{CR}$	0,002	0,019	0,176	3,162	0,002	0,029	0,176	3,162	0,002	0,042	0,176	3,162
60	$\widehat{s}_p$	0,001	0,013	0,137	3,102	0,001	0,020	0,137	3,102	0,001	0,029	0,137	3,102
	$\widehat{\breve{s}}_{BN}$	0,001	0,013	0,137	3,102	0,001	0,020	0,133	3,091	0,001	0,029	0,137	3,102
	$\widehat{\bar{s}}_{BN}$	0,001	0,013	0,137	3,100	0,001	0,020	0,133	3,088	0,001	0,029	0,137	3,100
	$\widehat{s}_{CR}$	0,001	0,013	0,137	3,103	0,001	0,020	0,137	3,103	0,001	0,029	0,137	3,103
100	$\widehat{s}_p$	0,001	0,007	0,155	3,052	0,001	0,012	0,155	3,052	0,001	0,017	0,155	3,052
	$\widehat{\breve{s}}_{BN}$	0,001	0,007	0,155	3,052	0,001	0,012	0,155	3,052	0,001	0,017	0,155	3,052
	$\widehat{\bar{s}}_{BN}$	0,001	0,007	0,155	3,051	0,001	0,012	0,155	3,050	0,001	0,017	0,155	3,050
	$\widehat{s}_{CR}$	0,001	0,007	0,156	3,052	0,001	0,012	0,156	3,052	0,001	0,017	0,156	3,052

Fonte: Próprio autor (2021).

Tabela 12 – Médias e variâncias amostrais de diferentes estatísticas de teste e os valores assintóticos destas, $H_0 : s = s_0$, distribuição χ^2_1.

			$s = 2,0$				$s = 2,5$				$s = 3,0$			
n		χ^2	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}
20	média	1,00	1,055	0,945	0,933	1,112	1,055	0,944	0,934	1,112	1,055	0,945	0,936	1,112
	var	2,00	2,312	1,847	1,839	2,571	2,312	1,847	1,820	2,571	2,312	1,848	1,884	2,571
30	média	1,00	1,044	0,971	0,966	1,082	1,044	0,972	0,967	1,082	1,044	0,972	0,967	1,082
	var	2,00	2,177	1,882	1,860	2,340	2,177	1,885	1,863	2,340	2,177	1,885	1,863	2,340
40	média	1,00	1,016	0,963	0,960	1,044	1,016	0,964	0,961	1,044	1,016	0,964	0,961	1,044
	var	2,00	2,123	1,907	1,889	2,240	2,123	1,908	1,899	2,240	2,123	1,908	1,902	2,240
60	média	1,00	1,025	0,990	0,989	1,044	1,025	0,989	0,988	1,044	1,025	0,990	0,989	1,044
	var	2,00	2,144	1,999	1,991	2,223	2,144	1,992	1,985	2,223	2,144	1,999	1,991	2,223
100	média	1,00	0,996	0,976	0,975	1,007	0,996	0,976	0,975	1,007	0,996	0,976	0,975	1,007
	var	2,00	1,997	1,916	1,914	2,041	1,997	1,916	1,914	2,041	1,997	1,916	1,914	2,041

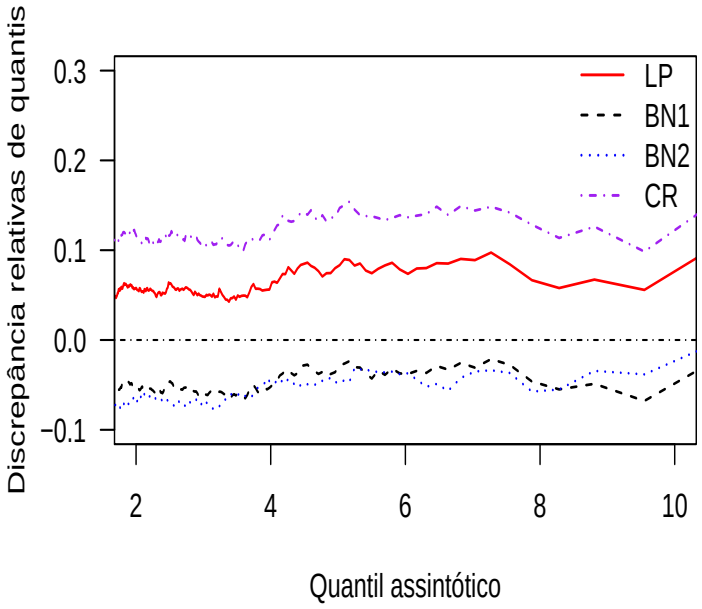
Fonte: Próprio autor (2021).

Tabela 13 – taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : s = s_0$, para $c = 4$.

		$n = 20$				$n = 30$				$n = 40$				$n = 60$				$n = 100$			
s	$\alpha(\%)$	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}
2,0	10	10,99	9,02	8,85	11,95	10,95	9,79	9,83	11,56	10,28	9,43	9,40	10,69	10,25	9,72	9,67	10,51	9,85	9,62	9,57	9,96
	5	5,58	4,47	4,45	6,30	5,73	4,78	4,63	6,15	5,53	4,92	4,88	5,83	5,26	4,90	4,83	5,41	4,95	4,64	4,61	5,05
	1	1,32	0,92	0,83	1,57	1,10	0,86	0,82	1,24	1,05	0,80	0,77	1,24	1,23	1,01	1,01	1,29	0,95	0,83	0,84	0,99
2,5	10	10,99	9,01	8,84	11,95	10,95	9,80	9,84	11,56	10,28	9,44	9,41	10,69	10,25	9,71	9,66	10,51	9,85	9,62	9,57	9,96
	5	5,58	4,47	4,44	6,30	5,73	4,79	4,64	6,15	5,43	4,93	4,89	5,83	5,26	4,89	4,82	5,41	4,95	4,64	4,61	5,05
	1	1,32	0,92	0,82	1,57	1,10	0,86	0,82	1,24	1,05	0,80	0,78	1,24	1,23	1,00	1,00	1,29	0,95	0,83	0,84	0,99
3,0	10	10,99	9,03	8,85	11,95	10,95	9,80	9,84	11,56	10,28	9,44	9,42	10,69	10,25	9,72	9,67	10,51	9,85	9,62	9,57	9,96
	5	5,58	4,47	4,45	6,30	5,73	4,79	4,64	6,15	5,53	4,93	4,90	5,83	5,26	4,90	4,83	5,41	4,95	4,64	4,61	5,05
	1	1,32	0,92	0,83	1,57	1,10	0,86	0,82	1,24	1,05	0,80	0,78	1,24	1,11	1,01	1,01	1,29	0,95	0,83	0,84	0,99

Fonte: Próprio autor (2021).

Figura 5 – Gráfico de discrepância relativa dos quantis da distribuição log-logística para $n = 20$, $c = 4$ e $s = 2$.



Fonte: Próprio autor (2021).

Tabela 15 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo log-logística indexado pelos parâmetros $c = 4$, $s = 2$ e $n = 20$.

$\alpha(\%)$	exato	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}
50	50000	51229	48981	48724	52344
15	15000	16407	14196	13954	17511
10	10000	11219	9359	9118	12151
5	5000	5873	4617	4511	6541
3	3000	3698	2718	2635	4204
1	1000	1286	856	814	1572
0,5	500	682	431	403	836
0,1	100	159	81	82	209
0,05	50	85	44	38	109
0,01	10	24	7	7	30

Fonte: Próprio autor (2021).

Tabela 14 – taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 20$ e nível nominal de 10%.

s	LR_p	LR_{BN1}	LR_{BN2}	LR_{CR}
2,1	13,75	13,77	14,09	13,85
2,2	25,72	25,70	25,85	25,80
2,3	41,22	41,21	41,31	41,49
2,4	56,77	56,70	56,80	56,96
2,5	72,77	72,73	72,26	72,86
2,6	83,13	83,15	82,40	83,22
2,7	90,29	90,30	89,89	90,47
2,8	94,69	94,71	94,02	94,70
2,9	97,65	97,64	97,09	97,66
3,0	98,73	98,73	98,35	98,75
3,1	99,53	99,54	99,27	99,53

Fonte: Próprio autor (2021).

3.7 CASO (iii)

No terceiro caso, estamos considerando c e k como os parâmetros de interesse e s como sendo o parâmetro de incômodo, ou seja, $\boldsymbol{\mu} = (c, k)^\top$ e $\phi = s$. Fixando (c e k) e resolvendo a equação $u_s = 0$, obtemos o EMV restrito de s ($\widehat{s}_{\boldsymbol{\mu}}$). Substituindo s por $\widehat{s}$ na log-verossimilhança original, isto é, em (3.2), obtemos a log-verossimilhança perfilada para $\boldsymbol{\mu}$. Desta maneira, temos que a log-verossimilhança perfilada para o parâmetro $\boldsymbol{\mu}$ pode ser expressa como

$$\ell_p(\boldsymbol{\mu}) = n\log(c) + n\log(k) - nc\log(\widehat{s}_{\boldsymbol{\mu}}) - (k+1)\sum_{i=1}^{n}\log[1 + (y_i/\widehat{s}_{\boldsymbol{\mu}})^c] + (c-1)\sum_{i=1}^{n}\log(y_i).$$

O EMV perfilado $\widehat{\boldsymbol{\mu}}_p$ não possui solução analítica, portanto, deverá ser obtido por um procedimento de otimização não-linear.

Derivamos agora, duas aproximações para a log-verossimilhança perfilada modificada proposta por Barndorff-Nielsen (1983) com base nas propostas de Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998).

A primeira, foi proposta por Fraser e Reid (1995) e Fraser, Reid e Wu (1999). A log-

verossimilhança perfilada é dada por

$$\breve{\ell}_{BN}(\boldsymbol{\mu}) = \ell_p(\boldsymbol{\mu}) + \frac{1}{2}\log|j_{ss}(\boldsymbol{\mu}, \widehat{s}_{\boldsymbol{\mu}})| - \log|\ell_{s;Y}(\boldsymbol{\mu}, \widehat{s}_{\boldsymbol{\mu}})\widehat{V}_s|, \tag{3.11}$$

em que

$$j_{ss}(\boldsymbol{\mu}, \widehat{s}_{\boldsymbol{\mu}}) = -\frac{nc}{\widehat{s}^2_{\boldsymbol{\mu}}} + (k+1)\sum_{i=1}^{n}\left[\frac{c(y_i/\widehat{s}_{\boldsymbol{\mu}})^c(1+c+(y_i/\widehat{s}_{\boldsymbol{\mu}})^c)}{\widehat{s}^2_{\boldsymbol{\mu}}(1+(y_i/\widehat{s}_{\boldsymbol{\mu}})^c)^2}\right]$$

e

$$\ell_{s;Y}(\boldsymbol{\mu}, \widehat{s}_{\boldsymbol{\mu}})\widehat{V}_s = \sum_{i=1}^{n}\left[\frac{y_i c^2(k+1)(y_i/\widehat{s}_{\boldsymbol{\mu}})^{c-1}}{\widehat{s}\ \widehat{s}^2_{\boldsymbol{\mu}}(1+(y_i/\widehat{s}_{\boldsymbol{\mu}})^c)^2}\right].$$

A segunda aproximação foi introduzida por Severini (1998) e a log-verossimilhança perfilada modificada pode ser expressa por

$$\bar{\ell}_{BN}(\boldsymbol{\mu}) = \ell_p(\boldsymbol{\mu}) + \frac{1}{2}\log|j_{ss}(\boldsymbol{\mu}, \widehat{s}_{\boldsymbol{\mu}})| - \log|I_s(\boldsymbol{\mu}, \widehat{s}_{\boldsymbol{\mu}}; \widehat{\boldsymbol{\mu}}, \widehat{s})|, \tag{3.12}$$

em que

$$I_s(\boldsymbol{\mu}, \widehat{s}_{\boldsymbol{\mu}}; \widehat{\boldsymbol{\mu}}, \widehat{s}) = \sum_{i=1}^{n}\left[\left(\frac{c(k(y_i/\widehat{s}_{\boldsymbol{\mu}})^c - 1)}{\widehat{s}_{\boldsymbol{\mu}}(1+(y_i/\widehat{s}_{\boldsymbol{\mu}})^c)}\right)\left(\frac{\widehat{c}(\widehat{k}(y_i/\widehat{s})^{\widehat{c}} - 1)}{\widehat{s}(1+(y_i/\widehat{s})^{\widehat{c}})}\right)\right].$$

Os EMV perfilados modificados de $\boldsymbol{\mu}$ obtidos de (3.11) e (3.12) são denotados, respectivamente, por $\widehat{\breve{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$. Estes estimadores não possuem solução analítica e devem ser encontrados maximizando numericamente as respectivas funções de log-verossimilhanças.

Os testes de hipóteses baseados nas estatísticas da razão de verossimilhanças com base na log-verossimilhança perfilada $\ell_p(\boldsymbol{\mu})$ e nas log-verossimilhanças perfiladas modificadas $\breve{\ell}_{BN}(\boldsymbol{\mu})$ e $\bar{\ell}_{BN}(\boldsymbol{\mu})$ para testar $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ contra $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$ são dadas, respectivamente, por

$$LR_p = 2\{\ell_p(\widehat{\boldsymbol{\mu}}) - \ell_p(\boldsymbol{\mu}_0)\},$$

$$LR_{BN1} = 2\{\breve{\ell}_{BN}(\widehat{\breve{\boldsymbol{\mu}}}) - \breve{\ell}_{BN}(\boldsymbol{\mu}_0)\}$$

e

$$LR_{BN2} = 2\{\bar{\ell}_{BN}(\widehat{\bar{\boldsymbol{\mu}}}) - \bar{\ell}_{BN}(\boldsymbol{\mu}_0)\}.$$

Satisfeitas as condições de regularidade e sob a hipótese nula, as estatísticas LR_p, LR_{BN1} e LR_{BN2} são distribuídas assintoticamente como χ^2_2. Mesmo com a modificação, a aproximação pela distribuição χ^2_2 tem erro de ordem $O(n^{-1})$.

3.7.1 Resultados numéricos

Avaliaremos numericamente a qualidade das inferências realizadas para os parâmetros de interesse $\boldsymbol{\mu} = (c, k)^\top$ da distribuição Burr XII baseadas nas funções de verossimilhança perfilada e perfiladas modificadas com base na proposta de Barndorff-Nielsen (1983) através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Os principais objetivos dos estudos de simulação são: (i) comparar os desempenhos dos EMV perfilado e perfilados modificados, considerando as seguintes medidas descritivas dos diferentes estimadores: viés relativo, erro quadrático médio (EQM), assimetria e curtose (Tabela 16), (ii) comparar a média e a variância das estatísticas LR_p, LR_{BN1} e LR_{BN2} com a média e variância da distribuição χ^2_2 (Tabela 17), (iii) verificar a qualidade das aproximações das estatísticas de teste em relação à distribuição χ^2_2 (Figura 6) e (iv) comparar os desempenhos dos testes LR_p, LR_{BN1} e LR_{BN2}, através dos tamanhos e poderes (Tabelas 18 e 19). As entradas das Tabelas 18 e 19 apresentadas são porcentagem. Para estimação pontual e os testes de hipóteses, o número de réplicas de Monte Carlo utilizadas em cada simulação foi 10.000, considerando os seguintes tamanhos amostrais: $n = 30$, 40, 50, 60 e 100. Os níveis nominais utilizados foram α= 1%, 5% e 10%. Os valores considerados para o parâmetro de interesse $\boldsymbol{\mu}$ foram $c = 1,5$ e variamos k em $0,5$ e $1,5$. Para o parâmetro de incômodo consideramos $s = 2$. As simulações foram realizadas usando a linguagem de programação Ox (DOORNIK, 2009). Para as maximizações das log-verossimilhanças, utilizamos os métodos numéricos BFGS e SQP. O gráfico foi gerado usando a linguagem de programação R (R Core Team, 2020).

Avaliamos os desempenhos dos estimadores baseados nas log-verossimilhanças perfilada e suas versões modificadas e os resultados são apresentados na Tabela 16. Note que, os estimadores $\widehat{\widetilde{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$, de modo geral, apresentaram desempenhos superiores ao estimador perfilado $\widehat{\boldsymbol{\mu}}_p$ com valores menores de viés relativo, erro quadrático médio, assimetria e curtose. Observamos também, que à medida que o valor de k aumenta, os desempenhos dos estimadores considerados melhoram. A assimetria dos estimadores são positivas, indicando que a distribuição dos estimadores estudados apresentam valores concentrados à esquerda. A curtose do estimador $\widehat{\boldsymbol{\mu}}_p$ é bem elevada em amostras menores quando comparado com as curtoses dos estimadores $\widehat{\widetilde{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$. Por exemplo, para $n = 30$ e $k = 0,5$, a curtose do estimador $\widehat{\boldsymbol{\mu}}_p$ é igual a (40,153; 3981,779), enquanto para os estimadores $\widehat{\widetilde{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$ são iguais a (32,403; 40,868) e (22,424; 45,968), respectivamente. Notamos também, que a curtose dos estimadores apresentam valores bem elevados. Deste modo, a distribuição desses estimadores apresentam curva leptocúrtica.

A Tabela 17 apresenta a média e a variância das estatísticas de teste LR_p, LR_{BN1} e LR_{BN2}, estimadas via simulação de Monte Carlo para o caso em que, $c = 1,5$, $s = 2$ e k variando de 0,5 e 1,5 para diferentes valores de n. Comparamos a média e a variância das estatísticas de teste com a média e variância da distribuição χ^2_2. Note que, de modo geral, quando $k = 0,5$ as médias e as variâncias estimadas das estatísticas LR_{BN1} e LR_{BN2} são as que mais se aproximam da média e variância da distribuição de referência. Por outro lado, quando $k = 1,5$ a média e a variância da estatística de teste LR_p é a que mais se aproxima das medidas da distribuição χ^2_2. Quando o tamanho amostral aumenta as médias e as variâncias das estatísticas apresentam comportamentos semelhantes e ficam mais próximas da média e da variância da distribuição χ^2_2.

Na Tabela 18, comparamos as taxas de rejeição nula dos testes baseados nas estatísticas da razão de verossimilhanças perfiladas LR_p e perfiladas modificadas LR_{BN1} e LR_{BN2}, para diferentes tamanhos amostrais e níveis nominais. Observamos que para $k = 0,5$, $c = 1,5$ e $s = 2$, os testes LR_{BN1} e LR_{BN2} têm desempenhos superiores ao teste da razão de verossimilhanças perfiladas LR_p. Por exemplo, para um nível nominal de $\alpha = 10\%$, $n = 40$ e $k = 0,5$, os testes LR_{BN1} e LR_{BN2} apresentam taxas iguais a $9,75\%$ e $9,69\%$, respectivamente, enquanto o teste LR_p apresentou taxa de $11,57\%$. Neste caso podemos dizer que o teste LR_p é liberal, apresentando taxa de rejeição acima do nível nominal. Já os testes LR_{BN1} e LR_{BN2} são conservativos até o tamanho de amostra $n = 40$ e passando a ser ligeiramente liberais a partir do tamanho de amostra $n = 50$. Quando o tamanho amostral aumenta, os testes apresentam desempenhos semelhantes. Por exemplo, para $n = 100$, $k = 0,5$ e nível nominal de $\alpha = 10\%$, os testes LR_p, LR_{BN1} e LR_{BN2} apresentam as seguintes taxas de rejeição: $10,94\%$, $10,52\%$ e $10,54\%$, respectivamente. Outro fato interessante é que, para valores elevados do parâmetro k, todos os testes passam a ser conservativos, isto é, apresentam taxas de rejeição abaixo do nível nominal e o teste LR_p passa a apresentar taxas de rejeição mais próxima do nível nominal, quando comparado com os testes LR_{BN1} e LR_{BN2}. Por exemplo, para $n = 40$, $k = 1,5$ e nível nominal de $\alpha = 5\%$, os testes apresentam as seguintes taxas de rejeição: $LR_p = 4,46\%$, $LR_{BN1} = 3,68\%$ e $LR_{BN2} = 3,73\%$.

A Figura 6 apresenta o gráfico de discrepâncias relativas de quantis das estatísticas LR_p, LR_{BN1} e LR_{BN2} contra os correspondentes quantis assintóticos para o tamanho de amostra de $n = 40$, $c = 1,5$, $s = 2$ e $k = 0,25$. Quanto mais próximo à curva estiver da ordenada zero, melhor é a aproximação nula da estatística de teste por sua distribuição assintótica χ^2_2. As curvas das estatísticas LR_{BN1} e LR_{BN2} são semelhantes e as que mais se aproximam da

ordenada zero quando comparada com a curva da estatística LR_p.

A Tabela 19 apresenta resultados de simulação obtidos sob a hipótese alternativa para $n = 30$, $\alpha = 10\%$ e com $c = k = \delta$ variando de 1,6 a 2,6 foram considerados. Novamente, os testes foram realizados usando valores críticos estimados para cada estatística, pois, os testes baseados nas estatísticas LR_p, LR_{BN1} e LR_{BN2} apresentam diferenças entre as taxas de rejeição para um nível nominal fixado. Foi considerado o teste de $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ contra $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, sendo $\boldsymbol{\mu}_0 = (c_0, k_0)^\top$ com $\delta_0 = c_0 = k_0 = 1,5$ e de modo geral, observamos que os testes LR_{BN1} e LR_{BN2} apresentam poderes similares e foram mais poderosos que o teste LR_p.

Na Tabela 20 apresentamos os resultados de simulação via Monte Carlo de 100.000 réplicas com o objetivo de avaliar os tamanhos empíricos dos testes para valores pequenos do nível nominal. Esta Tabela mostra o número de rejeições para os testes LR_p, LR_{BN1} e LR_{BN2} para o caso em que $n = 30$, $c = 1,5$, $s = 2$ e $k = 0,5$. Note que, de modo geral, as diferenças relativas entre as taxas de rejeição empíricas e os níveis nominais do teste LR_p é positiva e menos acentuadas no extremo da cauda da distribuição χ^2_2. Por exemplo, para o nível nominal de 50% esta diferença é de cerca de $5,8\%$ $((52907 - 50000)/50000) \times 100$, enquanto ao nível nominal de $0,1\%$ esta diferença é de 2%. Por outro lado, as diferenças relativas entre as taxas de rejeição empíricas e os níveis nominais dos testes LR_{BN1} e LR_{BN2} são negativas e mais acentuadas em módulo no extremo da cauda da distribuição de referência. Por exemplo, para o nível nominal de 50% as diferenças são de cerca de $-0,5\%$ e $-0,8\%$, respectivamente, enquanto ao nível nominal de $0,1\%$ as diferenças são de cerca de -50% para ambas estatísticas.

Tabela 16 – Medidas descritivas de diferentes estimadores de $\boldsymbol{\mu} = (c, k)^\top$ para $s = 2$.

		$k = 0,5$				$k = 1,5$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
30	$\widehat{\boldsymbol{\mu}}_p$	0,248; 3,683	1,610; 9350,766	4,860; 62,107	40,153; 3981,779	0,183; 24,796	0,807; 8344,829	9,149; 37,184	179,194; 1611,939
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,171; 0,276	0,898; 0,295	3,998; 4,271	32,403; 40,868	0,157; 0,304	0,484; 3,716	4,278; 3,181	40,474; 18,746
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,186; 0,258	0,948; 0,296	3,514; 4,624	22,424; 45,968	0,168; 0,342	0,557; 4,777	4,957; 3,404	53,112; 18,867
40	$\widehat{\boldsymbol{\mu}}_p$	0,170; 0,365	0,834; 34,627	4,978; 98,366	51,112; 9781,456	0,110; 19,843	0,308; 1351,690	4,255; 58,288	47,889; 3852,488
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,125; 0,210	0,565; 0,191	3,825; 3,515	31,824; 26,118	0,097; 0,377	0,249; 4,255	3,370; 3,345	32,505; 19,748
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,135; 0,198	0,616; 0,184	3,958; 3,407	33,492; 24,782	0,103; 0,425	0,288; 5,832	5,431; 4,171	96,690; 30,013
50	$\widehat{\boldsymbol{\mu}}_p$	0,125; 0,170	0,470; 0,256	4,151; 11,729	43,473; 267,071	0,075; 2,991	0,172; 414,529	1,952; 81,716	11,950; 730,091
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,095; 0,157	0,359; 0,137	3,239; 4,703	27,885; 63,356	0,066; 0,413	0,154; 4,514	1,747; 3,421	9,500; 19,660
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,100; 0,155	0,381; 0,143	3,286; 5,203	26,849; 72,874	0,068; 0,450	0,159; 5,776	1,854; 4,081	10,571; 26,919
60	$\widehat{\boldsymbol{\mu}}_p$	0,099; 0,130	0,378; 0,128	5,867; 6,483	84,248; 117,503	0,055; 7,258	0,127; 209,826	1,574; 67,646	9,057; 534,594
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,076; 0,137	0,301; 0,114	4,760; 4,788	62,114; 65,386	0,049; 0,396	0,117; 4,349	1,516; 4,019	8,365; 26,989
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,078; 0,134	0,302; 0,111	4,253; 4,996	48,817; 79,674	0,050; 0,428	0,119; 5,464	1,574; 4,597	8,908; 33,220
100	$\widehat{\boldsymbol{\mu}}_p$	0,048; 0,059	0,111; 0,038	1,360; 2,060	7,443; 18,102	0,023; 0,374	0,059; 8,376	0,851; 27,109	4,378; 137,390
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,037; 0,066	0,103; 0,037	1,278; 1,971	6,824; 16,300	0,019; 0,303	0,057; 2,771	0,848; 4,906	4,335; 40,310
	$\widehat{\widehat{\boldsymbol{\mu}}}_{BN}$	0,038; 0,065	0,105; 0,038	1,348; 2,025	7,418; 17,456	0,019; 0,315	0,057; 3,233	0,853; 5,743	4,373; 54,021

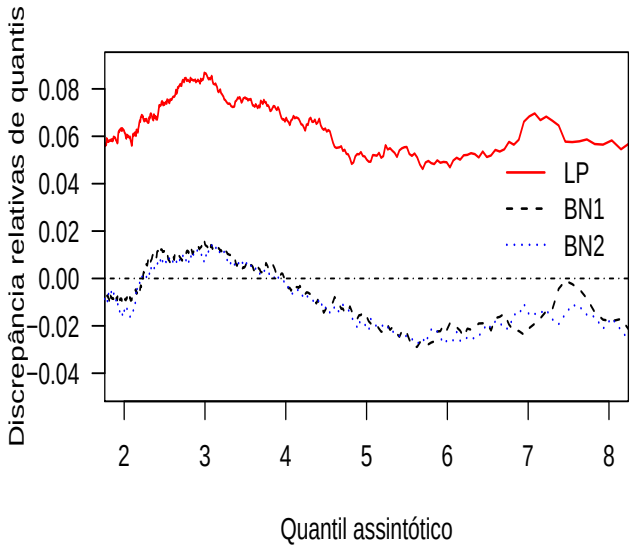
Fonte: Próprio autor (2021).

Tabela 17 – Médias e variâncias amostrais das estatísticas de teste e os valores assintóticos destas, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, distribuição χ^2_2.

			$k = 0,5$			$k = 1,5$		
n		χ^2_2	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
30	média	2,00	2,109	1,924	1,922	1,892	1,742	1,743
	var	4,00	4,211	3,524	3,544	3,899	3,424	3,470
40	média	2,00	2,127	1,991	1,987	1,906	1,783	1,787
	var	4,00	4,525	3,912	3,908	3,861	3,436	3,474
50	média	2,00	2,117	2,019	2,019	1,924	1,834	1,829
	var	4,00	4,410	3,975	3,979	3,624	3,371	3,358
60	média	2,00	2,129	2,057	2,053	1,950	1,869	1,863
	var	4,00	4,538	4,220	4,200	3,765	3,512	3,477
100	média	2,00	2,072	2,035	2,035	2,014	1,970	1,970
	var	4,00	4,323	4,167	4,172	3,878	3,720	3,726

Fonte: Próprio autor (2021).

Figura 6 – Gráfico de discrepância relativa dos quantis da distribuição Burr XII com $n = 40$, $\boldsymbol{\mu} = (0,5; 1,5)$, $s = 2$.



Fonte: Próprio autor (2021).

Tabela 18 – taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, $\boldsymbol{\mu} = (c, k)^\top$ para $(s = 2)$.

		$n = 30$			$n = 40$			$n = 50$			$n = 60$			$n = 100$		
k	$\alpha(\%)$	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
0,5	10	11,11	8,92	8,97	11,57	9,75	9,69	11,40	10,09	10,14	11,88	10,97	10,89	10,94	10,52	10,54
	5	5,76	4,13	4,11	5,74	4,70	4,64	5,97	5,13	5,11	5,99	5,40	5,50	5,86	5,63	5,66
	1	1,07	0,73	0,70	1,24	0,90	0,93	1,20	0,93	0,93	1,35	1,20	1,18	1,25	1,07	1,05
1,5	10	8,62	7,38	7,45	8,61	7,37	7,42	8,92	8,09	8,05	9,03	8,41	8,28	9,91	9,38	9,39
	5	4,60	3,75	3,79	4,46	3,38	3,73	4,11	3,54	3,57	4,52	4,11	4,04	5,02	4,72	4,75
	1	1,03	0,81	0,80	1,05	0,76	0,78	0,91	0,83	0,83	0,90	0,75	0,72	0,88	0,86	0,86

Fonte: Próprio autor (2021).

Tabela 19 – taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 30$ e nível nominal de 10%.

δ	LR_p	LR_{BN1}	LR_{BN2}
1,6	13,75	15,75	15,22
1,7	24,09	26,03	26,72
1,8	38,09	40,02	42,19
1,9	54,18	56,13	58,98
2,0	69,99	70,53	73,48
2,1	80,35	81,18	84,18
2,2	89,35	88,95	90,99
2,3	93,65	93,58	95,10
2,4	96,83	96,58	97,53
2,5	98,08	98,14	98,77
2,6	99,26	99,15	99,40

Fonte: Próprio autor (2021).

Tabela 20 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo Burr XII indexado pelos parâmetros $\boldsymbol{\mu} = (1,5; 0,5)$, $s = 2$ e $n = 30$.

$\alpha(\%)$	exato	LR_p	LR_{BN1}	LR_{BN2}
50	50000	52907	49727	49575
15	15000	16716	13828	13798
10	10000	11162	8800	8839
5	5000	5648	4069	4121
3	3000	3344	2339	2364
1	1000	1099	669	678
0,5	500	503	307	302
0,1	100	102	50	50
0,05	50	47	24	26
0,01	10	10	7	7

Fonte: Próprio autor (2021).

3.7.2 Caso particular: distribuição Lomax

Agora, considere o caso em que, $c = 1$. Dessa maneira, obtemos como caso particular da distribuição Burr XII, a distribuição Lomax (k, s). Assim, temos k como o parâmetro de interesse e s como parâmetro de incômodo, isto é, $\mu = k$ e parâmetro de incômodo dado por $\phi = s$.

Substituindo s por seu EMV restrito, $\widehat{s}_k$ na expressão (3.2), obtemos a log-verossimilhança perfilada para o parâmetro k que é dado por

$$\ell_p(k) = n\log(k) - n\log(\widehat{s}_k) - (k+1)\sum_{i=1}^{n}\log[1+(y_i/\widehat{s}_k)],$$

em que $\widehat{s}_k$ é encontrado como solução $\partial\ell(\boldsymbol{\theta})/\partial s = 0$ para k fixo. Entretanto, $\widehat{s}_k$ não possui solução analítica, a utilização de procedimentos de otimização não-linear restrita é necessária para sua obtenção. Resolvendo a expressão $u_p(k) = 0$ obtemos o EMV perfilado $\widehat{k}_p$. Como $\widehat{k}_p$ não tem solução em forma fechada, dever ser obtido por métodos de otimização não-linear restrita.

$$u_p(k) = \frac{n}{k} - \sum_{i=1}^{n}\log[1+(y_i/\widehat{s}_k)].$$

No que segue, apresentamos as log-verossimilhanças perfiladas modificadas para a distribuição Lomax.

A log-verossimilhança perfilada modificada proposta por Fraser e Reid (1995) e Fraser, Reid e Wu (1999) é dada por

$$\breve{\ell}_{BN}(k) = \ell_p(k) + \frac{1}{2}\log|j_{ss}(k,\widehat{s}_k)| - \log|\ell_{s;Y}(k,\widehat{s}_k)\widehat{V}_s|,$$

em que

$$j_{ss}(k,\widehat{s}_k) = -\frac{n}{\widehat{s}_k^2} + (k+1)\sum_{i=1}^{n}\left[\frac{y_i(2\widehat{s}_k+y)}{\widehat{s}_k^2(\widehat{s}_k+y_i)^2}\right]$$

e

$$\ell_{s;Y}(k,\widehat{s}_k)\widehat{V}_s = \sum_{i=1}^{n}\left[\frac{y_i(k+1)}{\widehat{s}(\widehat{s}_k+y_i)^2}\right].$$

O EMV perfilado modificado de k, neste caso, é denotado por $\breve{\widehat{k}}_{BN}$.

A log-verossimilhança perfilada modificada por Severini (1998) pode ser expressa como

$$\bar{\ell}_{BN}(k) = \ell_p(k) + \frac{1}{2}\log|j_{ss}(k,\widehat{s}_k)| - \log|I_s(k,\widehat{s}_k;\widehat{k},\widehat{s})|,$$

em que

$$I_s(k,\widehat{s}_k;\widehat{k},\widehat{s}) = \sum_{i=1}^{n}\left[\left(\frac{(k(y_i/\widehat{s}_k)-1)}{\widehat{s}_k(1+(y_i/\widehat{s}_k))}\right)\left(\frac{(k(y_i/\widehat{s})-1)}{\widehat{s}(1+(y_i/\widehat{s}))}\right)\right].$$

Neste caso, o EMV perfilado ajustado de k é denotado por $\bar{\widehat{k}}_{BN}$. Os estimadores $\breve{\widehat{k}}_{BN}$ e $\bar{\widehat{k}}_{BN}$ não têm solução em forma fechada e devem ser encontrados através de algoritmos de otimização que maximizam o logaritmo de suas respectivas funções de verossimilhanças.

Os testes baseados nas estatísticas da razão de verossimilhanças obtidas a partir das log-verossimilhanças perfilada e perfiladas modificadas para testar $H_0 : k = k_0$ contra $H_1 : k \neq k_0$ são, respectivamente, dadas por

$$LR_p = 2\{\ell_p(\widehat{k}) - \ell_p(k_0)\},$$
$$LR_{BN1} = 2\{\breve{\ell}_{BN}(\widehat{\breve{k}}) - \breve{\ell}_{BN}(k_0)\} \quad \text{e}$$
$$LR_{BN2} = 2\{\bar{\ell}_{BN}(\widehat{\bar{k}}) - \bar{\ell}_{BN}(k_0)\}.$$

Sob condições de regularidade e sob a hipótese nula, as estatísticas LR_p, LR_{BN1} e LR_{BN2} são distribuídas assintoticamente como χ^2_1. Mesmo com a modificação, a aproximação pela distribuição χ^2_1 tem erro de ordem $O(n^{-1})$.

3.7.3 Resultados numéricos

Avaliaremos numericamente a qualidade das inferências realizadas para o parâmetro de interesse $\mu = k$ e $\phi = s$, como parâmetro de incômodo da distribuição Lomax baseadas nas funções de verossimilhança perfilada e perfiladas modificadas com base na proposta de Barndorff-Nielsen (1983) através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Os principais objetivos dos estudos de simulação são: (i) comparar os desempenhos dos EMV perfilado e perfilados modificados através das medidas descritivas dos diferentes estimadores: viés relativo, erro quadrático médio (EQM), assimetria e curtose (Tabela 21), (ii) comparar a média e a variância das estatísticas LR_p, LR_{BN1} e LR_{BN2} com a média e a variância da distribuição χ^2_1 (Tabela 22), (iii) verificar a qualidade das aproximações das estatísticas de teste em relação à distribuição χ^2_1 (Figura 7) e (iv) comparar os desempenhos dos testes LR_p, LR_{BN1} e LR_{BN2}, através dos tamanhos e poderes (Tabelas 23 e 24). As entradas das Tabelas 23 e 24 são porcentagem. Para estimação pontual e os testes de hipóteses, o número de réplicas de Monte Carlo utilizadas em cada simulação foi 10.000, considerando os seguintes tamanhos amostrais: $n = 30$, 40, 50, 60 e 100. Os níveis nominais utilizados foram α= 1%, 5% e 10%. O valor considerado para o parâmetro de incômodo foi $s = 3$ e variamos o parâmetro de interesse k em 0,15; 0,25 e 0,35. As simulações foram realizadas usando a linguagem de programação Ox (DOORNIK, 2009). Para as maximizações das log-verossimilhanças, utilizamos os métodos numéricos BFGS e SQP. O gráfico foi gerado usando a linguagem de programação R (R Core Team, 2020).

Apresentamos os resultados de simulação para os EMV perfilado e perfilados modificados na Tabela 21. Observamos que os estimadores perfilados modificados apresentam vieses relativos menores quando comparado com o vies relativo do estimador $\widehat{k}_p$ em todos os tamanhos amostrais e diferentes valores do parâmetro de interesse k. Por exemplo, para $n = 30$ e $k = 0,15$, os vieses relativos dos estimadores $\widehat{\widehat{k}}_{BN}$ e $\widehat{\widehat{k}}_{BN}$ são iguais à $5,3\%$ e $5,1\%$, respectivamente, enquanto o viés relativo do estimador $\widehat{k}_p$ é de $6,9\%$. Note também que a assimetria e a curtose do estimador $\widehat{k}_p$ ficaram próximas dos valores de assimetria e curtose dos estimadores ajustados $\widehat{\widehat{k}}_{BN}$ e $\widehat{\widehat{k}}_{BN}$. A medida que o tamanho de amostra aumenta, a curtose dos estimadores se aproximam de três, isto é, a curtose desses estimadores ficam próximas da curtose da distribuição normal. Além disso, observa-se que os erros quadráticos médios dos estimadores $\widehat{\widehat{k}}_{BN}$ e $\widehat{\widehat{k}}_{BN}$ são semelhantes e menores quando comparado com o erro quadrático médio do estimador $\widehat{k}_p$. Como esperado, todos os vieses relativos, os erros quadráticos médios, as assimetrias e as curtoses diminuem conforme o tamanho de amostra aumenta.

A Tabela 22 apresenta a média e a variância das estatísticas de teste LR_p, LR_{BN1} e LR_{BN2}, estimadas por meio da simulação de Monte Carlo com o parâmetro de interesse k variando de $0,15$, $0,25$ e $0,35$ para diferentes tamanhos de amostras. Com o interesse em avaliar a aproximação da distribuição nula das estatísticas pela distribuição χ^2_1, comparamos a média e a variância estimadas das estatísticas LR_p, LR_{BN1} e LR_{BN2} com a média e variância da distribuição de referência χ^2_1. Note que as médias e variâncias estimadas das estatísticas LR_{BN1} e LR_{BN2} são as que mais se aproximam da média e variância da distribuição de referência mesmo em amostras menores comparada com a média e variância da estatística LR_p. Por exemplo, para $n = 30$, $k = 0,35$, a média e a variância das estatísticas LR_{BN1} e LR_{BN2} são de (1,025; 2,087) e (1,026; 2,091), respectivamente, enquanto a média e a variância da estatística LR_p são de 1,058 e 2,226. Entretanto, em amostras grandes, a média e a variância da estatística LR_p se aproxima das demais estatísticas. Isso acontece pelo fato de que assintoticamente essas estatísticas têm comportamentos semelhantes.

Apresentamos na Tabela 23 as taxas de rejeição nula dos testes baseados nas estatísticas da razão de verossimilhanças perfiladas LR_p e perfiladas modificadas LR_{BN1} e LR_{BN2} para testar a hipótese nula $H_0 : k = k_0$ contra a hipótese alternativa $H_1 : k \neq k_0$, diferentes tamanhos amostrais e níveis nominais foram considerados. De início, nota-se que o teste LR_p é bastante liberal, apresentando taxas de rejeição acima dos valores nominais, enquanto os testes LR_{BN1} e LR_{BN2} apresentam taxas de rejeição mais próximas dos rescpectivos níveis nominais. Também observamos que os testes LR_{BN1} e LR_{BN2} apresentam desempenhos superiores ao

teste LR_p. Por exemplo, para $n = 30$, $k = 0,15$ e nível nominal $\alpha = 10\%$, os tamanhos dos testes LR_{BN1} e LR_{BN2} são de $10,53\%$ e $10,44\%$, respectivamente, enquanto o tamanho do teste LR_p é de $10,97\%$. Quando o tamanho amostral aumenta, os testes apresentam desempenhos semelhantes. Por exemplo, para $n = 100$, $k = 0,25$ e nível nominal $\alpha = 10\%$, os testes LR_p, LR_{BN1} e LR_{BN2} apresentam as seguintes taxas de rejeição: $10,10\%$, $10,07\%$ e $10,11\%$, respectivamente.

A Figura 7 apresenta o gráfico de discrepância relativa de quantis das três estatísticas de teste contra os quantis assintóticos correspondentes. Lembrando que quanto mais próxima à curva estiver da ordenada zero, melhor é a aproximação utilizada no teste. As estatísticas de teste LR_{BN1} e LR_{BN2} são as que mais se aproximam da distribuição nula assintótica χ^2_1 quando comparada com a estatística LR_p. As curvas de discrepâncias relativas baseadas nas estatísticas LR_{BN1} e LR_{BN2} são muito semelhantes e encontra-se pouco acima da ordenada zero, confirmando que os testes baseados nessas estatísticas são ligeiramente liberais, enquanto a curva de discrepância relativa da estatística LR_p está muito acima da ordenada zero, indicando assim, que o teste LR_p é o teste mais liberal.

A Tabela 24 apresenta o estudo de poder dos testes considerando a hipótese alternativa para $n = 30$, $s = 3$ e o parâmetro de interesse k variando de 0,30 a 0,85, ao nível nominal de $\alpha = 10\%$. Note que, as simulações correspondem ao cenário apresentado na Tabela 23 para $n = 30$ e $s = 3$. Novamente, porque os testes LR_p, LR_{BN1} e LR_{BN2} apresentam diferenças entre as taxas de rejeição para um nível nominal fixado, os testes foram obtidos utilizando valores críticos estimados para cada estatística, fazendo com que os testes tenham o mesmo patamar. Observamos que os testes LR_{BN1} e LR_{BN2} são similares e apresentam poderes superiores ao teste LR_p.

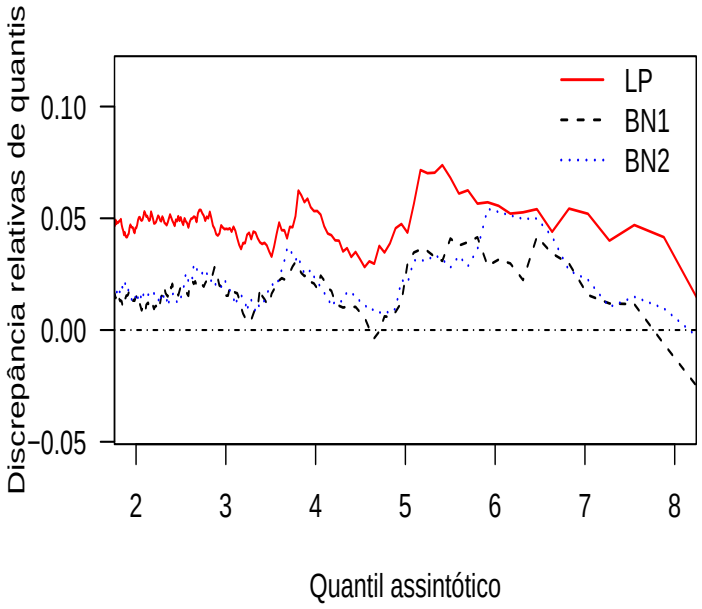
Na Tabela 25 apresentamos os resultados de simulação de 100.000 réplicas de Monte Carlo considerando o caso em que $n = 30$, $k = 0,25$ e $s = 3$. Observe que no extremo da cauda da distribuição χ^2_1 as diferenças relativas entre as taxas de rejeição empíricas dos testes da razão de verossimilhanças e os níveis nominais são mais acentuadas, diferente do que ocorre na região mediana da distribuição. Por exemplo, para $\alpha = 0,01\%$, a estatística LR_p apresenta uma diferença relativa de 60% $((16 - 10)/10) \times 100$, enquanto para o nível nominal de 50% a diferença relativa é de cerca de $1,97\%$. Observe que essa discrepância relativa é menor para as estatísticas LR_{BN1} e LR_{BN2}. Por exemplo, as estatísticas LR_{BN1} e LR_{BN2} apresentaram diferenças relativas para o nível nominal $\alpha = 0,01\%$ iguais à 30% e 40%, respectivamente, enquanto para o nível nominal de 50% estas diferenças são de cerca de $0,83\%$ e $0,76\%$.

Tabela 21 – Medidas descritivas de diferentes estimadores de k para $s = 3$.

		$k = 0,15$				$k = 0,25$				$k = 0,35$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
30	$\widehat{k}_p$	0,069	0,001	1,193	6,294	0,085	0,006	1,744	11,413	0,106	0,016	2,792	27,061
	$\widehat{\widehat{k}}_{BN}$	0,053	0,001	1,176	6,106	0,066	0,005	1,685	10,688	0,083	0,014	2,540	22,113
	$\widehat{\widehat{k}}_{BN}$	0,051	0,001	1,191	6,253	0,065	0,005	1,740	11,494	0,082	0,014	2,674	24,840
40	$\widehat{k}_p$	0,049	0,001	1,064	6,083	0,060	0,004	1,465	10,772	0,072	0,009	1,999	17,869
	$\widehat{\widehat{k}}_{BN}$	0,037	0,001	1,058	6,050	0,046	0,003	1,440	10,431	0,057	0,009	1,939	17,000
	$\widehat{\widehat{k}}_{BN}$	0,036	0,001	1,057	6,010	0,046	0,003	1,453	10,633	0,056	0,009	1,943	16,811
50	$\widehat{k}_p$	0,038	0,001	0,806	4,179	0,045	0,003	0,940	4,714	0,054	0,006	1,111	5,497
	$\widehat{\widehat{k}}_{BN}$	0,029	0,001	0,798	4,146	0,035	0,002	0,932	4,677	0,042	0,006	1,098	5,429
	$\widehat{\widehat{k}}_{BN}$	0,028	0,001	0,802	4,169	0,034	0,002	0,938	4,713	0,042	0,006	1,106	5,474
60	$\widehat{k}_p$	0,030	0,001	0,725	4,188	0,036	0,002	0,884	4,863	0,043	0,005	1,050	5,542
	$\widehat{\widehat{k}}_{BN}$	0,023	0,001	0,727	4,188	0,027	0,002	0,879	4,839	0,033	0,005	1,039	5,497
	$\widehat{\widehat{k}}_{BN}$	0,022	0,001	0,727	4,190	0,027	0,002	0,880	4,852	0,033	0,005	1,042	5,512
100	$\widehat{k}_p$	0,017	0,001	0,552	3,713	0,021	0,001	0,653	4,018	0,025	0,003	0,750	4,299
	$\widehat{\widehat{k}}_{BN}$	0,013	0,001	0,551	3,708	0,016	0,001	0,651	4,009	0,019	0,003	0,746	4,286
	$\widehat{\widehat{k}}_{BN}$	0,013	0,001	0,551	3,711	0,015	0,001	0,653	4,018	0,019	0,003	0,748	4,297

Fonte: Próprio autor (2021).

Figura 7 – Gráfico de discrepância relativa dos quantis da distribuição Lomax para $n = 30$, $k = 0,25$ e $s = 3$.



Fonte: Próprio autor (2021).

Tabela 22 – Médias e variâncias amostrais das estatísticas de teste e os valores assintóticos destas, $H_0 : k = k_0$, distribuição χ^2_1.

			$k = 0,15$			$k = 0,25$			$k = 0,35$		
n		χ^2	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
30	média	1,00	1,045	1,017	1,023	1,050	1,019	1,023	1,058	1,025	1,026
	var	2,00	2,162	2,047	2,078	2,192	2,069	2,085	2,226	2,087	2,091
40	média	1,00	1,028	1,008	1,010	1,037	1,015	1,017	1,039	1,016	1,017
	var	2,00	2,193	2,106	2,114	2,207	2,113	2,117	2,208	2,109	2,108
50	média	1,00	1,018	1,005	1,006	1,021	1,005	1,006	1,024	1,006	1,006
	var	2,00	2,072	2,011	2,016	2,046	2,982	1,986	2,042	1,973	1,975
60	média	1,00	1,021	1,009	1,010	1,020	1,008	1,008	1,020	1,006	1,006
	var	2,00	2,097	2,048	2,053	2,109	2,061	2,065	2,114	2,059	2,062
100	média	1,00	1,007	1,001	1,001	1,010	1,003	1,003	1,012	1,004	1,004
	var	2,00	2,029	2,003	2,003	2,051	2,022	2,022	2,066	2,034	2,035

Fonte: Próprio autor (2021).

Tabela 23 – taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : k = k_0$, para $s = 3$.

		$n = 30$			$n = 40$			$n = 50$			$n = 60$			$n = 100$		
k	$\alpha(\%)$	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
0,15	10	10,97	10,53	10,44	10,54	10,19	10,18	10,64	10,37	10,39	10,32	9,98	10,08	10,02	9,95	10,01
	5	5,41	5,15	5,15	5,46	5,26	5,26	5,32	5,11	5,14	5,16	5,05	5,08	4,95	4,81	4,85
	1	1,07	1,01	1,05	1,15	1,11	1,10	1,16	1,05	1,04	1,13	1,04	1,01	1,01	0,97	0,95
0,25	10	10,78	10,35	10,46	10,78	10,20	10,23	10,57	10,23	10,31	10,03	9,88	9,91	10,10	10,07	10,11
	5	5,50	5,33	5,42	5,45	5,28	5,28	5,64	5,48	5,41	5,28	5,18	5,18	5,19	5,18	5,20
	1	1,21	1,14	1,18	1,11	1,05	1,06	1,10	1,00	1,00	1,23	1,13	1,12	0,97	0,94	0,96
0,35	10	10,78	10,11	10,38	10,59	10,14	10,30	10,39	10,15	10,21	10,15	9,89	9,89	10,18	9,98	9,95
	5	5,64	5,34	5,21	5,36	5,17	5,15	5,66	5,44	5,36	5,34	5,06	5,07	5,18	5,13	5,15
	1	1,29	1,17	1,21	1,21	1,12	1,08	1,04	0,87	0,89	1,18	1,15	1,11	1,04	1,06	1,06

Fonte: Próprio autor (2021).

Tabela 24 – taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 30$ e nível nominal de 10%.

k	LR_p	LR_{BN1}	LR_{BN2}
0,30	19,14	20,93	21,04
0,35	38,31	41,44	41,56
0,40	58,64	61,73	61,85
0,45	73,44	75,74	75,69
0,50	83,72	85,31	85,26
0,55	90,37	91,41	91,36
0,60	94,17	94,93	94,83
0,65	96,04	96,48	96,45
0,70	97,39	97,75	97,80
0,80	98,39	98,56	98,57
0,85	99,05	99,20	99,19

Fonte: Próprio autor (2021).

Tabela 25 – Número de rejeições em 100.000 amostras de Monte Carlo para um modelo Lomax indexado pelos parâmetros $k = 0,25$, $s = 3$ e $n = 30$.

$\alpha(\%)$	exato	LR_p	LR_{BN1}	LR_{BN2}
50	50000	50983	50415	50382
15	15000	15903	15288	15368
10	10000	10713	10224	10323
5	5000	5530	5194	5244
3	3000	3408	3138	3192
1	1000	1204	1069	1070
0,5	500	619	550	559
0,1	100	131	103	108
0,05	50	64	53	56
0,01	10	16	13	14

Fonte: Próprio autor (2021).

4 MODELO DE REGRESSÃO LOG-BURR XII

Neste capítulo, consideraremos o problema em realizar inferências sobre os parâmetros do modelo de regressão log-Burr XII (LBXII) por meio da função de verossimilhança perfilada através de algumas covariáveis. Obteremos a função de verossimilhança perfilada proposta por Barndorff-Nielsen (1983) através das aproximações propostas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Vale ressaltar, que não existem trabalhos publicados sobre refinamentos assintóticos no modelo de regressão log-Burr XII. Neste sentido, nosso trabalho contribui para preencher essa lacuna. Apresentaremos os EMV perfilados modificados e como ficam os testes da razão de verossimilhanças baseados nestas funções considerando o caso em que k é parâmetro de interesse e $(\sigma, \boldsymbol{\beta}^\top)^\top$ são parâmetros de incômodo.

4.1 DEFINIÇÕES BÁSICAS

Um modelo log-linear é expresso como (LAWLESS, 2003).

$$Y = \mu + \sigma Z.$$

Considerando que Y pertence a família de distribuições que possui parâmetro de locação μ $(-\infty < \mu < \infty)$ e parâmetro de escala $\sigma > 0$. As distribuições que pertencem a essa familia têm função densidade de probabiliade dada na forma

$$f(y; \mu, \sigma) = \frac{1}{\sigma} g\left(\frac{y-\mu}{\sigma}\right) \quad -\infty < y < \infty.$$

Considere que $Y = \log(T)$ e que o parâmetro μ depende de um vetor de covariáveis $\boldsymbol{x}$ e σ é um parâmetro constante. Geralmente μ é escrito como $\mu(\boldsymbol{x}) = \boldsymbol{x}_i^\top \boldsymbol{\beta}$, sendo $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ o vetor de parâmetros desconhecidos e $\boldsymbol{x}_i^\top = (x_{i1}, \ldots, x_{ip})$ o vetor de variáveis explicativas.

Neste caso, um modelo de regressão que relaciona Y e o vetor de covariáveis $\boldsymbol{x}$ é o modelo linear representado por

$$Y = \mu(\boldsymbol{x}) + \sigma Z,$$

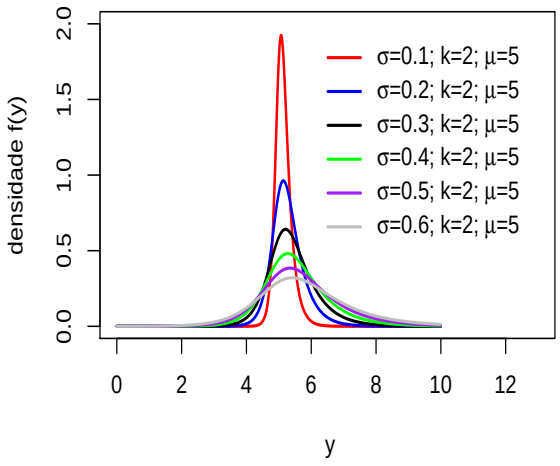
em que $Y = \log(T)$, Z é um erro aleatório e $\mu(\boldsymbol{x}) = \boldsymbol{x}_i^\top \boldsymbol{\beta}$. Note que o modelo é log-linear para T, então, um modelo de regressão linear para Y. Além disso, o vetor de covariáveis $\boldsymbol{x}$ tem efeito multiplicativo em T, isto é, $T = \exp(\mu(\boldsymbol{x}))\exp(\sigma Z)$, portanto, tem efeito linear em Y.

Seja T uma variável aleatória com distribuição dada por (3.1). Considere a variável $Y = \log(T)$ e a reparametrização $c = 1/\sigma$ e $s = \exp(\mu)$. A função densidade de Y pode ser escrita como

$$f(y; k, \sigma, \mu) = \frac{k}{\sigma}\left[1 + \exp\left(\frac{y-\mu}{\sigma}\right)\right]^{-(k+1)} \exp\left(\frac{y-\mu}{\sigma}\right), \tag{4.1}$$

em que $-\infty < y < \infty$, $-\infty < \mu < \infty$, $k > 0$ e $\sigma > 0$. Esta nova distribuição foi denominada de log-Burr XII (LBXII) por Silva et al. (2008). A Figura 8 mostra a função densidade para a variável aleatória Y do modelo LBXII. Note que, conforme o valor de σ aumenta, a densidade torna-se mais "achatada".

Figura 8 – Função densidade do modelo LBXII para diferentes valores de σ para $\mu = 5$ e $k = 2$.



Fonte: Próprio autor (2021).

A função de distribuição acumulada e função de risco são dadas, respectivamente, por

$$F(y; k, \sigma, \mu) = 1 - \left[1 + \exp\left(\frac{y-\mu}{\sigma}\right)\right]^{-k}$$

e

$$h(y; k, \sigma, \mu) = \frac{k \exp\left(\frac{y-\mu}{\sigma}\right)}{\sigma\left(1 + \exp\left(\frac{y-\mu}{\sigma}\right)\right)}.$$

Silva et al. (2008) mostra que a média de Y é dada por

$$E(Y) = s + \sigma[\psi(1) - \psi(k)], \quad \text{se} \quad k, \sigma > r,$$

em que $\psi(\cdot)$ é a função digama (LAWLESS, 2003).

Considerando que Y dado $\boldsymbol{x}$ tem distribuição LBXII. A expressão (4.1) pode ser representado por

$$y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \sigma z_i, \quad i = 1, \ldots, n,$$

em que $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ é um vetor de parâmetros desconhecidos, $k > 0$, $\sigma > 0$ e $\boldsymbol{x}_i^\top = (x_{i1}, \ldots, x_{ip})$ é um vetor de variáveis explicativas. A função densidade de Z para este caso é expressa por

$$f(z; k) = k\left[1 + \exp(z)\right]^{-(k+1)} \exp(z), \quad -\infty < z > \infty.$$

em que $z = \frac{y-\mu}{\sigma}$.

4.2 ESTIMAÇÃO DOS PARÂMETROS

O logaritmo da função de verossimilhança (log-verossimilhança) para $\boldsymbol{\theta} = (k, \sigma, \boldsymbol{\beta}^\top)^\top$ é dada por

$$\ell(\boldsymbol{\theta}) = n\log(k) - n\log(\sigma) - (k+1)\sum_{i=1}^{n} \log\left[1 + \exp(z_i)\right] + \sum_{i=1}^{n} z_i, \tag{4.2}$$

em que $z_i = \frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}$, $i = 1, \ldots, n$.

O vetor escore $u(\boldsymbol{\theta}) = (u_k, u_\sigma, u_{\boldsymbol{\beta}})^\top$ é dado por

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial k} = \frac{n}{k} - \sum_{i=1}^{n} \log\left[1 + \exp(z_i)\right],$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \sigma} = -\frac{n}{\sigma} + \frac{(k+1)}{\sigma}\sum_{i=1}^{n}\left[\frac{z_i \exp(z_i)}{1 + \exp(z_i)}\right] - \frac{1}{\sigma}\sum_{i=1}^{n} z_i$$

e

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta_t} = \frac{(k+1)}{\sigma}\sum_{i=1}^{n}\left[\frac{x_{it}\exp(z_i)}{1 + \exp(z_i)}\right] - \frac{1}{\sigma}\sum_{i=1}^{n} x_{it}, \quad t = 1, \ldots, p.$$

Os EMV são obtidos a partir da resolução do sistema de equações $\partial\ell(\boldsymbol{\theta})/\partial k = 0$, $\partial\ell(\boldsymbol{\theta})/\partial\sigma = 0$ e $\partial\ell(\boldsymbol{\theta})/\partial\beta_t = 0$. Porém, não é possível obter uma forma analítica fechada

para encontrar esses estimadores. Neste caso, temos que recorrer a métodos numéricos para resolver o sistema de equações, como, por exemplo, o algoritmo de Newton (Newton-Raphson, escore de Fisher) ou um algoritmo quasi-Newton (BFGS). Para maiores detalhes sobre tais algoritmos, ver Nocedal e Wright (2006).

A matriz de informação observada do modelo de regressão LBXII de dimensão $(p+2)\times(p+2)$ é dada por

$$j(\boldsymbol{\theta}) = \begin{pmatrix} j_{\sigma\sigma} & j_{\sigma\beta} & j_{\sigma k} \\ j_{\beta\sigma} & j_{\beta\beta} & j_{\beta k} \\ j_{k\sigma} & j_{k\beta} & j_{kk} \end{pmatrix},$$

em que

$$-\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\sigma\partial\sigma^\top} = -\frac{n}{\sigma^2} - \frac{2}{\sigma^2}\sum_{i=1}^{n} z_i + \frac{(k+1)}{\sigma^2}\sum_{i=1}^{n}\left[\frac{z_i^2\exp(z_i)}{(1+\exp(z_i))^2} + \frac{2z_i\exp(z_i)}{(1+\exp(z_i))^2} + \frac{2z_i\exp(2z_i)}{(1+\exp(z_i))^2}\right],$$

$$-\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\beta_t\partial\sigma} = -\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\sigma\partial\beta_t} = \frac{(k+1)}{\sigma^2}\sum_{i=1}^{n}\left[\frac{x_{it}z_i\exp(z_i)}{(1+\exp(z_i))^2} + \frac{x_{it}\exp(z_i)}{(1+\exp(z_i))^2} + \frac{x_{it}\exp(2z_i)}{(1+\exp(z_i))^2}\right] - \frac{1}{\sigma^2}\sum_{i=1}^{n} x_{it},$$

$$-\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial k\partial\sigma} = -\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\sigma\partial k} = -\frac{1}{\sigma}\sum_{i=1}^{n}\left[\frac{z_i\exp(z_i)}{1+\exp(z_i)}\right],$$

$$-\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\beta_t\partial\beta_s} = \frac{(k+1)}{\sigma^2}\sum_{i=1}^{n}\left[\frac{x_{it}x_{is}\exp(z_i)}{(1+\exp(z_i))^2}\right],$$

$$-\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial k\partial\beta_t} = -\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\beta_t\partial k} = -\frac{1}{\sigma}\sum_{i=1}^{n}\left[\frac{x_{it}\exp(z_i)}{1+\exp(z_i)}\right]$$

e

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

em que $s = 1,\ldots,p$.

Os elementos da matriz da informação observada são uteis para obtenção dos diferentes ajustes da função de verossimilhança perfilada.

Intervalos de confiança e testes de hipóteses podem ser obtidos usando o fato de que a distribuição assintótica para os EMV é a distribuição normal com média $\boldsymbol{\theta}$ e matriz de covariância dada pelo inverso da matriz de informação de Fisher (SEN; SINGER, 1993). A matriz de covariância assintótica é dada por $\boldsymbol{I}^{-1}(\boldsymbol{\theta})$ com $\boldsymbol{I}(\boldsymbol{\theta}) = E[\ddot{\boldsymbol{L}}(\boldsymbol{\theta})]$, sendo $\ddot{\boldsymbol{L}}(\boldsymbol{\theta}) = -\left[\partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top\right]$. Entretanto, em muitos casos é complicado obter a matriz de informação de Fisher $\boldsymbol{I}(\boldsymbol{\theta})$. Assim, pode-se usar a matriz hessiana $-\ddot{\boldsymbol{L}}(\boldsymbol{\theta})$, avalianda em $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ que é um estimador consistente. Desta maneira, a distribuição assintótica para $\widehat{\boldsymbol{\theta}}$ é dada por $\widehat{\boldsymbol{\theta}}^\top \sim N_{(p+2)}\left\{\widehat{\boldsymbol{\theta}}^\top; \ddot{\boldsymbol{L}}(\boldsymbol{\theta})^{-1}\right\}$, em que $\ddot{\boldsymbol{L}}(\boldsymbol{\theta})$ é a matriz de informação observada.

4.3 VEROSSIMILHANÇA PERFILADA MODIFICADA PARA O MODELO DE REGRESSÃO LBXII

Considere o caso em que k é parâmetro de interesse e $(\sigma, \boldsymbol{\beta}^\top)^\top$ são os parâmetros de incômodo, isto é, $\mu = k$ e $\boldsymbol{\phi} = (\sigma, \boldsymbol{\beta}^\top)^\top$.

Fixando k e resolvendo a equação $u(\boldsymbol{\phi}) = 0$, obtemos os estimadores de máxima verossimilhança (EMV) restritos de $\boldsymbol{\phi}$ $(\widehat{\boldsymbol{\phi}}_k)$. Substituindo $\boldsymbol{\phi}$ por $\widehat{\boldsymbol{\phi}}$ na log-verossimilhança original, isto é, em (4.2), obtemos a log-verossimilhança perfilada para k. A log-verossimilhança perfilada é dada por

$$\ell_p(k) = n\log(k) - n\log(\widehat{\sigma}_k) - (k+1)\sum_{i=1}^{n}\log\left[1+\exp(\widehat{z}_i)\right] + \sum_{i=1}^{n}\widehat{z}_i,$$

em que $\widehat{z}_i = \frac{y_i - \boldsymbol{x}_i^\top\widehat{\boldsymbol{\beta}}_k}{\widehat{\sigma}_k}$, $i = 1, \ldots, n$.

O EMV perfilado de $\widehat{k}_p$ é obtido como solução de $u_p(k) = 0$ e não tem solução em forma fechada. Logo, para encontrarmos o valor de $\widehat{k}_p$ que maximiza $\ell_p(k)$ é necessário a utilização de um procedimento de otimização não-linear.

Neste trabalho, daremos ênfase às aproximações para log-verossimilhança perfilada de Barndorff-Nielsen (1983) sugeridas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998) para o modelo de regressão LBXII.

A aproximação proposta por Fraser e Reid (1995) e Fraser, Reid e Wu (1999) é baseada em uma aproximação aproximadamente ancilar. A log-verossimilhança perfilada ajustada é dada por

$$\breve{\ell}_{BN}(k) = \ell_p(k) + \frac{1}{2}\log|j_{\phi\phi}(k, \widehat{\boldsymbol{\phi}}_k)| - \log|\ell_{\boldsymbol{\phi};Y}(k, \widehat{\boldsymbol{\phi}}_k)\widehat{V}_{\boldsymbol{\phi}}|, \tag{4.3}$$

em que $j_{\phi\phi}(k, \widehat{\boldsymbol{\phi}}_k)$ é o bloco da matriz de informação observada para os parâmetros de incômodo $(\sigma, \boldsymbol{\beta}^\top)^\top$, $\ell_{\boldsymbol{\phi};Y}(k, \widehat{\boldsymbol{\phi}}_k) = \partial \ell_{\boldsymbol{\phi}}(\mu, \boldsymbol{\phi})/\partial Y^\top$, $\ell_{\boldsymbol{\phi}}(\mu, \boldsymbol{\phi})$ é a função escore avaliada em $\boldsymbol{\phi} = (\sigma, \boldsymbol{\beta}^\top)^\top$, $Y^\top = (y_1, \ldots, y_n)$ e

$$\widehat{V}_{\boldsymbol{\phi}} = \left(-\frac{F(y_1; \widehat{k}, \widehat{\boldsymbol{\phi}})/\partial \widehat{\boldsymbol{\phi}}}{f(y_1; \widehat{k}, \widehat{\boldsymbol{\phi}})}, \ldots, -\frac{F(y_n; \widehat{k}, \widehat{\boldsymbol{\phi}})/\partial \widehat{\boldsymbol{\phi}}}{f(y_n; \widehat{k}, \widehat{\boldsymbol{\phi}})} \right)^\top,$$

em que $F(y_j; k, \phi)$ é a distribuição acumulada de y_j e $f(y_j; k, \phi)$ é a função densidade de probabilidade de y_j, $j = 1, \ldots, n$. As matrizes $j_{\phi\phi}(k, \widehat{\boldsymbol{\phi}}_k)$ e $\ell_{\boldsymbol{\phi};Y}(k, \widehat{\boldsymbol{\phi}}_k)\widehat{V}_{\boldsymbol{\phi}}$ têm dimensão $(p+1) \times (p+1)$.

No modelo de regressão LBXII, as quantidades acima são dadas por

$$j_{\phi\phi}(k, \widehat{\boldsymbol{\phi}}_k) = \begin{pmatrix} j_{\sigma\sigma}(k, \widehat{\boldsymbol{\phi}}_k) & j_{\sigma\beta_t}(k, \widehat{\boldsymbol{\phi}}_k) \\ j_{\beta_t\sigma}(k, \widehat{\boldsymbol{\phi}}_k) & j_{\beta_t\beta_s}(k, \widehat{\boldsymbol{\phi}}_k) \end{pmatrix},$$

sendo

$$j_{\sigma\sigma}(k, \widehat{\boldsymbol{\phi}}_k) = -\frac{n}{\widehat{\sigma}_k^2} - \frac{2}{\widehat{\sigma}_k^2}\sum_{i=1}^{n} \widehat{z}_i + \frac{(k+1)}{\widehat{\sigma}_k^2}\sum_{i=1}^{n}\left[\frac{\widehat{z}_i^{\,2}\exp(\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} + \frac{2\widehat{z}_i\exp(\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} + \frac{2\widehat{z}_i\exp(2\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} \right],$$

$$j_{\beta_t\sigma}(k, \widehat{\boldsymbol{\phi}}_k) = \frac{(k+1)}{\widehat{\sigma}_k^2}\sum_{i=1}^{n}\left[\frac{x_{it} zi\exp(\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} + \frac{x_{it}\exp(\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} + \frac{x_{it}\exp(2\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} \right] - \frac{1}{\widehat{\sigma}_k^2}\sum_{i=1}^{n} x_{it}$$

e

$$j_{\beta_t\beta_s}(k, \widehat{\boldsymbol{\phi}}_k) = \frac{(k+1)}{\widehat{\sigma}_k^2}\sum_{i=1}^{n}\left[\frac{x_{it}x_{is}\exp(\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} \right].$$

A quantidade $\ell_{\boldsymbol{\phi};Y}(k, \widehat{\boldsymbol{\phi}}_k)\widehat{V}_{\boldsymbol{\phi}}$ é dada por

$$\ell_{\boldsymbol{\phi};Y}(k, \widehat{\boldsymbol{\phi}}_k)\widehat{V}_{\boldsymbol{\phi}} = \begin{pmatrix} \ell_{\sigma;Y}(k, \widehat{\sigma}_k)\widehat{V}_\sigma & \ell_{\sigma;Y}(k, \widehat{\sigma}_k)\widehat{V}_{\beta_t} \\ \ell_{\beta_t;Y}(k, \widehat{\boldsymbol{\beta}}_k)\widehat{V}_\sigma & \ell_{\beta_t;Y}(k, \widehat{\boldsymbol{\beta}}_k)\widehat{V}_{\beta_t} \end{pmatrix},$$

em que

$$\ell_{\sigma;Y}(k, \widehat{\sigma}_k)\widehat{V}_\sigma = \sum_{i=1}^{n}\left[\frac{(k+1)}{\widehat{\sigma}_k^2}\left[\frac{\widehat{z}_i\exp(\widehat{z}_i)}{1+\exp(\widehat{z}_i)} - \frac{\widehat{z}_i\exp(2\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} + \frac{\exp(\widehat{z}_i)}{1+\exp(\widehat{z}_i)} \right] - \frac{1}{\widehat{\sigma}_k^2} \times \widehat{\xi}_i \right],$$

$$\ell_{\sigma;Y}(k, \widehat{\sigma}_k)\widehat{V}_{\beta_t} = \sum_{i=1}^{n}\left[\frac{(k+1)}{\widehat{\sigma}_k^2}\left[\frac{\widehat{z}_i\exp(\widehat{z}_i)}{1+\exp(\widehat{z}_i)} - \frac{\widehat{z}_i\exp(2\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} + \frac{\exp(\widehat{z}_i)}{1+\exp(\widehat{z}_i)} \right] - \frac{1}{\widehat{\sigma}_k^2} \times x_{it} \right],$$

$$\ell_{\beta_t;Y}(k, \widehat{\boldsymbol{\beta}}_k)\widehat{V}_\sigma = \sum_{i=1}^{n}\left[\frac{(k+1)}{\widehat{\sigma}_k^2}\left[\frac{x_{it}\exp(\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2} \right] \times \widehat{\xi}_i \right]$$

e

$$\ell_{\beta_t;Y}(k,\widehat{\boldsymbol{\beta}}_k)\widehat{V}_{\beta_t}=\sum_{i=1}^{n}\left[\frac{(k+1)}{\widehat{\sigma}_k^2}\left[\frac{x_{it}\exp(\widehat{z}_i)}{(1+\exp(\widehat{z}_i))^2}\right]\times x_{it}\right],$$

sendo $\widehat{\xi}_i=\frac{(y_i-\boldsymbol{x}_i^\top\widehat{\beta})}{\widehat{\sigma}}$.

A log-verossimilhança perfilada modificada proposta por Severini (1998) é dada por

$$\bar{\ell}_{BN}(k)=\ell_p(k)+\frac{1}{2}\log|j_{\phi\phi}(k,\widehat{\boldsymbol{\phi}}_k)|-\log|I_{\boldsymbol{\phi}}(k,\widehat{\boldsymbol{\phi}}_k;\widehat{k},\widehat{\boldsymbol{\phi}})|, \tag{4.4}$$

em que, $I_{\boldsymbol{\phi}}(k,\widehat{\boldsymbol{\phi}}_k;\widehat{k},\widehat{\boldsymbol{\phi}})=\sum_{i=1}^{n}\ell_{\boldsymbol{\phi}}^{(i)}(k,\widehat{\boldsymbol{\phi}}_k)\ell_{\boldsymbol{\phi}}^{(i)}(\widehat{k},\widehat{\boldsymbol{\phi}})^\top$, sendo $\ell_{\boldsymbol{\phi}}^{(i)}$ a função escore baseada na i-ésima observação, avaliado em $(k,\widehat{\boldsymbol{\phi}}_k)$ e $(\widehat{k},\widehat{\boldsymbol{\phi}})$, respectivamente. A matriz $I_{\boldsymbol{\phi}}(k,\widehat{\boldsymbol{\phi}}_k;\widehat{k},\widehat{\boldsymbol{\phi}})$ têm dimensão $(p+1)\times(p+1)$.

No modelo de regressão LBXII, temos que

$$I_{\boldsymbol{\phi}}(k,\widehat{\boldsymbol{\phi}}_k;\widehat{k},\widehat{\boldsymbol{\phi}})=\begin{pmatrix} I_{\sigma\sigma} & I_{\sigma\beta_t} \\ I_{\beta_t\sigma} & I_{\beta_t\beta_s}\end{pmatrix},$$

em que

$$I_{\sigma\sigma}=\sum_{i=1}^{n}\left[\left(-\frac{1}{\widehat{\sigma}_k}+\frac{(k+1)}{\widehat{\sigma}_k}\left[\frac{\widehat{z}_i\exp(\widehat{z}_i)}{1+\exp(\widehat{z}_i)}\right]-\frac{\widehat{z}_i}{\widehat{\sigma}_k}\right)\left(-\frac{1}{\widehat{\sigma}}+\frac{(\widehat{k}+1)}{\widehat{\sigma}}\left[\frac{\widehat{\xi}_i\exp(\widehat{\xi}_i)}{1+\exp(\widehat{\xi}_i)}\right]-\frac{\widehat{\xi}_i}{\widehat{\sigma}}\right)\right],$$

$$I_{\sigma\beta_t}=\sum_{i=1}^{n}\left[\left(-\frac{1}{\widehat{\sigma}_k}+\frac{(k+1)}{\widehat{\sigma}_k}\left[\frac{\widehat{z}_i\exp(\widehat{z}_i)}{1+\exp(\widehat{z}_i)}\right]-\frac{\widehat{z}_i}{\widehat{\sigma}_k}\right)\left(\frac{(\widehat{k}+1)}{\widehat{\sigma}}\left[\frac{x_{it}\exp(\widehat{\xi}_i)}{1+\exp(\widehat{\xi}_i)}\right]-\frac{x_{it}}{\widehat{\sigma}}\right)\right],$$

$$I_{\beta_t\sigma}=\sum_{i=1}^{n}\left[\left(\frac{(k+1)}{\widehat{\sigma}_k}\left[\frac{x_{it}\exp(\widehat{z}_i)}{1+\exp(\widehat{z}_i)}\right]-\frac{x_{it}}{\widehat{\sigma}_k}\right)\left(-\frac{1}{\widehat{\sigma}}+\frac{(\widehat{k}+1)}{\widehat{\sigma}}\left[\frac{\widehat{\xi}_i\exp(\widehat{\xi}_i)}{1+\exp(\widehat{\xi}_i)}\right]-\frac{\widehat{\xi}_i}{\widehat{\sigma}}\right)\right]$$

e

$$I_{\beta_t\beta_s}=\sum_{i=1}^{n}\left[\left(\frac{(k+1)}{\widehat{\sigma}_k}\left[\frac{x_{it}\exp(\widehat{z}_i)}{1+\exp(\widehat{z}_i)}\right]-\frac{x_{it}}{\widehat{\sigma}_k}\right)\left(\frac{(\widehat{k}+1)}{\widehat{\sigma}}\left[\frac{x_{is}\exp(\widehat{\xi}_i)}{1+\exp(\widehat{\xi}_i)}\right]-\frac{x_{is}}{\widehat{\sigma}}\right)\right].$$

Em que $\widehat{\sigma}$, $\widehat{k}$ e $\widehat{\boldsymbol{\beta}}^\top$ são, respectivamente, os EMV de σ, k e $\boldsymbol{\beta}^\top$. Os EMV perfilados modificados de k obtidos em (4.3) e (4.4) são denominados, respectivamente, por $\breve{\widehat{k}}_{BN}$ e $\bar{\widehat{k}}_{BN}$. Novamente, os EMV não possuem solução analítica e devem ser obtidos numericamente por meio da maximização do logaritmo das suas respectivas funções de log-verossimilhanças.

As estatísticas da razão de verossimilhanças baseadas na função de verossimilhança perfilada $\ell_p(k)$ e nas funções de verossimilhanças perfiladas modificadas $\breve{\ell}_{BN}(k)$ e $\bar{\ell}_{BN}(k)$ para testar $H_0:k=k_0$ contra $H_1:k\neq k_0$ são dadas, respectivamente, por

$$LR_p = 2\{\ell_p(\widehat{k}) - \ell_p(k_0)\},$$

$$LR_{BN1} = 2\{\breve{\ell}_{BN}(\widehat{\breve{k}}) - \breve{\ell}_{BN}(k_0)\}$$

e

$$LR_{BN2} = 2\{\bar{\ell}_{BN}(\widehat{\bar{k}}) - \bar{\ell}_{BN}(k_0)\}.$$

Satisfeitas as condições de regularidade e sob a hipótese nula, as estatísticas da razão de verossimilhanças LR_p, LR_{BN1} e LR_{BN2} são distribuídas assintoticamente como χ^2_1. Mesmo com a modificação, a aproximação pela distribuição χ^2_1 tem erro de ordem $O(n^{-1})$.

4.4 RESULTADOS NUMÉRICOS

Avaliaremos numericamente a qualidade das inferências realizadas para o parâmetro de interesse k do modelo de regressão LBXII baseadas nas funções de verossimilhança perfilada e perfiladas modificadas com base na proposta de Barndorff-Nielsen (1983) através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Os principais objetivos dos estudos de simulação são: (i) comparar os desempenhos dos EMV perfilado e perfilados modificados, considerando as seguintes medidas descritivas dos diferentes estimadores: viés relativo, erro quadrático médio (EQM), assimetria e curtose (Tabela 26), (ii) comparar a média e a variância das estatísticas da razão de verossimilhanças perfilada (LR_p) e razão de verossimilhanças perfiladas modificadas (LR_{BN1} e LR_{BN2}) com a média e a variância da distribuição χ^2_1 (Tabela 27), (iii) verificar a qualidade das aproximações das estatísticas de teste em relação à distribuição χ^2_1 (Figura 9) e (iv) comparar os desempenhos dos testes da razão de verossimilhanças baseados nas estatísticas LR_p, LR_{BN1} e LR_{BN2}, através dos tamanhos e poderes (Tabelas 28 e 29). As entradas das Tabelas 28 e 29 apresentadas são porcentagem. Para estimação pontual e os testes de hipóteses, o número de réplicas de Monte Carlo utilizadas em cada simulação foi de 10.000, considerando os seguintes tamanhos amostrais: $n = 35$, 45, 55, 65 e 100. Os valores considerados para o parâmetro de interesse k foram de $0,25$ e $0,5$. Para os parâmetros de incômodo foram considerados $\sigma = 0,5$ e assumimos $\mu_i = \beta_0 + \beta_1 x_i$ em que x_i foi gerado de uma distribuição uniforme no intervalo (0,1) com $\beta_0 = -1$ e $\beta_1 = 1$. Para cada tamanho de amostra e cada nível nominal considerado, calculamos as taxas de rejeição de cada teste, isto é, estimamos, via simulação $P(LR_p > x_\alpha)$, $P(LR_{BN1} > x_\alpha)$ e $P(LR_{BN2} > x_\alpha)$ em que x_α é o quantil $(1-\alpha)$ apropriado da distribuição

χ^2_p. As simulações foram realizadas usando a linguagem de programação Ox (DOORNIK, 2009). Para as maximizações dos logaritmos das funções de verossimilhanças, utilizamos os métodos numéricos BFGS e SQP. O gráfico foi realizado usando a linguagem de programação R (R Core Team, 2020).

A Tabela 26 contém estatísticas descritivas de diferentes estimadores. Observamos que os estimadores perfilados modificados $\widehat{\widehat{k}}_{BN}$ e $\widehat{\widehat{k}}_{BN}$ apresentam desempenhos semelhantes e superiores quando comparado com o EMV perfilado $\widehat{k}_p$. Por exemplo, para $n = 35$ e $k = 0,25$, o erro quadrático médio do estimador $\widehat{k}_p$ é de 0,072, enquanto os erros quadráticos médios dos estimadores perfilados ajustados $\widehat{\widehat{k}}_{BN}$ e $\widehat{\widehat{k}}_{BN}$ foram de 0,038 e 0,043, respectivamente. Satisfeita as condições de regularidade, os EMV têm distribuição assintótica normal. Nota-se que a curtose de todos os estimadores são maiores que três, isto implica que a distribuição desses estimadores têm uma curva leptocúrtica. A assimetria é positiva, isto é, a distribuição dos estimadores apresentam valores concentrados à esquerda. À medida que o valor de k aumenta, os vieses relativos, os erros quadráticos médios, as assimetrias e as curtoses dos estimadores ficam mais elevados. Observe também que, de modo geral, a assimetria e a curtose dos estimadores perfilados ajustados são menores quando comparado com a assimetria e a curtose do estimador perfilado. Como esperado, à medida que o tamanho da amostra aumenta, as medidas dos estimadores $\widehat{k}_p$, $\widehat{\widehat{k}}_{BN}$ e $\widehat{\widehat{k}}_{BN}$ passam a ter desempenhos semelhantes. Por exemplo, para $n = 100$ e $k = 0,25$, o viés relativo e o erro quadrático médio dos estimadores $\widehat{k}_p$, $\widehat{\widehat{k}}_{BN}$ e $\widehat{\widehat{k}}_{BN}$ são (0,037 e 0,008), (0,047 e 0,007) e (0,044 e 0,007), respectivamente.

A Tabela 27 apresenta a média e a variância das estatísticas de teste LR_p, LR_{BN1} e LR_{BN2}, obtidas via simulação de Monte Carlo para o caso em que $k = 0,25$ e $k = 0,50$, para diferentes valores de n. Aqui, comparamos a média e a variância estimadas das diferentes estatísticas de teste com a média e a variância da distribuição χ^2_1. Note que para $n = 35$, a média e a variância da estatística LR_p está mais próxima da média e variância da distribuição χ^2_1 quando comparado com a média e a variância das estatísticas LR_{BN1} e LR_{BN2}. Porém, a partir de $n = 45$ a média e variância das estatísticas LR_{BN1} e LR_{BN2} passam a ser mais próximas das medidas da distribuição de referência χ^2_1 quando comparado com a estatística LR_p. Por exemplo, para $n = 35$ e $k = 0,25$, a média e a variância da estatística LR_p é de 1,024 e 2,199, enquanto as mesmas medidas das estatísticas LR_{BN1} e LR_{BN2} são de (0,861 e 1,587) e (0,811 e 1,683), respectivamente. Para $n = 45$ e $k = 0,25$ a média e a variância da estatística LR_p é de 1,055 e 2,324, enquanto a média e a variância das estatísticas LR_{BN1} e LR_{BN2} são de (0,921 e 2,020) e (0,981 e 2,023), respectivamente.

Na Tabela 28 comparamos os desempenhos dos testes da razão de verossimilhanças baseados nas estatísticas LR_p, LR_{BN1} e LR_{BN2}. Foram considerados os mesmos tamanhos amostrais da estimação pontual e para cada tamanho de amostra foram calculadas as taxas de rejeição dos testes considerando a hipótese $H_0 : k = k_0$ contra a hipótese alternativa $H_1 : k \neq k_0$ com níveis nominais de 1%, 5% e 10%. Os desempenhos são avaliados em função da proximidade da probabilidade de rejeição da hipótese nula, sendo esta verdadeira, aos seus respectivos níveis nominais. Aumentamos os parâmetros de incômodo para avaliar o impacto nas taxas de rejeição dos testes estudados. Observamos que quando aumentamos a quantidade de parâmetros de incômodo, de modo geral, os testes baseados nas estatísticas LR_{BN1} e LR_{BN2} tornam-se mais conservativos, isto é, apresentando taxas de rejeição abaixo dos níveis nominais. Enquanto que os teste LR_p apresenta um comportamento oposto, isto é, mais liberal, apresentado taxas de rejeição acima dos níveis nominais. Mas, à medida que o tamanho amostral aumenta, as taxas de rejeição de todos os testes ficam mais próximas dos níveis nominais considerados. Assim, como esperado, o teste LR_p passa a apresentar desempenho similar aos testes LR_{BN1} e LR_{BN2} em amostras maiores. Por exemplo, para $n = 100$ e $k = 0,5$ e nível nominal de 5%, os testes LR_p, LR_{BN1} e LR_{BN2} apresentaram taxas de rejeição iguais a $5,28\%$, $4,94\%$ e $4,93$.

Avaliamos a aproximação assintótica das distribuições nula das estatísticas de teste pela distribuição χ^2_1. A Figura 9 apresenta o gráfico das discrepâncias relativas entre os quantis amostrais e os quantis assintóticos das estatísticas dos testes contra os quantis assintóticos, para $n = 55$ e $k = 0,25$. Mais precisamente, denotando por ST a estatística do teste (LR_p, LR_{BN1}, LR_{BN2}, conforme o caso), o quantil amostral de ordem $(1 - \alpha)$ do conjunto de valores simulados da estatística do teste sendo denotado por $ST(1 - \alpha)$ e o correspondente quantil da distribuição χ^2_1 sendo denotado por $\chi^2_1(1 - \alpha)$. A discrepância relativa é definida como

$$\frac{ST(1-\alpha) - \chi^2_1(1-\alpha)}{\chi^2_1(1-\alpha)}.$$

Quanto mais próxima à curva estiver da ordenada zero, melhor é a aproximação da distribuição nula da estatística de teste pela distribuição assintótica χ^2_1. Observe que as curvas relacionadas às estatísticas LR_{BN1} e LR_{BN2} são similares e estão abaixo da ordenada zero indicando que os testes baseados nas estatísticas LR_{BN1} e LR_{BN2} são conservativos. Por outro lado, a curva da estatística LR_p encontra-se acima da ordenada zero, indicando que o teste baseado na estatística LR_p é liberal.

A Tabela 29 apresenta resultados de simulação obtidos sob a hipótese alternativa para $n = 55$, k variando de $0,30$ até $0,95$ e nível nominal de 10%. Uma vez que os testes de tamanho baseados nas estatísticas LR_p, LR_{BN1} e LR_{BN2} apresentam diferenças entre as taxas de rejeição para um nível nominal fixado, os testes foram obtidos usando valores críticos estimados para cada estatística, de maneira que, todos os testes tivessem o mesmo patamar. De modo geral, observamos que os desempenhos dos testes baseados nas três estatísticas consideradas foram bastante similares. Notamos que os poderes dos testes LR_{BN1} e LR_{BN2} não perdem desempenho com relação ao poder do teste LR_p. Além disso, os testes tornam-se mais poderosos à medida que ocorre um afastamento do valor verdadeiro do parâmetro especificado na hipótese nula, conforme esperado.

Tabela 26 – Medidas descritivas de diferentes estimadores de k para $\sigma = 0,5$, $\beta_0 = -1$ e $\beta_1 = 1$.

		$k = 0,25$				$k = 0,5$			
n	Estimador	VR	EQM	assimetria	curtose	VR	EQM	assimetria	curtose
35	$\widehat{k}_p$	0,228	0,072	9,152	186,373	0,361	0,564	8,224	111,763
	$\widehat{\widehat{k}}_{BN}$	0,260	0,038	2,881	19,016	0,238	0,167	3,073	24,705
	$\widehat{\widehat{k}}_{BN}$	0,252	0,043	5,135	75,462	0,223	0,163	3,253	27,671
45	$\widehat{k}_p$	0,144	0,033	5,903	107,951	0,268	0,294	6,944	94,911
	$\widehat{\widehat{k}}_{BN}$	0,178	0,030	6,791	154,285	0,210	0,151	3,640	33,044
	$\widehat{\widehat{k}}_{BN}$	0,172	0,032	8,476	209,010	0,203	0,146	3,231	21,989
55	$\widehat{k}_p$	0,098	0,020	3,592	42,123	0,187	0,168	6,628	93,856
	$\widehat{\widehat{k}}_{BN}$	0,122	0,018	3,858	54,949	0,163	0,111	4,269	52,155
	$\widehat{\widehat{k}}_{BN}$	0,118	0,018	2,060	11,475	0,160	0,114	4,552	55,946
65	$\widehat{k}_p$	0,077	0,015	2,253	16,973	0,154	0,126	8,660	211,073
	$\widehat{\widehat{k}}_{BN}$	0,097	0,014	1,863	11,564	0,137	0,082	2,829	18,461
	$\widehat{\widehat{k}}_{BN}$	0,093	0,014	1,990	14,023	0,137	0,085	2,913	19,309
100	$\widehat{k}_p$	0,037	0,008	1,417	8,905	0,081	0,047	3,829	43,406
	$\widehat{\widehat{k}}_{BN}$	0,047	0,007	1,107	5,580	0,074	0,038	2,078	14,251
	$\widehat{\widehat{k}}_{BN}$	0,044	0,007	1,108	5,609	0,074	0,038	1,836	9,961

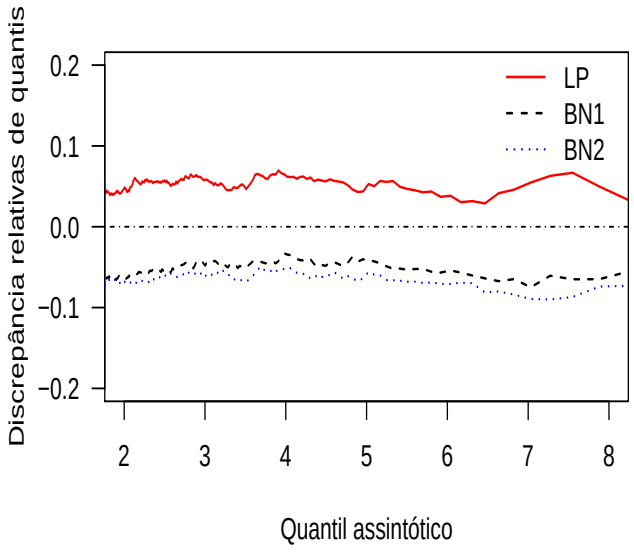
Fonte: Próprio autor (2021).

Tabela 27 – Médias e variâncias amostrais das estatísticas de teste e os valores assintóticos destas, $H_0 : k = k_0$, distribuição χ^2_1.

			$k = 0,25$			$k = 0,50$		
n		χ^2_2	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
35	média	1,00	1,024	0,861	0,811	0,984	0,839	0,821
	var	2,00	2,199	1,587	1,683	1,725	1,209	1,169
45	média	1,00	1,055	0,921	0,891	1,029	0,931	0,910
	var	2,00	2,324	2,020	2,023	2,040	1,628	1,596
55	média	1,00	1,033	0,932	0,920	1,001	0,935	0,926
	var	2,00	2,202	1,767	1,726	1,923	1,641	1,604
65	média	1,00	1,066	0,984	0,975	1,047	0,993	0,991
	var	2,00	2,322	1,900	1,863	2,090	1,845	1,818
100	média	1,00	1,038	0,993	0,990	1,033	0,991	0,991
	var	2,00	2,142	1,935	1,923	2,130	1,909	1,890

Fonte: Próprio autor (2021).

Figura 9 – Gráfico de discrepância relativa dos quantis da distribuição LBXII para $n = 55$, $\sigma = 0,5$, $\beta_0 = -1$, $\beta_1 = 1$ e $k = 0,25$.



Fonte: Próprio autor (2021).

Tabela 28 – Taxas de rejeição da hipótese nula dos diferentes testes da razão de verossimilhanças, $H_0 : k = k_0$, para $\sigma = 0,5$, $\beta_0 = -1$ e $\beta_1 = \beta_2 = \beta_3 = 1$.

$\mu_i = \beta_0 + \beta_1 x_1$.															
		$n = 35$			$n = 45$			$n = 55$			$n = 65$			$n = 100$	
k	$\alpha(\%)$	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
0,25	10	10,75	7,87	7,37	11,55	9,25	8,72	10,89	9,13	9,04	11,26	9,95	9,81	10,74	10,02	9,91
	5	5,64	3,57	3,31	5,90	4,31	4,05	5,75	4,63	4,44	5,88	4,93	4,77	5,15	4,99	4,93
	1	1,15	0,67	0,59	1,25	0,81	0,74	1,13	0,75	0,69	1,35	0,91	0,93	1,21	0,90	0,91
0,5	10	9,98	7,41	6,94	10,44	8,81	8,40	10,32	9,25	9,10	11,09	10,25	10,16	10,49	10,06	10,11
	5	4,44	2,74	2,58	5,17	3,93	3,64	5,01	4,12	3,99	5,54	4,92	4,77	5,28	4,94	4,93
	1	0,75	0,21	0,21	0,93	0,59	0,53	0,88	0,63	0,58	1,03	0,95	0,82	0,95	0,84	0,83
$\mu_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2$.																
k	$\alpha(\%)$	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
0,25	10	11,24	8,12	7,74	11,25	7,46	7,43	10,81	8,67	8,56	11,23	9,42	9,27	10,67	9,86	9,81
	5	5,83	3,84	3,61	5,88	3,46	3,45	5,71	4,11	4,10	5,91	4,65	4,54	5,34	4,82	4,71
	1	1,23	0,96	0,73	1,20	0,59	0,65	1,21	0,66	0,65	1,43	0,86	0,87	1,10	0,91	0,91
0,5	10	9,90	6,48	6,11	10,20	8,29	8,06	10,30	8,96	8,70	10,79	9,69	9,60	10,17	9,90	9,91
	5	4,40	2,27	2,13	5,01	3,49	3,33	4,84	3,94	3,74	5,28	4,56	4,42	5,30	4,92	4,89
	1	0,60	0,20	0,18	0,92	0,42	0,43	0,88	0,62	0,59	0,96	0,81	0,78	0,92	0,81	0,81
$\mu_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$.																
k	$\alpha(\%)$	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}	LR_p	LR_{BN1}	LR_{BN2}
0,25	10	11,70	8,67	7,73	11,99	6,26	7,45	12,10	7,50	8,52	11,06	8,48	9,28	10,37	9,25	9,80
	5	5,97	4,17	3,61	6,15	2,51	3,45	6,35	3,32	4,06	5,67	4,02	4,55	5,49	4,58	4,71
	1	1,19	0,71	0,73	1,36	0,45	0,65	1,48	0,54	0,63	1,23	0,81	0,87	1,26	0,90	0,91
0,5	10	9,53	5,45	6,13	9,83	6,63	8,06	10,50	8,38	8,69	10,69	9,22	9,60	10,02	9,55	9,90
	5	4,45	2,15	2,14	4,78	2,54	3,33	5,25	3,72	3,75	5,37	4,27	4,42	5,42	5,05	4,88
	1	0,58	0,20	0,18	0,66	0,24	0,42	1,02	0,66	0,59	0,91	0,60	0,78	1,14	1,01	0,80

Fonte: Próprio autor (2021).

Tabela 29 – Taxas de rejeição não nula dos diferentes testes da razão de verossimilhanças, $n = 55$ e nível nominal de 10%, indexados pelos parâmetros $\sigma = 0,5$, $\beta_0 = -1$ e $\beta_1 = 1$.

k	LR_p	LR_{BN1}	LR_{BN2}
0,30	12,96	12,32	12,29
0,35	19,92	18,96	19,37
0,40	30,21	29,55	29,83
0,45	40,07	40,21	40,92
0,50	50,00	50,28	50,85
0,55	59,49	59,92	60,29
0,60	66,73	67,24	67,85
0,65	72,31	73,00	73,40
0,70	77,14	77,89	78,01
0,75	81,25	81,82	81,88
0,80	83,24	83,58	83,93
0,95	86,57	87,05	86,72

Fonte: Próprio autor (2021).

5 APLICAÇÕES

Neste capítulo, ilustraremos através de dados reais, o uso das funções de verossimilhança perfilada e perfiladas modificadas na distribuição Burr XII e no modelo de regressão log-Burr XII. Para ilustrar o efeito do tamanho amostral, utilizamos dois conjuntos de dados reais para cada caso na distribuição Burr XII. Os exemplos considerados consistem de conjuntos de dados com tamanhos diferentes, um envolvendo uma quantidade pequena e outro, uma quantidade elevada de observações. Para as aplicações na distribuição Burr XII, consideramos dois casos a saber: caso (i) em que k é parâmetro de interesse e $\boldsymbol{\phi} = (s, c)^\top$ são parâmetros de incômodo e o caso (iii) em que o vetor de parâmetros de interesse é $\boldsymbol{\mu} = (c, k)^\top$ e s é parâmetro de incômodo, usando as funções de verossimilhança perfilada e perfiladas modificadas dadas nas seções 3.5 e 3.7, respectivamente. Para o modelo de regressão LBXII utilizamos um conjunto de dados reais para o caso em que k é parâmetro de interesse e $(\sigma, \boldsymbol{\beta}^\top)^\top$ são parâmetros de incômodo. Assumimos que os dados são independentes e seguem uma distribuição Burr XII.

As estimativas de máxima verossimilhanças perfiladas e perfiladas modificadas foram obtidas usando os comandos da linguagem de programação Ox: MaxBFGS e MaxSQP, respectivamente. O comando MaxBFGS maximiza funções usando o método quase-Newton desenvolvido por Bryden, Fletcher, Goldfarb e Shanno (BFGS), enquanto o MaxSQP maximiza funções com restrição nas variáveis usando o método SQP (*Sequential Quadratic Programming*). O algoritmo deste método no Ox é similar ao algoritmo 18.7 em Nocedal e Wright (1999).

5.1 EXEMPLO 1

Para este exemplo, utilizamos os dados analisados por Zaharim et al. (2009) referente a 36 observações sobre a velocidade mensal do vento em um distrito localizado em Pahang na Malásia, denominado de Cameron Highland nos anos de 2004 a 2006. Zaharim et al. (2009) mostrou, através de métodos de bondade de ajustes de Kolmogorov Smirnov (KS) e Anderson Darling (AD), que a distribuição Burr XII apresentou um melhor ajuste para os dados quando comparado com as distribuições log-normal e Frechet. Os dados são apresentados na tabela abaixo.

Tabela 30 – velocidade mensal do vento (m/s) em Cameron Highland nos anos de 2004 a 2006.

Anos	2004	2005	2006
Janeiro	2,76	3,00	2,31
Fevereiro	1,56	1,88	3,11
Março	2,00	2,25	1,92
Abril	1,38	1,55	1,73
Maio	0,91	1,66	1,61
junho	2,02	1,66	2,01
Julho	1,12	1,67	1,61
Agosto	1,34	1,66	1,98
Setembro	1,27	1,83	1,62
Outubro	1,95	1,69	1,59
Novembro	1,63	1,79	2,27
Dezembro	2,61	2,69	2,77

Fonte: ZAHARIM et al. (2009).

Algumas medidas descritivas, tais como, a média, a mediana, a variância, a assimetria, a curtose, o valor mínimo e máximo do conjunto de dados apresentadas na Tabela 30 estão disponíveis na Tabela 31. De início, notamos que as médias são superiores às medianas, indicando que, a cauda da distribuição é um pouco mais longa à direita, que também pode ser observado pelo coeficiente de assimetria positivo. As curtoses das distribuições dos dados para os anos de 2004 e 2006 são menores que a curtose da distribuição normal, indicando que as distribuições passam a ter uma forma platicúrticas. Para o ano de 2005 a curtose é maior que a curtose da distribuição normal, fazendo com que a distribuição tenha uma forma leptocúrtica.

Tabela 31 – Medidas descritivas.

Anos	Média	Mediana	Variância	Assimetria	Curtose	Mínimo	Máximo
2004	1,7125	1,5950	0,3005	0,5232	2,2880	0,91	2,76
2005	1,9442	1,7400	0,1952	1,3987	3,5310	1,55	3,00
2006	2,0442	1,9500	0,2222	0,9993	2,9212	1,59	3,11

Fonte: Próprio autor (2021).

As estimativas de máxima verossimilhança obtidas através da maximização das funções de verossimilhança perfilada e perfiladas modificadas com seus respectivos erros-padrão (em parênteses) para $\boldsymbol{\mu} = (c, k)^\top$ são: $\widehat{\boldsymbol{\mu}}_p = (7,588(2,2411); 0,756(0,4444))^\top$, $\widehat{\tilde{\boldsymbol{\mu}}}_{BN} =$

$(7,550(2,2281); 0,752(0,4279))^\top$ e $\widehat{\widetilde{\boldsymbol{\mu}}}_{BN} = (7,354(2,1920); 0,782(0,4716))^\top$. As estimativas pontuais são bastante similares. Note também, que a estimativa baseada na verossimilhança perfilada apresenta valores e erro-padrão para o parâmetro c ligeiramente superior do que as estimativas baseadas nas verossimilhanças perfiladas modificadas. Com relação ao parâmetro k, a estimativa baseada na verossimilhança perfilada modificada $\widehat{\widetilde{\boldsymbol{\mu}}}_{BN}$ apresenta valores mais elevados quando comparado com as estimativas $\widehat{\boldsymbol{\mu}}_p$ e $\widehat{\widetilde{\boldsymbol{\mu}}}_{BN}$, respectivamente.

Suponha, que estamos interessados em testar $H_0 : (c,k)^\top = (4,0;2,5)^\top$ contra $H_1 : (c,k)^\top \neq (4,0;2,5)^\top$. Para este teste, obtemos as estatísticas da razão de verossimilhanças perfiladas e perfiladas modificadas $LR_p = 4,9174$, $LR_{BN1} = 4,5411$ e $LR_{BN2} = 4,3951$, que conduzem aos respectivos níveis nominais $0,0855$, $0,1032$ e $0,1110$. Observamos que, a hipótese nula não é rejeitada ao nível nominal de 10% quando o teste é baseado nas estatísticas LR_{BN1} e LR_{BN2}, ao contrário do que ocorre com o teste baseado na estatística LR_p. Desse modo, os testes apresentam conclusões divergentes.

5.2 EXEMPLO 2

O segundo conjunto de dados representa os tempos de remissão não censurados (em meses) de uma amostra aleatória de 128 pacientes com câncer de bexiga analisados em Lee e Wang (2003), ver tabela abaixo.

Tabela 32 – Tempos de remissão (em meses) de pacientes com câncer de bexiga.

0,08	2,09	3,48	4,87	6,94	8,66	13,11	23,63	0,20	2,23
3,52	4,98	6,97	9,02	13,29	0,40	2,26	3,57	5,06	7,09
9,22	13,80	25,74	0,50	2,46	3,64	5,09	7,26	9,47	14,24
25,82	0,51	2,54	3,70	5,17	7,28	9,74	14,76	26,31	0,81
2,62	3,82	5,32	7,32	10,06	14,77	32,15	2,64	3,88	5,32
7,39	10,34	14,83	34,26	0,90	2,69	4,18	5,34	7,59	10,66
15,96	36,66	1,05	2,69	4,23	5,41	7,62	10,75	16,62	43,01
1,19	2,75	4,26	5,41	7,63	17,12	46,12	1,26	2,83	4,33
5,49	7,66	11,25	17,14	79,05	1,35	2,87	5,62	7,87	11,64
17,36	1,40	3,02	4,34	5,71	7,93	11,79	18,10	1,46	4,40
5,85	8,26	11,98	19,13	1,76	3,25	4,50	6,25	8,37	12,02
2,02	3,31	4,51	6,54	8,53	12,03	20,28	2,02	3,36	6,76
12,07	21,73	2,07	3,36	6,93	8,65	12,63	22,69		

Fonte: LEE; WANG (2003).

Estes dados também foram utilizados por Cordeiro et al. (2015) como exemplo prático para a família de distribuições Beta Odd log-logística generalizada. Para este conjunto de dados, Cordeiro et al. (2015) obtém os valores de AIC (*Akaike information criterion*) para cinco distribuições diferentes. Calculamos o valor de AIC da distribuição Burr XII, otendo, 825,479. Desta maneira, a distribuição Burr XII apresentou o menor valor de AIC quando comparado com os valores de AIC apresentados por Cordeiro et al. (2015), indicando, segundo esse critério, que a distribuição Burr XII tem um melhor ajuste para esse conjunto de dados.

Tabela 33 – Medidas descritivas.

Média	Mediana	Variância	Assimetria	Curtose	Mínimo	Máximo
9,3656	6,3950	109,56	3,2866	18,483	0,08	79,05

Fonte: Próprio autor (2021).

Na Tabela 33, notamos que o tempo médio de remissão em pacientes com câncer de bexiga é de 9,3656 meses. Novamente, a média é maior que a mediana. Neste caso, a distribuição é assimétrica à direita, confirmado pelo coeficiente de assimetria positivo. A curtose da distribuição dos dados é bem superior a curtose da distribuição normal, indicando que a distribuição tem forma leptocúrtica. Observe também que o menor tempo de remissão é de 0,08 meses e o maior tempo de remissão é de 79,05 meses.

As estimativas de máxima verossimilhança obtidas através da maximização das funções de verossimilhança perfilada e perfiladas modificadas com seus respectivos erros-padrão (em parênteses) para $\boldsymbol{\mu} = (c, k)^\top$ são: $\boldsymbol{\mu}_p = (1,427(0,1779); 2,070(0,9682))^\top$, $\widehat{\boldsymbol{\mu}}_{BN} = (1,424(0,1809); 2,052(0,9682))^\top$ e $\widehat{\widehat{\boldsymbol{\mu}}}_{BN} = (1,422(0,1803); 2,062(0,9758))^\top$. A estimativa baseada na verossimilhança perfilada é ligeiramente superior às estimativas baseadas nas verossimilhanças perfiladas modificadas.

Suponha, que estamos interessados em testar a hipótese nula $H_0 : (c, k)^\top = (1,5; 3,5)^\top$ contra a hipótese alternativa $H_1 : (c, k)^\top \neq (1,5; 3,5)^\top$. Os testes baseados nas estatísticas $LR_p = 5,7737$, $LR_{BN1} = 6,0582$ e $LR_{BN2} = 6,0360$ conduzem aos respectivos níveis descritivos $0,0557$, $0,0483$ e $0,0489$, respectivamente. Neste caso, o tamanho da amostra é grande, portanto, como esperado, os testes baseados nas estatísticas LR_{BN1} e LR_{BN2} apresentam desempenho similar ao teste LR_p. Ao nível nominal de 5%, os testes LR_{BN1} e LR_{BN2} rejeitam a hipótese nula, enquanto o teste LR_p não rejeita, dessa forma, os testes apresentam conclusões divergentes.

5.3 EXEMPLO 3

O terceiro conjunto de dados é constituído de 24 observações referentes aos intervalos de tempo (em dias) de sucessivos terremotos que ocorreram na zona de falha do Norte da Anatólia, localizada no leste da Turquia. Recentemente, foi utilizado por Kuş (2007) como exemplo prático para a distribuição Exponencial–Poisson (EP). Os dados são apresentados na tabela abaixo.

Tabela 34 – Intervalos de tempo (em dias) dos sucessivos terremotos.

1163	3258	323	159	756	409	501	616	398	67
896	8592	2039	217	9	633	461	1821	4863	143
182	2117	3709	979						

Fonte: KUŞ (2007).

Calculamos os valores de AIC correspondentes as distribuições Burr XII e Exponencial-Poisson, obtendo, respectivamente, 490,979 e 569,898. Portando, para esse conjunto de dados, a distribuição Burr XII possui o menor AIC, indicando, segundo esse critério, um melhor ajuste aos dados quando comparado com a distribuição Exponencial-Poisson.

Tabela 35 – Medidas descritivas.

Média	Mediana	Variância	Assimetria	Curtose	Mínimo	Máximo
1429,6	624,5	3923279	2,3223	8,3507	9	8592

Fonte: Próprio autor (2021).

Avaliando a Tabela 35, observamos que o tempo médio dos intervalos dos terremotos é de 1429,6 dias. A média é superior à mediana e com assimetria positiva, indicando que a distribuição é assimétrica à direita. A curtose da distribuição dos dados é maior que a curtose da distribuição normal. Neste caso, a distribuição tem forma leptocúrtica. Note também que, o menor intervalo de terremotos é de 9 dias e o maior intervalo é de 8592 dias.

Inicialmente, estamos interessados em fazer inferências para o parâmetro k, isto é, $\mu = k$ é parâmetro de interesse e $\boldsymbol{\phi} = (c, s)^\top$ são parâmetros de incômodo. As estimativas de máxima verossimilhança perfilada e perfiladas modificadas com seus respectivos erros-padrão em parenteses para o parâmetro k são: $\widehat{k}_p = 2,782(0,9044)$, $\widehat{\widehat{k}}_{BN} = 0,761(0,4499)$ e $\widehat{\widehat{k}}_{BN} =$

$0,733(0,5133)$. Aqui, as estimativas com base nas verossimilhanças perfiladas modificadas são menores e similares quando comparadas com a estimativa baseada na verossimilhança perfilada.

Suponha que estamos interessados em testar $H_0 : k = 0,36$ contra $H_1 : k \neq 0,36$. Para este teste, temos $LR_p = 4,9178$, $LR_{BN1} = 3,4089$ e $LR_{BN2} = 3,3260$, que conduzem aos níveis descritivos $0,0265$, $0,0648$ e $0,0681$, respectivamente. Assim, os testes baseados nas estatísticas LR_{BN1} e LR_{BN2} não rejeitam a hipótese nula ao nível nominal de 5%, ao contrário do que ocorre com o teste LR_p. Note que, os testes apresentam conclusões diferentes.

5.4 EXEMPLO 4

O quarto conjunto de dados reais é sobre a coleta de informações de 202 atletas do Australian Institute of Sport. A variável avaliada neste estudo é a concentração de ferritina plasmática (ver Tabela 37). O conjunto de dados foi analisado por Weisberg (2014) e recentemente foi estudado por Alizadeh et al. (2015) como exemplo prático para a família de distribuições beta Marshall-Olkin. Para este conjunto de dados Alizadeh et al. (2015) obteve os valores de AIC de nove distribuições. Aqui, calculamos o valor de AIC da distribuição Burr XII, obtendo, 2065,41. Dessa maneira, a distribuição Burr XII apresentou o menor valor de AIC quando comparado com os valores de AIC apresentado por Alizadeh et al. (2015). Assim, concluímos, que a distribuição Burr XII apresenta um melhor ajuste para este conjunto de dados.

Tabela 36 – Medidas descritivas.

Média	Mediana	Variância	Assimetria	Curtose	Mínimo	Máximo
76,876	65,5	2245,2	1,2806	4,4202	8	234

Fonte: Próprio autor (2021).

A Tabela 36, apresenta as medidas descritivas dos dados referentes a concentração de ferritina plasmáticas dada na Tabela 37. Observamos que a concentração média de ferritina plasmática é de 76,876 (ng/dl). A distribuição é assimétrica à direita, pois, apresenta assimetria positiva. A curtose da distribuição tem forma leptocúrtica, uma vez que, apresenta curtose maior que a curtose da distribuição normal. Note que, a menor concentração de ferritina é de 8 (ng/dl) e a maior concentração de ferritina é de 234 (ng/dl).

Novamente, o interesse reside em obter inferências para o caso em que o parâmetro $\mu = k$ é de interesse e $\boldsymbol{\phi} = (c, s)^\top$ são parâmetros de incômodo. As estimativas de máxima

verossimilhança perfilada e perfiladas modificadas seguidos pelos respectivos erros-padrão em parenteses são: $\widehat{k}_p = 1,527(0,5863)$, $\widehat{\widehat{k}}_{BN} = 1,525(0,6325)$ e $\widehat{\widehat{k}}_{BN} = 1,513(0,6155)$. As estimativas estudadas são similares, como esperado, visto que, o tamanho da amostra é grande. Ainda assim, as estimativas com base nas verossimilhanças perfiladas modificadas são ligeiramente menores que a estimativa baseada na verossimilhança perfilada.

Suponha que estamos interessados em testar $H_0 : k = 0,8$ contra $H_1 : k \neq 0,8$. As estatísticas da razão de verossimilhanças perfiladas e perfiladas modificadas são, respectivamente, dadas por $LR_p = 3,8548$, $LR_{BN1} = 3,8127$ e $LR_{BN2} = 3,7688$, que conduzem aos correspondentes níveis descritivos $0,0496$, $0,0508$ e $0,0522$, respectivamente. Observe que, os testes LR_{BN1} e LR_{BN2} não rejeitam a hipótese nula ao nível nominal de 5%, diferente do que ocorre com o teste LR_p. Sendo assim, os testes apresentam conclusões diferentes.

Tabela 37 – Concentração de ferritina plasmática (ng/dl) em 202 atletas.

60	68	21	69	29	42	73	44	41	44	38	26	30	48	30
29	43	34	53	59	43	40	92	48	77	71	37	71	73	85
64	19	39	41	13	20	59	22	30	78	21	109	102	71	64
68	78	107	39	58	127	102	86	40	50	58	33	51	82	25
86	22	30	27	60	115	124	54	29	164	50	36	40	62	90
12	36	45	43	51	22	44	26	16	58	46	43	34	41	57
73	88	182	80	88	109	69	41	66	63	34	97	55	76	93
46	155	99	35	124	176	107	177	130	64	109	125	150	115	89
93	183	84	70	118	61	72	91	58	97	110	72	36	53	72
39	61	49	35	8	106	32	41	20	44	103	50	41	101	73
56	74	58	87	139	82	87	80	67	132	43	212	94	213	122
76	53	91	36	101	184	44	66	220	191	40	189	141	212	97
53	126	234	50	94	156	124	87	97	102	55	117	52	133	214
143	65	90	38	122	233	32								

Fonte: WEISBERG (2014).

5.5 EXEMPLO 5

O quinto conjunto de dados reais é sobre o tempo de durabilidade de pastilhas de freio esquerdo dianteiro de 98 automóveis ver Tabela 39. O conjunto de dados foi estudado por Kalbfleisch e Lawless (1992), onde usou como exemplo prático para alguns métodos não paramétricos. Os dados foram obtidos usando a leitura atual do odômetro (em km) e a medida

da espessura restante das pastilhas acima do nível minímo determinado pelo fabricante com a espessura inicial da pastilha para estimar o tempo de durabilidade (em km).

Tabela 38 – Medidas descritivas.

Dados	Média	Mediana	Variância	Assimetria	Curtose	Mínimo	Máximo
tempo durabilidade (t)	68,263	65,300	682,14	0,9199	4,2470	20,8	165,5
leitura odômetro (v)	22,983	20,850	90,351	1,4577	5,3943	7	57,300

Fonte: Próprio autor (2021).

A Tabela 38, apresenta as medidas descritivas dos dados referentes ao tempo de durabilidade de pastilhas de freio e a leitura atual do odômetro apresentado na Tabela 39. Observamos que as pastilhas de freio duram em média 68,263 km e a quilometragem média apresentada nos odômetros dos carros é de 22,983. A distribuição é assimétrica à direita de ambas variáveis, pois, apresentam assimetria positiva. A curtose das distribuições têm forma leptocúrtica, uma vez que, apresentam curtose maior que a curtose da distribuição normal. Note que, a menor durabilidade da pastilha foi de 20,8 km e durabilidade máxima foi de 165,5 km.

Para este exemplo, é de interesse obter inferências considerando o caso em que $\mu = k$ é parâmetro de interesse e $\boldsymbol{\phi} = (\sigma, \boldsymbol{\beta}^\top)^\top$ são parâmetros de incômodo. As estimativas de máxima verossimilhança perfilada e perfiladas modificadas seguidos pelos respectivos erros-padrão em parenteses são: $\widehat{k}_p = 1,605\ (0,8175)$, $\widehat{\widehat{k}}_{BN} = 1,581\ (0,8482)$ e $\widehat{\widehat{k}}_{BN} = 1,596\ (0,8765)$. As estimativas estudadas são similares, como esperado, visto que, o tamanho da amostra é grande. Ainda assim, as estimativas com base nas verossimilhanças perfiladas modificadas são ligeiramente menores que a estimativa baseada na verossimilhança perfilada.

Suponha que estamos interessados em testar $H_0 : k = 0,7$ contra $H_1 : k \neq 0,7$. As estatísticas da razão de verossimilhanças perfiladas e perfiladas modificadas são, respectivamente, dadas por $LR_p = 3,9717$, $LR_{BN1} = 3,7637$ e $LR_{BN2} = 3,8312$, que conduzem aos correspondentes níveis descritivos $0,0473$, $0,0523$ e $0,0523$, respectivamente. Observe que, os testes LR_{BN1} e LR_{BN2} não rejeitam a hipótese nula ao nível nominal de 5%, diferente do que ocorre com o teste LR_p. Sendo assim, os testes apresentam conclusões divergentes.

Tabela 39 – Tempo de durabilidade (t) e leituras de odômetro (v) de 98 carros (por 1000 km).

v	t	v	t	v	t	v	t
22,2	38,7	16,5	69,6	18,4	86,7	10,9	79,5
23,0	49,2	15,7	74,8	18,2	43,8	25,5	55,0
24,0	42,4	28,0	32,9	15,9	100,6	12,4	46,8
28,6	73,8	13,3	51,5	16,4	67,6	39,9	124,5
21,8	46,7	16,5	31,8	23,6	89,5	17,7	92,5
17,0	44,1	24,2	77,6	19,2	60,3	26,3	110,0
26,0	61,9	17,6	63,7	23,3	103,6	14,1	101,2
23,2	39,3	27,8	83,0	20,4	82,6	21,0	59,4
18,9	49,8	18,3	24,8	20,9	88,0	11,2	27,8
21,9	46,3	17,7	68,8	28,5	42,4	10,8	33,6
27,3	56,2	20,0	68,8	23,2	68,9	25,7	69,0
13,8	50,5	13,2	89,1	17,9	95,7	32,4	75,2
24,0	54,9	16,9	65,0	46,1	78,1	13,6	58,4
20,1	54,0	14,9	65,1	39,3	83,6	19,1	105,6
15,7	49,2	15,5	59,3	11,8	18,6	16,1	56,2
26,8	44,8	7,0	53,9	17,7	92,6	53,3	55,9
27,9	72,2	15,8	79,4	30,9	42,4	57,3	83,8
15,3	107,8	15,0	47,4	22,4	34,3	36,5	123,5
28,8	81,6	38,3	61,4	45,0	105,6	19,7	69,0
16,0	45,2	11,2	72,8	18,2	20,8	20,8	101,9
23,6	124,6	38,2	54,0	30,2	52,0	30,8	87,6
53,8	64,0	26,7	37,2	21,8	77,2	20,0	38,8
21,7	83,0	17,1	44,2	18,2	68,9	39,6	74,7
28,8	143,6	29,0	50,8	23,0	78,7		
17,0	43,4	18,3	65,5	27,2	165,5		

Fonte: KALBFLEISCH; LAWLESS (1992).

6 CONCLUSÃO

Neste trabalho, apresentamos algumas características e propriedades da distribuição Burr XII e log-Burr XII, assim como, alguns casos particulares da distribuição Burr XII. Derivamos para estas distribuições as funções de verossimilhança perfilada e perfiladas modificadas baseadas na proposta de Barndorff-Nielsen (1983) através das aproximações introduzidas por Fraser e Reid (1995), Fraser, Reid e Wu (1999) e Severini (1998). Essas funções foram derivadas para três casos da distribuição Burr XII, a saber: caso (i): k como parâmetro de interesse e $(s, c)^\top$ como parâmetros de incômodo, caso (ii): $(k, s)^\top$ como parâmetros de interesse e c como parâmetro de incômodo e caso (iii): $(k, c)^\top$ como parâmetros de interesse e s como parâmetro de incômodo. Para o modelo de regressão LBXII derivamos o caso em que k é parâmetro de interesse e $(\sigma, \boldsymbol{\beta}^\top)^\top$ são parâmetros de incômodo.

Obtivemos também os estimadores de máxima verossimilhança perfilado e perfilados modificados a partir da maximização das versões modificadas da função de verossimilhança. Além disso, mostramos como ficam os testes da razão de verossimilhanças baseados nas funções de verossimilhanças perfiladas e perfiladas modificadas. Fizemos diversos estudos de simulação com o objetivo de avaliar os desempenhos dos estimadores pontuais e testes propostos, em amostras finitas. Apresentaremos, a seguir, as principais conclusões desta dissertação.

Caso (i): os estimadores perfilados modificados $\widehat{\overset{\circ}{k}}_{BN}$ e $\widehat{\bar{k}}_{BN}$ se mostraram eficientes em corrigir os vieses e reduzirem os erros quadráticos médios quando comparado com o estimador $\widehat{k}_p$, principalmente quando $k = 0,5$. Com relação aos tamanhos dos testes LR_p, LR_{BN1} e LR_{BN2}, podemos concluir, que os testes baseados nas versões modificadas LR_{BN1} e LR_{BN2} apresentaram uma melhor redução nas distorções de tamanho, produzindo taxas mais próximas dos níveis nominais quando comparados com o teste LR_p. Os poderes dos testes perfilados corrigidos são bastante similares ao poder do teste perfilado, indicando que os ajustes considerados não tiveram perda de poder.

Caso (ii): os estimadores $\widehat{\overset{\circ}{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$ apresentaram desempenhos superiores quando comparado com o estimador $\widehat{\boldsymbol{\mu}}_p$, embora, para valores elevados de k, observamos uma deterioração no desempenho dos estimadores, principalmente no estimador baseado na verossimilhança perfilada. Observamos que, em geral, os testes LR_{BN1} e LR_{BN2} apresentaram desempenhos similares e superiores ao teste LR_p. Notamos também que o teste LR_p foi bastante liberal, enquanto os testes LR_{BN1} e LR_{BN2} foram mais conservativos apresentando taxas mais próximas dos níveis

nominais. Novamente, os poderes dos testes estudados tiveram desempenhos semelhantes, indicando que, as modificações realizadas na verossimilhança perfilada não tiveram perda de poder.

Caso (iii): os estimadores perfilados modificados $\widehat{\breve{\boldsymbol{\mu}}}_{BN}$, $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$ mostraram desempenhos muito superiores que o estimador perfilado $\widehat{\boldsymbol{\mu}}_p$. Aqui, conforme o valor de k aumenta, os desempenhos dos estimadores melhoram. Os desempenhos dos testes $\widehat{\breve{\boldsymbol{\mu}}}_{BN}$, $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$ são superiores ao teste $\widehat{\boldsymbol{\mu}}_p$ para valores menores de k. Para valores maiores de k, o teste $\widehat{\boldsymbol{\mu}}_p$ passa a ter desempenho ligeiramente superior aos testes perfilados modificados $\widehat{\breve{\boldsymbol{\mu}}}_{BN}$ e $\widehat{\bar{\boldsymbol{\mu}}}_{BN}$ apresentando taxas mais próximas aos níveis nominais. Notamos que os testes baseados nas verossimilhanças perfiladas modificadas foram mais poderosos que o teste baseado na verossimilhança perfilada.

Para o modelo de regressão LBXII, os estimadores perfilados modificados $\widehat{\breve{k}}_{BN}$ e $\widehat{\bar{k}}_{BN}$ apresentaram desempenhos superiores como, viés relativo, erro quadrático médio, assimetria e curtose menores quando comparado com $\widehat{k}_p$. Com relação aos tamanhos dos testes LR_p, LR_{BN1} e LR_{BN2}, podemos concluir, de modo geral, que os testes baseados nas versões modificadas LR_{BN1} e LR_{BN2} são mais conservativos apresentando taxas mais próximas dos níveis nominais. Já o teste LR_p é mais liberal, apresentando taxas de rejeição nula mais elevadas aos níveis nominais. Os poderes dos testes perfilados corrigidos são bastante similares ao poder do teste perfilado, indicando que os ajustes considerados não tiveram perda de poder.

Concluímos que os resultados obtidos das simulações sugerem que as inferências realizadas utilizando as funções de verossimilhanças perfiladas modificadas são tipicamente superiores às obtidas através da função de verossimilhança perfilada usual e que ambos os ajustes melhoram as inferências de maneira semelhante. Além disso, as distribuições nula dessas estatísticas de teste, são melhor aproximadas pela distribuição χ^2_p. Os poderes dos testes perfilados corrigidos apresentam-se bastante semelhantes aos usuais, o que permite concluir que os ajustes considerados não acarretaram perda de poder para este teste. Vale salientar, que os ajustes desenvolvidos na função de verossimilhança perfilada apresentam algumas vantagens, como, por exemplo, são invariantes sob reparametrização e também não requerem a ortogonalidade entre parâmetros.

Por fim, recomendamos o uso da aproximação proposta por Severini (1998), por sua simplicidade de aplicação, que consiste basicamente no produto das respectivas funções escore correspondentes aos parâmetros de incômodo da verossimilhança perfilada e da verossimilhança genuína e não envolve derivadas que, em casos específicos, podem ser bastante trabalhosos.

REFERÊNCIAS

ALIZADEH, M.; CORDEIRO, G. M.; BRITO, E.; DEMÉTRIO, C. G. B. The beta Marshall-Olkin family of distributions. *Journal of Statistical Distributions and Applications*, 2015.

ARAÚJO, M. C.; CYSNEIROS, A. H. M. A.; MONTENEGRO, L. C. Improved heteroskedasticity likelihood ratio tests in symmetric nonlinear regression models. *Statistical Papers*, v. 61, n. 1, p. 167–188, February 2020. Disponível em: <https://ideas.repec.org/a/spr/stpapr/v61y2020i1d10.1007_s00362-017-0933-5.html>.

BARNDORFF-NIELSEN, O. E. On a formula to the distribution of the maximum likelihood estimator. *Biometrika*, 70, p. 343–365, 1983.

BARNDORFF-NIELSEN, O. E. Adjusted versions of proflle likelihood and directed likelihood, and extended likelihood. *Journal of the Royal Statistical Society*, B, n. 56, p. 125–140, 1994.

BARRETO, L.; CYSNEIROS, A.; CRIBARI-NETO, F. Improved Birnbaum–Saunders inference under type II censoring. *Computational Statistics & Data Analysis*, v. 57, p. 68–81, 01 2013.

BURR, I. W. Cumulative frequency functions. 215-232, Annals of Mathematical Statistics 13, 1942.

CORDEIRO, G.; ALIZADEH, M.; TAHIR, M.; MANSOOR, M.; BOURGUIGNON, M.; HAMEDANI, G. The beta odd log-logistic generalized family of distributions. *Hacettepe University Bulletin of Natural Sciences and Engineering Series B: Mathematics and Statistics*, v. 45, 06 2015.

COX, D.; REID, N. Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society*, B, 49, p. 1–39, 1987.

COX, D.; REID, N. A note on the calculation of adjusted profile likelihood. *Journal of the Royal Statistical Society*, B, 55, p. 467–471, 1993.

CYSNEIROS, A. H. M. A.; CRIBARI-NETO, F.; JúNIOR, C. A. G. A. On BirnbaumSaunders Inference. *Computational Statistics and Data Analysis*, 80, p. 1–21, 2008.

CYSNEIROS, A. H. M. A.; FERRARI, S. L. P. An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics and Probability Letters*, n. 76, p. 355–265, 2006.

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

EVANS, I. G.; RAGAB, A. S. Bayesian inference given a type-2 censored sample from Burr distribution. *Comm. Statist.-Theory Meth*, v. 12, p. 1569–1580, 1983.

FERGUSON, H.; REID, N. P.; COX, D. Estimating equations from modified profile likelihood. In: . [S.l.: s.n.], 1991.

FERRARI, S. L. P.; CRIBARI-NETO, F. Corrected modified profile likelihood inference for heteroskedastic models. *Statistics and Probability Letters*, n. 47, p. 353–361, 2002.

FERRARI, S. L. P.; SILVA, M. F. D.; CRIBARI-NETO, F. Adjusted profile likelihood for two-parameter exponential family models. *Communications in Statistics - Theory and Methods*, Taylor & Francis, v. 34, n. 2, p. 257–276, 2005. Disponível em: <https://doi.org/10.1080/03610920509342419>.

FERRARI, S. L. P.; SILVA, M. F. D.; CRIBARI–NETO, F. Adjusted profile likelihoods for the weibull shape parameter. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 77, n. 7, p. 531–548, 2007. Disponível em: <https://doi.org/10.1080/10629360600565160>.

FRASER, D. A. S.; REID, N. Ancillaries and third-order significance. *Utilitas Mathematica*, 47, p. 33–53, 1995.

FRASER, D. A. S.; REID, N.; WU, J. A simple formula for tail probabilities for frequentist and bayesian inference. *Biometrika*, 86, p. 655–661, 1999.

GENTLE, J. E. *Random Number Generation and Monte Carlo Methods*. 2nd. ed. Springer, 2004. ISBN 0387001786. Disponível em: <http://www.scribd.com/doc/6833846/JEGentle-Random-number-generation-and-Monte-Carlo-methods-2nd-edition>.

JALALI, A.; WATKINS, A. J. On Maximum Likelihood Estimation for the Two Parameter Burr XI Distribution. *Communications in Statistics: Theory and Methods*, v. 38, n. 11, p. 1916–1926, 2009.

KALBFLEISCH, J. D.; LAWLESS, J. F. Some useful statistical methods for truncated data. *Journal of Quality Technology*, Taylor & Francis, v. 24, n. 3, p. 145–152, 1992. Disponível em: <https://doi.org/10.1080/00224065.1992.11979385>.

KHAN, A. H.; KHAN, A. I. Moments of order statistics from Burr's distribution and its characterization. *Communications in Statistics: Theory and Methods*, v. 45, n. 45, p. 21–29, 1987.

KUŞ, C. A new lifetime distribution. *Computational Statistics & Data Analysis*, v. 51, n. 9, p. 4497 – 4509, 2007. ISSN 0167-9473. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0167947306002313>.

LAWLESS, J. F. *Statistical Models and Methods for Lifetime Data*. [S.l.]: New York: Wiley, 2003.

LEE, E.; WANG, J. *Statistical Methods for Survival Data Analysis*. [S.l.]: Third edition Wiley, New York, 2003.

LEWIS, A. *The Burr distribution as a general parametric family in survivorship and reliability theory applications*. Tese (Ph.D. Thesis) — Department of Biostatistics, University of North Carolina, Chapel Hill, 1981.

MCCULLAGH, P.; TIBISHIRANI, R. A simple method for the adjustment of profile likelihood. *Jounal of the Royal Statistical Society*, B, n. 52, p. 325–344, 1990.

MELO, T. F.; FERRARI, S. L.; CRIBARI-NETO, F. Improved testing inference in mixed linear models. *Computational Statistics & Data Analysis*, Elsevier BV, v. 53, n. 7, p. 2573–2582, May 2009. ISSN 0167-9473. Disponível em: <http://dx.doi.org/10.1016/j.csda.2008.12.007>.

NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. New York: Springer, 1999.

NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. second. New York, NY, USA: Springer, 2006.

OKASHA, M. K.; MATTER, M. Y. On the three-parameter burr type xii distribution and its application to heavy tailed lifetime data. *JOURNAL OF ADVANCES IN MATHEMATICS*, v. 10, n. 4, p. 3429–3442, Apr. 2015. Disponível em: <https://rajpub.com/index.php/jam/article/view/1242>.

PACE, L.; SAVAN, A. *Principles of Statistical Inference from a NeoFisherian Perspective*. [S.l.]: London: World scientific, 1997.

PAPADOPOULOS, A. S. The Burr distribution as a failure model from a Bayesian approach. *IEEE Transactions on Reliability*, v. 27, n. 5, p. 369–371, 1978.

PIRES, J.; CYSNEIROS, A.; CRIBARI-NETO, F. Improved inference for the generalized pareto distribution. *Brazilian Journal of Probability and Statistics*, v. 32, p. 69–85, 02 2018.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020. Disponível em: <http://www.R-project.org>.

RODRIGUEZ, R. *A Guide to the Burr XII distributions*. [S.l.]: Biometrika, London, 1977. v. 64. 129-134 p.

SEN, P. K.; SINGER, J. M. *Large sample methods in statistics*. [S.l.]: An introduction with applications. New York: Chapman and Hall, 1993. 382 p.

SEVERINI, T. A. An approximation to the modifled profile likelihood function. *Biometrika*, 85, p. 403–411, 1998.

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

SEVERINI, T. A. *Likelihood Methods in Statistics*. [S.l.]: Oxford: Oxford University Press., 2000.

SHAO, Q.; WONG, H.; XIA, J. Models for extremes using the extended three parameter Burr XII system with application to flood frequency analysis. *Hydrological Sciences Journal des Sciences Hydrologiques*, London, v. 49, p. 685–702, 2004.

SILVA, G. O.; ORTEGA, E. M. M.; CANCHO, V. G.; BARRETO, M. L. Log-Burr XII regression models with censored data. *Computational Statistics and Data Analysis*, 52, p. 3820–3842, 2008.

SINGH, S. K.; MODDOLA, G. S. A function for size distribution of incomes. *Econometrica*, 44, p. 963–969, 1976.

SOLIMAN, A. A. Reliability estimation in generalized life-model with application to the Burr-XII. *IEEE Transactions on Reliability*, v. 51, n. 3, p. 337–343, 2002.

STERN, S. A second-order adjustment to the profile likelihood in the case of a multidimensional parameter of interest. *Jounal of the Royal Statistical Society*, B, n. 59, p. 653–665, 1997.

TEJEDA, H. A.; GOODWIN, B. *Modeling Crop prices through a Burr distribution and Analysis of Correlation between Crop Prices and Yields using a Copula method*. [S.l.], 2008. Disponível em: <https://EconPapers.repec.org/RePEc:ags:aaea08:6061>.

WANG, F. K.; KEATS, J. B. Maximum likelihood estimation of the Burr XII parameters with censored and uncensored data. *Microelectronics Reliability*, v. 36, n. 3, p. 359–362, 1983.

WEISBERG, S. *Applied linear regression*. [S.l.]: 4rd edition. Wiley, New York, 2014.

ZAHARIM, A.; NAJID, S. K.; RAZALI, A. M.; SOPIAN, K. Analyzing Malaysian Wind Speed Data Using Statistical Distribution. *Proceedings of the 4th IASME/WSEAS international conference on Energy & environment*, p. 363–70, 2009.

ZIMMER, W.; KEATS, J.; WANG, F. The Burr XII distribution in reliability analysis. *Journal of Quality Technology*, Milwaukee, v. 30, p. 389–394, 1998.