

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS
CURSO DE PÓS-GRADUAÇÃO E ESPECIALIZAÇÃO EM
ESTATÍSTICA E MODELAGEM QUANTITATIVA**

**UMA NOVA CLASSE DE DISTRIBUIÇÕES
UNITÁRIAS BASEADA NA FAMÍLIA
WEIBULL ESTENDIDA COM APLICAÇÃO
EM DADOS ELEITORAIS**

MONOGRAFIA DE ESPECIALIZAÇÃO

Charles Peixoto Mafalda

Santa Maria, RS, Brasil

2020

**UMA NOVA CLASSE DE DISTRIBUIÇÕES UNITÁRIAS
BASEADA NA FAMÍLIA WEIBULL ESTENDIDA COM
APLICAÇÃO EM DADOS ELEITORAIS**

Charles Peixoto Mafalda

Monografia apresentada ao Curso de Pós-Graduação e Especialização em Estatística e Modelagem Quantitativa da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para a obtenção do grau de **Especialista em Estatística e Modelagem Quantitativa**

Orientadora: Prof. Dr^a. Renata Rojas Guerra

Co-orientador: Prof. Dr. Fernando Arturo Peña-Ramírez

Santa Maria, RS, Brasil

2020

Peixoto Mafalda, Charles

Uma Nova Classe de Distribuições Unitárias Baseada na Família Weibull Estendida com aplicação em dados eleitorais / por Charles Peixoto Mafalda. – 2020.

57 f.: il.; 30 cm.

Orientadora: Renata Rojas Guerra

Co-orientador: Fernando Arturo Peña-Ramírez

Monografia (Especialização) - Universidade Federal de Santa Maria, Centro de Ciências Naturais e Exatas, Curso de Pós-Graduação e Especialização em Estatística e Modelagem Quantitativa, RS, 2020.

1. Eleições presidenciais de 2018. 2. Dados eleitorais. 3. Distribuições unitárias. 4. Modelos probabilísticos. 5. Região Nordeste. I. Rojas Guerra, Renata. II. Arturo Peña-Ramírez, Fernando. III. Título.

© 2020

Todos os direitos autorais reservados a Charles Peixoto Mafalda. A reprodução de partes ou do todo deste trabalho só poderá ser feita mediante a citação da fonte.

E-mail: charles1995peixoto@hotmail.com

**Universidade Federal de Santa Maria
Centro de Ciências Naturais e Exatas
Curso de Pós-Graduação e Especialização em Estatística e Modelagem
Quantitativa**

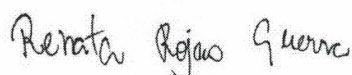
A Comissão Examinadora, abaixo assinada,
aprova a Monografia de Especialização

**UMA NOVA CLASSE DE DISTRIBUIÇÕES UNITÁRIAS BASEADA
NA FAMÍLIA WEIBULL ESTENDIDA COM APLICAÇÃO EM
DADOS ELEITORAIS**

elaborada por
Charles Peixoto Mafalda

como requisito parcial para obtenção do grau de
Especialista em Estatística e Modelagem Quantitativa

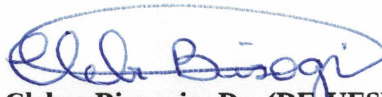
COMISSÃO EXAMINADORA:



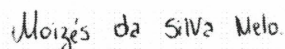
Renata Rojas Guerra, DR^a (DE-UFSM)
(Presidente/Orientadora)



Fernando Arturo Peña-Ramírez, Dr. (DE-UFSM)
(Co-orientador)



Cleber Bisognin, Dr. (DE-UFSM)



Moisés da Silva Melo, MC. (IME-USP)

Santa Maria, 17 de março de 2020.

RESUMO

Monografia de Especialização
Curso de Pós-Graduação e Especialização em Estatística e Modelagem Quantitativa
Universidade Federal de Santa Maria

UMA NOVA CLASSE DE DISTRIBUIÇÕES UNITÁRIAS BASEADA NA FAMÍLIA WEIBULL ESTENDIDA COM APLICAÇÃO EM DADOS ELEITORAIS

AUTOR: CHARLES PEIXOTO MAFALDA

ORIENTADORA: RENATA ROJAS GUERRA

CO-ORIENTADOR: FERNANDO ARTURO PEÑA-RAMÍREZ

Local da Defesa e Data: Santa Maria, 17 de março de 2020.

Neste trabalho é proposta uma família de distribuições chamada nova família Weibull estendida unitária (\mathcal{NWEU}), a qual é derivada a partir da transformação em uma variável aleatória que pertence à classe de distribuições Weibull estendida. São definidos e apresentados seis submodelos pertencentes à família, sendo eles: i) distribuição exponencial exponencializada unitária; ii) distribuição Gompertz unitária; iii) distribuição half-logistic unitária; iv) distribuição Lomax unitária; v) distribuição Nadarajah-Haghighi unitária e vi) Distribuição Rayleigh unitária. Foi proposta uma parametrização em termos dos quantis das distribuições. Apresentou-se um estudo detalhado sobre os estimadores de máxima verossimilhança dessa nova família. Realizou-se simulações de Monte Carlo para estudar o comportamento dos estimadores de máxima verossimilhança das distribuições Rayleigh unitária e Gompertz unitária. Verificou-se que o modelo Rayleigh unitário possui forma fechada para o estimador de máxima verossimilhança e que este é aproximadamente não viesado, mesmo para amostras pequenas. Por fim, os novos modelos foram ajustados à proporção de votos válidos para o segundo turno das eleições presidenciais brasileiras, considerando o ano de 2018 e a região nordeste. A escolha justifica-se pelo fato de que a região nordeste possui o segundo maior sítio eleitoral do Brasil. Além disso, esta é a única região em que o presidente eleito não obteve a maioria dos votos em todos os estados. Na aplicação verificou-se a superioridade dos modelos propostos ao comparar o ajuste com alguns modelos usuais no contexto de dados unitários, tais como as distribuições Beta, Kumaraswamy e gama unitária.

Palavras-chave: Eleições presidenciais de 2018. Dados eleitorais. Distribuições unitárias. Modelos probabilísticos. Região Nordeste.

ABSTRACT

Specialization Monograph

A New Class of Unit Distributions Based on the Extended Weibull Family with application in electoral data

Federal University of Santa Maria

UMA NOVA CLASSE DE DISTRIBUIÇÕES UNITÁRIAS BASEADA NA FAMÍLIA WEIBULL ESTENDIDA COM APLICAÇÃO EM DADOS ELEITORAIS

AUTHOR: CHARLES PEIXOTO MAFALDA

ADVISOR: RENATA ROJAS GUERRA

COADVISOR: FERNANDO ARTURO PEÑA-RAMÍREZ

Defense Place and Date: Santa Maria, March 17st, 2020.

We propose a family of distributions so-called the new unit extended Weibull family (\mathcal{NWEU}). It is derived from a transformation in an extended Weibull random variable. Six \mathcal{NWEU} submodels are studied, namely: i) exponential unitary exponential distribution; ii) unit Gompertz distribution; iii) unit half-logistic distribution; iv) unit Lomax distribution; v) unit Nadarajah-Haghighi distribution and vi) Unit Rayleigh distribution. We propose a quantile-parameterization for the new family. The maximum likelihood estimators (MLEs) are presented. A Monte Carlo study is performed to evaluate the behavior of the MLEs of unit Rayleigh and unit Gompertz. We note that the unit Rayleigh MLE has closed-form and is approximately unbiased MLE, even for small samples. Finally, the \mathcal{NWEU} submodels are adjusted to the proportion of the valid votes for the second round Brazilian presidential elections, considering the year 2018 and the northeast region. We focus on this region since it is the second-largest electoral site in Brazil. Besides, this is the only Brazilian region where the elected president did not receive the majority of electoral votes. The application shows the superiority of the new family when compared with other usual unit models, such as the Beta, Kumaraswamy, and unit gamma distributions.

Keywords: 2018 presidential elections. Election data. Unit distributions. Probability models. Northeast Region.

LISTA DE FIGURAS

Figura 3.1 – Gráficos das densidades para alguns casos particulares da família \mathcal{NWEU} . .	26
Figura 5.1 – VR(%) e EQM para o estimador de máxima verossimilhança da mediana da distribuição \mathcal{NRU}	37
Figura 5.2 – VR(%) total e EQM total para os estimadores de máxima verossimilhança da distribuição \mathcal{NUGO}	38
Figura 6.1 – Proporção dos votos válidos do 2º turno das eleições presidências no ano de 2018.	41
Figura 6.2 – Histograma dos votos válidos e alguns modelos ajustados.	45
Figura 1 – VR(%) e EQM para o estimador de máxima verossimilhança da distribuição \mathcal{NRU} , considerando $\tau = 0.25$	52
Figura 2 – VR(%) e EQM para o estimador de máxima verossimilhança da distribuição \mathcal{NRU} , considerando $\tau = 0.75$	53
Figura 3 – VR(%) e EQM para o estimador de máxima verossimilhança da distribuição \mathcal{NGOU} , considerando $\tau = 0.25$	55
Figura 4 – VR(%) e EQM para o estimador de máxima verossimilhança da distribuição \mathcal{NGOU} , considerando $\tau = 0.75$	57

LISTA DE TABELAS

Tabela 2.1 – Modelos especiais e as funções correspondentes $H(x; \xi)$, $H^{-1}(x; \xi)$ e $h(x; \xi)$	16
Tabela 2.2 – Algumas das principais propriedades das distribuições unitárias.....	19
Tabela 5.1 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para o parâmetro da distribuição \mathcal{NRU} , para diferentes cenários e tamanhos amostrais.....	36
Tabela 5.2 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para os parâmetros da distribuição \mathcal{NUGO} , para diferentes cenários e tamanhos amostrais.	39
Tabela 6.1 – Análise descritiva dos dados	42
Tabela 6.2 – Estimativas dos parâmetros obtidos via estimação de máxima verossimilhança dos modelos ajustados à proporção dos votos válidos da região nordeste no segundo turno das eleições presidenciais no ano de 2018 e seus correspondentes EQMs (dados entre parênteses) e os valores das estatísticas A^* e W^*	44
Tabela 1 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para o parâmetro da distribuição \mathcal{NRU} , para diferentes cenários e tamanhos amostrais, considerando $\tau = 0.25$	51
Tabela 2 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para o parâmetro da distribuição \mathcal{NRU} , para diferentes cenários e tamanhos amostrais, considerando $\tau = 0.75$	52
Tabela 3 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para os parâmetros da distribuição \mathcal{NUGO} , para diferentes cenários e tamanhos amostrais, considerando $\tau = 0.25$	54
Tabela 4 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para os parâmetros da distribuição \mathcal{NUGO} , para diferentes cenários e tamanhos amostrais, considerando $\tau = 0.75$	56

SUMÁRIO

1 INTRODUÇÃO	10
1.1 Objetivos	11
1.1.1 Objetivo geral	11
1.1.2 Objetivos específicos	11
1.2 Estrutura do trabalho	12
2 REVISÃO DE LITERATURA	13
2.1 Família de distribuições Weibull estendida	13
2.2 Distribuições com suporte duplamente limitado	17
3 NOVA FAMÍLIA WEIBULL ESTENDIDA UNITÁRIA	20
3.1 Definição da nova família	20
3.2 Alguns casos particulares	21
3.2.1 Distribuição exponencial exponencializada unitária	22
3.2.2 Distribuição Gompertz unitária	23
3.2.3 Distribuição Half-Logistic unitária	23
3.2.4 Distribuição Rayleigh unitária	24
3.2.5 Distribuição Lomax unitária	24
3.2.6 Distribuição Nadarajah-Haghighi unitária	25
4 ESTIMAÇÃO VIA MÁXIMA VEROSSIMILHANÇA	27
4.1 Estimado de máxima verossimilhança para a distribuição Rayleigh unitária	30
4.2 Estimador de máxima verossimilhança para a distribuição Gompertz unitária ...	31
5 ESTUDO DE SIMULAÇÃO	34
5.1 Resultados numéricos para a distribuição \mathcal{NRU}	35
5.2 Resultados numéricos para a distribuição \mathcal{NGOU}	37
6 APLICAÇÃO EM DADOS ELEITORAIS	40
6.1 Panorama Geral das Eleições	40
6.2 Aplicação do banco de dados referente a região nordeste	41
7 CONCLUSÃO	46
REFERÊNCIAS	47
APÊNDICES	50

1 INTRODUÇÃO

No contexto de dados com suporte limitado, isto é, que estão restritos a um intervalo do tipo (a, b) , em que a e b são escalares e $a < b$, não há uma gama tão ampla de distribuições quanto no caso dos reais positivos. Especificamente para o intervalo unitário, pode-se citar a distribuição Beta (\mathcal{BETA}) como uma das distribuições mais usuais. Na parametrização proposta por Ferrari e Cribari-Neto (2004), os parâmetros da distribuição \mathcal{BETA} são interpretados como a média e precisão da variável aleatória. Entretanto, a distribuição \mathcal{BETA} não tem função de distribuição acumulada em forma fechada. Uma alternativa é a distribuição Kumaraswamy (\mathcal{KW}) (KUMARASWAMY, 1980) que tem função de distribuição acumulada e, consequentemente, quantis em forma explícita. Por outro lado, este modelo não possui forma fechada para a média.

Devido ao crescente interesse por variáveis do tipo taxa ou proporção e à importância de se levar em consideração a natureza dos dados, outras distribuições unitárias têm sido consideradas além dos modelos \mathcal{BETA} e \mathcal{KW} . Algumas contribuições que podem ser citadas são os modelos Topp-Leone (\mathcal{TL}) (TOPP; LEONE, 1955), gama unitária (\mathcal{GU}) (GRASSIA, 1977), simplex unitária (\mathcal{SU}) (BARNDORFF-NIELSEN; JORGENSEN, 1991) e Birnbaum-Saunders unitária (\mathcal{BSU}) (MAZUCHELI; MENEZES; DEY, 2018), entre outras. Entretanto, os modelos usuais nem sempre são adequados às características dos dados. Nestas situações, é necessário considerar outras alternativas que permitam flexibilizar a escolha e que possuam propriedades desejáveis, tais como caudas pesadas e que modelem dados assimétricos.

Desta forma, o presente trabalho tem como propósito introduzir uma nova família de distribuições unitárias, aqui denominada como nova família Weibull estendida unitária (\mathcal{NWEU}). A família de distribuições proposta é obtida a partir da transformação em uma variável aleatória que pertence à classe Weibull Estendida (\mathcal{WE}) e é introduzida com foco no ajuste da proporção de voto em eleições presidenciais.

O resultado de eleições presidenciais desempenha um papel importante para a política nacional e de governança (ALEMÁN; KELLAM, 2017). Nesse sentido, a proporção de votos válidos permite mensurar o desempenho eleitoral dos candidatos. Conforme (ALEMÁN; KELLAM, 2017), o uso da proporção de votos justifica-se uma vez que o tamanho da população varia consideravelmente entre diferentes distritos eleitorais¹, tipicamente definidos como

¹ “O distrito eleitoral é a unidade territorial onde os votos são contabilizados para efeito de distribuição das cadeiras em disputa.”(NICOLAU, 2004)

municípios, estados e/ou regiões de um país. Além disso, o estudo desta variável permite mensurar outras características do cenário político, tais como a volatilidade eleitoral (POWELL; TUCKER, 2013) e nacionalização em mudanças eleitorais (ALEMÁN; KELLAM, 2017).

Nesse trabalho, é apresentado um estudo de caso referente à proporção de votos válidos para o segundo turno das eleições presidenciais brasileiras. Foi considerado o ano de 2018 e a região nordeste. No caso do Brasil, foi em 2018 que ocorreu a última eleição presidencial e, segundo Lancet (2018), esta foi uma disputa com um enfoque particular em temas como violência e corrupção. A escolha da região nordeste justifica-se pelo fato desta possuir o segundo maior sítio eleitoral do Brasil. Além disso, o nordeste é a única região em que o presidente eleito não obteve a maioria dos votos em todos os seus estados (ELEITORAL-TSE, 2018).

1.1 Objetivos

1.1.1 Objetivo geral

Propor uma nova família de distribuições unitárias obtida a partir de uma transformação na classe \mathcal{WE} , que modelem dados eleitorais.

1.1.2 Objetivos específicos

- Apresentar formas gerais para função densidade acumulada e função densidade de probabilidade da nova família.
- Propor uma reparametrização da nova família em termos dos quantis.
- Obter os estimadores dos parâmetros das distribuições da família \mathcal{NWEU} a partir do método de máxima verossimilhança.
- Realizar estudos de simulação para avaliar o desempenho dos estimadores de máxima verossimilhança para alguns casos particulares da família.
- Avaliar a viabilidade da utilização dos modelos propostos através de um estudo de caso das eleições presidenciais de 2018.

1.2 Estrutura do trabalho

O trabalho está dividido da seguinte forma. No Capítulo 2 é realizada uma revisão de literatura. No Capítulo 3, é definida a família $\mathcal{NW}\mathcal{EU}$. Na Seção 3.1 são definidas as formas gerais para a função densidade acumulada e função densidade de probabilidade. A Seção 3.2 apresenta seis casos especiais da família $\mathcal{NW}\mathcal{EU}$. No Capítulo 4, são obtidos os estimadores de máxima verossimilhança da família $\mathcal{NW}\mathcal{EU}$. São apresentadas expressões gerais para família e desenvolvidos resultados específicos para dois casos particulares. No Capítulo 5 é apresentado um estudo de simulação de Monte Carlo, em que avalia-se a acurácia e precisão das estimativas de máxima verossimilhança. A simulação foi realizada com duas distribuição que são casos especiais da família $\mathcal{NW}\mathcal{EU}$. Já no Capítulo 6, é realizada uma aplicação com dados reais do tipo taxas e proporção. Com o objetivo de encontrar um melhor modelo probabilístico que modele dados eleitorais. É realizada a comparação de seis submodelos da classe proposta neste trabalho com mais cinco distribuições que são usuais quando se trabalha com dados no intervalo unitário. O banco de dados utilizado trata-se do resultado do segundo turno das eleições presidenciais no ano de 2018 para a região nordeste do Brasil. Por fim, no Capítulo 7 são apresentadas as considerações finais desse trabalho.

2 REVISÃO DE LITERATURA

Neste capítulo, será apresentada uma revisão bibliográfica sobre os modelos probabilísticos relacionados a classe de distribuições proposta. Dentre esses, destaca-se a família de distribuições Weibull estendida, a partir da qual desenvolveu-se a transformação proposta. Além disso, é apresentada uma abordagem geral sobre os modelos probabilísticos que tem como suporte o intervalo unitário.

2.1 Família de distribuições Weibull estendida

Gurvich et. al. (1997) introduziram a família de distribuições Weibull Estendida (\mathcal{WE}) como uma alternativa ao modelo Weibull, clássico nos contextos de confiabilidade e análise de sobrevivência. Assim como no caso da Weibull, uma das vantagens da classe \mathcal{WE} em relação a outros modelos utilizados nestas áreas é o fato de que esta apresenta forma fechada para a sobrevida, bem como para a função taxa de falha. Além disso, Cordeiro e Silva (2014) utilizaram a classe \mathcal{WE} para obter distribuições de probabilidade mais flexíveis, como por exemplo, a distribuição Geométrica Weibull modificada complementar (\mathcal{GWMC}) e a distribuição de Poisson Pareto complementar (\mathcal{PPC}). Esses modelos foram introduzidos com o intuito de trabalhar com problemas de análise de sobrevivência que surgem em várias áreas, como a saúde pública, estudos biomédicos, demografia e confiabilidade industrial.

Seja X uma variável aleatória que segue uma distribuição da classe \mathcal{WE} , então sua função densidade acumulada é dada por

$$G(x) = 1 - \exp \{-\alpha H(x; \boldsymbol{\xi})\}, \quad (2.1)$$

em que $x > 0$, $\alpha > 0$ é um parâmetro de forma e $H(x; \boldsymbol{\xi})$ é uma função não-negativa, monótona crescente e que depende do vetor de parâmetros $\boldsymbol{\xi}$. Tem-se que, se $H(x; \boldsymbol{\xi})$ for uma função potência, a equação (2.1) reduz-se à distribuição Weibull clássica. A função densidade de probabilidade de X é dada por

$$g(x) = \alpha h(x; \boldsymbol{\xi}) \exp \{-\alpha H(x; \boldsymbol{\xi})\},$$

em que $h(x; \boldsymbol{\xi})$ é a derivada de $H(x; \boldsymbol{\xi})$ em relação a x .

Diversas distribuições conhecidas podem ser obtidas a partir de diferentes expressões de $H(x; \boldsymbol{\xi})$. Santos-Neto et. al. (2014) apresentam diversos modelos especiais que fazem parte da

classe \mathcal{WE} de distribuições, os quais são apresentados na Tabela 2.1 com suas funções $H(x; \boldsymbol{\xi})$, $H^{-1}(x; \boldsymbol{\xi})$, $h(x; \boldsymbol{\xi})$ e vetor paramétrico correspondentes. A função $H^{-1}(\cdot; \boldsymbol{\xi})$ é a função inversa de $H(\cdot; \boldsymbol{\xi})$. A função quantílica da família \mathcal{WE} é obtida pela inversa de (2.1) como

$$Q_Y(u|\alpha, \boldsymbol{\xi}) = H^{-1} \left[-\frac{1}{\alpha} \log(1 - u) \right].$$

Note que a função quantílica da família \mathcal{WE} depende apenas da função $H^{-1}(\cdot; \boldsymbol{\xi})$.

Novas classes de distribuições também têm sido propostas com o objetivo de proporcionar uma maior flexibilidade na modelagem de dados com suporte nos reais positivos. Dentre estas, cabe destacar alguns avanços obtidos através de composições entre a classe \mathcal{WE} e outras famílias existentes. A seguir, apresenta-se algumas dessas famílias.

- Família beta Weibull estendida (CORDEIRO; ORTEGA; SILVA, 2012), com função densidade de probabilidade dada por

$$f_{BWE}(y) = \frac{\alpha h(y; \boldsymbol{\xi})}{B(a,b)} [1 - \exp\{-\alpha H(y; \boldsymbol{\xi})\}]^{a-1} \exp\{-\alpha b H(y; \boldsymbol{\xi})\}, \quad 0 < y,$$

em que $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. Ainda, tem-se que $a > 0$, $b > 0$ e $\alpha > 0$ são três parâmetros de forma adicionais. De acordo com Cordeiro, Ortega e Silva (2012) essa nova família, pode ser usada em diversos problemas na modelagem de dados de sobrevivência devido a sua flexibilidade em acomodar diferentes formas para a função taxa de falha.

- Família Weibull estendida em série de potência (SILVA et al., 2013), com função densidade de probabilidade dada por

$$f_{WESP}(y) = \theta \alpha h(y; \boldsymbol{\xi}) \exp\{-\alpha H(y; \boldsymbol{\xi})\} \frac{C'(\theta \exp\{-\alpha H(y; \boldsymbol{\xi})\})}{C(\theta)}, \quad 0 < y,$$

em que α e $\theta > 0$. $C(\theta)$ é uma função de θ e $C'(\cdot)$ é a derivada de $C(\cdot)$ em relação a θ . Os autores realizaram aplicações de alguns casos especiais dessa família a dados de sobrevivência, em particular para o ajuste de concentrações de fósforo nas folhas de 128 plantas.

- Família Weibull estendida em série de potência complementar (CORDEIRO; SILVA, 2014), com sua função densidade de probabilidade dada por

$$f_{WESPC}(y) = \theta \alpha h(y; \boldsymbol{\xi}) \exp\{-\alpha H(y; \boldsymbol{\xi})\} \frac{C'[\theta(1 - \exp\{-\alpha H(y; \boldsymbol{\xi})\})]}{C(\theta)}, \quad 0 < y,$$

em que α e $\theta > 0$. Os autores realizaram uma aplicação dessa família a banco de dados que refere-se aos pontos fortes de fibras de vidro de 1,5 cm.

- Família Weibull estendida em série de potência exponencializada (TAHMASEBI; JAFARI, 2015), com função densidade de probabilidade dada por

$$f_{WESPE}(y) = \frac{\alpha \beta \lambda h(y; \boldsymbol{\xi}) \exp\{-\alpha H(y; \boldsymbol{\xi})\} [1 - \exp\{-\alpha H(y; \boldsymbol{\xi})\}]^{\beta-1} C' [\lambda [1 - \exp\{-\alpha H(y; \boldsymbol{\xi})\}]^{\beta}]}{C'(\lambda)}, \quad 0 < y,$$

em que α, β e $\lambda > 0$. $C(\lambda)$ é uma função de λ e $C'(\cdot)$ é a derivada de $C(\cdot)$ em relação a λ . Os autores aplicaram essa classe de distribuições a um conjunto de dados que consiste no tempo de falha de 20 componentes elétricos.

- Família Gama Weibull estendida (NASCIMENTO et al., 2014), com função densidade de probabilidade dada por

$$f_{GWE}(y) = \frac{\alpha^{\delta}}{\Gamma(\delta)} h(y; \boldsymbol{\xi}) H(y; \boldsymbol{\xi})^{\delta-1} \exp\{-\alpha H(y; \boldsymbol{\xi})\}, \quad 0 < y,$$

em que α e $\delta > 0$. Os autores realizaram uma aplicação, utilizando um banco de dados que consiste nos tempos de falha de 20 componentes mecânicos.

- Família Marshall-Olkin Weibull estendida (SANTOS-NETO et al., 2014), com função densidade de probabilidade dada por

$$f_{MOWE}(y) = \frac{\delta \alpha h(y; \boldsymbol{\xi}) \exp\{-\alpha H(y; \boldsymbol{\xi})\}}{[1 - (1 - \delta) \exp\{-\alpha H(y; \boldsymbol{\xi})\}]^2}, \quad 0 < y,$$

em que α e $\delta > 0$. Ainda, destaca-se que os autores realizaram uma aplicação a um banco de dados em que os dados apresentam a porcentagem de gordura corporal determinada pela pesagem sub-aquática para 250 homens.

Tabela 2.1 – Modelos especiais e as funções correspondentes $H(x; \xi)$, $H^{-1}(x; \xi)$ e $h(x; \xi)$

Distribuição	$H(x; \xi)$	$H^{-1}(x; \xi)$	$h(x; \xi)$	α	ξ	Referências
Exponencial ($x \geq 0$)	x	x	1	α	\emptyset	Johnson et al. (1994)
Pareto ($x \geq K$)	$\log(x/K)$	$K \exp(x)$	$1/x$	α	K	Johnson et al. (1994)
Burr XII ($x \geq 0$)	$\log(1 - x^c)$	$[1 - \exp(x)]^{1/c}$	$cx^{c-1}/(1 + x^c)$	α	c	Rodriguez (1977)
Lomax ($x \geq 0$)	$\log(1 + x)$	$\exp(x) - 1$	$1/(1 + x)$	α	\emptyset	Lomax (1954)
Log-logistic ($x \geq 0$)	$\log(1 + x^c)$	$[1 - \exp(x)]^{1/c}$	$cx^{c-1}/(1 + x^c)$	1	c	Fisk (1961)
Rayleigh ($x \geq 0$)	x^2	$x^{1/2}$	$2x$	α	\emptyset	Rayleigh (1880)
Weibull ($x \geq 0$)	x^y	$x^{1/y}$	yx^{y-1}	α	y	Johnson et al. (1994)
Fréchet ($x \geq 0$)	x^{-y}	$x^{-1/y}$	$-yx^{-(y+1)}$	α	y	Fréchet (1927)
Taxa de falha linear ($x \geq 0$)	$ax + bx^2/2$	(*)	$a + bx$	1	$[a, b]$	Bain (1974)
Weibull Modificada ($x \geq 0$)	$x^y \exp(\lambda x)$	(*)	$x^{y-1} \exp(\lambda x)(y + \lambda x)$	α	$[y, \lambda]$	Lai et al. (2003)
Weibull Estendida ($x \geq 0$)	$\lambda[\exp(x/\lambda)^\beta - 1]$	$\lambda \left[\log\left(\frac{x}{y} + 1\right) \right]^{1/\beta}$	$\beta \exp(x/\lambda)^\beta (x/\lambda)^{\beta-1}$	α	$[y, \lambda, \beta]$	Xie et al. (2002)
Phani ($0 < \mu < x < \sigma < \infty$)	$[(x - \mu) - (\sigma - x)]^\beta$	$(x^{1/\beta} + \mu + \sigma)/2$	$\beta[(x - \mu) - (\sigma - x)]^{\beta-1} [(\sigma - \mu)/(\sigma - t)]^2$	α	$[\mu, \sigma, \beta]$	Phani (1987)
Kies Weibull ($0 < \mu < x < \sigma < \infty$)	$(x - \mu)^{\beta_1} / (\sigma - x)^{\beta_2}$	(*)	$(x - \mu)^{\beta_1 - 1} / (\sigma - x)^{\beta_2 - 1} [\beta_1(\sigma - x) + \beta_2(x - \mu)]$	α	$[\mu, \sigma, \beta_1, \beta_2]$	Kies (1958)
Weibull Aditiva ($x \geq 0$)	$(x/\beta_1)^{\alpha_1} + (x/\beta_2)^{\alpha_2}$	(*)	$(\alpha_1/\beta_1)(x/\beta_1)^{\alpha_1 - 1} + (\alpha_2/\beta_2)(x/\beta_2)^{\alpha_2 - 1}$	1	$[\alpha_1, \alpha_2, \beta_1, \beta_2]$	Xie and Lai (1995)
Weibull Gen. Potência ($x \geq 0$)	$[1 + (x/\beta)^{\alpha_1}]^\theta - 1$	$\beta[(x + 1)^{1/\theta} - 1]^{1/\alpha_1}$	$(\theta \alpha_1 / \beta) [1 + (x/\beta)^{\alpha_1}]^{\theta-1} (x/\beta)^{\alpha_1}$	1	$[\alpha_1, \beta, \theta]$	Nikulin and Haghighi (2006)
Weibull Estendida Flexível ($x \geq 0$)	$\exp(\lambda x - \beta/x)$	(*)	$\exp(\lambda x - \beta/x)(\lambda + \beta/x^2)$	1	$[\lambda, \beta]$	Bebbington et al. (2007)
Gompertz ($x \geq 0$)	$\beta^{-1} [\exp(\beta x) - 1]$	$\beta^{-1} \log(\beta x + 1)$	$\exp(\beta x)$	α	β	Gompertz (1825)
Poder Exponencial ($x \geq 0$)	$\exp[(\lambda x)^\beta] - 1$	$\lambda^{-1} [\log(x + 1)]^{1/\beta}$	$\beta \lambda \exp[(\lambda x)^\beta] (\lambda x)^{\beta-1}$	1	$[\lambda, \cdot]$	Smith and Bain (1975)
Chen ($x \geq 0$)	$\exp(x^b) - 1$	$[\log(x + 1)]^{1/b}$	$b x^{b-1} \exp(x^b)$	α	b	Chen (2000)
Pham ($x \geq 0$)	$(a^x)^\beta - 1$	$\log(x + 1) / \beta \log(a)$	$\beta (a^x)^\beta \log(a)$	1	$[a, \beta]$	Pham (2002)

Fonte: Adaptado de Santos et. al. (2014). * Distribuições sem forma fechada para $H^{-1}(\cdot; \xi)$.

2.2 Distribuições com suporte duplamente limitado

Nesta seção, será apresentada uma breve revisão de literatura na qual se aborda algumas das principais distribuições que assumem valores no intervalo unitário. Nesse sentido, pode-se citar a distribuição \mathcal{BETA} como uma das distribuições mais usuais. Na reparametrização proposta por Ferrari e Cribari-Neto (2004), os parâmetros da distribuição beta são interpretados como média e precisão. A sua função densidade de probabilidade é dada por

$$f_{Beta}(y) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1, \quad (2.2)$$

com média $\mu \in (0,1)$ e parâmetro de precisão $\phi > 0$.

Entretanto, a função densidade acumulada da distribuição \mathcal{BETA} não tem uma forma fechada. Uma alternativa é a distribuição \mathcal{KWU} (KUMARASWAMY, 1980) que tem função densidade acumulada e, conseqüentemente, quantis em forma explícita. Por outro lado, não é possível obter uma reparametrização para este modelo em relação a sua média, por não ter expressão fechada para seus momentos. Diante deste fato, Mitnik e Baek (2013) apresentaram uma reparametrização do modelo \mathcal{KWU} em termos de sua mediana. A sua função densidade de probabilidade é

$$f_{KWU}(y) = \frac{\phi \log(0.5)}{\log(1-\mu^\phi)} y^{\phi-1} (1-y^\phi)^{\log(0.5)/(\log(1-\mu^\phi)-1)}, \quad 0 < y < 1, \quad (2.3)$$

em que $\mu \in (0,1)$ é a mediana e $\phi > 0$ é um parâmetro de precisão.

Devido ao crescente interesse por variáveis do tipo taxa ou proporção e à importância de se levar em consideração a natureza dos dados, outras distribuições unitárias têm sido consideradas além dos modelos \mathcal{BETA} e \mathcal{KWU} . Por exemplo, têm-se as seguintes distribuições:

- Distribuição Gama Unitária (\mathcal{GU}).

A família de distribuições \mathcal{GU} , introduzida por Grassia (1977), recebeu esse nome pelo fato de que se X segue uma distribuição gama com parâmetros de forma α e $\tau > 0$, então $Y = e^{-X}$ segue uma distribuição \mathcal{GU} . Ainda, foi introduzida uma reparametrização em termos de sua média, considerando $\mu = E(Y)$ e $\phi = \tau$. Seja Y uma variável aleatória com distribuição \mathcal{GU} com média μ e parâmetro de precisão ϕ , sua função densidade de probabilidade é dada por

$$f_{GU}(y) = \left[\frac{\mu^{1/\phi}}{1-\mu^{1/\phi}} \right]^\phi \frac{1}{\Gamma(\phi)} y^{\mu^{1/\phi}/[(1-\mu^{1/\phi})-1]} [-\log(y)]^{\phi-1}, \quad 0 < y < 1. \quad (2.4)$$

Salienta-se, que assim como a distribuição \mathcal{KWU} , a distribuição \mathcal{GU} não tem a propriedade de reflexão. O único caso de perfeita simetria é quando avalia-se os parâmetros $\mu = 0.5$ e $\phi = 1$ que corresponde à distribuição uniforme. Além disso, é possível obter as formas assimétrica, U, J, J invertido e til.

- Distribuição Simplex Unitária (SU).

Seja Y uma variável aleatória com distribuição SU , $Y \sim S(\mu, \phi)$, sua função densidade de probabilidade é dada por

$$f_{SU} = \left[\frac{\phi}{2\pi (y - y^2)^3} \right]^{1/2} \exp \left\{ -\frac{\phi}{2} \frac{(y - \mu)^2}{y(1 - y) \mu^2(1 - \mu)^2} \right\} \quad 0 < y < 1.$$

em que $\mu \in (0,1)$ é a média e $\phi > 0$ é o parâmetro de precisão. Essa distribuição foi proposta por Bardorff-Nielsen e Jorgensen (1991) e sua função densidade acumulada não tem forma fechada. Assim como a distribuição \mathcal{BETA} , a distribuição SU tem a propriedade de reflexão, ou seja, se $Y \sim S(\mu, \phi)$, então $Y \sim S(1 - \mu, \phi)$. Além disso, essa distribuição pode apresentar as formas unimodal ou bimodal e assimétrica (LIMA, 2018).

- Distribuição Birnbaum-Saunders unitária (BSU).

A distribuição BSU , introduzida por Mazucheli et. al. (2018), tem função densidade de probabilidade dada por

$$f_{BSU}(y) = \frac{1}{2y\alpha\beta\sqrt{2\pi}} \left[\left(-\frac{\beta}{\log(y)} \right)^{\frac{1}{2}} + \left(-\frac{\beta}{\log(y)} \right)^{\frac{3}{2}} \right] \exp \left\{ \frac{1}{2\alpha^2} \left(\frac{\log(y)}{\beta} + \frac{\beta}{\log(y)} + 2 \right) \right\}, \quad 0 < y < 1, \quad (2.5)$$

em que $\beta > 0$ é um parâmetro de escala e $\alpha > 0$ um parâmetro de forma. A distribuição BSU é flexível para modelar conjuntos de dados assimétricos positivos, que podem exibir taxa de falha crescente ou em forma de banheira. Mazucheli (2018) ajustou esse modelo a um banco de dados composto por observações mensais da capacidade de água do reservatório de Shasta na Califórnia, EUA. O autor ainda, fez um estudo comparativo com as distribuições \mathcal{BETA} e \mathcal{KWU} .

Por fim, temos a distribuição Weibull Unitária (WU) reparametrizada em termos dos seus quantis. Essa reparametrização foi proposta por Mazucheli et. al. (2019). A sua função densidade de probabilidade é dada por

$$f_{WEU}(y) = \frac{\beta}{y} \left(\frac{\log(\tau)}{\log(\mu)} \right) \left(\frac{\log(y)}{\log(\mu)} \right)^{\beta-1} \tau^{(\log(y)/\log(\mu))^\beta}, \quad 0 \leq 0 < y < 1, \quad (2.6)$$

em que $\mu \in (0,1)$ é o parâmetro associado ao τ -ésimo quantil da distribuição, $\beta > 0$ é o parâmetro de forma e τ é assumido como conhecido. Além disso, segundo Mazucheli (2019), a função densidade de probabilidade da distribuição \mathcal{WU} pode assumir formas diferentes (decrecente, crescente, unimodal, anti-modal) de acordo com os valores que seus parâmetros assumem.

Segundo Lima (2018), é possível estimar de maneira relativamente simples os parâmetros dessas distribuições pelo método de máxima verossimilhança. Além disso, o autor ressalta que pode ser desejável que a função densidade acumulada da distribuição tenha forma fechada. Assim, utilizando a propriedade de invariância dos estimadores de máxima verossimilhança facilmente consegue-se estimar os quantis (LIMA, 2018). A seguir, apresenta-se na Tabela 2.2 um resumo de algumas características das distribuições apresentadas nesta seção. Destaca-se que essa tabela foi adaptada de Lima (2018), acrescentando algumas novas distribuições citadas nesse trabalho.

Tabela 2.2 – Algumas das principais propriedades das distribuições unitárias.

	<i>BETA</i>	<i>KWU</i>	<i>GU</i>	<i>SU</i>	<i>BSU</i>	<i>WU</i>
Simétrica	sim	não	sim	sim	sim	sim
Assimetria	sim	sim	sim	sim	sim	sim
J e J invertido	sim	sim	sim	não	sim	sim
U	sim	sim	sim	não	sim	sim
Bimodal	não	não	não	sim	sim	não
Til e til invertido	sim	sim	sim	não	sim	sim
Reflexão	sim	não	não	sim	não	não
Fda fechada	não	sim	não	não	sim	sim

Fonte: Adaptado de Lima (2018).

3 NOVA FAMÍLIA WEIBULL ESTENDIDA UNITÁRIA

Neste capítulo é definida a nova família de distribuições unitárias através das formas gerais para sua função densidade acumulada e função densidade de probabilidade. Além disso, são desenvolvidas expressões explícitas para seis casos especiais da nova família.

3.1 Definição da nova família

Seja X uma variável aleatória com suporte nos reais positivos. Aplicando a transformação $Y = X/(1 + X)$, obtém-se uma nova variável aleatória que assumirá valores no intervalo $(0,1)$. Neste trabalho, a família de distribuições $\mathcal{NW}\mathcal{E}\mathcal{U}$ é obtida tomando a referida transformação, em que X é uma variável aleatória pertencendo à família \mathcal{WE} . Desta forma, seja Y uma variável aleatória da família de distribuições $\mathcal{NW}\mathcal{E}\mathcal{U}$, sua função densidade acumulada é dada por

$$F_Y(y) = 1 - \exp \left\{ -\alpha H \left(\frac{y}{1-y}; \boldsymbol{\xi} \right) \right\}, \quad (3.1)$$

em que $0 < y < 1$, $\alpha > 0$ é o parâmetro de escala e $\boldsymbol{\xi}$ é o vetor paramétrico associado à função $H(\cdot; \boldsymbol{\xi})$ definida como em (2.1). A função densidade de probabilidade correspondente a (3.1) é dada por

$$f_Y(y) = \alpha (1-y)^{-2} h \left(\frac{y}{1-y}; \boldsymbol{\xi} \right) \exp \left\{ -\alpha H \left(\frac{y}{1-y}; \boldsymbol{\xi} \right) \right\}. \quad (3.2)$$

A função quantílica, definida como $Q(u) = F_Y^{-1}(x)$, está expressa em termos de $H^{-1}(\cdot; \boldsymbol{\xi})$, que é a função inversa de $H(\cdot; \boldsymbol{\xi})$. Portanto, a função quantílica da nova família é dada por

$$Q_Y(u|\alpha; \boldsymbol{\xi}) = \frac{H^{-1} \left(\frac{-\log(1-u)}{\alpha}; \boldsymbol{\xi} \right)}{1 + H^{-1} \left(\frac{-\log(1-u)}{\alpha}; \boldsymbol{\xi} \right)}. \quad (3.3)$$

Pode-se verificar que função quantílica da nova família apresenta uma expressão analítica simples, isto é, sua estrutura não envolve nenhuma função especial. Desta forma, é possível obter facilmente uma reparametrização alternativa para a família $\mathcal{NW}\mathcal{E}\mathcal{U}$ em termos dos quantis da distribuição. A seguir é apresentada uma proposta de reparametrização.

Proposição 3.1.1 *Seja Y uma variável aleatória com distribuição $\mathcal{NW}\mathcal{E}\mathcal{U}$. Então a função de distribuição acumulada de Y pode ser reescrita como*

$$F_Y(y) = 1 - \left(\frac{1}{1-\tau} \right)^{-H\left(\frac{y}{1-y}; \boldsymbol{\xi}\right)/H\left(\frac{\mu}{1-\mu}; \boldsymbol{\xi}\right)}, \quad y \in (0,1), \quad (3.4)$$

em que $\mu \in (0,1)$ é o τ -ésimo quantil de Y e $\boldsymbol{\xi}$ é o vetor paramétrico associado à função $H(\cdot; \boldsymbol{\xi})$.

Prova: Substituindo $\alpha = \log(1 - \tau)^{-1} / H\left(\frac{\mu}{1-\mu}; \boldsymbol{\xi}\right)$ na equação (3.1), obtemos diretamente a função densidade acumulada dada em (3.4). Neste caso, a função quantílica de Y é dada por

$$Q_Y(u|\mu, \boldsymbol{\xi}) = \frac{H^{-1}\left(-H\left(\frac{\mu}{1-\mu}; \boldsymbol{\xi}\right) \frac{\log(1-u)}{\log(1-\tau)^{-1}}; \boldsymbol{\xi}\right)}{1 + H^{-1}\left(-H\left(\frac{\mu}{1-\mu}; \boldsymbol{\xi}\right) \frac{\log(1-u)}{\log(1-\tau)^{-1}}; \boldsymbol{\xi}\right)}.$$

Tomando $u = \tau$ na equação acima, tem-se que $Q_Y(\tau|\mu, \boldsymbol{\xi}) = \mu$. Com isso, a prova fica concluída.

A função densidade de probabilidade da $\mathcal{NW}\mathcal{EU}$ baseada na reparametrização anterior é dada por

$$f_Y(y) = \frac{(1-y)^{-2} \log(1-\tau)^{-1}}{H\left(\frac{\mu}{1-\mu}; \boldsymbol{\xi}\right)} h\left(\frac{y}{1-y}; \boldsymbol{\xi}\right) \left(\frac{1}{1-\tau}\right)^{-H\left(\frac{y}{1-y}; \boldsymbol{\xi}\right)/H\left(\frac{\mu}{1-\mu}; \boldsymbol{\xi}\right)}, \quad (3.5)$$

em que $y \in (0,1)$. Têm-se que τ é um valor fixo, a ser definido de acordo com o interesse do pesquisador. Por exemplo, pode-se obter uma reparametrização em termos do primeiro quartil fixando $\tau = 0,25$. Analogamente, ao fixar $\tau = 0,5$ obtém-se uma parametrização em termos da mediana da nova família. Uma vantagem de utilizar uma parametrização em termos da mediana é o fato de que, de acordo com Rocke (1989), esta medida é pouco sensível na presença de *outliers*, isto é, apresenta resistência a valores extremos.

Pode-se encontrar na literatura abordagens análogas à introduzida na Proposição (3.1.1), porém considerando outras distribuições unitárias. Com respeito a reparametrizações em termos da mediana, pode-se mencionar as propostas de Mitnik e Baek (2013) para a distribuição \mathcal{KW} , Lemonte e Bazán (2016) para o modelo Johnson SB e Guerra et. al. (2020) para a família Weibull estendida unitária. Para uma reparametrização mais geral, em termos de qualquer quantil, tem-se Mazucheli et. al. (2019) como referência.

3.2 Alguns casos particulares

Cada diferente forma funcional de $H(x; \boldsymbol{\xi})$ em (3.1), gera uma nova distribuição na classe $\mathcal{NW}\mathcal{EU}$ de distribuições. A Tabela (2.1) pode ser utilizada para a obtenção de modelos da nova família unitária. Nessa seção, serão apresentados em detalhes alguns destes

casos especiais, a saber: i) Nova distribuição exponencial exponencializada unitária ($\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}$); ii) Nova distribuição Gompertz unitária ($\mathcal{N}\mathcal{G}\mathcal{O}\mathcal{U}$); iii) Nova distribuição Half-Logistic unitária ($\mathcal{N}\mathcal{H}\mathcal{L}\mathcal{U}$); iv) Nova distribuição Lomax unitária ($\mathcal{N}\mathcal{L}\mathcal{U}$); v) Nova distribuição Nadarajah-Haghighi unitária ($\mathcal{N}\mathcal{N}\mathcal{H}\mathcal{U}$); vi) Nova distribuição Rayleigh unitária ($\mathcal{N}\mathcal{R}\mathcal{U}$). Serão desenvolvidas as propriedades matemáticas de cada caso, bem como gráficos para ilustrar a flexibilidade das novas densidades unitárias. Ressalta-se que as propriedades apresentadas são baseadas na reparametrização dada na Proposição (3.1.1).

3.2.1 Distribuição exponencial exponencializada unitária

Considere a distribuição exponencial exponencializada como base na família \mathcal{WE} . Aplicando a transformação proposta na seção (3.1), a $\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}$ possui função densidade acumulada dada por

$$F_{\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}}(y) = \left[1 - \left(\frac{1}{1 - \tau^{1/\beta}} \right)^{-y/[(1-\mu)(1-y)]} \right]^\beta, \quad (3.6)$$

em que $\mu \in (0,1)$ é o τ -ésimo quantil e $\beta > 0$ é um parâmetro de forma. A função densidade de probabilidade correspondente é

$$f_{\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}}(y) = \frac{\beta \log(1 - \tau^{1/\beta})^{-1}}{(1 - \mu)(1 - y)^2} \left(\frac{1}{1 - \tau^{1/\beta}} \right)^{-y/[(1-\mu)(1-y)]} \left[1 - \left(\frac{1}{1 - \tau^{1/\beta}} \right)^{-y/[(1-\mu)(1-y)]} \right]^{\beta-1}. \quad (3.7)$$

A Figura 3.1(a) apresenta o gráfico da função densidade de probabilidade da distribuição $\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}$, considerando diferentes valores dos parâmetros μ e β . Sem perda de generalidade, foi fixado τ em 0,5 dado que a ordem do quantil não interfere nas formas das densidades da distribuição. Pela Figura 3.1(a), é possível analisar que a função densidade de probabilidade da distribuição $\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}$, assume forma unimodal, crescente e decrescente.

Uma das características importante da nova família é que a sua função quantílica apresenta uma expressão simples. A função quantílica é dada por

$$Q_{\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}}(u|\mu, \beta, \tau) = \frac{(1 - \mu) \log(1 - \mu^{1/\beta})}{(1 - \mu) \log(1 - u^{1/\beta}) - \log(1 - \tau^{1/\beta})^{-1}}. \quad (3.8)$$

3.2.2 Distribuição Gompertz unitária

Considere a distribuição Gompertz como base na família \mathcal{WE} . De forma análoga ao caso da \mathcal{NEEU} , a \mathcal{NGOU} é caracterizada pela função densidade acumulada

$$F_{\mathcal{NGOU}}(y) = 1 - \left(\frac{1}{1-\tau} \right)^{-(\exp\{\beta\frac{y}{1-y}\}-1)/(\exp\{\frac{\mu\beta}{1-\mu}\}-1)}, \quad (3.9)$$

em que $\beta > 0$ é um parâmetro de forma, $\mu \in (0,1)$ é o τ -ésimo quantil da distribuição. A função densidade de probabilidade correspondente é dada por

$$f_{\mathcal{NGOU}}(y) = \frac{\beta \log(1-\tau)^{-1} \exp\left\{\beta\frac{y}{1-y}\right\}}{(1-y)^2 \exp\left\{\frac{\mu\beta}{1-\mu}\right\} - (1-y)^2} \left(\frac{1}{1-\tau} \right)^{-(\exp\{\beta\frac{y}{1-y}\}-1)/(\exp\{\frac{\mu\beta}{1-\mu}\}-1)}. \quad (3.10)$$

A Figura 3.1(b) apresenta o gráfico da função densidade de probabilidade da \mathcal{NGOU} , considerando diferentes valores do vetor de parâmetros $\theta = (\mu, \beta)^T$ e $\tau = 0,5$. Analisando o comportamento gráfico apresentado na Figura 3.1(b) da função densidade de probabilidade da distribuição \mathcal{NGOU} , é possível verificar que a mesma assume forma unimodal.

A função quantílica é dada por

$$Q_{\mathcal{NGOU}}(u|\mu, \beta, \tau) = \frac{\log\left[-\frac{\log(1-u)}{\log(1-\tau)^{-1}} \left[\exp\left\{\frac{\beta\mu}{1-\mu}\right\} - 1\right] + 1\right]}{\beta + \log\left[-\frac{\log(1-u)}{\log(1-\tau)^{-1}} \left[\exp\left\{\frac{\beta\mu}{1-\mu}\right\} - 1\right] + 1\right]}. \quad (3.11)$$

3.2.3 Distribuição Half-Logistic unitária

Considere a distribuição Half-Logistic como base na família \mathcal{WE} . De forma análoga ao caso da \mathcal{NEEU} , a distribuição \mathcal{NHLU} é caracterizada pela função densidade acumulada e função densidade de probabilidade dadas por

$$F_{\mathcal{NHLU}}(y) = \frac{1 - \left(\frac{1+\tau}{1-\tau}\right)^{-y(1-\mu)/[\mu(1-y)]}}{1 + \left(\frac{1+\tau}{1-\tau}\right)^{-y(1-\mu)/[\mu(1-y)]}} \quad (3.12)$$

e

$$f_{\mathcal{NHLU}}(y) = \frac{2(1-\mu) \log\left(\frac{1+\tau}{1-\tau}\right)}{\mu(1-y)^2} \left(\frac{1+\tau}{1-\tau}\right)^{-y(1-\mu)/[\mu(1-y)]} \times \left[1 + \left(\frac{1+\tau}{1-\tau}\right)^{-y(1-\mu)/[\mu(1-y)]}\right]^{-2}, \quad (3.13)$$

respectivamente, em que $\mu \in (0,1)$ é o τ -ésimo quantil. A Figura 3.1(c), apresenta o gráfico da função densidade de probabilidade da distribuição \mathcal{NHLU} , considerando diferentes valores do parâmetro μ e $\tau = 0,5$. Analisando o comportamento gráfico da função densidade de probabilidade, é possível verificar que a mesma assume forma unimodal e crescente.

A função quantílica é dada por

$$Q_{\mathcal{NHLU}}(u|\mu, \tau) = \frac{\mu \log\left(\frac{1-u}{1+u}\right) - \mu \log\left(\frac{1-\tau}{1+\tau}\right)}{\mu \log\left(\frac{1-u}{1+u}\right) - \log\left(\frac{1-\tau}{1+\tau}\right)}. \quad (3.14)$$

3.2.4 Distribuição Rayleigh unitária

Considere a distribuição Rayleigh como base na família \mathcal{WE} . De forma análoga ao caso da \mathcal{NEEU} , a \mathcal{NRU} é caracterizada pela função densidade acumulada e função densidade de probabilidade dadas como

$$F_{\mathcal{NRU}}(y) = 1 - \left(\frac{1}{1-\tau}\right)^{-[y^2(1-\mu)^2]/[\mu^2(1-y)^2]}, \quad (3.15)$$

e

$$f_{\mathcal{NRU}}(y) = \frac{2y(1-\mu)^2}{\mu^2(1-y)^3} \log(1-\tau)^{-1} \left(\frac{1}{1-\tau}\right)^{-[y^2(1-\mu)^2]/[\mu^2(1-y)^2]}, \quad (3.16)$$

respectivamente, em que $\mu \in (0,1)$ é o τ -ésimo quantil da distribuição.

A Figura 3.1(d) apresenta o gráfico da função densidade de probabilidade da \mathcal{NRU} , considerando diferentes valores do parâmetro μ e $\tau = 0,5$. Analisando o comportamento gráfico da função densidade de probabilidade, é possível verificar que a mesma assume forma unimodal.

A função quantílica da \mathcal{NRU} é dada por

$$Q_{\mathcal{NRU}}(u|\mu, \tau) = \frac{\left[\frac{-\mu^2 \log(1-u)}{(1-\mu)^2 \log(1-\tau)^{-1}}\right]^{1/2}}{1 + \left[\frac{-\mu^2 \log(1-u)}{(1-\mu)^2 \log(1-\tau)^{-1}}\right]^{1/2}}. \quad (3.17)$$

3.2.5 Distribuição Lomax unitária

Considere a distribuição Lomax como base na família \mathcal{WE} . De forma análoga ao caso da \mathcal{NEEU} , a \mathcal{NLU} é caracterizada pela função densidade acumulada e função densidade de probabilidade dadas por

$$F_{\mathcal{NLU}}(y) = 1 - \left[\frac{\beta(1-y)}{\beta - \beta y + y}\right]^{\log(1-\tau)/[\log[\beta(1-\mu)] - \log[\beta + \beta(1-\mu)]]} \quad (3.18)$$

e

$$f_{\mathcal{N}\mathcal{L}\mathcal{U}}(y) = \frac{\log(1-\tau)}{\beta \log[\beta(1-\mu)] - \log[\beta + \beta(1-\mu)]} \left[\frac{\beta - \beta y}{\beta - \beta y + y} \right]^{\log(1-\tau)/[\log[\beta(1-\mu)] - \log[\beta + \beta(1-\mu)]] + 1} \frac{\beta}{(\beta - \beta y + y)^2}, \quad (3.19)$$

respectivamente, em que $\beta > 0$ é um parâmetro de forma e $\mu \in (0,1)$ é o τ -ésimo quantil da distribuição.

A Figura 3.1(e), apresenta o gráfico da função densidade de probabilidade da distribuição $\mathcal{N}\mathcal{L}\mathcal{U}$, considerando diferentes valores do parâmetro μ e $\tau = 0,5$. Analisando o comportamento gráfico da função densidade de probabilidade da distribuição $\mathcal{N}\mathcal{L}\mathcal{U}$, é possível verificar que a mesma assume forma decrescente.

A função quantílica é dada por

$$Q_{\mathcal{N}\mathcal{L}\mathcal{U}}(u|\mu, \beta, \tau) = \frac{\beta - \beta(1-u)^{[\log[\beta(1-\mu)] - \log[\beta + \beta(1-\mu)]]/\log(1-\tau)}}{(1-\beta)(1-u)^{[\log[\beta(1-\mu)] - \log[\beta + \beta(1-\mu)]]/\log(1-\tau)} - \beta}. \quad (3.20)$$

3.2.6 Distribuição Nadarajah-Haghighi unitária

Considere a distribuição Nadarajah-Haghighi como base na família $\mathcal{W}\mathcal{E}$. De forma análoga ao caso da $\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}$, a $\mathcal{N}\mathcal{N}\mathcal{H}\mathcal{U}$ é caracterizada pela função densidade acumulada

$$F_{\mathcal{N}\mathcal{N}\mathcal{H}\mathcal{U}}(y) = 1 - \exp \left\{ 1 - \left[1 + \frac{y [1 - \log(1-\tau)]^{1/\beta} (1-\mu) + \mu y - y}{\mu (1-y)} \right]^\beta \right\}, \quad (3.21)$$

a função densidade de probabilidade é dada por

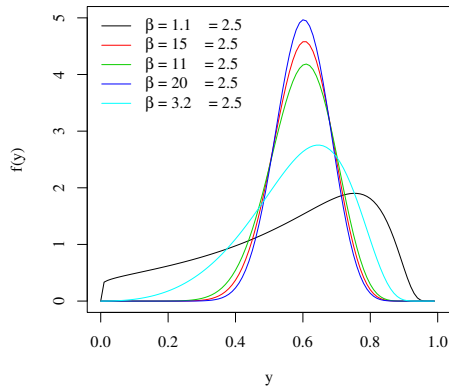
$$f_{\mathcal{N}\mathcal{N}\mathcal{H}\mathcal{U}}(y) = \frac{\beta [1 - \log(1-\tau)]^{1/\beta} (1-\mu) + \mu \beta - \beta}{\mu (1-y)^2} \left[1 + \frac{y [1 - \log(1-\tau)]^{1/\beta} (1-\mu) + \mu y - y}{\mu (1-y)} \right]^{\beta-1} \exp \left\{ 1 - \left[1 + \frac{y [1 - \log(1-\tau)]^{1/\beta} (1-\mu) + \mu y - y}{\mu (1-y)} \right]^\beta \right\}, \quad (3.22)$$

em que $\beta > 0$ é um parâmetro de forma e $\mu \in (0,1)$ é o τ -ésimo quantil da distribuição.

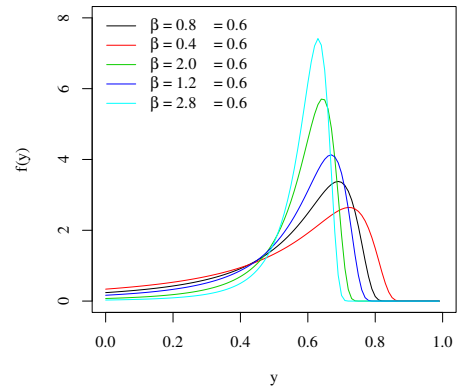
A Figura 3.1(f), apresenta o gráfico da função densidade de probabilidade da $\mathcal{N}\mathcal{N}\mathcal{H}\mathcal{U}$, considerando diferentes valores do parâmetro μ e $\tau = 0,5$. Analisando o comportamento gráfico da função densidade de probabilidade da $\mathcal{N}\mathcal{N}\mathcal{H}\mathcal{U}$, é possível verificar que a mesma assume forma unimodal e decrescente.

A função quantílica é dada por

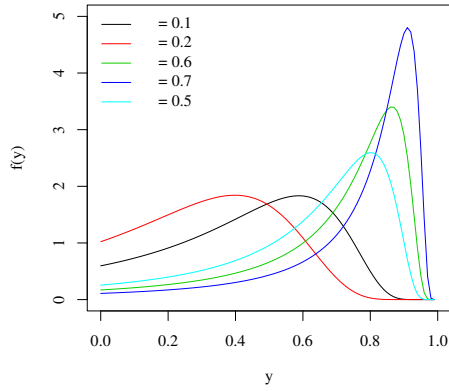
$$Q_{\mathcal{N}\mathcal{N}\mathcal{H}\mathcal{U}}(u|\mu, \beta, \tau) = \frac{\mu [1 - \log(1-u)]^{1/\beta} - \mu}{(1-\mu)[1 - \log(1-\tau)]^{1/\beta} + \mu[1 - \log(1-u)]^{1/\beta} - 1}. \quad (3.23)$$



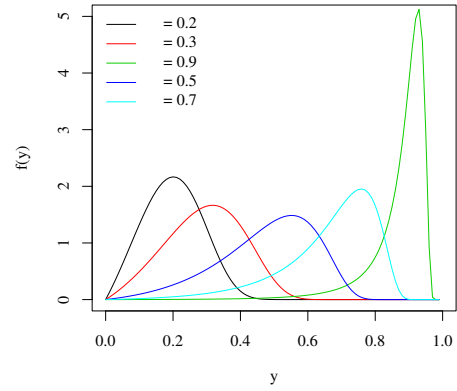
(a) $\mathcal{N}\mathcal{E}\mathcal{E}\mathcal{U}$



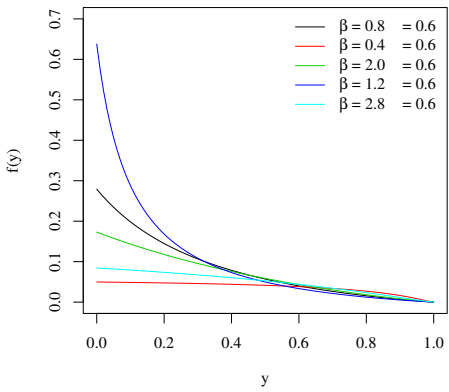
(b) $\mathcal{N}\mathcal{G}\mathcal{E}\mathcal{U}$



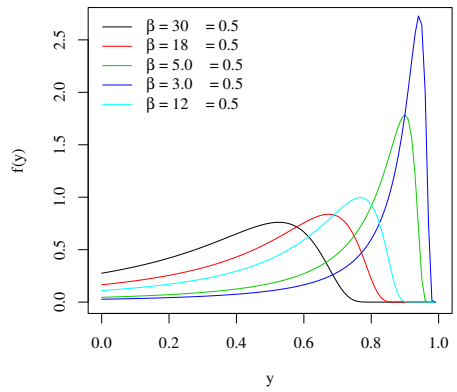
(c) $\mathcal{N}\mathcal{H}\mathcal{L}\mathcal{U}$



(d) $\mathcal{N}\mathcal{R}\mathcal{U}$



(e) $\mathcal{N}\mathcal{L}\mathcal{U}$



(f) $\mathcal{N}\mathcal{N}\mathcal{H}\mathcal{U}$

Figura 3.1 – Gráficos das densidades para alguns casos particulares da família $\mathcal{N}\mathcal{W}\mathcal{E}\mathcal{U}$.

4 ESTIMAÇÃO VIA MÁXIMA VEROSSIMILHANÇA

Neste trabalho, a estimação dos parâmetros dos modelos é feita utilizando o método de máxima verossimilhança, que é baseado na função de verossimilhança. A partir de um conjunto de dados e da suposição de uma distribuição de probabilidade para a variável de interesse, é possível obter os estimadores de máxima verossimilhança para os diferentes parâmetros associados à cada modelo. Os estimadores de máxima verossimilhança apresentam as seguintes propriedades assintóticas (BOLFARINE; SANDOVAL, 2001):

- **Consistência:** os estimadores de máxima verossimilhança são consistentes, isto é, elas convergem em probabilidade para o valor do parâmetro. Ou seja, para grandes amostras ($n \rightarrow \infty$) os estimadores de máxima verossimilhança, para efeitos práticos, são não-viesadas.
- **Eficiência Assintótica:** O Teorema Cramer-Rao afirma que, para um dado parâmetro, existe um limite inferior para a variância de qualquer estimador não viesado. Esse limite é chamado de Limite Inferior de Cramer-Rao, o qual garante que nenhum estimador não viesado terá variância menor que este limite. Para grandes amostras, o estimador de máxima verossimilhança atinge o limite de Cramer-Rao e, portanto, tem variância assintótica mínima dentre todos os estimadores não viesados. Isso mostra que o estimador de máxima verossimilhança é assintoticamente eficiente.
- **Normalidade Assintótica:** À medida que o tamanho de amostra crescer, o estimador de máxima verossimilhança converge em distribuição para distribuição Gaussiana. Ou seja, o estimador de máxima verossimilhança segue, assintoticamente, distribuição normal.
- **Invariância:** O estimador de máxima verossimilhança é invariante sob transformações monótonas, isto é, se $\hat{\mu}$ é um estimador de μ e $g(\mu)$ é uma função bijetora, então $g(\hat{\mu})$ é um estimador para $g(\mu)$.

Estas propriedades constituem vantagens da utilização do estimador de máxima verossimilhança. A partir do método de máxima verossimilhança também é possível obter estimativas intervalares. Segundo Bolfarine e Sandoval (2001), os intervalos de confiança clássicos são obtidos a partir de variáveis aleatórias especiais, denominadas de quantidades pivotais. Os autores ainda ressaltam que, um vetor aleatório $Q(Y_1, \dots, Y_n; \theta) = Q(\mathbf{Y}; \theta)$ é dito ser uma quantidade pivotal para o vetor de parâmetros θ se sua distribuição independe de θ .

Bolfarine e Sandoval (2001) destacam que é possível obter intervalos de confiança aproximados baseados na distribuição assintótica dos estimadores de máxima verossimilhança, através do seguinte resultado:

$$\frac{\hat{\theta} - \theta}{\sqrt{(n I_F(\theta))^{-1}}} \sim^a N(0,1). \quad (4.1)$$

Em que a quantidade $I_F(\theta)$ é a informação de Fisher de θ , que é dada por

$$I_F(\theta) = E \left[\left(\frac{\partial \log f(Y|\theta)}{\partial \theta} \right)^2 \right].$$

Como, $I_F(\theta)$ pode depender de θ , que não é conhecido, essa quantidade pode ser substituída pelo seu estimador de máxima verossimilhança $I_F(\hat{\theta})$. Assim tem-se,

$$Q(\mathbf{Y}, \theta) = \frac{\hat{\theta} - \theta}{\sqrt{((n I_F(\hat{\theta}))^{-1})}} \sim^a N(0,1). \quad (4.2)$$

De modo que $Q(\mathbf{Y}, \theta)$ é uma quantidade pivotal com distribuição aproximadamente igual à distribuição $N(0,1)$ (BOLFARINE; SANDOVAL, 2001).

Bolfarine e Sandoval (2001) destacam que é possível construir teste de hipóteses a partir dos estimadores de máxima verossimilhança. Os autores definem que uma hipótese estatística é qualquer afirmação acerca da distribuição de probabilidade de uma ou mais variáveis aleatórias.

Diante disso, apresenta-se alguns conceitos essenciais para a obtenção dos estimadores de máxima verossimilhança. Seja Y uma variável aleatória populacional, com função densidade de probabilidade denotada por $f(y, \theta)$, em que $\theta = (\theta_1, \dots, \theta_r)^\top$ é o vetor de parâmetros desconhecido, sendo $\theta \in \Theta$, em que $\Theta \in \mathbb{R}$ é o espaço paramétrico (BOLFARINE; SANDOVAL, 2001).

Seja Y_1, \dots, Y_n uma amostra aleatória de tamanho n com função densidade ou função de probabilidade $f(y, \theta)$. A função de verossimilhança $L(\theta)$ para o vetor de parâmetros é dada por

$$L(\theta; y_1, \dots, y_n) = f(y_1; \theta) \times \dots \times f(y_n; \theta) = \prod_{i=1}^n f(y_i; \theta),$$

em que y_1, \dots, y_n são os valores observados da amostra aleatória. Os estimadores de máxima verossimilhança de θ são os valores $\hat{\theta} \in \Theta$ que maximiza a função de verossimilhança $L(\theta; y_1, \dots, y_n)$. Vale ressaltar que, pelo fato da função logaritmo ser monótona crescente, procurar os valores de θ que maximiza $L(\theta)$, equivale a procurar os valores de θ que maximiza o logaritmo de $L(\theta)$, isto é, $\ell(\theta, y_1, \dots, y_n) = \log[L(\hat{\theta}; y_1, \dots, y_n)] = \sum_{i=1}^n \log[f(y_i; \theta)]$. Geralmente é mais conveniente trabalhar com o logaritmo da função de verossimilhança.

De acordo com Bolfarine e Sandoval (2001), em alguns casos, principalmente quando a verossimilhança está associada a modelos mais complexos, a função de verossimilhança não apresenta solução analítica explícita. Sendo assim, os estimadores de máxima verossimilhança podem ser obtidos via métodos numéricos.

Para isso, define-se a função escore da seguinte maneira:

$$U(\theta) = \frac{\partial \ell(\theta)}{\partial \theta},$$

em que para obter os estimadores de máxima verossimilhança de θ , precisa-se obter $U(\hat{\theta}) = 0$.

O método de escore de Fisher é um procedimento iterativo baseado na expansão do vetor escore em série de Taylor até primeira ordem, em torno de um valor inicial θ_0 . Expandindo $U(\hat{\theta})$, tem-se que

$$0 \cong U(\hat{\theta}) = U(\theta_0) + (\hat{\theta} - \theta_0)U'(\theta_0).$$

Desta forma, chega-se ao seguinte processo iterativo

$$\theta^{(i+1)} = \theta^{(i)} - \frac{U(\theta^{(i)})}{U'(\theta^{(i)})}. \quad (4.3)$$

O estimador de máxima verossimilhança é obtido quando a distância entre $\theta^{(i+1)}$ e $\theta^{(i)}$, isto é, $|\theta^{(i+1)} - \theta^{(i)}|$ é menor que uma tolerância especificada. Esse método é conhecido como método escore (BOLFARINE; SANDOVAL, 2001).

O método escore de Fisher apresenta algumas propriedades importantes, que são dadas a seguir: (i) $\mathbb{E}(U) = 0$ e (ii) $\text{Cov}(U) = \mathbb{E}\left(-\frac{\partial U^\top}{\partial \theta}\right) = \mathbb{E}\left(-\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta^\top}\right) = J$. Em que J é a matriz de informação esperada. Já a matriz de informação observada é dada por:

$$\mathcal{I} = -\frac{\partial U^\top}{\partial \theta} = -\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta^\top}.$$

A seguir, são apresentadas as quantidades necessárias para a obtenção dos estimadores de máxima verossimilhança dos parâmetros associados à família proposta. A função de log-verossimilhança da família de distribuições \mathcal{NWEU} baseado no vetor de parâmetros $\theta = (\mu, \boldsymbol{\xi}^\top)^\top$ é dada por

$$\begin{aligned} \ell \equiv \ell(\theta|y_1, \dots, y_n) = & -2 \sum_{i=1}^n \log(1 - y_i) + n \log [\log(1 - \tau)^{-1}] \\ & - n \log \left[\text{H} \left(\frac{\mu}{1 - \mu}; \boldsymbol{\xi} \right) \right] + \sum_{i=1}^n \log \left[\text{h} \left(\frac{y_i}{1 - y_i}; \boldsymbol{\xi} \right) \right] - \frac{\log(1 - \tau)^{-1}}{\text{H} \left(\frac{\mu}{1 - \mu}; \boldsymbol{\xi} \right)} \sum_{i=1}^n \text{H} \left(\frac{y_i}{1 - y_i}; \boldsymbol{\xi} \right). \end{aligned} \quad (4.4)$$

Os componentes do vetor escore $U_{\theta} = (U_{\mu}, U_{\xi})^T$ são dados por

$$U_{\mu} = \frac{h\left(\frac{\mu}{1-\mu}; \xi\right)}{(1-\mu)^2 H\left(\frac{\mu}{1-\mu}; \xi\right)} \left[-n + \frac{\log(1-\tau)^{-1}}{H\left(\frac{\mu}{1-\mu}; \xi\right)} \sum_{i=1}^n H\left(\frac{y_i}{1-y_i}; \xi\right) \right]$$

e

$$U_{\xi} = \frac{\partial \ell}{\partial \xi} = -n \frac{h\left(\frac{\mu}{1-\mu}; \xi\right)}{H\left(\frac{\mu}{1-\mu}; \xi\right)} + \sum_{i=1}^n \frac{1}{h\left(\frac{y_i}{1-y_i}; \xi\right)} \frac{\partial \ell}{\partial \xi} h\left(\frac{y_i}{1-y_i}; \xi\right) - \frac{\log(1-\tau)^{-1}}{H\left(\frac{\mu}{1-\mu}; \xi\right)} \sum_{i=1}^n \frac{\partial \ell}{\partial \xi} h\left(\frac{y_i}{1-y_i}; \xi\right) + \frac{\log(1-\tau)^{-1} h\left(\frac{\mu}{1-\mu}; \xi\right)}{\left[H\left(\frac{\mu}{1-\mu}; \xi\right)\right]^2} \sum_{i=1}^n \frac{\partial \ell}{\partial \xi} H\left(\frac{y_i}{1-y_i}; \xi\right),$$

respectivamente. Deve-se salientar que, τ é um valor fixo que representa a ordem do quantil da distribuição da família \mathcal{NWEU} a ser modelado.

Para valores fixos de ξ , é possível obter uma forma fechada para o estimador de máxima verossimilhança de μ . Fazendo $U_{\mu} = 0$ e isolando μ obtemos a seguinte expressão

$$\hat{\mu}(\hat{\xi}) = \frac{H^{-1}\left[\frac{1}{n} \log\left(\frac{1}{1-\tau}\right) \sum_{i=1}^n \frac{y_i}{1-y_i}; \xi\right]}{1 + H^{-1}\left[\frac{1}{n} \log\left(\frac{1}{1-\tau}\right) \sum_{i=1}^n \frac{y_i}{1-y_i}; \xi\right]}.$$

Portanto, é possível obter o estimador de máxima verossimilhança de μ em forma fechada quando $\xi = \emptyset$. Caso contrário, para obter os estimadores de máxima verossimilhança dos parâmetros μ e ξ é necessário o uso de algum procedimento iterativo tal como Newton-Raphson para estimar os parâmetros do modelo.

4.1 Estimado de máxima verossimilhança para a distribuição Rayleigh unitária

Seja y_1, \dots, y_n uma amostra aleatória observada de tamanho n da distribuição $\mathcal{NRU}(\mu)$.

Sua função de log-verossimilhança é dada por

$$\begin{aligned} \ell \equiv \ell(\mu; y) &= n \log(2) + \sum_{i=1}^n \log(y_i) + n \log((1-\mu)^2) + n \log\left[\log\left(\frac{1}{1-\tau}\right)\right] - 2n \log(\mu) \\ &- 3 \sum_{i=1}^n \log(1-y_i) - \frac{\log\left(\frac{1}{1-\tau}\right) (1-\mu)^2}{\mu^2} \sum_{i=1}^n \frac{y_i^2}{(1-y_i)^2}. \end{aligned} \quad (4.5)$$

O componente da função escore $U_{\theta} = (U_{\mu})^T$ é dada por

$$U_{\mu} = \frac{\partial \ell}{\partial \mu} = \frac{2}{\mu(1-\mu)} \left[-n + \frac{(1-\mu)^2 \log\left(\frac{1}{1-\tau}\right)}{\mu^2} \sum_{i=1}^n \frac{y_i^2}{(1-y_i)^2} \right]. \quad (4.6)$$

A partir da equação (4.6), pode-se encontrar uma expressão analítica de forma fechada para o estimador de máxima verossimilhança de μ . O estimador de máxima verossimilhança de μ , $\hat{\mu}$, é dado por

$$\hat{\mu} = \frac{\sqrt{\log(1-\tau)^{-1} \sum_{i=1}^n \frac{y_i^2}{(1-y_i)^2}}}{\sqrt{n} + \sqrt{\log(1-\tau)^{-1} \sum_{i=1}^n \frac{y_i^2}{(1-y_i)^2}}}. \quad (4.7)$$

A informação de Fisher observada é dada por

$$\frac{\partial^2 \ell}{\partial \mu^2} = \frac{2n(1-2\mu)}{(\mu-1)^2 \mu^2} + \frac{2(2\mu-3)(\mu-1)^2}{\mu^4(\mu-1)^2} \log(1-\tau)^{-1} \sum_{i=1}^n \frac{y_i^2}{(1-y_i)^2}.$$

4.2 Estimador de máxima verossimilhança para a distribuição Gompertz unitária

Seja y_1, \dots, y_n uma amostra aleatória de tamanho n da distribuição $\mathcal{NGOU}(\beta, \mu)$ com vetor de parâmetros $\theta = (\beta, \mu)^\top$. A função de log-verossimilhança de θ é dada por

$$\begin{aligned} \ell \equiv \ell(\theta; y) = & -n \log(\beta) + n \log \left[\log \left(\frac{1}{1-\tau} \right) \right] - n \frac{\beta \mu}{1-\mu} - \\ & \sum_{i=1}^n \log \left[(1-y_i)^2 \exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - (1-y_i)^2 \right] - \frac{\log \left[\log \left(\frac{1}{1-\tau} \right) \right]}{\exp \left\{ \beta \frac{\mu}{1-\mu} \right\} - 1} \\ & \sum_{i=1}^n \exp \left\{ \beta \frac{y_i}{1-y_i} \right\} + n \frac{\log \left[\log \left(\frac{1}{1-\tau} \right) \right]}{\exp \left\{ \beta \frac{\mu}{1-\mu} \right\} - 1}. \end{aligned} \quad (4.8)$$

Os componentes do vetor escore $U_\theta = (U_\beta, U_\mu)^\top$ são dados por

$$U_\beta = \frac{\partial \ell}{\partial \beta} = -n \frac{\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - 1}{\beta} + n \frac{\mu}{1-\mu} + \frac{\mu}{1-\mu} \log \left(\frac{1}{1-\tau} \right) \sum_{i=1}^n \exp \left\{ \frac{\beta y_i}{1-y_i} \right\}$$

$$\log \left(\frac{1}{1-\tau} \right) \sum_{i=1}^n \frac{y_i}{1-y_i} \exp \left\{ \frac{\beta y_i}{1-y_i} \right\},$$

$$\begin{aligned} U_\mu = \frac{\partial \ell}{\partial \mu} = & -n - \frac{\exp \left\{ \frac{\beta \mu}{1-\mu} \right\}}{\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - 1} + \frac{\log \left[\log \left(\frac{1}{1-\tau} \right) \right]}{\left[\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - 1 \right]^2} \sum_{i=1}^n \exp \left\{ \frac{\beta y_i}{1-y_i} \right\} \\ & - n \log \left[\log \left(\frac{1}{1-\tau} \right) \right]. \end{aligned}$$

Os estimadores de máxima verossimilhança não possuem forma fechada, sendo necessário o uso de algoritmos de otimização não-linear para a maximização da função de log-verossimilhança.

Para o modelo \mathcal{NGO} , os elementos da matriz de informação de Fisher observada são dados por

$$\begin{aligned}
U_{\mu^2} &= \frac{\partial^2 \ell}{\partial \mu^2} = - \frac{\beta \log [\log (1 - \tau)^{-1}] \exp \left\{ \frac{\beta \mu}{1 - \mu} \right\}}{(\mu - 1)^4 \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)^2} [n(-2 + \beta + 2\mu) - (2 + \beta - 2\mu) \\
&\quad \sum_{i=1}^n \exp \left\{ \frac{y_i}{1 - y_i} \right\}] - \frac{2n\beta}{(\mu - 1)^2} - \frac{2\beta^2 \log [\log (1 - \tau)^{-1}] \exp \left\{ \frac{\beta \mu}{1 - \mu} \right\}}{(\mu - 1)^4 \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)^3} \\
&\quad \left(n - \sum_{i=1}^n \exp \left\{ \frac{y_i}{1 - y_i} \right\} \right) + \frac{2n\beta \exp \left\{ \frac{\beta \mu}{1 - \mu} \right\}}{(\mu - 1)^3 \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)} - \frac{2n\mu\beta}{(\mu - 1)^3} \\
&\quad + \frac{n\beta \exp \left\{ \frac{2\beta \mu}{1 - \mu} \right\}}{(\mu - 1)^2 \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)^2} \left[\frac{\beta}{1 - \mu} + \frac{\mu\beta}{(1 - \mu)^2} \right] - \frac{n\beta \exp \left\{ \frac{\beta \mu}{1 - \mu} \right\}}{(\mu - 1)^2 \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)} \\
&\quad \left[\frac{\beta}{1 - \mu} + \frac{\mu\beta}{(1 - \mu)^2} \right],
\end{aligned}$$

$$\begin{aligned}
U_{\beta^2} &= \frac{\partial^2 \ell}{\partial \beta^2} = \frac{-2\mu^2 \log [\log (1 - \tau)^{-1}]}{(\mu - 1)^2 \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)^3} \left(n - \sum_{i=1}^n \exp \left\{ \frac{y_i}{1 - y_i} \right\} \right) \\
&\quad - \frac{\mu^2 \log [\log (1 - \tau)^{-1}]}{(\mu - 1)^2 \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)^2} \left(n - \sum_{i=1}^n \exp \left\{ \frac{y_i}{1 - y_i} \right\} \right) + \frac{n}{\beta^2} \\
&\quad - \frac{n\mu^2 \exp \left\{ \frac{2\beta \mu}{1 - \mu} \right\}}{\left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)^2 (1 - \mu)(\mu - 1)} - \frac{\log [\log (1 - \tau)^{-1}]}{\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1} \\
&\quad \sum_{i=1}^n \frac{y_i^2}{(1 - y_i)^2} \exp \left\{ \frac{y_i}{1 - y_i} \right\} + \frac{n\mu^2 \exp \left\{ \frac{\beta \mu}{1 - \mu} \right\}}{(1 - \mu)(\mu - 1) \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)} \\
&\quad + \frac{2\mu \log [\log (1 - \tau)^{-1}] \exp \left\{ \frac{\beta \mu}{1 - \mu} \right\}}{(1 - \mu) \left(\exp \left\{ \frac{\beta \mu}{1 - \mu} \right\} - 1 \right)^2} \sum_{i=1}^n \frac{y_i}{1 - y_i} \exp \left\{ \frac{y_i}{1 - y_i} \right\}
\end{aligned}$$

e

$$\begin{aligned}
U_{\mu\beta} = \frac{\partial^2 \ell}{\partial \mu \partial \beta} = & - \frac{\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} \log [\log (1-\tau)^{-1}]}{(1-\mu)^2 \left(\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - 1 \right)^2} \left(n - \sum_{i=1}^n \exp \left\{ \frac{y_i}{1-y_i} \right\} - \beta \sum_{i=1}^n \frac{y_i}{1-y_i} \right. \\
& \left. \exp \left\{ \frac{y_i}{1-y_i} \right\} \right) + \frac{2\mu\beta \exp \left\{ \frac{\beta \mu}{\mu-1} \right\} \log [\log (1-\tau)^{-1}]}{(\mu-1)^3 \left(\exp \left\{ \frac{\beta \mu}{\mu-1} \right\} - 1 \right)^3} \left(n - \sum_{i=1}^n \exp \left\{ \frac{y_i}{1-y_i} \right\} \right) \\
& + \frac{\mu\beta \exp \left\{ \frac{\beta \mu}{1-\mu} \right\} \log [\log (1-\tau)^{-1}]}{(\mu-1)^3 \left(\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - 1 \right)^2} \left(n - \sum_{i=1}^n \exp \left\{ \frac{y_i}{1-y_i} \right\} \right) - \frac{n}{1-\mu} - \frac{n\mu}{(1-\mu)^2} \\
& - \frac{n \exp \left\{ \frac{\beta \mu}{1-\mu} \right\}}{(\mu-1)^2 \left(\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - 1 \right)} + \frac{n\mu\beta \exp \left\{ \frac{2\beta \mu}{1-\mu} \right\}}{(1-\mu)(\mu-1)^2 \left(\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - 1 \right)^2} \\
& - \frac{n\mu\beta \exp \left\{ \frac{\beta \mu}{1-\mu} \right\}}{(1-\mu)(\mu-1)^2 \left(\exp \left\{ \frac{\beta \mu}{1-\mu} \right\} - 1 \right)}.
\end{aligned}$$

5 ESTUDO DE SIMULAÇÃO

Nesta seção, é apresentado um estudo de simulações Monte Carlo para avaliar as propriedades assintóticas dos estimadores de máxima verossimilhança nos submodelos \mathcal{NRU} e \mathcal{NGOU} , que pertencem ao modelo \mathcal{NWEU} .

Todas as rotinas de simulações para os dois submodelos foram implementadas usando a linguagem \mathbb{R} (R Core Team, 2018), em que foi utilizada uma rotina de otimização não-linear quase-Newton da BFGS. Todos os resultados são baseados em 10.000 réplicas de Monte Carlo, com tamanhos amostrais de $n \in (10, 25, 50, 75, 100)$. Além disso, é fixado o valor de $\tau = 0,5$ para ambos os casos.

Para avaliar o desempenho dos estimadores de máxima verossimilhança, algumas medidas podem ser necessária. Baseadas na réplicas de Monte Carlo dos estimadores de máxima verossimilhança, aqui, calcula-se as seguintes medidas

- A Média do estimador de máxima verossimilhança, dada por

$$\bar{\hat{\theta}} = \frac{1}{B} \sum_{i=1}^B \hat{\theta}_i,$$

em que $\hat{\theta}_i$ é o valor estimado do parâmetro na i -ésima réplica e B é o número de réplicas de Monte Carlo.

- Variância do estimador de máxima verossimilhança

$$\text{Var}(\hat{\theta}) = \frac{1}{B-1} \sum_i^B [\hat{\theta}_i - \bar{\hat{\theta}}]^2.$$

- Viés relativo percentual (VR%).

$$VR(\hat{\theta}) = \left(\frac{\bar{\hat{\theta}} - \theta}{\theta} \right) * 100.$$

- Erro quadrático médio. O EQM desse estimador é dado da seguinte forma:

$$\text{EQM}(\hat{\theta}) = \text{var}(\hat{\theta}) + [B(\hat{\theta})]^2.$$

A seguir, apresenta-se os resultados numéricos das simulações Monte Carlo para as duas distribuições abordadas.

5.1 Resultados numéricos para a distribuição \mathcal{NRU}

Como foi possível encontrar uma forma analítica fechada para o estimador de máxima verossimilhança de μ da distribuição \mathcal{NRU} que é dado pela equação (4.7). Realizou-se, as simulações de Monte Carlo a partir dessa equação. Sendo assim, considerou-se cinco cenários com valores diferentes para o parâmetro μ da distribuição \mathcal{NRU} .

Na Tabela 5.1, apresenta-se as estimativas do parâmetro μ , bem como a sua variância, VR(%), EQM e a taxa de cobertura do intervalo de confiança aproximado. Foi utilizada uma taxa de cobertura nominal de 0,95. Analisando a Tabela 5.1, verifica-se que o VR(%) diminui a medida que o tamanho amostral aumenta. Esse comportamento já era esperado, devido as propriedades assintóticas dos estimadores de máxima verossimilhança.

O cenário 1, $\mu = 0,15$, apresentou o menor VR(%) entre todos os cenários abordados. Salienta-se, que não apenas o VR(%) foi o menor, mas também a variância e o EQM apresentaram os menores valores. Com exceção da taxa de cobertura, que para os cinco cenários apresenta um valor percentual próximo ao valor da taxa de cobertura nominal.

No cenário 3, para $n = 10$ pode-se verificar que os valores foram mais altos para o VR(%), a variância e o EQM. Em relação a taxa de cobertura temos um percentual mais próximo do nominal. Destaca-se, que a medida que n aumenta esses valores diminuem relativamente. Cabe ressaltar que para $n = 25$ já foi possível verificar que o VR(%) caiu de 7,5% para 0,2% e o EQM diminuiu de 0,0015 para 0,0005.

Outro ponto relevante a se destacar, que para o cenário 4, onde $\mu = 0,7$. Obteve-se os maiores valores do VR(%), da variância e do EQM quando $n = 100$. Apesar disso, esses valores podem ser considerados muito baixos uma vez que para esse tamanho de amostra o maior VR(%) está abaixo de 1%.

Na Figura 5.1(a) ilustra o comportamento do VR(%) para todos os tamanhos de amostra. Conforme exposto anteriormente, essa medida tende a aproximar-se de zero para todos os cenários. Portanto, a partir dos resultados do estudo de simulação realizado, pode-se concluir que o estimador de máxima verossimilhança para o τ -ésimo quantil da distribuição \mathcal{NRU} é aproximadamente não viesado. O Apêndice A.1 apresenta os resultados para τ diferente de 0,5. A Figura 5.1(b) apresenta o gráfico do EQM dos cinco cenários, a qual permite observar o decaimento à medida que o tamanho amostral aumenta, sendo muito próximo de zero para tamanhos amostrais maiores que $n = 25$.

Tabela 5.1 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para o parâmetro da distribuição \mathcal{NRU} , para diferentes cenários e tamanhos amostrais.

Cenário	μ	n	MÉDIA($\hat{\mu}$)	VARIÂNCIA($\hat{\mu}$)	VR %($\hat{\mu}$)	EQM($\hat{\mu}$)	TAXA($\hat{\mu}$)
1	0.15	10	0.1475	0.0003937	-0.2432	0.0003997	0.949
		25	0.1492	0.0001724	-0.0741	0.0001730	0.943
		50	0.1493	0.0000842	-0.0617	0.0000846	0.959
		75	0.1494	0.0000569	-0.0505	0.0000572	0.948
		100	0.1497	0.0000392	-0.0286	0.0000393	0.958
2	0.25	10	0.2463	0.0008927	-0.3669	0.0009061	0.9470
		25	0.2484	0.0003442	-0.1584	0.0003467	0.9579
		50	0.2492	0.0001743	-0.0773	0.0001749	0.9533
		75	0.2495	0.0001156	-0.0446	0.0001158	0.9537
		100	0.2497	0.00008894	-0.0287	0.00008902	0.9518
3	0.4	10	0.3943	0.0014	-0.5635	0.0015	0.9552
		25	0.3976	0.0005	-0.2385	0.0005	0.9594
		50	0.3988	0.0002	-0.1170	0.0002	0.9538
		75	0.3993	0.0002	-0.0689	0.0001	0.9536
		100	0.3995	0.0001	-0.0459	0.0001	0.9508
4	0.7	10	0.6933	0.0012	-0.6654	0.0012	0.9518
		25	0.6972	0.0004	-0.2725	0.0004	0.9535
		50	0.6986	0.0002	-0.1341	0.0002	0.9506
		75	0.6991	0.0001	-0.0812	0.0001	0.9521
		100	0.6994	0.0001	-0.0562	0.0001	0.9469
5	0.9	10	0.8966	0.0002382	-0.3384	0.0002496	0.9405
		25	0.8986	0.0000843	-0.1356	0.0000862	0.9489
		50	0.8993	0.0000413	-0.0666	0.0000417	0.9497
		75	0.8995	0.0000271	-0.0408	0.0000272	0.9515
		100	0.8997	0.0000208	-0.0287	0.0000208	0.9454

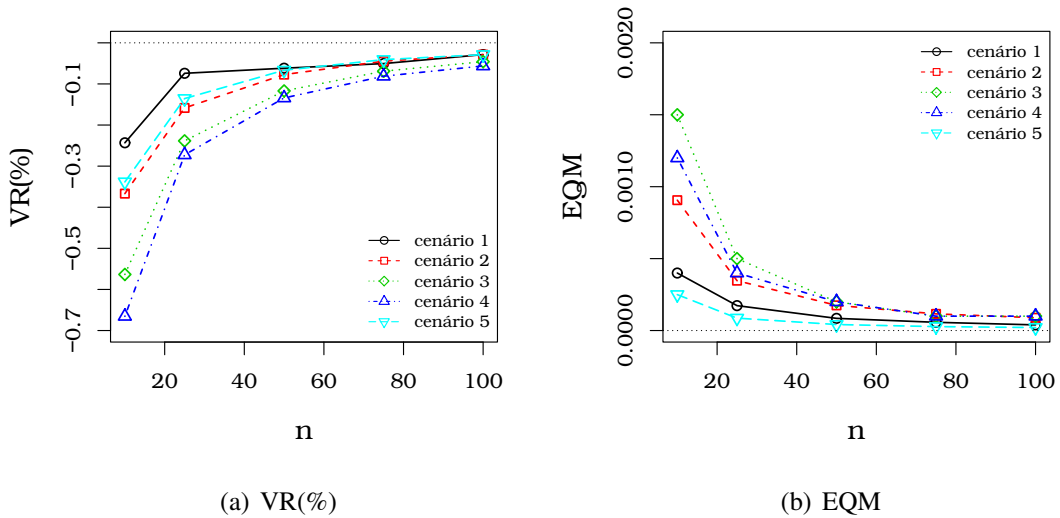


Figura 5.1 – VR(%) e EQM para o estimador de máxima verossimilhança da mediana da distribuição \mathcal{NRU} .

5.2 Resultados numéricos para a distribuição \mathcal{NGOU}

Para a distribuição \mathcal{NGOU} , considerou-se seis cenários com diferentes combinações de valores para os parâmetros μ e β . Buscou-se, combinar valores de $\mu \in (0,1)$ com valores de β entre $(0,1)$ e maiores que 1.

A Tabela 5.2 apresenta os resultados da avaliação numérica dos estimadores de máxima verossimilhança dos parâmetros na distribuição \mathcal{NGOU} . Ao analisar o tamanho amostra $n = 10$, pode-se observar valores elevados para as estimativas das medidas de precisão e variabilidade relacionadas a $\hat{\beta}$. Em contrapartida, as estimativas para o estimador de máxima verossimilhança de μ apresentaram melhores resultados para essas medidas. Em relação ao VR(%), as estimativas para o parâmetro μ apresentam valores inferiores a 0,5% em todos os cenários. Como esperado, devido às propriedades assintóticas do estimador de máxima verossimilhança; o VR(%), a variância e o EQM dos estimadores de ambos os parâmetros diminuem à medida que o tamanho amostral aumenta.

Na comparação entre cenários, evidencia-se que o cenário 4 apresenta os maiores valores para a variância, VR(%) e EQM de $\hat{\beta}$. Em contrapartida, esse foi o cenário com menor VR(%) para o estimador de máxima verossimilhança de μ . A Figura 5.2(a) apresenta o gráfico do VR(%) total, definido como a soma dos valores absolutos dos VR(%) de ambos os parâmetros. Como observado na Tabela 5.2, a medida que o tamanho amostral aumento o VR(%) total

aproxima-se de zero, isso ocorre para os seis cenários.

Ainda, a Figura 5.2(b) o gráfico do EQM dos seis cenários. Verifica-se que a medida que o tamanho amostral aumenta o EQM total aproxima-se de zero. A partir da figura, pode-se concluir que os estimadores de máxima verossimilhança da distribuição \mathcal{NGOU} apresentam estimativas acuradas para tamanhos amostrais acima de 25, uma vez que o EQM total foi inferior a um em todos os cenários.

Com relação aos intervalos assintóticos, considerou-se uma taxa de cobertura nominal de 0,95. Observa-se que a taxa de cobertura de $\hat{\beta}$ fica próximo ao valor nominal já quando $n = 10$, alcançando resultados melhores quando comparada à taxa de cobertura para $\hat{\mu}$.

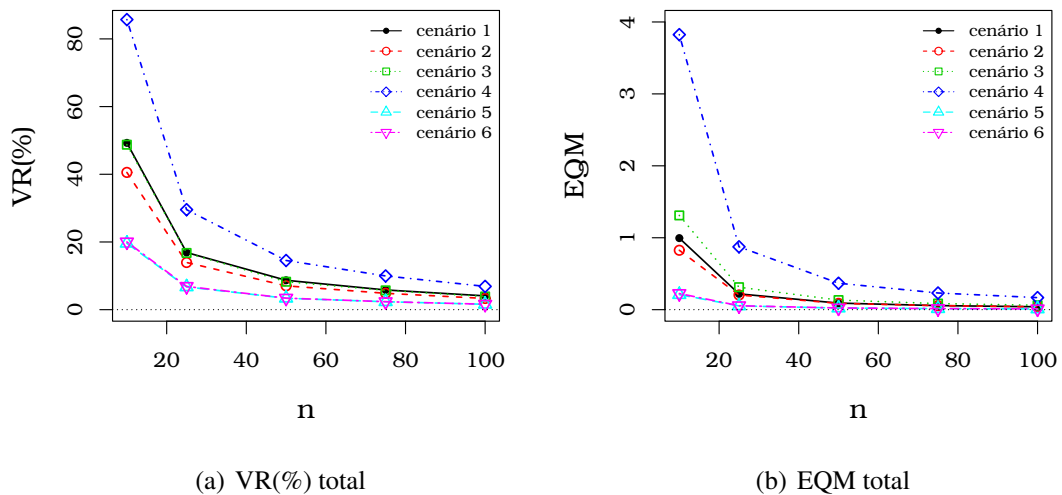


Figura 5.2 – VR(%) total e EQM total para os estimadores de máxima verossimilhança da distribuição \mathcal{NGOU} .

Tabela 5.2 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para os parâmetros da distribuição \mathcal{NUGO} , para diferentes cenários e tamanhos amostrais.

Cenário	μ	β	n	MÉDIA($\hat{\mu}$)		VARIÂNCIA($\hat{\mu}$)		VR %($\hat{\mu}$)		EQM($\hat{\mu}$)		TAXA($\hat{\mu}$)	
				$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$
1	0.4	0.8	10	0.4038	1.2889	0.0046	0.7510	0.3819	48.8994	0.0047	0.9902	0.8852	0.9311
			25	0.4014	0.9674	0.0019	0.1895	0.1444	16.7461	0.0019	0.2176	0.9268	0.9433
			50	0.4011	0.8856	0.0010	0.0827	0.1146	8.5658	0.0010	0.0901	0.9377	0.9474
			75	0.4009	0.8571	0.0006	0.0526	0.0995	5.7184	0.0006	0.0559	0.9410	0.9489
			100	0.4006	0.8400	0.0005	0.0384	0.0653	4.0097	0.0005	0.0400	0.9423	0.9462
2	0.5	1.5	10	0.4971	1.9084	0.0025	0.6566	-0.2869	40.8414	0.0025	0.8234	0.8941	0.9427
			25	0.4985	1.6408	0.0010	0.1832	-0.1444	14.0812	0.0010	0.2031	0.9283	0.9480
			50	0.4994	1.5709	0.0004	0.0824	-0.0544	7.0981	0.0004	0.0874	0.9406	0.9499
			75	0.4998	1.5485	0.0003	0.0534	-0.0195	4.8538	0.0003	0.0558	0.9412	0.9498
			100	0.4997	1.5333	0.0002	0.0393	-0.0226	3.3385	0.0002	0.0404	0.9444	0.9476
3	0.6	2.8	10	0.5980	3.2892	0.000489	1.0709	-0.1989	48.9285	0.000493	1.3103	0.9042	0.9526
			25	0.5991	2.9682	0.000184	0.2841	-0.0875	16.8275	0.000185	0.3124	0.9339	0.9508
			50	0.5996	2.8829	0.000892	0.1265	-0.0388	8.2924	0.000894	0.1333	0.9430	0.9513
			75	0.5998	2.8571	0.0000606	0.0816	-0.0182	5.7180	0.0000606	0.0849	0.9437	0.9510
			100	0.5998	2.8390	0.0000459	0.0601	-0.0167	3.9059	0.0000459	0.0616	0.9456	0.9475
4	0.7	5	10	0.6996	5.8574	0.0000463	3.0886	-0.0362	85.7439	0.0000465	3.8238	0.9016	0.9499
			25	0.6998	5.2956	0.0000179	0.7867	-0.0164	29.5603	0.0000179	0.8741	0.9350	0.9507
			50	0.6999	5.1454	0.00000881	0.3466	-0.0066	14.5412	0.00000880	0.3677	0.9436	0.9504
			75	0.6999	5.0996	0.00000597	0.2220	-0.0018	9.9651	0.00000597	0.2319	0.9435	0.9505
			100	0.6999	5.0688	0.00000454	0.1630	-0.0022	6.8801	0.00000454	0.1677	0.9449	0.9481
5	0.7	0.9	10	0.6961	1.1001	0.0012	0.1686	-0.3809	20.0002	0.0012	0.2086	0.8990	0.9485
			25	0.6983	0.9690	0.0004	0.0471	-0.1693	6.9054	0.0004	0.0519	0.9297	0.9507
			50	0.6992	0.9345	0.0002	0.0212	-0.0738	3.4534	0.0002	0.0224	0.9420	0.9513
			75	0.6996	0.9238	0.0001	0.0138	-0.0374	2.3819	0.0001	0.0143	0.9415	0.9507
			100	0.6996	0.9162	0.0001	0.0101	-0.0337	1.6285	0.0001	0.0104	0.9446	0.9486
6	0.75	1.1	10	0.7475	1.3025	0.000489	0.1828	-0.2473	20.2578	0.000495	0.2239	0.9037	0.9520
			25	0.7489	1.1697	0.000180	0.0496	-0.1068	6.9781	0.000182	0.0545	0.9330	0.9515
			50	0.7495	1.1345	0.0000865	0.0222	-0.0477	3.4546	0.0000867	0.0234	0.9432	0.9516
			75	0.7497	1.1238	0.0000587	0.0144	-0.0244	2.3864	0.0000588	0.0149	0.9429	0.9515
			100	0.7497	1.1163	0.0000443	0.0106	-0.0214	1.6317	0.0000444	0.0108	0.9452	0.9470

6 APLICAÇÃO EM DADOS ELEITORAIS

Nessa seção, apresenta-se uma aplicação a dados eleitorais. Desta forma, realizou-se um estudo de caso referente ao segundo turno das eleições presidenciais no ano de 2018, considerando a região nordeste. Foi realizado o ajuste dos seis modelos propostos no Capítulo 3. Também foram considerados cinco distribuições que são clássicas na literatura acerca de taxas e proporções, as quais são descritas no Capítulo 2.

6.1 Panorama Geral das Eleições

Nas eleições presidenciais brasileiras no ano de 2018 os números de votos são significativos. Totalizaram 147 milhões de eleitores. Nesse caso, está sendo considerado apenas o segundo turno das eleições, em que foram computados 104.838.753 milhões de votos válidos. Os candidatos que concorreram a presidência da república no segundo turno das eleições no ano de 2018, foram: (i) Jair Messias Bolsonaro, candidato pelo Partido Social Liberal - PSL. Que foi o candidato que conseguiu se eleger com 57.797.466 milhões de votos, totalizando 55,13% dos votos válidos; (ii) Fernando Haddad, candidato pelo Partido dos Trabalhadores - PT, que obteve 47.040.859 milhões de votos, totalizando 44,87% dos votos válidos (ELEITORAL-TSE, 2018).

Algo que chamou a atenção para o resultado das eleições presidenciais no ano de 2018, foi o percentual de votos brancos e nulos. Que segundo Scorsim (2018), aproximadamente um terço do total de votantes, que equivale a 42.1 milhões de eleitores (30,87% dos votos), não escolheu nem Jair Messias Bolsonaro nem Fernando Haddad para presidente. Ainda segundo o autor, desses 30,87% dos votos, (i) optaram pela abstenção: 31.371.704 milhões de eleitores (21,30%); (ii) optaram pelo voto branco: 2.486.593 milhões de eleitores (2,14%) e (iii) optaram por anular o voto: 8.608.105 milhões de eleitores (7,43%).

Nesse sentido, a Figura 6.1 apresenta o mapa geral referente aos votos válidos do segundo turno das eleições presidenciais de 2018. No qual é possível verificar como ficou dividido a proporção de votos para os dois candidatos. Os pontos que estão com um tom de azul mais forte referem-se as cidades em que o candidato do PSL obteve uma maior porcentagem dos votos, já o tom de azul mais claro o candidato obteve uma menor porcentagem.

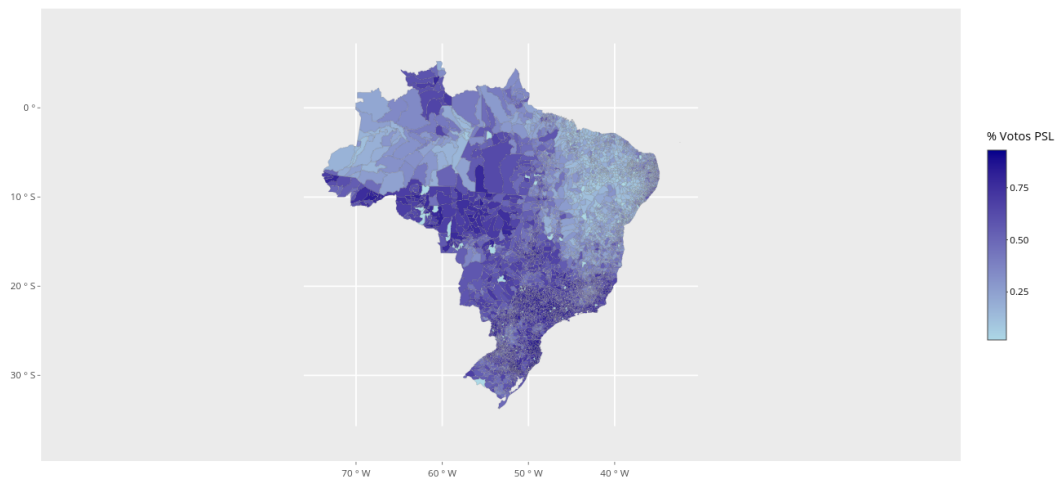


Figura 6.1 – Proporção dos votos válidos do 2º turno das eleições presidências no ano de 2018.

Salienta-se, que a região em que o candidato Bolsonaro obteve a menor porcentagem de votos foi na região nordeste, como pode-se verificar no gráfico. Com isso, essa foi a região em que o candidato Fernando Haddad conquistou as suas maiores porcentagens de votos.

6.2 Aplicação do banco de dados referente a região nordeste

Para a aplicação em dados eleitorais buscou-se encontrar novos modelos probabilísticos que se ajustem a esse tipo de dados de forma mais adequada. Utilizou-se a base de dados fornecida no site do Tribunal Superior Eleitoral (TSE) através do repositório de dados eleitorais disponível ao público na URL www.tse.jus.br. O banco de dados é composto pela proporcionalidade dos votos válidos do segundo turno das eleições presidenciais fixando-se os municípios da região nordeste como distritos eleitorais.

Ressalta-se, que para essa aplicação foram considerados dados da região nordeste do país, que segundo o site do TSE (2010) é o segundo maior sítio eleitoral no Brasil. Devido à relevância e às peculiaridades da região, outros autores já realizaram trabalhos com dados eleitorais tendo como foco o nordeste. Abensur et. al. (2007) realizam um estudo de regressão sobre os impactos do programa Bolsa Família nos resultados das eleições presidenciais no Brasil em 2006, em que o candidato Lula se reelegeu como presidente da república. Nesse trabalho os autores destacam o grande impacto que a região nordeste teve nos resultados eleitorais. Almeida Junior e Souza (2015) desenvolveram um trabalho semelhante analisando se o programa Bolsa Família influenciou para que a presidente Dilma Rousseff ganhasse a eleição presidencial no ano de 2010.

No gráfico da Figura 6.1, foi possível verificar a divisão dos votos válidos por região do país. Com isso, verificou-se que a região nordeste foi a única região em que o candidato Fernando Haddad conseguiu ganhar em todos os estados. Nessa região o candidato do PT registrou 69,3% dos votos válidos. Já o candidato do PSL, ficou com 30,7% dos votos válidos. Ainda, de acordo com o TSE (2018) o candidato Haddad bateu o recorde de votação na região nordeste. Até então, a maior votação que o partido trabalhista tinha obtido foi no ano de 2014, quando 20,17 milhões de nordestinos votaram em Dilma Rousseff no segundo turno. Apesar disso, em percentual de votos válidos, a vantagem petista em relação ao segundo colocado foi a menor desde 2006, com 68% dos votos válidos para Haddad. Isso pode ser explicado pelo crescimento do eleitorado no período e o maior comparecimento às urnas neste ano (ELEITORAL-TSE, 2018).

A região Nordeste é composta por 1795 municípios. A Tabela 6.1 apresenta as estatísticas descritivas dos votos válidos de ambos os candidatos que concorreram a presidência da república no segundo turno.

Tabela 6.1 – Análise descritiva dos dados

Candidato	Média	Mediana	Moda	Variância	Assimetria	Curtose	Mínimo	Máximo
Haddad	0,6929	0,7569	0,8250	0,0338	-1,0329	-0,0046	0,1326	0,9392
Bolsonaro	0,3071	0,2431	0,1750	0,0338	1,0329	-0,0046	0,0608	0,8774

Nesta aplicação foram ajustadas onze distribuições aos banco de dados utilizado, das quais seis pertencem a $\mathcal{NW}\mathcal{EU}$. As distribuições ajustadas em termos da mediana são: (i) $\mathcal{N}\mathcal{E}\mathcal{EU}$ com função densidade de probabilidade definida na equação (3.7); (ii) $\mathcal{N}\mathcal{G}\mathcal{OU}$ com função densidade de probabilidade definida na equação (3.10); (iii) $\mathcal{N}\mathcal{H}\mathcal{LU}$ com função densidade de probabilidade definida na equação (3.13); (iv) $\mathcal{N}\mathcal{RU}$ com função densidade de probabilidade definida na equação (3.16); (v) $\mathcal{N}\mathcal{LU}$ com função densidade de probabilidade definida na equação (3.19) e (vi) $\mathcal{N}\mathcal{N}\mathcal{HU}$ com função densidade de probabilidade definida na equação (3.22).

As outras cinco distribuições que foram utilizadas no ajuste são: (i) $\mathcal{B}\mathcal{E}\mathcal{T}\mathcal{A}$ com função densidade de probabilidade definida na equação (2.2); (ii) \mathcal{KWU} com função densidade de probabilidade definida na equação (2.3); (iii) \mathcal{GU} com função densidade de probabilidade definida na equação (2.4); (iv) \mathcal{BSU} com função densidade de probabilidade definida na equação (2.5) e (v) \mathcal{WU} com função densidade de probabilidade definida na equação (2.6).

Levou-se em consideração as estatísticas corrigidas de Anderson-Darling (A^*) (CHEN;

BALAKRISHNAN, 1995) e Cramér-von Misses (W^*) (DURBIN; KNOTT, 1972) para avaliar a qualidade do ajuste dos diferentes modelos ajustados ao conjunto de dados. Quanto menor o seu valor, melhor é o ajuste aos dados. As análises foram realizadas usando o pacote `AdequacyModel` (MARINHO; BOURGUIGNON; DIAS, 2016) da linguagem de programação R.

A Tabela 6.2 apresenta os resultados do ajuste das distribuições em relação aos votos válidos dos dois candidatos. Pode-se verificar que em relação aos votos válidos do candidato Bolsonaro, a distribuição que obteve o menor valor em relação as estatísticas A^* e W^* foi a distribuição \mathcal{NGOU} que pertence à família \mathcal{NWU} . Além disso, as estimativas dos estimadores de máxima verossimilhança dessa distribuição foram precisas, com EQM pequeno.

Com o intuito de se ter um comparativo com outras distribuições bastante utilizadas na literatura para modelar dados que pertencem ao intervalo unitário, comparamos o modelo \mathcal{NGOU} com as distribuições \mathcal{BETA} e \mathcal{WU} . A Figura 6.2(a) apresenta o histograma dos votos válidos do candidato Bolsonaro, bem como o ajuste das três distribuições supracitadas.

Na mesma Figura, tem-se o diagrama de caixa do candidato Bolsonaro, em que é possível verificar que sua mediana está próximo ao primeiro quartil, pois a proporcionalidade de votos deste candidato nesta região foi menor.

Analisando a Figura 6.2(a), pode-se perceber que a distribuição que obteve um melhor ajuste ao histograma é a \mathcal{NGOU} . A mesma mostrou-se superior as outras duas distribuições, tanto em relação as estatísticas A^* e W^* , quanto ao ajuste do histograma dos dados.

Já analisando a Tabela 6.2, em relação aos votos válidos do candidato Haddad, a distribuição que obteve o menor valor em relação as estatísticas A^* e W^* foi a distribuição \mathcal{NRU} que também pertence à família \mathcal{NWU} . Além disso, podemos verificar que as estimativas dos estimadores de máxima verossimilhança foram precisas, com um EQM bastante baixo.

O gráfico 6.2(b) apresenta o histograma dos votos válidos em relação ao candidato Haddad, bem como o ajuste de três distribuições, analogamente como foi realizado em relação aos votos válidos do candidato Bolsonaro. Comparou-se o modelo \mathcal{NRU} com as distribuições \mathcal{BETA} e \mathcal{WU} . É possível observar que a distribuição \mathcal{NRU} teve um melhor ajuste ao histograma dos votos válidos. Ressalta-se, que por mais que o modelo \mathcal{NRU} tenha um pico maior que o histograma, ele apresenta um melhor ajuste na cauda da distribuição. Isso fica mais evidente analisando os valores das estatísticas A^* e W^* , em que a distribuição \mathcal{NRU} foi superior.

Nesse mesmo gráfico, é apresentado o diagrama de caixa do candidato Haddad, em que é possível verificar que sua mediana está mais próxima do último quartil, fato evidenciado pelo percentual alto de votos que este candidato obteve nessa região.

Tabela 6.2 – Estimativas dos parâmetros obtidos via estimação de máxima verossimilhança dos modelos ajustados à proporção dos votos válidos da região nordeste no segundo turno das eleições presidenciais no ano de 2018 e seus correspondentes EQMs (dados entre parênteses) e os valores das estatísticas A^* e W^* .

Distribuição	Candidatos				Candidatos			
	Bolsonaro				Haddad			
	$\hat{\beta}$	$\hat{\mu}$	W^*	A^*	$\hat{\beta}$	$\hat{\mu}$	W^*	A^*
\mathcal{BETA}	6,382 (0,201)	0,314 (0,004)	8,21	47,67	6,382 (0,201)	0,685 (0,004)	8,20	47,68
\mathcal{BSU}	1,147 (0,015)	0,582 (0,009)	12,81	71,80	0,752 (0,012)	0,325 (0,325)	3,37	20,07
\mathcal{GU}	4,104 (0,131)	0,312 (0,003)	13,90	77,51	1,710 (0,052)	0,489 (0,017)	1,39	9,23
\mathcal{KWU}	0,738 (0,019)	0,393 (0,008)	9,18	53,05	2,283 (0,057)	0,815 (0,005)	7,29	42,70
\mathcal{WU}	2,556 (0,050)	0,268 (0,003)	3,22	19,66	1,364 (0,023)	0,704 (0,005)	10,28	58,89
\mathcal{NEEU}	1,332 (0,045)	0,533 (0,010)	13,51	75,48	1,752 (0,059)	0,398 (0,010)	1,64	10,68
\mathcal{NGOU}	1,605 (0,033)	0,2799 (0,005)	0,41	4,25	0,1420 (0,008)	0,749 (0,004)	11,35	64,24
\mathcal{NHLU}		0,216 (0,033)	5,98	35,38		0,822 (0,072)	10,52	60,48
\mathcal{NLU}	0,485 (0,016)	0,637 (0,065)	7,67	44,66	3,560 (0,119)	0,840 (0,075)	7,44	43,57
\mathcal{NNHU}	0,794 (0,031)	0,284 (0,005)	3,39	20,49	6,259 (1,679)	0,747 (0,004)	11,21	63,73
\mathcal{NRU}		0,907 (0,025)	21,59	116,02		0,781 (0,002)	0,33	3,82

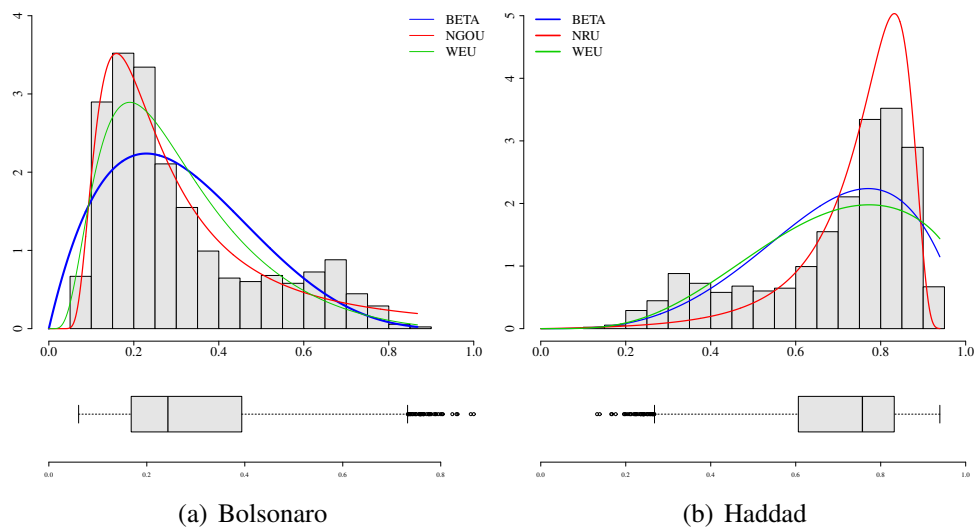


Figura 6.2 – Histograma dos votos válidos e alguns modelos ajustados.

Assim, pode-se concluir que em ambas as Figuras, as duas distribuições propostas obtiveram um melhor ajuste em relação as distribuições $BETA$ e WEU . Desta forma, conclui-se que as distribuições NRU e $NGOU$ são as mais apropriadas para modelar dados eleitorais do segundo turno das eleições presidenciais de 2018. Destaca-se que esses modelos podem ser utilizados para modelar outros tipos de dados que sejam do tipo taxas e proporção, isto é, que pertençam ao intervalo unitário.

7 CONCLUSÃO

Esse trabalho deriva uma nova família de distribuições que modelam dados duplamente limitados, particularmente ao intervalo unitário. Isso foi possível aplicando uma transformação em uma variável aleatória com distribuição Weibull estendida. Um dos propósitos do trabalho é desenvolver alternativas viáveis às distribuições unitárias já existentes. Foram apresentados seis casos especiais dessa nova família, incluindo suas funções quantílicas e gráficos das densidades. Os casos particulares foram desenvolvidos a partir de uma parametrização nos quantis da família. A proposta da reparametrização justifica-se pela melhor interpretabilidade e maior resistência a observações discrepantes.

A seguir foram desenvolvidas expressões explícitas para a obtenção dos estimadores de máxima verossimilhança para a nova família e, particularmente, para as distribuições Rayleigh unitária (\mathcal{NRU}) e Gompertz unitária (\mathcal{NGOU}). Foi encontrada a expressão fechada para o estimador de máxima verossimilhança do parâmetro associado ao modelo \mathcal{NRU} e, através de simulações de Monte Carlo, verificou-se que este é aproximadamente não viesado. Também foram realizados estudos de simulação para o modelo \mathcal{NGOU} , os quais apresentaram resultados adequados em termos de precisão e acurácia, exceto em amostras de tamanho amostral pequeno, como no caso de $n = 10$, que apresentou inferências distorcidas.

Com relação à sua aplicabilidade, os submodelos da família proposta mostraram-se adequados para modelar variáveis contínuas no intervalo $(0,1)$, tais como taxas e proporções. Esse fato foi ilustrado através de um estudo de caso com dados referentes à proporção de votos válidos do segundo turno das eleições presidenciais no ano de 2018, considerando a região nordeste do Brasil. O destaque a essa região justifica-se pelo tamanho do sítio eleitoral, que é o segundo maior do país e abrange 1795 municípios. Além disso, o nordeste apresentou a peculiaridade de ser a única região em que o candidato eleito não alcançou 50% dos votos em nenhum estado.

A partir desse estudo de caso, verificou-se a adequabilidade das distribuições propostas com relação à base de dados considerada. Os modelos \mathcal{NGOU} e \mathcal{NRU} foram superiores no ajuste da proporção de votos dos candidatos Jair Bolsonaro e Fernando Haddad, respectivamente. Essas duas distribuições obtiveram um melhor ajuste a dados eleitorais do que a distribuição \mathcal{BETA} e Kumaraswamy, que são as mais usuais no contexto de dados unitários. Assim, pode-se concluir que os submodelos da família proposta são apropriados para o ajuste da proporção de votos.

REFERÊNCIAS

- ABENSUR, T. C. et al. Impactos do Programa Bolsa Família nos resultados das eleições presidenciais no Brasil em 2006. **Anais do XXXV Encontro Nacional de Economia**, v.51, 2007.
- ALEMÁN, E.; KELLAM, M. The nationalization of presidential elections in the Americas. **Electoral Studies**, v.47, p.125 – 135, 2017.
- ALMEIDA JUNIOR, P. M. d.; SOUZA, T. C. Estimativas de votos da presidente Dilma Rousseff nas eleições presidenciais de 2010 sob o âmbito do Bolsa Família. **Ciência e Natura**, v.37, n.1, p.12–22, 2015.
- BARNDORFF-NIELSEN, O. E.; JORGENSEN, B. Some parametric models on the simplex. **Journal of Multivariate Analysis**, v.39, n.1, p.106–116, 1991.
- BOLFARINE, H.; SANDOVAL, M. C. **Introdução à inferência estatística**. [S.l.]: SBM, 2001. v.2.
- BRASIL. Instruções do TSE: eleições 2010. **Tribunal Superior Eleitoral**, p.409, 2010.
- CHEN, G.; BALAKRISHNAN, N. A general purpose approximate goodness-of-fit test. **Journal of Quality Technology**, v.27, n.2, p.154–161, 1995.
- CORDEIRO, G. M.; ORTEGA, E.; SILVA, G. The Beta extended Weibull family. **Journal of Probability and Statistical Science**, v.10, p.15–40, 2012.
- CORDEIRO, G. M.; SILVA, R. B. The complementary extended Weibull power series class of distributions. **Ciência e Natura**, v.36, p.1–13, 2014.
- DURBIN, J.; KNOTT, M. Components of Cramér–von Mises statistics. I. **Journal of the Royal Statistical Society: Series B (Methodological)**, v.34, n.2, p.290–307, 1972.
- ELEITORAL-TSE, T. S. Repositório de Dados Eleitorais. **Resultados. Apresenta resultados das eleições gerais brasileiras. Disponível em: <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>. Acesso em: maio de, 2018.**
- FERRARI, S. L. P.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. **Journal of Applied Statistics**, v.7, p.799–815, 2004.

- GRASSIA, A. On a family of distributions with argument between 0 and 1 obtained by transformation of the Gamma distribution and derived compound distributions. **Australian Journal of Statistics** ., v.19, p.108–114, 1977.
- GUERRA, R. R.; PEÑA-RAMÍREZ, F. A.; BOURGUIGNON, M. The unit extended Weibull families of distributions and its applications. **Journal of Applied Statistics**, v.submetido, 2020.
- GURVICH, M.; DIBENEDETTO, A.; RANADE, S. A new statistical distribution for characterizing the random strength of brittle materials. **Journal of Materials Science**, v.32, p.2559–2564, 1997.
- KUMARASWAMY, P. A generalized probability density function for double-bounded random processes. **Journal of Hydrology**, v.46, p.79–88, 1980.
- LANCET, T. Are Brazilian elections healthy without a plan for UHC? **The Lancet**, v.392, n.10155, p.1280, 2018.
- LEMONTE, A. J.; BAZÁN, J. L. New class of Johnson distributions and its associated regression model for rates and proportions. **Biometrical Journal**, v.58, n.4, p.727–746, 2016.
- LIMA, F. A. d. **Distribuições de probabilidade no intervalo unitário**. 2018. Tese (Doutorado em Ciência da Computação) — Universidade de São Paulo.
- MARINHO, P. R. D.; BOURGUIGNON, M.; DIAS, C. R. B. **AdequacyModel**: adequacy of probabilistic models and general purpose optimization. [S.l.: s.n.], 2016. R package version 2.0.0.
- MATIAS, A.; JR., V. O mapa da votação presidencial nos municípios. **Valor Econômico**, 2018.
- MAZUCHELI, J. et al. The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. **Journal of Applied Statistics**, p.1–21, 2019.
- MAZUCHELI, J.; MENEZES, A. F. B.; DEY, S. The unit-Birnbaum-Saunders distribution with applications. **Chilean Journal of Statistics**, v.In press, 2018.
- MITNIK, P. A.; BAEK, S. The Kumaraswamy distribution: median-dispersion reparameterizations for regression modeling and simulation-based estimation. **Statistical Papers**, v.54, p.177–192, 2013.

NASCIMENTO, A. D. C. et al. The Gamma extended Weibull family of distributions. **Journal of Statistical Theory and Applications**, v.13, p.1–16, 2014.

NICOLAU, J. **História do Voto no Brasil**. Rio de Janeiro: Jorge Zahar, 2004.

POWELL, E. N.; TUCKER, J. A. Revisiting electoral volatility in post-communist countries: new data, new results and new approaches. **British Journal of Political Science**, v.44, p.123–147, 2013.

R Core Team. **R: a language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2018.

ROCKE, D. M. Robust control charts. **Technometrics**, v.31, n.2, p.173–184, 1989.

RODOLFO, A.; GABRIEL, Z. Como o Brasil votou no segundo turno das eleições de 2018 para presidente, em mapas. **NEXO**, 2020.

SANTOS-NETO, M. et al. The Marshall-Olkin extended Weibull family of distributions. **Journal of Statistical Distributions and Applications**, v.1, p.9, 2014.

SCORSIM, E. M. Brasil, eleições 2018: a análise dos números de votos do vencedor, do perdedor e brancos, nulos e abstenções. **Migalhas**, 2018.

SILVA, R. B. et al. The compound class of extended Weibull power series distributions. **Computational Statistics and Data Analysis**, v.58, p.352–367, 2013.

TAHMASEBI, S.; JAFARI, A. A. Exponentiated extended Weibull-power series class of distributions. **Ciência e Natura**, v.7, p.183–193, 2015.

TOPP, C. W.; LEONE, F. C. A family of J-shaped frequency functions. **Journal of the American Statistical Association**, v.50, p.209–219, 1955.

APÊNDICES

APÊNDICE A

A.1. Resultados numéricos para a distribuição \mathcal{NRU} considerando outros valores de τ

Tabela 1 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para o parâmetro da distribuição \mathcal{NRU} , para diferentes cenários e tamanhos amostrais, considerando $\tau = 0.25$.

Cenário	μ	n	MÉDIA($\hat{\mu}$)	VARIÂNCIA($\hat{\mu}$)	VR %($\hat{\mu}$)	EQM($\hat{\mu}$)	TAXA($\hat{\mu}$)
1	0.15	10	0.1475	0.0003937	-0.2432	0.0595	0.949
		25	0.1493	0.0000842	-0.0617	0.0039	0.95
		50	0.1493	0.0000842	-0.0617	0.0039	0.959
		75	0.1494	0.0000569	-0.0505	0.0026	0.948
		100	0.1497	0.0000392	-0.0286	0.0008	0.958
2	0.25	10	0.2463	0.000892	-0.3669	0.1355	0.9470
		25	0.2484	0.000344	-0.1584	0.0254	0.9579
		50	0.2492	0.000174	-0.0773	0.0061	0.9533
		75	0.2495	0.000115	-0.0446	0.0021	0.9537
		100	0.2497	0.0000889	-0.0287	0.0009	0.9518
3	0.4	10	0.3943	0.0014	-0.5635	0.3190	0.9552
		25	0.3976	0.0005	-0.2385	0.0574	0.9594
		50	0.3988	0.0002	-0.1170	0.0139	0.9538
		75	0.3993	0.0001	-0.0689	0.0049	0.9536
		100	0.3995	0.0001	-0.0459	0.0022	0.9508
4	0.7	10	0.6933	0.0012	-0.6654	0.4440	0.9518
		25	0.6972	0.0004	-0.2725	0.0747	0.9535
		50	0.6986	0.0002	-0.1341	0.0182	0.9506
		75	0.6991	0.0001	-0.0812	0.0067	0.9521
		100	0.6994	0.0001	-0.0562	0.0032	0.9469
5	0.9	10	0.8966	0.000238	-0.3384	0.1148	0.9405
		25	0.8986	0.0000843	-0.1356	0.0184	0.9489
		50	0.8993	0.0000413	-0.0666	0.0044	0.9497
		75	0.8995	0.0000271	-0.0408	0.0016	0.9515
		100	0.8997	0.0000208	-0.0287	0.0008	0.9454

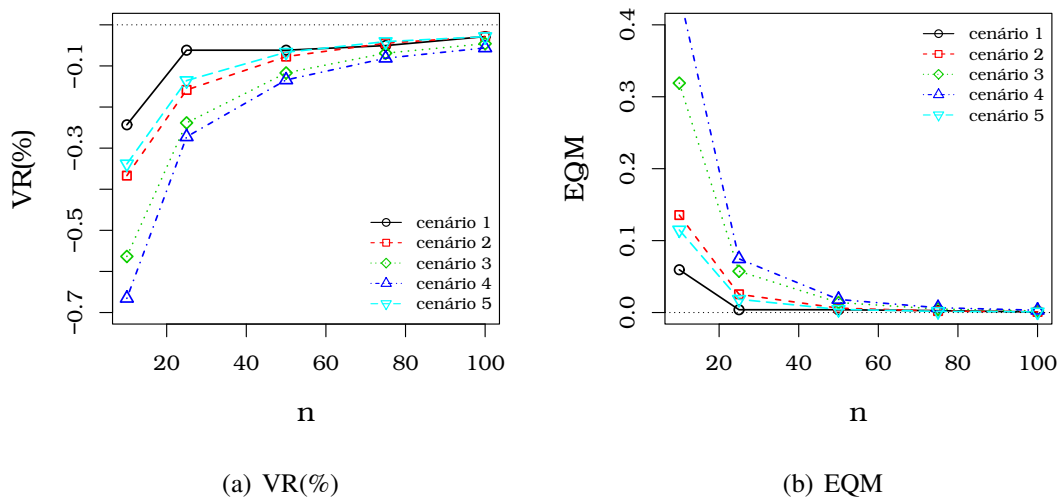


Figura 1 – VR(%) e EQM para o estimador de máxima verossimilhança da distribuição \mathcal{NRU} , considerando $\tau = 0.25$.

Tabela 2 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para o parâmetro da distribuição \mathcal{NRU} , para diferentes cenários e tamanhos amostrais, considerando $\tau = 0.75$.

Cenário	μ	n	MÉDIA($\hat{\mu}$)	VARIÂNCIA($\hat{\mu}$)	VR %($\hat{\mu}$)	EQM($\hat{\mu}$)	TAXA($\hat{\mu}$)
1	0.15	10	0.1475	0.000393	-0.2432	0.0595	0.949
		25	0.1492	0.000172	-0.0741	0.0056	0.943
		50	0.1493	0.0000842	-0.0617	0.0039	0.959
		75	0.1494	0.0000569	-0.0505	0.0026	0.948
		100	0.1497	0.0000392	-0.0286	0.0008	0.958
2	0.25	10	0.2463	0.000892	-0.3669	0.1355	0.9470
		25	0.2484	0.000344	-0.1584	0.0254	0.9579
		50	0.2492	0.000174	-0.0773	0.0061	0.9533
		75	0.2495	0.000115	-0.0446	0.0021	0.9537
		100	0.2497	0.000889	-0.0287	0.0009	0.9518
3	0.4	10	0.3943	0.0014	-0.5635	0.3190	0.9552
		25	0.3976	0.0005	-0.2385	0.0574	0.9594
		50	0.3988	0.0002	-0.1170	0.0139	0.9538
		75	0.3993	0.0001	-0.0689	0.0049	0.9536
		100	0.3995	0.0001	-0.0459	0.0022	0.9508
4	0.7	10	0.6933	0.0012	-0.6654	0.4440	0.9518
		25	0.6972	0.0004	-0.2725	0.0747	0.9535
		50	0.6986	0.0002	-0.1341	0.0182	0.9506
		75	0.6991	0.0001	-0.0812	0.0067	0.9521
		100	0.6994	0.0001	-0.0562	0.0032	0.9469
5	0.9	10	0.8966	0.000238	-0.3384	0.1148	0.9405
		25	0.8986	0.0000843	-0.1356	0.01847	0.9489
		50	0.8993	0.0000413	-0.0666	0.0044	0.9497
		75	0.8995	0.0000271	-0.0408	0.0016	0.9515
		100	0.8997	0.0000208	-0.0287	0.0008	0.945

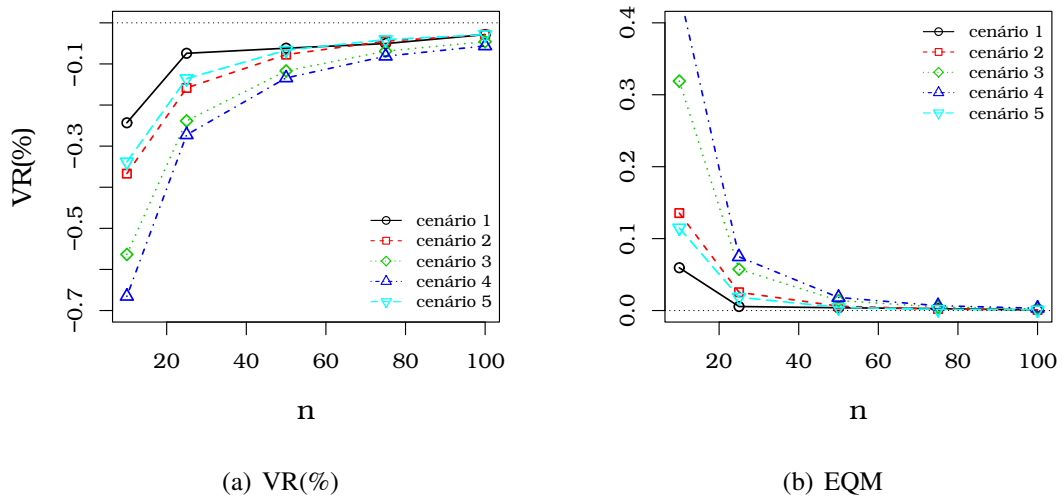


Figura 2 – VR(%) e EQM para o estimador de máxima verossimilhança da distribuição \mathcal{NRU} , considerando $\tau = 0.75$.

A.2. Resultados numéricos para a distribuição \mathcal{NGOU} considerando outros valores de τ

Tabela 3 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para os parâmetros da distribuição \mathcal{NUGO} , para diferentes cenários e tamanhos amostrais, considerando $\tau = 0.25$.

Cenário	μ	β	n	MÉDIA($\hat{\mu}$)		VARIÂNCIA($\hat{\mu}$)		VR %($\hat{\mu}$)		EQM($\hat{\mu}$)		TAXA($\hat{\mu}$)	
				$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$
1	0.4	0.8	10	0.4176	1.0934	0.0072	0.3085	1.7653	29.3480	0.0075	0.3946	0.8745	0.9375
			25	0.4068	0.9008	0.0031	0.0837	0.6851	10.0894	0.0032	0.0939	0.9200	0.9459
			50	0.4039	0.8512	0.0015	0.0373	0.3933	5.1278	0.0016	0.0400	0.9355	0.9484
			75	0.4029	0.8346	0.0011	0.0241	0.2958	3.4678	0.0011	0.0253	0.9383	0.9496
			100	0.4019	0.8240	0.0008	0.0176	0.1990	2.4019	0.0008	0.0182	0.9414	0.9471
2	0.5	1.5	10	0.5050	1.8224	0.0041	0.4432	0.5026	32.2471	0.0041	0.5472	0.8704	0.9498
			25	0.5015	1.6112	0.0017	0.1237	0.1558	11.1275	0.0017	0.1360	0.9149	0.9505
			50	0.5010	1.5555	0.0008	0.0557	0.1009	5.5596	0.0008	0.0588	0.9333	0.9516
			75	0.5009	1.5383	0.0006	0.0362	0.0923	3.8320	0.0006	0.0376	0.9367	0.9515
			100	0.5004	1.5262	0.0004	0.0266	0.0497	2.6217	0.0004	0.0273	0.9393	0.9485
3	0.6	2.8	10	0.6010	3.2802	0.0009	1.0210	0.1034	48.0217	0.000906	1.2516	0.8816	0.9524
			25	0.6002	2.9648	0.0003	0.2661	0.0260	16.4886	0.000355	0.2932	0.9223	0.9516
			50	0.6002	2.8812	0.0001	0.1179	0.0213	8.1212	0.000171	0.1245	0.9367	0.9510
			75	0.6002	2.8557	0.0001	0.0757	0.0249	5.5727	0.000117	0.0788	0.9408	0.9519
			100	0.6001	2.8382	0.00008	0.0558	0.0101	3.8236	0.0000883	0.0572	0.9409	0.9488
4	0.7	5	10	0.7007	5.8588	0.0000874	3.0950	0.0745	85.8877	0.0000880	3.8327	0.8891	0.9505
			25	0.7003	5.2968	0.0000350	0.7870	0.0266	29.6807	0.0000351	0.8751	0.9246	0.9507
			50	0.7002	5.1453	0.0000171	0.3462	0.01563	14.5381	0.0000171	0.3673	0.9388	0.9503
			75	0.7001	5.0991	0.0000117	0.2226	0.0136	9.9168	0.0000117	0.2325	0.9413	0.9505
			100	0.7000	5.0696	0.00000884	0.1632	0.0084	6.9653	0.00000884	0.1680	0.9426	0.9481
5	0.7	0.9	10	0.7004	1.0727	0.0021	0.1316	0.0489	17.2732	0.0021	0.1615	0.8699	0.9520
			25	0.6997	0.9595	0.0008	0.0361	-0.0129	5.9559	0.0008	0.0397	0.9160	0.9512
			50	0.70008	0.9295	0.0004	0.0162	0.0083	2.9576	0.0004	0.0171	0.9336	0.9518
			75	0.7002	0.9204	0.0002	0.0105	0.0225	2.0456	0.0002	0.0109	0.9374	0.9513
			100	0.70003	0.9139	0.0003	0.0077	0.0033	1.3979	0.0003	0.0079	0.9397	0.9476
6	0.75	1.1	10	0.7501	1.2926	0.000894	0.1659	0.0100	19.2602	0.000894	0.2030	0.8747	0.9525
			25	0.7498	1.1662	0.000344	0.0440	-0.0121	6.6210	0.000344	0.0484	0.9186	0.9506
			50	0.75002	1.1326	0.000164	0.01962	0.0026	3.2657	0.000164	0.0206	0.9355	0.9512
			75	0.7501	1.1224	0.000112	0.01266	0.0120	2.2490	0.000112	0.0131	0.9399	0.9510
			100	0.7500	1.1153	0.0000845	0.0093	0.0006	1.5397	0.0000845	0.0095	0.9402	0.9475

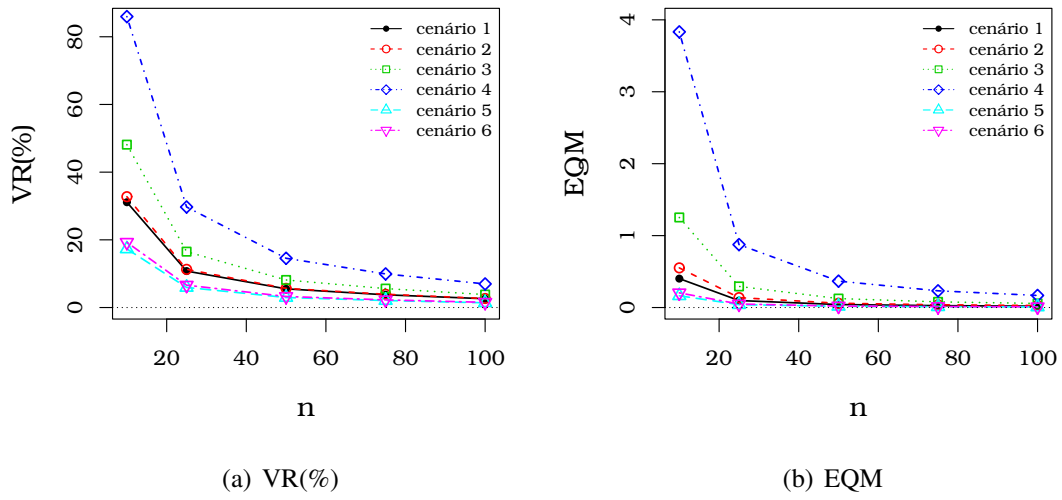


Figura 3 – VR(%) e EQM para o estimador de máxima verossimilhança da distribuição \mathcal{NGOU} , considerando $\tau = 0.25$.

Tabela 4 – Resultados da avaliação numérica dos estimadores de máxima verossimilhança para os parâmetros da distribuição \mathcal{NUGO} , para diferentes cenários e tamanhos amostrais, considerando $\tau = 0.75$.

Cenário	μ	β	n	MÉDIA($\hat{\mu}$)		VARIÂNCIA($\hat{\mu}$)		VR %($\hat{\mu}$)		EQM($\hat{\mu}$)		TAXA($\hat{\mu}$)	
				$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\beta}$
1	0.4	0.8	10	0.3889	1.6129	0.0036	1.8802	-1.1010	81.2902	0.0037	2.5410	0.8580	0.9260
			25	0.3954	1.0811	0.0014	0.4361	-0.4579	28.1194	0.0014	0.5152	0.9204	0.9447
			50	0.3978	0.9410	0.0007	0.1857	-0.2129	14.1035	0.0007	0.2056	0.9371	0.9464
			75	0.3987	0.8948	0.0004	0.1174	-0.1254	9.4821	0.0004	0.1264	0.9438	0.9476
			100	0.3990	0.8666	0.0003	0.0848	-0.0925	6.6619	0.0003	0.0893	0.9396	0.9469
2	0.5	1.5	10	0.4889	2.0435	0.0019	1.0623	-1.1072	54.3567	0.0020	1.3578	0.8751	0.9378
			25	0.4954	1.6869	0.0007	0.2889	-0.4545	18.6901	0.0007	0.3238	0.9286	0.9463
			50	0.4978	1.5949	0.0003	0.1289	-0.2188	9.4927	0.0003	0.1379	0.9362	0.9487
			75	0.4986	1.5642	0.0002	0.0831	-0.1352	6.4243	0.0002	0.0873	0.9436	0.9496
			100	0.4990	1.5445	0.0001	0.0610	-0.0992	4.4502	0.0001	0.0630	0.9412	0.9474
3	0.6	2.8	10	0.5952	3.3067	0.000371	1.1475	-0.4797	50.6705	0.000394	1.4043	0.8895	0.9525
			25	0.5980	2.9744	0.000132	0.3098	-0.1926	17.4452	0.000136	0.3402	0.9342	0.9507
			50	0.5990	2.8862	0.0000649	0.1386	-0.0939	8.6236	0.0000658	0.1461	0.9395	0.9517
			75	0.5994	2.8596	0.0000428	0.0897	-0.0578	5.9635	0.0000432	0.0932	0.9466	0.9515
			100	0.5995	2.8407	0.0000327	0.0661	-0.0417	4.0764	0.0000329	0.0678	0.9422	0.9470
4	0.7	5	10	0.6986	5.8552	0.0000359	3.0852	-0.1310	85.5202	0.0000376	3.8166	0.8885	0.9499
			25	0.6994	5.2945	0.0000132	0.7860	-0.0527	29.4549	0.0000134	0.8728	0.9321	0.9506
			50	0.6997	5.1438	0.00000655	0.3464	-0.0256	14.3895	0.00000662	0.3671	0.9381	0.9502
			75	0.6998	5.0987	0.00000433	0.2215	-0.0155	9.8796	0.00000435	0.2313	0.9467	0.9501
			100	0.6998	5.0679	0.00000332	0.1630	-0.0109	6.7998	0.00000333	0.1676	0.9409	0.9480
5	0.7	0.9	10	0.6913	1.1426	0.000975	0.2325	-0.8659	24.2641	0.000105	0.2914	0.8863	0.9431
			25	0.6965	0.9836	0.000336	0.0649	-0.3463	8.3684	0.000348	0.0719	0.9314	0.9479
			50	0.6983	0.9421	0.000161	0.0292	-0.1676	4.2128	0.000164	0.0309	0.9400	0.9502
			75	0.6989	0.9288	0.000107	0.0189	-0.1041	2.8846	0.000108	0.0197	0.9448	0.9498
			100	0.6992	0.9198	0.0000814	0.0139	-0.0764	1.9841	0.0000819	0.0143	0.9419	0.9479
6	0.75	1.1	10	0.7448	1.3190	0.000371	0.2095	-0.5193	21.9045	0.000398	0.2575	0.8913	0.9515
			25	0.7479	1.1755	0.000128	0.0579	-0.2065	7.5579	0.000132	0.0637	0.9343	0.9511
			50	0.7489	1.1375	0.0000621	0.0260	-0.1004	3.7591	0.0000631	0.0275	0.9403	0.9514
			75	0.7493	1.1259	0.0000409	0.0169	-0.0622	2.5967	0.0000413	0.0176	0.9467	0.9512
			100	0.7495	1.1178	0.0000312	0.0124	-0.0452	1.7836	0.0000314	0.0128	0.9423	0.9475

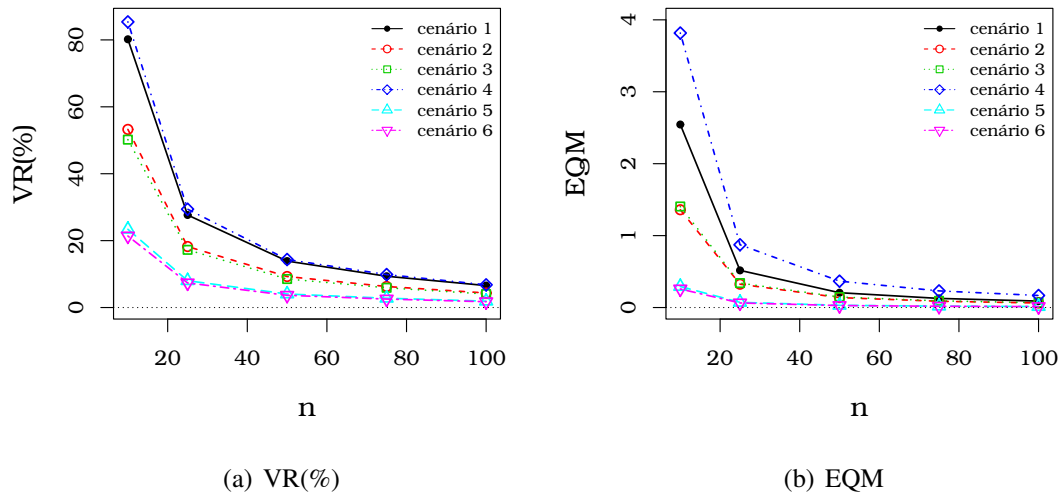


Figura 4 – VR(%) e EQM para o estimador de máxima verossimilhança da distribuição \mathcal{NGOU} , considerando $\tau = 0.75$.