



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

Benedito Vicente dos Santos

**Regressão Quantílica para Modelos na Família de Distribuições  
G-exponencializada Reparametrizada e suas Aplicações**

Recife  
2025

Benedito Vicente dos Santos

**Regressão Quantílica para Modelos na Família de Distribuições  
G-exponencializada Reparametrizada e suas Aplicações**

Tese apresentada ao programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutor em Estatística.

**Área de Concentração:** Estatística Aplicada

**Orientador:** Prof. Dr. Francisco José de Azevêdo Cysneiros

**Coorientador:** Prof. Dr. Manoel Ferreira dos Santos Neto

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Santos, Benedito Vicente Dos.

Regressão quantílica para modelos na família de distribuições G-exponencializada reparametrizada e suas aplicações / Benedito Vicente Dos Santos. - Recife, 2025.

105f.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Ciências Exatas e da Natureza, Programa de Pós-Graduação em Estatística, 2025.

Orientação: Francisco José de Azevêdo Cysneiros.

Inclui referências e apêndices.

1. Análise de diagnóstico; 2. Distribuição Baseline; 3. Distribuição G-exponencializada reparametrizada; 4. Quantil; 5. Regressão quantílica; 6. Simulação de Monte Carlo. I. Cysneiros, Francisco José de Azevêdo. II. Título.

UFPE-Biblioteca Central

**Benedito Vicente dos Santos**

**Regressão Quantílica para Modelos na Família de Distribuições  
G-exponencializada Reparametrizada e suas Aplicações**

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

Aprovada em: 27 de fevereiro de 2025.

**BANCA EXAMINADORA**

Prof. Dr. Manoel Ferreira dos Santos Neto  
UFC

Prof. Dr. Roberto Ferreira Manghi  
UFPE

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Maria do Carmo Soares de Lima  
UFPE

Prof. Dr. Juvêncio Santos Nobre  
UFC

Prof. Dr. Helton Saulo Bezerra dos Santos  
UNB

Dedico esta tese aos meus pais e irmãos.

## AGRADECIMENTOS

Aos meus pais, Aniceto e Lúcia, e aos meus irmãos, Iris, Celson, Maísa, Cristina, Rafaela e Karl Marx por todo o apoio e compreensão.

Agradeço aos professores Francisco José de Azevêdo Cysneiros e Manoel Ferreira dos Santos Neto por me orientar na construção desta tese, pelo apoio, ajuda, dedicação e compreensão em todos os momentos da realização desta tese. Muito obrigado!

Agradeço aos membros da banca examinadora pelas correções, sugestões e contribuições, que foram fundamentais para melhorar a versão final desta tese.

Aos docentes do Programa de Pós-Graduação em Estatística da UFPE, que contribuíram na minha formação acadêmica no doutorado.

A todos os colegas do doutorado pelos momentos de estudo.

À UFPE pela oportunidade.

À Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) pelo apoio financeiro.

A todos que contribuíram para o desenvolvimento desta tese. Muito obrigado!

*“A alegria está na luta, na tentativa, no sofrimento envolvido e não na vitória propriamente dita.” (Mahatma Gandhi)*

## RESUMO

Neste trabalho, propomos uma nova família de distribuições denominada família de distribuições G-exponencializada reparametrizada. Esta família reparametrizada tem uma vantagem em comparação com a distribuição G-exponencializada. Ela é indexada por um parâmetro  $\eta_\tau$ , e permite uma interpretação desse parâmetro em termos do  $\tau$ -ésimo quantil da distribuição. Sendo assim, a família de distribuições G-exponencializada reparametrizada é importante quando o interesse é modelar os quantis da distribuição. Além disso, apresentamos vários novos modelos probabilísticos baseados na família de distribuições G-exponencializada reparametrizada. Para exemplificar como são obtidos os submodelos da família de distribuições G-exponencializada reparametrizada foi utilizado o modelo Weibull reparametrizado como *baseline*, resultando no modelo Weibull exponencializado reparametrizado (WER). Através de simulações de Monte Carlo realizadas utilizando a linguagem R foi avaliado numericamente o desempenho dos estimadores de máxima verossimilhança e seus respectivos intervalos de confiança assintóticos. Para ilustrar, aplicou-se a família de distribuições proposta a um conjunto de dados reais. Também, propomos modelos de regressão quantílica de modo geral para modelos pertencentes à família de distribuições G-exponencializada reparametrizada. Foi introduzido o novo modelo de regressão quantílica WER e algumas propriedades matemáticas desse modelo. Adicionalmente, foi utilizado o método de máxima verossimilhança para estimar os parâmetros dos modelos propostos. Foi realizado um estudo de simulação de Monte Carlo para avaliar o desempenho e as propriedades dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de regressão quantílica WER. Além disso, propomos alguns resíduos e técnicas de análise de diagnóstico, com o objetivo de detectar observações potencialmente influentes e examinar o ajuste do modelo ao conjunto de dados. Por fim, para exemplificar, foram aplicados os novos modelos de regressão quantílica a um conjunto de dados reais na área de economia.

**Palavras-chave:** Análise de diagnóstico; Distribuição *Baseline*; Distribuição G-exponencializada reparametrizada; Quantil; Regressão quantílica; Simulação de Monte Carlo.

## ABSTRACT

In this work, we propose a new distributions family called reparameterized exponentialized-G distributions family. This reparameterized family has an advantage compared to the exponentialized-G distribution. She is indexed by a parameter  $\eta_\tau$ , and allows an interpretation of this parameter in terms of the  $\tau$ -th quantile of the distribution. Therefore, the reparameterized exponentialized-G distributions family is important when the interest is to model the quantiles of the distribution. Furthermore, we present several new probabilistic models based on the reparameterized exponentialized-G distribution family. To exemplify how the submodels of the reparameterized exponentialized-G distributions family are obtained, the reparameterized Weibull model was used as a baseline, resulting in the reparameterized exponentialized Weibull (REW) model. Through Monte Carlo simulations performed using the R language, the performance of the maximum likelihood estimators and their respective asymptotic confidence intervals was numerically evaluated. To illustrate, the proposed distributions family was applied to a real dataset. We also propose general form quantile regression models for models belonging to the reparameterized exponentialized-G distributions family. The new REW quantile regression model and some mathematical properties of this model were introduced. Additionally, the maximum likelihood method was used to estimate the parameters of the proposed models. A Monte Carlo simulation study was carried out to evaluate the performance and properties of maximum likelihood estimators of the parameters indexing the REW quantile regression model. Furthermore, we propose some residuals and diagnostic analysis techniques, aiming to detect potentially influential observations and examine the fit of the model to the dataset. Finally, to illustrate, the new quantile regression models were applied to a real dataset in the area of economics.

**Keywords:** Diagnostic analysis; Baseline distribution; Reparametrized exponentialized-G distribution; Quantile; Quantile regression; Monte Carlo simulations.

## LISTA DE FIGURAS

Figura 1 – Densidades WER para diferentes valores de $\kappa$ e $\tau$ com $\eta_\tau = 1$ fixo. . . . .	31
Figura 2 – Densidades WER para diferentes valores de $\eta_\tau$ e $\tau$ com $\kappa = 5$ fixo. . . . .	32
Figura 3 – Densidades WER para diferentes valores de $\kappa$ e $\tau$ considerando valores de $\kappa$ menores que 1, para $\eta_\tau = 1$ fixo. . . . .	33
Figura 4 – Distribuição empírica de $\hat{\kappa}$ com $\tau = 0,25$ e $\eta_\tau = 1$ fixo. . . . .	42
Figura 5 – Distribuição empírica de $\hat{\kappa}$ com $\tau = 0,5$ e $\eta_\tau = 1$ fixo. . . . .	43
Figura 6 – Distribuição empírica de $\hat{\kappa}$ com $\tau = 0,75$ e $\eta_\tau = 1$ fixo. . . . .	44
Figura 7 – Histograma com densidades estimadas para os modelos WER e LLER para o conjunto de dados de tensão de ruptura de fibras de carbono (em GBa). . . . .	48
Figura 8 – qqplots da distribuição WER para o conjunto de dados de tensão de ruptura de fibras de carbono (em GBa). . . . .	49
Figura 9 – qqplots da distribuição LLER para o conjunto de dados de tensão de ruptura de fibras de carbono (em GBa). . . . .	50
Figura 10 – Histograma da renda (a), gráfico de dispersão da renda e educação (b) e gráfico de dispersão da renda e porcentagem de mulheres no cargo (c) para o conjunto de dados Prestige. . . . .	73
Figura 11 – Envelopes simulados com faixas de confiança para o resíduo quantílico e modelos de RQ WER, LLER e LER com $\tau = 0,25$ . . . . .	78
Figura 12 – Envelopes simulados com faixas de confiança para o resíduo generalizado de Cox-Snell e modelos de RQ WER, LLER e LER com $\tau = 0,25$ . . . . .	79
Figura 13 – Envelopes simulados com faixas de confiança para o resíduo tipo Martingal e modelos de RQ WER, LLER e LER com $\tau = 0,25$ . . . . .	80
Figura 14 – Distância de Cook generalizada para os modelos de RQ WER, LLER e LER com $\tau = 0,25$ . . . . .	81
Figura 15 – Gráfico de $C_i(\theta)$ versus o índice das observações para o modelo de RQ WER para $\tau = 0,25$ e utilizando o esquema de perturbação de casos (a), perturbação na resposta (b) e perturbação na covariável contínua (c). . . . .	83

Figura 16 – Gráfico de  $C_i(\theta)$  versus o índice das observações para o modelo de RQ LLER para  $\tau = 0,25$  e utilizando o esquema de perturbação de casos (a), perturbação na resposta (b) e perturbação na covariável contínua (c). 84

Figura 17 – Gráfico de  $C_i(\theta)$  versus o índice das observações para o modelo de RQ LLER para  $\tau = 0,25$  e utilizando o esquema de perturbação de casos (a), perturbação na resposta (b) e perturbação na covariável contínua (c). . 85

## LISTA DE TABELAS

Tabela 1 – Alguns submodelos da família de distribuições $\alpha_\tau$ -G reparametrizada. . . . .	27
Tabela 2 – Estimativas de máxima verossimilhança e medidas de desempenho estimadas pelo método BFGS para $\kappa$ : 1,5, 2,5 e 5, $\eta_\tau = 1$ fixo, e $\tau$ : 0,25, 0,5 e 0,75. . . . .	40
Tabela 3 – Estimativas intervalares e probabilidades de cobertura (PC) empíricas dos IC assintótico e PC empíricas das caudas esquerda (LIE) e direita (LSD) para $\kappa$ : 1,5, 2,5 e 5, $\eta_\tau = 1$ fixo, e $\tau$ : 0,25, 0,5 e 0,75, com nível de confiança $1 - \gamma = 0,95$ . . . . .	41
Tabela 4 – Medidas descritivas para o conjunto de dados de tensão de ruptura de fibras de carbono. . . . .	45
Tabela 5 – EMV dos parâmetros para vários modelos ajustados, com as estimativas dos erros-padrão correspondentes entre parênteses. . . . .	46
Tabela 6 – Estatísticas AIC, CAIC, BIC, $A^*$ , $W^*$ e $K-S$ (com seus $p$ -valores) para o conjunto de dados de estresse de ruptura das fibras de carbono (em Gba). . . . .	47
Tabela 7 – Estimativas do CS, CK, VAR, VR e EQM para $\psi_0(\tau) = 0,5$ , $\psi_1(\tau) = 1,5$ fixos, $\kappa$ : 0,5, 2,5 e 3,5 e $\tau$ : 0,25, 0,5 e 0,75. . . . .	70
Tabela 8 – Estatísticas descritivas para o conjunto de dados Prestige. . . . .	72
Tabela 9 – Estatísticas AIC, CAIC, BIC e função log-verossimilhança para os modelos RQ WER, LLER e LER com $\tau = 0,25$ fixo e funções de ligações FLL, FLR e FLI para o conjunto de dados Prestige. . . . .	75
Tabela 10 – EMV dos parâmetros, estimativas dos erros-padrão (EP) correspondentes e seus respectivos $p$ -valores para os modelos de RQ ajustados utilizando o conjunto de dados Prestige com $\tau = 0,25$ fixo e função de ligação FLL. . . . .	75
Tabela 11 – Estatísticas AIC, CAIC, BIC, $R_\tau^2$ e função log-verossimilhança para os modelos de RQ WER, LLER e LER com $\tau = 0,25$ fixo e função de ligação FLL utilizando o conjunto de dados Prestige. . . . .	76
Tabela 12 – Medidas RCs (em %) com seus correspondentes se's e respectivos $p$ -valores para o modelo de RQ WER com $\tau = 0,25$ e utilizado o conjunto de dados Prestige. . . . .	87

Tabela 13 – Medidas RCs (em %) com seus correspondentes se's e respectivos p- valores para o modelo de RQ LLER com $\tau = 0,25$ e utilizado o conjunto de dados Prestige. . . . .	87
Tabela 14 – Medidas RCs (em %) com seus correspondentes se's e respectivos p- valores para o modelo de RQ LER com $\tau = 0,25$ e utilizado o conjunto de dados Prestige. . . . .	88

## SUMÁRIO

<b>1</b>	<b>PRELIMINARES</b> . . . . .	<b>15</b>
1.1	INTRODUÇÃO E REVISÃO BIBLIOGRÁFICA . . . . .	15
1.2	MOTIVAÇÃO E JUSTIFICATIVA DA TESE . . . . .	18
1.3	OBJETIVOS . . . . .	18
<b>1.3.1</b>	<b>Geral</b> . . . . .	<b>18</b>
<b>1.3.2</b>	<b>Específicos</b> . . . . .	<b>19</b>
1.4	PRODUTOS DA TESE . . . . .	19
1.5	ORGANIZAÇÃO DA TESE . . . . .	19
1.6	FERRAMENTAS COMPUTACIONAIS . . . . .	20
<b>2</b>	<b>FAMÍLIA DE DISTRIBUIÇÕES G-EXPONENCIALIZADA REPA- RAMETRIZADA: PROPRIEDADES E APLICAÇÕES</b> . . . . .	<b>21</b>
2.1	RESUMO . . . . .	21
2.2	INTRODUÇÃO . . . . .	21
2.3	FAMÍLIA DE DISTRIBUIÇÕES $\alpha$ -G REPARAMETRIZADA . . . . .	23
<b>2.3.1</b>	<b>Estimação via Método de Máxima Verossimilhança</b> . . . . .	<b>24</b>
2.4	DISTRIBUIÇÃO WEIBULL REPARAMETRIZADA PELA MEDIANA . . . . .	28
2.5	DISTRIBUIÇÃO WER . . . . .	28
<b>2.5.1</b>	<b>Estimação via Método de Máxima Verossimilhança</b> . . . . .	<b>34</b>
2.6	SIMULAÇÃO . . . . .	35
2.7	APLICAÇÃO . . . . .	45
2.8	OBSERVAÇÕES FINAIS . . . . .	51
<b>3</b>	<b>REGRESSÃO QUANTÍLICA PARA MODELOS BASEADOS NA FAMÍLIA DE DISTRIBUIÇÕES G-EXPONENCIALIZADA (<math>\alpha_\tau</math>-G) REPARAMETRIZADA</b> . . . . .	<b>52</b>
3.1	RESUMO . . . . .	52
3.2	INTRODUÇÃO . . . . .	52
3.3	MODELO DE REGRESSÃO QUANTÍLICA BASEADO NA FAMÍLIA DE DISTRIBUIÇÕES $\alpha_\tau$ -G REPARAMETRIZADA . . . . .	55
<b>3.3.1</b>	<b>Modelo de Regressão Quantílica Baseado no Modelo WER</b> . . . . .	<b>58</b>

3.4	ANÁLISE DE RESÍDUOS E DIAGNÓSTICO DE INFLUÊNCIA . . . . .	61
3.4.1	<b>Distância de Cook Generalizada . . . . .</b>	<b>61</b>
3.4.2	<b>Resíduos Quantílicos (<math>RQ_{\tau}</math>) . . . . .</b>	<b>62</b>
3.4.3	<b>Resíduo Generalizado de Cox-Snell . . . . .</b>	<b>63</b>
3.4.4	<b>Resíduo Tipo Martingal . . . . .</b>	<b>63</b>
3.4.5	<b>Pseudo-<math>R^2</math> . . . . .</b>	<b>63</b>
3.4.6	<b>Influência Local . . . . .</b>	<b>64</b>
3.4.7	<b>Esquemas de Perturbação . . . . .</b>	<b>65</b>
3.4.7.1	<i>Perturbação de Casos . . . . .</i>	<i>65</i>
3.4.7.2	<i>Perturbação na Resposta . . . . .</i>	<i>66</i>
3.4.7.3	<i>Perturbação na Covariável Contínua . . . . .</i>	<i>66</i>
3.5	SIMULAÇÃO . . . . .	67
3.6	APLICAÇÃO . . . . .	71
3.6.1	<b>Análise de Resíduos para os Modelos de RQ WER, LLER e LER</b>	<b>76</b>
3.6.2	<b>Análise da Distância de Cook Generalizada para os Modelos de RQ WER, LLER e LER . . . . .</b>	<b>77</b>
3.6.3	<b>Análise de Influência Local para os Modelos de RQ WER, LLER e LER . . . . .</b>	<b>82</b>
3.6.4	<b>Análise Confirmatória . . . . .</b>	<b>86</b>
3.7	OBSERVAÇÕES FINAIS . . . . .	88
4	<b>CONCLUSÕES E TRABALHOS FUTUROS . . . . .</b>	<b>90</b>
4.1	TRABALHOS FUTUROS . . . . .	91
	<b>REFERÊNCIAS . . . . .</b>	<b>92</b>
	<b>APÊNDICE A – ESQUEMAS DE PERTURBAÇÃO PARA O MODELO DE REGRESSÃO QUANTÍLICA WER</b>	<b>96</b>
	<b>APÊNDICE B – ESQUEMAS DE PERTURBAÇÃO PARA O MODELO DE REGRESSÃO QUANTÍLICA LLER</b>	<b>99</b>
	<b>APÊNDICE C – ESQUEMAS DE PERTURBAÇÃO PARA O MODELO DE REGRESSÃO QUANTÍLICA LER</b>	<b>102</b>

## 1 PRELIMINARES

### 1.1 INTRODUÇÃO E REVISÃO BIBLIOGRÁFICA

Um gerador de distribuições, também conhecido na literatura por família de distribuições G-exponencializada, foi proposto por Lehmann (1953). A ideia da família de distribuições G-exponencializada é adicionar um parâmetro de forma à distribuição *baseline*, obtendo um modelo que pode ser útil em situações para os quais o modelo *baseline* expressa algum tipo de limitação.

A partir da família de distribuições G-exponencializada, vários modelos probabilísticos foram propostos na literatura com o objetivo de generalizar a distribuição *baseline*. Por exemplo, o modelo exponencial exponencializado (EE), proposto por Gupta e Kundu (2001), é uma extensão do modelo exponencial. Este modelo é frequentemente utilizado na análise de dados de sobrevivência, em aplicações médicas e de engenharia.

Já a distribuição Weibull exponencializada (WE) foi proposta por Mudholkar e Srivastava (1993) e estende a distribuição Weibull. O modelo WE tem sido muito utilizado para análise de dados de sobrevivência uma vez que pode apresentar função de taxa de falha em forma constante, monótona crescente, monótona decrescente, unimodal e banheira (forma de U). Também encontramos aplicações do modelo WE em engenharia de confiabilidade e em diversas outras áreas. Adicionalmente, o modelo EE é um caso especial do modelo WE quando o parâmetro de forma  $\kappa$  assume o valor 1. Além disso, a distribuição WE tem amplo uso teórico e prático na estatística.

Nadarajah e Kotz (2003) propuseram o modelo Fréchet exponencializado (FE) e estudaram suas propriedades estatísticas. O modelo Rayleigh exponencializado (RE), proposto por Kundu e Raqab (2005) é uma generalização do modelo Rayleigh. Além disso, Kundu e Raqab (2005) consideraram diferentes estimadores e realizaram um estudo comparativo desses estimadores. O modelo Birnbaum-Saunders exponencializado (BSE) foi introduzido por Cordeiro, Lemonte e Ortega (2013). Além disso, esses autores analisaram várias propriedades estatísticas desse modelo.

O modelo Lomax exponencializado (LE), foi proposto por Abdul-Moniem e Abdel-Hameed (2012) é uma generalização do modelo Lomax. Adicionalmente, Abdul-Moniem e Abdel-Hameed (2012) examinaram algumas propriedades estatísticas desse modelo. El-

Gohary, Ahmad e Adel (2013) introduziram o modelo Gompertz exponencializado (GE) e também discutiram a flexibilidade do modelo GE e suas propriedades estatísticas. Nadarajah (2006) introduziu a distribuição Gumbel exponencializada (GmE), que generaliza a distribuição Gumbel, e desenvolveu algumas propriedades estatísticas e aplicações para essa distribuição.

Rosaiah, Kantam e Kumar (2006) consideraram o modelo log-logístico exponencializado (LLE), que é uma extensão do modelo log-logístico. Os autores também apresentaram um plano de teste de confiabilidade e aplicaram o modelo LLE usando um exemplo numérico. O modelo Burr XII exponencializado (BXIIE) foi proposto por AL-Hussaini e Hussein (2011) e generaliza o modelo Burr XII. O modelo Birnbaum-Saunders generalizado exponencializado (BSGE) foi introduzido por Gallardo e Santos-Neto (2021) para generalizar o modelo baseline Birnbaum-Saunders generalizado proposto por Owen (2006).

Nesta tese, a família de distribuições G-exponencializada reparametrizada é proposta e algumas propriedades estatísticas gerais dessa família são derivadas. Também são apresentadas algumas distribuições que pertencem à essa família, em que, especificando a distribuição *baseline*,  $G(y)$ , da família de distribuições G-exponencializada reparametrizada, obtem-se estas distribuições. A família de distribuições G-exponencializada reparametrizada é uma alternativa à família de distribuições G-exponencializada, quando o interesse consiste em modelar os quantis da distribuição. A família de distribuições G-exponencializada reparametrizada, proposta nesta tese, preenche uma lacuna em teoria de novas distribuições no contexto da abordagem quantílica. Essa proposta apresenta uma abordagem diferenciada e tem o potencial de gerar uma quantidade imensurável de novas distribuições, úteis em diversos contextos.

A regressão quantílica, proposta por Koenker e Jr (1978), é um método que teve um crescimento significativo nos últimos anos. A ideia da regressão quantílica é permitir a modelagem dos quantis condicionais, ao invés da média condicional da variável resposta. A regressão quantílica paramétrica pode ser aplicada em diferentes áreas, tais como, estatística espacial (SÁNCHEZ et al., 2020), análise de sobrevivência (RODRIGUES; BORGES; SANTOS, 2021), confiabilidade (SÁNCHEZ et al., 2021b) e economia (GALLARDO; SANTOS-NETO, 2021; SÁNCHEZ et al., 2021a; SAULO et al., 2022).

Neste trabalho, propomos uma classe geral de modelos de regressão quantílica para distribuições pertencentes à família de distribuições G-exponencializada reparametrizada. Dessa

forma, especificando a distribuição *baseline* da família de distribuições G-exponencializada reparametrizada podemos obter o correspondente modelo de regressão quantílica (RQ) para a distribuição pertencente à família de distribuições G-exponencializada reparametrizada. Por meio dos submodelos da família de distribuições G-exponencializada reparametrizada, pode-se modelar os quantis da distribuição, utilizando conceito similar à dos modelos lineares generalizados (MLGs), ver McCullagh e Nelder (1989) e Sánchez et al. (2021a).

Além disso, de acordo com Koenker (2005), seja  $Y$  uma variável aleatória e, para qualquer  $\tau \in (0, 1)$ , o  $\tau$ -ésimo quantil de  $Y$  é definido da seguinte forma:

$$F^{-1}(\tau) = \inf\{y | F(y) \geq \tau\},$$

sendo  $F(y) = P(Y \leq y)$  a função de distribuição acumulada de  $Y$ . Então, por exemplo, a mediana é dada por  $F^{-1}(1/2)$ , o primeiro quartil e o terceiro quartil são definidos por  $F^{-1}(1/4)$  e  $F^{-1}(3/4)$  respectivamente. Ademais, recentemente Gallardo e Santos-Neto (2021), propuseram um modelo de regressão quantílica baseado no modelo Birnbaum-Saunders exponencializado reparametrizado (BSER). Este modelo de regressão quantílica BSER (GALLARDO; SANTOS-NETO, 2021) é um caso particular de modelo de regressão quantílica pertencente à classe de regressão quantílica baseada na família de distribuições G-exponencializada reparametrizada proposta nesta tese. Santos-Neto et al. (2025) introduziram um modelo de regressão quantílica baseado no modelo Birnbaum-Saunders generalizado exponencializado reparametrizado (BSGER). De modo análogo, o modelo de regressão quantílica proposto por Santos-Neto et al. (2025) é um caso especial da classe geral de modelos de regressão quantílica baseada na família de distribuições G-exponencializada reparametrizada. Portanto, pode-se gerar uma quantidade imensurável de novos modelos de regressão quantílica utilizando esta nova metodologia apresentada nesta tese.

Adicionalmente, os aspectos metodológicos e teóricos apresentados neste estudo generalizam vários trabalhos existentes nas linhas de pesquisa de novas distribuições e regressão quantílica paramétrica. Sendo assim, esse estudo visa contribuir para o desenvolvimento das linhas de pesquisa de novas distribuições e regressão quantílica paramétrica.

---

## 1.2 MOTIVAÇÃO E JUSTIFICATIVA DA TESE

Neste material de tese, uma nova família de distribuições, denominada família de distribuições G-exponencializada reparametrizada, é apresentada. Ademais, são propostos modelos de regressão quantílica, de forma geral, para as distribuições pertencentes à essa família. Espera-se que esta nova abordagem e os modelos propostos encontrem ampla aplicabilidade teórica e prática em diversas áreas de pesquisa, incluindo economia, análise de sobrevivência, confiabilidade, estatística espacial, séries temporais, medicina, entre outras.

Nesta tese, introduzimos a família de distribuições G-exponencializada reparametrizada. Esta é a primeira vez, na literatura, salvo o nosso desconhecimento que um gerador de distribuições foi reparametrizado utilizando essa nova metodologia. A ideia desta nova família reparametrizada é apresentar uma família de distribuições indexada por um parâmetro que representa o quantil de interesse. Logo, por meio da família de distribuições G-exponencializada reparametrizada pode ser introduzido na literatura uma quantidade imensa de novos modelos probabilísticos indexados pelo quantil, em que, selecionando apenas a distribuição *baseline*, da família de distribuições G-exponencializada reparametrizada, obtém-se estes modelos.

Neste trabalho, também são propostos, de forma geral, novos modelos de regressão quantílica baseados em modelos pertencentes à família de distribuições G-exponencializada reparametrizada. Portanto, a importância dos modelos propostos está no fato de que eles são recomendados para modelagem quantílica, uma área promissora da estatística, que se desenvolveu muito nos últimos anos. Portanto, esta tese visa contribuir significativamente para o desenvolvimento da regressão quantílica paramétrica, bem como da linha de pesquisa de novas distribuições, devido aos aspectos metodológicos e teóricos apresentados.

## 1.3 OBJETIVOS

### 1.3.1 Geral

- Sugere-se a proposição da família de distribuições G-exponencializada reparametrizada, bem como a formulação de modelos de regressão quantílica fundamentados nessa família de distribuições.

### 1.3.2 Específicos

- Propor modelos alternativos para família de distribuições G-exponencializada;
- Propor modelos alternativos para regressão quantílica;
- Apresentar a nova família de distribuições G-exponencializada reparametrizada;
- Apresentar novos modelos de regressão quantílica baseados na família de distribuições G-exponencializada reparametrizada;
- Aplicar a teoria proposta através de dados reais ou de simulações;

### 1.4 PRODUTOS DA TESE

Através desta tese podemos obter os seguintes produtos.

- SANTOS, B. Vicente dos; SANTOS-NETO, M.; CYSNEIROS, F. J. A. Reparametrized exponentiated-G distribution family: Properties and applications: Accepted - september 2024. REVSTAT-Statistical Journal, Nov. 2024. Disponível em: <<https://revstat.ine.pt/index.php/REVSTAT/article/view/791>>.
- SANTOS, B. Vicente dos; SANTOS-NETO, M.; CYSNEIROS, F. J. A. Quantile regression for models based on the reparametrized exponentiated-G distribution family (em produção).

### 1.5 ORGANIZAÇÃO DA TESE

A tese está estruturada em capítulos no formato de artigo da seguinte forma: no Capítulo 2, é introduzida a família de distribuições G-exponencializada reparametrizada, juntamente com algumas propriedades estatísticas gerais dessa família. Também são apresentados alguns submodelos dessa família caracterizados pela função de distribuição acumulada (FDA), além do modelo Weibull exponencializado reparametrizado (WER) e algumas de suas principais propriedades estatísticas.

No Capítulo 3, são propostos de forma geral modelos de regressão quantílica baseados em modelos pertencentes à família de distribuições G-exponencializada reparametrizada.

---

Adicionalmente, é apresentado um novo modelo de regressão quantílica baseado no modelo WER, acompanhado de algumas propriedades estatísticas desse modelo. Também são introduzidos alguns resíduos para analisar o ajuste dos modelos propostos ao conjunto de dados e identificar possíveis observações atípicas. Além disso, propõem-se técnicas de análise de diagnósticos com o objetivo de detectar observações potencialmente influentes. No Capítulo 4, são apresentadas as considerações gerais sobre o trabalho e são discutidos possíveis trabalhos futuros a serem desenvolvidos com base neste estudo.

## 1.6 FERRAMENTAS COMPUTACIONAIS

Neste trabalho, utilizou-se a linguagem de programação R em sua versão 3.5.1 R (TEAM, 2020) disponível gratuitamente no endereço [www.r-project.org/](http://www.r-project.org/). Sendo assim, os códigos das simulações e os gráficos foram escritos e gerados, respectivamente, utilizando o software R. Nas simulações de Monte Carlo, utilizou-se o método numérico BFGS (PRESS et al., 1992). Os modelos probabilísticos considerados neste estudo foram ajustados por meio da função `goodness.fit()` do pacote `AdequacyModel` (MARINHO et al., 2019) versão 1.0.8, disponível para a linguagem de programação R em <http://cran.rstudio.com/web/packages/AdequacyModel/index.html> e utilizou-se o método numérico do tipo quase-Newton BFGS. Além disso, os modelos de regressão quantílica analisados neste trabalho foram ajustados utilizando o método numérico BFGS através da função `optim()`. Adicionalmente, utilizou-se para escrever a tese o  $\text{\LaTeX}$  (GRATZER, 1996).

## 2 FAMÍLIA DE DISTRIBUIÇÕES G-EXPONENCIALIZADA REPARAMETRIZADA: PROPRIEDADES E APLICAÇÕES

### 2.1 RESUMO

O objetivo principal deste capítulo é apresentar a família de distribuições G-exponencializada reparametrizada e delinear suas propriedades estatísticas gerais. Em contraste com a distribuição G-exponencializada, esta família reparametrizada é caracterizada por um parâmetro  $\eta_\tau$ , que oferece uma interpretação baseada no  $\tau$ -ésimo quantil da distribuição. Para ilustrar a obtenção de submodelos dentro da família de distribuições G-exponencializada reparametrizada, empregamos o modelo Weibull reparametrizado como *baseline*, resultando no modelo Weibull exponencializado reparametrizado. Através de simulações de Monte Carlo, implementadas utilizando a linguagem R, conduzimos uma avaliação numérica do desempenho dos estimadores de máxima verossimilhança e seus intervalos de confiança assintóticos associados. Além disso, aplicamos a família de distribuições proposta a um conjunto de dados reais para ilustrar a respeito de sua flexibilidade e aplicabilidade em problemas de dados reais.

**Palavras-chave:** Distribuição *baseline*; Distribuição G-exponencializada reparametrizada; Estimador de máxima verossimilhança; Quantil; Simulações de Monte Carlo.

### 2.2 INTRODUÇÃO

Seja  $Y$  uma variável aleatória absolutamente contínua com uma função de distribuição acumulada (FDA) arbitrária, denotada como  $G(y)$ . Se esta FDA for elevada ao expoente  $\alpha$  ( $\alpha > 0$ ), denotada por  $G^\alpha(y)$ , então obtemos uma distribuição G-exponencializada com parâmetro  $\alpha$  (LEHMANN, 1953). Para simplificar a notação, esta distribuição será denotada por  $\alpha$ -G. A função de densidade de probabilidade (FDP) da distribuição  $\alpha$ -G é dada por:

$$f(y) = \alpha g(y) G^{\alpha-1}(y),$$

em que  $g(y) = dG(y)/dy$  corresponde à FDP de  $G(\cdot)$ , chamada de distribuição *baseline*. É importante observar que quando o parâmetro  $\alpha$  assume o valor 1, obtém-se como caso particular o modelo *baseline* da distribuição  $\alpha$ -G.

Vários modelos probabilísticos pertencentes à família de distribuições  $\alpha$ -G foram propostos na literatura. Por exemplo, a distribuição exponencial exponencializada (EE), introduzida em Gupta e Kundu (2001), estende a distribuição exponencial e encontra aplicação na análise de dados de sobrevivência nos domínios médico e de engenharia. Da mesma forma, a distribuição Weibull exponencializada (WE), proposta em Mudholkar e Srivastava (1993), estende a distribuição Weibull e oferece versatilidade na representação de funções de taxa de falha.

Além disso, as distribuições Fréchet exponencializada (FE) (NADARAJAH; KOTZ, 2003), Rayleigh exponencializada (RE) (KUNDU; RAQAB, 2005), Birnbaum-Saunders exponencializada (BSE) (CORDEIRO; LEMONTE; ORTEGA, 2013), Lomax exponencializada (LE) (ABDUL-MONIEM; ABDEL-HAMEED, 2012), Gompertz exponencializada (GE) (EL-GOHARY; AHMAD; ADEL, 2013), Gumbel exponencializada (GUE) (NADARAJAH, 2006), log-logística exponencializada (LLE) (ROSAIAH; KANTAM; KUMAR, 2006) e Birnbaum-Saunders generalizada exponencializada (BSGE) (GALLARDO; SANTOS-NETO, 2021) foram introduzidas para generalizar suas respectivas distribuições *baseline* e explorar suas propriedades estatísticas. Esses modelos encontram aplicações em vários campos na modelagem estatística como por exemplo, na economia, análise de confiabilidade e estudos de sobrevivência. Além disso, estudos investigaram suas propriedades teóricas e as aplicaram a conjuntos de dados reais.

Neste capítulo, foi apresentada a família de distribuições  $\alpha$ -G reparametrizada e delineou-se suas propriedades estatísticas gerais. Esta família mostra-se valiosa para modelar quantis de distribuição, pois é indexada por um parâmetro  $\eta_\tau$ , representando o quantil de interesse. Nossa proposta generalizará uma ampla variedade de distribuições existentes, proporcionando, assim, flexibilidade significativa para modelagem de dados.

Este capítulo está estruturado da seguinte forma: a Seção 2.3, apresenta a FDP e a FDA da família de distribuições  $\alpha$ -G reparametrizada, juntamente com seus submodelos obtidos da alteração da distribuição *baseline*,  $G(y)$ . Também apresenta-se a função de sobrevivência,  $S(y)$ , a função de taxa de falha,  $h(y)$ , e a função de taxa de falha acumulada,  $H(y)$ , para esta família. Além disso, o processo de estimação dos parâmetros foi delineado usando máxima verossimilhança e apresentando propriedades estatísticas gerais aplicáveis a vários submodelos selecionando diferentes distribuições *baseline*. Na Seção 2.4, foi revisado o modelo Weibull e suas principais propriedades estatísticas, incluindo uma reparametrização com base na mediana. Na Seção 2.5, é introduzido um caso específico da família de distribuições  $\alpha$ -G

reparametrizada, conhecido como modelo Weibull exponencializado reparametrizado (WER), e detalhamos o procedimento de estimação dos parâmetros. A Seção 2.6 apresenta simulações de Monte Carlo para avaliar o desempenho dos estimadores de máxima verossimilhança para o modelo WER. Também avalia-se numericamente os intervalos de confiança assintóticos. Na Seção 2.7 realizou-se um ajuste a um conjunto de dados reais. Na Seção 2.8, apresentamos as observações finais.

### 2.3 FAMÍLIA DE DISTRIBUIÇÕES $\alpha$ -G REPARAMETRIZADA

Se a distribuição *baseline* for indexada por sua mediana ou puder ser reparametrizada em termos da mediana,  $\eta$ , e se  $\tau$  ( $0 < \tau < 1$ ) for a ordem do quantil de interesse, então o parâmetro  $\alpha$  da distribuição  $\alpha$ -G pode ser reescrito da seguinte forma (GALLARDO; SANTOS-NETO, 2021):

$$\begin{aligned} [G(\eta)]^\alpha &= \tau \\ \left(\frac{1}{2}\right)^\alpha &= \tau \\ \log\left(\frac{1}{2}\right)^\alpha &= \log(\tau) \\ -\alpha \log(2) &= \log(\tau) \\ \alpha &= -\frac{\log(\tau)}{\log(2)}, \end{aligned}$$

em que  $G(\cdot)$  representa a FDA da distribuição *baseline*, e  $\alpha = \alpha_\tau$  passa a ser um valor conhecido que depende de  $\tau$ . Assim, o parâmetro  $\eta_\tau = \eta$  representa o  $\tau$ -ésimo quantil dessa família de distribuições.

Agora, a FDP e FDA da família de distribuições  $\alpha$ -G reparametrizada são dadas, respectivamente, por:

$$f(y) = \alpha_\tau g(y) G^{\alpha_\tau - 1}(y) \quad (2.1)$$

e

$$F(y) = G^{\alpha_\tau}(y). \quad (2.2)$$

Denotaremos essa família de distribuições como  $\alpha_\tau$ -G. Observe que a família de distribuições  $\alpha_\tau$ -G reparametrizada inclui todos os submodelos da família de distribuições  $\alpha$ -G que são indexados pela mediana ou cuja mediana possui uma expressão em forma

fechada. Adicionalmente, podemos observar que quando  $\tau$  assume o valor 0,5, obtemos como caso especial o modelo *baseline* da família de distribuições  $\alpha_\tau$ -G reparametrizada.

Adicionalmente, as funções de sobrevivência,  $S(y)$ , de taxa de falha,  $h(y)$ , e de taxa de falha acumulada,  $H(y)$ , da família de distribuições  $\alpha_\tau$ -G reparametrizada são dadas, respectivamente, por:

$$S(y) = 1 - F(y) = 1 - G^{\alpha_\tau}(y), \quad h(y) = \alpha_\tau g(y) G^{\alpha_\tau - 1}(y) [1 - G^{\alpha_\tau}(y)]^{-1}$$

e

$$H(y) = -\log[1 - G^{\alpha_\tau}(y)].$$

Na Tabela 1 são apresentados alguns submodelos da família de distribuições  $\alpha_\tau$ -G reparametrizada. Esses submodelos são obtidos a partir das Equações (2.1) e (2.2), substituindo  $G(y)$  e  $g(y)$ , que correspondem à função de distribuição acumulada (FDA) e à função de densidade de probabilidade (FDP), respectivamente, da distribuição *baseline*. Por exemplo, o modelo WER é obtido ao considerar a distribuição Weibull reparametrizada como *baseline*.

### 2.3.1 Estimação via Método de Máxima Verossimilhança

Seja  $(y_1, \dots, y_n)$  uma amostra aleatória de tamanho  $n$  da família de distribuições  $\alpha_\tau$ -G reparametrizada, e seja  $\boldsymbol{\theta}$  o vetor de parâmetros de dimensão  $q$ , em que  $\eta_\tau \in \boldsymbol{\theta}$ . A função de verossimilhança para essa família pode ser expressa como:

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{i=1}^n f(y_i; \boldsymbol{\theta}) = \prod_{i=1}^n \alpha_\tau g(y_i; \boldsymbol{\theta}) G^{\alpha_\tau - 1}(y_i; \boldsymbol{\theta}) \\ &= (\alpha_\tau)^n \prod_{i=1}^n g(y_i; \boldsymbol{\theta}) \prod_{i=1}^n G^{\alpha_\tau - 1}(y_i; \boldsymbol{\theta}). \end{aligned}$$

O logaritmo da função de verossimilhança,  $\ell(\boldsymbol{\theta}) = \log[L(\boldsymbol{\theta})]$ , é dado por:

$$\ell(\boldsymbol{\theta}) = n \log(\alpha_\tau) + \sum_{i=1}^n \log g(y_i; \boldsymbol{\theta}) + (\alpha_\tau - 1) \sum_{i=1}^n \log G(y_i; \boldsymbol{\theta}).$$

O vetor escore é definido por:

$$\mathbf{U}(\boldsymbol{\theta}) = \left( \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^\top,$$

e seus elementos, para  $j = 1, \dots, q$ , são:

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_j} = \sum_{i=1}^n \frac{1}{g(y_i; \boldsymbol{\theta})} \frac{\partial g(y_i; \boldsymbol{\theta})}{\partial \theta_j} + (\alpha_\tau - 1) \sum_{i=1}^n \frac{1}{G(y_i; \boldsymbol{\theta})} \frac{\partial G(y_i; \boldsymbol{\theta})}{\partial \theta_j}.$$

Os estimadores de máxima verossimilhança (EMVs) de  $\boldsymbol{\theta}$  são as soluções do sistema de equações  $\mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}$ . No entanto, nesse caso, o sistema não possui solução em forma fechada. Portanto, os EMVs devem ser obtidos por métodos numéricos. Aqui, empregamos o método quase-Newton, especificamente o algoritmo BFGS, conforme descrito em Press et al. (1992). Na linguagem de programação R (TEAM, 2020), o método BFGS pode ser implementado por meio da função `optim()`.

A matriz de informação de Fisher correspondente é uma matriz  $q \times q$  denotada por:

$$I(\boldsymbol{\theta}) = \mathbb{E} \left[ \mathbf{U}(\boldsymbol{\theta}) \mathbf{U}(\boldsymbol{\theta})^\top \right].$$

Se certas condições de regularidade (COX; HINKLEY, 1974) para a função de verossimilhança forem satisfeitas, então

$$\mathbb{E} [\mathbf{U}(\boldsymbol{\theta})] = \mathbf{0}$$

e

$$I(\boldsymbol{\theta}) = -\mathbb{E} \left[ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right].$$

Portanto, seus elementos podem ser expressos como:

$$\begin{aligned} \mathbb{E} \left[ -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right] &= n \mathbb{E} \left[ \frac{1}{g^2(Y; \boldsymbol{\theta})} \frac{\partial g(Y; \boldsymbol{\theta})}{\partial \theta_i} \frac{\partial g(Y; \boldsymbol{\theta})}{\partial \theta_j} \right] \\ &\quad + n(\alpha_\tau - 1) \mathbb{E} \left[ \frac{1}{G^2(Y; \boldsymbol{\theta})} \frac{\partial G(Y; \boldsymbol{\theta})}{\partial \theta_i} \frac{\partial G(Y; \boldsymbol{\theta})}{\partial \theta_j} \right] \\ &\quad - n \mathbb{E} \left[ \frac{1}{g(Y; \boldsymbol{\theta})} \frac{\partial^2 g(Y; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right] - n \mathbb{E} \left[ \frac{1}{G(Y; \boldsymbol{\theta})} \frac{\partial^2 G(Y; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right]. \end{aligned}$$

De acordo com Lehmann e Casella (1998), quando certas condições de regularidade são satisfeitas, os EMVs  $\hat{\boldsymbol{\theta}}$  seguem, assintoticamente, uma distribuição normal multivariada de dimensão  $q$ . Em outras palavras:

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \stackrel{a}{\sim} N_q(\mathbf{0}, I^{-1}(\boldsymbol{\theta})),$$

onde  $\stackrel{a}{\sim}$  denota “assintoticamente distribuído como”. Esse resultado pode ser utilizado para construir intervalos de confiança assintóticos e realizar testes de hipóteses para os parâmetros da família de distribuições  $\alpha_\tau$ -G reparametrizada.

---

Adicionalmente, quando for difícil obter a matriz de informação de Fisher  $I(\boldsymbol{\theta})$ , podemos substituí-la pela matriz de informação de Fisher observada,

$$-\ddot{L}(\boldsymbol{\theta}) = - \left[ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right],$$

avaliada em  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ .

Os intervalos de confiança assintóticos com nível de confiança  $(1 - \gamma) \times 100\%$  para o parâmetro  $\theta_j$  são dados por:

$$\hat{\theta}_j \pm z_{(1-\gamma/2)} \sqrt{v_j},$$

onde  $z_{(1-\gamma/2)}$  representa o  $(1 - \gamma/2)$ -ésimo quantil da distribuição normal padrão. A variância assintótica  $v_j$  de  $\hat{\theta}_j$  corresponde ao  $j$ -ésimo elemento da diagonal da matriz  $I(\boldsymbol{\theta})^{-1}$ , avaliada em  $\hat{\boldsymbol{\theta}}$ .

Tabela 1: Alguns submodelos da família de distribuições  $\alpha_\tau$ -G reparametrizada.

Modelos	$F(y)$	$\eta$	Reparametrização
EER	$(1 - 2^{-\frac{y}{\eta_\tau}})^{\alpha_\tau}$	$\beta \log(2)$	$\beta = \eta / \log(2)$
WER	$(1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}})^{\alpha_\tau}$	$\beta [\log(2)]^{1/\kappa}$	$\beta = \eta / [\log(2)]^{\frac{1}{\kappa}}$
FER	$\left\{ \exp \left[ - \left( \frac{\eta_\tau \sqrt[\lambda]{\ln(2)}}{y} \right)^\lambda \right] \right\}^{\alpha_\tau}$	$\frac{\sigma}{\sqrt[\lambda]{\ln(2)}}$	$\sigma = \eta \sqrt[\lambda]{\ln(2)}$
RER	$\left[ 1 - \exp \left( - \frac{y^2 \ln(2)}{\eta_\tau^2} \right) \right]^{\alpha_\tau}$	$\sigma \sqrt{2 \ln(2)}$	$\sigma = \frac{\eta}{\sqrt{2 \ln(2)}}$
GER	$\left\{ 1 - \exp \left[ - \lambda \left( \frac{y}{\eta_\tau} \right) \ln \left[ - \left( \frac{1}{\lambda} \right) \ln \left( \frac{1}{2} \right) + 1 \right] - 1 \right] \right\}^{\alpha_\tau}$	$\left( \frac{1}{\theta} \right) \ln \left[ - \left( \frac{1}{\lambda} \right) \ln \left( \frac{1}{2} \right) + 1 \right]$	$\theta = \left( \frac{1}{\eta} \right) \ln \left[ - \left( \frac{1}{\lambda} \right) \ln \left( \frac{1}{2} \right) + 1 \right]$
LER	$\left\{ 1 - \left[ 1 + \frac{y (\sqrt[2]{2} - 1)}{\eta_\tau} \right]^{-\kappa} \right\}^{\alpha_\tau}$	$\lambda (\sqrt[2]{2} - 1)$	$\lambda = \frac{\eta}{\sqrt[2]{2} - 1}$
LLER	$\left\{ 1 - \left[ 1 + \left( \frac{y}{\eta_\tau} \right)^{\kappa-1} \right] \right\}^{\alpha_\tau}$	$\gamma$	$\gamma = \eta$
BSER	$\left\{ \Phi \left[ \frac{1}{\lambda} \left( \sqrt{\frac{y}{\eta_\tau}} - \sqrt{\frac{\eta_\tau}{y}} \right) \right] \right\}^{\alpha_\tau}$	$\beta$	$\beta = \eta$
BSGER	$\left\{ \Phi \left[ \frac{1}{\lambda} \left( \frac{y^{1-k}}{\sqrt{\eta_\tau}} - \frac{\sqrt{\eta_\tau}}{y^k} \right) \right] \right\}^{\alpha_\tau}$	$\beta$	$\beta = \eta$
BXIIER	$\left\{ 1 - \left[ 1 + \left( \frac{y \left( (2)^{\frac{1}{k}} - 1 \right)^{\frac{1}{c}}}{\eta_\tau} \right)^{c-k} \right] \right\}^{\alpha_\tau}$	$s \left( (2)^{\frac{1}{k}} - 1 \right)^{\frac{1}{c}}$	$s = \eta / \left( (2)^{\frac{1}{k}} - 1 \right)^{\frac{1}{c}}$

## 2.4 DISTRIBUIÇÃO WEIBULL REPARAMETRIZADA PELA MEDIANA

Nesta seção, a distribuição *baseline* para a distribuição WER é apresentada. A distribuição Weibull, proposta por Weibull (1939), é um modelo probabilístico indexado por um parâmetro de forma ( $\kappa > 0$ ) e um parâmetro de escala ( $\beta > 0$ ). Sua FDP é dada por:

$$g(y; \kappa, \beta) = \frac{\kappa}{\beta^\kappa} y^{(\kappa-1)} \exp \left[ - \left( \frac{y}{\beta} \right)^\kappa \right] \mathbb{I}_{\{y>0\}},$$

em que  $\mathbb{I}_{\{y>0\}}$  é a função indicadora que vale 1 quando  $y > 0$  e 0 caso contrário. A FDA é definida como:

$$G(y; \kappa, \beta) = \{1 - \exp[-(y/\beta)^\kappa]\} \mathbb{I}_{\{y>0\}}.$$

Além disso, sua média e variância são:

$$\mathbb{E}(Y) = \beta \Gamma \left( \frac{1}{\kappa} + 1 \right) \quad \text{e} \quad \text{Var}(Y) = \beta^2 \left[ \Gamma \left( 1 + \frac{2}{\kappa} \right) - \Gamma^2 \left( \frac{1}{\kappa} + 1 \right) \right],$$

respectivamente. Em que,  $\Gamma(\cdot)$  é a função gama. A mediana da distribuição Weibull é  $\eta = \beta[\log(2)]^{1/\kappa}$ .

Reparametrizando a distribuição Weibull em termos de mediana, utilizou-se uma nova parametrização em que  $\beta = \eta/[\log(2)]^{1/\kappa}$ . Assim, a FDP da distribuição Weibull reparametrizada é dada por:

$$g(y; \kappa, \eta) = \frac{\kappa \log(2) y^{(\kappa-1)} 2^{-\left(\frac{y}{\eta}\right)^\kappa}}{\eta^\kappa} \mathbb{I}_{\{y>0\}}, \quad (2.3)$$

e sua FDA é:

$$G(y; \kappa, \eta) = \left[ 1 - 2^{-\left(\frac{y}{\eta}\right)^\kappa} \right] \mathbb{I}_{\{y>0\}}. \quad (2.4)$$

## 2.5 DISTRIBUIÇÃO WER

A FDP e FDA da distribuição WER são obtidas aplicando as Equações (2.3) e (2.4) nas Equações (2.1) e (2.2), respectivamente. Consequentemente,

$$f(y; \kappa, \eta_\tau, \alpha_\tau) = \frac{\log(2) \alpha_\tau \kappa y^{\kappa-1} 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}}{\eta_\tau^\kappa} \left( 1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}} \right)^{\alpha_\tau-1} \mathbb{I}_{\{y>0\}},$$

e

$$F(y; \kappa, \eta_\tau, \alpha_\tau) = \left(1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}\right)^{\alpha_\tau} \mathbb{I}_{\{y>0\}},$$

respectivamente. Nota-se que, quando  $\kappa = 1$ , obtém-se como caso especial a distribuição EER. Vamos verificar se a distribuição WER é identificável. Sabe-se da literatura que, a identificabilidade é definida considerando que diferentes valores dos parâmetros geram funções de densidades diferentes. Então para o modelo WER se:

$$f(y; \kappa, \eta_\tau, \alpha_\tau) = f(y; \kappa', \eta'_\tau, \alpha'_\tau), \quad \mathbb{I}_{\{y>0\}},$$

consequentemente, o conjunto de parâmetros  $(\kappa, \eta_\tau, \alpha_\tau) = (\kappa', \eta'_\tau, \alpha'_\tau)$ . Supondo que:

$$f(y; \kappa, \eta_\tau, \alpha_\tau) = f(y; \kappa', \eta'_\tau, \alpha'_\tau), \quad \mathbb{I}_{\{y>0\}}.$$

Então,

$$\frac{\alpha_\tau \kappa y^{\kappa-1} 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}}{\eta_\tau^\kappa} \left(1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}\right)^{\alpha_\tau-1} = \frac{\alpha'_\tau \kappa' y^{\kappa'-1} 2^{-\frac{y^{\kappa'}}{\eta'^{\kappa'}}}}{\eta'^{\kappa'}} \left(1 - 2^{-\frac{y^{\kappa'}}{\eta'^{\kappa'}}}\right)^{\alpha'_\tau-1}.$$

Verificando os expoentes de  $y$  nota-se que, a igualdade acima é verdadeira se para todo  $y > 0$ , os expoentes dominantes de  $y$  precisam ser iguais. Observa-se que,  $y^{\kappa-1}$  depende de  $\kappa$  e o termo  $2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}$  decresce exponencialmente para  $y^\kappa$ . Então, se  $\kappa \neq \kappa'$  os comportamentos assintóticos de  $y^{\kappa-1}$  e  $y^{\kappa'-1}$  são diferentes para  $y \rightarrow \infty$  isto é uma contradição com a igualdade de densidades acima, logo,  $\kappa = \kappa'$ .

Agora vamos verificar se  $\eta_\tau = \eta'_\tau$ , já temos que  $\kappa = \kappa'$ , logo, a igualdade de densidades pode ser reescrita como:

$$\frac{\alpha_\tau 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}}{\eta_\tau^\kappa} \left(1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}\right)^{\alpha_\tau-1} = \frac{\alpha'_\tau 2^{-\frac{y^\kappa}{\eta'^{\kappa'}}}}{\eta'^{\kappa'}} \left(1 - 2^{-\frac{y^\kappa}{\eta'^{\kappa'}}}\right)^{\alpha'_\tau-1},$$

Nota-se que,  $2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}$  e  $2^{-\frac{y^\kappa}{\eta'^{\kappa'}}$  são iguais para todo  $y > 0$ , quando os expoentes são iguais, ou seja:

$$\frac{y^\kappa}{\eta_\tau^\kappa} = \frac{y^\kappa}{\eta'^{\kappa'}},$$

isto implica que  $\eta_\tau = \eta'_\tau$ . Já tem-se que,  $\kappa = \kappa'$  e  $\eta_\tau = \eta'_\tau$ , dessa forma a igualdade de densidades pode ser reescrita como:

$$\alpha_\tau \left(1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}\right)^{\alpha_\tau-1} = \alpha'_\tau \left(1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}\right)^{\alpha'_\tau-1},$$

multiplicando ambos os lados da equação acima por:

$$\left(1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}\right),$$

logo a igualdade se reduz a:

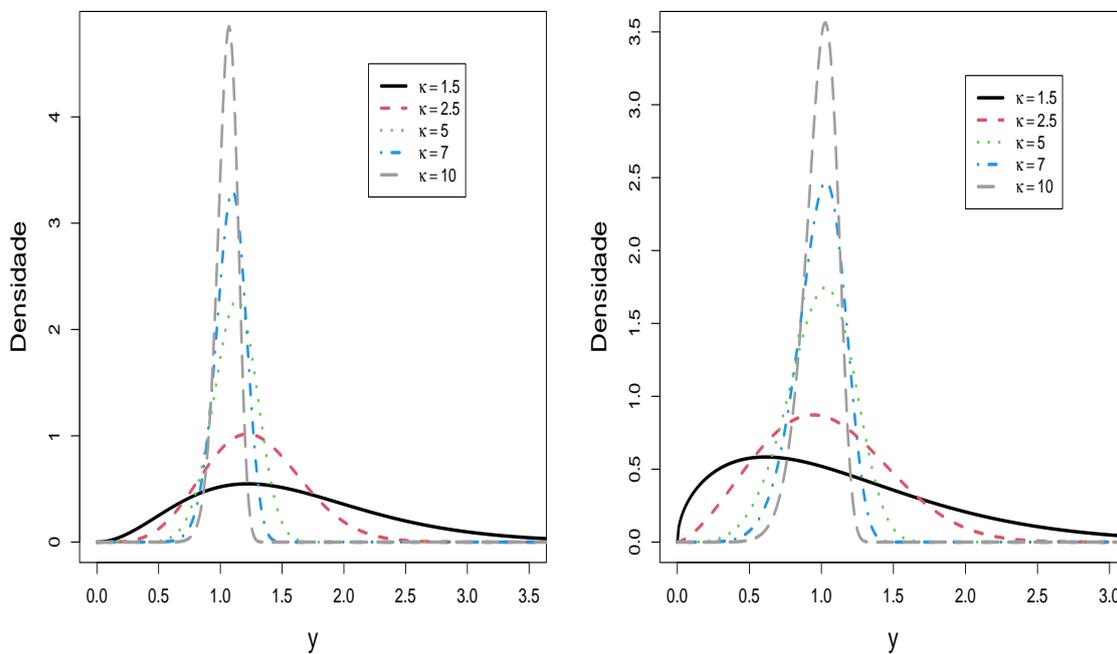
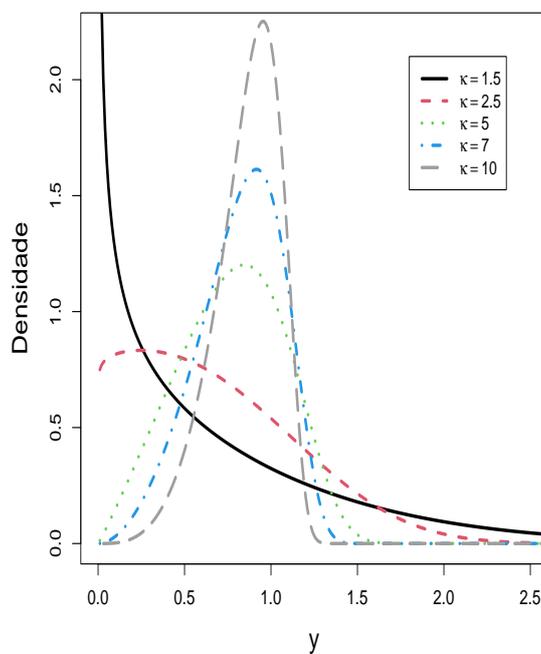
$$\alpha_\tau \left(1 - 2^{-\frac{y^\kappa}{\eta_\tau^\kappa}}\right)^{\alpha_\tau} = \alpha'_\tau \left(1 - 2^{-\frac{y^{\kappa'}}{\eta_{\tau'}^{\kappa'}}}\right)^{\alpha'_\tau},$$

para a igualdade acima ser verdadeira para todo  $y > 0$ , isto implica que,  $\alpha_\tau$  tem que ser igual a  $\alpha'_\tau$ . Portanto, provamos que  $\kappa = \kappa'$ ,  $\eta_\tau = \eta'_\tau$  e  $\alpha_\tau = \alpha'_\tau$ , então o modelo WER é identificável.

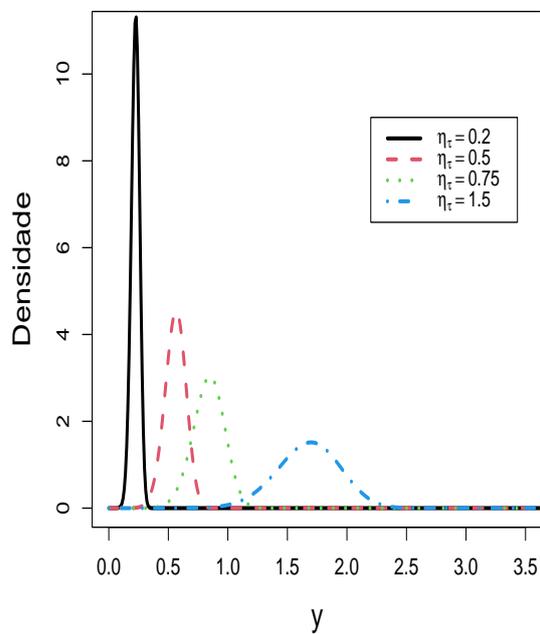
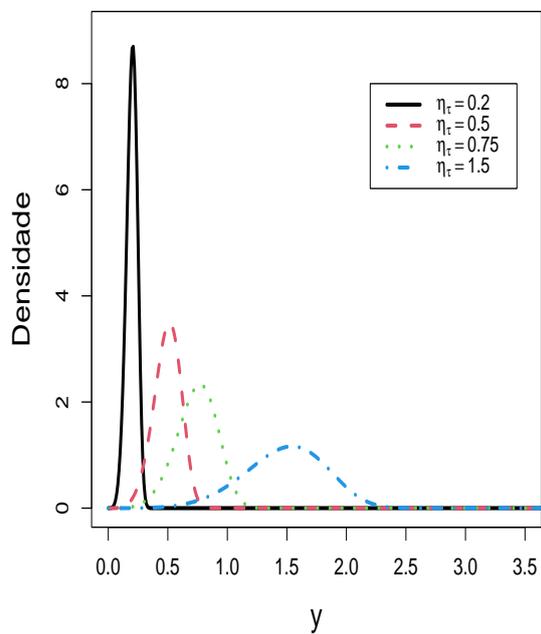
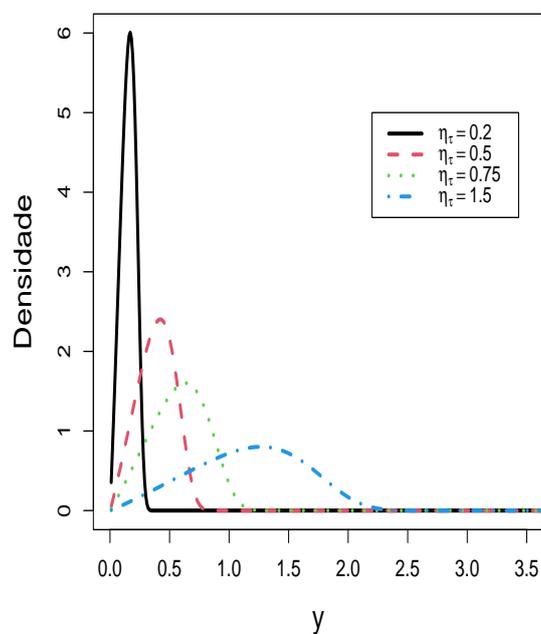
Na Figura 1 são exibidas as curvas de densidade da distribuição WER para vários valores do parâmetro  $\kappa$ : 1,5, 2,5, 5, 7 e 10 e  $\tau$ : 0,25, 0,5 e 0,75, com  $\eta_\tau = 1$ . O parâmetro que representar o quantil,  $\eta_\tau$ , foi fixado em 1, para observar como mudanças no parâmetro  $\kappa$  e em  $\tau$  afetam a densidade do modelo WER. Observe que o parâmetro  $\kappa$  controla o nível de curtose da distribuição. Quanto maior o valor de  $\kappa$ , mais pontiaguda se torna a curva. Além disso,  $\kappa$  é um parâmetro de forma. Logo, mudando os valores de  $\kappa$ , a forma da distribuição também muda. Por outro lado,  $\tau$  parece influenciar o grau de assimetria da distribuição. Aumentando o valor de  $\tau$  aumenta a assimetria da curva. Nota-se também que  $\eta_\tau$  foi fixado em 1, pois este parâmetro possui uma interpretação, ou seja, representa o quantil da distribuição. Por exemplo, para o valor de  $\tau = 0,5$ , à medida que  $\kappa$  aumenta, a distribuição torna-se mais simétrica em torno do parâmetro  $\eta_\tau$ .

Na Figura 2 são apresentadas as curvas de densidade da distribuição WER para vários valores do parâmetro  $\eta_\tau$ : 0,2, 0,5, 0,75 e 1,5 e  $\tau$ : 0,25, 0,5 e 0,75, com o parâmetro de forma  $\kappa = 5$ , que indica forma unimodal. Nota-se que, o aumento do valor de  $\eta_\tau$  esta deslocando a densidade para a direita. Este resultado já era esperado, uma vez que, o parâmetro  $\eta_\tau$  representa o quantil da distribuição WER.

A Figura 3 apresenta as curvas de densidade da distribuição WER para vários valores do parâmetro  $\kappa$ : 0,1, 0,5 e 0,95 e  $\tau$ : 0,25, 0,5 e 0,75, com  $\eta_\tau = 1$ . Percebe-se, nesta figura, de modo geral, que, para valores do parâmetro de forma  $\kappa$  menores que 1, a densidade apresenta uma forma decrescente.

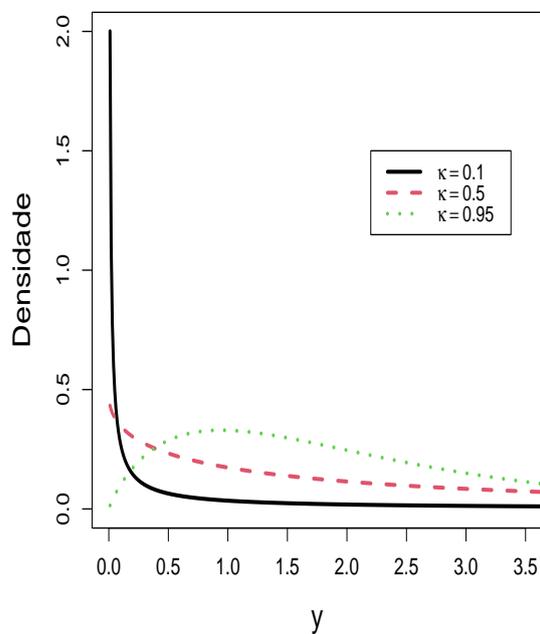
Figura 1 – Densidades WER para diferentes valores de  $\kappa$  e  $\tau$  com  $\eta_\tau = 1$  fixo.(a)  $\tau = 0,25$ (b)  $\tau = 0,50$ (c)  $\tau = 0,75$ 

Fonte: O autor (2025).

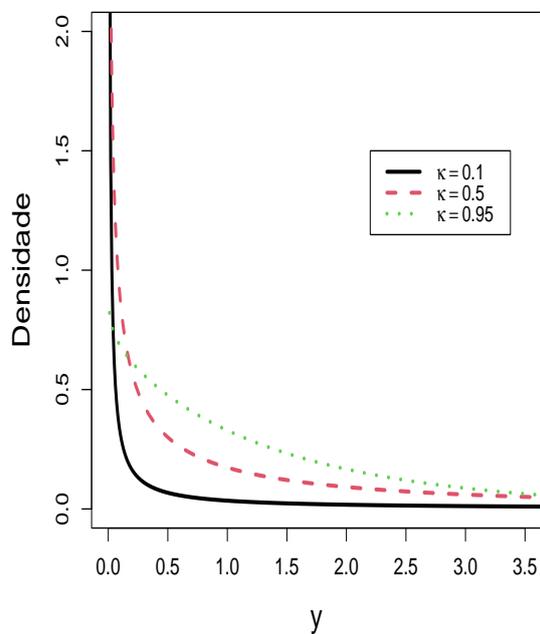
Figura 2 – Densidades WER para diferentes valores de  $\eta_\tau$  e  $\tau$  com  $\kappa = 5$  fixo.(a)  $\tau = 0,25$ (b)  $\tau = 0,50$ (c)  $\tau = 0,75$ 

Fonte: O autor (2025).

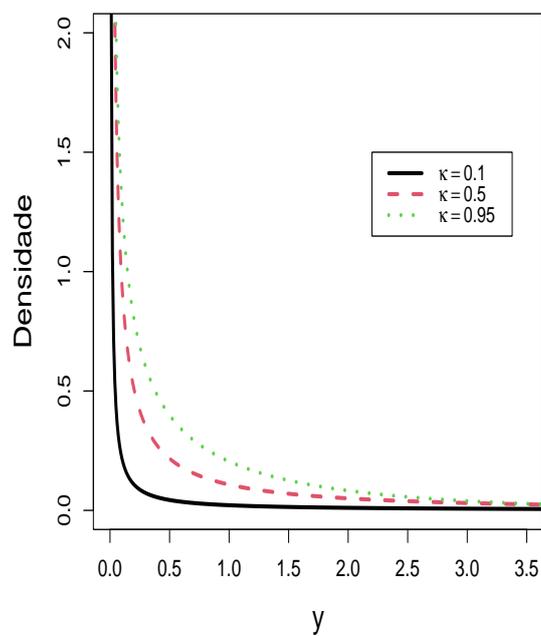
Figura 3 – Densidades WER para diferentes valores de  $\kappa$  e  $\tau$  considerando valores de  $\kappa$  menores que 1, para  $\eta_\tau = 1$  fixo.



(a)  $\tau = 0,25$



(b)  $\tau = 0,50$



(c)  $\tau = 0,75$

Fonte: O autor (2025).

### 2.5.1 Estimação via Método de Máxima Verossimilhança

Seja  $(y_1, \dots, y_n)$  uma amostra aleatória de tamanho  $n$  da distribuição WER com vetor de parâmetros  $\boldsymbol{\theta} = (\kappa, \eta_\tau, \alpha_\tau)^\top$ , em que,  $\alpha_\tau$  é conhecido. A função de verossimilhança pode ser escrita como:

$$L(\boldsymbol{\theta}) = \frac{[\log(2)]^n (\alpha_\tau)^n \kappa^n (\prod_{i=1}^n y_i)^{\kappa-1} 2^{-\frac{\sum_{i=1}^n y_i^\kappa}{\eta_\tau^\kappa}}}{\eta_\tau^{n\kappa}} \prod_{i=1}^n \left(1 - 2^{-\frac{y_i^\kappa}{\eta_\tau^\kappa}}\right)^{\alpha_\tau-1}.$$

O logaritmo da função de verossimilhança  $\ell(\kappa, \eta_\tau, \alpha_\tau) = \log [L(\kappa, \eta_\tau, \alpha_\tau)]$  é dado por:

$$\begin{aligned} \ell(\kappa, \eta_\tau, \alpha_\tau | y_i) = & n \log [\log(2)] + n \log (\alpha_\tau) + n \log(\kappa) + (\kappa - 1) \sum_{i=1}^n \log (y_i) \\ & - \frac{\log(2)}{\eta_\tau^\kappa} \sum_{i=1}^n y_i^\kappa - n\kappa \log(\eta_\tau) + (\alpha_\tau - 1) \sum_{i=1}^n \log \left(1 - 2^{-\frac{y_i^\kappa}{\eta_\tau^\kappa}}\right). \end{aligned}$$

A função escore para  $\boldsymbol{\theta}$  é definida por:

$$\mathbf{U}(\boldsymbol{\theta}) = \left( \frac{\partial \ell(\kappa, \eta_\tau, \alpha_\tau | y_i)}{\partial \kappa}, \frac{\partial \ell(\kappa, \eta_\tau, \alpha_\tau | y_i)}{\partial \eta_\tau} \right)^\top,$$

em que,

$$\begin{aligned} \frac{\partial \ell(\kappa, \eta_\tau, \alpha_\tau | y_i)}{\partial \kappa} = & \frac{n}{\kappa} + \sum_{i=1}^n \log (y_i) - n \log(\eta_\tau) + \log(2) \eta_\tau^{-\kappa} \sum_{i=1}^n y_i^\kappa [\ln(\eta_\tau) - \ln(y_i)] \\ & + (\alpha_\tau - 1) \sum_{i=1}^n \left[ \frac{\ln(2) \left(\frac{y_i}{\eta_\tau}\right)^\kappa \ln \left(\frac{y_i}{\eta_\tau}\right)}{2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} - 1} \right] \end{aligned}$$

e

$$\begin{aligned} \frac{\partial \ell(\kappa, \eta_\tau, \alpha_\tau | y_i)}{\partial \eta_\tau} = & \log(2) \kappa \eta_\tau^{-(\kappa+1)} \sum_{i=1}^n y_i^\kappa - \frac{n\kappa}{\eta_\tau} \\ & + (\alpha_\tau - 1) \sum_{i=1}^n \left[ \frac{\ln(2) y_i^\kappa \kappa \eta_\tau^{-(\kappa+1)}}{1 - 2^{-\frac{y_i^\kappa}{\eta_\tau^\kappa}}} \right]. \end{aligned}$$

Os EMVs para os parâmetros do modelo WER são soluções do sistema de equações  $\mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}$ . Porém, este sistema de equações não possui solução em forma fechada. Sendo assim, as estimativas de máxima verossimilhança dos parâmetros que indexam o modelo WER foram obtidas com o software R, utilizando o método BFGS através da função `optim()`.

Os elementos da matriz  $\ddot{L}$  correspondente são:

$$\begin{aligned} \frac{\partial^2 \ell(\kappa, \eta_\tau, \alpha_\tau | y_i)}{\partial \kappa^2} &= -\frac{n}{\kappa^2} - \log(2) \sum_{i=1}^n \left[ \ln \left( \frac{\eta_\tau}{y_i} \right) \right]^2 \left( \frac{y_i}{\eta_\tau} \right)^\kappa \\ &\quad + (\alpha_\tau - 1) \sum_{i=1}^n \left\{ \ln(2) \left[ \ln \left( \frac{y_i}{\eta_\tau} \right) \right]^2 \left( \frac{y_i}{\eta_\tau} \right)^\kappa \left( 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} - 1 \right) \right. \\ &\quad \left. - [\ln(2)]^2 \left( \frac{y_i}{\eta_\tau} \right)^{2\kappa} \left[ \ln \left( \frac{y_i}{\eta_\tau} \right) \right]^2 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} \right\} \times \frac{1}{\left( 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} - 1 \right)^2}, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\kappa, \eta_\tau, \alpha_\tau | y_i)}{\partial \eta_\tau^2} &= \frac{n\kappa}{\eta_\tau^2} - \log(2) \sum_{i=1}^n y_i^\kappa \kappa (\kappa + 1) \eta_\tau^{-(\kappa+2)} \\ &\quad + (\alpha_\tau - 1) \sum_{i=1}^n \left\{ -\ln(2) y_i^\kappa \kappa (\kappa + 1) \eta_\tau^{-(\kappa+2)} \left( 1 - 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} \right) \right. \\ &\quad \left. - [\ln(2)]^2 (y_i)^{2\kappa} \kappa^2 \eta_\tau^{-2(\kappa+1)} 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} \right\} \times \frac{1}{\left( 1 - 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} \right)^2}, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\kappa, \eta_\tau, \alpha_\tau | y_i)}{\partial \eta_\tau \partial \kappa} &= -\frac{n}{\eta_\tau} + \log(2) \sum_{i=1}^n y_i^\kappa \eta_\tau^{-(\kappa+1)} \left[ -\kappa \ln \left( \frac{\eta_\tau}{y_i} \right) + 1 \right] \\ &\quad + (\alpha_\tau - 1) \sum_{i=1}^n \left\{ -\ln(2) y_i^\kappa \eta_\tau^{-(\kappa+1)} \left[ \kappa \ln \left( \frac{y_i}{\eta_\tau} \right) + 1 \right] \times \right. \\ &\quad \left( 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} - 1 \right) + [\ln(2)]^2 \left( \frac{y_i}{\eta_\tau} \right)^\kappa \ln \left( \frac{y_i}{\eta_\tau} \right) \times \\ &\quad \left. \left( 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} \right) y_i^\kappa \kappa \eta_\tau^{-(\kappa+1)} \right\} \times \frac{1}{\left( 2^{\frac{y_i^\kappa}{\eta_\tau^\kappa}} - 1 \right)^2}. \end{aligned}$$

## 2.6 SIMULAÇÃO

Nesta seção, são apresentados as análises e os resultados numéricos das simulações de Monte Carlo. Os códigos das simulações foram escritos usando a linguagem de programação R em sua versão 3.5.1 R (TEAM, 2020), disponível gratuitamente no endereço <[www.r-project.org/](http://www.r-project.org/)>. Nas simulações de Monte Carlo, utilizou-se o método numérico BFGS (PRESS et al., 1992) e a função `optim()`. Neste trabalho, foi usado o método BFGS, pois ele é bastante aplicado na literatura para maximizar numericamente o logaritmo da função de verossimilhança. De acordo com Cordeiro, Ortega e Silva (2011) os chutes iniciais para

os parâmetros podem ser obtidos do ajuste do modelo *baseline*. Portanto, os chutes iniciais para o parâmetro de forma  $\kappa$  do modelo WER foram dados pelas estimativas do método de momentos da distribuição *baseline* Weibull, e para o parâmetro  $\eta_\tau$ , os chutes iniciais foram dados pelo quantil do modelo Weibull. As medidas utilizadas para avaliar o comportamento das estimativas pontuais obtidas pelo método de máxima verossimilhança para distribuição WER foram: viés relativo percentual em valor absoluto (VR, em que  $VR_\kappa$  e  $VR_{\eta_\tau}$  são os vieses relativos percentual em valor absoluto para  $\kappa$  e  $\eta_\tau$ , respectivamente), variância (VAR, com  $VAR_\kappa$  e  $VAR_{\eta_\tau}$  são as variâncias para a distribuição amostral dos EMVs de  $\kappa$  e  $\eta_\tau$ , respectivamente), erro quadrático médio (EQM, em que  $EQM_\kappa$  e  $EQM_{\eta_\tau}$  são os erros quadráticos médios para a distribuição amostral dos EMVs de  $\kappa$  e  $\eta_\tau$ , respectivamente), coeficiente de assimetria (CS, com  $CS_\kappa$  e  $CS_{\eta_\tau}$  representando os coeficientes de assimetria para a distribuição amostral dos EMVs de  $\kappa$  e  $\eta_\tau$ , respectivamente) e coeficiente de curtose (CK, em que  $CK_\kappa$  e  $CK_{\eta_\tau}$  são os coeficientes de curtose para a distribuição amostral dos EMVs de  $\kappa$  e  $\eta_\tau$ , respectivamente).

Na simulação, foram realizadas 5000 réplicas de Monte Carlo, considerando tamanhos amostrais de  $n \in \{10, 30, 50, 100\}$ . Neste trabalho, foram utilizados os seguintes valores para o parâmetro de forma  $\kappa$ : 1,5, 2,5 e 5. Observa-se, na Figura 1, que diferentes valores para o parâmetro de forma  $\kappa$  podem indicar os graus de curtose dos dados. Além disso, os valores da curtose, assimetria, entre outras medidas descritivas baseadas nos valores dos parâmetros do modelo WER podem ser obtidos no software R por meio do pacote *dprop*, que utiliza uma parametrização para o modelo Weibull exponencializado que pode ser adaptado para calcular essas estatísticas descritivas para o modelo WER. Sendo assim, os valores fixados para  $\kappa$  foram selecionados com o objetivo de obter modelos com diferentes níveis de curtose. Sem perda de generalidade, o parâmetro  $\eta_\tau$  foi considerado fixo sendo  $\eta_\tau = 1$  em todos os experimentos de Monte Carlo. Também consideramos  $\tau$ : 0,25, 0,5, e 0,75 com o objetivo de estudar o comportamento dos parâmetros para diferentes valores de  $\tau$ . Para valores de  $\tau$  mais extremos a interpretação dos resultados é análoga. Além disso, apresentamos os resultados numéricos das simulações para os intervalos de confiança (IC) assintótico.

Na Tabela 2, é possível observar que, independentemente dos valores  $\kappa$  e  $\tau$ , à medida que o tamanho da amostra aumenta, a variância, o erro quadrático médio e o viés relativo percentual em valor absoluto diminuem. Logo, à medida que o tamanho da amostra cresce, os

valores das estimativas dos parâmetros  $\kappa$  e  $\eta_\tau$  estão se aproximando dos valores verdadeiros dos parâmetros do modelo WER, ou seja, apresentaram melhor desempenho com o aumento da amostra. Além disso, percebe-se que para valores de  $\kappa$  e  $n$  fixos à medida que o valor de  $\tau$  aumenta o viés relativo percentual em valor absoluto, a variância e o erro quadrático médio para  $\kappa$  também aumentam. Sendo assim, as estimativas do parâmetro  $\kappa$  apresentaram melhor desempenho para valores menores de  $\tau$ . Adicionalmente, para valores de  $\kappa$  e  $n$  fixos o estimador  $\hat{\eta}_\tau$  apresentou melhor desempenho para valores de  $\tau = 0,5$ . Também, nota-se que, para valores de  $\tau$  e  $n$  fixos, à medida que o valor do parâmetro de forma  $\kappa$  aumenta, a variância e o erro quadrático médio também aumentam e o estimador  $\hat{\kappa}$  apresentou comportamento similar quando comparado às estimativas do viés relativo percentual em valor absoluto. Isto mostra que, à medida que o valor do parâmetro  $\kappa$  cresce, ou seja com o aumento do grau de curtose, as estimativas do parâmetro  $\kappa$  pioram. Além disso, podemos observar que, para  $n = 10$  o estimador  $\hat{\kappa}$  apresentou curtose alta, isto indica que há fortes indícios de que o tamanho da amostra é pequeno neste cenário. Adicionalmente, para  $n = 10$ ,  $\kappa = 5$  e  $\tau: 0,25, 0,5$  e  $0,75$  o estimador  $\hat{\kappa}$  apresentou variância e erro quadrático médio alto, isto ocorreu por que há fortes indícios de que o tamanho da amostra é pequeno para  $n = 10$ . Por fim, podemos observar que com o aumento do tamanho da amostra a assimetria vai diminuindo e se aproximando de zero e a curtose está se aproximando de 3. Portanto, este resultado indica que há fortes indícios de que os estimadores dos parâmetros do modelo WER são distribuídos assintoticamente como uma variável aleatória normal, ou seja, é possível aproximar a distribuição do estimador pela normal para grandes amostras.

De acordo com a Tabela 3, observa-se que todos os intervalos de confiança contêm valores no espaço paramétrico de  $\kappa$  e  $\eta_\tau$ . Além disso, todos os intervalos de confiança (*IC*) foram calculados com nível nominal de cobertura  $(1 - \gamma)$  igual a 0,95. Adicionalmente, o melhor estimador intervalar para os parâmetros  $\kappa$  e  $\eta_\tau$  é o que apresentar o menor comprimento do *IC*. Além disso, observe que os comprimentos médios dos intervalos de confiança (*CMIC*) em relação ao parâmetro  $\kappa$ , quando  $\kappa = 1,5$  e variando os valores de  $\tau$  apresentaram *CMIC* similares para tamanhos amostrais fixos. Por exemplo, para  $\kappa = 1,5$ ,  $n = 100$  e  $\tau$  assumindo os valores 0,25, 0,5 e 0,75, os *CMIC* foram iguais a 0,428, 0,451 e 0,496, respectivamente. De modo análogo, nota-se o mesmo comportamento no caso em que  $\kappa = 2,5$  e  $\tau$  assumindo os valores 0,25, 0,5 e 0,75, ou seja, os *IC* apresentaram *CMIC* próximos para tamanhos amostrais fixos. Em termos práticos significa que para

valores de  $\kappa$  e  $n$  fixos, os valores de  $\tau$  representam uma influência pequena no CMIC.

Nota-se ainda na Tabela 3 que, com  $\kappa = 5$  e variando os valores de  $\tau$ , os *IC* apresentaram *CMIC* similares para tamanhos amostrais fixo. Além disso, podemos observar que para tamanhos amostrais e  $\tau$  fixos, à medida que o valor do parâmetro  $\kappa$  aumenta o *CMIC* também aumenta. Por exemplo, para  $\kappa = 1,5$ ,  $\tau = 0,25$  e  $n = 100$  o *CMIC* foi igual a 0,428. Por outro lado, para  $\kappa = 5$ ,  $\tau = 0,25$  e  $n = 100$  o *CMIC* foi igual a 1,428. Isso indica que quanto menor o valor do parâmetro  $\kappa$ , melhor é a estimativa intervalar em relação ao parâmetro  $\kappa$ . Adicionalmente, considerando valores de  $\kappa$  fixo e tamanhos amostrais fixos, à medida que o valor de  $\tau$  aumenta, o *CMIC* para o parâmetro  $\eta_\tau$  também aumenta. Isso mostra que quanto menor o valor de  $\tau$ , melhor é a estimativa intervalar para o parâmetro  $\eta_\tau$ . Também percebe-se que, para todos os intervalos de confiança, com o aumento do tamanho da amostra, os *CMIC* diminuem, ou seja, à medida que o tamanho da amostra cresce, as estimativas intervalares para os parâmetros  $\kappa$  e  $\eta_\tau$  melhoram.

Nas Figuras 4-6, ilustra-se os histogramas das 5000 estimativas de máxima verossimilhança para os parâmetros do modelo WER. Os segmentos de reta indicam o comprimento do intervalo de confiança. Nota-se que, à medida que o tamanho da amostra aumenta, independentemente dos valores de  $\kappa$  e  $\tau$ , a distribuição empírica torna-se mais simétrica em torno do parâmetro  $\kappa$ , uma vez que há indícios de que os estimadores dos parâmetros do modelo WER são distribuídos assintoticamente como uma variável aleatória normal. Para avaliar as estimativas intervalares, foi calculada a probabilidade de cobertura empírica dos intervalos assintóticos, denotada na Tabela 3 por PC, a probabilidade de cobertura empírica à esquerda, denotada por LIE, e a probabilidade de cobertura empírica à direita denotada por LSD, com nível nominal de cobertura  $(1 - \gamma)$  igual a 0,95 para todos os tamanhos amostrais. De acordo com a Tabela 3, nota-se, que para todos os tamanhos amostrais, a probabilidade de cobertura para o parâmetro de forma  $\kappa$  está próxima de 0,95. Logo, a estimativa intervalar para  $\kappa$  é precisa, ou seja, com pequeno erro de cobertura. As probabilidades de cobertura observadas para o parâmetro  $\eta_\tau$  foram melhores para tamanhos amostrais maiores. Dessa forma, quando o tamanho da amostra aumenta, as probabilidades de cobertura observadas para o parâmetro  $\eta_\tau$  se aproximam da probabilidade de cobertura nominal de 0,95. Além disso, quando o tamanho da amostra aumenta, os intervalos de confiança tendem a se equilibrar em relação à probabilidade de cobertura. Adicionalmente, os histogramas tendem à distribuição normal, ou seja, suas formas estão tendendo a um

formato semelhante ao de um sino.

Nota-se ainda na Tabela 3 que, para valores fixos de  $\tau$  e  $n$ , à medida que o valor do parâmetro de forma  $\kappa$  aumenta, a PC, LIE e LSD apresentaram comportamentos similares nesse cenário. No entanto, ao considerar esse cenário, observa-se uma pequena variação na PC, LIE e LSD a partir da quinta casa decimal. Há fortes evidências de que esse comportamento ocorreu devido ao fato de o viés relativo percentual em valor absoluto ter apresentado comportamentos similares nesse contexto.

Portanto, nesta seção foi realizado um estudo de simulação de Monte Carlo para analisar empiricamente o desempenho dos estimadores de máxima verossimilhança e intervalos de confiança assintóticos dos parâmetros que indexam o modelo WER, e investigar suas propriedades. Os resultados numéricos das simulações indicam que, com o aumento do tamanho da amostra, os valores das estimativas se aproximam dos valores verdadeiros dos parâmetros do modelo WER. Também se nota que, quanto maior o tamanho da amostra a variância, o erro quadrático médio e o viés relativo percentual em valor absoluto diminuem, mostrando que os estimadores de máxima verossimilhança dos parâmetros do modelo WER são assintoticamente não enviesados. Nota-se ainda que, o estimador  $\hat{\kappa}$  tem um melhor desempenho para valores menores de  $\tau$ . Também se observa que os intervalos de confiança assintóticos para o parâmetro  $\kappa$  apresentaram melhores taxas de cobertura. Além disso, com o aumento do tamanho da amostra a assimetria vai diminuindo e se aproximando de zero e a curtose está se aproximando de 3. Logo, há fortes indícios de que os estimadores dos parâmetros do modelo WER são distribuídos assintoticamente como uma variável aleatória normal.

Tabela 2: Estimativas de máxima verossimilhança e medidas de desempenho estimadas pelo método BFGS para  $\kappa$ : 1,5, 2,5 e 5,  $\eta_\tau = 1$  fixo, e  $\tau$ : 0,25, 0,5 e 0,75.

$\kappa$	$\tau$	$n$	$\hat{\kappa}$	$\hat{\eta}_\tau$	$CS_{\hat{\kappa}}$	$CS_{\hat{\eta}_\tau}$	$CK_{\hat{\kappa}}$	$CK_{\hat{\eta}_\tau}$	$VR_{\hat{\kappa}}$	$VR_{\hat{\eta}_\tau}$	$VAR_{\hat{\kappa}}$	$VAR_{\hat{\eta}_\tau}$	$EQM_{\hat{\kappa}}$	$EQM_{\hat{\eta}_\tau}$
1,5	0,25	10	1,74416	1,04598	1,55313	0,42791	7,44962	3,30610	16,27730	4,59805	0,26157	0,05024	0,32118	0,05236
		30	1,56682	1,01217	0,68042	0,16606	3,73955	3,07211	4,45464	1,21665	0,05043	0,01551	0,05489	0,01567
		50	1,54004	1,00743	0,48242	0,12505	3,42835	3,04584	2,66918	0,74353	0,02808	0,00922	0,02969	0,00923
		100	1,52014	1,00417	0,28948	0,08392	3,00650	3,06001	1,34291	0,41713	0,01297	0,00458	0,01338	0,00459
	0,50	10	1,75799	1,01934	1,56941	0,46148	7,61157	3,25464	17,19933	1,93435	0,29188	0,06301	0,35844	0,06338
		30	1,56992	1,00331	0,68509	0,17866	3,81571	3,02799	4,66122	0,33115	0,05540	0,02016	0,06029	0,02017
		50	1,54166	1,00193	0,47964	0,13299	3,42529	3,00156	2,77768	0,19300	0,03070	0,01194	0,03244	0,01194
		100	1,52094	1,00139	0,28208	0,09837	3,04650	3,00244	1,39617	0,13921	0,01416	0,00592	0,01459	0,00592
	0,75	10	1,78913	0,97652	1,60713	0,52118	7,88335	3,19977	19,27528	2,34760	0,35425	0,10894	0,43786	0,10949
		30	1,57697	0,98853	0,70374	0,27094	4,00811	3,05832	5,13143	1,14698	0,06528	0,03706	0,07121	0,03719
		50	1,54569	0,99287	0,48182	0,18158	3,42195	2,93302	3,04628	0,71334	0,03593	0,02197	0,03801	0,02202
		100	1,52282	0,99674	0,28543	0,10589	3,10116	2,98055	1,52172	0,32624	0,01651	0,01079	0,01703	0,01079
2,5	0,25	10	2,90694	1,02169	1,55314	0,16482	7,44959	3,08347	16,27746	2,16993	0,72658	0,01733	0,89217	0,01780
		30	2,61137	1,00545	0,68042	0,01668	3,73948	3,01627	4,45467	0,54465	0,14009	0,00554	0,15249	0,00557
		50	2,56673	1,00335	0,48247	0,00889	3,42853	3,03291	2,66909	0,33530	0,07800	0,00330	0,08246	0,00331
		100	2,53359	1,00196	0,28951	0,00122	3,00675	3,04443	1,34346	0,19568	0,03603	0,00164	0,03715	0,00165
	0,50	10	2,92998	1,00422	1,56939	0,16933	7,61164	3,01978	17,19928	0,42191	0,81079	0,02218	0,99568	0,02220
		30	2,61652	0,99956	0,68517	0,00992	3,81598	2,98036	4,66096	0,04347	0,15389	0,00726	0,16746	0,00726
		50	2,56946	0,99973	0,47961	0,00254	3,42516	2,99125	2,77836	0,02731	0,08529	0,00429	0,09011	0,00429
		100	2,53490	1,00012	0,28211	0,00698	3,04651	2,97963	1,39612	0,01238	0,03933	0,00213	0,04054	0,00213
	0,75	10	2,98189	0,97220	1,60720	0,14555	7,88371	2,87057	19,27544	2,77957	0,98409	0,03993	1,21631	0,04071
		30	2,62829	0,98855	0,70370	0,04221	4,00806	2,94927	5,13164	1,14454	0,18134	0,01350	0,19779	0,01363
		50	2,57617	0,99304	0,48183	0,00975	3,42190	2,91469	3,04664	0,69634	0,09980	0,00798	0,10560	0,00803
		100	2,53805	0,99674	0,28518	-0,01703	3,10086	2,95879	1,52191	0,32638	0,04586	0,00390	0,04731	0,00391
5	0,25	10	5,81393	1,00868	1,55316	-0,03559	7,44910	3,07556	16,27869	0,86844	2,90648	0,00426	3,56897	0,00433
		30	5,22278	1,00203	0,68069	-0,09561	3,73972	3,02725	4,45562	0,20308	0,56050	0,00138	0,61013	0,00139
		50	5,13347	1,00126	0,48244	-0,07915	3,42828	3,05559	2,66949	0,12643	0,31203	0,00082	0,32984	0,00083
		100	5,06717	1,00077	0,28958	-0,06093	3,00645	3,04916	1,34344	0,07728	0,14423	0,00041	0,14864	0,00041
	0,50	10	5,85994	0,99933	1,56890	-0,05756	7,60946	3,04905	17,19880	0,06706	3,24268	0,00556	3,98218	0,00556
		30	5,23306	0,99887	0,68509	-0,11826	3,81559	3,01482	4,66121	0,11306	0,61558	0,00182	0,66990	0,00183
		50	5,13889	0,99932	0,47967	-0,09643	3,42533	3,02086	2,77788	0,06771	0,34115	0,00108	0,36044	0,00108
		100	5,06982	0,99979	0,28211	-0,06159	3,04649	2,98199	1,39649	0,02038	0,15731	0,00053	0,16218	0,00053
	0,75	10	5,96378	0,98067	1,60701	-0,15768	7,88137	2,99322	19,27571	1,93316	3,93644	0,01049	4,86532	0,01087
		30	5,25657	0,99252	0,70372	-0,13303	4,00799	2,99999	5,13144	0,74746	0,72537	0,00345	0,79119	0,00350
		50	5,15232	0,99549	0,48180	-0,12205	3,42197	2,96697	3,04641	0,45025	0,39919	0,00202	0,42224	0,00204
		100	5,07608	0,99787	0,28533	-0,10954	3,10101	2,97506	1,52167	0,21259	0,18345	0,00098	0,18924	0,00099

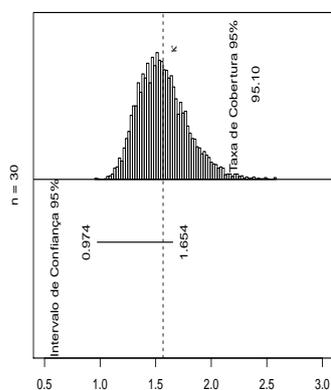
Fonte: O autor (2025).

Tabela 3: Estimativas intervalares e probabilidades de cobertura (PC) empíricas dos IC assintótico e PC empíricas das caudas esquerda (LIE) e direita (LSD) para  $\kappa$ : 1,5, 2,5 e 5,  $\eta_\tau = 1$  fixo, e  $\tau$ : 0,25, 0,5 e 0,75, com nível de confiança  $1 - \gamma = 0,95$ .

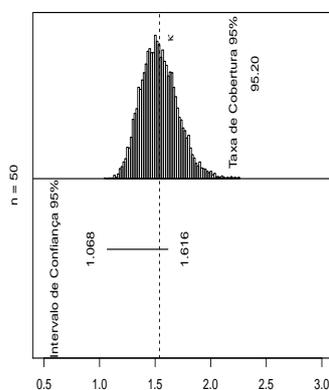
$\kappa$	$\tau$	$n$	$\kappa$			$\eta_\tau$			$\kappa$			$\eta_\tau$		
			<i>LI</i>	<i>LS</i>	<i>CMIC</i>	<i>LI</i>	<i>LS</i>	<i>CMIC</i>	<i>PC</i>	<i>LIE</i>	<i>LSD</i>	<i>PC</i>	<i>LIE</i>	<i>LSD</i>
1,5	0,25	10	0,841	2,057	1,216	0,536	1,334	0,798	0,9422	0,0420	0,0158	0,9046	0,0676	0,0278
		30	0,974	1,654	0,680	0,788	1,394	0,606	0,9510	0,0300	0,0190	0,9426	0,0320	0,0254
		50	1,068	1,616	0,548	0,752	1,156	0,404	0,9520	0,0276	0,0204	0,9460	0,0302	0,0238
		100	1,260	1,688	0,428	0,864	1,139	0,274	0,9508	0,0268	0,0224	0,9502	0,0284	0,0214
	0,50	10	0,788	1,994	1,206	0,434	1,394	0,959	0,9500	0,0340	0,0160	0,8942	0,0430	0,0628
		30	0,988	1,718	0,730	0,836	1,589	0,752	0,9524	0,0272	0,0204	0,9368	0,0222	0,0410
		50	1,057	1,626	0,569	0,750	1,229	0,479	0,9534	0,0254	0,0212	0,9412	0,0232	0,0356
		100	1,251	1,702	0,451	0,854	1,171	0,317	0,9552	0,0226	0,0222	0,9464	0,0210	0,0326
	0,75	10	0,683	1,975	1,292	0,215	1,660	1,446	0,9556	0,0260	0,0184	0,8446	0,0190	0,1364
		30	1,010	1,842	0,832	0,861	2,015	1,154	0,9546	0,0224	0,0230	0,9138	0,0102	0,0760
		50	1,034	1,641	0,607	0,708	1,404	0,696	0,9544	0,0238	0,0218	0,9272	0,0120	0,0608
		100	1,242	1,737	0,496	0,821	1,254	0,433	0,9550	0,0220	0,0230	0,9398	0,0130	0,0472
2,5	0,25	10	1,402	3,431	2,029	0,714	1,206	0,492	0,9422	0,0420	0,0158	0,9020	0,0786	0,0194
		30	1,623	2,758	1,134	0,878	1,229	0,351	0,9510	0,0300	0,0190	0,9436	0,0368	0,0196
		50	1,779	2,693	0,914	0,848	1,096	0,247	0,9520	0,0276	0,0204	0,9452	0,0352	0,0196
		100	2,100	2,814	0,713	0,918	1,083	0,165	0,9508	0,0268	0,0224	0,9496	0,0308	0,0196
	0,50	10	1,314	3,323	2,009	0,649	1,246	0,596	0,9500	0,0340	0,0160	0,8986	0,0512	0,0502
		30	1,647	2,864	1,216	0,914	1,331	0,418	0,9524	0,0272	0,0204	0,9388	0,0274	0,0338
		50	1,761	2,709	0,948	0,849	1,138	0,289	0,9534	0,0254	0,0212	0,9414	0,0270	0,0316
		100	2,084	2,836	0,752	0,913	1,102	0,189	0,9552	0,0226	0,0222	0,9472	0,0234	0,0294
	0,75	10	1,138	3,291	2,153	0,517	1,407	0,889	0,9556	0,0260	0,0184	0,8578	0,0272	0,1150
		30	1,683	3,070	1,387	0,944	1,543	0,599	0,9546	0,0224	0,0230	0,9214	0,0158	0,0628
		50	1,724	2,736	1,012	0,829	1,238	0,409	0,9544	0,0238	0,0218	0,9324	0,0180	0,0496
		100	2,069	2,896	0,827	0,894	1,150	0,256	0,9552	0,0220	0,0228	0,9412	0,0162	0,0426
5	0,25	10	2,804	6,862	4,058	0,854	1,105	0,251	0,9422	0,0420	0,0158	0,8980	0,0876	0,0144
		30	3,247	5,515	2,268	0,941	1,112	0,171	0,9510	0,0300	0,0190	0,9436	0,0410	0,0154
		50	3,559	5,386	1,827	0,923	1,049	0,125	0,9520	0,0276	0,0204	0,9430	0,0390	0,0180
		100	4,200	5,628	1,428	0,959	1,041	0,082	0,9508	0,0268	0,0224	0,9496	0,0330	0,0174
	0,50	10	2,628	6,647	4,019	0,820	1,127	0,306	0,9500	0,0340	0,0160	0,9010	0,0600	0,0390
		30	3,294	5,728	2,433	0,961	1,158	0,197	0,9524	0,0272	0,0204	0,9424	0,0314	0,0262
		50	3,522	5,419	1,897	0,924	1,069	0,145	0,9534	0,0254	0,0212	0,9420	0,0302	0,0278
		100	4,169	5,672	1,504	0,957	1,051	0,094	0,9552	0,0226	0,0222	0,9504	0,0254	0,0242
	0,75	10	2,276	6,583	4,307	0,754	1,208	0,453	0,9556	0,0260	0,0184	0,8674	0,0344	0,0982
		30	3,366	6,138	2,772	0,981	1,249	0,269	0,9546	0,0224	0,0230	0,9284	0,0204	0,0512
		50	3,448	5,471	2,023	0,916	1,117	0,201	0,9542	0,0240	0,0218	0,9348	0,0222	0,0430
		100	4,139	5,792	1,653	0,948	1,074	0,127	0,9548	0,0222	0,0230	0,9440	0,0180	0,0380

Fonte: O autor (2025).

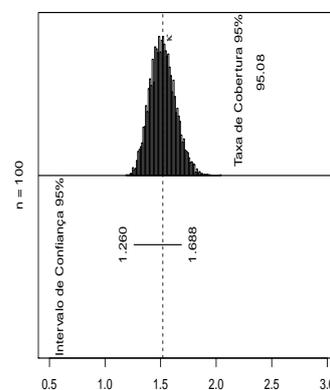
Figura 4 – Distribuição empírica de  $\hat{\kappa}$  com  $\tau = 0,25$  e  $\eta_{\tau} = 1$  fixo.



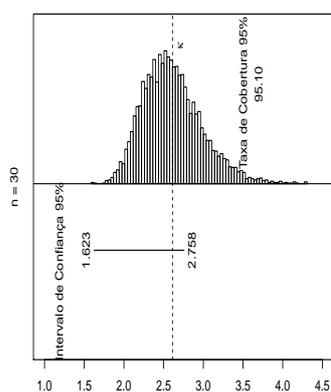
(a)  $\kappa = 1,5$



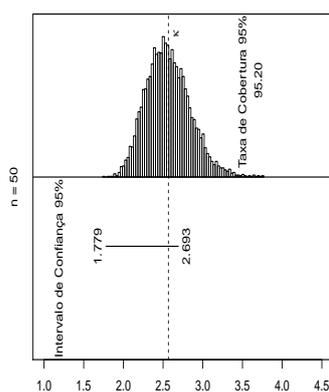
(b)  $\kappa = 1,5$



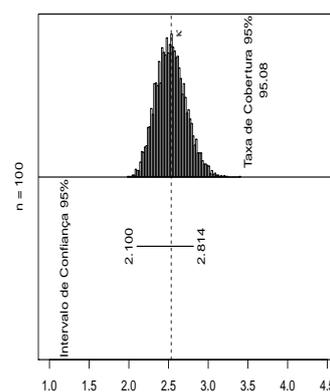
(c)  $\kappa = 1,5$



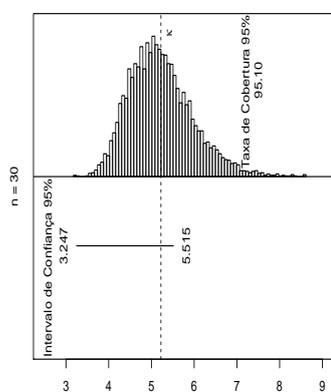
(d)  $\kappa = 2,5$



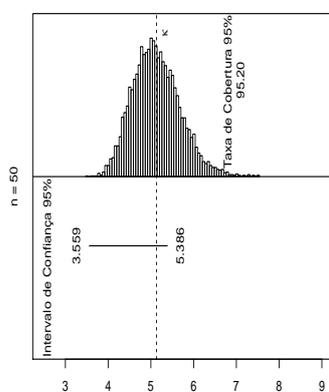
(e)  $\kappa = 2,5$



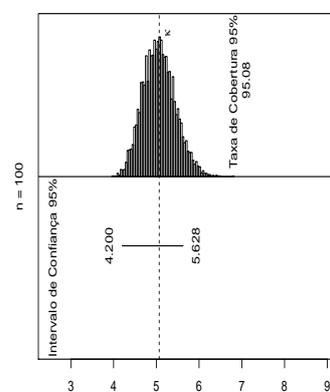
(f)  $\kappa = 2,5$



(g)  $\kappa = 5$



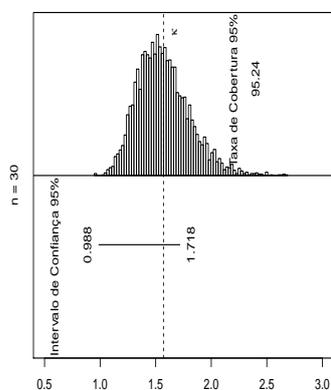
(h)  $\kappa = 5$



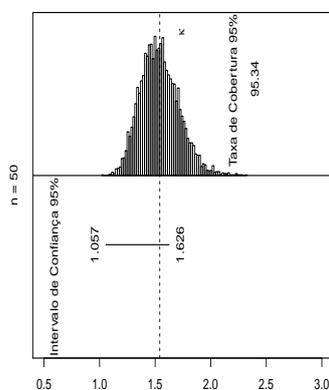
(i)  $\kappa = 5$

Fonte: O autor (2025).

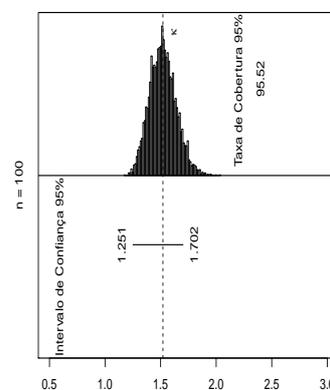
Figura 5 – Distribuição empírica de  $\hat{\kappa}$  com  $\tau = 0,5$  e  $\eta_\tau = 1$  fixo.



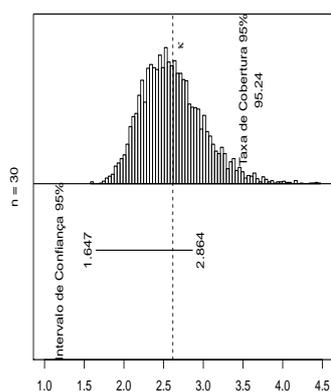
(a)  $\kappa = 1,5$



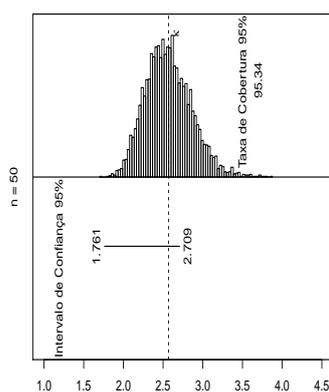
(b)  $\kappa = 1,5$



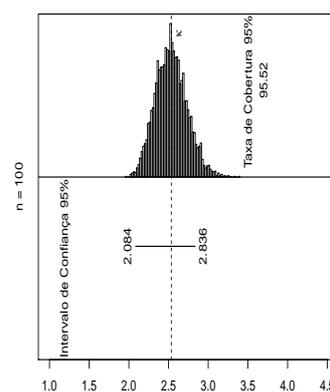
(c)  $\kappa = 1,5$



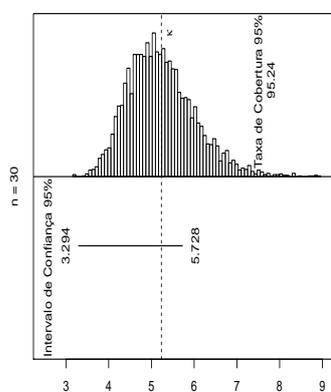
(d)  $\kappa = 2,5$



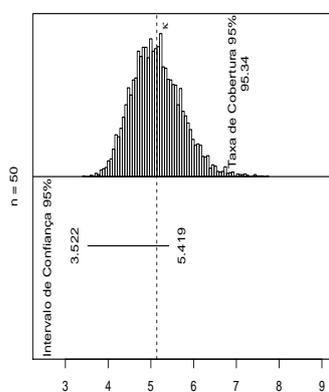
(e)  $\kappa = 2,5$



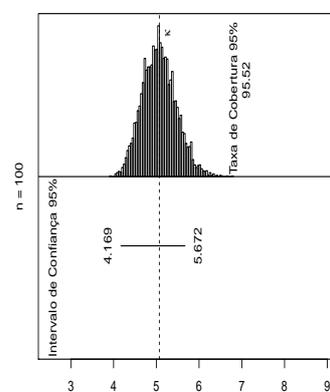
(f)  $\kappa = 2,5$



(g)  $\kappa = 5$



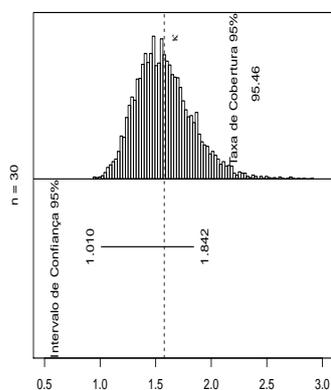
(h)  $\kappa = 5$



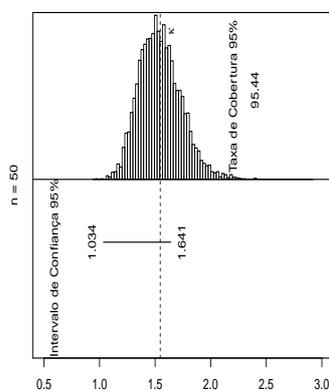
(i)  $\kappa = 5$

Fonte: O autor (2025).

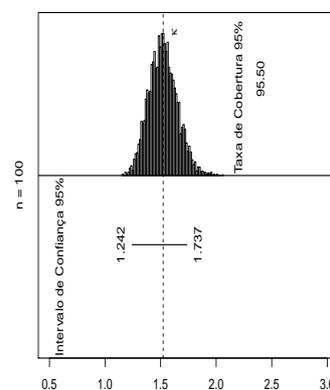
Figura 6 – Distribuição empírica de  $\hat{\kappa}$  com  $\tau = 0,75$  e  $\eta_\tau = 1$  fixo.



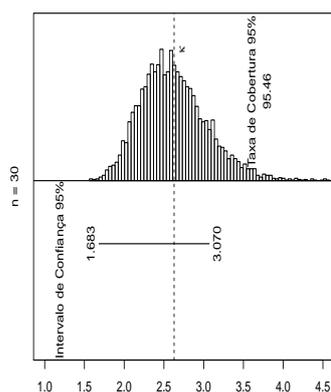
(a)  $\kappa = 1,5$



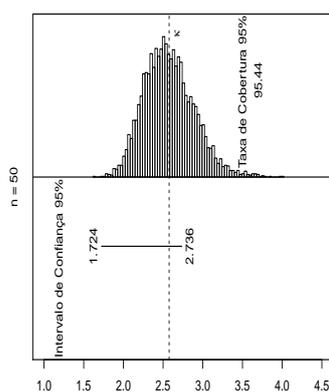
(b)  $\kappa = 1,5$



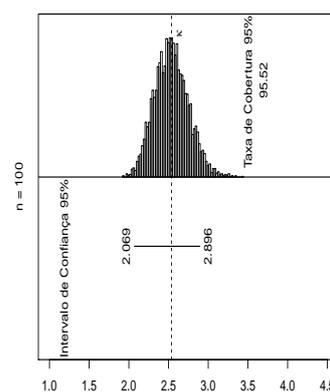
(c)  $\kappa = 1,5$



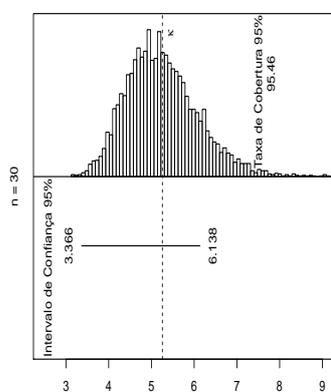
(d)  $\kappa = 2,5$



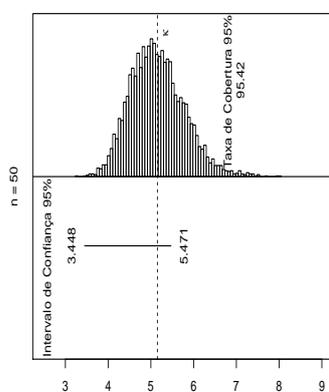
(e)  $\kappa = 2,5$



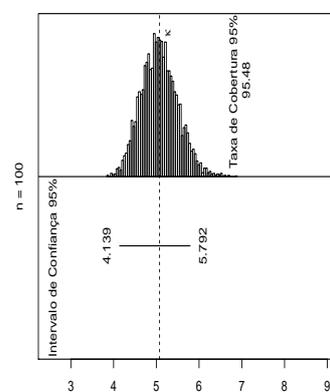
(f)  $\kappa = 2,5$



(g)  $\kappa = 5$



(h)  $\kappa = 5$



(i)  $\kappa = 5$

Fonte: O autor (2025).

## 2.7 APLICAÇÃO

Nesta aplicação, o conjunto de dados utilizado é referente à tensão de ruptura de fibras de carbono (em GBa) (NICHOLS; PADGETT, 2006). O conjunto de dados é composto por 100 medições e está disponível no pacote `AdequacyModel` (MARINHO et al., 2019) do R. Para acessá-lo, basta executar o comando `data(carbone)`.

Na Tabela 4, fornecemos diversas estatísticas descritivas para este conjunto de dados. Essas medidas foram obtidas utilizando a função `basicStats()` do pacote `fbasic` no R. Pode-se observar que a tensão média de ruptura das fibras de carbono é de 2,62. A distribuição dos dados mostra uma ligeira assimetria positiva. Além disso, a distribuição apresenta uma forma mesocúrtica. Nota-se, ainda, que o primeiro quartil é igual a 1,84, o segundo quartil, que representa a mediana, é igual a 2,7 e o terceiro quartil é igual a 3,22. Nota-se também que, o coeficiente de variação percentual (CV%), foi igual a 38,68.

Tabela 4: Medidas descritivas para o conjunto de dados de tensão de ruptura de fibras de carbono.

Média	$Q_1$	Mediana	$Q_3$	Variância	Assimetria	Curtose	Mínimo	Máximo	CV%	$n$
2,62	1,84	2,70	3,22	1,03	0,36	0,04	0,39	5,56	38,68	100

Fonte: O autor (2025).

Para ajustar os dados, usou-se a função `goodness.fit()` do pacote `AdequacyModel`. As estimativas de máxima verossimilhança e seus erros-padrão são apresentados na Tabela 5. Observa-se, portanto, que todos os modelos ajustados apresentaram erros-padrão pequenos em comparação com os valores das estimativas para seus respectivos parâmetros. O conjunto de dados foi ajustado considerando os modelos: EER, WER, BSER e LLER, descritos na Tabela 1. Adicionalmente, consideramos os seguintes valores de  $\tau$ : 0,25, 0,5 e 0,75.

Para analisar qual modelo proporcionou o melhor ajuste, consideramos os seguintes critérios: critério de informação de Akaike (AIC) (SAKAMOTO; ISHIGURO; KITAGAWA, 1983), critério de informação de Akaike consistente (CAIC) (AKAIKE, 1973), critério de informação Bayesiano (BIC) (SCHWARZ, 1978), Anderson-Darling ( $A^*$ ) (ANDERSON; DARLING, 1954), Cramér-Von Mises ( $W^*$ ) (CRAMÉR, 1928; MISES, 1928) e Kolmogorov-Sminorv ( $K-S$ ) (KOLMOGOROV, 1933) com seu  $p$ -valor associado. Portanto, o processo de seleção do modelo determina o modelo mais adequado entre os modelos ajustados com base nos valores mais

baixos dessas estatísticas. Os chutes iniciais para os parâmetros foram obtidos do ajuste do modelo *baseline*. Os valores dessas estatísticas são apresentados na Tabela 6.

Tabela 5: EMV dos parâmetros para vários modelos ajustados, com as estimativas dos erros-padrão correspondentes entre parênteses.

$\tau$	Modelos	$\eta_\tau$	$\kappa$	$\lambda$
0,25	EER	1,2429 (0,0919)		
	WER	1,8827 (0,1003)	1,9429 (0,1471)	
	BSER	1,7280 (0,0872)		0,5797 (0,0417)
	LLER	1,8357 (0,0889)	3,1394 (0,2505)	
0,50	EER	1,8170 (0,1817)		
	WER	2,5816 (0,1085)	2,7928 (0,2141)	
	BSER	2,3660 (0,1064)		0,4621 (0,0327)
	LLER	2,4984 (0,1054)	4,1178 (0,3441)	
0,75	EER	3,2339 (0,4919)		
	WER	3,4316 (0,1205)	4,6236 (0,3639)	
	BSER	3,2074 (0,1519)		0,3425 (0,0246)
	LLER	3,2382 (0,1191)	6,5582 (0,5854)	

Fonte: O autor (2025).

No geral, a distribuição WER apresentou o melhor ajuste de acordo com os critérios utilizados. Porém, para o quantil 0,75, a distribuição LLER apresentou os menores valores para as estatísticas utilizadas e o maior  $p$ -valor para o teste de Kolmogorov-Smirnov. O ajuste desses modelos ao conjunto de dados estudado é apresentado na Figura 7.

Nas Figuras 8-9, apresentam-se os gráficos qqplots para as distribuições WER e LLER

respectivamente. Nota-se que, para o modelo WER com  $\tau$  assumindo os valores 0,25 e 0,75 os dados estão mais próximos da bissetriz, representada pela linha em preto, este resultado indica que a distribuição WER é mais adequada para se ajustar a esse conjunto de dados em comparação com a distribuição LLER. Por outro lado, para o quantil 0,75, os dados estão mais próximos da linha em preto para a distribuição LLER em comparação com a distribuição WER. Este resultado aponta que a distribuição LLER apresentou melhor ajuste para o quantil 0,75 em comparação com a distribuição WER.

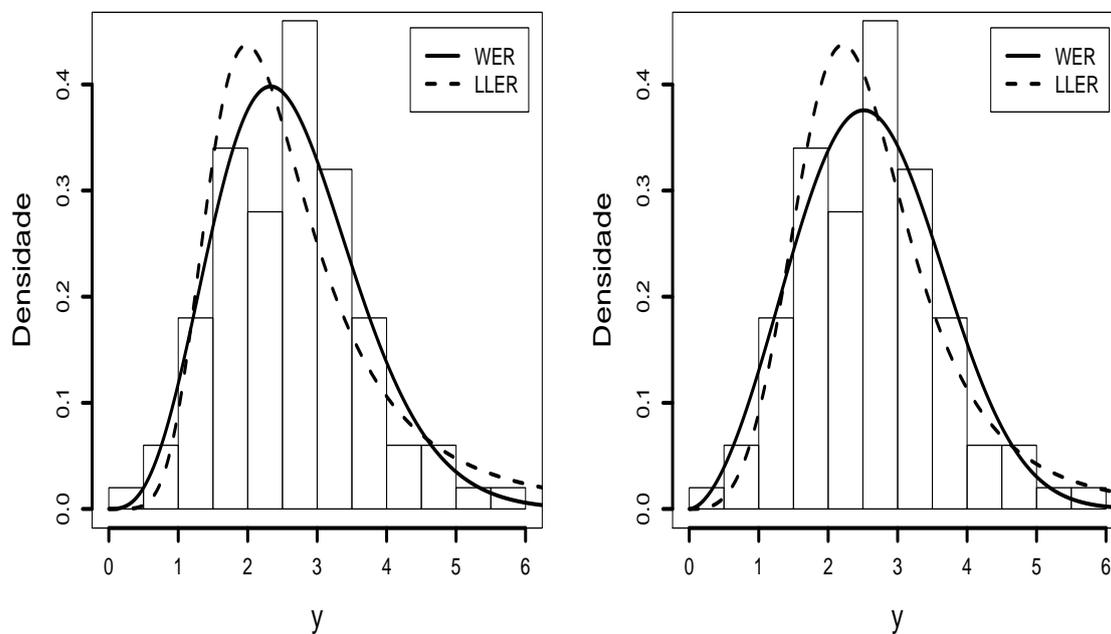
Nota-se que, especificando diferentes quantis, há certa intervenção nos resultados da análise. Portanto, na modelagem quantílica diferentes modelos podem apresentar um bom ajuste ao conjunto de dados para diferentes quantis. Pois, essa técnica permite capturar as características dos dados em diferentes partes da distribuição dos dados.

Tabela 6: Estatísticas AIC, CAIC, BIC,  $A^*$ ,  $W^*$  e  $K-S$  (com seus  $p$ -valores) para o conjunto de dados de estresse de ruptura das fibras de carbono (em Gba).

$\tau$	Models	AIC	CAIC	BIC	$A^*$	$W^*$	KS	$p$ -value
0,25	EER	340,209	340,249	342,814	0,830	0,162	0,210	< 0,001
	WER	<b>287,420</b>	<b>287,544</b>	<b>292,631</b>	<b>0,503</b>	<b>0,096</b>	<b>0,077</b>	<b>0,585</b>
	BSER	304,785	304,909	309,995	1,670	0,307	0,129	0,070
	LLER	308,544	308,667	313,754	2,161	0,399	0,108	0,191
0,50	EER	394,742	394,782	397,347	0,764	0,149	0,320	< 0,001
	WER	<b>287,059</b>	<b>287,182</b>	<b>292,269</b>	<b>0,416</b>	<b>0,062</b>	<b>0,060</b>	<b>0,858</b>
	BSER	304,122	304,245	309,332	1,618	0,298	0,130	0,066
	LLER	296,559	296,683	301,769	1,241	0,239	0,090	0,389
0,75	EER	489,636	489,676	492,241	0,750	0,146	0,480	< 0,001
	WER	294,857	294,981	300,067	0,869	0,109	0,092	0,363
	BSER	302,183	302,307	307,393	1,467	0,271	0,129	0,071
	LLER	<b>286,662</b>	<b>286,786</b>	<b>291,873</b>	<b>0,364</b>	<b>0,066</b>	<b>0,059</b>	<b>0,879</b>

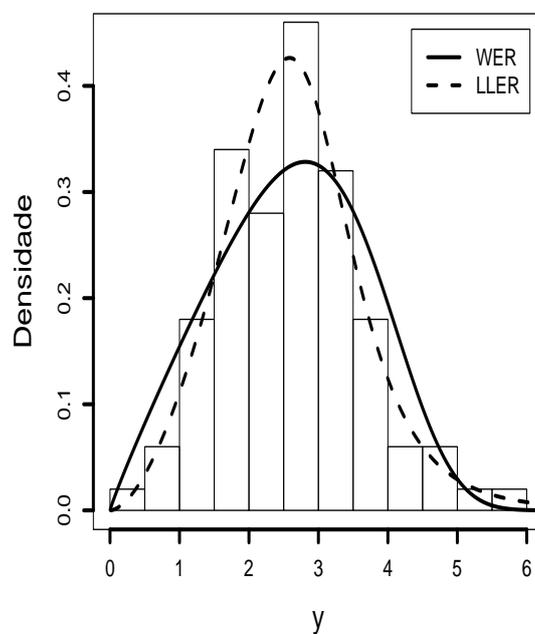
Fonte: O autor (2025).

Figura 7 – Histograma com densidades estimadas para os modelos WER e LLER para o conjunto de dados de tensão de ruptura de fibras de carbono (em GBa).



(a)  $\tau = 0,25$

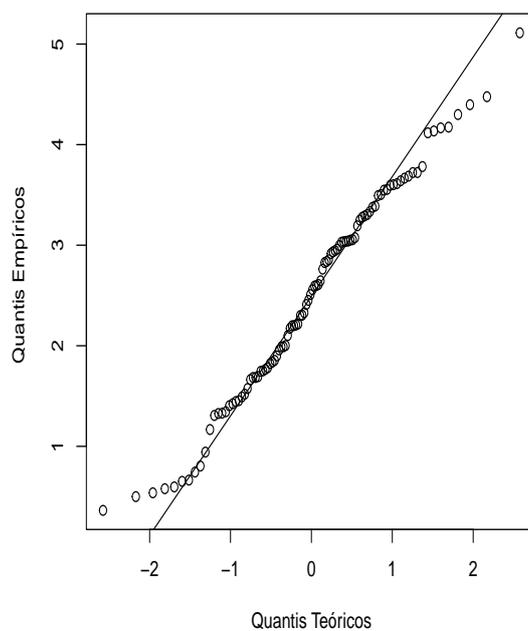
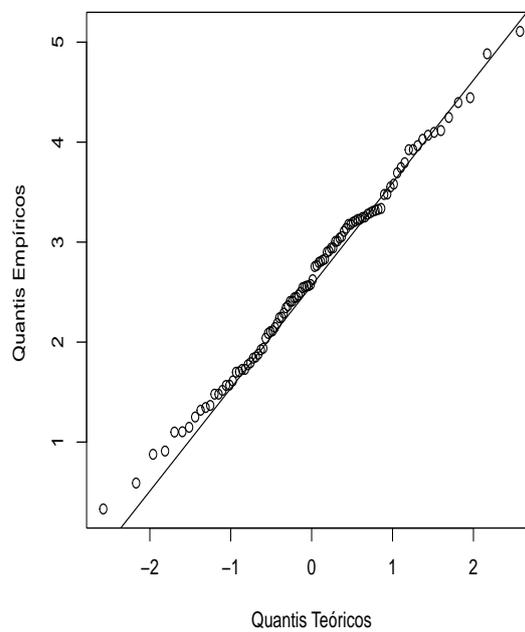
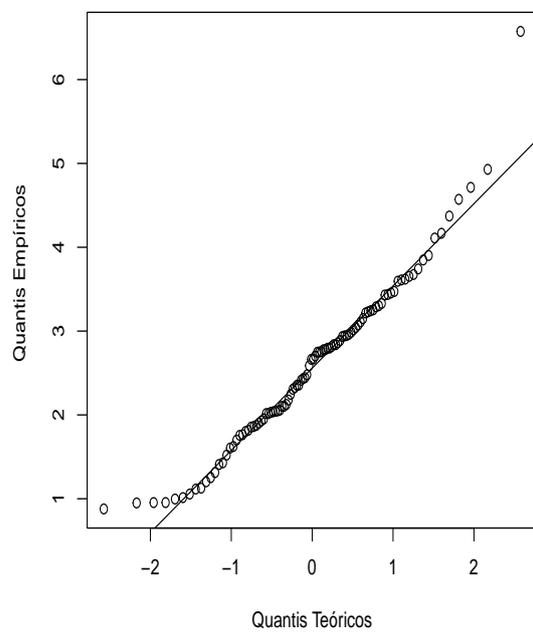
(b)  $\tau = 0,50$



(c)  $\tau = 0,75$

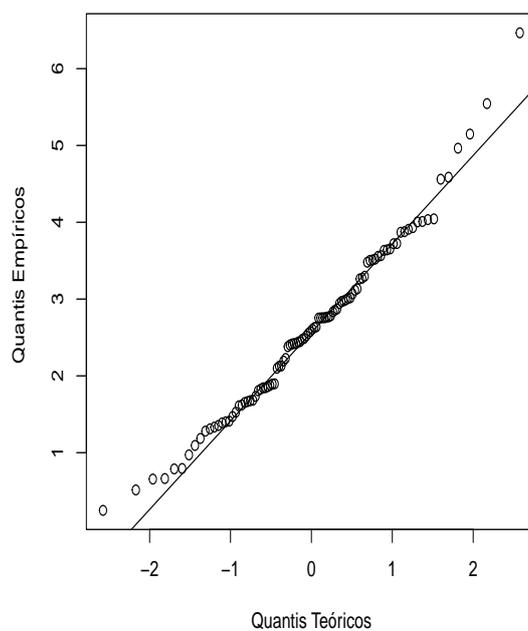
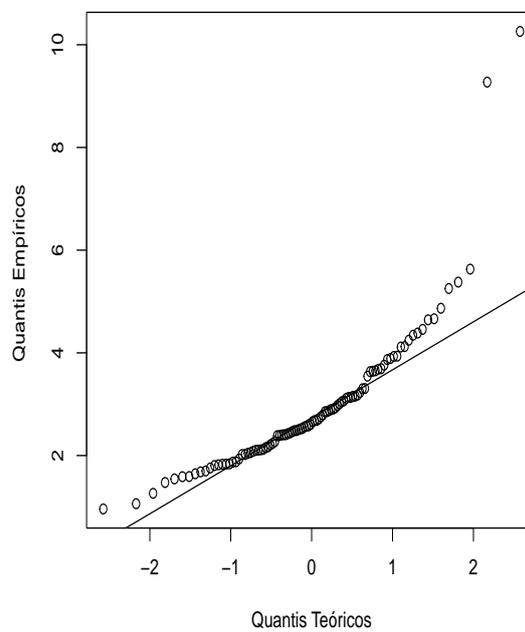
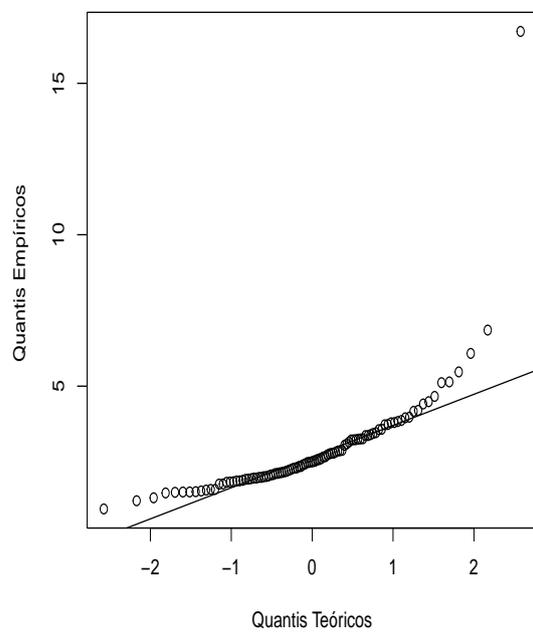
Fonte: O autor (2025).

Figura 8 – qqplots da distribuição WER para o conjunto de dados de tensão de ruptura de fibras de carbono (em GBa).



Fonte: O autor (2025).

Figura 9 – qqplots da distribuição LLER para o conjunto de dados de tensão de ruptura de fibras de carbono (em GBa).



Fonte: O autor (2025).

---

## 2.8 OBSERVAÇÕES FINAIS

Neste capítulo, propomos uma nova família de distribuições indexada por um parâmetro representando o  $(\tau \times 100)$ -ésimo quantil. Isto é altamente útil em vários campos, pois muitas vezes há interesse na modelagem de quantis. Além disso, os resultados aqui apresentados generalizam vários trabalhos existentes. Realizou-se estudos de simulação de Monte Carlo para investigar as propriedades dos estimadores de máxima verossimilhança e intervalos de confiança assintóticos para a distribuição WER. Esses estudos mostraram que o estimador  $\hat{\kappa}$  tem um melhor desempenho para quantis inferiores. Além disso, notamos que os intervalos de confiança assintóticos para o parâmetro  $\kappa$  apresentaram melhores taxas de cobertura. Finalmente, aplicamos a família de distribuições proposta a um conjunto de dados reais. Nesta aplicação, observou-se que a distribuição WER apresentou melhor ajuste para os quantis inferiores, enquanto para o quantil 0,75 a distribuição LLER apresentou ajuste superior.

### 3 REGRESSÃO QUANTÍLICA PARA MODELOS BASEADOS NA FAMÍLIA DE DISTRIBUIÇÕES G-EXPONENCIALIZADA ( $\alpha_T$ -G) REPARAMETRIZADA

#### 3.1 RESUMO

O objetivo deste capítulo é apresentar de forma geral novos modelos de regressão quantílica baseados na família de distribuições  $\alpha_T$ -G reparametrizada. Dessa forma, especificando o modelo *baseline* da família de distribuições  $\alpha_T$ -G reparametrizada podemos obter o correspondente modelo de regressão quantílica para o submodelo da família de distribuições  $\alpha_T$ -G reparametrizada. Para exemplificar a metodologia proposta foi introduzido o novo modelo de regressão quantílica Weibull exponencializado reparametrizado (WER) e alguns de seus aspectos estatísticos. Através dos modelos propostos neste estudo, pode-se modelar os quantis da distribuição, de forma equivalente aos conceitos utilizados nos modelos lineares generalizados. Adicionalmente, é apresentado um estudo de simulação de Monte Carlo para avaliar o desempenho e as propriedades dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de regressão quantílica WER. Resíduos são propostos para verificar a adequabilidade do ajuste do modelo proposto ao conjunto de dados e detectar possíveis observações atípicas. Além disso, foram derivadas medidas de diagnóstico sob enfoque de influência local com o objetivo de identificar observações potencialmente influentes. Para ilustrar, aplicou-se a metodologia proposta a um conjunto de dados reais.

**Palavras-chave:** Análise de diagnóstico; Distribuição *baseline*; Regressão quantílica; Reparametrização; Simulação de Monte Carlo.

#### 3.2 INTRODUÇÃO

De acordo com, Montgomery, Peck e Vining (2012), a análise de regressão é a área da estatística que tem como objetivo estudar e modelar a relação entre a variável resposta e as variáveis independentes. Alternativamente, a regressão quantílica é um ramo da estatística que proporciona analisar a relação entre a variável resposta e as variáveis independentes para qualquer quantil de interesse. Dessa forma, a regressão quantílica oferece um estudo detalhado das características dessa relação. A regressão quantílica é recomendada quando o interesse está em modelar um quantil específico de uma determinada população em

---

estudo como, por exemplo, para modelar variáveis assimétricas, situação comum na área de economia. Neste caso, a regressão quantílica apresenta uma análise mais adequada, detalhada e representativa em comparação com modelos de regressão que modelam a média. É possível notar que a modelagem quantílica proporciona uma análise mais detalhada, pois, em vez de considerar apenas medidas de tendência central, como a média, essa técnica permite analisar diferentes partes da distribuição dos dados por meio dos quantis de interesse. Quando a variável resposta segue uma distribuição assimétrica, é mais adequado modelar os quantis do que a média da variável resposta, pois medidas como a média não são representativas nessa situação, uma vez que a média é muito influenciada por valores extremos. De acordo com Gallardo e Santos-Neto (2021), a modelagem quantílica da distribuição é uma alternativa para evitar esse problema de representatividade dos dados.

Dessa forma, as principais vantagens da regressão quantílica (RQ) em comparação com os modelos de regressão que modelam a média da variável resposta são: flexibilidade e robustez. A primeira proporciona uma investigação mais detalhada da relação entre a variável resposta e as variáveis independentes, e é obtida especificando o quantil de interesse. Quanto à robustez, a RQ é menos influenciada por valores atípicos no conjunto de dados.

A utilização da regressão quantílica, proposta por Koenker e Jr (1978), vem crescendo muito nos últimos anos e tem recebido considerável atenção devido ao seu amplo uso teórico e prático. Tem larga aplicabilidade em diversas áreas de pesquisas, tais como, economia, análise de sobrevivência, confiabilidade, estatística espacial, médica, entre outras. Exemplos de uso dessa metodologia utilizando uma abordagem paramétrica podem ser encontrados em vários trabalhos.

Sánchez et al. (2020) propuseram um modelo de regressão quantílica espacial com base no modelo Birnbaum–Saunders. Além disso, esse modelo foi aplicado a um conjunto de dados espaciais. Gallardo et al. (2020) introduziram um novo modelo de regressão quantílica com base na distribuição normal assimétrica de potência. Adicionalmente, esses autores discutiram o método de máxima verossimilhança e apresentaram medidas de influência local sob três diferentes esquemas de perturbação.

Gallardo e Santos-Neto (2021) propuseram uma classe geral de modelos de regressão quantílica para dados contínuos positivos. Além disso, também forneceram ferramentas de inferência e diagnóstico para esta classe de modelos. Sánchez et al. (2021a) introduziram um modelo de regressão quantílica baseado na distribuição Birnbaum–Saunders, realizaram uma

---

análise de diagnóstico e aplicaram esse modelo de regressão quantílica a um conjunto de dados econômicos. Mazucheli et al. (2021) propuseram um modelo de regressão quantílica baseado na distribuição Birnbaum–Saunders unitária. Esse modelo foi obtido reparametrizando a distribuição Birnbaum–Saunders unitária em termos do  $\tau$ -ésimo quantil por meio da função quantílica. Os autores também desenvolveram um pacote no *R* para estimação dos parâmetros e análise da metodologia proposta.

Sánchez et al. (2021b) desenvolveram um modelo de regressão quantílica baseado no modelo Weibull. Esse modelo é obtido reparametrizando a distribuição Weibull através da função quantílica e introduzindo a estrutura de regressão quantílica. Esses autores também apresentaram medidas de diagnóstico para detectar observações potencialmente influentes. Rodrigues, Borges e Santos (2021) desenvolveram um modelo de regressão quantílica para dados de sobrevivência com base na distribuição Gompertz generalizada. Adicionalmente, esses autores reparametrizaram a distribuição Gompertz generalizada em termos do  $\tau$ -ésimo quantil e apresentaram o modelo de regressão quantílica. Segundo os autores, esse modelo permite modelar diretamente os diferentes quantis condicionais da variável resposta, detalhando como cada variável independente influencia os tempos de sobrevivência em distintos quantis.

Gallardo, Gómez e Segovia (2021) introduziram um modelo de regressão quantílica baseado no modelo Maxwell de potência exponencializado. Além disso, foram estudadas algumas propriedades estatísticas desse modelo e realizaram a estimação dos parâmetros usando o método de máxima verossimilhança. Saulo et al. (2022) propuseram modelos de regressão quantílica paramétrica baseados nas distribuições Dagum e Singh-Maddala. Os autores apresentaram as reparametrizações das distribuições Dagum e Singh-Maddala indexadas por um parâmetro quantílico e introduziram seus respectivos modelos de regressão quantílica. Ademais, esses modelos de regressão quantílica foram aplicados a um conjunto de dados econômicos.

Neste capítulo, estendemos o trabalho proposto por Santos, Santos-Neto e Cysneiros (2024) para a classe de regressão quantílica baseada na família de distribuições  $\alpha_\tau$ -G reparametrizada. Sendo assim, o modelo de *RQ* para o submodelo da família de distribuições  $\alpha_\tau$ -G reparametrizada é obtido especificando o correspondente modelo *baseline* da família de distribuições  $\alpha_\tau$ -G reparametrizada e introduzindo a estrutura de regressão quantílica. Dessa forma, para exemplificar, foi apresentado um caso especial da classe geral de modelos

de regressão quantílica baseada na família de distribuições  $\alpha_\tau$ -G reparametrizada, chamado de modelo de regressão quantílica WER. Por meio da regressão quantílica, proposta neste trabalho é possível analisar como cada variável independente pode influenciar os diferentes quantis condicionais da variável resposta. Neste trabalho, utilizou-se a família de distribuições  $\alpha_\tau$ -G reparametrizada como suporte para modelos de regressão quantílica, uma vez que esta família reparametrizada é indexada pelo parâmetro  $\eta_\tau$ , que representa o  $\tau$ -ésimo quantil.

Este capítulo está organizado da seguinte maneira: na Seção 3.3, foi apresentado de maneira geral o modelo de regressão quantílica baseado em modelos na família de distribuições  $\alpha_\tau$ -G reparametrizada. Também, foi apresentado o processo de estimação dos parâmetros desse modelo utilizando o método de máxima verossimilhança. Além disso, foi introduzido um caso particular de modelo de regressão quantílica pertencente à classe de regressão quantílica baseada na família de distribuições  $\alpha_\tau$ -G reparametrizada. Esse novo modelo foi denominado de modelo de regressão quantílica WER, e suas principais propriedades estatísticas, derivadas da classe geral de modelos de regressão quantílica proposta foram desenvolvidas. Na Seção 3.4, foram apresentados alguns resíduos para avaliar o ajuste do modelo e identificar observações atípicas. Também realizou-se uma análise de diagnóstico para detectar observações potencialmente influentes. Na Seção 3.5, foi realizado um estudo de simulação de Monte Carlo para observar o comportamento e as propriedades dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de regressão quantílica WER. Na Seção 3.6, aplicaram-se os novos modelos de regressão quantílica a um conjunto de dados reais na área de economia. Por fim, na Seção 3.7, são apresentadas as considerações finais sobre o estudo.

### 3.3 MODELO DE REGRESSÃO QUANTÍLICA BASEADO NA FAMÍLIA DE DISTRIBUIÇÕES $\alpha_\tau$ -G REPARAMETRIZADA

Nesta seção, estendemos a família de distribuições  $\alpha_\tau$ -G reparametrizada introduzida por Santos, Santos-Neto e Cysneiros (2024), para a classe de regressão quantílica. Sejam  $Y_1, \dots, Y_n$  variáveis aleatórias independentes. Mais especificamente, assuma que  $Y_i$  segue uma distribuição pertencente à família  $\alpha_\tau$ -G reparametrizada proposta por Santos, Santos-Neto e Cysneiros (2024). Assuma, ainda, que  $Y_i$  é indexado por um parâmetro denotado por  $\eta_{\tau_i}$ , para  $i = 1, \dots, n$ , que representa o  $\tau$ -ésimo quantil da família de distribuições  $\alpha_\tau$ -G

reparametrizada. O modelo de regressão quantílica (RQ) baseado na família de distribuições  $\alpha_\tau$ -G reparametrizada é definido pelo componente sistemático dado por:

$$h(\eta_{\tau_i}) = \phi_i(\tau) = \mathbf{x}_i^\top \boldsymbol{\psi}(\tau), \quad i = 1, \dots, n, \quad (3.1)$$

em que,  $\mathbf{x}_i^\top = (1, x_{i1}, \dots, x_{i(p-1)})$  corresponde aos valores de  $p$  covariáveis,  $\eta_{\tau_i} = h^{-1}(\cdot)$ , sendo  $h^{-1}$  a função inversa de  $h$ ,  $\boldsymbol{\psi}(\tau) = [\psi_0(\tau), \psi_1(\tau), \dots, \psi_{(p-1)}(\tau)]^\top$ , com  $p < n$ , é o vetor de parâmetros de regressão desconhecidos. Adicionalmente,  $h(\cdot)$  é uma função de ligação duas vezes diferenciável, estritamente monótona e invertível. Por exemplo, podemos considerar  $h(\eta_{\tau_i}) = \log(\eta_{\tau_i})$ . Portanto, escolhendo o modelo *baseline* da família de distribuições  $\alpha_\tau$ -G reparametrizada obtém-se, por meio do componente sistemático dado por (3.1), o correspondente modelo de RQ para o submodelo da família de distribuições  $\alpha_\tau$ -G reparametrizada.

A função de log-verossimilhança para o modelo (3.1), com vetor de parâmetros  $\Theta = (\boldsymbol{\psi}^\top(\tau), \boldsymbol{\theta})^\top$ , em que  $\boldsymbol{\theta}$  é o vetor de parâmetros de  $G$  de dimensão  $q - 1$ , pode ser escrita como  $\ell(\Theta | \mathbf{y}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta} | y_i)$ , em que

$$\ell_i(\boldsymbol{\theta} | y_i) = \log(\alpha_\tau) + \log[g(y_i; \boldsymbol{\theta})] + (\alpha_\tau - 1) \log[G(y_i; \boldsymbol{\theta})].$$

Os elementos da função escore para  $\psi_l(\tau)$ , com  $l = 1, \dots, p - 1$ , e  $\theta_j$ ,  $j = 1, \dots, q - 1$  são definidos respectivamente por:

$$U_{\psi_l(\tau)}(\Theta) = \frac{\partial \ell(\Theta)}{\partial \psi_l(\tau)} = \sum_{i=1}^n \underbrace{\left[ \frac{1}{g(y_i; \boldsymbol{\theta})} \frac{\partial g(y_i; \boldsymbol{\theta})}{\partial \eta_{\tau_i}} + (\alpha_\tau - 1) \frac{1}{G(y_i; \boldsymbol{\theta})} \frac{\partial G(y_i; \boldsymbol{\theta})}{\partial \eta_{\tau_i}} \right]}_{z_i} a_i x_{il}$$

e

$$U_{\theta_j}(\Theta) = \frac{\partial \ell(\Theta)}{\partial \theta_j} = \sum_{i=1}^n \underbrace{\left[ \frac{1}{g(y_i; \boldsymbol{\theta})} \frac{\partial g(y_i; \boldsymbol{\theta})}{\partial \theta_j} + (\alpha_\tau - 1) \frac{1}{G(y_i; \boldsymbol{\theta})} \frac{\partial G(y_i; \boldsymbol{\theta})}{\partial \theta_j} \right]}_{b_i},$$

em que  $a_i = \frac{d\eta_{\tau_i}}{d\phi_i(\tau)} = \frac{1}{h'(\eta_{\tau_i})}$ ,  $x_{il} = \frac{\partial \phi_i(\tau)}{\partial \psi_l(\tau)}$ ,  $h'(\eta_{\tau_i}) = \frac{dh}{d\eta_{\tau_i}}$  e, para  $U_{\theta_j}(\Theta)$  considere o parâmetro  $\eta_{\tau_i}$  constante. Em forma matricial temos, a função escore dada por:

$$\mathbf{U}_{\psi(\tau)}(\Theta) = \mathbf{X}^\top \mathbf{D}(\mathbf{a})\mathbf{z} \quad \text{e} \quad \mathbf{U}_{\boldsymbol{\theta}}(\Theta) = \text{tr}(\mathbf{D}(\mathbf{b})),$$

em que  $\mathbf{X}^\top = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ , com  $\mathbf{x}_i$  dado em (3.1),  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$ ,  $\mathbf{z} = (z_1, \dots, z_n)^\top$  e  $\mathbf{D}(\mathbf{b}) = \text{diag}\{b_1, b_2, \dots, b_n\}$ . Portanto, a função escore pode ser representada por:

$$\mathbf{U}(\Theta) = [\mathbf{U}_{\psi(\tau)}(\Theta), \mathbf{U}_{\boldsymbol{\theta}}(\Theta)]^\top.$$

Os estimadores de máxima verossimilhança (EMVs) para os parâmetros do modelo de regressão quantílica baseado na família de distribuições  $\alpha_\tau$ -G reparametrizada são soluções do sistema de equações  $\mathbf{U}(\Theta) = \mathbf{0}$ . Nota-se que esse sistema de equações não lineares não possui solução em forma fechada. Então, para estimar os parâmetros do modelo de regressão quantílica baseado na família de distribuições  $\alpha_\tau$ -G reparametrizada pode-se utilizar o método de máxima verossimilhança e o método numérico do tipo quase-Newton BFGS (PRESS et al., 1992), por exemplo. Dessa forma, as estimativas de máxima verossimilhança dos parâmetros que indexam o modelo de regressão quantílica baseado na família de distribuições  $\alpha_\tau$ -G reparametrizada foram obtidas com o software R, utilizando o método BFGS através da função `optim()`.

A matriz hessiana para o modelo de regressão quantílica baseado na família de distribuições  $\alpha_\tau$ -G reparametrizada é dada por:

$$\ddot{\mathbf{L}}(\Theta) = \begin{bmatrix} \frac{\partial^2 \ell(\Theta)}{\partial \psi(\tau) \partial \psi(\tau)^\top} & \frac{\partial^2 \ell(\Theta)}{\partial \psi(\tau) \partial \theta^\top} \\ \frac{\partial^2 \ell(\Theta)}{\partial \theta \partial \psi(\tau)^\top} & \frac{\partial^2 \ell(\Theta)}{\partial \theta \partial \theta^\top} \end{bmatrix} = \begin{bmatrix} \ddot{L}_{\psi(\tau)\psi(\tau)} & \ddot{L}_{\psi\theta} \\ \ddot{L}_{\theta\psi} & \ddot{L}_{\theta\theta} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{D}(\mathbf{c}) \mathbf{X} & \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{m} \\ \mathbf{m}^\top \mathbf{D}(\mathbf{a}) \mathbf{X} & \text{tr}(\mathbf{D}(\mathbf{d})) \end{bmatrix},$$

em que

$$\frac{\partial^2 \ell(\Theta)}{\partial \psi_l(\tau) \partial \psi_k(\tau)} = \sum_{i=1}^n c_i x_{il} x_{ik},$$

sendo  $c_i = d_{\eta_\tau^2}^{(i)} (a_i)^2 + d_{\eta_\tau}^{(i)} a'_i a_i$  e

$$d_{\eta_\tau}^{(i)} = \frac{\partial \ell(\Theta)}{\partial \eta_{\tau_i}} = \frac{1}{g(y_i; \theta)} \frac{\partial g(y_i; \theta)}{\partial \eta_{\tau_i}} + (\alpha_\tau - 1) \frac{1}{G(y_i; \theta)} \frac{\partial G(y_i; \theta)}{\partial \eta_{\tau_i}},$$

$$d_{\eta_\tau^2}^{(i)} = \frac{\partial^2 \ell(\Theta)}{\partial \eta_{\tau_i}^2} = - \left( \frac{1}{g^2(y_i; \theta)} \frac{\partial g(y_i; \theta)}{\partial \eta_{\tau_i}} \frac{\partial g(y_i; \theta)}{\partial \eta_{\tau_i}} \right) - (\alpha_\tau - 1) \left( \frac{1}{G^2(y_i; \theta)} \frac{\partial G(y_i; \theta)}{\partial \eta_{\tau_i}} \frac{\partial G(y_i; \theta)}{\partial \eta_{\tau_i}} \right) + \left( \frac{1}{g(y_i; \theta)} \frac{\partial^2 g(y_i; \theta)}{\partial \eta_{\tau_i} \partial \eta_{\tau_i}} \right) + \left( \frac{1}{G(y_i; \theta)} \frac{\partial^2 G(y_i; \theta)}{\partial \eta_{\tau_i} \partial \eta_{\tau_i}} \right) \quad e$$

$$a'_i = \frac{\partial}{\partial \eta_{\tau_i}} \left( \frac{d \eta_{\tau_i}}{\phi_i(\tau)} \right) = - \frac{h''(\eta_{\tau_i})}{[h'(\eta_{\tau_i})]^2}.$$

Aqui, denota-se a segunda derivada de  $h$  por  $h''$  e  $\ddot{L}_{\psi(\tau)\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{c}) \mathbf{X}$  foi escrito na forma matricial, em que  $\mathbf{D}(\mathbf{c}) = \text{diag}\{c_1, c_2, \dots, c_n\}$ . Além disso,

$$\frac{\partial^2 \ell(\Theta)}{\partial \psi(\tau) \partial \theta^\top} = \sum_{i=1}^n m_i a_i x_{i1},$$

em que

$$m_i = - \left( \frac{1}{g^2(y_i; \boldsymbol{\theta})} \frac{\partial g(y_i; \boldsymbol{\theta})}{\partial \eta_{\tau_i}} \frac{\partial g(y_i; \boldsymbol{\theta})}{\partial \theta_j} \right) - (\alpha_{\tau} - 1) \left( \frac{1}{G^2(y_i; \boldsymbol{\theta})} \frac{\partial G(y_i; \boldsymbol{\theta})}{\partial \eta_{\tau_i}} \frac{\partial G(y_i; \boldsymbol{\theta})}{\partial \theta_j} \right) + \left( \frac{1}{g(y_i; \boldsymbol{\theta})} \frac{\partial^2 g(y_i; \boldsymbol{\theta})}{\partial \eta_{\tau_i} \partial \theta_j} \right) + \left( \frac{1}{G(y_i; \boldsymbol{\theta})} \frac{\partial^2 G(y_i; \boldsymbol{\theta})}{\partial \eta_{\tau_i} \partial \theta_j} \right).$$

Consequentemente, obtém-se o elemento  $\ddot{L}_{\boldsymbol{\theta}\boldsymbol{\psi}} = \frac{\partial^2 \ell(\boldsymbol{\Theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\psi}(\tau)^\top}$ , uma vez que  $\ddot{\mathbf{L}}(\boldsymbol{\Theta})$  é uma matriz transposta. Na forma matricial, tem-se que  $\ddot{L}_{\boldsymbol{\psi}\boldsymbol{\theta}} = \mathbf{X}^\top \mathbf{D}(\mathbf{a})\mathbf{m}$  e  $\ddot{L}_{\boldsymbol{\theta}\boldsymbol{\psi}} = \mathbf{m}^\top \mathbf{D}(\mathbf{a})\mathbf{X}$ , com  $\mathbf{m} = (m_1, \dots, m_n)^\top$ . Adicionalmente,

$$\frac{\partial^2 \ell(\boldsymbol{\Theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \sum_{i=1}^n d_i,$$

em que

$$d_i = \left[ - \left( \frac{1}{g^2(y_i; \boldsymbol{\theta})} \frac{\partial g(y_i; \boldsymbol{\theta})}{\partial \theta_j} \frac{\partial g(y_i; \boldsymbol{\theta})}{\partial \theta_j} \right) - (\alpha_{\tau} - 1) \left( \frac{1}{G^2(y_i; \boldsymbol{\theta})} \frac{\partial G(y_i; \boldsymbol{\theta})}{\partial \theta_j} \frac{\partial G(y_i; \boldsymbol{\theta})}{\partial \theta_j} \right) + \left( \frac{1}{g(y_i; \boldsymbol{\theta})} \frac{\partial^2 g(y_i; \boldsymbol{\theta})}{\partial \theta_j \partial \theta_j} \right) + \left( \frac{1}{G(y_i; \boldsymbol{\theta})} \frac{\partial^2 G(y_i; \boldsymbol{\theta})}{\partial \theta_j \partial \theta_j} \right) \right].$$

Escrevendo na forma matricial, tem-se que  $\ddot{L}_{\boldsymbol{\theta}\boldsymbol{\theta}} = \text{tr}(\mathbf{D}(\mathbf{d}))$ , com  $\mathbf{D}(\mathbf{d}) = \text{diag}\{d_1, d_2, \dots, d_n\}$ .

### 3.3.1 Modelo de Regressão Quantílica Baseado no Modelo WER

O modelo WER foi proposto por Santos, Santos-Neto e Cysneiros (2024), denotado aqui por  $\text{WER}(\kappa, \eta_{\tau}, \alpha_{\tau})$ . Sejam  $Y_1, \dots, Y_n$  variáveis aleatórias independentes, em que cada  $Y_i \sim \text{WER}(\kappa, \eta_{\tau_i}, \alpha_{\tau})$ , para  $i = 1, \dots, n$  e  $\mathbf{y}_i = (y_1, \dots, y_n)$  são as observações correspondentes. Quando  $Y_i \sim \text{WER}(\kappa, \eta_{\tau_i}, \alpha_{\tau})$ , e considerando o componente sistemático dado por (3.1), temos o modelo de regressão quantílica baseado no modelo WER. A função de log-verossimilhança para o modelo de regressão quantílica baseado no modelo WER para  $\boldsymbol{\Theta} = (\boldsymbol{\psi}^\top(\tau), \kappa)^\top$ , pode ser escrita como  $\ell(\boldsymbol{\Theta} | \mathbf{y}) = \sum_{i=1}^n \ell_i(\kappa, \eta_{\tau_i}, \alpha_{\tau} | y_i)$ , em que

$$\begin{aligned} \ell_i(\kappa, \eta_{\tau_i}, \alpha_{\tau} | y_i) &= \log[\log(2)] + \log(\alpha_{\tau}) + \log(\kappa) + (\kappa - 1) \log(y_i) \\ &\quad - \frac{1}{\eta_{\tau_i}^{\kappa}} y_i^{\kappa} \log(2) - \kappa \log(\eta_{\tau_i}) + (\alpha_{\tau} - 1) \log\left(1 - 2^{-\frac{y_i^{\kappa}}{\eta_{\tau_i}^{\kappa}}}\right). \end{aligned}$$

Os elementos da função escore para  $\psi_l(\tau)$ , com  $l = 1, \dots, p - 1$ , e  $\kappa$ , são definidos respectivamente por:

$$U_{\psi_l(\tau)}(\boldsymbol{\Theta}) = \frac{\partial \ell(\boldsymbol{\Theta})}{\partial \psi_l} = \sum_{i=1}^n \mathbf{z}_i a_i x_{il}, \quad \text{em que,}$$

$$z_i = \log(2)y_i^\kappa \kappa \eta_{\tau_i}^{-(\kappa+1)} - \frac{\kappa}{\eta_{\tau_i}} + (\alpha_\tau - 1) \left[ \frac{\ln(2)y_i^\kappa \kappa \eta_{\tau_i}^{-(\kappa+1)}}{1 - 2 \frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right],$$

e

$$U_\kappa(\Theta) = \frac{\partial \ell(\Theta)}{\partial \kappa} = \sum_{i=1}^n b_i, \quad \text{com}$$

$$b_i = \frac{1}{\kappa} + \log(y_i) - \log(\eta_{\tau_i}) + \log(2)\eta_{\tau_i}^{-\kappa} y_i^\kappa [\ln(\eta_{\tau_i}) - \ln(y_i)] \\ + (\alpha_\tau - 1) \left[ \frac{\ln(2) \left( \frac{y_i}{\eta_{\tau_i}} \right)^\kappa \ln \left( \frac{y_i}{\eta_{\tau_i}} \right)}{2 \frac{y_i^\kappa}{\eta_{\tau_i}^\kappa} - 1} \right],$$

em que  $a_i = \frac{d\eta_{\tau_i}}{d\phi_i(\tau)} = \frac{1}{h'(\eta_{\tau_i})}$ ,  $x_{il} = \frac{\partial \phi_i(\tau)}{\partial \psi_l(\tau)}$  e  $h'(\eta_{\tau_i}) = \frac{dh}{d\eta_{\tau_i}}$ . Na forma matricial a função escore pode ser expressa por:

$$U_{\psi(\tau)}(\Theta) = \mathbf{X}^\top \mathbf{D}(\mathbf{a})\mathbf{z} \quad \text{e} \quad U_\kappa(\Theta) = \text{tr}(\mathbf{D}(\mathbf{b})),$$

em que  $\mathbf{X}^\top = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ , com  $\mathbf{x}_i$  dado em (3.1),  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$ ,  $\mathbf{z} = (z_1, \dots, z_n)^\top$  e  $\mathbf{D}(\mathbf{b}) = \text{diag}\{b_1, b_2, \dots, b_n\}$ . Portanto, a função escore pode ser escrita como:

$$\mathbf{U}(\Theta) = [U_{\psi(\tau)}(\Theta), U_\kappa(\Theta)]^\top.$$

Os elementos da matriz hessiana para o modelo de regressão quantílica WER são dados por:

$$\frac{\partial^2 \ell(\Theta)}{\partial \psi_l(\tau) \partial \psi_k(\tau)} = \sum_{i=1}^n c_i x_{il} x_{ik},$$

em que,  $c_i = d_{\eta_\tau^2}^{(i)}(a_i)^2 + d_{\eta_\tau}^{(i)} a'_i a_i$ , sendo:

$$d_{\eta_\tau}^{(i)} = \frac{\partial \ell(\Theta)}{\partial \eta_{\tau_i}} = \log(2)y_i^\kappa \kappa \eta_{\tau_i}^{-(\kappa+1)} - \frac{\kappa}{\eta_{\tau_i}} \\ + (\alpha_\tau - 1) \left[ \frac{\ln(2)y_i^\kappa \kappa \eta_{\tau_i}^{-(\kappa+1)}}{1 - 2 \frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right],$$

$$d_{\eta_\tau^2}^{(i)} = \frac{\partial^2 \ell(\Theta)}{\partial \eta_{\tau_i}^2} = \frac{\kappa}{\eta_{\tau_i}^2} - \log(2)y_i^\kappa \kappa(\kappa + 1)\eta_{\tau_i}^{-(\kappa+2)} \\ + (\alpha_\tau - 1) \left\{ -\ln(2)y_i^\kappa \kappa(\kappa + 1)\eta_{\tau_i}^{-(\kappa+2)} \left( 1 - 2 \frac{y_i^\kappa}{\eta_{\tau_i}^\kappa} \right) \right. \\ \left. - [\ln(2)]^2 (y_i)^{2\kappa} \kappa^2 \eta_{\tau_i}^{-2(\kappa+1)} 2 \frac{y_i^\kappa}{\eta_{\tau_i}^\kappa} \right\} \times \frac{1}{\left( 1 - 2 \frac{y_i^\kappa}{\eta_{\tau_i}^\kappa} \right)^2}$$

e

$$a' = \frac{\partial}{\partial \eta_{\tau_i}} \left( \frac{d\eta_{\tau_i}}{\phi_i(\tau)} \right) = -\frac{h''(\eta_{\tau_i})}{[h'(\eta_{\tau_i})]^2}.$$

Escrevendo na forma matricial, tem-se que  $\ddot{L}_{\psi(\tau)\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{c}) \mathbf{X}$ , em que  $\mathbf{D}(\mathbf{c}) = \text{diag}\{c_1, c_2, \dots, c_n\}$ . Adicionalmente,

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\Theta})}{\partial \psi(\tau) \partial \kappa} &= \sum_{i=1}^n m_i a_i x_{il}, \quad \text{em que,} \\ m_i &= -\frac{1}{\eta_{\tau_i}} + \log(2) y_i^\kappa \eta_{\tau_i}^{-(\kappa+1)} \left[ -\kappa \ln \left( \frac{\eta_{\tau_i}}{y_i} \right) + 1 \right] \\ &\quad + (\alpha_\tau - 1) \left\{ -\ln(2) y_i^\kappa \eta_{\tau_i}^{-(\kappa+1)} \left[ \kappa \ln \left( \frac{y_i}{\eta_{\tau_i}} \right) + 1 \right] \times \right. \\ &\quad \left. \left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} - 1 \right) + [\ln(2)]^2 \left( \frac{y_i}{\eta_{\tau_i}} \right)^\kappa \ln \left( \frac{y_i}{\eta_{\tau_i}} \right) \times \right. \\ &\quad \left. \left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right) y_i^\kappa \kappa \eta_{\tau_i}^{-(\kappa+1)} \right\} \times \frac{1}{\left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} - 1 \right)^2}, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\Theta})}{\partial \kappa \partial \psi(\tau)^\top} &= \sum_{i=1}^n m_i a_i x_{il} \quad \text{e} \\ m_i &= -\frac{1}{\eta_{\tau_i}} + \log(2) y_i^\kappa \eta_{\tau_i}^{-(\kappa+1)} \left[ -\kappa \ln \left( \frac{\eta_{\tau_i}}{y_i} \right) + 1 \right] \\ &\quad + (\alpha_\tau - 1) \left\{ -\ln(2) y_i^\kappa \eta_{\tau_i}^{-(\kappa+1)} \left[ \kappa \ln \left( \frac{y_i}{\eta_{\tau_i}} \right) + 1 \right] \times \right. \\ &\quad \left. \left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} - 1 \right) + [\ln(2)]^2 \left( \frac{y_i}{\eta_{\tau_i}} \right)^\kappa \ln \left( \frac{y_i}{\eta_{\tau_i}} \right) \times \right. \\ &\quad \left. \left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right) y_i^\kappa \kappa \eta_{\tau_i}^{-(\kappa+1)} \right\} \times \frac{1}{\left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} - 1 \right)^2}. \end{aligned}$$

Na forma matricial,  $\ddot{L}_{\psi\kappa} = \mathbf{X}^\top \mathbf{D}(\mathbf{a})\mathbf{m}$  e  $\ddot{L}_{\kappa\psi} = \mathbf{m}^\top \mathbf{D}(\mathbf{a})\mathbf{X}$ , com  $\mathbf{m} = (m_1, \dots, m_n)^\top$ . Finalmente,

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\Theta})}{\partial \kappa^2} &= \sum_{i=1}^n d_i, \quad \text{em que,} \\ d_i &= -\frac{1}{\kappa^2} - \log(2) \left[ \ln \left( \frac{\eta_{\tau_i}}{y_i} \right) \right]^2 \left( \frac{y_i}{\eta_{\tau_i}} \right)^\kappa \\ &\quad + (\alpha_\tau - 1) \left\{ \ln(2) \left[ \ln \left( \frac{y_i}{\eta_{\tau_i}} \right) \right]^2 \left( \frac{y_i}{\eta_{\tau_i}} \right)^\kappa \left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} - 1 \right) \right. \\ &\quad \left. - [\ln(2)]^2 \left( \frac{y_i}{\eta_{\tau_i}} \right)^{2\kappa} \left[ \ln \left( \frac{y_i}{\eta_{\tau_i}} \right) \right]^2 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right\} \times \frac{1}{\left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} - 1 \right)^2}. \end{aligned}$$

Na forma matricial, tem-se  $\ddot{L}_{\kappa\kappa} = \text{tr}(\mathbf{D}(\mathbf{d}))$ , com  $\mathbf{D}(\mathbf{d}) = \text{diag}\{d_1, d_2, \dots, d_n\}$ .

### 3.4 ANÁLISE DE RESÍDUOS E DIAGNÓSTICO DE INFLUÊNCIA

Nesta seção, apresentamos a distância de Cook generalizada (GCD), que é uma generalização da distância de Cook proposta por Cook (1977) e Cook (1986). A medida GCD é utilizada para detectar observações potencialmente influentes. Também apresentamos três tipos de resíduos, a saber: resíduos quantílicos, resíduo generalizado de Cox-Snell e resíduo tipo martingal. Eles são utilizados para analisar se o modelo é adequado e identificar observações potencialmente discrepantes. Além disso, esses três tipos de resíduos foram considerados com o objetivo de oferecer uma investigação mais completa do desempenho e ajuste dos modelos e verificar qual resíduo apresenta o melhor desempenho e ajuste. Também apresentamos o pseudo- $R^2$  para examinar o ajuste do modelo. Por fim, estudou-se a influência local e os esquemas de perturbação de casos, perturbação na resposta e perturbação na covariável contínua foram incluídos.

#### 3.4.1 Distância de Cook Generalizada

Nesta subseção, uma generalização da distância de Cook, descrita por Gallardo e Santos-Neto (2021), é usada para analisar a influência de observações individuais na análise do conjunto de dados utilizando RQ. Essa generalização da distância de Cook é dada por:

$$GCD_i(\boldsymbol{\Theta}) = \frac{1}{b} \left[ (\hat{\boldsymbol{\Theta}} - \hat{\boldsymbol{\Theta}}_{(i)})^\top \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\Theta}}}^{-1} (\hat{\boldsymbol{\Theta}} - \hat{\boldsymbol{\Theta}}_{(i)}) \right], \quad (i = 1, \dots, n),$$

em que  $\Theta = (\boldsymbol{\psi}(\tau), \boldsymbol{\theta})$ ,  $\boldsymbol{\theta}$  é o vetor de parâmetros de  $G$  de dimensão  $q - 1$ ,  $b = \dim(\Theta)$ ,  $\hat{\Sigma}_{\hat{\Theta}}$  representa a estimativa da matriz de variância-covariância de  $\hat{\Theta}$  e  $\hat{\Theta}_{(i)}$  é o estimador de máxima verossimilhança de  $\hat{\Theta}$  obtido quando excluído o caso  $i$ . O limiar utilizado neste trabalho para destacar as observações potencialmente influentes foi determinado pelos maiores valores da distância de Cook generalizada, nos quais os pontos estão mais destacados em relação aos demais. De acordo com Gallardo e Santos-Neto (2021), quando o interesse for só no vetor  $p \times 1$  de coeficientes de regressão,  $\boldsymbol{\psi}(\tau) = [\psi_0(\tau), \psi_1(\tau), \dots, \psi_{(p-1)}(\tau)]^\top$ , podemos reescrever a equação anterior como:

$$GCD_i(\hat{\boldsymbol{\psi}}(\tau)) = \frac{1}{p} \left[ (\hat{\boldsymbol{\psi}}(\tau) - \hat{\boldsymbol{\psi}}(\tau)_{(i)})^\top \hat{\Sigma}_{\hat{\boldsymbol{\psi}}(\tau)}^{-1} (\hat{\boldsymbol{\psi}}(\tau) - \hat{\boldsymbol{\psi}}(\tau)_{(i)}) \right] \quad (i = 1, \dots, n),$$

em que  $\hat{\Sigma}_{\hat{\boldsymbol{\psi}}(\tau)}$  representa a estimativa da matriz de variância-covariância de  $\hat{\boldsymbol{\psi}}(\tau)$ .

Além disso, foi obtida a alteração relativa percentual em valor absoluto (RC), para avaliar o impacto nas estimativas dos parâmetros do modelo para os casos influentes. A medida RC é obtida removendo os casos influentes e estimando novamente os parâmetros do modelo e seus correspondentes erros-padrão (s.e). A medida RC pode ser calculada como:

$$RC_{\Theta_{(i)}} = \left| \frac{\hat{\Theta}_j - \hat{\Theta}_{j(i)}}{\hat{\Theta}_j} \right| \times 100\% \quad \text{e} \quad RC_{se(\hat{\Theta}_{j(i)})} = \left| \frac{se(\hat{\Theta}_j) - se(\hat{\Theta}_{j(i)})}{se(\hat{\Theta}_j)} \right| \times 100\%,$$

em que  $\hat{\Theta}_{j(i)}$  é o EMV de  $\hat{\Theta}$  e  $se(\hat{\Theta}_{j(i)})$  representa seu s.e correspondente, calculado removendo o caso  $i$ .

### 3.4.2 Resíduos Quantílicos (RQ<sub>τ</sub>)

Nesta subseção, inclui-se os resíduos quantílicos aleatórios (RQ<sub>τ</sub>), que foram propostos por Dunn e Smyth (1996). Então, os RQ<sub>τ</sub> para a família de distribuições  $\alpha_\tau$ -G reparametrizada podem ser escritos como:

$$RQ_\tau = \Phi^{-1} \left[ G^{\alpha_\tau}(y_i; \hat{\boldsymbol{\theta}}) \right], \quad (i = 1, \dots, n),$$

em que  $\Phi(\cdot)$  é a função de distribuição acumulada da normal padrão,  $\hat{\boldsymbol{\theta}}$  é o vetor de estimativas de máxima verossimilhança dos parâmetros de  $G$  de dimensão  $q$  e  $G(y)$  é a distribuição *baseline* da família de distribuições  $\alpha_\tau$ -G reparametrizada. Quando o modelo ajustado é adequado para o conjunto de dados, os RQ<sub>τ</sub> seguem aproximadamente uma distribuição normal padrão. Sendo assim, pode-se utilizar o gráfico dos RQ<sub>τ</sub> para investigar a adequação do modelo ajustado e identificar observações atípicas.

### 3.4.3 Resíduo Generalizado de Cox-Snell

Cox e Snell (1968) recomendaram o resíduo generalizado de Cox-Snell (GCS). Esse resíduo para a família de distribuições  $\alpha_\tau$ -G reparametrizada é dado por:

$$R_{GCS_\tau} = -\log [1 - G^{\alpha_\tau}(y_i; \hat{\theta})], \quad (i = 1, \dots, n).$$

Quando o modelo é apropriado, o  $R_{GCS_\tau}$  tem aproximadamente uma distribuição exponencial padrão (com o parâmetro de taxa igual a um). Dessa forma, pode-se utilizar o gráfico dos  $R_{GCS_\tau}$  para examinar a adequação do modelo ajustado e identificar observações atípicas.

### 3.4.4 Resíduo Tipo Martingal

Therneau, Grambsch e Fleming (1990) propuseram o resíduo tipo martingal denotado, neste trabalho, por  $R_{MT_\tau}$ . O  $R_{MT_\tau}$  pode ser escrito da seguinte maneira:

$$R_{MT_\tau} = \text{sgn}(1 - R_{GCS_\tau}) \{-2[1 - R_{GCS_\tau} + \log(R_{GCS_\tau})]\}^{0,5},$$

em que  $\text{sgn}(\cdot)$  representa a função sinal expressa por:

$$\text{sgn}(x) = \begin{cases} -1, & \text{se } x < 0, \\ 0, & \text{se } x = 0, \\ 1, & \text{se } x > 0. \end{cases}$$

Quando o modelo é adequado, o  $R_{MT_\tau}$  tem aproximadamente uma distribuição normal padrão. O gráfico do resíduo tipo martingal é utilizado para analisar a adequação do modelo ajustado e detectar observações atípicas.

Portanto, especificando a função de distribuição acumulada do modelo *baseline*, podemos obter os resíduos  $RQ_\tau$ ,  $R_{GCS_\tau}$  e  $R_{MT_\tau}$  para o correspondente submodelo da família de distribuições  $\alpha_\tau$ -G reparametrizada.

### 3.4.5 Pseudo- $R^2$

O pseudo- $R^2$  foi introduzido por Maddala (1983). Para fins de simplificação da notação, ele é denotado neste trabalho por  $R_\tau^2$  e pode ser escrito como:

$$R_\tau^2 = 1 - \exp \left\{ \frac{2}{n} [\ell(\tilde{\theta}) - \ell(\hat{\theta})] \right\},$$

em que  $\ell(\tilde{\theta})$  representa a log-verossimilhança maximizada para o modelo de regressão sem covariável e  $\ell(\hat{\theta})$  é a log-verossimilhança maximizada para o modelo completo com todas as covariáveis. Além disso, o valor de  $R_\tau^2$  pertence ao intervalo  $[0, 1]$  e quanto mais próximo de um o valor de  $R_\tau^2$ , indica melhor ajuste do modelo.

### 3.4.6 Influência Local

A ideia de influência local foi introduzida por Cook (1986), para analisar a alteração nos resultados da análise quando se adicionam pequenas perturbações ao modelo.

Defina o vetor de perturbação  $\omega$ , em que  $\omega \in \mathbb{R}^n$ . Seja,  $\omega_0 \in \mathbb{R}^n$  o vetor que indica a ausência de perturbação no modelo. Então, a medida de afastamento da verossimilhança pode ser expressa por:

$$LD(\omega) = 2[\ell(\hat{\theta}) - \ell(\hat{\theta}_\omega)],$$

em que  $\ell(\theta)$  representa a função de log-verossimilhança do modelo apresentado na Equação (3.1), sem perturbação no modelo, ou seja,  $\ell(\theta_{\omega_0}) = \ell(\theta)$ ,  $\hat{\theta}$  simboliza o estimador de máxima verossimilhança de  $\theta$ ,  $\ell(\theta_\omega)$  representa a função de log-verossimilhança do modelo perturbado e  $\hat{\theta}_\omega$  é o estimador de máxima verossimilhança de  $\theta_\omega$  obtido de  $\ell(\hat{\theta}_\omega)$ . Além disso, suponha que  $\ell(\theta_\omega)$  possui as duas primeiras derivadas contínuas em  $(\hat{\theta}^\top, \omega_0^\top)$ . Então, Cook (1986) recomendou analisar o comportamento do afastamento da verossimilhança na vizinhança de  $\omega_0$ , considerando  $LD(\omega_0) = 0$ . Cook (1986) também propôs estudar a curvatura normal na direção de um vetor  $\mathbf{d}$  com  $\|\mathbf{d}\| = 1$  e fazer o gráfico de  $LD(\omega_0 + a\mathbf{d})$  versus  $a$ , com  $a \in \mathbb{R}$ . Adicionalmente, o autor demonstrou que a curvatura normal para  $\theta$  na direção de  $\mathbf{d}$  com  $\|\mathbf{d}\| = 1$  pode ser escrita como:

$$C_d(\theta) = 2|\mathbf{d}^\top \Delta^\top (\ddot{\mathbf{I}}(\theta))^{-1} \Delta \mathbf{d}|, \quad (3.2)$$

sendo  $\ddot{\mathbf{I}}(\theta)$  a matriz de informação de Fisher observada e  $\Delta$  a matriz  $(p+q) \times n$  associada ao esquema de perturbação com elementos definidos por:

$$\Delta_{ji} = \left. \frac{\partial^2 \ell(\theta; \omega)}{\partial \theta_j \partial \omega_i} \right|_{\theta=\hat{\theta}, \omega=\omega_0}, \quad i = 1, \dots, n \quad \text{e} \quad j = 1, \dots, p+q,$$

em que  $\ell(\theta; \omega)$  representa a função log-verossimilhança do modelo perturbado pelo vetor  $\omega$ . Considerando o modelo (3.1), os cálculos de  $\ddot{\mathbf{I}}(\theta) = -\ddot{\mathbf{L}}(\theta)$  são apresentados na Seção 3.3.

Cook (1986) ainda apresentou a curvatura máxima  $C_{d_{max}}$  na direção de  $d_{max}$ . De acordo com Cordeiro, Ortega e Silva (2011),  $C_{d_{max}}$  na direção de  $d_{max}$  revela qual tipo de perturbação no modelo gera a maior alteração local em  $LD(\omega)$ . Considerando  $d = d_{max}$  na Equação (3.2) obtemos a curvatura máxima  $C_{d_{max}}$ , em que  $d_{max}$  simboliza o autovetor normalizado correspondente ao maior autovalor da matriz  $B = \Delta^T \ddot{I}(\theta)^{-1} \Delta$ . Sendo assim, o gráfico mais comum de se encontrar na literatura é o gráfico dos elementos de  $d_{max}$  contra a ordem dos casos, utilizado para detectar casos que são potencialmente influentes.

Lesaffre e Verbeke (1998), introduziram a influência local total,  $C_i = C_{d_i}(\theta)$ , sendo  $d_i$  um vetor,  $(n \times 1)$ , de zeros com um na  $i$ -ésima posição. Dessa forma, a curvatura na direção de  $d_i$  é calculada por  $C_i = 2|\Delta_i^T (\ddot{I}(\theta))^{-1} \Delta_i|$ , em que  $\Delta_i^T$  representa a  $i$ -ésima linha de  $\Delta$ . Neste trabalho, utilizou-se o gráfico de  $C_i$  versus a ordem dos casos para identificar elementos que são potencialmente influentes. Adicionalmente, observações com  $C_i(\hat{\theta}) > 2\bar{C}(\hat{\theta})$ , em que,  $\bar{C}(\hat{\theta}) = \sum_{i=1}^n C_i(\hat{\theta})/n$  são classificadas como potencialmente influentes.

### 3.4.7 Esquemas de Perturbação

Nesta subseção, apresentamos alguns esquemas de perturbação, a saber: os esquemas de perturbação de casos, perturbação na resposta e perturbação na covariável contínua. Adicionalmente, esses esquemas de perturbação são mostrados para a família de distribuições  $\alpha_\tau$ -G reparametrizada. Então, escolhendo o modelo *baseline*, pode-se obter o correspondente esquema de perturbação para o submodelo da família de distribuições  $\alpha_\tau$ -G reparametrizada.

#### 3.4.7.1 Perturbação de Casos

Defina o vetor de pesos  $\omega = (\omega_1, \dots, \omega_n)^T$ . Então, nesse esquema de perturbação, o logaritmo da função de verossimilhança perturbada para a família de distribuições  $\alpha_\tau$ -G reparametrizada pode ser escrito como:

$$\ell(\theta | \omega) = \sum_{i=1}^n \omega_i \ell_i(\theta),$$

em que  $0 \leq \omega_i \leq 1$  e  $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$ . A matriz de perturbação pode ser escrita como:

$$\Delta = \left( \begin{array}{c} \Delta_{\psi(\tau)} \\ \Delta_{\theta_j} \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

em que os elementos de  $\Delta$  podem ser escritos como,  $\Delta_{\psi(\tau)} = \mathbf{x}_i a_i \mathbf{z}_i$  e  $\Delta_{\theta_j} = b_i$  para  $i = 1, \dots, n$ . Note que,  $a_i$ ,  $z_i$  e  $b_i$  para a família de distribuições  $\alpha_\tau$ -G reparametrizada foram apresentados na Seção 3.3. Na forma matricial  $\Delta_{\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{z}$  e  $\Delta_{\theta_j} = (b_1, \dots, b_n)^\top$ , sendo  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$  e  $\mathbf{z} = (z_1, \dots, z_n)^\top$ .

### 3.4.7.2 Perturbação na Resposta

A perturbação na resposta permite detectar a sensibilidade do modelo quando é inserida uma perturbação aditiva na variável resposta dada por:

$$y_i(\omega_i) = y_i + \omega_i s_y, \quad (i = 1, \dots, n),$$

em que  $\omega_i \in \mathbb{R}$  e  $s_y$  é o desvio padrão amostral de  $Y$ . Portanto, o logaritmo da função de verossimilhança perturbada,  $\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega})$ , para a família de distribuições  $\alpha_\tau$ -G reparametrizada pode ser calculado substituindo  $y_i$  por  $y_i(\omega_i)$  em  $\ell(\boldsymbol{\Theta} \mid \mathbf{y}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta} \mid y_i)$ . Além disso, o vetor de não perturbação é denotado por  $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . A matriz  $\Delta$  pode ser escrita da seguinte forma:

$$\Delta = \left( \begin{array}{c} \Delta_{\psi(\tau)} \\ \Delta_{\theta_j} \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

com  $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . Os elementos de  $\Delta$  são dados por  $\Delta_{\psi(\tau)} = \mathbf{x}_i a_i \Psi_i$  e  $\Delta_{\theta_j} = \nu_i$  para  $i = 1, \dots, n$ , em que:

$$\Psi_i = s_{y_i} \left( \frac{\partial z_i}{\partial y_i} \right) \quad \text{e} \quad \nu_i = s_{y_i} \left( \frac{\partial b_i}{\partial y_i} \right).$$

Escrevendo na forma matricial, ou seja,  $\Delta_{\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{D}(\boldsymbol{\Psi})$  e  $\Delta_{\theta_j} = (\nu_1, \dots, \nu_n)^\top$ , em que  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$  e  $\mathbf{D}(\boldsymbol{\Psi}) = \text{diag}\{\Psi_1, \Psi_2, \dots, \Psi_n\}$ .

### 3.4.7.3 Perturbação na Covariável Contínua

A perturbação em uma covariável contínua permite analisar a sensibilidade do modelo quando é inserida uma perturbação aditiva em uma covariável contínua,  $X_t$  para  $t =$

$1, \dots, p-1$ , dada por:

$$x_{ti}(\omega_i) = x_{ti} + \omega_i s_{X_t}, \quad (i = 1, \dots, n),$$

em que  $s_{X_t}$  é o desvio padrão amostral da covariável contínua  $X_t$ . Além disso,  $X_t, \omega_i \in \mathbb{R}$ . Então, o logaritmo da função de verossimilhança perturbada,  $\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega})$ , para a família de distribuições  $\alpha_\tau$ -G reparametrizada pode ser calculado substituindo  $x_{ti}$  por  $x_{ti}(\omega_i)$  em  $\ell(\boldsymbol{\Theta} \mid \mathbf{y}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta} \mid y_i)$ . Sendo assim, o vetor de não perturbação é denotado por  $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . Note que, considerando a Equação (3.1), tem-se que  $\eta_{\tau_i}(\omega_i) = h^{-1} [\mathbf{x}_i^\top(\omega_i) \boldsymbol{\psi}(\tau)]$  e  $\mathbf{x}_i^\top(\omega_i) = (1, x_{i1}, \dots, x_{ti}(\omega_i), \dots, x_{i(p-1)})$ . Portanto, a matriz  $\Delta$  é dada por:

$$\Delta = \left( \begin{array}{c} \Delta_{\psi(\tau)} \\ \Delta_{\theta_j} \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

em que  $\Delta_{\psi(\tau)} = (\Delta_{\psi(\tau)_{il}})$  é uma matriz  $p \times n$  expressa por:

$$\Delta_{\psi(\tau)_{il}} = \begin{cases} s_{X_t} \psi(\tau)_t a'_i a_i x_{il} z_i + s_{X_t} \psi(\tau)_t x_{il} a_i^2 c_i & \text{se } l \neq t, \quad i = 1, \dots, n, \\ s_{X_t} a_i z_i + s_{X_t} \psi(\tau)_t a'_i a_i x_{it} z_i + s_{X_t} \psi(\tau)_t x_{it} a_i^2 c_i & \text{se } l = t, \quad i = 1, \dots, n. \end{cases}$$

Adicionalmente, os elementos de  $\Delta_{\theta_j}$  são dados por  $\Delta_{\theta_j} = (\zeta_1, \dots, \zeta_n)$ , em que  $\zeta_i = s_{X_t} \psi(\tau)_t a_i m_i$ . Observa-se que,  $a_i, z_i, a', m_i$  e  $c_i$  para a família de distribuições  $\alpha_\tau$ -G reparametrizada foram desenvolvidos na Seção 3.3.

Nos Apêndices A, B e C são apresentados os cálculos desses esquemas de perturbação para os modelos de regressão quantílica WER, LLER e LER respectivamente.

### 3.5 SIMULAÇÃO

Nesta seção, são apresentados os resultados numéricos de um estudo de simulação de Monte Carlo, para avaliar o desempenho e as propriedades dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de RQ baseado no modelo WER. Esta simulação foi realizada usando o software R em sua versão 3.5.1, e utilizou-se o método numérico BFGS por meio da função `optim()`. Os valores dos chutes iniciais para os parâmetros do modelo de RQ WER foram obtidos do ajuste do modelo *baseline*. As medidas utilizadas para analisar empiricamente o desempenho e as propriedades dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de RQ WER foram:

coeficiente de assimetria (CS), coeficiente de curtose (CK), variância (VAR), viés relativo percentual em valor absoluto (VR) e erro quadrático médio (EQM). Na simulação foram realizadas 10000 réplicas de Monte Carlo. Adicionalmente, para observar o comportamento dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de RQ WER quando  $n$  cresce, utilizou-se os tamanhos amostrais  $n \in \{10, 30, 50, 100, 200\}$ . Também, considere a estrutura de regressão quantílica baseada na Equação (3.1), para o  $\tau$ -ésimo quantil do modelo WER dada por:

$$\log(\eta_{\tau_i}) = \psi_0(\tau) + \psi_1(\tau)x_{i1}, \quad i = 1, \dots, n, \quad n \in \{10, 30, 50, 100, 200\},$$

em que  $x_{i1}$  é distribuído uniformemente no intervalo  $(0, 1)$  e os valores gerados nesta variável são fixos em cada replicação. É importante observar que o modelo de RQ Weibull, proposto por Sánchez et al. (2021b) é um caso particular do modelo de RQ WER quando  $\tau = 0,5$ . Foi utilizada a função de ligação logarítmica, uma vez que Sánchez et al. (2021b), mostrou, por meio de um conjunto de dados reais, que a função de ligação logarítmica tem um melhor ajuste ao conjunto de dados em comparação com a função de ligação raiz quadrada para regressão mediana ( $\tau = 0,5$ ) do modelo de RQ Weibull. Os verdadeiros valores dos parâmetros foram fixados em  $\psi_0(\tau) = 0,5$ ,  $\psi_1(\tau) = 1,5$  e  $\kappa \in \{0,5, 2,5, 3,5\}$ , que indica modelos com diferentes níveis de curtose. Por exemplo, para  $\tau = 0,5$  fixo, e  $\kappa \in \{0,5, 2,5, 3,5\}$  os valores da curtose foram 0,1516, 3,6613 e 4,2233 para os respectivos valores de  $\kappa$ . Também considere  $\tau \in \{0,25, 0,5, 0,75\}$ . O interesse, neste caso, é observar o comportamento dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de RQ WER para diferentes valores de  $\tau$ .

Utilizou-se um algoritmo para gerar números pseudo-aleatórios para o modelo WER usando o método da transformação inversa. O algoritmo consiste nos seguintes passos:

**Passo 1:** Gere um número aleatório  $U \sim (0,1)$ ;

**Passo 2:** Defina valores para  $\kappa$ ,  $\eta_\tau$  e  $\tau$  de  $Y \sim \text{WER}(\kappa, \eta_\tau, \alpha_\tau)$ ;

**Passo 3:** Faça  $Y = \left[ \frac{\eta_\tau}{\log(2)^{1/\kappa}} \right] \times \left[ -\log(1 - U^{1/\alpha_\tau}) \right]^{1/\kappa}$  e pare.

Aqui,  $\alpha_\tau = -\log(\tau)/\log(2)$  fixo e  $U \sim (0,1)$  indica que  $U$  é uma variável aleatória com distribuição uniforme no intervalo  $(0, 1)$ . Por meio do algoritmo foram produzidos números pseudo-aleatórios para gerar variáveis aleatórias independentes  $Y_i$ . Além disso, foi calculado

o valor de  $\eta_{\tau_i}$  usando o fato que  $\eta_{\tau_i} = \exp\{\psi_0(\tau) + \psi_1(\tau)x_{i1}\}$ . Para cada uma das 10000 réplicas de Monte Carlo obtemos as observações  $\mathbf{y} = (y_1, \dots, y_n)^\top$  para o modelo WER indexado pelos parâmetros  $\kappa$  e  $\eta_{\tau_i}$ .

Na Tabela 7, é possível observar de forma geral que, independentemente dos valores de  $\kappa$  e  $\tau$ , à medida que o tamanho da amostra aumenta, a variância, o viés relativo percentual em valor absoluto e o erro quadrático médio diminuem. Sendo assim, com o aumento do tamanho da amostra, os valores das estimativas se aproximam dos valores verdadeiros dos parâmetros do modelo de RQ WER, ou seja, os estimadores dos parâmetros apresentaram melhor desempenho com o aumento da amostra. Este resultado também indica que, há fortes indícios de que os estimadores de máxima verossimilhança dos parâmetros do modelo de RQ WER são consistentes.

Além disso, percebe-se que, para valores de  $n$  e  $\kappa$  fixos, à medida que o valor de  $\tau$  aumenta, a variância, o viés relativo percentual em valor absoluto e o erro quadrático médio para o parâmetro  $\kappa$  também aumentam. Então, o estimador  $\hat{\kappa}$  apresentou melhor desempenho para valores menores de  $\tau$ .

Nota-se ainda que, para valores de  $n$  e  $\kappa$  fixos os estimadores  $\hat{\psi}_0(\tau)$  e  $\hat{\psi}_1(\tau)$  apresentaram pior desempenho para  $\tau = 0,75$ . Adicionalmente, pode-se observar na Tabela 7, que para valores de  $n$  e  $\tau$  fixos, à medida que o valor do parâmetro  $\kappa$  aumenta o estimador  $\hat{\kappa}$  apresentou comportamento semelhante com relação ao viés relativo percentual em valor absoluto. Por outro lado, a variância e o erro quadrático médio aumentaram. Então, quando o valor do parâmetro  $\kappa$  cresce, as estimativas do parâmetro  $\kappa$  pioram.

Também, nota-se que há fortes indícios de que o tamanho da amostra é pequeno para  $n = 10$ , uma vez que as estimativas dos parâmetros apresentaram variância e erro quadrático médio grandes em alguns cenários da simulação considerando  $n = 10$ , isso está relacionado ao fato de que há fortes indícios de que o tamanho da amostra é pequeno para  $n = 10$ . Por fim, nota-se ainda que, à medida que o tamanho da amostra aumenta, o coeficiente de assimetria vai se aproximando de zero (assimetria da distribuição normal) e o coeficiente de curtose está se aproximando de 3 (curtose da distribuição normal). Logo, há fortes indícios de que os estimadores dos parâmetros do modelo de RQ WER são distribuídos assintoticamente como uma variável aleatória normal.

Tabela 7: Estimativas do CS, CK, VAR, VR e EQM para  $\psi_0(\tau) = 0,5$ ,  $\psi_1(\tau) = 1,5$  fixos,  $\kappa$ : 0,5, 2,5 e 3,5 e  $\tau$ : 0,25, 0,5 e 0,75.

$\kappa$	$\tau$	$n$	CS			CK			VAR			VR			EQM		
			$\hat{\kappa}$	$\hat{\psi}_0(\tau)$	$\hat{\psi}_1(\tau)$												
0,5	0,25	10	1,61916	-0,22388	0,00699	8,08444	3,07211	3,11126	0,03802	1,97733	4,31600	25,25838	20,82622	2,29585	0,05397	1,98818	4,31719
		30	0,68029	-0,14464	0,00814	3,94168	3,02806	3,07052	0,00608	0,39213	0,84832	6,83885	6,11159	1,26367	0,00726	0,39306	0,84868
		50	0,47646	-0,09976	-0,03411	3,39204	3,00443	3,01585	0,00329	0,21186	0,46408	3,95125	4,95509	0,03053	0,00368	0,21247	0,46408
		100	0,29594	-0,07538	-0,02132	3,14363	2,98499	3,04971	0,00146	0,11648	0,28572	1,80553	2,20995	0,21719	0,00154	0,11660	0,28574
	200	0,24571	-0,08668	0,00716	3,18957	2,98234	2,91936	0,00072	0,05504	0,13660	0,87981	0,89135	0,02275	0,00074	0,05506	0,13660	
	0,50	10	1,59707	-0,35788	0,04124	7,90168	3,35751	3,32171	0,04324	3,72632	8,66423	27,33463	20,29526	6,02210	0,06192	3,73662	8,67239
		30	0,69772	-0,22043	0,01616	4,06746	3,07134	3,17081	0,00674	0,66539	1,59909	7,30519	6,97238	2,41447	0,00807	0,66661	1,60041
		50	0,47435	-0,16869	-0,01384	3,39441	3,11748	3,09157	0,00363	0,34753	0,85741	4,22383	1,54237	0,02917	0,00408	0,34759	0,85741
		100	0,30676	-0,11157	-0,02464	3,19465	2,99258	3,07135	0,00161	0,19064	0,52112	1,90743	0,83508	0,34849	0,00169	0,19066	0,52114
	200	0,24818	-0,10112	0,00808	3,22351	3,01818	2,93088	0,00079	0,08873	0,24864	0,93871	0,52132	0,12910	0,00081	0,08874	0,24865	
	0,75	10	1,62102	-0,61018	0,13577	8,05478	4,16043	3,92663	0,05435	10,51835	25,16381	31,63200	125,00470	18,70670	0,07936	10,90901	25,24255
		30	0,72974	-0,38249	0,03929	4,21753	3,37804	3,46276	0,00804	1,57844	3,99128	8,25825	37,67104	5,46170	0,00975	1,61392	3,99799
50		0,48685	-0,31277	0,02151	3,43115	3,35460	3,24566	0,00431	0,77650	2,03489	4,77368	16,54027	0,11279	0,00488	0,78334	2,03490	
100		0,33055	-0,18691	-0,02143	3,26016	3,06326	3,10533	0,00190	0,41508	1,19973	2,14270	7,60691	0,58465	0,00201	0,41656	1,19981	
200	0,25402	-0,14636	0,01617	3,24360	3,11327	2,98546	0,00093	0,18892	0,56723	1,05963	3,7339	0,38546	0,00096	0,18927	0,56727		
2,5	0,25	10	1,61932	-0,22378	0,00707	8,08503	3,07230	3,11114	0,95054	0,07908	0,17263	25,25747	4,16576	0,45874	1,34926	0,07952	0,17268
		30	0,68090	-0,14401	0,00815	3,94343	3,02815	3,07093	0,15217	0,01568	0,03392	6,83313	1,21373	0,25349	0,18135	0,01572	0,03394
		50	0,47661	-0,09988	-0,03417	3,39288	3,00369	3,01574	0,08213	0,00847	0,01856	3,95162	0,99276	0,00521	0,09189	0,00849	0,01856
		100	0,29780	-0,07565	-0,02084	3,14544	2,98642	3,05289	0,03643	0,00465	0,01141	1,79638	0,43720	0,04713	0,03845	0,00466	0,01142
	200	0,24548	-0,08764	0,00703	3,18810	2,98302	2,91933	0,01794	0,00220	0,00546	0,88081	0,18002	0,00510	0,01843	0,00220	0,00546	
	0,50	10	1,59721	-0,35795	0,04132	7,90262	3,35819	3,32228	1,08099	0,14902	0,34651	27,33458	4,05738	1,20379	1,54798	0,14944	0,34684
		30	0,69758	-0,21992	0,01570	4,06760	3,07153	3,17160	0,16844	0,02661	0,06396	7,30738	1,39798	0,48697	0,20182	0,02667	0,06401
		50	0,47420	-0,16856	-0,01356	3,39373	3,11788	3,09256	0,09077	0,01389	0,03428	4,22318	0,31099	0,00607	0,10192	0,01390	0,03429
		100	0,30704	-0,11212	-0,02375	3,19513	2,99154	3,07182	0,04017	0,00762	0,02083	1,90617	0,16706	0,07284	0,04245	0,00762	0,02083
	200	0,24794	-0,10204	0,00867	3,22354	3,01882	2,93259	0,01977	0,00354	0,00994	0,93761	0,10434	0,02762	0,02032	0,00355	0,00994	
	0,75	10	1,62107	-0,61048	0,13611	8,05508	4,16183	3,92738	1,35876	0,42070	1,00647	31,63153	24,99840	3,74005	1,98411	0,43633	1,00962
		30	0,72983	-0,38260	0,03939	4,21778	3,37801	3,46271	0,20121	0,06313	0,15964	8,25818	7,53434	1,09275	0,24384	0,06456	0,15992
50		0,48668	-0,31240	0,02112	3,43084	3,35432	3,24545	0,10777	0,03106	0,08140	4,77468	3,31017	0,02488	0,12202	0,03134	0,08140	
100		0,33047	-0,18744	-0,02193	3,25933	3,06309	3,10483	0,04752	0,01659	0,04800	2,14224	1,54844	0,11344	0,05039	0,01666	0,04801	
200	0,25373	-0,14660	0,01642	3,24419	3,11407	2,98601	0,02327	0,00755	0,02268	1,06005	0,74538	0,07863	0,02397	0,00757	0,02268		
3,5	0,25	10	1,61931	-0,22383	0,00709	8,08546	3,07247	3,11128	1,86312	0,04034	0,08807	25,25900	2,97723	0,32713	2,64469	0,04057	0,08809
		30	0,68094	-0,14400	0,00822	3,94314	3,02811	3,07106	0,29830	0,00800	0,01731	6,83189	0,86701	0,18060	0,35548	0,00802	0,01732
		50	0,47638	-0,09988	-0,03408	3,39208	3,00376	3,01574	0,16103	0,00432	0,00947	3,95119	0,70885	0,00364	0,18015	0,00434	0,00947
		100	0,29768	-0,07560	-0,02084	3,14509	2,98625	3,05295	0,07140	0,00237	0,00582	1,79695	0,31286	0,03378	0,07536	0,00238	0,00583
	200	0,24526	-0,08783	0,00690	3,18901	2,98293	2,91936	0,03514	0,00112	0,00278	0,87972	0,12806	0,00378	0,03609	0,00112	0,00279	
	0,50	10	1,59729	-0,35801	0,04138	7,90191	3,35828	3,32229	2,11878	0,07603	0,17679	27,33311	2,89762	0,85932	3,03398	0,07624	0,17695
		30	0,69757	-0,22009	0,01581	4,06767	3,07148	3,17156	0,33015	0,01358	0,03263	7,30626	0,99866	0,34744	0,39554	0,01360	0,03266
		50	0,47450	-0,16861	-0,01346	3,39567	3,11780	3,09232	0,17787	0,00709	0,01749	4,22216	0,22159	0,00380	0,19971	0,00709	0,01749
		100	0,30696	-0,11214	-0,02368	3,19526	2,99140	3,07178	0,07874	0,00388	0,01062	1,90584	0,11908	0,05228	0,08319	0,00389	0,01063
	200	0,24705	-0,10261	0,00937	3,22097	3,01930	2,93274	0,03876	0,00181	0,00507	0,94144	0,07134	0,02067	0,03985	0,00181	0,00507	
	0,75	10	1,62118	-0,61046	0,13610	8,05482	4,16188	3,92747	2,66336	0,21464	0,51351	31,63258	17,85483	2,67090	3,88913	0,22262	0,51512
		30	0,73016	-0,38247	0,03922	4,21862	3,37789	3,46273	0,39434	0,03221	0,08144	8,25643	5,38234	0,78041	0,47785	0,03293	0,08159
50		0,48662	-0,31232	0,02113	3,43064	3,35385	3,24538	0,21124	0,01584	0,04153	4,77480	2,36308	0,01704	0,23917	0,01599	0,04153	
100		0,33001	-0,18752	-0,02185	3,25818	3,06318	3,10518	0,09314	0,00846	0,02448	2,14249	1,10606	0,08089	0,09877	0,00849	0,02449	
200	0,25330	-0,14583	0,01594	3,24375	3,11270	2,98541	0,04561	0,00385	0,01157	1,06166	0,52626	0,05907	0,04699	0,00386	0,01158		

Fonte: O autor (2025).

Adicionalmente, os resultados obtidos nas simulações deste capítulo da tese corroboram com o estudo de Sánchez et al. (2021b). Esses autores realizaram um estudo de simulação para o modelo de RQ Weibull, que é um caso especial do modelo de RQ WER e revelaram que, há fortes indícios de que os estimadores dos parâmetros do modelo de RQ Weibull são consistentes e assintoticamente distribuídos como uma variável aleatória normal.

### 3.6 APLICAÇÃO

A regressão quantílica tem larga aplicabilidade em diversas áreas de pesquisas, inclusive na área da economia. Nesta seção, aplicam-se os modelos de RQ WER, LLER e LER a um conjunto de dados reais, na área de economia. O conjunto de dados possui 102 observações, pode ser obtido pelo software R (TEAM, 2020), utilizando o pacote carData. Esse conjunto de dados é denominado de Prestige e contém informações sobre ocupações (cargo ou profissão) no Canadá. O conjunto de dados Prestige foi analisado por Fox (2015), no contexto de modelos de regressão, em que modelou-se a média da variável resposta. **Esta é a primeira vez na literatura que o conjunto de dados Prestige é modelado utilizando regressão quantílica.** A variável resposta,  $Y$ , corresponde a renda média das ocupações em 1971 (em dólares), já as variáveis explicativas são,  $X_1$ , que representa a educação média de cada profissão em 1971 (medida em anos) e  $X_2$ , que indica a porcentagem de mulheres no cargo.

O objetivo desta aplicação foi modelar a renda dos mais pobres que pertencem aos 25% das rendas mais baixas, utilizando os modelos de RQ WER, LLER e LER. Então, considere  $\tau = 0,25$  fixo. Também foi realizado um estudo comparativo para conhecer o modelo que apresenta o melhor ajuste ao conjunto de dados.

Na Tabela 8, mostramos algumas estatísticas descritivas para o conjunto de dados Prestige. Nota-se que, a renda média é 6797,90 dólares. A distribuição da renda é assimétrica positiva e apresenta uma forma leptocúrtica. Além disso, a menor renda observada foi 611 dólares e a maior renda foi 25879 dólares. Nota-se ainda que, o primeiro quartil  $Q_1$  é igual a 4106,00, ou seja 25% dos valores da renda são menores ou iguais a 4106,00 dólares. Observa-se que, o coeficiente de variação percentual (CV%), foi igual a 62,46. Além disso, para a variável independente  $X_1$ , a média de anos de educação de cada profissão é 10,74 anos. O primeiro quartil  $Q_1$  foi igual a 8,44, isto é 25% dos valores da educação média de

cada profissão são menores ou iguais a 8,44 anos. O menor valor da educação média de cada profissão foi 6,38 anos e o maior valor foi 15,97 anos de estudo. Nota-se ainda que, o coeficiente de variação percentual (CV%), foi igual a 25,41. Já para a variável independente  $X_2$ , a porcentagem média de mulheres nos cargos foi 28,98. Percebe-se que, o primeiro quartil  $Q_1$  foi igual a 3,59, ou seja 25% dos valores da porcentagem de mulheres nos cargos são menores ou iguais a 3,59. Nota-se também que, o coeficiente de variação percentual (CV%), foi igual a 109,47.

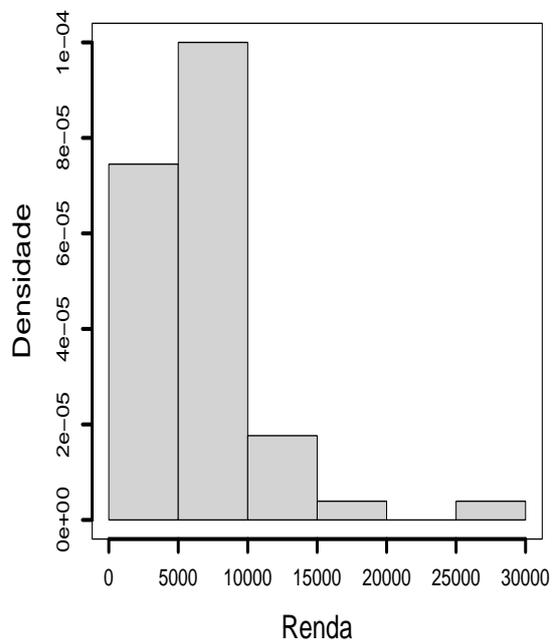
Tabela 8: Estatísticas descritivas para o conjunto de dados Prestige.

Variável	Média	$Q_1$	Mediana	$Q_3$	Variância	Assimetria	Curtose	Mínimo	Máximo	CV%
$Y$	6797,90	4106,00	5930,50	8187,25	18027860	2,13	6,29	611	25879	62,46
$X_1$	10,74	8,44	10,54	12,65	7,44	0,32	-1,03	6,38	15,97	25,41
$X_2$	28,98	3,59	13,60	52,20	1006,47	0,89	-0,67	00,00	97,51	109,47

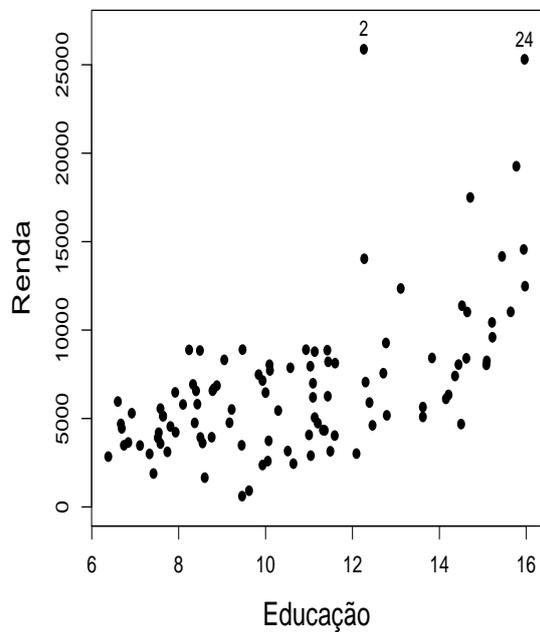
Fonte: O autor (2025).

Na Figura 10, apresentamos o histograma da renda e os gráficos de dispersão. Percebe-se na Figura 10-(a), que o histograma da renda apresenta comportamento assimétrico com a maioria dos valores da renda entre 0 e 10000 dólares. Nota-se que, na Figura 10-(b) existe uma correlação positiva entre renda e educação, ou seja, à medida que a educação média de cada profissão aumenta, a renda também aumenta. Nota-se também que, na Figura 10-(c) existe uma correlação negativa entre renda e porcentagem de mulheres no cargo, ou seja quando aumenta a porcentagem de mulheres no cargo a renda diminui. Além disso, os gráficos de dispersão detectaram dois casos atípicos que são as observações #2 e #24 desse conjunto de dados. Nota-se ainda que, as observações #2 e #24 são as maiores rendas observadas neste conjunto de dados.

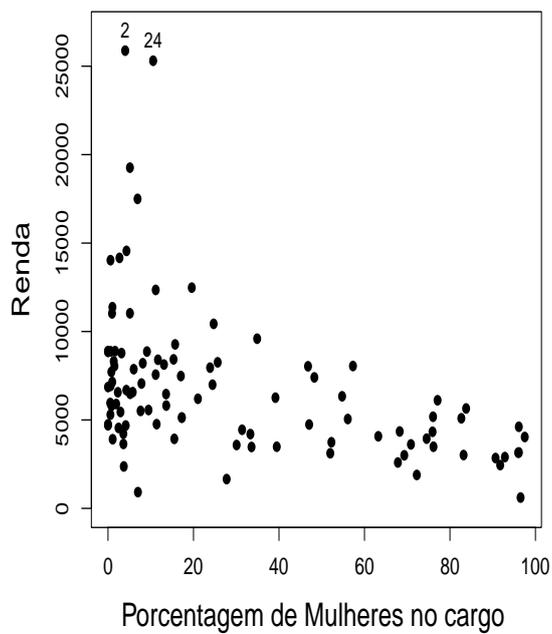
Figura 10 – Histograma da renda (a), gráfico de dispersão da renda e educação (b) e gráfico de dispersão da renda e porcentagem de mulheres no cargo (c) para o conjunto de dados Prestige.



(a)



(b)



(c)

Fonte: O autor (2025).

Para selecionar o modelo de RQ que apresentou o melhor ajuste ao conjunto de dados utilizou-se os seguintes critérios: critério de informação de Akaike (AIC) (SAKAMOTO; ISHIGURO; KITAGAWA, 1983), critério de informação de Akaike consistente (CAIC) (AKAIKE, 1973) e critério de informação Bayesiano (BIC) (SCHWARZ, 1978). Portanto, o modelo que apresenta os menores valores dessas estatísticas é selecionado entre os modelos ajustados, como o modelo mais adequado para descrever o conjunto de dados. Também considerou-se os resíduos quantílicos, resíduo generalizado de Cox-Snell e resíduo tipo Martingal nesta seleção. Além disso, apresentou-se os valores de  $R^2_\tau$  para analisar o quanto o modelo de RQ explica da variabilidade da variável resposta.

Na Tabela 9, por meio da estrutura de regressão quantílica baseada na Equação (3.1), foi comparado o desempenho das funções de ligações mais utilizadas na literatura, a saber função de ligação logarítmica (FLL) dada por  $\log(\eta_{\tau_i}) = \mathbf{x}_i^\top \boldsymbol{\psi}(\tau)$ , função de ligação raiz quadrada (FLR), expressa por  $\sqrt{\eta_{\tau_i}} = \mathbf{x}_i^\top \boldsymbol{\psi}(\tau)$ , e função de ligação identidade (FLI), escrita como  $\eta_{\tau_i} = \mathbf{x}_i^\top \boldsymbol{\psi}(\tau)$ . Os resultados mostrados na Tabela 9, indicam que o modelo de RQ com a função de ligação FLL deve ser utilizado para detalhar o primeiro quartil da renda. Portanto, considere a estrutura de regressão quantílica baseada na Equação (3.1), para o  $\tau$ -ésimo quantil dos modelos WER, LLER e LER expressa por:

$$\log(\eta_{\tau_i}) = \psi_0(\tau) + \psi_1(\tau)x_{i1} + \psi_2(\tau)x_{i2}, \quad i = 1, \dots, 102.$$

Além disso, o conjunto de dados foi ajustado para os modelos de RQ WER, LLER e LER, utilizando o software R (TEAM, 2020) com o método numérico BFGS (PRESS et al., 1992) por meio da função `optim()`. Os valores dos chutes iniciais para os parâmetros dos modelos de RQ utilizados nesta aplicação foram dados pelas estimativas do método de máxima verossimilhança da distribuição *baseline*.

Na Tabela 10, apresentamos as estimativas de máxima verossimilhança dos parâmetros (EMV) e estimativas dos erros-padrão (EP) correspondentes e seus respectivos p-valores para os modelos de RQ ajustados utilizando o conjunto de dados Prestige com  $\tau = 0,25$  fixo. Nota-se que, para os modelos de RQ WER e LLER as estimativas de todos os parâmetros apresentaram erros-padrão baixo, em comparação com os valores das estimativas. Logo, utilizando o teste de Wald todos os parâmetros dos modelos de RQ WER e LLER foram significativos ao nível de significância de 5%. Nota-se ainda que, o modelo de RQ LER apresentou erro-padrão alto para a estimativa do parâmetro  $\kappa$ , ou seja, estimativa não

precisa para o parâmetro  $\kappa$  e o p-valor indica que esse parâmetro não é significativo. Na Tabela 10, o parâmetro não significativo está destacado com um asterisco (\*). Este resultado indica que o conjunto de dados não segue uma distribuição LER. Mas, percebe-se que todos os parâmetros de regressão do modelo de RQ LER foram significativos ao nível de significância de 5% usando o teste de Wald. Portanto, como o parâmetro de forma  $\kappa$  do modelo LER é positivo e todos os parâmetros de regressão desse modelo foram significativos, então, continua-se a análise para os três modelos de RQ ajustados nesta aplicação.

Tabela 9: Estatísticas AIC, CAIC, BIC e função log-verossimilhança para os modelos RQ WER, LLER e LER com  $\tau = 0,25$  fixo e funções de ligações FLL, FLR e FLI para o conjunto de dados Prestige.

Modelos	FL	AIC	CAIC	BIC	Log-verossimilhança
WER	FLL	<b>1834, 259</b>	<b>1834, 672</b>	<b>1844, 759</b>	<b>-913, 129</b>
	FLR	1847, 326	1847, 738	1857, 826	-919, 663
	FLI	1872, 079	1872, 491	1882, 579	-932, 039
LLER	FLL	<b>1851, 335</b>	<b>1851, 747</b>	<b>1861, 835</b>	<b>-921, 667</b>
	FLR	2111, 547	2111, 960	2122, 047	-1051, 774
	FLI	1868, 956	1869, 369	1879, 456	-930, 4781
LER	FLL	<b>1929, 774</b>	<b>1930, 186</b>	<b>1940, 274</b>	<b>-960, 887</b>
	FLR	1931, 639	1932, 052	1942, 139	-961, 819
	FLI	2426, 12	2426, 533	2436, 62	-1209, 060

Fonte: O autor (2025).

Tabela 10: EMV dos parâmetros, estimativas dos erros-padrão (EP) correspondentes e seus respectivos p-valores para os modelos de RQ ajustados utilizando o conjunto de dados Prestige com  $\tau = 0,25$  fixo e função de ligação FLL.

Modelo	Estatística	$\hat{\psi}_0(\tau)$	$\hat{\psi}_1(\tau)$	$\hat{\psi}_2(\tau)$	$\hat{\kappa}$
WER	Estimativa	7, 3996	0, 1288	-0, 0110	2, 4409
	EP	0, 1264	0, 0109	0, 0009	0, 1699
	p-valor	< 0, 001	< 0, 001	< 0, 001	< 0, 001
LLER	Estimativa	7, 4927	0, 1206	-0, 0100	4, 4342
	EP	0, 1283	0, 0114	0, 0010	0, 3769
	p-valor	< 0, 001	< 0, 001	< 0, 001	< 0, 001
LER	Estimativa	6, 9306	0, 1256	-0, 0106	128, 2961*
	EP	0, 3015	0, 0269	0, 0023	131, 0815
	p-valor	< 0, 001	< 0, 001	< 0, 001	0, 327

Fonte: O autor (2025).

Na Tabela 11, nota-se que o modelo de RQ WER apresentou os menores valores das estatísticas AIC, CAIC, BIC, log-verossimilhança e maiores valores de  $R^2_\tau$  em comparação com os modelos de RQ LLER e LER. Logo, o modelo de RQ WER é o mais adequado para se ajustar ao conjunto de dados Prestige considerando  $\tau = 0,25$  fixo. Portanto, considere a estrutura de regressão quantílica baseada na Equação (3.1), para o  $\tau$ -ésimo quantil do modelo WER expressa por:

$$\log(\eta_{\tau_i}) = \psi_0(\tau) + \psi_1(\tau)x_{i1} + \psi_2(\tau)x_{i2}, \quad i = 1, \dots, 102.$$

em que,  $\psi_0(\tau) = 7,3996$  é o intercepto do modelo, ou seja, é o logaritmo do quantil 0,25 da renda obtido quando  $X_1$  e  $X_2$  são iguais a zero. Além disso,  $\psi_1(\tau) = 0,1288$ , isto indica que para cada aumento de 1 ano de educação, há um aumento no quantil 0,25 da renda em  $\exp(0,1288) = 1,137$ . Então, há um aumento percentual de aproximadamente 13,7% no quantil 0,25 da renda para cada aumento de 1 ano de educação considerando a variável independente  $X_2$  constante. Nota-se também que,  $\psi_2(\tau) = -0,0110$  isto mostra que, para cada aumento de 1% na porcentagem de mulheres em um cargo, diminui no quantil 0,25 da renda em  $\exp(-0,0110) = 0,989$ . Logo, a renda para o quantil 0,25 diminui aproximadamente 1,1% para cada aumento de 1% na porcentagem de mulheres em um cargo considerando a variável independente  $X_1$  constante.

Tabela 11: Estatísticas AIC, CAIC, BIC,  $R^2_\tau$  e função log-verossimilhança para os modelos de RQ WER, LLER e LER com  $\tau = 0,25$  fixo e função de ligação FLL utilizando o conjunto de dados Prestige.

Modelos	AIC	CAIC	BIC	$R^2_\tau$	Log-verossimilhança
WER	<b>1834,259</b>	<b>1834,672</b>	<b>1844,759</b>	<b>0,699</b>	<b>-913,129</b>
LLER	1851,335	1851,747	1861,835	0,642	-921,667
LER	1929,774	1930,186	1940,274	0,323	-960,887

Fonte: O autor (2025).

### 3.6.1 Análise de Resíduos para os Modelos de RQ WER, LLER e LER

As Figuras 11-13, ilustram os resíduos quantílicos, resíduo generalizado de Cox-Snell e resíduo tipo Martingal, com envelopes simulados e faixas de confiança para os modelos de RQ WER, LLER e LER. Nota-se, nessas figuras, que o resíduo tipo Martingal,  $R_{MT_\tau}$ , apresentou melhor ajuste. Percebe-se que, para o  $R_{MT_\tau}$  Figura 13-(a), o modelo de RQ WER

apresenta um melhor ajuste em comparação com os modelos de RQ LLER e LER. Além disso, na Figura 13, nota-se que, dentre os gráficos de eventos apresentados, o gráfico referente ao modelo de RQ WER não apresenta nenhum comportamento não usual, indicando que o modelo WER é adequado. Portanto, o modelo de RQ WER apresentou melhor ajuste para descrever o conjunto de dados Prestige.

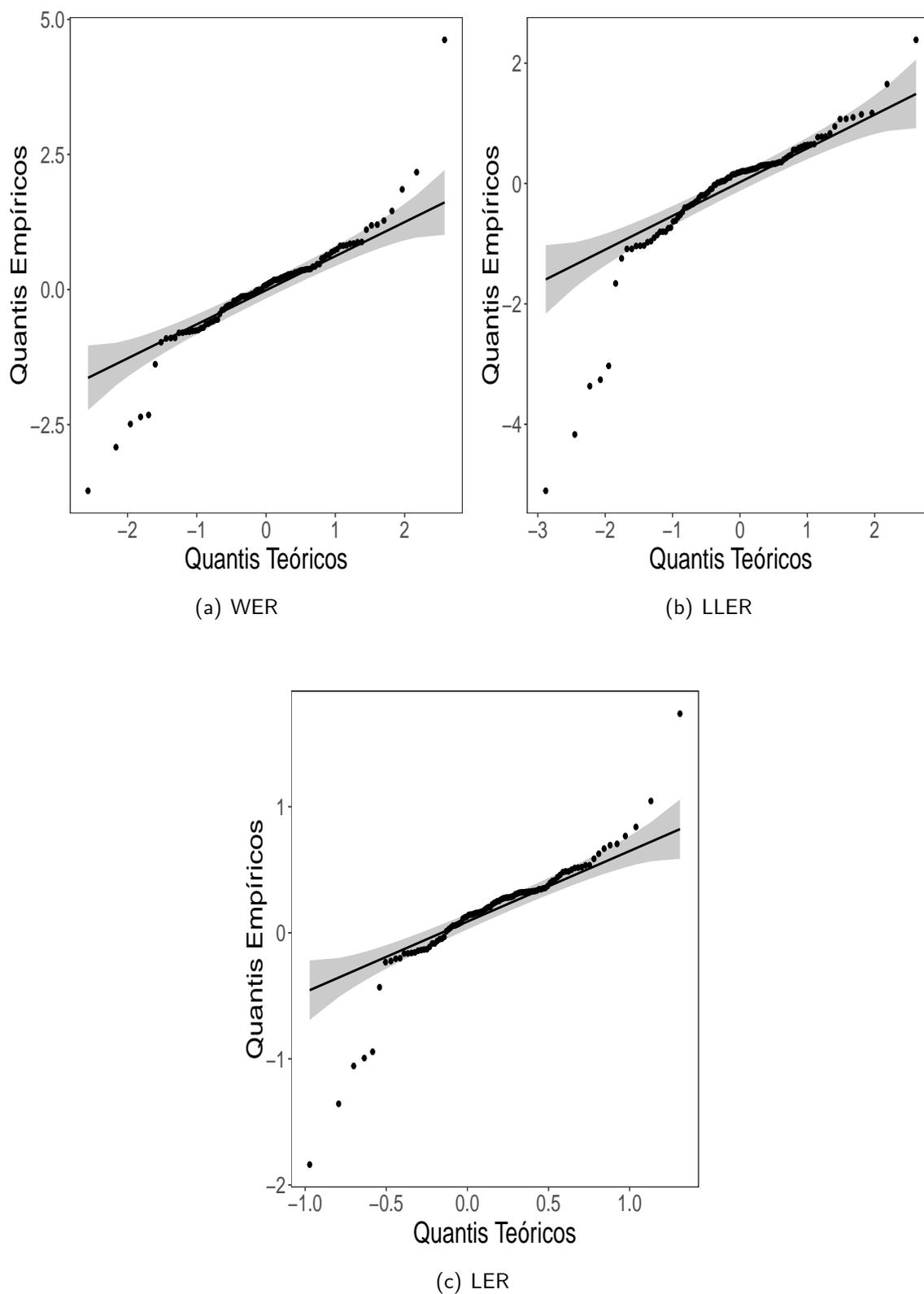
### 3.6.2 Análise da Distância de Cook Generalizada para os Modelos de RQ WER, LLER e LER

A Figura 14 apresenta os gráficos da distância de Cook generalizada para os modelos de RQ WER, LLER e LER com  $\tau = 0,25$  fixo. Esses gráficos são utilizados para identificar possíveis observações influentes.

Nota-se que, para o modelo de RQ WER a Figura 14-(a) revela que a observação #2 é classificada como possível observação influente. A observação #2 refere-se ao cargo de gerentes gerais, tem a maior renda observada no conjunto de dados (25879 USD) com uma porcentagem de 4,02% de mulheres neste cargo e educação média de 12,26 anos de estudo. Portanto, essa observação foi destacada pelo motivo de representar a maior renda observada no conjunto de dados, e por apresentar uma porcentagem pequena de mulheres neste cargo.

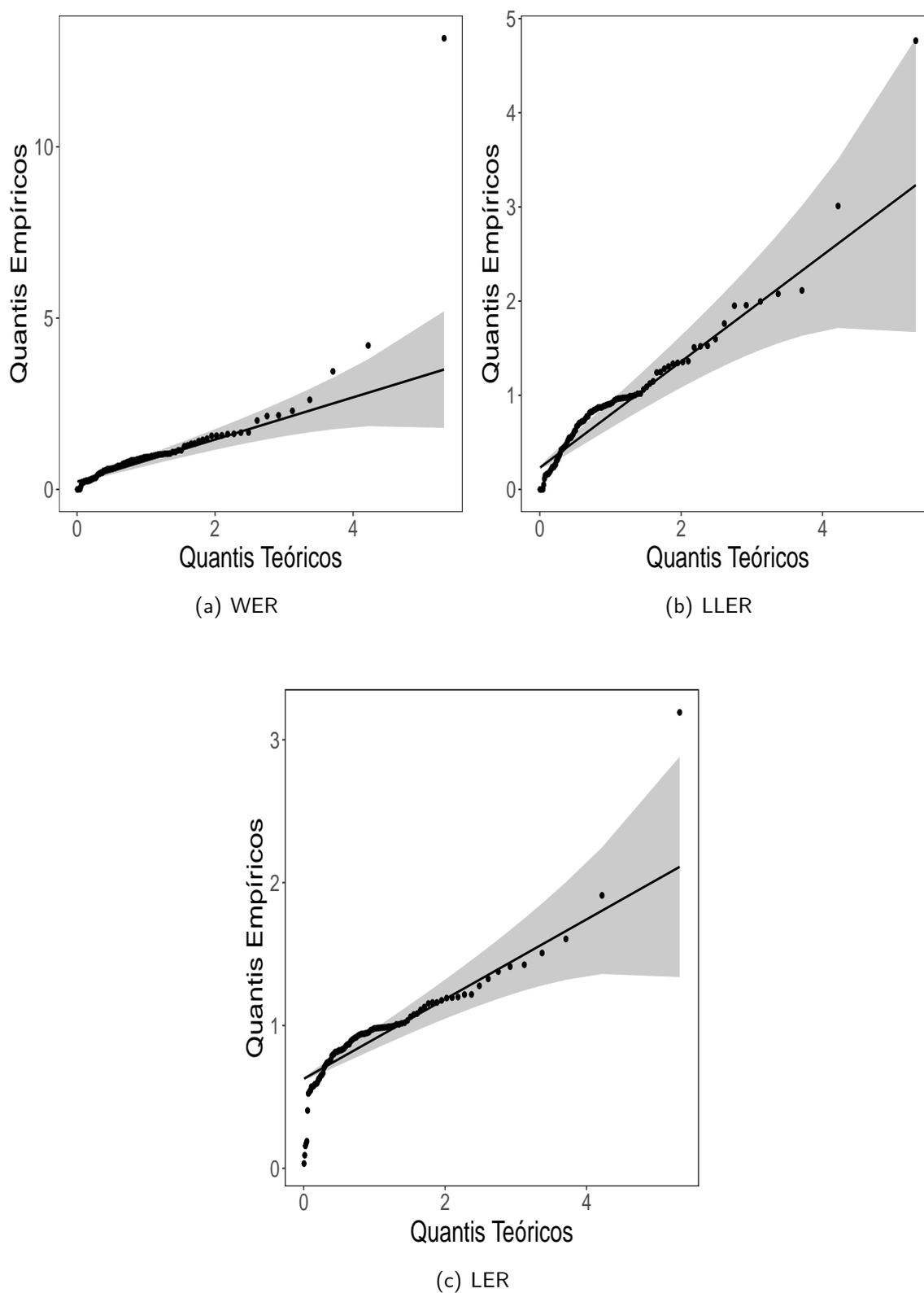
Percebe-se que, para o modelo de RQ LLER a Figura 14-(b) mostra que as observações #53 e #63 são identificadas como possíveis observações influentes. A observação #63 refere-se ao cargo de babás, tem a menor renda observada no conjunto de dados (611 USD) com uma porcentagem de 96,53% de mulheres neste cargo e educação média de 9,46 anos de estudo. Logo, a observação #63 foi destacada pelo motivo de representar a menor renda observada no conjunto de dados, e por apresentar uma porcentagem alta de mulheres neste cargo. A observação #53 representa o cargo de jornalheiros apresenta a segunda menor renda observada neste conjunto de dados (918 USD) com uma porcentagem de 7% de mulheres neste cargo e educação média de 9,62 anos de estudo. Dessa forma, essa observação foi destacada pelo motivo de representar a segunda menor renda observada no conjunto de dados, e por apresentar uma porcentagem pequena de mulheres neste cargo. Nota-se ainda que, para o modelo de RQ LER a Figura 14-(c) destaca a observação #2 como possível observação influente, assim como no modelo de RQ WER.

Figura 11 – Envelopes simulados com faixas de confiança para o resíduo quantílico e modelos de RQ WER, LLER e LER com  $\tau = 0,25$ .



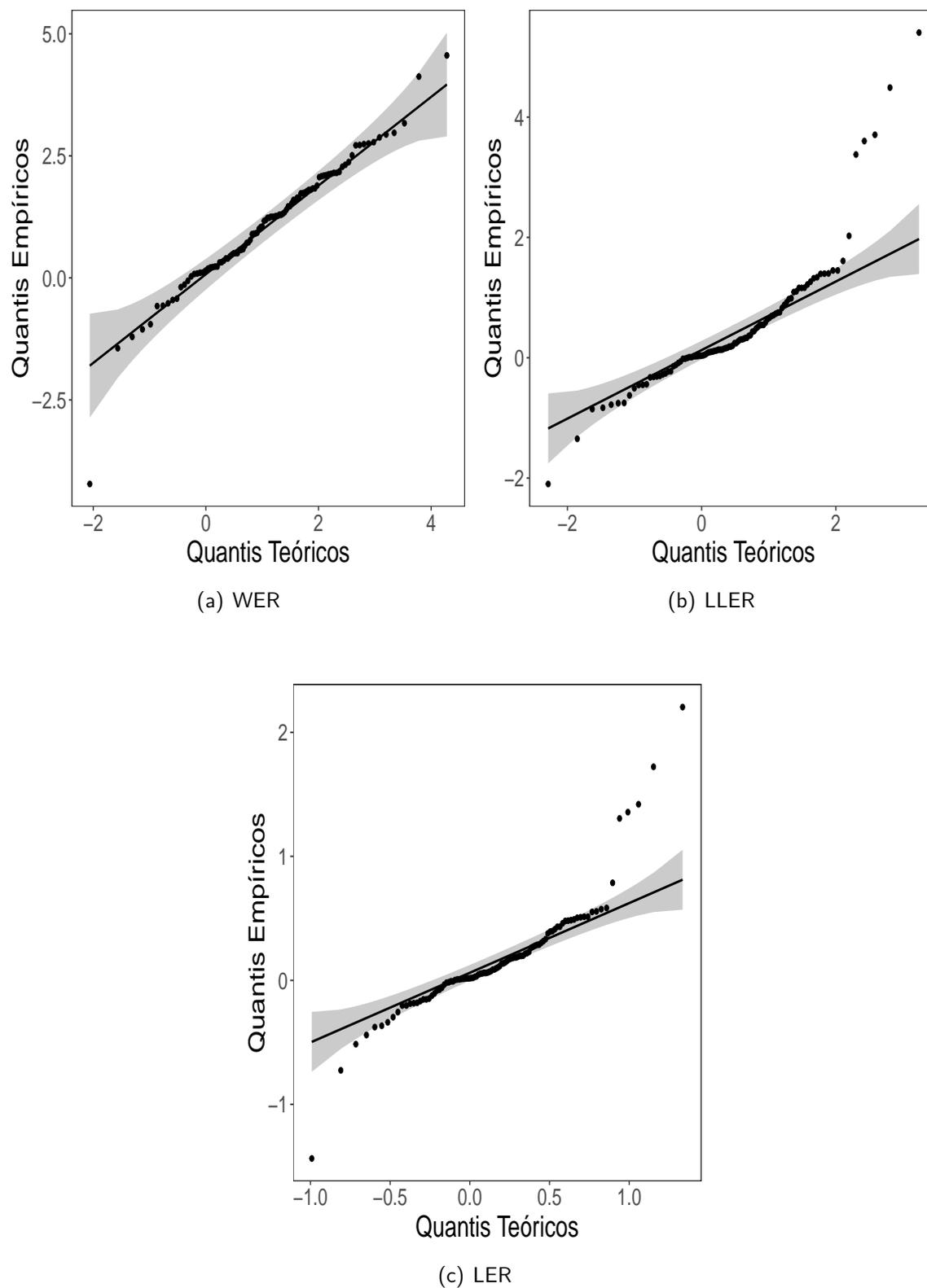
Fonte: O autor (2025).

Figura 12 – Envelopes simulados com faixas de confiança para o resíduo generalizado de Cox-Snell e modelos de RQ WER, LLER e LER com  $\tau = 0,25$ .

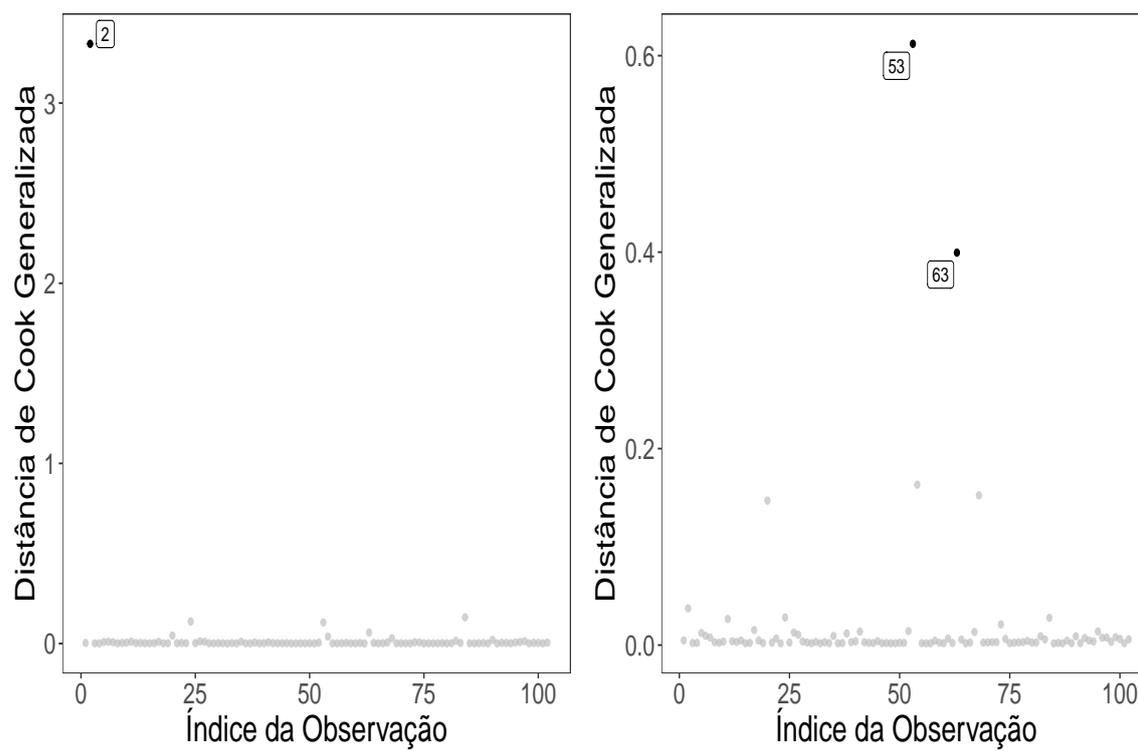


Fonte: O autor (2025).

Figura 13 – Envelopes simulados com faixas de confiança para o resíduo tipo Martingal e modelos de RQ WER, LLER e LER com  $\tau = 0,25$ .

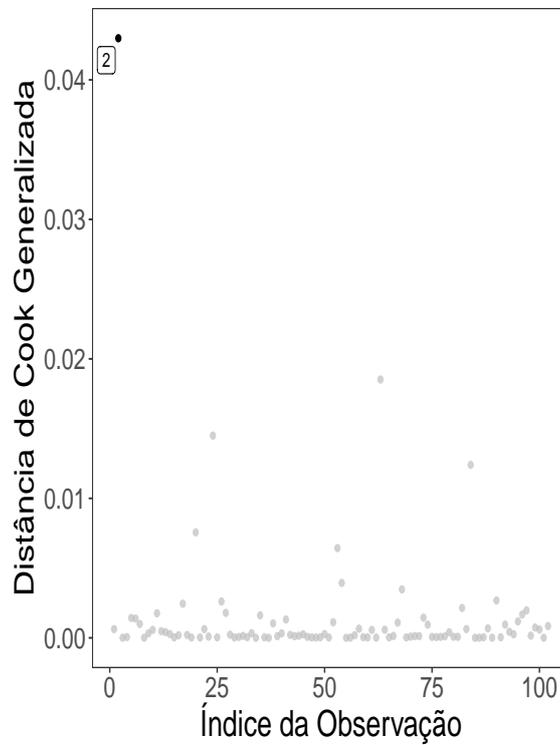


Fonte: O autor (2025).

Figura 14 – Distância de Cook generalizada para os modelos de RQ WER, LLER e LER com  $\tau = 0,25$ .

(a) WER

(b) LLER



(c) LER

Fonte: O autor (2025).

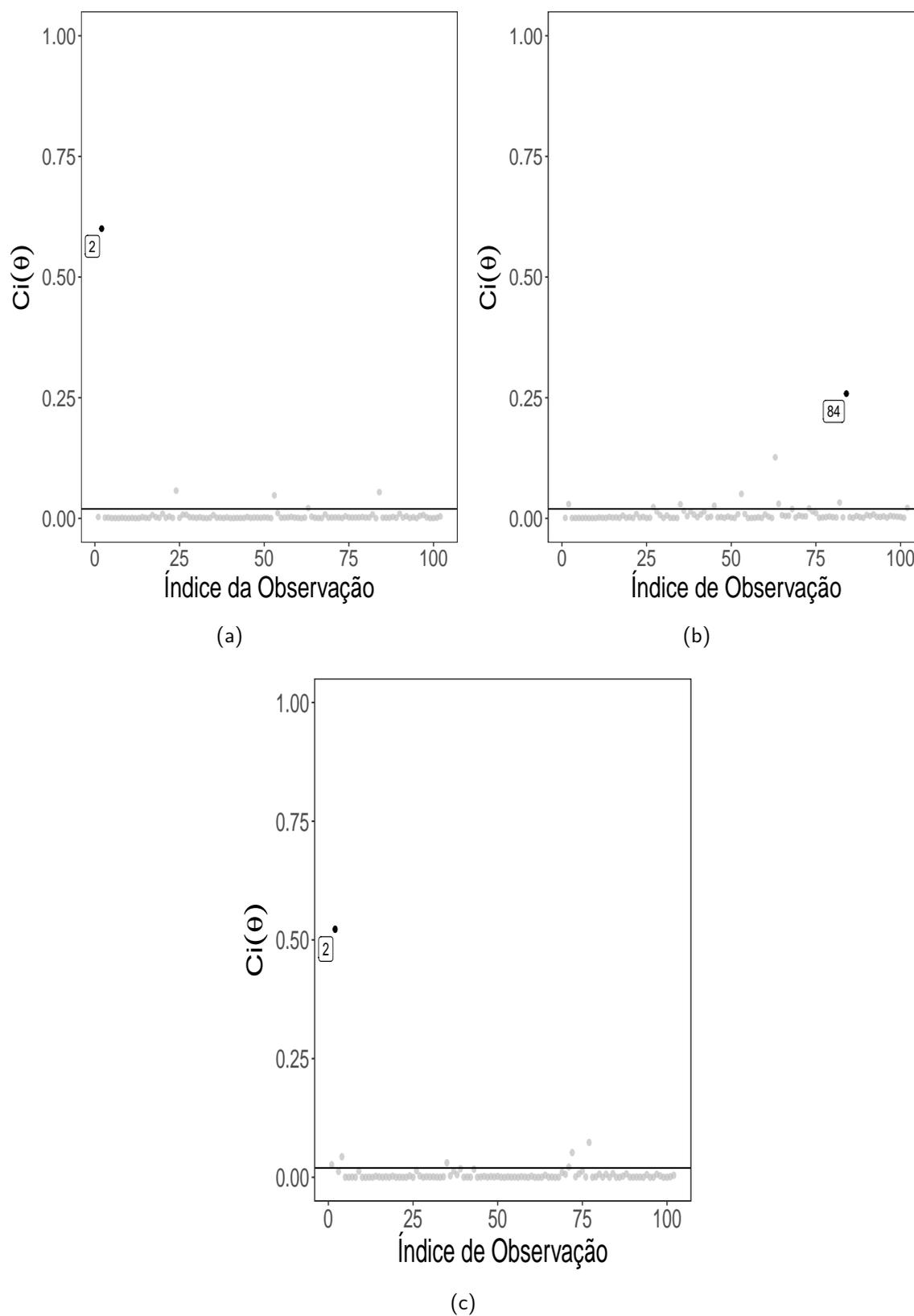
### 3.6.3 Análise de Influência Local para os Modelos de RQ WER, LLER e LER

A Figura 15 ilustra os gráficos de  $C_i(\theta)$  para o modelo de RQ WER utilizando os esquemas de perturbação de casos, perturbação na resposta e perturbação na covariável contínua respectivamente. Nota-se que, as observações #2 e #84 são destacadas como possíveis observações influentes. A observação #84 corresponde à observação com menor educação média (6,38 anos de estudo) e representa o cargo de operadores de máquina de costura com uma porcentagem de 90,67% de mulheres neste cargo e renda (2847 USD). Portanto, a observação #84 foi destacada pelo motivo de representar à observação com menor educação média no conjunto de dados, e por apresentar uma porcentagem alta de mulheres neste cargo.

A Figura 16, mostra os gráficos de  $C_i(\theta)$  para o modelo de RQ LLER aplicando os esquemas de perturbação de casos, perturbação na resposta e perturbação na covariável contínua respectivamente. Percebe-se que, as observações #53 e #63 são classificadas como possíveis observações influentes.

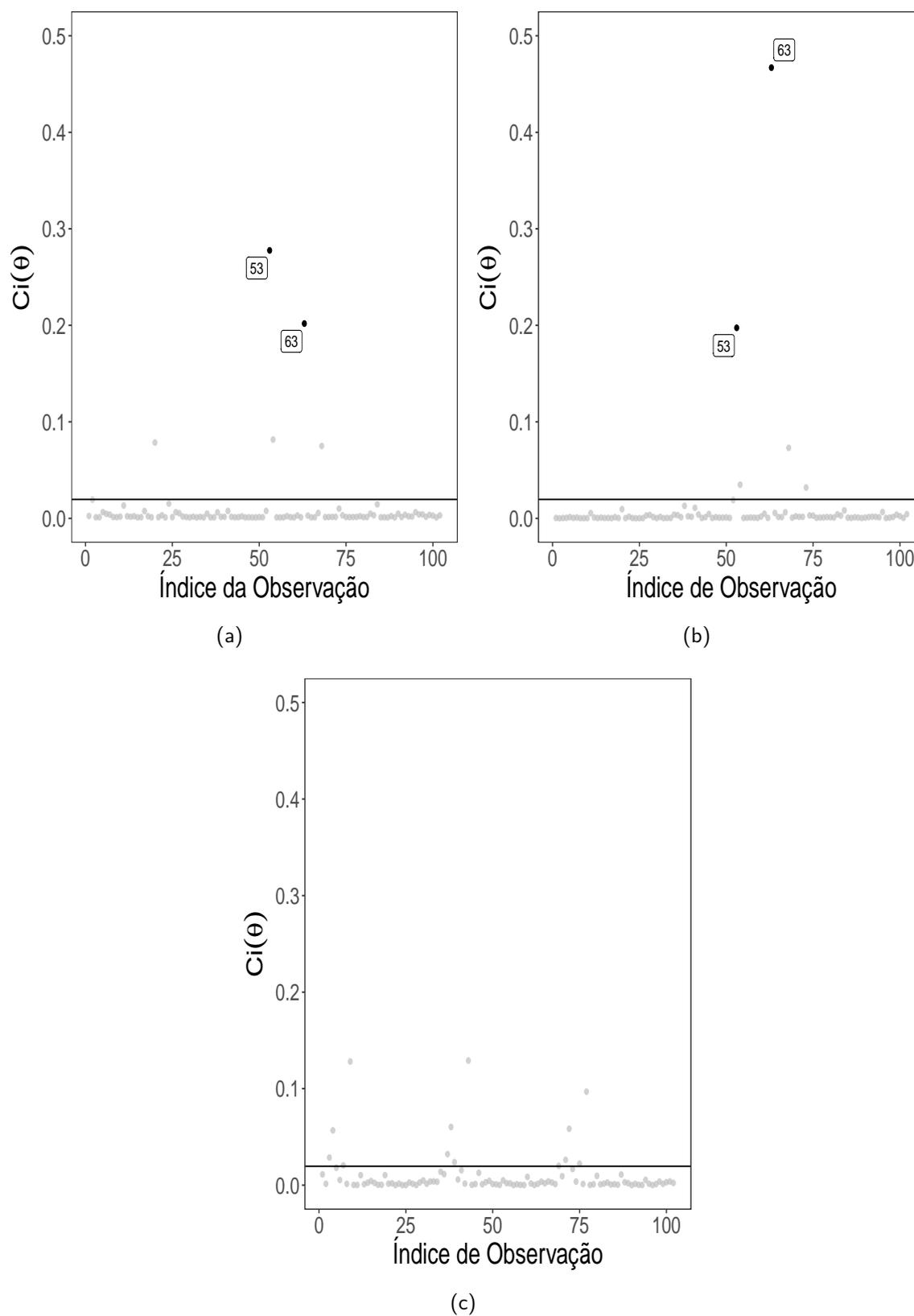
A Figura 17, apresenta os gráficos de  $C_i(\theta)$  para o modelo de RQ LER usando os esquemas de perturbação de casos, perturbação na resposta e perturbação na covariável contínua respectivamente. Observa-se, nesta figura que as observações #2, #63 e #77 são consideradas como possíveis observações influentes. A observação #77 representa o cargo de mecânicos com uma porcentagem de 4,28% de mulheres neste cargo, renda (6686 USD) e educação média (8,81 anos de estudo). Portanto, a observação #77 pode ter sido destacada devido à baixa porcentagem de mulheres neste cargo.

Figura 15 – Gráfico de  $C_i(\theta)$  versus o índice das observações para o modelo de RQ WER para  $\tau = 0, 25$  e utilizando o esquema de perturbação de casos (a), perturbação na resposta (b) e perturbação na covariável contínua (c).



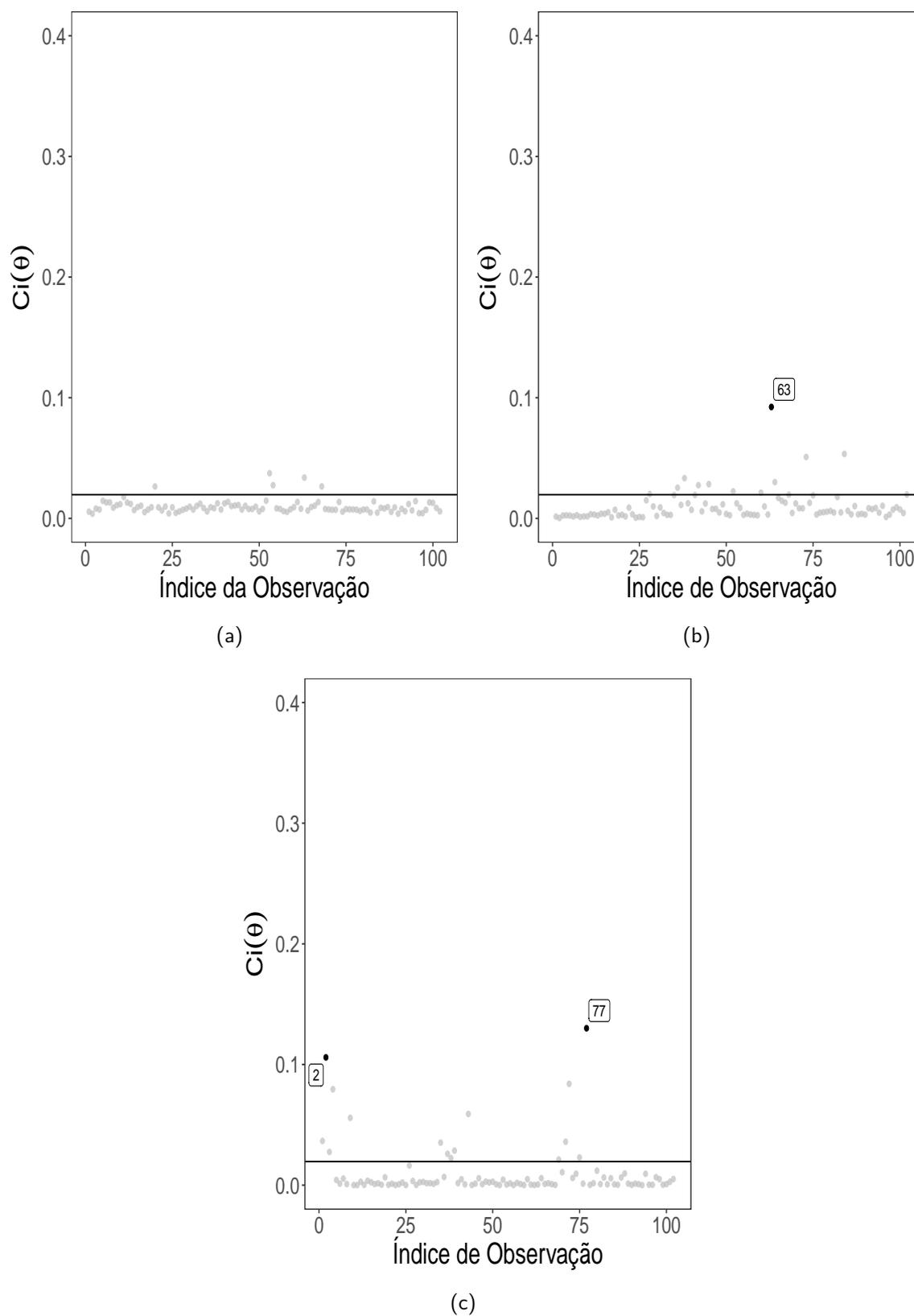
Fonte: O autor (2025).

Figura 16 – Gráfico de  $C_i(\theta)$  versus o índice das observações para o modelo de RQ LLER para  $\tau = 0, 25$  e utilizando o esquema de perturbação de casos (a), perturbação na resposta (b) e perturbação na covariável contínua (c).



Fonte: O autor (2025).

Figura 17 – Gráfico de  $C_i(\theta)$  versus o índice das observações para o modelo de RQ LER para  $\tau = 0, 25$  e utilizando o esquema de perturbação de casos (a), perturbação na resposta (b) e perturbação na covariável contínua (c).



Fonte: O autor (2025).

### 3.6.4 Análise Confirmatória

Nesta seção, foi realizado um estudo para investigar o impacto das observações potencialmente influentes nas estimativas dos parâmetros dos modelos propostos. Dessa forma, foram removidos os casos potencialmente influentes e calculadas as estimativas dos parâmetros e seus correspondentes erros-padrão (se's) para os modelos de RQ WER, LLER e LER. A Tabela 12, mostra os valores de RCs (em %) com seus correspondentes se's e respectivos p-valores para o modelo de RQ WER com  $\tau = 0,25$ . Nota-se que a maior mudança relativa percentual em valor absoluto nas estimativas e erros-padrão foi de RCs igual a 23,582 para o parâmetro  $\kappa$  quando foram excluídos os casos  $\{\#2, \#84\}$ . De forma geral essa análise não revela nenhuma mudança inferencial, ou seja, não impactar na significância dos parâmetros. Logo, a análise de influência local identificou casos potencialmente influentes para o modelo de RQ WER, porém esses casos não impactam na inferência do modelo. Sendo assim, o modelo de RQ WER é robusto aos casos atípicos identificados e apresentou um bom ajuste ao conjunto de dados Prestige.

A Tabela 13 apresenta os valores de RCs (em %) com seus correspondentes se's e respectivos p-valores para o modelo de RQ LLER com  $\tau = 0,25$ . Percebe-se que, o maior valor de RCs foi 24,858 para o parâmetro  $\kappa$  quando foram removidos os casos  $\{\#53, \#63\}$ . Nota-se que, esse estudo não apresenta nenhuma mudança inferencial. A Tabela 14 apresenta os valores de RCs (em %) com seus correspondentes erros-padrão e respectivos p-valores para o modelo de RQ LER com  $\tau = 0,25$ . Observa-se que, de modo geral, não existe mudança inferencial.

Tabela 12: Medidas RCs (em %) com seus correspondentes se's e respectivos p-valores para o modelo de RQ WER com  $\tau = 0,25$  e utilizado o conjunto de dados Prestige.

Casos removidos	Componente	Parâmetros			
		$\hat{\psi}_0(\tau)$	$\hat{\psi}_1(\tau)$	$\hat{\psi}_2(\tau)$	$\hat{\kappa}$
{#2}	$RC_{\Theta_{(i)}}$	0,950	4,149	6,355	21,029
	$RC_{se(\hat{\Theta}_{j(i)})}$	19,904	20,053	14,420	33,741
	p-valor	< 0,0001	< 0,0001	< 0,0001	< 0,0001
{#84}	$RC_{\Theta_{(i)}}$	0,575	3,911	4,561	1,442
	$RC_{se(\hat{\Theta}_{j(i)})}$	0,243	3,014	4,592	2,683
	p-valor	< 0,0001	< 0,0001	< 0,0001	< 0,0001
{#2, #84}	$RC_{\Theta_{(i)}}$	0,508	0,867	2,165	23,582
	$RC_{se(\hat{\Theta}_{j(i)})}$	20,285	18,683	13,424	38,734
	p-valor	< 0,0001	< 0,0001	< 0,0001	< 0,0001

Fonte: O autor (2025).

Tabela 13: Medidas RCs (em %) com seus correspondentes se's e respectivos p-valores para o modelo de RQ LLER com  $\tau = 0,25$  e utilizado o conjunto de dados Prestige.

Casos removidos	Componente	Parâmetros			
		$\hat{\psi}_0(\tau)$	$\hat{\psi}_1(\tau)$	$\hat{\psi}_2(\tau)$	$\kappa$
{#53}	$RC_{\Theta_{(i)}}$	0,633	1,226	2,006	13,098
	$RC_{se(\hat{\Theta}_{j(i)})}$	10,767	10,3895	8,771	11,777
	p-valor	< 0,0001	< 0,0001	< 0,0001	< 0,0001
{#63}	$RC_{\Theta_{(i)}}$	0,423	2,067	5,532	9,341
	$RC_{se(\hat{\Theta}_{j(i)})}$	7,929	7,670	7,508	8,610
	p-valor	< 0,0001	< 0,0001	< 0,0001	< 0,0001
{#53, #63}	$RC_{\Theta_{(i)}}$	0,993	3,161	3,850	24,858
	$RC_{se(\hat{\Theta}_{j(i)})}$	18,253	17,811	17,083	22,319
	p-valor	< 0,0001	< 0,0001	< 0,0001	< 0,0001

Fonte: O autor (2025).

Tabela 14: Medidas RCs (em %) com seus correspondentes se's e respectivos p-valores para o modelo de RQ LER com  $\tau = 0,25$  e utilizado o conjunto de dados Prestige.

Casos removidos	Componente	Parâmetros			
		$\hat{\psi}_0(\tau)$	$\hat{\psi}_1(\tau)$	$\hat{\psi}_2(\tau)$	$\kappa$
{#2}	$RC_{\Theta_{(i)}}$	0,137	2,706	4,013	29,543
	$RC_{se}(\hat{\Theta}_{j(i)})$	0,761	0,721	1,445	44,700
	p-valor	< 0,0001	< 0,0001	< 0,0001	0,3809
{#63}	$RC_{\Theta_{(i)}}$	0,157	1,410	5,534	0,076
	$RC_{se}(\hat{\Theta}_{j(i)})$	0,246	0,067	2,668	0,073
	p-valor	< 0,0001	< 0,0001	< 0,0001	0,3269
{#77}	$RC_{\Theta_{(i)}}$	0,0507	0,150	0,414	0,420
	$RC_{se}(\hat{\Theta}_{j(i)})$	0,542	0,057	1,288	0,145
	p-valor	< 0,0001	< 0,0001	< 0,0001	0,3290
{#2, #63}	$RC_{\Theta_{(i)}}$	0,295	4,135	9,392	26,695
	$RC_{se}(\hat{\Theta}_{j(i)})$	0,518	0,505	3,134	39,919
	p-valor	< 0,0001	< 0,0001	< 0,0001	0,3754
{#2, #77}	$RC_{\Theta_{(i)}}$	0,074	2,512	4,247	30,706
	$RC_{se}(\hat{\Theta}_{j(i)})$	0,245	0,539	1,736	47,677
	p-valor	< 0,0001	< 0,0001	< 0,0001	0,3863
{#63, #77}	$RC_{\Theta_{(i)}}$	0,104	1,243	5,727	0,388
	$RC_{se}(\hat{\Theta}_{j(i)})$	0,788	0,276	2,955	0,048
	p-valor	< 0,0001	< 0,0001	< 0,0001	0,3293
{#2, #63, #77}	$RC_{\Theta_{(i)}}$	0,230	3,934	9,632	27,485
	$RC_{se}(\hat{\Theta}_{j(i)})$	0,003	0,319	3,424	41,830
	p-valor	< 0,0001	< 0,0001	< 0,0001	0,3789

Fonte: O autor (2025).

### 3.7 OBSERVAÇÕES FINAIS

Neste capítulo, foram propostos modelos de regressão quantílica para a família de distribuições  $\alpha_\tau$ -G reparametrizada. Dessa forma, ao escolher o modelo *baseline* da família de distribuições  $\alpha_\tau$ -G reparametrizada, pode ser gerado o correspondente modelo de regressão quantílica para o submodelo da família de distribuições  $\alpha_\tau$ -G reparametrizada. Para exemplificar, utilizando essa metodologia foi apresentado o novo modelo de regressão quantílica WER e seus principais aspectos estatísticos. Quanto à simulação, foi realizado um estudo de simulação de Monte Carlo para investigar numericamente o desempenho e as propriedades dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de regressão quantílica WER. Esta simulação mostrou que, o estimador  $\hat{\kappa}$  apresentou melhor desempenho para valores menores de  $\tau$ . Notou-se que, à medida que o valor do parâmetro  $\kappa$  aumenta, as estimativas do parâmetro  $\kappa$  pioram. Percebeu-se ainda que há

fortes indícios de que os estimadores dos parâmetros do modelo de RQ WER são consistentes e assintoticamente distribuídos como uma variável aleatória normal. Além disso, aplicou-se os modelos de regressão quantílica propostos a um conjunto de dados reais. Os resultados da aplicação mostram que, o modelo de regressão quantílica WER apresentou melhor ajuste.

## 4 CONCLUSÕES E TRABALHOS FUTUROS

Neste estudo, foi proposto uma nova família de distribuições indexada por um parâmetro representando o  $\tau$ -ésimo quantil, denominada família de distribuições G-exponencializada ( $\alpha_\tau$ -G) reparametrizada e apresentamos alguns de seus submodelos. Por exemplo, escolhendo o modelo *baseline* podemos obter os submodelos da família de distribuições  $\alpha_\tau$ -G reparametrizada. Essa família de distribuições reparametrizada é uma alternativa a família de distribuições G-exponencializada, quando o interesse for modelar os quantis da distribuição. Sendo assim, recomendamos a família de distribuições  $\alpha_\tau$ -G reparametrizada quando o interesse consiste em modelar os quantis da distribuição. Também foi realizado um estudo de simulação de Monte Carlo para examinar as propriedades dos estimadores de máxima verossimilhança e intervalos de confiança assintóticos para a distribuição WER. Os resultados desse estudo de simulação indicaram que o estimador  $\hat{\kappa}$  tem um melhor desempenho para quantis inferiores. Além disso, os intervalos de confiança assintóticos para o parâmetro  $\kappa$  apresentaram melhores taxas de cobertura. Também se observa que há fortes indícios de que os estimadores dos parâmetros do modelo WER são consistentes e distribuídos assintoticamente como uma variável aleatória normal. Quanto à aplicação, os resultados mostraram que, para  $\tau$ : 0,25 e 0,5, o modelo WER foi selecionado como o mais adequado para se ajustar ao conjunto de dados sobre o estresse de ruptura das fibras de carbono (em Gba). Por outro lado, para  $\tau = 0,75$ , o modelo LLER foi selecionado como o mais adequado para se ajustar a esse conjunto de dados.

Também, foi proposto de modo geral modelos de regressão quantílica para modelos na família de distribuições  $\alpha_\tau$ -G reparametrizada. Então, especificando o modelo *baseline* da família de distribuições  $\alpha_\tau$ -G reparametrizada podemos obter o correspondente modelo de regressão quantílica para o submodelo da família de distribuições  $\alpha_\tau$ -G reparametrizada. Para exemplificar, utilizando essa metodologia foi apresentado o novo modelo de regressão quantílica WER. Em relação à simulação, foi apresentado um estudo de simulação de Monte Carlo para investigar numericamente o desempenho e as propriedades dos estimadores de máxima verossimilhança dos parâmetros que indexam o modelo de regressão quantílica WER. Esta simulação revelou que, o estimador  $\hat{\kappa}$  apresentou melhor desempenho para valores menores de  $\tau$ . Além disso, quando o valor do parâmetro  $\kappa$  cresce, as estimativas

---

do parâmetro  $\kappa$  pioram. Notou-se ainda que há fortes indícios de que os estimadores dos parâmetros do modelo de RQ WER são consistentes e assintoticamente distribuídos como uma variável aleatória normal. Por fim, aplicou-se os novos modelos de regressão quantílica a um conjunto de dados reais. Os resultados desta aplicação indicam que, o modelo de regressão quantílica WER apresentou melhor ajuste.

#### 4.1 TRABALHOS FUTUROS

Como possíveis trabalhos futuros, é de interesse:

- Inserir censura na família de distribuições  $\alpha_\tau$ -G reparametrizada;
- Propor modelos de RQ baseados na família de distribuições  $\alpha_\tau$ -G reparametrizada considerando a presença de censura;
- Implementar um pacote no R para produzir várias ferramentas para ajuste dos modelos pertencentes à família de distribuições  $\alpha_\tau$ -G reparametrizada;
- Implementar um pacote no R para fornece várias ferramentas para ajuste e diagnóstico dos modelos de RQ baseados na família de distribuições  $\alpha_\tau$ -G reparametrizada;
- Apresentar modelos de regressão quantílica parcialmente linear para modelos pertencentes à família de distribuições  $\alpha_\tau$ -G reparametrizada;
- Propor modelos de séries temporais quantílica baseados na família de distribuições  $\alpha_\tau$ -G reparametrizada;
- Propor modelos de regressão quantílica espacial fundamentados na família de distribuições  $\alpha_\tau$ -G reparametrizada.

## REFERÊNCIAS

- ABDUL-MONIEM, I.; ABDEL-HAMEED, H. On exponentiated lomax distribution. *International Journal of Mathematical Education*, v. 33, p. 1–7, 04 2012.
- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: AKADEMIKI KIADO. *2nd International Symposium on Information Theory, 1973*. [S.l.], 1973. p. 267–281.
- AL-HUSSAINI, E. K.; HUSSEIN, M. Estimation using censored data from exponentiated burr type xii population. *American Open Journal of Statistics*, v. 1, p. 33–45, 2011.
- ANDERSON, T. W.; DARLING, D. A. A test of goodness of fit. *Journal of the American Statistical Association*, Taylor Francis, v. 49, n. 268, p. 765–769, 1954.
- COOK, R. D. Detection of influential observation in linear regression. *Technometrics*, vol. 19, no. 1, 19, pp. 15–18., 1977.
- COOK, R. D. Assessment of local influence. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(2) 133–155., 1986.
- CORDEIRO, G. M.; LEMONTE, A. J.; ORTEGA, E. M. An extended fatigue life distribution. *Statistics*, Taylor & Francis, v. 47, n. 3, p. 626–653, 2013.
- CORDEIRO, G. M.; ORTEGA, E. M. M.; SILVA, G. O. Modelos de regressão de sobrevivência estendidos. In: *Escola de Modelos de Regressão*. [S.l.]: UFC/DEMA/ABE, 2011.
- COX, D.; HINKLEY, D. *Theoretical Statistics (1st ed.)*. [S.l.]: Chapman and Hall, UK, 1974.
- COX, D. R.; SNELL, E. J. A general definition of residuals. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 30, n. 2, p. 248–265, 1968.
- CRAMÉR, H. On the composition of elementary errors. *Scandinavian Actuarial Journal*, Taylor Francis, v. 1928, n. 1, p. 13–74, 1928.
- DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, v. 5, n. 3, p. 236–244, 1996.
- EL-GOHARY, A.; AHMAD, A.; ADEL, N. A.-O. The generalized gompertz distribution. *Applied Mathematical Modelling*, v. 37, n. 1, p. 13–24, 2013.
- FOX, J. *Applied regression analysis and generalized linear models*. [S.l.]: Sage publications, 2015.
- GALLARDO, D. I.; BOURGUIGNON, M.; GALARZA, C. E.; GÓMEZ, H. W. A parametric quantile regression model for asymmetric response variables on the real line. *Symmetry*, v. 12, n. 12, 2020. ISSN 2073-8994. Disponível em: <<https://www.mdpi.com/2073-8994/12/12/1938>>.

- 
- GALLARDO, D. I.; GÓMEZ, Y. M.; SEGOVIA, F. A. Exponentiated power maxwell distribution with quantile regression and applications. *SORT-Statistics and Operations Research Transactions*, v. 45, n. 2, p. 181–200, 2021.
- GALLARDO, D. I.; SANTOS-NETO, M. Quantile regression for positive data using a general class of distributions. *arXiv preprint arXiv:2109.09281*, 2021.
- GRATZER, G. *Math into LaTeX : an introduction to LaTeX and AMS-LaTeX*. Boston: Birkhä user, 1996. ISBN 978-0817638054.
- GUPTA, R. D.; KUNDU, D. Exponentiated exponential family: An alternative to gamma and weibull distributions. *Biometrical Journal*, v. 43, n. 1, p. 117–130, 2001.
- KOENKER, R. *Quantile Regression*. [S.l.]: Cambridge University Press, 2005. (Econometric Society Monographs).
- KOENKER, R.; JR, G. B. "Regression quantiles", *econometrica: Journal of the economic society*, vol. 46, no. 1. 1978.
- KOLMOGOROV, A. N. Sulla determinazione empirica di una legge didistribuzione. *Giorn Dell'inst Ital Degli Att*, v. 4, p. 89–91, 1933.
- KUNDU, D.; RAQAB, M. Z. Generalized rayleigh distribution: different methods of estimations. *Computational Statistics and Data Analysis*, v. 49, n. 1, p. 187–200, 2005. ISSN 0167-9473.
- LEHMANN, E. L. The Power of Rank Tests. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 24, n. 1, p. 23 – 43, 1953. Disponível em: <<https://doi.org/10.1214/aoms/1177729080>>.
- LEHMANN, E. L.; CASELLA, G. *Theory of Point Estimation*. Second. New York, NY, USA: Springer-Verlag, 1998.
- LESAFFRE, E.; VERBEKE, G. Local influence in linear mixed models. *Biometrics*, JSTOR, p. 570–582, 1998.
- MADDALA, G. S. Limited-dependent and qualitative variables in econometrics. In: \_\_\_\_\_. [S.l.]: Cambridge University Press, 1983.
- MARINHO, P. R. D.; SILVA, R. B.; BOURGUIGNON, M.; CORDEIRO, G. M.; NADARAJAH, S. Adequacy model: An r package for probability distributions and general purpose optimization. *PLOS ONE*, Public Library of Science, v. 14, n. 8, p. 1–30, 08 2019.
- MAZUCHELI, J.; LEIVA, V.; ALVES, B.; MENEZES, A. F. B. A new quantile regression for modeling bounded data under a unit birnbaum–saunders distribution with applications in medicine and politics. *Symmetry*, v. 13, n. 4, 2021. ISSN 2073-8994. Disponível em: <<https://www.mdpi.com/2073-8994/13/4/682>>.
- MCCULLAGH, P.; NELDER, J. *Generalized Linear Models, Second Edition*. Chapman & Hall, 1989. (Chapman and Hall/CRC Monographs on Statistics and Applied Probability Series). ISBN 9780412317606. Disponível em: <<http://books.google.com/books?id=h9kFH2\FfBkC>>.

- MISES, R. V. Statistik und wahrheit. *Julius Springer*, Springer, v. 20, 1928.
- MONTGOMERY, D.; PECK, E.; VINING, G. *Introduction to Linear Regression Analysis*. Wiley, 2012. (Wiley Series in Probability and Statistics). ISBN 9780470542811. Disponível em: <<https://books.google.fr/books?id=0yR4KUL4VDkC>>.
- MUDHOLKAR, G. S.; SRIVASTAVA, D. K. Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, v. 42, n. 2, p. 299–302, 1993.
- NADARAJAH, S. The exponentiated gumbel distribution with climate application. *Environmetrics*, v. 17, n. 1, p. 13–23, 2006.
- NADARAJAH, S.; KOTZ, S. The exponentiated fréchet distribution. *Interstat Electronic Journal*, Citeseer, v. 14, p. 1–7, 2003.
- NICHOLS, M. D.; PADGETT, W. J. A bootstrap control chart for weibull percentiles. *Quality and Reliability Engineering International*, v. 22, n. 2, p. 141–151, 2006.
- OWEN, W. A new three-parameter extension to the birnbaum-saunders distribution. *IEEE Transactions on Reliability*, v. 55, n. 3, p. 475–479, 2006.
- PRESS, W. H.; FLANNERY, B. P.; TEUKOLSKY, S. A.; VETTERLING, W. T. Numeric recipes in c: the art of scientific computing. *Camb. Univ. Press Camb*, 1992.
- RODRIGUES, A.; BORGES, P.; SANTOS, B. A defective cure rate quantile regression model for male breast cancer data. *arXiv preprint arXiv:2105.03699*, 2021.
- ROSAIAH, K.; KANTAM, R.; KUMAR, S. Reliability test plans for exponentiated log-logistic distribution. Walter de Gruyter GmbH & Co. KG, 2006.
- SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. *Akaike Information Criterion Statistics*. D. Reidel Publishing Company, 1983. (Mathematics and its applications). Disponível em: <<https://books.google.com.br/books?id=w6wQnQEACAAJ>>.
- SÁNCHEZ, L.; LEIVA, V.; GALEA, M.; SAULO, H. Birnbaum-saunders quantile regression models with application to spatial data. *Mathematics*, v. 8, n. 6, 2020.
- SÁNCHEZ, L.; LEIVA, V.; GALEA, M.; SAULO, H. Birnbaum-saunders quantile regression and its diagnostics with application to economic data. *Applied Stochastic Models in Business and Industry*, v. 37, n. 1, p. 53–73, 2021.
- SÁNCHEZ, L.; LEIVA, V.; SAULO, H.; MARCHANT, C.; SARABIA, J. M. A new quantile regression model and its diagnostic analytics for a weibull distributed response with applications. *Mathematics*, v. 9, n. 21, 2021. ISSN 2227-7390.
- SANTOS, B. Vicente dos; SANTOS-NETO, M.; CYSNEIROS, F. J. A. Reparametrized exponentiated-G distribution family: Properties and applications: Accepted - september 2024. *REVSTAT-Statistical Journal*, Nov. 2024. Disponível em: <<https://revstat.ine.pt/index.php/REVSTAT/article/view/791>>.
- SANTOS-NETO, M.; GALLARDO, D. I.; COSTA, E.; MARCHANT, C.; RENAN, I. A new parametric quantile regression model based on an owen distribution. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 95, n. 1, p. 156–185, 2025. Disponível em: <<https://doi.org/10.1080/00949655.2024.2422921>>.

---

SAULO, H.; VILA, R.; BORGES, G. V.; BOURGUIGNON, M. *Parametric quantile regression for income data*. 2022. Disponível em: <<https://arxiv.org/abs/2207.06558>>.

SCHWARZ, G. Estimating the dimension of a model. *The annals of statistics*, JSTOR, p. 461–464, 1978.

TEAM, R. D. C. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020.

THERNEAU, T. M.; GRAMBSCH, P. M.; FLEMING, T. R. Martingale-based residuals for survival models. *Biometrika*, Oxford University Press, v. 77, n. 1, p. 147–160, 1990.

WEIBULL, W. *The Phenomenon of Rupture in Solids*. Generalstabena Litografiska Anstalts Forlag, 1939. (Handlingar / Ingeniörsvetenskapsakademien). Disponível em: <<https://books.google.com.br/books?id=G81cvwEACAAJ>>.

## APÊNDICE A – ESQUEMAS DE PERTURBAÇÃO PARA O MODELO DE REGRESSÃO QUANTÍLICA WER

**Perturbação de casos:** Nesse esquema de perturbação, a função de log-verossimilhança perturbada para o modelo de regressão quantílica WER é expressa por:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta}),$$

em que

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) = & \log[\log(2)] + \log(\alpha_\tau) + \log(\kappa) + (\kappa - 1) \log(y_i) \\ & - \frac{1}{\eta_{\tau_i}^\kappa} y_i^\kappa \log(2) - \kappa \log(\eta_{\tau_i}) + (\alpha_\tau - 1) \log\left(1 - 2^{-\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}}\right), \end{aligned}$$

com  $0 \leq \omega_i \leq 1$  e  $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$ . A matriz de perturbação pode ser escrita como:

$$\Delta = \left( \begin{array}{c} \Delta_{\psi(\tau)} \\ \Delta_{\kappa} \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

em que os elementos de  $\Delta$  podem ser escritos como,  $\Delta_{\psi(\tau)} = \mathbf{x}_i a_i z_i$  e  $\Delta_{\kappa} = b_i$  para  $i = 1, \dots, n$ . Note que,  $a_i$ ,  $z_i$  e  $b_i$  para o modelo de RQ WER foram apresentados na Seção 3.3.1. Na forma matricial  $\Delta_{\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{z}$  e  $\Delta_{\kappa} = (b_1, \dots, b_n)^\top$ , sendo  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$  e  $\mathbf{z} = (z_1, \dots, z_n)^\top$ .

**Perturbação na resposta:** Utilizando o esquema de perturbação na resposta  $\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega})$  pode ser escrito como:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}),$$

em que

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) = & \log[\log(2)] + \log(\alpha_\tau) + \log(\kappa) + (\kappa - 1) \log[y_i(\omega_i)] \\ & - \frac{1}{\eta_{\tau_i}^\kappa} [y_i(\omega_i)]^\kappa \log(2) - \kappa \log(\eta_{\tau_i}) + (\alpha_\tau - 1) \log\left(1 - 2^{-\frac{[y_i(\omega_i)]^\kappa}{\eta_{\tau_i}^\kappa}}\right). \end{aligned}$$

A matriz  $\Delta$  pode ser escrita da seguinte forma:

$$\Delta = \left( \begin{array}{c} \Delta_{\psi(\tau)} \\ \Delta_{\kappa} \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

com,  $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . Os elementos de  $\boldsymbol{\Delta}$  são dados por  $\boldsymbol{\Delta}_{\boldsymbol{\psi}(\tau)} = \mathbf{x}_i a_i \Psi_i$  e  $\boldsymbol{\Delta}_{\boldsymbol{\kappa}} = \nu_i$  para  $i = 1, \dots, n$ , em que:

$$\Psi_i = s_{y_i} \left\{ \kappa^2 y_i^{\kappa-1} \log(2) \eta_{\tau_i}^{-(\kappa+1)} + (\alpha_\tau - 1) \left\{ \kappa^2 y_i^{\kappa-1} \ln(2) \eta_{\tau_i}^{-(\kappa+1)} \left( 1 - 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right) + \kappa^2 [\ln(2)]^2 y_i^{(2\kappa-1)} \eta_{\tau_i}^{-(2\kappa+1)} 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right\} \times \frac{1}{\left( 1 - 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right)^2} \right\},$$

e

$$\nu_i = s_{y_i} \left\{ \frac{1}{y_i} + \log(2) y_i^{\kappa-1} \eta_{\tau_i}^{-\kappa} \left[ \kappa \ln \left( \frac{\eta_{\tau_i}}{y_i} \right) - 1 \right] + (\alpha_\tau - 1) \left\{ \ln(2) y_i^{\kappa-1} \eta_{\tau_i}^{-\kappa} \left[ \kappa \ln \left( \frac{y_i}{\eta_{\tau_i}} \right) + 1 \right] \times \left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} - 1 \right) - [\ln(2)]^2 \ln \left( \frac{y_i}{\eta_{\tau_i}} \right) \eta_{\tau_i}^{-2\kappa} \kappa y_i^{(2\kappa-1)} 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} \right\} \times \frac{1}{\left( 2^{\frac{y_i^\kappa}{\eta_{\tau_i}^\kappa}} - 1 \right)^2} \right\}.$$

Escrevendo na forma matricial, tem-se que  $\boldsymbol{\Delta}_{\boldsymbol{\psi}(\tau)} = \mathbf{X}^\top \mathbf{D}(\boldsymbol{a}) \mathbf{D}(\boldsymbol{\Psi})$  e  $\boldsymbol{\Delta}_{\boldsymbol{\kappa}} = (\nu_1, \dots, \nu_n)^\top$ , em que,  $\mathbf{D}(\boldsymbol{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$  e  $\mathbf{D}(\boldsymbol{\Psi}) = \text{diag}\{\Psi_1, \Psi_2, \dots, \Psi_n\}$ .

**Perturbação na covariável contínua:** Utilizando o esquema de perturbação na covariável contínua  $\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega})$  é definido por:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}),$$

em que

$$\ell_i(\boldsymbol{\theta}) = \log[\log(2)] + \log(\alpha_\tau) + \log(\kappa) + (\kappa - 1) \log(y_i) - \frac{1}{[\eta_{\tau_i}(\boldsymbol{\omega})]^\kappa} y_i^\kappa \log(2) - \kappa \log[\eta_{\tau_i}(\boldsymbol{\omega})] + (\alpha_\tau - 1) \log \left( 1 - 2^{-\frac{y_i^\kappa}{[\eta_{\tau_i}(\boldsymbol{\omega})]^\kappa}} \right),$$

em que  $\eta_{\tau_i}(\boldsymbol{\omega}) = h^{-1}[\mathbf{x}_i^\top(\boldsymbol{\omega}_i) \boldsymbol{\psi}(\tau)]$ ,  $\mathbf{x}_i^\top(\boldsymbol{\omega}_i) = (1, x_{i1}, \dots, x_{ti}(\boldsymbol{\omega}_i), \dots, x_{i(p-1)})$  e  $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . Sendo assim, a matriz  $\boldsymbol{\Delta}$  é dada por:

$$\boldsymbol{\Delta} = \left( \begin{array}{c} \boldsymbol{\Delta}_{\boldsymbol{\psi}(\tau)} \\ \boldsymbol{\Delta}_{\boldsymbol{\kappa}} \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

em que  $\Delta_{\psi(\tau)} = (\Delta_{\psi(\tau)_{il}})$  é uma matriz  $p \times n$  expressa por:

$$\Delta_{\psi(\tau)_{il}} = \begin{cases} s_{X_t} \psi(\tau)_t a'_i a_i x_{il} z_i + s_{X_t} \psi(\tau)_t x_{il} a_i^2 c_i & \text{se } l \neq t, \quad i = 1, \dots, n, \\ s_{X_t} a_i z_i + s_{X_t} \psi(\tau)_t a'_i a_i x_{it} z_i + s_{X_t} \psi(\tau)_t x_{it} a_i^2 c_i & \text{se } l = t, \quad i = 1, \dots, n. \end{cases}$$

Adicionalmente, os elementos de  $\Delta_{\kappa}$  são dados por,  $\Delta_{\kappa} = (\zeta_1, \dots, \zeta_n)$ , em que  $\zeta_i = s_{X_t} \psi(\tau)_t a_i m_i$ . Percebe-se que,  $a_i$ ,  $z_i$ ,  $a'$ ,  $m_i$  e  $c_i$  para o modelo de RQ WER foram calculados na Seção 3.3.1.

## APÊNDICE B – ESQUEMAS DE PERTURBAÇÃO PARA O MODELO DE REGRESSÃO QUANTÍLICA LLER

**Perturbação de casos:** No esquema de perturbação de casos, a função de log-verossimilhança perturbada para o modelo de regressão quantílica LLER é dado por:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta}),$$

em que

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) = & \log(\alpha_\tau) + \log(\kappa) + (\kappa - 1) \log(y_i) - \kappa \log(\eta_{\tau_i}) - 2 \log \left[ 1 + \left( \frac{y_i}{\eta_{\tau_i}} \right)^\kappa \right] \\ & + (\alpha_\tau - 1) \log \left\{ 1 - \left[ 1 + \left( \frac{y_i}{\eta_{\tau_i}} \right)^\kappa \right]^{-1} \right\}, \end{aligned}$$

com  $0 \leq \omega_i \leq 1$  e  $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$ . A matriz de perturbação pode ser escrita como:

$$\boldsymbol{\Delta} = \left( \begin{array}{c} \boldsymbol{\Delta}_{\psi(\tau)} \\ \boldsymbol{\Delta}_\kappa \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

em que, os elementos de  $\boldsymbol{\Delta}$  são dados por  $\boldsymbol{\Delta}_{\psi(\tau)} = \mathbf{x}_i a_i \mathbf{z}_i$  e  $\boldsymbol{\Delta}_\kappa = b_i$  para  $i = 1, \dots, n$ , em que:

$$\mathbf{z}_i = \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \eta_{\tau_i}} = -\frac{\kappa}{\eta_{\tau_i}} + \frac{2\kappa y_i^\kappa}{\eta_{\tau_i}^{\kappa+1} + \eta_{\tau_i} y_i^\kappa} - (\alpha_\tau - 1) \left( \frac{\kappa \eta_{\tau_i}^{\kappa-1}}{\eta_{\tau_i}^\kappa + y_i^\kappa} \right)$$

e

$$\begin{aligned} b_i = \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \kappa} = & \frac{1}{\kappa} + \log(y_i) - \log(\eta_{\tau_i}) + \frac{2y_i^\kappa [\log(\eta_{\tau_i}) - \log(y_i)]}{\eta_{\tau_i}^\kappa + y_i^\kappa} \\ & + (\alpha_\tau - 1) \frac{\eta_{\tau_i}^\kappa [\log(y_i) - \log(\eta_{\tau_i})]}{\eta_{\tau_i}^\kappa + y_i^\kappa}, \end{aligned}$$

com  $a_i = \frac{d\eta_{\tau_i}}{d\phi_i(\tau)} = \frac{1}{h'(\eta_{\tau_i})}$ ,  $x_{il} = \frac{\partial \phi_i(\tau)}{\partial \psi_l(\tau)}$  e  $h'(\eta_{\tau_i}) = \frac{dh}{d\eta_{\tau_i}}$ . Na forma matricial  $\boldsymbol{\Delta}_{\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{z}$  e  $\boldsymbol{\Delta}_\kappa = (b_1, \dots, b_n)^\top$ , sendo  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$  e  $\mathbf{z} = (z_1, \dots, z_n)^\top$ .

**Perturbação na resposta:** Utilizando este esquema de perturbação,  $\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega})$  pode ser escrito como:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}),$$

em que

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) &= \log(\alpha_\tau) + \log(\kappa) + (\kappa - 1) \log[y_i(\omega_i)] - \kappa \log(\eta_{\tau_i}) - 2 \log \left[ 1 + \left( \frac{y_i(\omega_i)}{\eta_{\tau_i}} \right)^\kappa \right] \\ &\quad + (\alpha_\tau - 1) \log \left\{ 1 - \left[ 1 + \left( \frac{y_i(\omega_i)}{\eta_{\tau_i}} \right)^\kappa \right]^{-1} \right\}, \end{aligned}$$

em que  $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . A matriz  $\boldsymbol{\Delta}$  assume a seguinte forma:

$$\boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{\psi(\tau)} \\ \boldsymbol{\Delta}_\kappa \end{pmatrix} \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

com,  $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . Os elementos de  $\boldsymbol{\Delta}$  são dados por  $\boldsymbol{\Delta}_{\psi(\tau)} = \mathbf{x}_i a_i \Psi_i$  e  $\boldsymbol{\Delta}_\kappa = \nu_i$  para  $i = 1, \dots, n$ , em que:

$$\Psi_i = s_{y_i} \left\{ \frac{2\beta^2 y_i^{\beta-1} \eta_{\tau_i}^{\beta-1}}{(\eta_{\tau_i}^{\beta+1} + \eta_{\tau_i} y_i^\beta)^2} + (\alpha_\tau - 1) \frac{\beta^2 \eta_{\tau_i}^{\beta-1} y_i^{\beta-1}}{(\eta_{\tau_i}^\beta + y_i^\beta)^2} \right\}$$

e

$$\begin{aligned} \nu_i = s_{y_i} \left\{ \frac{1}{y_i} + \frac{2y_i^{\beta-1} \left[ \beta \log \left( \frac{\eta_{\tau_i}}{y_i} \right) \eta_{\tau_i}^\beta - (\eta_{\tau_i}^\beta + y_i^\beta) \right]}{(\eta_{\tau_i}^\beta + y_i^\beta)^2} \right. \\ \left. + (\alpha_\tau - 1) \frac{\eta_{\tau_i}^\beta y_i^{-1} (\eta_{\tau_i}^\beta + y_i^\beta) - \eta_{\tau_i}^\beta \beta y_i^{\beta-1} [\log(y_i) - \log(\eta_{\tau_i})]}{(\eta_{\tau_i}^\beta + y_i^\beta)^2} \right\}. \end{aligned}$$

Escrevendo na forma matricial, tem-se  $\boldsymbol{\Delta}_{\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{D}(\boldsymbol{\Psi})$  e  $\boldsymbol{\Delta}_\kappa = (\nu_1, \dots, \nu_n)^\top$ , em que  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$  e  $\mathbf{D}(\boldsymbol{\Psi}) = \text{diag}\{\Psi_1, \Psi_2, \dots, \Psi_n\}$ .

**Perturbação na covariável contínua:** Nesse esquema de perturbação  $\ell(\boldsymbol{\theta} | \boldsymbol{\omega})$  é dado por:

$$\ell(\boldsymbol{\theta} | \boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}),$$

em que

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) &= \log(\alpha_\tau) + \log(\kappa) + (\kappa - 1) \log(y_i) - \kappa \log[\eta_{\tau_i}(\omega)] - 2 \log \left[ 1 + \left( \frac{y_i}{\eta_{\tau_i}(\omega)} \right)^\kappa \right] \\ &\quad + (\alpha_\tau - 1) \log \left\{ 1 - \left[ 1 + \left( \frac{y_i}{\eta_{\tau_i}(\omega)} \right)^\kappa \right]^{-1} \right\}, \end{aligned}$$

com  $\eta_{\tau_i}(\omega) = h^{-1} [\mathbf{x}_i^\top(\omega_i) \boldsymbol{\psi}(\tau)]$  e  $\mathbf{x}_i^\top(\omega_i) = (1, x_{i1}, \dots, x_{ti}(\omega_i), \dots, x_{i(p-1)})$  e  $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . Sendo assim, a matriz  $\boldsymbol{\Delta}$  é dada por:

$$\boldsymbol{\Delta} = \left( \begin{array}{c} \boldsymbol{\Delta}_{\psi(\tau)} \\ \boldsymbol{\Delta}_\kappa \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

em que  $\boldsymbol{\Delta}_{\psi(\tau)} = (\Delta_{\psi(\tau)il})$  é uma matriz  $p \times n$  expressa por:

$$\Delta_{\psi(\tau)il} = \begin{cases} s_{X_t} \psi(\tau)_t a'_i a_i x_{il} z_i + s_{X_t} \psi(\tau)_t x_{il} a_i^2 c_i & \text{se } l \neq t, \quad i = 1, \dots, n, \\ s_{X_t} a_i z_i + s_{X_t} \psi(\tau)_t a'_i a_i x_{it} z_i + s_{X_t} \psi(\tau)_t x_{it} a_i^2 c_i & \text{se } l = t, \quad i = 1, \dots, n, \end{cases}$$

em que  $c_i = d_{\eta_\tau^2}^{(i)}(a_i)^2 + d_{\eta_\tau}^{(i)} a'_i a_i$ , tal que

$$d_{\eta_\tau}^{(i)} = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \eta_{\tau_i}} = -\frac{\kappa}{\eta_{\tau_i}} + \frac{2\kappa y_i^\kappa}{\eta_{\tau_i}^{\kappa+1} + \eta_{\tau_i} y_i^\kappa} - (\alpha_\tau - 1) \left( \frac{\kappa \eta_{\tau_i}^{\kappa-1}}{\eta_{\tau_i}^\kappa + y_i^\kappa} \right),$$

$$d_{\eta_\tau^2}^{(i)} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \eta_{\tau_i}^2} = \frac{\kappa}{\eta_{\tau_i}^2} - \frac{2\kappa y_i^\kappa [(\kappa + 1) \eta_{\tau_i}^\kappa + y_i^\kappa]}{(\eta_{\tau_i}^{\kappa+1} + \eta_{\tau_i} y_i^\kappa)^2} - (\alpha_\tau - 1) \left[ \kappa(\kappa - 1) \eta_{\tau_i}^{\kappa-2} (\eta_{\tau_i}^\kappa + y_i^\kappa) - \kappa^2 \eta_{\tau_i}^{2(\kappa-1)} \right] \times \frac{1}{(\eta_{\tau_i}^\kappa + y_i^\kappa)^2},$$

e

$$a' = \frac{\partial}{\partial \eta_{\tau_i}} \left( \frac{d\eta_{\tau_i}}{\phi_i(\tau)} \right) = -\frac{h''(\eta_{\tau_i})}{[h'(\eta_{\tau_i})]^2}.$$

Adicionalmente, os elementos de  $\boldsymbol{\Delta}_\kappa$  são dados por,  $\boldsymbol{\Delta}_\kappa = (\zeta_1, \dots, \zeta_n)$ , em que  $\zeta_i = s_{X_t} \psi(\tau)_t a_i m_i$ , em que:

$$m_i = -\frac{1}{\eta_{\tau_i}} + \frac{2y_i^\kappa \eta_{\tau_i} (\eta_{\tau_i}^\kappa + y_i^\kappa) + 2\kappa y_i^\kappa \eta_{\tau_i}^{\kappa+1} \ln \left( \frac{y_i}{\eta_{\tau_i}} \right)}{(\eta_{\tau_i}^{\kappa+1} + \eta_{\tau_i} y_i^\kappa)^2} - (\alpha_\tau - 1) \left[ \frac{\eta_{\tau_i}^{2\kappa-1} + \eta_{\tau_i}^{\kappa-1} y_i^\kappa + \kappa \eta_{\tau_i}^{\kappa-1} y_i^\kappa \ln \left( \frac{\eta_{\tau_i}}{y_i} \right)}{(\eta_{\tau_i}^\kappa + y_i^\kappa)^2} \right].$$

## APÊNDICE C – ESQUEMAS DE PERTURBAÇÃO PARA O MODELO DE REGRESSÃO QUANTÍLICA LER

**Perturbação de casos:** Neste caso, a função de log-verossimilhança perturbada para o modelo de regressão quantílica LER é dada por:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega}) = \sum_{i=1}^n \omega_i \ell_i(\boldsymbol{\theta}),$$

em que

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) = & \log(\alpha_\tau) + \log(\kappa) + \log(\sqrt[\kappa]{2} - 1) - \log(\eta_{\tau_i}) \\ & - (\kappa + 1) \log \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right] + (\alpha_\tau - 1) \log \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\}, \end{aligned}$$

em que  $0 \leq \omega_i \leq 1$  e  $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$ . A matriz de perturbação pode ser escrita como:

$$\boldsymbol{\Delta} = \left( \begin{array}{c} \boldsymbol{\Delta}_{\psi(\tau)} \\ \boldsymbol{\Delta}_\kappa \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

em que os elementos de  $\boldsymbol{\Delta}$  podem ser escritos como  $\boldsymbol{\Delta}_{\psi(\tau)} = \mathbf{x}_i a_i \mathbf{z}_i$  e  $\boldsymbol{\Delta}_\kappa = b_i$  para  $i = 1, \dots, n$ , em que:

$$\begin{aligned} \mathbf{z}_i = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \eta_{\tau_i}} = & -\frac{1}{\eta_{\tau_i}} + \frac{(\kappa + 1)y_i (\sqrt[\kappa]{2} - 1)}{\left[ \eta_{\tau_i} + y_i (\sqrt[\kappa]{2} - 1) \right] \eta_{\tau_i}} \\ & - (\alpha_\tau - 1) \frac{(\sqrt[\kappa]{2} - 1) \kappa y_i \left[ \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} + 1 \right]^{-(\kappa+1)}}{\eta_{\tau_i}^2 \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\}} \end{aligned}$$

e

$$\begin{aligned} b_i = \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \kappa} = & \frac{1}{\kappa} - \frac{\ln(2) \sqrt[\kappa]{2}}{(\sqrt[\kappa]{2} - 1) \kappa^2} - \log \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right] + \frac{(\kappa + 1)y_i \ln(2) \sqrt[\kappa]{2}}{\kappa^2 \eta_{\tau_i} \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]} \\ & + (\alpha_\tau - 1) \left\{ \frac{\frac{-y_i \log(2) \sqrt[\kappa]{2}}{\eta_{\tau_i} \kappa} \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-(\kappa+1)}}{1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa}} \right\}, \end{aligned}$$

com  $a_i = \frac{d\eta_{\tau_i}}{d\phi_i(\tau)} = \frac{1}{h'(\eta_{\tau_i})}$ ,  $x_{il} = \frac{\partial\phi_i(\tau)}{\partial\psi_l(\tau)}$  e  $h'(\eta_{\tau_i}) = \frac{dh}{d\eta_{\tau_i}}$ . Na forma matricial  $\Delta_{\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{a})\mathbf{z}$  e  $\Delta_\kappa = (b_1, \dots, b_n)^\top$ , sendo  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$  e  $\mathbf{z} = (z_1, \dots, z_n)^\top$ .

**Perturbação na resposta:** Usando o esquema de perturbação na resposta,  $\ell(\boldsymbol{\theta} | \boldsymbol{\omega})$  pode ser escrito da seguinte maneira:

$$\ell(\boldsymbol{\theta} | \boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}),$$

em que

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) &= \log(\alpha_\tau) + \log(\kappa) + \log(\sqrt[\kappa]{2} - 1) - \log(\eta_{\tau_i}) \\ &\quad - (\kappa + 1) \log \left[ 1 + \frac{y_i(\omega_i) (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right] \\ &\quad + (\alpha_\tau - 1) \log \left\{ 1 - \left[ 1 + \frac{y_i(\omega_i) (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\}, \end{aligned}$$

em que,  $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . A matriz  $\Delta$  pode ser escrita da seguinte forma:

$$\Delta = \left( \begin{array}{c} \Delta_{\psi(\tau)} \\ \Delta_\kappa \end{array} \right) \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0},$$

com  $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . Os elementos de  $\Delta$  são dados por  $\Delta_{\psi(\tau)} = \mathbf{x}_i a_i \Psi_i$  e  $\Delta_\kappa = \nu_i$  para  $i = 1, \dots, n$ , em que:

$$\begin{aligned} \Psi_i = s_{y_i} &\left\{ \frac{(\kappa + 1) (\sqrt[\kappa]{2} - 1) \eta_{\tau_i}^2}{[\eta_{\tau_i}^2 + y_i (\sqrt[\kappa]{2} - 1) \eta_{\tau_i}]^2} - (\alpha_\tau - 1) \left\{ (\sqrt[\kappa]{2} - 1) \kappa \left\{ \left[ \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} + 1 \right]^{-(\kappa+1)} \right. \right. \right. \\ &\quad \left. \left. - y_i (\kappa + 1) \left[ \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} + 1 \right]^{-(\kappa+2)} \frac{(\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right\} \eta_{\tau_i}^2 \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\} \right. \\ &\quad \left. \left. - [(\sqrt[\kappa]{2} - 1)]^2 \kappa^2 \eta_{\tau_i} y_i \left[ \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} + 1 \right]^{-2(\kappa+1)} \right\} \times \right. \\ &\quad \left. \left. \frac{1}{\eta_{\tau_i}^4 \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\}^2} \right\} \right\} \end{aligned}$$

e

$$\nu_i = s_{y_i} \left\{ \begin{aligned} & -\frac{(\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i} + y_i (\sqrt[\kappa]{2} - 1)} + \frac{(\kappa + 1) \ln(2) \sqrt[\kappa]{2}}{\kappa^2 \eta_{\tau_i} \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^2} \\ & + (\alpha_\tau - 1) \left\{ \left\{ -\frac{\log(2) \sqrt[\kappa]{2}}{\eta_{\tau_i} \kappa} \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa+1} \right. \right. \\ & \left. \left. + \frac{y_i \log(2) \sqrt[\kappa]{2}}{\eta_{\tau_i}^2 \kappa} (\kappa + 1) \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa+2} (\sqrt[\kappa]{2} - 1) \right\} \right. \\ & \left. \times \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\} + \frac{y_i \log(2) \sqrt[\kappa]{2}}{\eta_{\tau_i}^2} \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-2(\kappa+1)} (\sqrt[\kappa]{2} - 1) \right\} \\ & \left. \times \frac{1}{\left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\}^2} \right\}. \end{aligned} \right.$$

Escrevendo na forma matricial, tem-se  $\Delta_{\psi(\tau)} = \mathbf{X}^\top \mathbf{D}(\mathbf{a}) \mathbf{D}(\Psi)$  e  $\Delta_\kappa = (\nu_1, \dots, \nu_n)^\top$ , em que  $\mathbf{D}(\mathbf{a}) = \text{diag}\{a_1, a_2, \dots, a_n\}$  e  $\mathbf{D}(\Psi) = \text{diag}\{\Psi_1, \Psi_2, \dots, \Psi_n\}$ .

**Perturbação na covariável contínua:** Nesse esquema de perturbação,  $\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega})$  pode ser expresso por:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}),$$

em que

$$\begin{aligned} \ell_i(\boldsymbol{\theta}) &= \log(\alpha_\tau) + \log(\kappa) + \log(\sqrt[\kappa]{2} - 1) - \log[\eta_{\tau_i}(\omega)] \\ &\quad - (\kappa + 1) \log \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}(\omega)} \right] + (\alpha_\tau - 1) \log \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}(\omega)} \right]^{-\kappa} \right\}, \end{aligned}$$

em que  $\eta_{\tau_i}(\omega) = h^{-1} [\mathbf{x}_i^\top(\omega_i) \boldsymbol{\psi}(\tau)]$  e  $\mathbf{x}_i^\top(\omega_i) = (1, x_{i1}, \dots, x_{ti}(\omega_i), \dots, x_{i(p-1)})$  e  $\boldsymbol{\omega}_0 = (0, \dots, 0)^\top$ . Sendo assim, a matriz  $\Delta$  é dada por:

$$\Delta = \left( \begin{array}{c} \Delta_{\psi(\tau)} \\ \Delta_\kappa \end{array} \right) \Bigg|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0},$$

em que  $\Delta_{\psi(\tau)} = (\Delta_{\psi(\tau)_{il}})$  é uma matriz  $p \times n$  determinada por:

$$\Delta_{\psi(\tau)_{il}} = \begin{cases} s_{X_t} \psi(\tau)_t a'_i a_i x_{il} z_i + s_{X_t} \psi(\tau)_t x_{il} a_i^2 c_i & \text{se } l \neq t, \quad i = 1, \dots, n, \\ s_{X_t} a_i z_i + s_{X_t} \psi(\tau)_t a'_i a_i x_{it} z_i + s_{X_t} \psi(\tau)_t x_{it} a_i^2 c_i & \text{se } l = t, \quad i = 1, \dots, n, \end{cases}$$

tal que  $c_i = d_{\eta_\tau^2}^{(i)}(a_i)^2 + d_{\eta_\tau}^{(i)} a'_i a_i$ , com

$$d_{\eta_\tau}^{(i)} = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \eta_{\tau_i}} = -\frac{1}{\eta_{\tau_i}} + \frac{(\kappa + 1)y_i (\sqrt[\kappa]{2} - 1)}{[\eta_{\tau_i} + y_i (\sqrt[\kappa]{2} - 1)] \eta_{\tau_i}} - (\alpha_\tau - 1) \frac{(\sqrt[\kappa]{2} - 1) \kappa y_i \left[ \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} + 1 \right]^{-(\kappa+1)}}{\eta_{\tau_i}^2 \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\}},$$

$$\begin{aligned} d_{\eta_\tau^2}^{(i)} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \eta_{\tau_i}^2} = \frac{1}{\eta_{\tau_i}^2} - \frac{(\kappa + 1)y_i (\sqrt[\kappa]{2} - 1) [2\eta_{\tau_i} + y_i (\sqrt[\kappa]{2} - 1)]}{\left\{ [\eta_{\tau_i} + y_i (\sqrt[\kappa]{2} - 1)] \eta_{\tau_i} \right\}^2} \\ &\quad - (\alpha_\tau - 1) \left\{ \frac{(\sqrt[\kappa]{2} - 1)^2 y_i^2 \kappa (\kappa + 1) \left[ \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} + 1 \right]^{-(\kappa+2)}}{\eta_{\tau_i}^2} \right. \\ &\quad \times \eta_{\tau_i}^2 \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\} - (\sqrt[\kappa]{2} - 1) \kappa y_i \left[ \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} + 1 \right]^{-(\kappa+1)} \\ &\quad \times \left\{ 2\eta_{\tau_i} - 2\eta_{\tau_i} \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} - \kappa \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-(\kappa+1)} \right. \\ &\quad \left. \times y_i (\sqrt[\kappa]{2} - 1) \right\} \left. \right\} \times \frac{1}{\eta_{\tau_i}^4 \left\{ 1 - \left[ 1 + \frac{y_i (\sqrt[\kappa]{2} - 1)}{\eta_{\tau_i}} \right]^{-\kappa} \right\}^2} \end{aligned}$$

e

$$d' = \frac{\partial}{\partial \eta_{\tau_i}} \left( \frac{d\eta_{\tau_i}}{\phi_i(\tau)} \right) = -\frac{h''(\eta_{\tau_i})}{[h'(\eta_{\tau_i})]^2}.$$

Adicionalmente, os elementos de  $\Delta_\kappa$  são dados por,  $\Delta_\kappa = (\zeta_1, \dots, \zeta_n)$ , em que

$\zeta_i = s_{X_t} \psi(\tau)_t a_i m_i$ , com:

$$\begin{aligned}
m_i &= \left\{ y_i \left( \sqrt[\kappa]{2} - 1 \right) - (\kappa + 1) y_i \ln(2) \sqrt[\kappa]{2} \right\} \eta_{\tau_i}^2 + y_i^2 \left( \sqrt[\kappa]{2} - 1 \right)^2 \eta_{\tau_i} \Big\} \\
&\times \frac{1}{\left\{ \left[ \eta_{\tau_i} + y_i \left( \sqrt[\kappa]{2} - 1 \right) \right] \eta_{\tau_i} \right\}^2} - (\alpha_\tau - 1) \left\{ y_i \left\{ \left[ -\frac{\sqrt[\kappa]{2} \ln(2)}{\kappa} + \left( \sqrt[\kappa]{2} - 1 \right) \right] \right. \right. \\
&\times \left[ 1 + \frac{y_i \left( \sqrt[\kappa]{2} - 1 \right)}{\eta_{\tau_i}} \right]^{-(\kappa+1)} + \left( \sqrt[\kappa]{2} - 1 \right) (\kappa + 1) \left[ 1 + \frac{y_i \left( \sqrt[\kappa]{2} - 1 \right)}{\eta_{\tau_i}} \right]^{-(\kappa+2)} \\
&\times \left. \left. \frac{y_i \log(2) \sqrt[\kappa]{2}}{\eta_{\tau_i} \kappa} \right\} \eta_{\tau_i}^2 \left\{ 1 - \left[ 1 + \frac{y_i \left( \sqrt[\kappa]{2} - 1 \right)}{\eta_{\tau_i}} \right]^{-\kappa} \right\} + \left( \sqrt[\kappa]{2} - 1 \right) \right. \\
&\times \left. \left[ 1 + \frac{y_i \left( \sqrt[\kappa]{2} - 1 \right)}{\eta_{\tau_i}} \right]^{-2(\kappa+1)} \eta_{\tau_i} y_i^2 \log(2) \sqrt[\kappa]{2} \right\} \\
&\times \frac{1}{\eta_{\tau_i}^4 \left\{ 1 - \left[ 1 + \frac{y_i \left( \sqrt[\kappa]{2} - 1 \right)}{\eta_{\tau_i}} \right]^{-\kappa} \right\}^2}.
\end{aligned}$$