



Universidade Federal de Pernambuco  
Centro de Ciências Exatas e da Natureza  
Departamento de Estatística  
Programa de Pós-Graduação em Estatística

Victor José Araújo de Carvalho

ANÁLISE DE DIAGNÓSTICO EM MODELOS DE REGRESSÃO KUMARASWAMY

Recife  
2020

Victor José Araújo de Carvalho

ANÁLISE DE DIAGNÓSTICO EM MODELOS DE REGRESSÃO KUMARASWAMY

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

**Área de Concentração:** Estatística Matemática.

Orientador: **Prof. Dr. Francisco Cribari Neto**  
Coorientadora: **Prof<sup>a</sup>. Dr<sup>a</sup>. Patrícia Leone Espinheira Ospina**

Recife  
2020

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

C331a Carvalho, Victor José Araújo de  
Análise de diagnóstico em modelos de regressão Kumaraswamy / Victor José Araújo de Carvalho. – 2020.  
116 f.: il., fig., tab.

Orientador: Francisco Cribari Neto.  
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2020.  
Inclui referências e apêndices.

1. Estatística matemática. 2. Análise de influência. I. Cribari Neto, Francisco (orientador). II. Título.

519.5

CDD (23. ed.)

UFPE- CCEN 2020 - 57

Victor José Araújo de Carvalho

**Análise de diagnóstico em modelos de regressão Kumaraswamy**

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

Aprovada em 17 de fevereiro de 2020

**Banca Examinadora**

---

Prof.<sup>(o)</sup> Francisco Cribari Neto  
UFPE

---

Prof.<sup>(o)</sup> Aldo Willian Medina Garay  
UFPE

---

Prof.<sup>(o)</sup> Raydonal Ospina Martínez  
UFPE

---

Prof.<sup>(a)</sup> Tatiene Correia de Souza UFPB  
UFPB

---

Prof.<sup>(a)</sup>. Tarciana Liberal Pereira de Araújo  
UFPB

# AGRADECIMENTOS

Gostaria de agradecer primeiramente a Deus, por ter me dado a força que eu precisava para conseguir concluir mais esta etapa.

Aos meus pais, Valdete Araujo de Carvalho e Arnon Ferreira de Carvalho, a minha amada esposa Vanessa Furtado Xavier e o meu irmão Adeilton Vinicius Araujo de Carvalho, por serem a melhor família que eu poderia ter, dando apoio, amor e força nos momentos mais difíceis.

Ao professor Dr. Francisco Cribari Neto, pela competente orientação, por seu profissionalismo e pela paciência que teve comigo durante esses quatro anos de doutorado. A professora Dr<sup>a</sup>. Patrícia Leone Espinheira, pela exelente co-orientação.

Aos amigos, Suelena, Mariana, Leon, Thiago Luis, Cristiane, Wanessa e Rayssa Caju por cada momento que tivemos, sendo eles de estudos, de distração ou simplesmente de um ombro amigo nas horas de angústia.

Aos professores do Programa de Pós-Graduação em Estatística Universidade Federal de Pernambuco (UFPE), em especial os que tive a honra e o prazer de aprender e crescer mais a cada dia. São eles: Prof. Dr. Francisco Cribari, Prof<sup>a</sup>. Dr<sup>a</sup>. Patrícia Leone, Prof. Dr. Klaus Vasconcelos Prof<sup>a</sup>. Dr<sup>a</sup>. Maria do Carmo, Prof. Dr. Gaus Cordeiro.

Aos membros da banca examinadora, pelas contribuições que enriquecem o nosso trabalho.

Aos meus professores da graduação, por serem exemplos de dedicação e comprometimento com o trabalho e por despertarem em mim o desejo pela pesquisa científica, em especial, Prof. Dr. João Bezerra do Ó, Prof<sup>a</sup>. Dr<sup>a</sup> Flavia Jeronimo.

Aos servidores da UFPE, principalmente à Valéria Bittencourt, pela paciência, prestatividade e também pela sensibilidade com a qual nos trata.

Quero agradecer também à UFPE, pelo ensino de qualidade e suporte para conseguir realizar o curso e ao apoio financeiro fornecido pela Fundação de Amparo à Ciência e Tecnologia do Estado Pernambuco (FACEPE), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Conselho Nacional de desenvolvimento Científico e Tecnológico (Cnpq) que foram decisivos no curso, permitindo dedicação exclusiva às atividades.

Aos demais colegas que fiz no Doutorado.

À todos aqueles que contribuíram de forma direta ou indireta para que eu pudesse crescer e concluir o curso.

# RESUMO

A distribuição Kumaraswamy é uma lei de probabilidade contínua com suporte em  $(a, b)$ ,  $a < b$  com  $a, b \in \mathbb{R}$ . Aqui consideramos apenas o intervalo  $(0, 1)$ . Trata-se de uma distribuição bastante flexível, podendo assim ser usada para modelar uma ampla gama de variáveis duplamente limitadas. Diversos autores vêm trabalhando com essa distribuição, entre eles destacam-se (MITNIK; BAEK, 2013) que propuseram uma reparametrização possibilitando a criação do modelo de regressão Kumaraswamy. No presente trabalho desenvolvemos várias ferramentas de diagnóstico para esta classe de modelos. Em particular, propomos dois novos resíduos e desenvolvemos medidas de influência global, local e conforme. Adicionalmente, propomos um chute inicial para o processo iterativo scoring de Fisher a ser utilizado no método de estimação por máxima verossimilhança. O método de influência local foi decisivo para a modelagem da proporção de umidade relativa do ar em função da temperatura na cidade de Recife no mês de setembro de 2019, problemática relevante tanto do ponto de vista de saúde pública, econômico e ambiental.

**Palavras-chave:** Análise de influência. Modelo de regressão Kumaraswamy. Resíduos. Variáveis duplamente limitadas.

# ABSTRACT

The Kumaraswamy distribution is a continuous probability law supported by  $(a, b)$ ,  $a < b$  with  $a, b \in \mathbb{R}$ . Here we consider only the range  $(0, 1)$ . It is a very flexible distribution, so it can be used to model a wide range of doubly limited variables. Several authors have been working with this distribution, among which (MITNIK; BAEK, 2013) stand out, who proposed a reparametrization enabling the creation of the Kumaraswamy regression model. In the present work we have developed several diagnostic tools for this class of models. In particular, we propose two new wastes and develop measures of global influence, local influence and conformal influence. In addition, we propose an initial guess for Fisher's iterative scoring process to be used in the maximum likelihood estimation method. The local influence method was decisive for modeling the proportion of relative humidity of the air as a function of temperature in the city of Recife in September 2019, a relevant issue from the point of view of public, economic and environmental health.

**Keywords:** Influence analysis. Kumaraswamy regression model. Residual. Limited double variables.

# LISTA DE FIGURAS

Figura 1 – Densidades Kumaraswamy . . . . .	20
Figura 2 – Gráficos QQ; Modelo: $\log(w_t/1-w_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, 2, \dots, 40$ , $w \in (0.0169, 0.3570)$ , $\phi = 2, 4$ e $7$ . . . . .	37
Figura 3 – Gráficos QQ; Modelo: $\log(w_t/1-w_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, 2, \dots, 40$ , $w \in (0.2353, 0.7758)$ , $\phi = 2, 10$ e $20$ . . . . .	38
Figura 4 – Gráficos QQ; Modelo: $\log(w_t/1-w_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, 2, \dots, 40$ , $w \in (0.8499, 0.9994)$ , $\phi = 3, 50$ e $150$ . . . . .	39
Figura 5 – Gráficos QQ; Modelo: $\log(w_t/1-w_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, 2, \dots, 120$ , $w \in (0.0169, 0.3570)$ , $\phi = 2, 4$ e $7$ . . . . .	40
Figura 6 – Gráficos QQ; Modelo: $\log(w_t/1-w_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, 2, \dots, 120$ , $w \in (0.2353, 0.7758)$ , $\phi = 2, 10$ e $20$ . . . . .	41
Figura 7 – Gráficos QQ; Modelo: $\log(w_t/1-w_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, 2, \dots, 120$ , $w \in (0.8499, 0.9994)$ , $\phi = 3, 50$ e $150$ . . . . .	42
Figura 8 – Boxplot(a) e histograma(b) das observações da variável resposta. . . . .	43
Figura 9 – Boxplot das observações da variável nível de aflatoxina. . . . .	44
Figura 10 – Gráficos de resíduos do modelo $\log(w_t/(1-w_t)) = \beta_1 + \beta_2 x_{t2}$ e $\log(\phi_t) =$ $\gamma_1 + \gamma_2 \log(x_{t2})$ , $t = 1, 2, \dots, 34$ ; Contaminação do Amendoim. . . . .	45
Figura 11 – Gráficos normais de probabilidade com envelopes simulados para o modelo $\log(w_t/(1-w_t)) = \beta_1 + \beta_2 x_{t2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 \log(x_{t2})$ , $t = 1, 2, \dots, 34$ ; Contaminação do Amendoim. . . . .	46
Figura 12 – Gráficos de resíduos do modelo $\log(w_t/(1-w_t)) = \beta_0 + \beta_1 x_t$ , $t =$ $1, 2, \dots, 34$ , $\phi$ constante; Contaminação do Amendoim. . . . .	48
Figura 13 – Gráficos normais de probabilidade com envelopes simulados para o modelo $\log(w_t/(1-w_t)) = \beta_0 + \beta_1 x_t$ , $t = 1, 2, \dots, 34$ , $\phi$ constante; Contaminação do Amendoim. . . . .	48
Figura 14 – Gráfico de Dispersão; Contaminação do Amendoim. . . . .	49
Figura 15 – Boxplot(a) e histograma(b) das observações da variável resposta. . . . .	50
Figura 16 – Boxplot das observações da variável temperatura. . . . .	51
Figura 17 – Gráficos normais de probabilidade com envelopes simulados para o modelo $\log(w_t/(1-w_t)) = \beta_0 + \beta_1 x_t$ , $t = 1, 2, \dots, 30$ , $\phi$ constante; Umidade Relativa do Ar. . . . .	52
Figura 18 – Gráficos de resíduos do modelo $\log(w_t/(1-w_t)) = \beta_0 + \beta_1 x_t$ , $t =$ $1, 2, \dots, 30$ , $\phi$ constante; Umidade Relativa do Ar. . . . .	52
Figura 19 – (a) Superfície $G(\delta)$ , (b) Secção normal. . . . .	57
Figura 20 – Linhas projetadas $LD_{\delta_0+a_1l} \times l$ e $LD_{\delta_0+a_2l} \times l$ . . . . .	57

Figura 21 – Gráficos influência local; Contaminação do Amendoim. . . . .	82
Figura 22 – Gráficos influência local total; Contaminação do Amendoim. . . . .	83
Figura 23 – Gráficos de influência local conforme sob perturbação de ponderação de casos; Contaminação do Amendoim. . . . .	84
Figura 24 – Gráficos de influência local conforme sob perturbação da variável res- posta; Contaminação do Amendoim. . . . .	85
Figura 25 – Gráficos de influência local conforme sob perturbação da covariada; Contaminação do Amendoim. . . . .	86
Figura 26 – Gráficos influência local; Umidade Relativa do Ar. . . . .	88
Figura 27 – Gráficos influência local total; Umidade Relativa do Ar. . . . .	89
Figura 28 – Gráficos de influência local conforme sob perturbação de ponderação de casos; Umidade Relativa do Ar. . . . .	91
Figura 29 – Gráficos de influência local conforme sob perturbação da variável res- posta; Umidade Relativa do Ar. . . . .	91
Figura 30 – Gráficos de influência local conforme sob perturbação da covariada; Umidade Relativa do Ar. . . . .	92
Figura 31 – Gráficos de influência local conforme sob perturbação de ponderação de casos; Modelo de regressão beta; Contaminação do Amendoim. . . . .	97
Figura 32 – Gráficos de influência local conforme sob perturbação da variável res- posta; Modelo de regressão beta; Contaminação do Amendoim. . . . .	98
Figura 33 – Gráficos de influência local conforme sob perturbação da covariada; Modelo de regressão beta; Contaminação do Amendoim. . . . .	99

# LISTA DE TABELAS

Tabela 1 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos	
quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:	
$\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, \dots, 5,$	
$t = 1, 2, \dots, 20, w \in (0.0258, 0.3452)$ e $\phi = 2$ . . . . .	31
Tabela 2 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos	
quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:	
$\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, \dots, 5,$	
$t = 1, 2, \dots, 20, w \in (0.0258, 0.3452)$ e $\phi = 4$ . . . . .	32
Tabela 3 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos	
quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:	
$\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, \dots, 5,$	
$t = 1, 2, \dots, 20, w \in (0.0258, 0.3452)$ e $\phi = 7$ . . . . .	32
Tabela 4 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos	
quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:	
$\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, \dots, 5,$	
$t = 1, 2, \dots, 20, w \in (0.2440, 0.8237)$ e $\phi = 3$ . . . . .	33
Tabela 5 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos	
quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:	
$\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, \dots, 5,$	
$t = 1, 2, \dots, 20, w \in (0.2440, 0.8237)$ e $\phi = 10$ . . . . .	33
Tabela 6 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos	
quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:	
$\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, \dots, 5,$	
$t = 1, 2, \dots, 20, w \in (0.2440, 0.8237)$ e $\phi = 20$ . . . . .	34
Tabela 7 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos	
quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:	
$\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, \dots, 5,$	
$t = 1, 2, \dots, 20, w \in (0.8855, 0.9983)$ e $\phi = 3$ . . . . .	35
Tabela 8 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos	
quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:	
$\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, \dots, 5,$	
$t = 1, 2, \dots, 20, w \in (0.8855, 0.9983)$ e $\phi = 50$ . . . . .	35

Tabela 9 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo: $\log(\frac{w_t}{1-w_t}) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $x_{ti} \sim U(0, 1)$ , $i = 2, \dots, 5$ , $t = 1, 2, \dots, 20$ , $w \in (0.8855, 0.9983)$ e $\phi = 150$ .	36
Tabela 10 – Estimativas dos parâmetros, erros-padrão e $p$ -valores do modelo $\log(w_t/(1-w_t)) = \beta_1 + \beta_2 x_{t2}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 \log(x_{t2})$ ; Contaminação do Amendoim.	46
Tabela 11 – Estimativas dos parâmetros, erros-padrão e $p$ -valores do modelo $\log(w_t/(1-w_t)) = \beta_1 + \beta_2 x_{t2}$ e $\phi$ constante; Contaminação do Amendoim.	49
Tabela 12 – Estimativas dos parâmetros, erros-padrão e $p$ -valores do modelo $\log(w_t/(1-w_t)) = \beta_1 + \beta_2 x_{t2}$ e $\phi$ constante; Umidade Relativa do Ar.	53
Tabela 13 – Estimativas dos parâmetros, erros-padrão e $p$ -valores do modelo $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2}$ e $\phi$ constante; Contaminação do Amendoim.	81
Tabela 14 – Variações percentuais nas estimativas dos parâmetros após a retirada de observações influentes; Contaminação do Amendoim.	83
Tabela 15 – Variações percentuais nas estimativas dos parâmetros após a retirada das observações influentes utilizando o método de influência local conforme; Contaminação do Amendoim.	86
Tabela 16 – Estimativas dos parâmetros, erros-padrão e $p$ -valores do modelo $\log(w_t/(1-w_t)) = \beta_1 + \beta_2 x_{t2}$ e $\phi$ constante; Umidade Relativa do Ar.	87
Tabela 17 – Variação percentual nas estimativas dos parâmetros retirando as observações influentes; Umidade Relativa do Ar.	90
Tabela 18 – Variação percentual nas estimativas dos parâmetros retirando as observações influentes; Umidade Relativa do Ar.	92
Tabela 19 – Estimativas dos parâmetros, erros-padrão e $p$ -valores do modelo $\log(\mu_t/(1-\mu_t)) = \beta_0 + \beta_1 \sqrt{x_{t2}}$ e $\phi$ constante; Modelo de regressão beta; Contaminação do Amendoim.	96
Tabela 20 – Variação percentual nas estimativas dos parâmetros retirando as observações influentes utilizando o método de influência local conforme; Contaminação do Amendoim.	99

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>2</b>	<b>ANÁLISE DE RESÍDUOS</b>	<b>17</b>
2.1	Introdução	17
2.2	Modelo de regressão Kumaraswamy	20
2.3	Resíduo quantílico	27
2.4	Resíduo ponderado	28
2.5	Resíduo ponderado padronizado	29
2.6	Avaliação numérica	30
2.7	Aplicações	42
2.8	Aplicação I: Contaminação do amendoim	43
2.9	Aplicação II: Umidade relativa do ar	49
2.10	Conclusão	53
<b>3</b>	<b>ANÁLISE DE INFLUÊNCIA</b>	<b>54</b>
3.1	Introdução	54
3.2	Modelo de regressão Kumaraswamy	57
3.3	Modelo de regressão beta	65
3.4	Influência global	70
3.5	Influência local	70
3.6	Influência local conforme	72
3.7	Propriedades da curvatura normal conforme	73
3.8	Influência de um único autovetor	74
3.9	Contribuição agregada de vetores básicos de perturbação	75
3.10	Cálculo das perturbações para o modelo de regressão Kumaraswamy	77
3.11	Perturbação de ponderação de casos	77
3.12	Perturbação da variável resposta	78
3.13	Perturbação individual de covariáveis	78
3.14	Matriz $Z$ totalmente diferente da matriz $X$	78
3.15	Matriz $Z$ igual à matriz $X$	79
3.16	A $p'$ -ésima coluna da matriz $Z$ igual à $p$ -ésima coluna da matriz $X$	79
3.17	A $p'$ -ésima coluna da matriz $Z$ é função da $p$ -ésima coluna da matriz $X$	80
3.18	Aplicações: Análise de Influência	80

3.19	Aplicação I: Contaminação do amendoim, modelo de regressão Ku- maraswamy	81
3.20	Aplicação II: Umidade relativa do ar, modelo de regressão Kuma- raswamy	86
3.21	Cálculo das perturbações para o modelo de regressão beta	93
3.22	Perturbação de ponderação de casos	93
3.23	Perturbação da variável resposta	93
3.24	Perturbação individual de covariáveis	94
3.25	Matriz $Z$ totalmente diferente da matriz $X$	94
3.26	Matriz $Z$ igual à matriz $X$	95
3.27	A $p'$ -ésima coluna da matriz $Z$ igual à $p$ -ésima coluna da matriz $X$	95
3.28	A $p'$ -ésima coluna da matriz $Z$ é função da $p$ -ésima coluna da matriz $X$	95
3.29	Aplicação: contaminação do amendoim, modelo de regressão beta	96
3.30	Conclusão	100
4	CONCLUSÃO	101
	REFERÊNCIAS	103
	APÊNDICE A – RESÍDUOS	106
	APÊNDICE B – INFLUÊNCIA LOCAL	109

# 1 INTRODUÇÃO

Modelos de regressão são comumente utilizados para analisar o comportamento de variáveis de interesse que são relacionadas com outras variáveis. Quando a relação entre a variável dependente (resposta) e as variáveis independentes (regressoras ou covariáveis) é linear o modelo de regressão frequentemente utilizado é o modelo de regressão normal linear. Tal modelo, contudo, impõe outras suposições, além da suposição de linearidade; por exemplo, considera-se que a variável resposta segue distribuição normal. Por esse motivo, o modelo de regressão normal linear não é apropriado em situações em que a variável resposta, por exemplo, assume valores no intervalo  $(0, 1)$ , já que a suposição de normalidade pode conduzir a valores ajustados para a variável dependente fora dos limites de tal intervalo. Adicionalmente, muitas vezes os dados apresentam assimetria e heteroscedasticidade, não satisfazendo as suposições do modelo de regressão normal linear. Uma possível solução para o problema seria utilizar uma transformação na variável resposta fazendo com que a nova variável assumira valores na reta real. No entanto, esta solução acarreta uma grande desvantagem ao dificultar a interpretação dos parâmetros do modelo com relação à variável resposta original.

Em muitas situações práticas, existe a necessidade de modelar dados no intervalo  $(0, 1)$ , por exemplo taxas e proporções. Vários modelos de regressão foram propostos nas últimas décadas para variáveis dependentes que assumem valores no intervalo  $(0, 1)$ . Por exemplo, (BARNDORFF-NIELSEN; JØRGENSEN, 1991) propuseram o modelo de regressão simplex, (FERRARI; CRIBARI-NETO, 2004) propuseram o modelo de regressão beta e (LEMONTE; BAZÁN, 2016) propuseram o modelo de regressão Johnson  $S_B$ . Em particular, (FERRARI; CRIBARI-NETO, 2004) propuseram uma reparametrização para a distribuição beta de modo a permitir a modelagem da média da variável resposta. Tal modelo está implementado no software estatístico **R** distribuído por (<https://www.r-project.org/>) através do pacote **betareg**.

A distribuição beta é uma das leis mais usadas para modelar experimentos aleatórios que produzem resultados no intervalo  $(0, 1)$ . É uma família de distribuições probabilidade bastante flexível, devido a possuir dois parâmetros de forma,  $p$  e  $q$ .

Segundo (BALAKRISHNAN; NEVZOROV, 2003) a distribuição beta com parâmetros  $p$  e  $q$ , possui propriedade de quase-simetria, isto é, se  $Y \sim \text{Beta}(p, q)$  então  $1 - Y \sim \text{Beta}(q, p)$ . Adicionalmente,  $Y \sim \text{Beta}(p, q)$ , então  $-\log(Y) \sim \text{Exponencial}(p)$ . Dentre as diversas forma assumidas por sua função de densidade de probabilidade, encontra-se a uniforme quando  $p = q = 1$ . A densidade beta é unimodal quando  $p > 0$  e  $q > 1$ .

Outra opção ainda pouco explorada é o modelo de regressão Kumaraswamy proposto

por (MITNIK; BAEK, 2013) em que os autores fazem uma reparametrização da distribuição Kumaraswamy permitindo a modelagem da mediana da variável resposta. Esta distribuição foi proposta em (KUMARASWAMY, 1976) e reformulada em (KUMARASWAMY, 1980). Assim como a distribuição beta, ela pode ser usada para modelar variáveis dependentes que assumem valores no intervalo  $(0, 1)$ . A distribuição Kumaraswamy é uma distribuição de probabilidade contínua com suporte em  $(a, b)$ ,  $a < b$  com  $a, b \in \mathbb{R}$ . No presente trabalho, consideraremos apenas o intervalo  $(0, 1)$ . Esta distribuição se assemelha em muitos aspectos à distribuição beta. Trata-se de uma distribuição bastante flexível. Sua densidade assume formas muito distintas dependendo dos valores de seus parâmetros, podendo assim ser usada para modelar uma grande variedade de dados.

Diversos autores vêm trabalhando com essa distribuição. (JONES, 2009) faz um comparativo com a distribuição beta, exibindo algumas vantagens no uso da distribuição Kumaraswamy tais como: função de distribuição e função quantil em forma fechada, (LEMONTE, 2011) desenvolveu uma correção de viés para os estimadores de máxima verossimilhança dos parâmetros que indexam o modelo. (MITNIK, 2013) desenvolveu muitas propriedades importantes, por exemplo: semelhante ao que ocorre com variáveis que seguem distribuição beta, qualquer transformação linear de uma variável com distribuição Kumaraswamy também segue esta distribuição e possui os mesmos parâmetros de forma que a distribuição original. (MITNIK; BAEK, 2013) propuseram o modelo de regressão Kumaraswamy a partir de uma reparametrização dos parâmetros de forma originais.

A presente tese de doutorado objetiva desenvolver ferramentas de diagnóstico para a classe de modelos de regressão Kumaraswamy. Em particular, propomos dois novos resíduos e desenvolvemos medidas de influência global, local e conforme.

A análise de resíduos é uma importante ferramenta usada para identificar se um modelo está ou não bem ajustado em relação aos dados. Na maioria dos casos temos conhecimento limitado sobre a relação entre a variável dependente (resposta) e as variáveis independentes (regressoras ou covariáveis). Se tomarmos os valores observados de tais variáveis como resultantes de um experimento devemos ter uma ferramenta teórica, isto é, um modelo matemático, através do qual estas variáveis estejam relacionadas, para atuar como base do processo gerador de dados. Entretanto, todos os modelos são, por definição, simplificações da realidade, isto é, modelos matemáticos que aproximam os dados obtidos do fenômeno em estudo. Desta forma é extremamente importante analisar o modelo, no sentido de avaliar a qualidade de sua aderência à realidade.

A análise de resíduos tem como principal objetivo avaliar possíveis afastamentos das suposições admitidas para o modelo, além da detecção de pontos mal ajustados ou aberrantes. As técnicas gráficas que utilizam resíduos são ferramentas de amplo uso na análise de diagnóstico. Por exemplo, o uso de envelopes simulados conforme proposto por (ATKINSON, 1981), inicialmente desenvolvidos para uso em modelos de regressão normais,

permite avaliar a qualidade do ajuste do modelo postulado.

Na classe de modelos para variáveis distribuídas continuamente no intervalo  $(0,1)$  vários resultados de diagnóstico foram obtidos nos últimos anos, em especial para a classe de modelos de regressão beta. Em (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) são desenvolvidos resíduos e medidas de influência local para a classe de modelos de regressão beta linear. Adicionalmente, (ESPINHEIRA; SANTOS; CRIBARI-NETO, 2017) apresentam um novo resíduo para a classe de modelos de regressão beta não linear e (LEMONTE; BAZÁN, 2016) também propõem resíduos e medidas de influência local para a classe de modelos Johnson  $S_B$ . Recentemente, (ESPINHEIRA; SILVA, 2019) desenvolveram um resíduo baseado no processo iterativo scoring de Fisher para  $\beta$  e medidas de influência local, considerando diversos esquemas de perturbação para os modelos de regressão simplex em que, tanto a média quanto a dispersão podem ser explicadas com base em preditores não lineares.

Inicialmente, apresentaremos a distribuição Kumaraswamy e algumas de suas principais propriedades. Em seguida, exibiremos as reparametrizações propostas por (MITNIK; BAEK, 2013), o modelo de regressão por eles apresentado e o resíduo quantílico aleatorizado apresentado por (DUNN; SMYTH, 1996), que é o resíduo frequentemente utilizado com o modelo de regressão Kumaraswamy. Em seguida, proporemos dois novos resíduos para este modelo de regressão. Eles seguem da ideia que está por trás dos resíduos ponderado e ponderado padronizado propostos por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) para o modelo de regressão beta com dispersão constante.

Para completar a análise de diagnóstico utilizaremos medidas de influência. Elas fornecem uma maneira eficaz de detectar observações influentes. A primeira técnica desenvolvida para avaliar o impacto individual de casos no processo de estimação é a exclusão de casos, que se tornou uma ferramenta muito popular entre estatísticos aplicados. Essa técnica é conhecida como análise de influência global. Deve-se notar, contudo, que o que ocorre na exclusão de casos é que se exclui toda a informação de uma observação e dificilmente podemos dizer se esta observação possui alguma influência sobre um aspecto específico do modelo. Para evitar isso utilizaremos análise de influência local, onde se investiga a sensibilidade das inferências realizadas sobre o modelo sob pequenas perturbações. Nesse contexto, (COOK, 1986) propôs uma estrutura geral para detectar observações influentes. Diversos autores consideraram o método de influência local em modelos de regressão; ver, por exemplo, (LAWRANCE, 1988), (THOMAS; COOK, 1990), (PAULA, 1996), (LESAFFRE; VERBEKE, 1998) e, mais recentemente, (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008), (ROCHA; SIMAS, 2011), (LEMONTE; BAZÁN, 2016) e (ESPINHEIRA; SILVA, 2019).

(POON; POON, 1999) chamam atenção para o fato de que a curvatura normal pode assumir qualquer valor real e não ser invariante sob uma mudança uniforme de

escala, ocasionando perda de objetividade no julgamento da grandeza da curvatura. Com o objetivo de solucionar esse problema e conseqüentemente aperfeiçoar o método de influência local, (POON; POON, 1999) propõem que façamos uso da curvatura conforme, que está relacionada com a curvatura normal, mas assume valores em um intervalo limitado e é invariante sob uma classe de reparametrizações que definiremos mais a seguir. A curvatura normal conforme e suas propriedades nos dão suporte para construirmos valores de referência a serem utilizados para a detecção de pontos influentes.

### Organização da tese

A presente tese é constituída de 3 capítulos e dois apêndices. No Capítulo 1, apresentamos uma introdução relativa à motivação da tese e também um breve histórico de trabalhos relacionados. No Capítulo 2, apresentamos a distribuição Kumaraswamy e o modelo de regressão associado, assim como algumas propriedades tais como a função escore e matriz de informação observada. Definimos ainda os resíduos ponderado e ponderado padronizado, fazemos uma avaliação numérica destes resíduos e, por fim, apresentamos aplicações dos resíduos a dados reais. No Capítulo 3, apresentamos medidas de influência global, local e conforme para o modelo de regressão Kumaraswamy e aplicamos essas medidas a um conjunto de dados reais. Por fim, apresentamos a medida de influência local conforme para o modelo de regressão beta e aplicamos essa medida a dois conjuntos de dados reais. Nos Apêndices A e B detalhamos alguns cálculos realizados ao longo da tese.

### Suporte computacional

Utilizamos a linguagem de programação matricial **Ox**, criada por Jungen Doornik em 1994, atualmente na versão 8.02, distribuída de forma gratuita para uso acadêmico pelo site <http://www.doornik.com/download.html> e disponível para as principais plataformas (Linux, Mac OS-X e Windows). Esta linguagem permite a implementação de técnicas estatísticas com extrema facilidade, precisão e velocidade. Para mais detalhes sobre esta linguagem de programação, ver (DOORNIK, 2006).

Para a análise gráfica utilizamos o ambiente computacional **R** em sua versão 3.6.2 que está disponível para as principais plataformas (Linux, Mac OS-X e Windows). O **R** é uma ferramenta de programação, análise de dados e geração de gráficos que recebe contribuições de pessoas em todo o mundo e encontra-se disponível gratuitamente em <https://www.r-project.org/>. Para mais detalhes, ver (VENABLES; RIPLEY, 2002), (CRIBARI-NETO; ZARKOS, 1999) e (VENABLES; SMITH, 2009).

Esta tese foi digitada através do sistema tipográfico  $\text{\LaTeX}$ , desenvolvido por Leslie Lamport em meados de 1980 e que é uma coleção de macros ou rotinas do sistema  $\text{\TeX}$ , criado por Donald Knuth na Universidade de Stanford. Mais detalhes podem ser encontrados em (LAMPART, 1994) ou através do site <http://www.tex.ac.uk/CTAN/latex>.

## 2 ANÁLISE DE RESÍDUOS

### 2.1 Introdução

Um modelo estatístico nada mais é que uma simplificação de algum aspecto da realidade, que se realiza sob incerteza e sobre o qual se dispõe de uma quantidade limitada de informações, tais como informações sobre as variáveis supostamente envolvidas nesta realidade. Adicionalmente, para tentar explicar tal realidade é necessário um experimento com base em valores observados destas variáveis, formando um conjunto de dados. Uma vez que se trata de valores observados, tem-se que considerar o caráter aleatório de alguma destas variáveis. Este é o contexto de um modelo de regressão, o qual se baseia em um conjunto de suposições, desde as variáveis envolvidas no problema real até a expressão matemática que supostamente descreve o problema e ainda a distribuição de probabilidade de interesse. Neste sentido, uma etapa de extrema importância na análise de regressão é a validação do modelo, isto é, a avaliação da qualidade desta simplificação da realidade. É disso que trata a análise de diagnóstico.

A análise de diagnóstico teve início a partir da análise de resíduos com o objetivo de detectar pontos mal ajustados pelo modelo ou aberrantes e avaliar indícios de afastamento das hipóteses do modelo, sendo a principal delas a adequação da distribuição postulada para a variável resposta. A análise de resíduos pode ser baseada nos resíduos ordinários, em suas possíveis padronizações ou em resíduos construídos a partir dos componentes da função desvio; ver, por exemplo, (MCCULLAGH; NELDER, 1989). As técnicas gráficas que utilizam resíduos são ferramentas de ampla utilização na análise de diagnóstico. Por exemplo, o uso de envelopes simulados, conforme proposto por (ATKINSON, 1981) inicialmente para modelos de regressão com erros normais, permite avaliar a qualidade do ajuste do modelo postulado a um conjunto de dados reais.

A importância da análise de resíduos também é destacada pela substancial representatividade do tema na produção científica relacionada à modelagem de dados, como por exemplo, os trabalhos de (MCCULLAGH; NELDER, 1989), (WILLIAMS, 1984), (WILLIAMS, 1987), (DAVISON; GIGLI, A, 1989), (FAHRMEIR; TUTZ, 1994), (DUNN; SMYTH, 1996), (PAULA, 1996), (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008), (ROCHA; SIMAS, 2011), (LEMONTE; BAZÁN, 2016) e (ESPINHEIRA; SILVA, 2019).

Nosso interesse aqui recai sobre o modelo de regressão Kumaraswamy. Inicialmente, apresentamos a distribuição Kumaraswamy, algumas de suas principais propriedades e seu modelo de regressão. Em seguida apresentaremos o resíduo quantílico aleatorizado apresentado em (DUNN; SMYTH, 1996), que é o resíduo que frequentemente é utilizado

com o modelo de regressão Kumaraswamy. Proporemos então dois novos resíduos para este modelo de regressão, que seguem a ideia que está por trás dos resíduos padronizados propostos por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) para o modelo de regressão beta com dispersão constante. Avaliaremos a distribuição empírica destes resíduos e apresentaremos aplicações a dados reais em que realizamos análise de resíduos.

A distribuição Kumaraswamy foi proposta por (KUMARASWAMY, 1976) e reformulada no artigo de (KUMARASWAMY, 1980). Assim como a distribuição beta, ela pode ser usada para modelar variáveis que assumem valores no intervalo  $(0, 1)$  e apresenta propriedades muito favoráveis, em especial em relação à facilidade de obtenção de expressões para os quantis populacionais em termos de seus parâmetros. Em particular, o interesse desta tese de doutorado recai sobre a modelagem da mediana da variável resposta  $Y$  cujo comportamento probabilístico pode ser explicado com base na distribuição Kumaraswamy.

A distribuição Kumaraswamy é uma distribuição de probabilidade contínua com suporte em  $(a, b)$ ,  $a < b$  com  $a, b \in \mathbb{R}$ . Nós consideraremos apenas o intervalo  $(0, 1)$ . Esta distribuição se assemelha em muito aspectos à distribuição beta. Trata-se de uma distribuição bastante flexível. Sua densidade assume formas muito distintas dependendo dos valores dos parâmetros, podendo assim ser usada para modelar uma gama de variáveis duplamente limitadas.

Entretanto, existem importantes diferenças entre essas duas distribuições. Por um lado, a distribuição Kumaraswamy possui função de distribuição cumulativa e função quantil em formas fechadas, o que a torna mais adequada do que a distribuição beta para tarefas que necessitam da geração de ocorrências de variáveis aleatórias, por exemplo, reamostragem bootstrap. Por outro lado, a disponibilidade de expressões simples em forma fechada para a média e para a variância da distribuição beta em termos de seus parâmetros facilita seu uso em aplicações.

Uma variável aleatória  $Z$  segue distribuição Kumaraswamy se sua função de densidade pode ser escrita como

$$f(z; p, q, a, b) = \frac{1}{b-a} pq \left( \frac{z-a}{b-a} \right)^{p-1} \left[ 1 - \left( \frac{z-a}{b-a} \right)^p \right]^{q-1}, \quad a < z < b,$$

com parâmetros de forma  $p > 0$  e  $q > 0$  e tal que  $z \in (a, b)$ ,  $a < b \in \mathbb{R}$ . Denotaremos essa forma geral por  $Z \sim Kum(p, q, a, b)$ .

Uma variável aleatória  $Y$  segue distribuição Kumaraswamy com parâmetros  $p, q$  e suporte no intervalo  $(0, 1)$ , denotada por  $Y \sim K(p, q)$ , se a sua função de densidade é da forma

$$f(y; p, q) = pqy^{p-1}(1-y^p)^{q-1}, \quad y \in (0, 1), \quad (2.1)$$

com  $p > 0$  e  $q > 0$ . A função de distribuição acumulada é dada por

$$F(y; p, q) = 1 - (1 - y^p)^q. \quad (2.2)$$

A média e a variância de  $Y$  são dadas, respectivamente, por

$$\mathbb{E}(Y) = \frac{q\Gamma(1 + \frac{1}{p})\Gamma(q)}{\Gamma(1 + \frac{1}{p} + q)} \quad (2.3)$$

e

$$\text{Var}(Y) = \mu_2 - \mu_1^2,$$

em que

$$\mu_r = \frac{q\Gamma(1 + \frac{r}{p})\Gamma(q)}{\Gamma(1 + \frac{r}{p} + q)} = qB\left(1 + \frac{r}{p}, q\right), \quad r = 1, 2, \dots,$$

é o  $r$ -ésimo momento, e  $\Gamma(\cdot)$  e  $B(\cdot, \cdot)$  são as funções gama e beta, respectivamente.

Como notado anteriormente, uma vantagem da distribuição Kumaraswamy é possuir função quantil em forma fechada. Tal função é dada por

$$y = F^{-1}(u) = [1 - (1 - u)^{\frac{1}{q}}]^{\frac{1}{p}}, \quad (2.4)$$

em que  $0 < u < 1$ .

É importante destacar algumas características importantes de variáveis aleatórias que seguem distribuição Kumaraswamy. Semelhantemente ao que ocorre com variáveis aleatórias que seguem distribuição beta, qualquer transformação linear de uma variável com distribuição Kumaraswamy também segue distribuição Kumaraswamy e sua distribuição tem os mesmos parâmetros de forma que a distribuição original. Diferentemente de variáveis aleatórias que seguem distribuição beta, variáveis aleatórias que seguem distribuição Kumaraswamy são fechadas sob exponenciação, isto é, para uma variável aleatória que segue distribuição Kumaraswamy, qualquer potência positiva dela também segue distribuição Kumaraswamy. Em suma, temos as propriedades elementares a seguir.

**Proposição 1.** *Seja  $Y = \beta + \alpha X$ , com  $\alpha \neq 0$ , então  $X \sim Kum(p, q, c, b) \Leftrightarrow Y \sim Kum(p, q, \beta + \alpha c, \beta + \alpha b)$ .*

Um resultado que segue como consequência imediata desta proposição está enunciado a seguir.

**Corolário 1.** *Seja  $X \sim Kum(p, q)$  e  $Y$  uma padronização de  $X$ , isto é,  $Y = \frac{X - \mu(X)}{[\text{var}(X)]^{1/2}}$ , então  $Y \sim Kum\left(p, q, -\frac{\mu(X)}{[\text{var}(X)]^{1/2}}, \frac{1 - \mu(X)}{[\text{var}(X)]^{1/2}}\right)$ .*

Outra consequência está enunciada abaixo.

**Corolário 2.** *Se  $Y = d - X$ ,  $d$  sendo um número real qualquer, então  $X \sim Kum(p, q, c, b) \Leftrightarrow Y \sim Kum(p, q, d - c, d - b)$ .*

Outra propriedade importante está enunciada a seguir

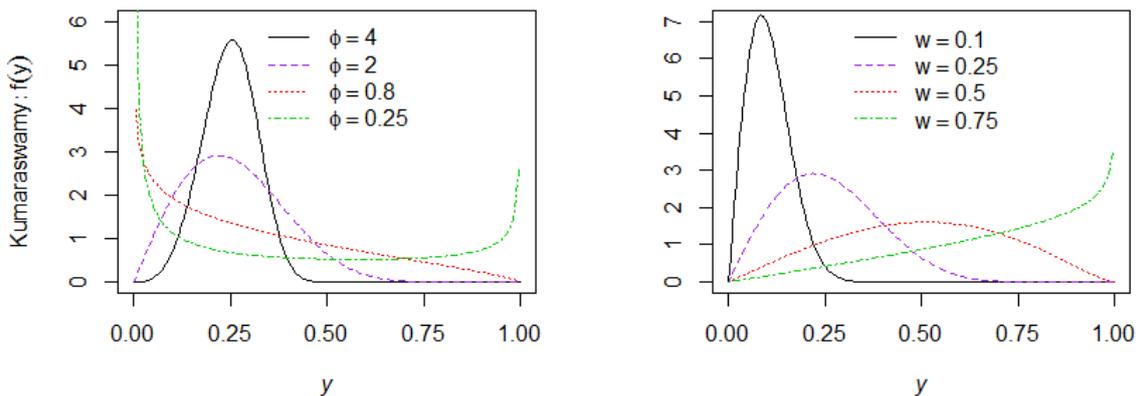
**Proposição 2.** *Sejam  $Y = X^m$  e  $m > 0$ , então  $X \sim Kum(p, q) \Leftrightarrow Y \sim Kum\left(\frac{p}{m}, q\right)$ .*

As demonstrações destas proposições e de outras propriedades da distribuição Kumaraswamy poder ser encontradas em (MITNIK, 2013). Pode-se demonstrar que a distribuição Kumaraswamy possui algumas propriedades básicas que são características da distribuição beta, a saber:

$$\begin{aligned} p > 1, q > 1 &\Rightarrow \text{unimodal}; & p < 1, q < 1 &\Rightarrow \text{antimodal}; \\ p > 1, q \geq 1 &\Rightarrow \text{crescente}; & p \geq 1, q > 1 &\Rightarrow \text{decrecente}; \\ p = q = 1 &\Rightarrow \text{constante}. \end{aligned}$$

Como notado anteriormente, a distribuição Kumaraswamy é bastante flexível por sua densidade assumir diversas formas, dependendo dos valores de seus parâmetros. Algumas densidades Kumaraswamy estão apresentadas na Figura 1. A densidade pode assumir, as formas unimodal, “banheira”, crescente, decrescente e constante.

Figura 1 – Densidades Kumaraswamy



Fonte: O autor (2020).

## 2.2 Modelo de regressão Kumaraswamy

Como notado anteriormente, a distribuição Kumaraswamy é bastante flexível e por este motivo torna-se atrativa para o trabalho de modelagem de dados. Porém, a expressão para a média dada na Equação (2.3) torna inviável uma reparametrização simples baseada na média da distribuição.

Uma solução foi dada por (MITNIK; BAEK, 2013), que, a partir da Equação (2.4), obtiveram uma expressão para a mediana da distribuição:

$$md(Y) = w = (1 - 0.5^{\frac{1}{q}})^{\frac{1}{p}}. \quad (2.5)$$

Tal resultado fornece uma base para a reparametrização da mediana e da dispersão. Ainda a partir da Equação (2.4) podemos expressar a faixa interquantílica como  $IQR = [(1 - 0.75^{\frac{1}{q}})^{\frac{1}{q}} - (1 - 0.25^{\frac{1}{q}})^{\frac{1}{q}}]$ . (MITNIK, 2013) mostrou que o desvio médio absoluto em torno da mediana é dado por  $\delta_2(y) = \int_0^1 |y - w| f(y) dy$ , que pode também ser expresso por  $\delta_2(y) = [2qB(2^{-\frac{1}{q}}, q, 1 + \frac{1}{p}) - qB(q, 1 + \frac{1}{p})]$ , em que  $B(z, p, q) = \int_0^z S^{p-1}(1 - S)^{q-1} dS$  é a função beta incompleta.

Exibiremos agora as reparametrizações apresentadas por (MITNIK; BAEK, 2013) que decorrem de forma imediata do resultado em (2.5), para os parâmetros de forma  $p = \phi$  e  $q$ . Elas são dadas por

$$q = \frac{\log(0.5)}{\log(1 - w^\phi)} \quad (2.6)$$

e

$$\phi = \frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)}. \quad (2.7)$$

Substituindo as expressões em (2.6) e (2.7) em (2.1), (2.2) e (2.4), obtemos duas reparametrizações da distribuição Kumaraswamy. A primeira reparametrização é denotada por  $K(w, \phi)$  e é dada por

$$f(y; w, \phi) = \frac{\phi \log(0.5)}{\log(1 - w^\phi)} y^{\phi-1} (1 - y^\phi)^{\frac{\log(0.5)}{\log(1 - w^\phi)} - 1}, \quad (2.8)$$

$$F(y) = 1 - (1 - y^\phi)^{\frac{\log(0.5)}{\log(1 - w^\phi)}},$$

$$F^{-1}(u) = y = \left[ 1 - (1 - u)^{\frac{\log(1 - w^\phi)}{\log(0.5)}} \right]^{\frac{1}{\phi}}.$$

A segunda reparametrização é denotada por  $K(w, q)$  e é dada por

$$f(y; w, q) = \frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)} q y^{\frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)} - 1} \left( 1 - y^{\frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)}} \right)^{q-1},$$

$$F(y) = 1 - \left( 1 - y^{\frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)}} \right)^q,$$

$$F^{-1}(u) = y = \left[ 1 - (1 - u)^{\frac{1}{q}} \right]^{\frac{\log(w)}{\log(1 - 0.5^{\frac{1}{q}})}}.$$

(MITNIK; BAEK, 2013) mostraram que os parâmetros  $\frac{1}{\phi}$  e  $\frac{1}{q}$  são parâmetros de dispersão. As reparametrizações introduzidas por (MITNIK; BAEK, 2013) possibilitaram o uso da distribuição Kumaraswamy para desenvolver um modelo de regressão que inclui dois submodelos, um para a mediana e outro para a dispersão. Os autores sugerem o uso da primeira reparametrização, por ser mais vantajosa em termo de interpretabilidade dos parâmetros.

Sejam  $Y_1, Y_2, \dots, Y_n$  variáveis aleatórias independentes, em que cada  $Y_t$ ,  $t = 1, 2, \dots, n$ , possui densidade (2.8) com mediana  $w_t$  e parâmetro de precisão  $\phi_t$ . O modelo de regressão Kumaraswamy proposto por (MITNIK; BAEK, 2013) é dado por

$$\begin{aligned} g(w_t) &= \sum_{i=1}^k x_{ti} \beta_i = \eta_{1t}, \\ h(\phi_t) &= \sum_{j=1}^l z_{tj} \gamma_j = \eta_{2t}, \end{aligned} \tag{2.9}$$

em que  $\beta = (\beta_1, \beta_2, \dots, \beta_k)^\top$  e  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_l)^\top$  são vetores de parâmetros desconhecidos ( $\beta \in \mathbb{R}^k$  e  $\gamma \in \mathbb{R}^l$ ),  $x_{t1}, x_{t2}, \dots, x_{tk}$ ,  $z_{t1}, z_{t2}, \dots, z_{tl}$ ,  $t = 1, 2, \dots, n$ , são observações de covariadas fixas e conhecidas ( $k + l < n$ ). As funções  $g(\cdot)$  e  $h(\cdot)$  são funções de ligação, estritamente monótonas e duas vezes diferenciáveis, em que  $g$  tem domínio no intervalo  $(0,1)$  e imagem no conjunto dos números reais, enquanto  $h$  tem domínio no intervalo  $(0, +\infty)$  e imagem no conjunto dos números reais.

Como exemplos de funções de ligação para o modelo da mediana podemos citar: a logito,  $g(w) = \log(w/(1 - w))$ , a probito,  $g(w) = \Phi^{-1}(w)$ , em que  $\Phi(\cdot)$  é a função de distribuição acumulada normal padrão, a log-log,  $g(w) = -\log(-\log(w))$ , dentre outras. Já para o modelo de precisão a função  $h(\phi) = \log(\phi)$  é a ligação mais comumente utilizada.

A função de verossimilhança para o modelo de regressão Kumaraswamy é dada por

$$L(w, \phi) = \prod_{t=1}^n f(y_t; w_t, \phi_t) = \prod_{t=1}^n \frac{\phi_t \log(0.5)}{\log(1 - w_t^{\phi_t})} y_t^{\phi_t - 1} (1 - y_t^{\phi_t})^{\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1}.$$

A função de log-verossimilhança é

$$\ell(\beta, \gamma) = \sum_{t=1}^n \ell_t(w_t, \phi_t), \tag{2.10}$$

em que

$$\begin{aligned}\ell_t(w_t, \phi_t) &= \log(\phi_t) + \log\left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})}\right) + (\phi_t - 1) \log(y_t) \\ &\quad + \left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right) \log(1 - y_t^{\phi_t}).\end{aligned}$$

A seguir obteremos a função escore através da diferenciação da função de log-verossimilhança com respeito aos seus parâmetros, representada por  $U(\beta, \gamma) = (\frac{\partial \ell}{\partial \beta^\top}, \frac{\partial \ell}{\partial \gamma^\top})^\top$ , que é um vetor de dimensão  $(k + l) \times 1$ . A função escore para o parâmetro  $\beta$  é dada por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial \ell_t(w_t, \phi_t)}{\partial w_t} \frac{dw_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \quad i = 1, 2, \dots, k, \quad (2.11)$$

com

$$\frac{\partial \ell_t(w_t, \phi_t)}{\partial w_t} = \frac{\phi_t w_t^{\phi_t}}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} + \frac{\log(0.5) \phi_t w_t^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})}.$$

A Equação (2.11) reduz-se a

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n (w_t^* + y_t^*) \frac{1}{g'(w_t)} x_{ti},$$

em que

$$w_t^* = \frac{\phi_t w_t^{\phi_t}}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} \text{ e } y_t^* = \frac{\log(0.5) \phi_t w_t^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})}. \quad (2.12)$$

Adicionalmente, verificamos no Apêndice A que

$$\mathbb{E}\left(\frac{\partial \ell_t(w_t, \phi_t)}{\partial w_t}\right) = 0 \Rightarrow \mathbb{E}(w_t^* + y_t^*) = 0 \Rightarrow \mathbb{E}(y_t^*) = -w_t^*.$$

Em forma matricial, a função escore para  $\beta$  é um vetor de dimensão  $k \times 1$  que é dado por

$$U_\beta(\beta, \phi) = X^\top T(w^* + y^*),$$

em que  $X$  é a matriz  $n \times k$  de covariadas,  $T = \text{diag}(1/g'(w_1), 1/g'(w_2), \dots, 1/g'(w_n))$ ,  $y^* = (y_1^*, y_2^*, \dots, y_n^*)^\top$  e  $w^* = (w_1^*, w_2^*, \dots, w_n^*)^\top$ .

A  $j$ -ésima componente do vetor escore de  $\gamma$  é

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n \frac{\partial \ell_t(w_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j}, \quad j = 1, 2, \dots, l, \quad (2.13)$$

com

$$\frac{\partial \ell_t(w_t, \phi_t)}{\partial \phi_t} = \frac{1}{\phi_t} + \frac{\log(0.5) w_t^{\phi_t} \log(w_t) \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})} + \log(y_t)$$

$$+ \frac{w_t^{\phi_t} \log(w_t)}{\log(1 - w_t^{\phi_t})(1 - w_t^{\phi_t})} - \frac{\left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right) y_t^{\phi_t} \log(y_t)}{1 - y_t^{\phi_t}}. \quad (2.14)$$

A Expressão em (2.14) pode ser reescrita como

$$\frac{\partial \ell_t(w_t, \phi_t)}{\partial \phi_t} = \sum_{t=1}^n [(y_t^* + w_t^*) \alpha_{1t} + \alpha_{2t}],$$

em que

$$\alpha_{1t} = \frac{w_t \log(w_t)}{\phi_t},$$

$$\alpha_{2t} = \frac{1}{\phi_t} + \log(y_t) - \frac{\left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right) y_t^{\phi_t} \log(y_t)}{1 - y_t^{\phi_t}}.$$

Assim, a Equação (2.13) pode ser reescrita como

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n [(y_t^* + w_t^*) \alpha_{1t} + \alpha_{2t}] \frac{1}{h'(\phi_t)} z_{tj}.$$

Em forma matricial, a função escore para  $\gamma$  é um vetor de dimensão  $l \times 1$  que pode ser expresso como

$$U_\gamma(\beta, \gamma) = Z^\top H[(w^* + y^*) \alpha_1 + \alpha_2],$$

em que  $Z$  é a matriz de covariadas de ordem  $n \times l$ ,  $H = \text{diag}(1/h'(\phi_1), 1/h'(\phi_2), \dots, 1/h'(\phi_n))$ ,  $y^* = (y_1^*, y_2^*, \dots, y_n^*)^\top$  e  $w^* = (w_1^*, w_2^*, \dots, w_n^*)^\top$ ,  $\alpha_1 = (\alpha_{11}, \alpha_{12}, \dots, \alpha_{1n})^\top$  e  $\alpha_2 = (\alpha_{21}, \alpha_{22}, \dots, \alpha_{2n})^\top$ .

Os estimadores de máxima verossimilhança para os parâmetros  $\beta$  e  $\gamma$  são obtidos resolvendo o sistema de equações não lineares dado por  $U = 0$ , que não possui solução em forma fechada. Portanto, utilizam-se algoritmos de otimização não linear, tais como algoritmos de Newton ou quase-Newton, para encontrar as estimativas de máxima verossimilhança através da maximização numérica da função de log-verossimilhança. Para tanto, há a necessidade de especificar valores iniciais para serem utilizados no esquema iterativo. Seguindo (FERRARI; CRIBARI-NETO, 2004), propomos como valores iniciais para  $\beta$  a estimativa de mínimos quadrados ordinários, obtidas de uma regressão linear das respostas transformadas pela função de ligação do modelo,  $g(y) = (g(y_1), g(y_2), \dots, g(y_n))^\top$  em  $X$ , isto é,

$$\beta^{(0)} = (X^\top X)^{-1} X^\top g(y).$$

Por outro lado, de (JONES, 2009) sabemos que o estimador de máxima verossimilhança  $\widehat{\phi}_t$  de  $\phi_t$  é obtido resolvendo a equação

$$f(\phi_t) = \frac{1}{\phi_t} \left( 1 + T_1(\phi_t) + \frac{T_2(\phi_t)}{T_3(\phi_t)} \right) = 0,$$

em que

$$T_1(\phi_t) = \frac{\log(y_t^{\phi_t})}{1 - y_t^{\phi_t}}, \quad T_2(\phi_t) = \frac{y_t^{\phi_t} \log(y_t^{\phi_t})}{1 - y_t^{\phi_t}} \quad \text{e} \quad T_3(\phi_t) = \log(1 - y_t^{\phi_t}).$$

Não é possível expressar a solução da equação acima de forma analítica. Considerando o método iterativo de Newton, temos

$$\phi_t^{(m+1)} = \phi_t^{(m)} - \frac{f(\phi_t^{(m)})}{f'(\phi_t^{(m)})}, \quad m \in \mathbb{N},$$

em que  $m$  indexa as iterações do algoritmo. Tomando  $\phi_t^{(0)} = 1$ , temos

$$\phi_t^{(1)} = 1 - \frac{f(1)}{f'(1)}.$$

Propomos como valores iniciais para  $\gamma$  a estimativa de uma regressão linear sobre  $\phi_t^{(1)}$  transformada pela função de ligação para o modelo de precisão,  $(h(\phi_1^{(1)}), h(\phi_2^{(1)}), \dots, h(\phi_n^{(1)}))^\top$ , isto é,

$$\gamma^{(0)} = (Z^\top Z)^{-1} Z h(\phi^{(1)}).$$

O passo seguinte é a obtenção da expressão da matriz de informação observada. Para tanto, são necessárias as derivadas de segunda ordem do logaritmo da função de verossimilhança em relação aos parâmetros. A partir da Equação (2.11), a derivada de segunda ordem de  $\ell(\beta, \gamma)$  em relação a  $\beta_i$  e  $\beta_j$ , com  $i, j = 1, 2, \dots, k$ , é dada por

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_j \partial \beta_i} &= \sum_{t=1}^n \frac{\partial}{\partial w_t} \left[ \frac{\partial \ell(w_t, \phi_t)}{\partial w_t} \frac{\partial w_t}{\partial \eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \right] \frac{\partial w_t}{\partial \eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_j} \\ &= \sum_{t=1}^n \frac{\partial^2 \ell(w_t, \phi_t)}{\partial w_t^2} \left( \frac{\partial w_t}{\partial \eta_{1t}} \right)^2 \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i} + \sum_{t=1}^n \frac{\partial \ell(w_t, \phi_t)}{\partial w_t} \left[ \frac{\partial^2 w_t}{\partial \eta_{1t}^2} \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i} \right], \end{aligned}$$

em que

$$\begin{aligned} \frac{\partial^2 \ell(w_t, \phi_t)}{\partial w_t^2} &= \frac{w_t^{\phi_t} \phi_t^2}{w_t^2 \log(1 - w_t^{\phi_t})(1 - w_t^{\phi_t})} + \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1 - w_t^{\phi_t})^2 w_t^2 (1 - w_t^{\phi_t})^2} \\ &\quad - \frac{w_t^{\phi_t} \phi_t}{\log(1 - w_t^{\phi_t}) w_t^2 (1 - w_t^{\phi_t})} + \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1 - w_t^{\phi_t}) w_t^2 (1 - w_t^{\phi_t})^2} \\ &\quad + \frac{\log(0.5) w_t^{\phi_t} \phi_t^2 \log(1 - y_t^{\phi_t})}{w^2 \log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})} + \frac{2 \log(0.5) (w_t^{\phi_t})^2 \phi_t^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^3 w_t^2 (1 - w_t^{\phi_t})^2} \\ &\quad - \frac{\log(0.5) w_t^{\phi_t} \phi_t \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t^2 (1 - w_t^{\phi_t})} + \frac{\log(0.5) (w_t^{\phi_t})^2 \phi_t^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t^2 (1 - w_t^{\phi_t})^2}. \end{aligned}$$

Em forma matricial, temos a seguinte expressão:

$$\ddot{L}_{\beta\beta} = X^\top S_1 X + X^\top S X, \quad (2.15)$$

em que  $S_1 = \text{diag} \left\{ \left( \frac{\partial^2 \ell(w_1, \phi_1)}{\partial w_1^2} \left( \frac{\partial w_1}{\partial \eta_{11}} \right)^2, \dots, \frac{\partial^2 \ell(w_n, \phi_n)}{\partial w_n^2} \left( \frac{\partial w_n}{\partial \eta_{1n}} \right)^2 \right) \right\}$  e  $S = \text{diag} \left\{ \left( \frac{\partial \ell(w_1, \phi_1)}{\partial w_1} \frac{\partial^2 w_1}{\partial \eta_{11}^2}, \dots, \frac{\partial \ell(w_n, \phi_n)}{\partial w_n} \frac{\partial^2 w_n}{\partial \eta_{1n}^2} \right) \right\}$ .

A partir da Equação (2.11), a derivada de segunda ordem de  $\ell(\beta, \gamma)$  em relação a  $\beta_i$  e  $\gamma_j$ , com  $i = 1, 2, \dots, k$  e  $j = 1, 2, \dots, l$ , é dada por

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_j \partial \beta_i} &= \sum_{t=1}^n \frac{\partial}{\partial \gamma_j} \left( \frac{\partial \ell(w_t, \phi_t)}{\partial w_t} \frac{\partial w_t}{\partial \eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \right) \\ &= \sum_{t=1}^n \frac{\partial^2 \ell(w_t, \phi_t)}{\partial \phi_t \partial w_t} \frac{\partial \phi_t}{\partial \eta_{2t}} \frac{\partial w_t}{\partial \eta_{1t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} \frac{\partial \eta_{1t}}{\partial \beta_i}, \end{aligned}$$

em que

$$\begin{aligned} \frac{\partial^2 \ell(w_t, \phi_t)}{\partial \phi_t \partial w_t} &= \frac{w_t^{\phi_t} \log(w_t) \phi_t}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} + \frac{w_t^{\phi_t}}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} \\ &+ \frac{(w_t^{\phi_t})^2 \phi_t \log(w_t)}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})^2} + \frac{(w_t^{\phi_t})^2 \phi_t \log(w_t)}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})^2} \\ &+ \frac{\log(0.5) w_t^{\phi_t} \log(w_t) \phi_t \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})} + \frac{\log(0.5) w_t^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})} \\ &- \frac{\log(0.5) w_t^{\phi_t} \phi_t y_t^{\phi_t} \log(y_t)}{(1 - y_t^{\phi_t}) \log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})} \\ &+ \frac{2 \log(0.5) (w_t^{\phi_t})^2 \phi_t \log(1 - y_t^{\phi_t}) \log(w_t)}{\log(1 - w_t^{\phi_t})^3 w_t (1 - w_t^{\phi_t})^2} \\ &+ \frac{\log(0.5) (w_t^{\phi_t})^2 \phi_t \log(1 - y_t^{\phi_t}) \log(w_t)}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})^2}. \end{aligned}$$

Em forma matricial, temos a seguinte expressão:

$$\ddot{L}_{\beta\phi} = X^\top A Z, \quad (2.16)$$

em que  $A = \text{diag} \left\{ \left( \frac{\partial^2 \ell(w_1, \phi_1)}{\partial \phi_1 \partial w_1} \frac{\partial w_1}{\partial \eta_{11}} \frac{\partial \phi_1}{\partial \eta_{21}}, \dots, \frac{\partial^2 \ell(w_n, \phi_n)}{\partial \phi_n \partial w_n} \frac{\partial w_n}{\partial \eta_{1n}} \frac{\partial \phi_n}{\partial \eta_{2n}} \right) \right\}$ .

A partir da Equação (2.13), a derivada de segunda ordem de  $\ell(\beta, \gamma)$  em relação a  $\gamma_i$ ,  $\gamma_j$ , em que  $i, j = 1, 2, \dots, l$ , é da seguinte forma:

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_j \partial \gamma_i} &= \sum_{t=1}^n \frac{\partial}{\partial \gamma_t} \left( \frac{\partial \ell(w_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_i} \right) \\ &= \sum_{t=1}^n \frac{\partial^2 \ell(w_t, \phi_t)}{\partial \phi_t^2} \left( \frac{\partial \phi_t}{\partial \eta_{2t}} \right)^2 \frac{\partial \eta_{2t}}{\partial \gamma_j} \frac{\partial \eta_{2t}}{\partial \gamma_i} + \sum_{t=1}^n \frac{\partial \ell(w_t, \phi_t)}{\partial \phi_t} \frac{\partial^2 \phi_t}{\partial \eta_{2t}^2} \frac{\partial \eta_{2t}}{\partial \gamma_j} \frac{\partial \eta_{2t}}{\partial \gamma_i}, \end{aligned}$$

em que

$$\begin{aligned}
\frac{\partial^2 \ell(w_t, \phi_t)}{\partial \phi_t^2} &= \frac{w_t^{\phi_t} \log(w_t)^2}{\log(1 - w_t^{\phi_t})(1 - w_t^{\phi_t})} - \frac{2 \log(0.5) w_t^{\phi_t} \log(w_t) y_t^{\phi_t} \log(y_t)}{(1 - y_t^{\phi_t}) \log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})} \\
&+ \frac{(w_t^{\phi_t})^2 \log(w_t)^2}{\log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})^2} + \frac{\log(0.5) w_t^{\phi_t} \log(w_t)^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})} \\
&+ \frac{(w_t^{\phi_t})^2 \log(w_t)^2}{\log(1 - w_t^{\phi_t})(1 - w_t^{\phi_t})^2} + \frac{2 \log(0.5) (w_t^{\phi_t})^2 \log(w_t)^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^3 (1 - w_t^{\phi_t})^2} \\
&+ \frac{\log(0.5) (w_t^{\phi_t})^2 \log(w_t)^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})^2} - \frac{1}{\phi_t^2} \\
&- \frac{\left[ \frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1 \right] y_t^{\phi_t} \log(y_t)^2}{1 - y_t^{\phi_t}} - \frac{\left[ \frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1 \right] (y_t^{\phi_t})^2 \log(y_t)^2}{(1 - y_t^{\phi_t})^2}.
\end{aligned}$$

Portanto,

$$\ddot{L}_{\phi\phi} = Z^\top B_1 Z + Z^\top B Z, \quad (2.17)$$

em que  $B_1 = \text{diag} \left\{ \left( \frac{\partial^2 \ell(w_1, \phi_1)}{\partial \phi_1^2} \left( \frac{\partial \phi_1}{\partial \eta_{21}} \right)^2, \dots, \frac{\partial^2 \ell(w_n, \phi_n)}{\partial \phi_n^2} \left( \frac{\partial \phi_n}{\partial \eta_{2n}} \right)^2 \right) \right\}$  e  $B = \text{diag} \left\{ \left( \frac{\partial \ell(w_1, \phi_1)}{\partial \phi_1} \frac{\partial^2 \phi_1}{\partial \eta_{21}^2}, \dots, \frac{\partial \ell(w_n, \phi_n)}{\partial \phi_n} \frac{\partial^2 \phi_n}{\partial \eta_{2n}^2} \right) \right\}$ .

Finalmente, das expressões em (2.15), (2.16) e (2.17) obtemos a matriz de informação observada, que é dada por

$$\ddot{L} = \begin{bmatrix} -\ddot{L}_{\beta\beta} & -\ddot{L}_{\beta\gamma} \\ -\ddot{L}_{\gamma\beta} & -\ddot{L}_{\gamma\gamma} \end{bmatrix} \quad (2.18)$$

$$= \begin{bmatrix} X^\top S_1 X + X^\top S X & X^\top A Z \\ Z^\top A X & Z^\top B_1 Z + Z^\top B Z \end{bmatrix}, \quad (2.19)$$

em que cada uma das matrizes acima já foi definida anteriormente.

## 2.3 Resíduo quantílico

O resíduo quantílico de (DUNN; SMYTH, 1996) pode ser utilizado em modelos de regressão cujas variáveis respostas são independentes e possuem função de distribuição em forma fechada. Sua distribuição é normal. A abordagem utilizada está relacionada à adotada por (COX; SNELL, 1968), mas, enquanto estes se concentraram em correções da média e variância, (DUNN; SMYTH, 1996) concentraram-se na transformação para a normalidade.

Seja  $F(y; \mu_i, \phi)$  a função de distribuição acumulada de uma variável contínua, o resíduo quantílico aleatorizado é definido por

$$r_{q,i} = \Phi^{-1}(F(y_i, \hat{\mu}_i, \hat{\phi}_i)), \quad (2.20)$$

em que  $\Phi(\cdot)$  é a função de distribuição acumulada normal padrão. Um aspecto atrativo do resíduo quantílico reside no fato de sua distribuição ser, para qualquer tamanho amostral,

normal padrão. Esse fato é muito favorável dado que se tem a garantia que o resíduo tem média zero, variância um e que cerca de 95% dos valores dos resíduos devem encontrar-se entre os limites  $-2$  e  $2$ .

Outro fator extremamente importante na análise de resíduos é a identificação de casos individualmente mal ajustados pelo modelo. Assim, um critério de avaliação de desempenho de resíduos é a sua habilidade de identificar tais casos e fazê-lo de forma a quantificar a magnitude desta falta de qualidade de ajuste.

Um ponto mal ajustado é conhecido como um ponto aberrante, que pode afetar de forma desproporcional o ajuste do modelo a ponto de interferir nas conclusões inferenciais sobre seus parâmetros, comprometendo, em especial, a expressão matemática postulada para descrever o fenômeno estudado. Além de investigar a aderência da distribuição empírica do resíduo quantílico aleatorizado à distribuição normal padrão avaliaremos sua habilidade de identificar, com magnitude adequada pontos aberrantes.

## 2.4 Resíduo ponderado

Define-se como resíduo uma medida que quantifica a discrepância entre o modelo ajustado e os dados. A maioria das formulações dos resíduos baseia-se na diferença  $y_t - \widehat{\mathbb{E}}(y_t)$ . No entanto, respeitando o formato da distribuição de probabilidade da variável resposta, é mais interessante pensar em uma função  $r(y_t, \mathbb{E}(y_t))$ , definição geral de resíduos proposta por (COOK, 1986). Sob esse ponto de vista, (FERRARI; CRIBARI-NETO, 2004) e (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) propuseram resíduos para o modelo de regressão beta. Nós adaptaremos o resíduo ponderado proposto por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) para o modelo de regressão Kumaraswamy.

A ideia proposta por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) de resíduos para o modelo de regressão beta foi baseada no processo iterativo Scoring de Fisher. Para o modelo de regressão Kumaraswamy, o processo iterativo scoring de Fisher para  $\beta$  é

$$\beta^{(m+1)} = \beta^{(m)} + \left(\kappa_{\beta\beta}^{(m)}\right)^{-1} U_{\beta}(\beta, \phi), \quad m = 1, 2, \dots,$$

em que  $m$  indexa as iterações do algoritmo. Usando as expressões para o vetor score e para  $\kappa_{\beta\beta}$  na expressão acima, obtemos a seguinte expressão para o método iterativo:

$$\beta^{m+1} = \beta^{(m)} + (X^{\top} W_{\beta\beta}^{(m)} X)^{-1} X^{\top} T^{(m)}(w^* + y^*)^{(m)},$$

em que a matriz  $\kappa_{\beta\beta}$  está definida no Apêndice **A**.

O processo iterativo acima pode ser representado como um processo de mínimos quadrados, ou seja,

$$\beta^{m+1} = (X^{\top} W_{\beta\beta}^{(m)} X)^{-1} X^{\top} W_{\beta\beta}^{(m)} a_1^{(m)},$$

em que  $a_1^{(m)} = \eta^{(m)} + (W_{\beta\beta}^{(m)})^{-1}T^{(m)}(w^* + y^*)^{(m)}$  com  $\eta^{(m)} = X\beta$ .

Após convergência, temos que

$$\hat{\beta} = (X^\top \widehat{W}_{\beta\beta} X)^{-1} X^\top \widehat{W}_{\beta\beta} \hat{a}_1, \quad (2.21)$$

com

$$\hat{a}_1 = \hat{\eta} + \widehat{W}_{\beta\beta}^{-1} \widehat{T}(\hat{w}^* + \hat{y}^*),$$

ou seja,  $\hat{\beta}$  pode ser visto como o estimador de mínimos quadrados da regressão linear de  $W_{\beta\beta}^{1/2} \hat{a}_1$  em  $W_{\beta\beta}^{1/2} X$ .

O resíduo de mínimos quadrados desta regressão é dado por

$$\begin{aligned} r^\beta &= \widehat{W}_{\beta\beta}^{1/2} \hat{a}_1 - \widehat{W}_{\beta\beta}^{1/2} X \hat{\beta} = \widehat{W}_{\beta\beta}^{1/2} (\hat{a}_1 - \hat{\eta}) \\ &= \widehat{W}_{\beta\beta}^{1/2} (\hat{\eta} + \widehat{W}_{\beta\beta}^{-1} \widehat{T}(\hat{w}^* + \hat{y}^*) - \hat{\eta}) \\ &= \widehat{W}_{\beta\beta}^{1/2} (\widehat{W}_{\beta\beta}^{-1} \widehat{T}(\hat{w}^* + \hat{y}^*)) \\ &= \widehat{W}_{\beta\beta}^{-1/2} \widehat{T}(\hat{w}^* + \hat{y}^*). \end{aligned}$$

Assim, o  $t$ -ésimo resíduo pode ser escrito como

$$r_t^\beta = \frac{1}{\widehat{W}_{\beta\beta(t,t)}^{1/2}} \frac{1}{g'(\hat{w}_t)} (\hat{w}_t^* + \hat{y}_t^*).$$

Sabemos que

$$\widehat{W}_{\beta\beta(t,t)} = \frac{(\hat{w}_t^\phi)^2 \hat{\phi}^2}{\log(1 - \hat{w}_t^\phi)^2 (\hat{w}_t)^2 (1 - \hat{w}_t^\phi)^2} \frac{1}{g'(\hat{w}_t)^2}.$$

Então,

$$\frac{1}{\widehat{W}_{\beta\beta(t,t)}^{1/2}} = \frac{\log(1 - \hat{w}_t^\phi) (\hat{w}_t) (1 - \hat{w}_t^\phi) g'(\hat{w}_t)}{\hat{w}_t^\phi \hat{\phi}}.$$

Portanto, o  $t$ -ésimo resíduo ponderado é dado por

$$r_t^\beta = \frac{1}{\widehat{W}_{\beta\beta(t,t)}^{1/2}} \frac{1}{g'(\hat{w}_t)} (\hat{w}_t^* + \hat{y}_t^*) = \frac{(\hat{w}_t^* + \hat{y}_t^*) (\log(1 - \hat{w}_t^\phi) (\hat{w}_t) (1 - \hat{w}_t^\phi))}{\hat{w}_t^\phi \hat{\phi}}. \quad (2.22)$$

## 2.5 Resíduo ponderado padronizado

Uma possível padronização do resíduo dado em (2.22) baseia-se na variância de  $\hat{a}_1$ . Considerando que  $\text{Cov}(\hat{\beta}) \approx (X^\top \widehat{W} X)^{-1}$  e  $\widehat{W} \approx W$  e reescrevendo a Equação (2.21) como

$$(X^\top \widehat{W} X) \hat{\beta} = X^\top \widehat{W} \hat{a}_1,$$

tem-se que

$$\begin{aligned}
 (X^\top W X) \text{Cov}(\widehat{\beta})(X^\top W X)^\top &\approx (X^\top W) \text{Cov}(\widehat{a}_1)(X^\top W)^\top \\
 (X^\top W X)(X^\top W X)^{-1}(X^\top W X) &\approx (X^\top W) \text{Cov}(\widehat{a}_1)(X^\top W)^\top \\
 (X^\top W) \text{Cov}(\widehat{a}_1)(W^\top X) &\approx (X^\top W X) \\
 (X X^\top) W \text{Cov}(\widehat{a}_1) W^\top (X X^\top) &\approx (X X^\top) W (X X^\top) \\
 \text{Cov}(\widehat{a}_1) &\approx W^{-1}.
 \end{aligned}$$

Logo,  $\widehat{\text{Cov}}(\widehat{a}_1) \approx \widehat{W}^{-1}$ . Em seguida, expressamos os resíduos como

$$\begin{aligned}
 r^\beta &= \widehat{W}_{\beta\beta}^{1/2} \widehat{a}_1 - \widehat{W}_{\beta\beta}^{1/2} X \widehat{\beta} = \widehat{W}_{\beta\beta}^{1/2} (\widehat{a}_1 - X (X^\top \widehat{W}_{\beta\beta} X)^{-1} X^\top \widehat{W}_{\beta\beta} \widehat{a}_1) \\
 &= (\widehat{W}_{\beta\beta}^{1/2} - \widehat{W}_{\beta\beta}^{1/2} X (X^\top \widehat{W}_{\beta\beta} X)^{-1} X^\top \widehat{W}_{\beta\beta}) \widehat{a}_1.
 \end{aligned}$$

Considerando que  $\text{Cov}(\widehat{a}_1) \approx W^{-1}$ , segue que

$$\begin{aligned}
 \text{Cov}(r^\beta) &\approx (W_{\beta\beta}^{1/2} - W_{\beta\beta}^{1/2} X (X^\top W_{\beta\beta} X)^{-1} X^\top W_{\beta\beta}) \text{Cov}(\widehat{a}_1) \\
 &\quad \times (W_{\beta\beta}^{1/2} - W_{\beta\beta}^{1/2} X (X^\top W_{\beta\beta} X)^{-1} X^\top W_{\beta\beta})^\top \\
 &\approx (I - W_{\beta\beta}^{1/2} X (X^\top W_{\beta\beta} X)^{-1} X^\top W_{\beta\beta}^{1/2}) \\
 &\approx (I - H).
 \end{aligned}$$

Logo,  $\widehat{\text{Cov}}(r^\beta) \approx (I - H)$ , com  $H = \widehat{W}_{\beta\beta}^{1/2} X (X^\top \widehat{W}_{\beta\beta} X)^{-1} X^\top \widehat{W}_{\beta\beta}$ , em que  $H$  é a matriz de projeção de  $W_{\beta\beta}^{1/2} a_1$  em  $W_{\beta\beta}^{1/2} X$ , que é simétrica e idempotente.

Assim, o resíduo ponderado padronizado é definido como

$$r_{pt}^\beta = \frac{r_t^\beta}{\sqrt{\widehat{\text{Var}}(r_t^\beta)}} = \frac{r_t^\beta}{\sqrt{1 - h_{tt}}}, \quad (2.23)$$

em que  $h_{tt}$  é o  $t$ -ésimo elemento diagonal de  $H$ .

## 2.6 Avaliação numérica

Definidos os resíduos ponderado e ponderado padronizado para o modelo de regressão Kumaraswamy é importante avaliar seus comportamentos, isto é, avaliar as distribuições empíricas dos resíduos definidos em (2.22) e (2.23). Com essa finalidade, realizamos simulações de Monte Carlo com 5000 réplicas sob diferentes cenários. As simulações foram realizadas utilizando a linguagem de programação matricial **Ox**, disponibilizada de forma gratuita para uso acadêmico em <http://www.doornik.com/> atualmente na versão 8.02. Consideramos o modelo de regressão Kumaraswamy com precisão constante dado por

$$\log\left(\frac{w_t}{1 - w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, \quad t = 1, 2, \dots, n. \quad (2.24)$$

Os valores das covariadas  $x_{ti}$ , com  $i = 2, 3, 4, 5$ , foram obtidos de forma independente da distribuição uniforme e mantidos fixos para cada réplica. Para este modelo consideramos cenários onde as medianas da variável resposta pertencem aos intervalos  $(0.0258, 0.3452)$ ,  $(0.2440, 0.8237)$  e  $(0.8855, 0.9983)$ . Para cada um destes intervalos consideramos cenários diferentes para a precisão.

Nas Tabelas 1, 2 e 3 são apresentados valores das médias, desvios-padrão, assimetrias, curtoses dos 5000 resíduos quantílico, ponderado e ponderado padronizado dados respectivamente em (2.20), (2.22) e (2.23), considerando  $n = 20$  observações. Para este cenário, as medianas da variável resposta pertencem ao intervalo  $(0.0258, 0.3452)$  e variamos o parâmetro de precisão  $\phi = 2, 4$  e  $7$ . Os valores verdadeiros dos parâmetros da regressão são  $\beta_1 = -2.0$ ,  $\beta_2 = 2.0$ ,  $\beta_3 = -1.8$ ,  $\beta_4 = 1.3$  e  $\beta_5 = 1.0$ .

Observamos nas Tabelas 1, 2 e 3 que os resíduos ponderado e ponderado padronizado possuem médias próximas de zero e desvios-padrão próximos de um. Adicionalmente, observamos assimetria negativa para os três valores de  $\phi$ . Os valores da curtose dos resíduos propostos estão muito próximos entre si, porém longe do valor de referência da distribuição normal.

Tabela 1 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log(\frac{wt}{1-wt}) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.0258, 0.3452)$  e  $\phi = 2$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.080	0.027	0.029	1.070	1.000	1.061	-0.013	-1.535	-1.535	2.540	5.220	5.217
2	-0.023	0.020	0.023	1.016	0.880	1.022	-0.292	-1.102	-1.101	2.626	3.651	3.644
3	-0.062	0.020	0.022	1.063	0.953	1.057	-0.125	-1.286	-1.285	2.430	4.221	4.211
4	0.056	-0.008	-0.010	0.948	0.794	1.005	-0.486	-0.696	-0.696	2.679	2.560	2.560
5	-0.019	-0.030	-0.032	1.081	1.027	1.092	-0.082	-1.380	-1.380	2.560	4.518	4.517
6	-0.038	0.043	0.050	1.003	0.855	1.011	-0.318	-1.137	-1.137	2.716	3.851	3.850
7	0.006	-0.064	-0.069	1.098	1.060	1.130	-0.095	-1.381	-1.380	2.553	4.558	4.554
8	0.021	-0.027	-0.031	1.028	0.915	1.064	-0.291	-1.031	-1.028	2.588	3.386	3.373
9	0.079	0.041	0.063	0.823	0.640	0.980	-0.811	-0.452	-0.454	3.404	2.422	2.422
10	-0.039	-0.022	-0.023	1.097	1.035	1.102	-0.084	-1.444	-1.444	2.561	4.815	4.815
11	-0.037	0.003	0.003	1.060	0.955	1.041	-0.200	-1.327	-1.322	2.630	4.520	4.486
12	0.064	-0.022	-0.028	0.962	0.815	1.036	-0.486	-0.765	-0.763	2.720	2.768	2.764
13	-0.005	-0.046	-0.050	1.088	1.039	1.120	-0.102	-1.375	-1.374	2.596	4.484	4.482
14	0.019	-0.018	-0.022	1.016	0.905	1.058	-0.281	-1.077	-1.077	2.564	3.630	3.626
15	0.015	0.001	0.001	0.995	0.842	1.022	-0.420	-0.875	-0.875	2.718	2.970	2.969
16	-0.018	0.063	0.080	0.943	0.769	0.982	-0.472	-0.953	-0.955	2.932	3.364	3.363
17	0.001	-0.079	-0.083	1.123	1.116	1.168	-0.016	-1.478	-1.478	2.516	4.919	4.918
18	-0.072	0.080	0.097	0.986	0.827	1.007	-0.262	-1.146	-1.145	2.588	3.845	3.835
19	0.038	-0.026	-0.030	1.006	0.871	1.016	-0.384	-0.921	-0.918	2.597	3.195	3.183
20	-0.027	0.046	0.054	0.975	0.839	0.980	-0.262	-1.083	-1.082	2.559	3.623	3.621

Fonte: O autor (2020).

Nas Tabelas 4, 5 e 6 são apresentados resultados obtidos para casos em que as medianas da variável resposta pertencem ao intervalo  $(0.2440, 0.8237)$  e variamos o parâmetro de precisão  $\phi = 3, 10$  e  $20$ . Os valores verdadeiros dos parâmetros de regressão

Tabela 2 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log(\frac{w_t}{1-w_t}) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.0258, 0.3452)$  e  $\phi = 4$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.077	0.076	0.088	0.994	0.854	0.991	-0.186	-1.235	-1.236	2.537	4.128	4.131
2	0.010	0.005	0.006	0.993	0.851	1.016	-0.370	-0.962	-0.963	2.667	3.254	3.259
3	-0.057	0.039	0.045	1.030	0.890	1.031	-0.205	-1.132	-1.132	2.417	3.673	3.670
4	-0.009	0.022	0.027	0.991	0.847	1.007	-0.317	-0.951	-0.951	2.544	3.120	3.119
5	0.004	-0.056	-0.059	1.087	1.052	1.118	-0.090	-1.393	-1.393	2.607	4.550	4.550
6	-0.013	-0.014	-0.016	1.057	0.951	1.060	-0.248	-1.232	-1.232	2.671	4.107	4.107
7	0.041	-0.089	-0.096	1.089	1.052	1.132	-0.145	-1.277	-1.276	2.550	4.150	4.148
8	0.019	-0.020	-0.024	1.021	0.899	1.056	-0.314	-0.987	-0.988	2.593	3.253	3.252
9	0.019	0.045	0.060	0.914	0.732	0.970	-0.573	-0.733	-0.733	2.921	2.829	2.828
10	-0.002	-0.050	-0.053	1.088	1.040	1.110	-0.099	-1.372	-1.372	2.537	4.503	4.503
11	-0.023	0.057	0.071	0.959	0.783	0.977	-0.441	-0.928	-0.926	2.794	3.214	3.195
12	-0.015	0.009	0.010	1.024	0.887	1.036	-0.308	-1.055	-1.055	2.607	3.440	3.440
13	-0.016	-0.045	-0.048	1.099	1.061	1.127	-0.069	-1.446	-1.446	2.608	4.773	4.772
14	0.044	-0.070	-0.078	1.056	0.991	1.107	-0.189	-1.155	-1.155	2.482	3.757	3.756
15	-0.015	0.022	0.026	1.005	0.850	1.010	-0.376	-0.994	-0.994	2.721	3.350	3.350
16	-0.019	0.068	0.087	0.935	0.759	0.975	-0.476	-0.937	-0.938	2.912	3.332	3.333
17	0.025	-0.097	-0.103	1.118	1.107	1.169	-0.052	-1.370	-1.370	2.473	4.417	4.417
18	-0.067	0.059	0.069	1.010	0.874	1.023	-0.202	-1.229	-1.229	2.558	4.093	4.093
19	0.011	0.025	0.031	0.960	0.794	0.981	-0.446	-0.829	-0.829	2.667	2.970	2.969
20	-0.004	0.032	0.038	0.967	0.834	0.989	-0.311	-1.049	-1.049	2.669	3.526	3.525

Fonte: O autor (2020).

Tabela 3 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log(\frac{w_t}{1-w_t}) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.0258, 0.3452)$  e  $\phi = 7$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.065	0.056	0.065	1.013	0.886	1.027	-0.206	-1.304	-1.304	2.669	4.378	4.377
2	0.015	-0.004	-0.005	1.004	0.867	1.034	-0.363	-0.982	-0.983	2.652	3.322	3.323
3	-0.073	0.057	0.066	1.024	0.875	1.015	-0.232	-1.193	-1.192	2.547	3.987	3.983
4	-0.022	0.039	0.046	0.988	0.823	0.979	-0.386	-0.952	-0.951	2.682	3.170	3.169
5	0.017	-0.072	-0.077	1.092	1.076	1.143	-0.053	-1.434	-1.434	2.552	4.770	4.770
6	-0.003	-0.026	-0.029	1.062	0.964	1.074	-0.240	-1.256	-1.256	2.635	4.233	4.233
7	0.063	-0.106	-0.114	1.083	1.058	1.138	-0.154	-1.272	-1.272	2.578	4.143	4.143
8	-0.006	0.003	0.003	1.021	0.883	1.036	-0.316	-1.017	-1.017	2.585	3.369	3.368
9	0.011	0.050	0.067	0.920	0.734	0.972	-0.570	-0.741	-0.740	2.924	2.791	2.789
10	-0.031	-0.032	-0.034	1.099	1.055	1.126	-0.048	-1.470	-1.470	2.571	4.898	4.898
11	-0.041	0.070	0.088	0.961	0.788	0.985	-0.386	-0.976	-0.975	2.711	3.303	3.302
12	0.027	-0.009	-0.011	0.990	0.875	1.022	-0.314	-1.021	-1.021	2.663	3.414	3.414
13	0.022	-0.083	-0.088	1.100	1.085	1.153	-0.062	-1.389	-1.389	2.539	4.501	4.501
14	0.035	-0.066	-0.074	1.064	0.999	1.115	-0.186	-1.178	-1.178	2.510	3.815	3.814
15	-0.039	0.046	0.055	0.998	0.842	0.999	-0.342	-1.064	-1.065	2.733	3.543	3.544
16	-0.019	0.089	0.114	0.895	0.722	0.927	-0.498	-0.968	-0.968	3.032	3.550	3.551
17	0.012	-0.093	-0.099	1.127	1.137	1.201	-0.005	-1.479	-1.479	2.567	4.805	4.805
18	-0.063	0.053	0.062	1.014	0.888	1.039	-0.189	-1.280	-1.280	2.606	4.296	4.296
19	0.034	0.013	0.016	0.946	0.794	0.980	-0.458	-0.841	-0.840	2.764	3.031	3.030
20	-0.010	0.035	0.042	0.968	0.838	0.993	-0.290	-1.077	-1.077	2.654	3.641	3.642

Fonte: O autor (2020).

são  $\beta_1 = 1.1$ ,  $\beta_2 = 1.0$ ,  $\beta_3 = -1.0$ ,  $\beta_4 = 1.4$  e  $\beta_5 = -1.2$ .

Tabela 4 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log(\frac{w_t}{1-w_t}) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.2440, 0.8237)$  e  $\phi = 3$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.075	0.045	0.048	1.037	0.947	1.016	-0.071	-1.514	-1.512	2.570	5.331	5.310
2	-0.003	0.007	0.007	1.012	0.868	1.009	-0.366	-1.040	-1.036	2.691	3.569	3.537
3	-0.059	0.028	0.030	1.051	0.919	1.030	-0.199	-1.236	-1.231	2.473	4.183	4.142
4	0.050	-0.009	-0.012	0.959	0.807	1.016	-0.450	-0.709	-0.709	2.584	2.559	2.554
5	-0.031	-0.021	-0.022	1.083	1.022	1.086	-0.078	-1.394	-1.394	2.523	4.640	4.639
6	-0.027	0.030	0.034	1.008	0.860	1.004	-0.335	-1.100	-1.095	2.703	3.734	3.704
7	-0.017	-0.035	-0.037	1.089	1.026	1.093	-0.127	-1.406	-1.406	2.603	4.812	4.805
8	0.006	-0.012	-0.014	1.026	0.903	1.050	-0.294	-1.040	-1.033	2.572	3.463	3.420
9	0.076	0.036	0.052	0.841	0.649	0.978	-0.830	-0.459	-0.459	3.408	2.452	2.437
10	-0.044	-0.013	-0.013	1.092	1.011	1.075	-0.119	-1.432	-1.432	2.537	4.921	4.921
11	-0.020	0.025	0.027	1.007	0.866	0.983	-0.352	-1.186	-1.167	2.797	4.219	4.094
12	0.024	0.003	0.005	0.979	0.819	1.029	-0.452	-0.838	-0.829	2.673	2.939	2.916
13	-0.013	-0.050	-0.053	1.106	1.046	1.124	-0.120	-1.350	-1.350	2.543	4.425	4.422
14	0.017	-0.024	-0.028	1.027	0.913	1.054	-0.275	-1.058	-1.055	2.493	3.589	3.566
15	0.000	0.005	0.005	1.012	0.851	1.031	-0.394	-0.877	-0.871	2.607	2.958	2.940
16	-0.013	0.052	0.063	0.954	0.778	0.988	-0.472	-0.931	-0.928	2.834	3.358	3.324
17	-0.010	-0.058	-0.061	1.110	1.082	1.133	-0.052	-1.521	-1.521	2.552	5.324	5.323
18	-0.039	0.050	0.060	0.988	0.844	1.016	-0.282	-1.140	-1.128	2.606	3.899	3.839
19	0.026	-0.016	-0.020	1.009	0.859	1.007	-0.405	-0.895	-0.887	2.549	3.191	3.152
20	-0.009	0.024	0.027	0.987	0.852	0.992	-0.301	-1.039	-1.036	2.587	3.499	3.481

Fonte: O autor (2020).

Tabela 5 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log(\frac{w_t}{1-w_t}) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.2440, 0.8237)$  e  $\phi = 10$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.080	0.027	0.029	1.070	1.000	1.061	-0.013	-1.535	-1.535	2.540	5.220	5.217
2	-0.023	0.020	0.023	1.016	0.880	1.022	-0.292	-1.102	-1.101	2.626	3.651	3.644
3	-0.062	0.020	0.022	1.063	0.953	1.057	-0.125	-1.286	-1.285	2.430	4.221	4.211
4	0.056	-0.008	-0.010	0.948	0.794	1.005	-0.486	-0.696	-0.696	2.679	2.560	2.560
5	-0.019	-0.030	-0.032	1.081	1.027	1.092	-0.082	-1.380	-1.380	2.560	4.518	4.517
6	-0.038	0.043	0.050	1.003	0.855	1.011	-0.318	-1.137	-1.137	2.716	3.851	3.850
7	0.006	-0.064	-0.069	1.098	1.060	1.130	-0.095	-1.381	-1.380	2.553	4.558	4.554
8	0.021	-0.027	-0.031	1.028	0.915	1.064	-0.291	-1.031	-1.028	2.588	3.386	3.373
9	0.079	0.041	0.063	0.823	0.640	0.980	-0.811	-0.452	-0.454	3.404	2.422	2.422
10	-0.039	-0.022	-0.023	1.097	1.035	1.102	-0.084	-1.444	-1.444	2.561	4.815	4.815
11	-0.037	0.003	0.003	1.060	0.955	1.041	-0.200	-1.327	-1.322	2.630	4.520	4.486
12	0.064	-0.022	-0.028	0.962	0.815	1.036	-0.486	-0.765	-0.763	2.720	2.768	2.764
13	-0.005	-0.046	-0.050	1.088	1.039	1.120	-0.102	-1.375	-1.374	2.596	4.484	4.482
14	0.019	-0.018	-0.022	1.016	0.905	1.058	-0.281	-1.077	-1.077	2.564	3.630	3.626
15	0.015	0.001	0.001	0.995	0.842	1.022	-0.420	-0.875	-0.875	2.718	2.970	2.969
16	-0.018	0.063	0.080	0.943	0.769	0.982	-0.476	-0.953	-0.955	2.932	3.364	3.363
17	0.001	-0.079	-0.083	1.123	1.116	1.168	-0.016	-1.478	-1.478	2.516	4.919	4.918
18	-0.072	0.080	0.097	0.986	0.827	1.007	-0.262	-1.146	-1.145	2.588	3.845	3.835
19	0.038	-0.026	-0.030	1.006	0.871	1.016	-0.384	-0.921	-0.918	2.597	3.195	3.183
20	-0.027	0.046	0.054	0.975	0.839	0.980	-0.262	-1.083	-1.082	2.559	3.623	3.621

Fonte: O autor (2020).

Tabela 6 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.2440, 0.8237)$  e  $\phi = 20$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.074	0.013	0.014	1.086	1.017	1.081	-0.048	-1.546	-1.547	2.605	5.322	5.328
2	-0.028	0.021	0.025	1.019	0.888	1.033	-0.258	-1.109	-1.108	2.576	3.615	3.616
3	-0.068	0.038	0.043	1.043	0.937	1.041	-0.136	-1.396	-1.396	2.589	4.680	4.683
4	0.041	0.013	0.017	0.938	0.765	0.967	-0.564	-0.723	-0.722	2.883	2.699	2.698
5	0.007	-0.062	-0.066	1.091	1.061	1.128	-0.059	-1.398	-1.399	2.506	4.614	4.615
6	0.011	-0.012	-0.014	1.018	0.912	1.077	-0.274	-1.158	-1.158	2.655	3.877	3.875
7	0.004	-0.068	-0.073	1.104	1.081	1.153	-0.056	-1.408	-1.408	2.557	4.573	4.573
8	0.016	-0.023	-0.026	1.025	0.918	1.068	-0.251	-1.057	-1.058	2.503	3.428	3.433
9	0.084	0.038	0.057	0.818	0.647	0.987	-0.760	-0.489	-0.488	3.356	2.465	2.465
10	-0.053	-0.008	-0.008	1.092	1.025	1.092	-0.058	-1.453	-1.453	2.499	4.845	4.845
11	-0.075	0.024	0.026	1.073	0.980	1.068	-0.080	-1.395	-1.394	2.533	4.575	4.564
12	0.076	-0.009	-0.012	0.918	0.781	0.989	-0.506	-0.771	-0.772	2.855	2.861	2.860
13	0.042	-0.093	-0.100	1.094	1.065	1.146	-0.140	-1.353	-1.352	2.588	4.499	4.498
14	0.010	-0.010	-0.012	1.016	0.905	1.058	-0.267	-1.089	-1.088	2.597	3.587	3.585
15	0.019	0.007	0.009	0.974	0.836	1.014	-0.364	-0.909	-0.910	2.624	3.083	3.084
16	-0.004	0.057	0.073	0.927	0.766	0.979	-0.451	-0.938	-0.939	2.894	3.290	3.294
17	-0.035	-0.042	-0.044	1.116	1.100	1.152	0.005	-1.616	-1.616	2.592	5.547	5.548
18	-0.097	0.100	0.122	0.985	0.818	0.996	-0.236	-1.177	-1.176	2.557	3.961	3.960
19	0.026	-0.013	-0.015	1.002	0.865	1.010	-0.378	-0.952	-0.951	2.631	3.271	3.268
20	-0.035	0.041	0.048	0.993	0.866	1.012	-0.231	-1.151	-1.150	2.601	3.781	3.780

Fonte: O autor (2020).

Das Tabelas 4, 5 e 6 observamos que as médias estão próximas de zero e que os desvios-padrão estão próximos de um. Quanto à assimetria, observamos que os resíduos ponderado e ponderado padronizado apresentam assimetria negativa nos três cenários. As curtoses dos resíduos são semelhantes, porém seus valores não estão próximo do valor de referência da distribuição normal.

Nas Tabelas 7, 8 e 9 apresentamos os resultados referentes às distribuições empíricas dos resíduos quando as medianas da variável resposta pertencem ao intervalo (0.8855, 0.9983) enquanto variamos  $\phi = 3, 50$  e  $150$ . Os valores verdadeiros dos parâmetros de regressão são  $\beta_1 = 4.6$ ,  $\beta_2 = -2.6$ ,  $\beta_3 = -3.1$ ,  $\beta_4 = 2.6$  e  $\beta_5 = 3.0$ .

Assim como anteriormente, as médias estão próximas de zero e os desvios-padrão estão próximos de um para os resíduos propostos. A assimetria é negativa em todos os cenários e as curtoses dos resíduos propostos não estão próximo do valor de referência da distribuição normal.

Como visto nas Tabelas 1 a 9, os resíduos ponderado e ponderado padronizado têm comportamento bem semelhante, com médias próximas de zero, desvios-padrão próximos de um, assimetrias negativas e curtoses que não estão próximas do valor de referência da distribuição normal. A aderência da distribuição dos resíduos à distribuição normal não é, assim, satisfatória.

Com o objetivo de comparar os quantis empíricos dos resíduos com os quantis

Tabela 7 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log(\frac{w_t}{1-w_t}) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.8855, 0.9983)$  e  $\phi = 3$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.009	0.049	0.087	0.957	0.788	0.967	-0.516	-1.190	-0.994	3.052	4.589	4.068
2	-0.083	0.056	0.054	1.053	0.871	0.951	-0.382	-1.288	-1.297	2.783	4.625	4.662
3	0.092	-0.009	0.021	0.899	0.739	1.057	-0.717	-0.863	-0.537	3.305	3.585	3.035
4	-0.023	0.020	0.022	1.026	0.854	0.974	-0.434	-1.065	-1.010	2.751	3.785	3.684
5	-0.160	0.033	0.033	1.174	1.038	1.102	-0.066	-1.478	-1.489	2.327	5.057	5.104
6	-0.073	0.046	0.047	1.057	0.866	0.953	-0.407	-1.165	-1.132	2.724	4.122	4.067
7	-0.102	0.031	0.030	1.114	0.964	1.014	-0.238	-1.429	-1.440	2.604	5.085	5.132
8	0.069	0.037	0.076	0.861	0.655	1.012	-0.941	-0.698	-0.456	3.811	3.369	2.769
9	0.133	0.004	0.012	0.807	0.627	0.989	-1.100	-0.621	-0.280	4.561	3.385	3.014
10	-0.112	-0.007	-0.009	1.175	1.061	1.126	-0.100	-1.459	-1.468	2.353	4.935	4.980
11	0.052	0.033	0.049	0.891	0.721	0.956	-0.742	-1.045	-0.741	3.644	4.442	3.938
12	-0.101	0.084	0.094	1.032	0.830	0.919	-0.407	-1.217	-1.172	2.791	4.290	4.233
13	-0.096	0.020	0.019	1.118	0.977	1.041	-0.199	-1.392	-1.403	2.483	4.831	4.863
14	-0.125	0.096	0.100	1.043	0.844	0.891	-0.349	-1.306	-1.314	2.738	4.674	4.705
15	-0.068	0.048	0.054	1.050	0.858	0.943	-0.456	-1.259	-1.224	2.882	4.740	4.632
16	0.022	0.010	0.009	0.980	0.805	0.992	-0.585	-1.036	-0.835	3.069	4.043	3.700
17	-0.108	0.005	0.004	1.155	1.013	1.078	-0.164	-1.345	-1.353	2.364	4.538	4.573
18	0.087	0.000	0.030	0.895	0.728	1.024	-0.810	-1.015	-0.736	3.721	4.417	3.468
19	0.038	0.016	0.028	0.946	0.753	1.026	-0.685	-0.835	-0.629	3.140	3.425	2.981
20	-0.036	0.011	0.006	1.057	0.896	0.993	-0.367	-1.147	-1.136	2.666	4.009	3.951

Fonte: O autor (2020).

Tabela 8 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log(\frac{w_t}{1-w_t}) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.8855, 0.9983)$  e  $\phi = 50$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.099	0.069	0.085	1.042	0.881	1.034	-0.218	-1.227	-1.186	2.505	4.130	3.969
2	-0.086	0.025	0.025	1.093	1.006	1.032	-0.133	-1.715	-1.717	2.782	6.408	6.418
3	0.107	0.026	0.069	0.801	0.644	1.108	-0.823	-0.595	-0.437	3.721	2.744	2.521
4	-0.051	-0.006	-0.008	1.093	1.014	1.065	-0.144	-1.578	-1.575	2.674	5.890	5.856
5	-0.066	-0.055	-0.060	1.178	1.116	1.201	-0.028	-1.307	-1.311	2.326	3.961	3.976
6	-0.087	0.016	0.016	1.109	1.013	1.044	-0.151	-1.689	-1.689	2.729	6.591	6.591
7	-0.097	0.021	0.022	1.113	1.014	1.054	-0.140	-1.644	-1.646	2.716	6.319	6.323
8	0.242	0.000	0.007	0.550	0.442	1.109	-1.522	-0.003	0.001	7.126	2.502	2.274
9	0.132	0.014	0.025	0.790	0.592	1.034	-1.167	-0.284	-0.141	4.506	2.487	2.421
10	-0.053	-0.076	-0.081	1.195	1.126	1.205	-0.070	-1.293	-1.293	2.296	4.022	4.024
11	-0.060	0.006	0.006	1.090	0.990	1.043	-0.200	-1.580	-1.567	2.741	5.956	5.861
12	-0.102	0.030	0.031	1.108	1.007	1.050	-0.138	-1.662	-1.659	2.730	6.087	6.066
13	-0.104	0.018	0.018	1.126	1.010	1.092	-0.119	-1.400	-1.399	2.510	4.640	4.625
14	-0.088	0.023	0.023	1.096	1.019	1.042	-0.106	-1.803	-1.804	2.771	7.189	7.189
15	-0.073	0.015	0.015	1.094	1.000	1.034	-0.188	-1.736	-1.734	2.818	6.952	6.936
16	-0.120	0.077	0.083	1.060	0.908	1.004	-0.218	-1.473	-1.450	2.729	5.274	5.149
17	-0.054	-0.050	-0.055	1.163	1.072	1.168	-0.123	-1.268	-1.268	2.368	4.003	3.998
18	0.116	0.037	0.066	0.761	0.606	1.068	-0.934	-0.562	-0.525	4.239	2.756	2.619
19	0.043	0.024	0.035	0.920	0.725	1.050	-0.661	-0.577	-0.534	2.945	2.435	2.335
20	-0.087	0.034	0.034	1.080	0.973	1.035	-0.165	-1.593	-1.589	2.735	5.899	5.845

Fonte: O autor (2020).

Tabela 9 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos quantílico ( $r_q$ ), ponderado ( $r^\beta$ ) e ponderado padronizado ( $r_p^\beta$ ). Modelo:  $\log\left(\frac{w_t}{1-w_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, \dots, 5$ ,  $t = 1, 2, \dots, 20$ ,  $w \in (0.8855, 0.9983)$  e  $\phi = 150$ .

t	Média			Desvio-padrão			Assimetria			Curtose		
	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$	$r_q$	$r^\beta$	$r_p^\beta$
1	-0.128	0.095	0.111	1.034	0.874	1.026	-0.147	-1.268	-1.266	2.457	4.141	4.131
2	-0.071	-0.001	-0.001	1.110	1.063	1.080	-0.071	-1.906	-1.905	2.803	7.806	7.794
3	0.124	0.025	0.053	0.766	0.615	1.124	-0.852	-0.422	-0.373	3.755	2.394	2.325
4	-0.048	-0.031	-0.032	1.125	1.082	1.117	-0.091	-1.739	-1.736	2.729	6.700	6.675
5	-0.035	-0.053	-0.057	1.136	1.092	1.185	-0.049	-1.369	-1.370	2.501	4.248	4.253
6	-0.065	-0.027	-0.027	1.139	1.109	1.133	-0.050	-1.878	-1.879	2.751	7.406	7.409
7	-0.070	-0.021	-0.022	1.136	1.092	1.131	-0.044	-1.711	-1.712	2.660	6.343	6.344
8	0.258	-0.000	-0.002	0.498	0.416	1.111	-1.265	0.159	0.140	5.446	2.359	2.287
9	0.126	0.031	0.059	0.759	0.568	1.030	-1.072	-0.139	-0.101	3.936	2.168	2.166
10	0.010	-0.123	-0.132	1.177	1.153	1.238	-0.067	-1.243	-1.243	2.350	3.735	3.734
11	-0.050	-0.034	-0.035	1.131	1.091	1.122	-0.083	-1.699	-1.696	2.717	6.254	6.232
12	-0.097	0.024	0.025	1.105	1.037	1.065	-0.086	-1.873	-1.870	2.828	7.524	7.500
13	-0.097	0.021	0.022	1.110	1.003	1.098	-0.103	-1.425	-1.421	2.510	4.757	4.733
14	-0.076	0.004	0.004	1.103	1.065	1.081	-0.013	-1.878	-1.878	2.729	7.561	7.560
15	-0.040	-0.013	-0.013	1.088	1.047	1.071	-0.106	-1.883	-1.883	2.865	7.961	7.955
16	-0.112	0.055	0.060	1.077	0.963	1.044	-0.116	-1.540	-1.532	2.671	5.450	5.393
17	-0.089	0.011	0.012	1.110	1.013	1.118	-0.054	-1.325	-1.327	2.382	4.141	4.148
18	0.126	0.029	0.052	0.754	0.612	1.120	-0.867	-0.528	-0.531	4.024	2.560	2.531
19	0.067	0.004	0.004	0.914	0.734	1.063	-0.655	-0.532	-0.527	2.987	2.260	2.246
20	-0.099	0.033	0.035	1.089	1.018	1.071	-0.029	-1.692	-1.691	2.671	6.120	6.110

Fonte: O autor (2020).

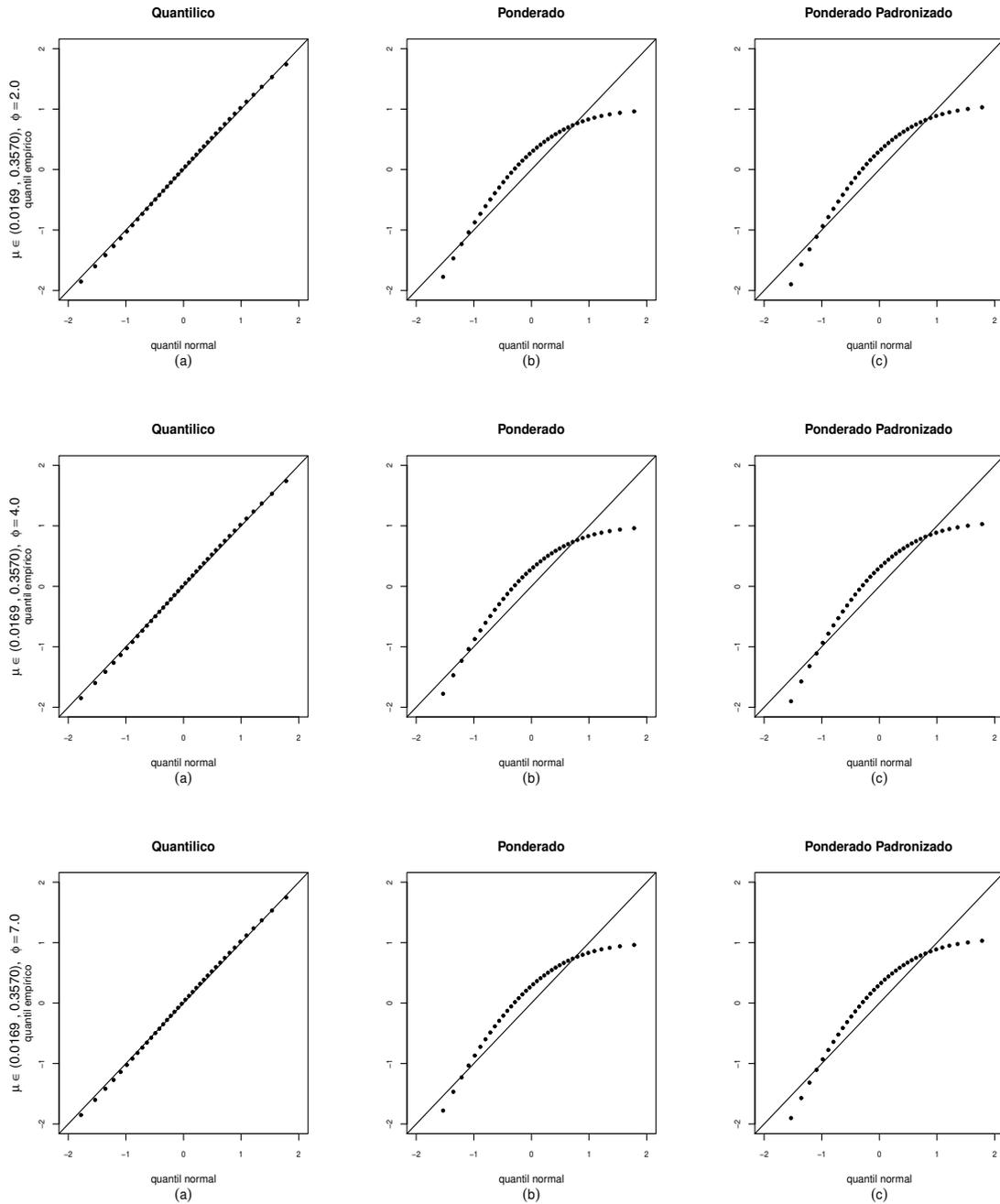
da distribuição normal padrão, construímos gráficos normais de probabilidade, que estão apresentados nas Figuras 2 a 7. Consideramos dois tamanhos amostrais  $n = 40$  e  $120$ . Consideramos ainda cenários em que as medianas da variável resposta pertencem aos intervalos  $(0.0169, 0.3570)$ ,  $(0.2353, 0.7758)$  e  $(0.8499, 0.9994)$ . Para cada um destes intervalos consideramos cenários diferentes para a precisão.

Na Figura 2 estão apresentados os gráficos normais de probabilidade para os cenários em que as medianas da variável resposta pertencem ao intervalo  $(0.0169, 0.3570)$  e variamos o parâmetro de precisão  $\phi = 2, 4, 7$ , com tamanho amostral  $n = 40$ . Os valores verdadeiros dos parâmetros de regressão são  $\beta_1 = -2$ ,  $\beta_2 = 2$ ,  $\beta_3 = -1.8$ ,  $\beta_4 = 1.3$  e  $\beta_5 = 1$ . Observamos nos gráficos normais de probabilidade da Figura 2 a existência de assimetria à esquerda nos gráficos referentes aos resíduos ponderado e ponderado padronizado.

Na Figura 3 são apresentados os gráficos normais de probabilidade para os cenários em que a mediana da variável resposta pertence ao intervalo  $(0.2353, 0.7758)$  e variamos o parâmetro de precisão  $\phi = 2, 10, 20$ , com tamanho amostral  $n = 40$ . Os valores verdadeiros dos parâmetros de regressão são  $\beta_1 = 1$ ,  $\beta_2 = 1$ ,  $\beta_3 = -1$ ,  $\beta_4 = 1.4$  e  $\beta_5 = -1.2$ . Notamos que há assimetria à esquerda nos gráficos referentes aos resíduos ponderado e ponderado padronizado nos gráficos normais de probabilidade.

Na Figura 4 estão apresentados os gráficos normais de probabilidade para os cenários em que as medianas da variável resposta pertencem ao intervalo  $(0.2353, 0.7758)$  e

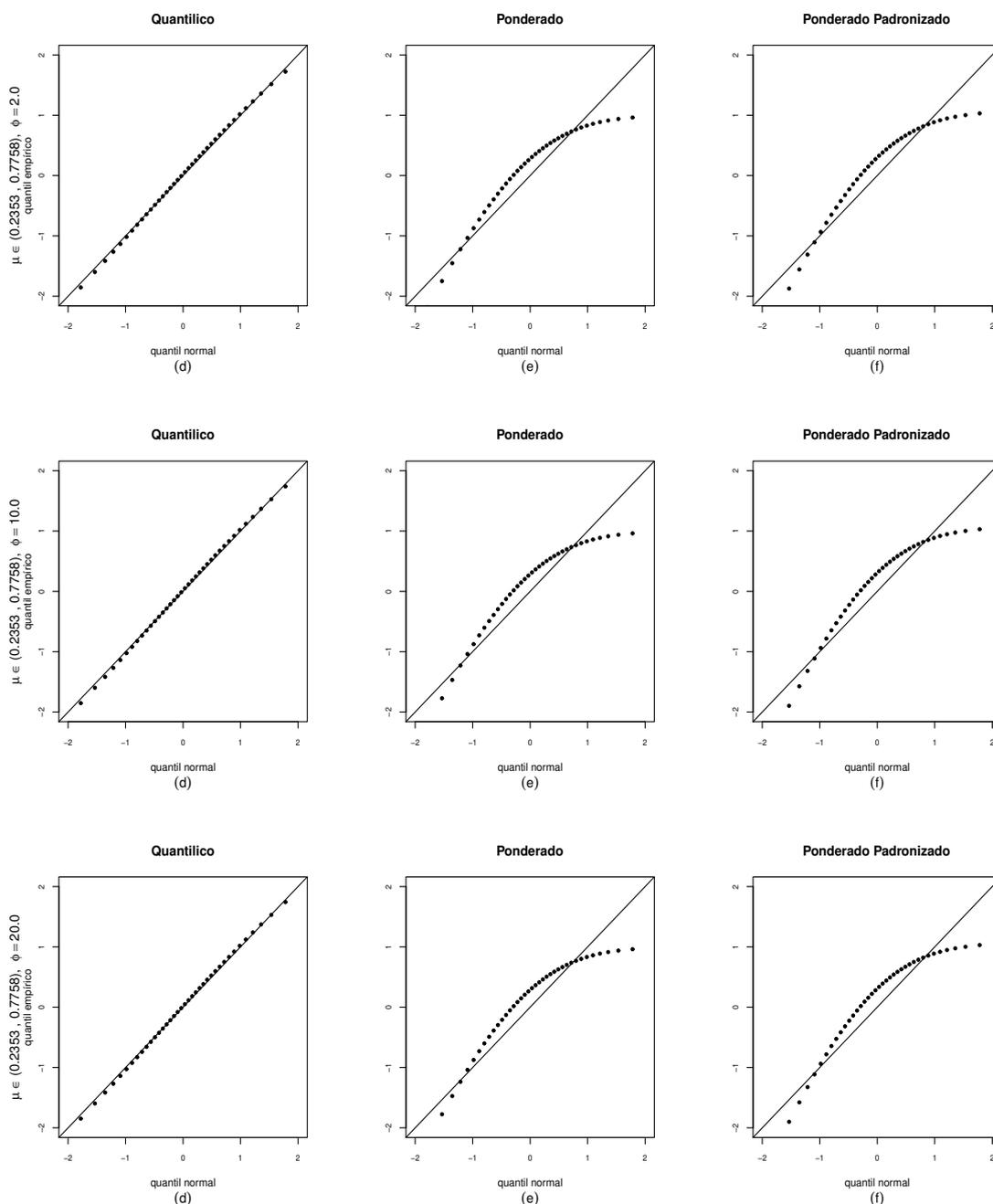
Figura 2 – Gráficos QQ; Modelo:  $\log(w_t/1 - w_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, 2, \dots, 40$ ,  $w \in (0.0169, 0.3570)$ ,  $\phi = 2, 4$  e  $7$ .



Fonte: O autor (2020).

variamos o parâmetro de precisão  $\phi = 3, 50, 150$ , com tamanho amostral  $n = 40$ . Os valores verdadeiros dos parâmetros de regressão são  $\beta_1 = 4.6$ ,  $\beta_2 = -2.6$ ,  $\beta_3 = -3.1$ ,  $\beta_4 = 2.6$  e  $\beta_5 = 3$ . Assim como nos casos anteriores notamos a existência de assimetria à esquerda nos gráficos referentes aos resíduos ponderado e ponderado padronizado.

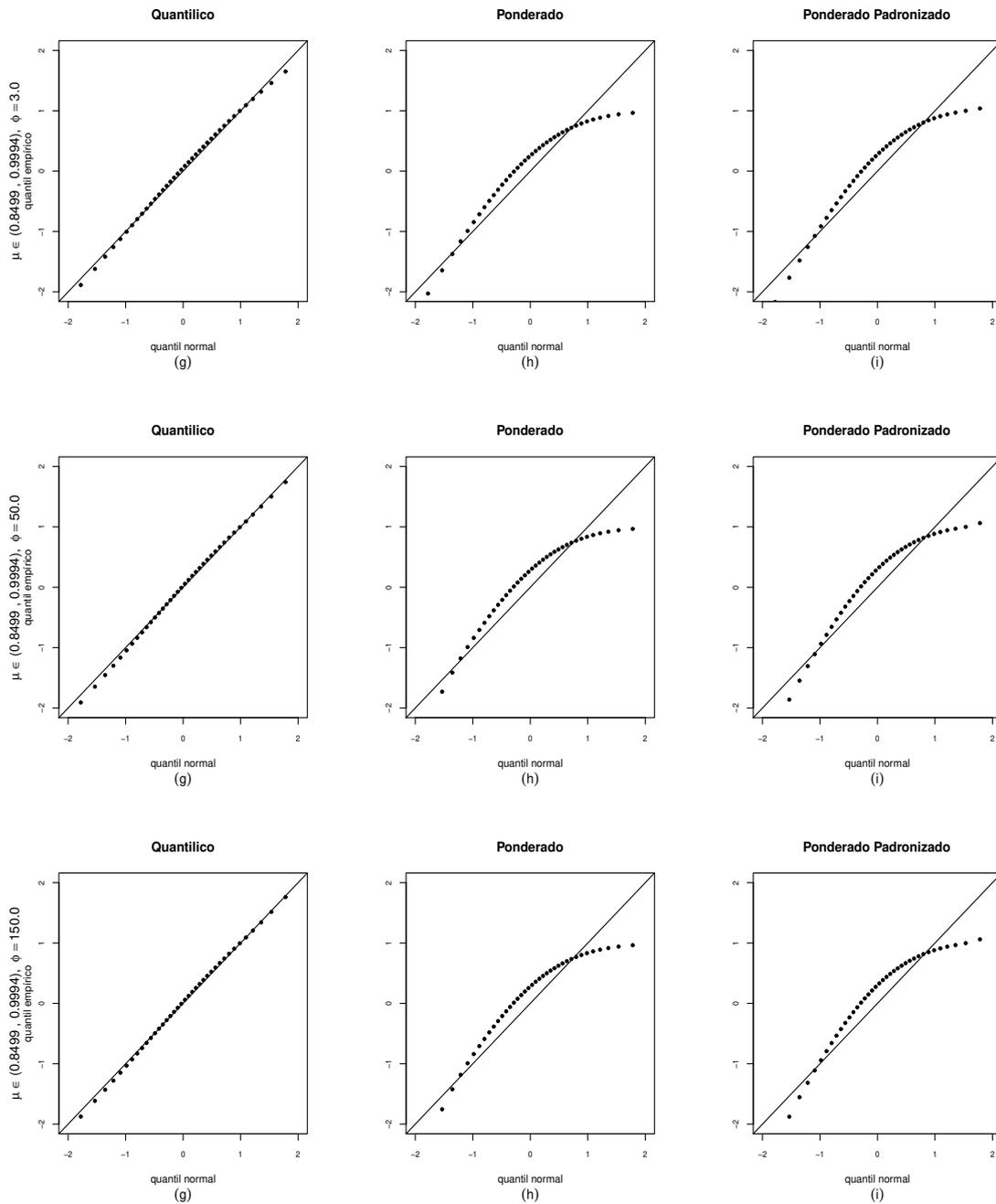
Figura 3 – Gráficos QQ; Modelo:  $\log(w_t/1 - w_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, 2, \dots, 40$ ,  $w \in (0.2353, 0.7758)$ ,  $\phi = 2, 10$  e  $20$ .



Fonte: O autor (2020).

Nas Figura 5, 6 e 7 estão apresentados os gráficos normais de probabilidade para os cenários em que as medianas da variável resposta pertencem aos intervalos  $(0.2353, 0.7758)$ ,  $(0.2353, 0.7758)$  e  $(0.8499, 0.9994)$ , respectivamente. Variamos o parâmetro de precisão de maneira análoga ao que foi feito anteriormente, com tamanho amostral  $n = 120$ . Podemos

Figura 4 – Gráficos QQ; Modelo:  $\log(w_t/1 - w_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, 2, \dots, 40$ ,  $w \in (0.8499, 0.9994)$ ,  $\phi = 3, 50$  e  $150$ .

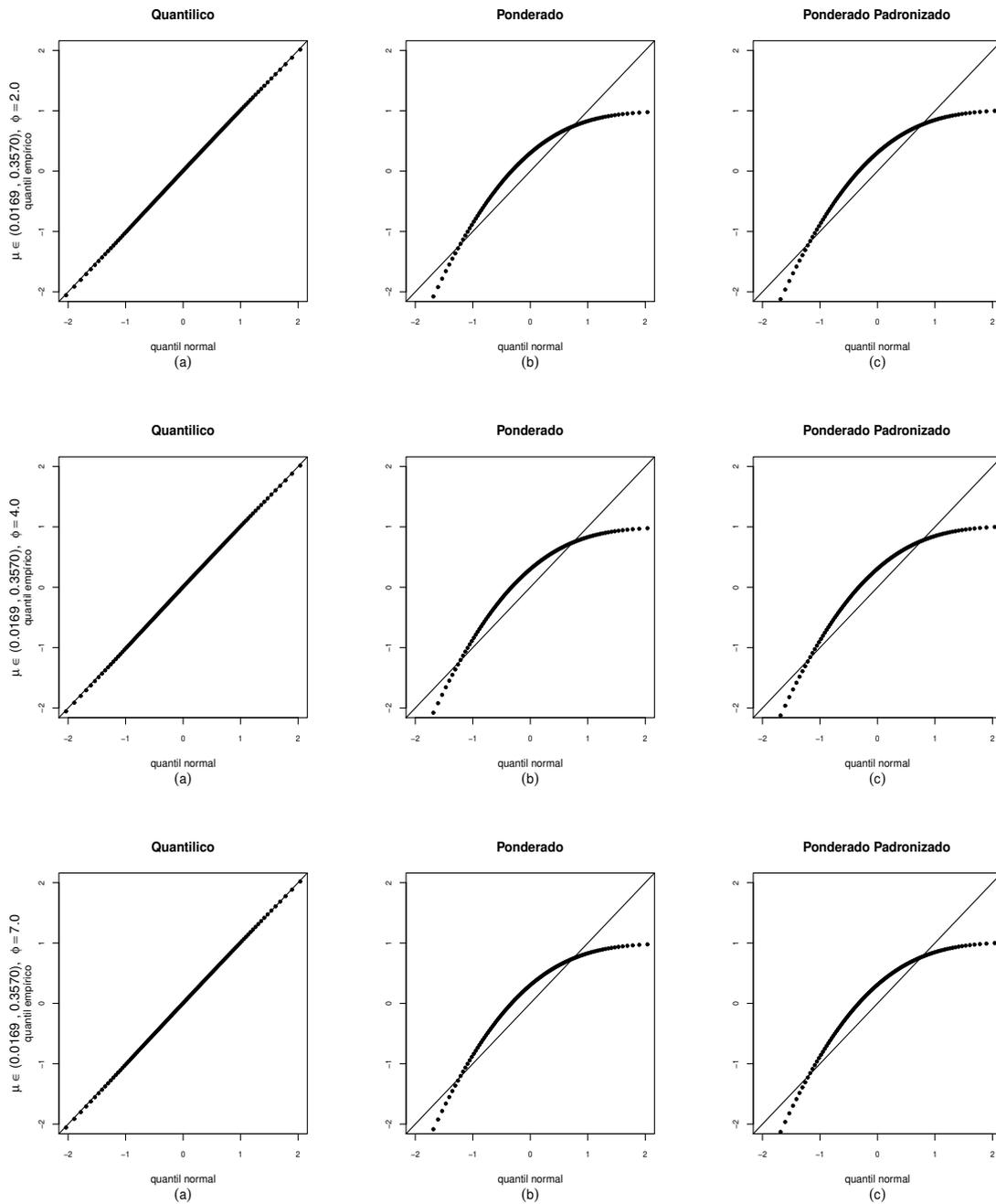


Fonte: O autor (2020).

observar que em todos do cenários há assimetria à esquerda nos gráficos referentes aos resíduos ponderado e ponderado padronizado.

Em todos os cenários considerados, os resíduos ponderado e ponderado padronizado apresentam assimetria à esquerda, o que nos leva a concluir que para estes resíduos

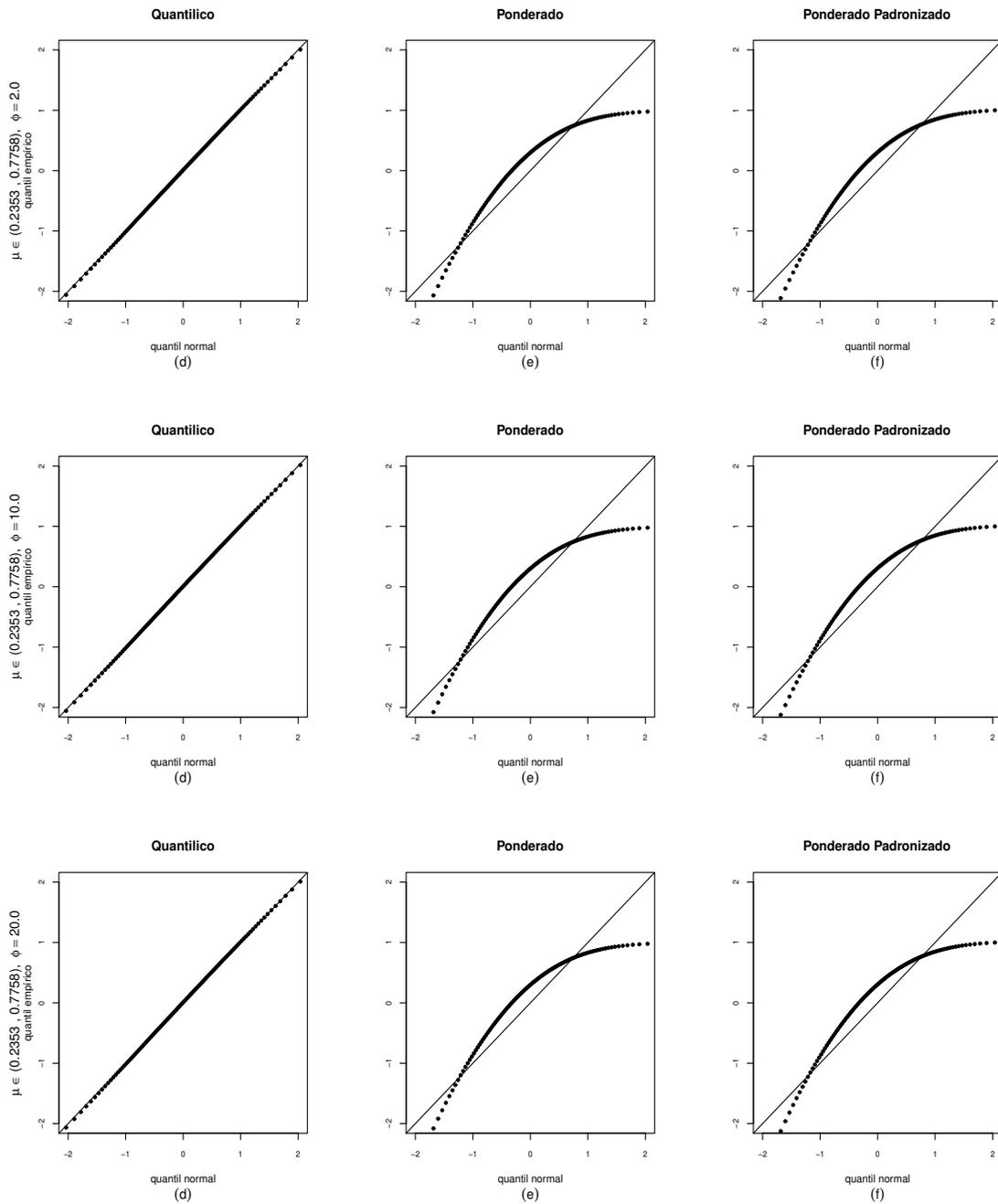
Figura 5 – Gráficos QQ; Modelo:  $\log(w_t/1 - w_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, 2, \dots, 120$ ,  $w \in (0.0169, 0.3570)$ ,  $\phi = 2, 4$  e  $7$ .



Fonte: O autor (2020).

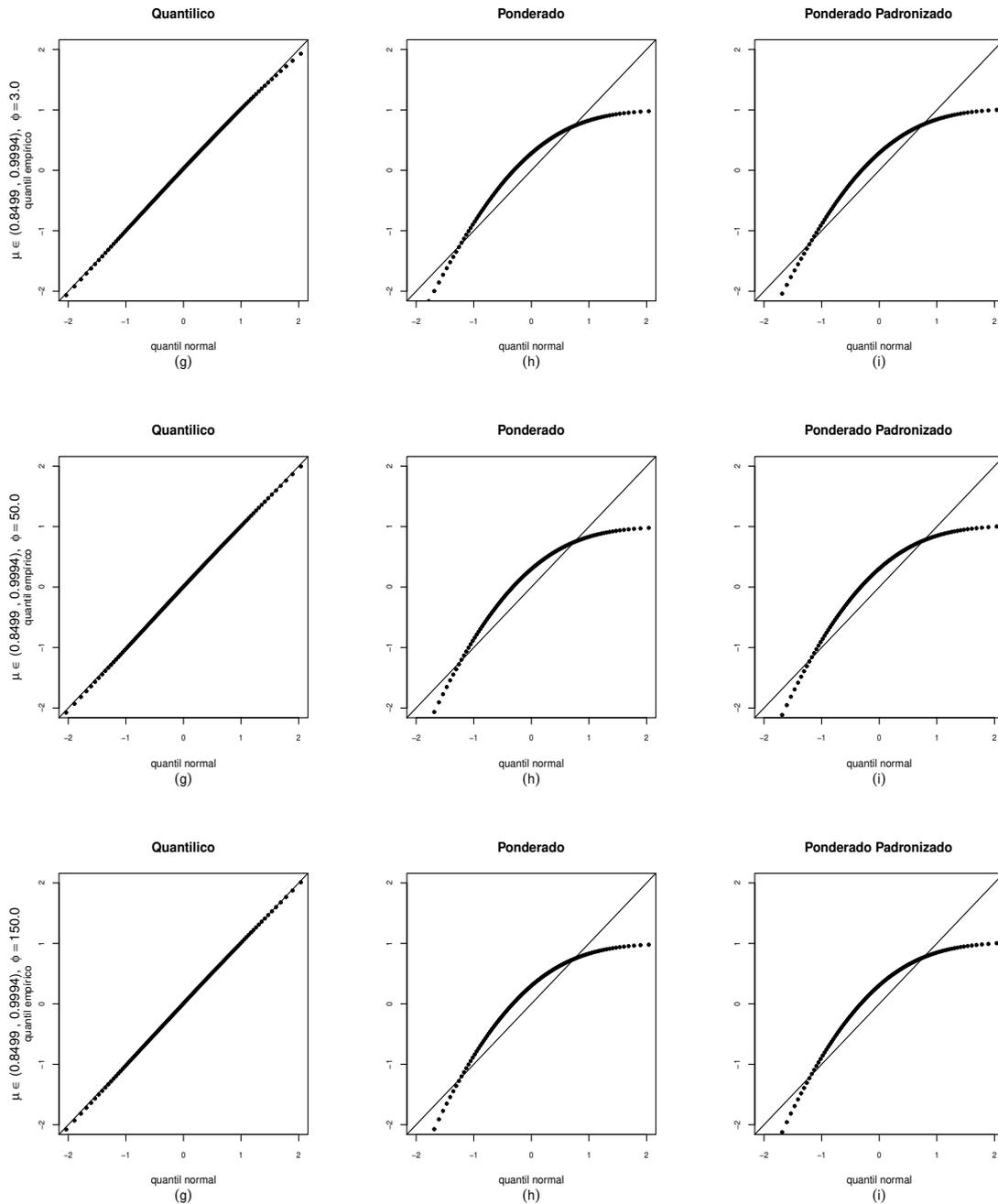
os limites usuais para detecção de pontos aberrantes  $(-2, 2)$  não devem ser utilizados. Assim, para os resíduos ponderados devemos utilizar como limites para detecção de pontos aberrantes quantis empíricos.

Figura 6 – Gráficos QQ; Modelo:  $\log(w_t/1 - w_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, 2, \dots, 120$ ,  $w \in (0.2353, 0.7758)$ ,  $\phi = 2, 10$  e  $20$ .



Fonte: O autor (2020).

Figura 7 – Gráficos QQ; Modelo:  $\log(w_t/1 - w_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, 2, \dots, 120$ ,  $w \in (0.8499, 0.9994)$ ,  $\phi = 3, 50$  e  $150$ .



Fonte: O autor (2020).

## 2.7 Aplicações

Nesta seção faremos duas aplicações do modelo de regressão Kumaraswamy a dados reais. O primeiro conjunto de dados refere-se à aplicação denominada Contaminação do

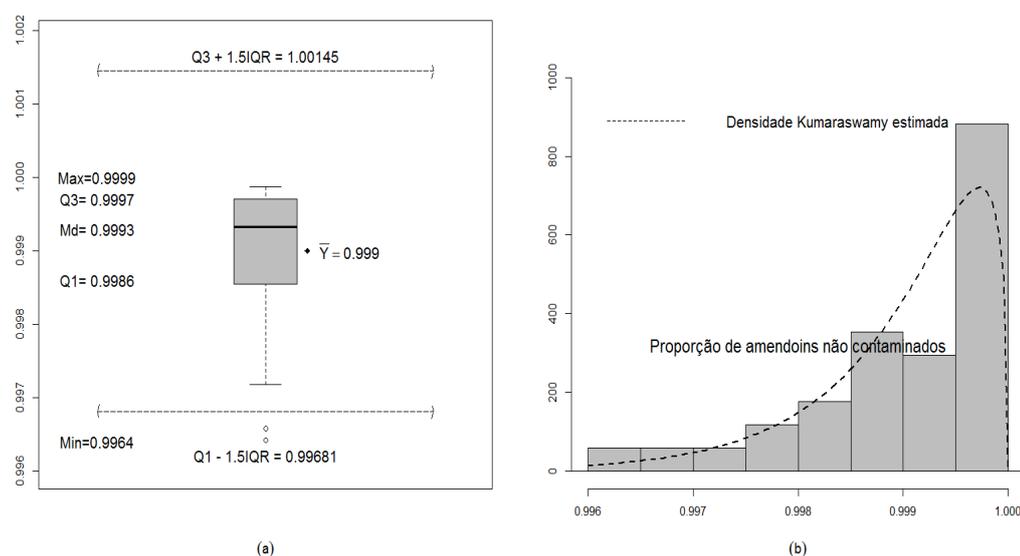
Amendoim em que utilizamos os dados fornecidos por (DRAPER; SMITH, 1998), já o segundo conjunto de dados representam a medição da umidade relativa do ar e temperatura de cada dia do mês de setembro de 2019 pela estação meteorológica Recife(Curado)-PE. Os dados estão disponíveis gratuitamente no site do INMET(<http://www.inmet.gov.br/portal/>).

## 2.8 Aplicação I: Contaminação do amendoim

Nesta aplicação utilizamos os dados fornecidos por (DRAPER; SMITH, 1998). Há 34 lotes de 120 libras de amendoim. De cada um destes lotes nos é fornecido o nível médio de aflatoxina e a proporção de amendoins não contaminados. As aflatoxinas são compostas de quatro substâncias principais que são produzidas por fungos. Esses fungos se desenvolvem sobre alimentos, por exemplo, amendoim, milho e grãos em geral, quando as condições são favoráveis. Dependendo da quantidade e concentração, a aflatoxina pode causar diversos efeitos à saúde humana.

Utilizaremos o modelo de regressão Kumaraswamy para avaliar o comportamento da proporção de amendoins não contaminados ( $y$ ) através da covariável níveis de aflatoxina ( $x$ ).

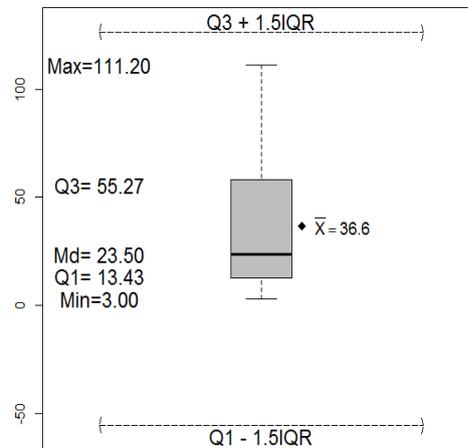
Figura 8 – Boxplot(a) e histograma(b) das observações da variável resposta.



Fonte: O autor (2020).

Inicialmente realizaremos uma análise descritiva dos dados. Construímos boxplots das observações da variável resposta e das observações da variável níveis de aflatoxina (Figuras 8 e 9). Nos boxplots consideramos algumas medidas descritivas destas variáveis. Adicionalmente, na Figura 8 apresentamos o histograma das observações da variável

Figura 9 – Boxplot das observações da variável nível de aflatoxina.



Fonte: O autor (2020).

resposta com a curva de densidade estimada Kumaraswamy. Os parâmetros da distribuição Kumaraswamy foram estimados com base no modelo Kumaraswamy sem covariadas.

Com base no gráfico boxplot das observações da variável resposta é interessante observar a concentração de valores extremamente próximos de um, limite superior do intervalo  $(0, 1)$ . Ainda com base na Figura 8, notamos como a densidade estimada Kumaraswamy se ajusta bem à resposta. Para esses dados, obtemos,  $\hat{\phi} = 1174.4$ , um valor alto.

Inicialmente consideramos um modelo com precisão variável, função de ligação logito para o submodelo da mediana e função de ligação logaritmo para o submodelo da precisão. Modelo  $M1$ :

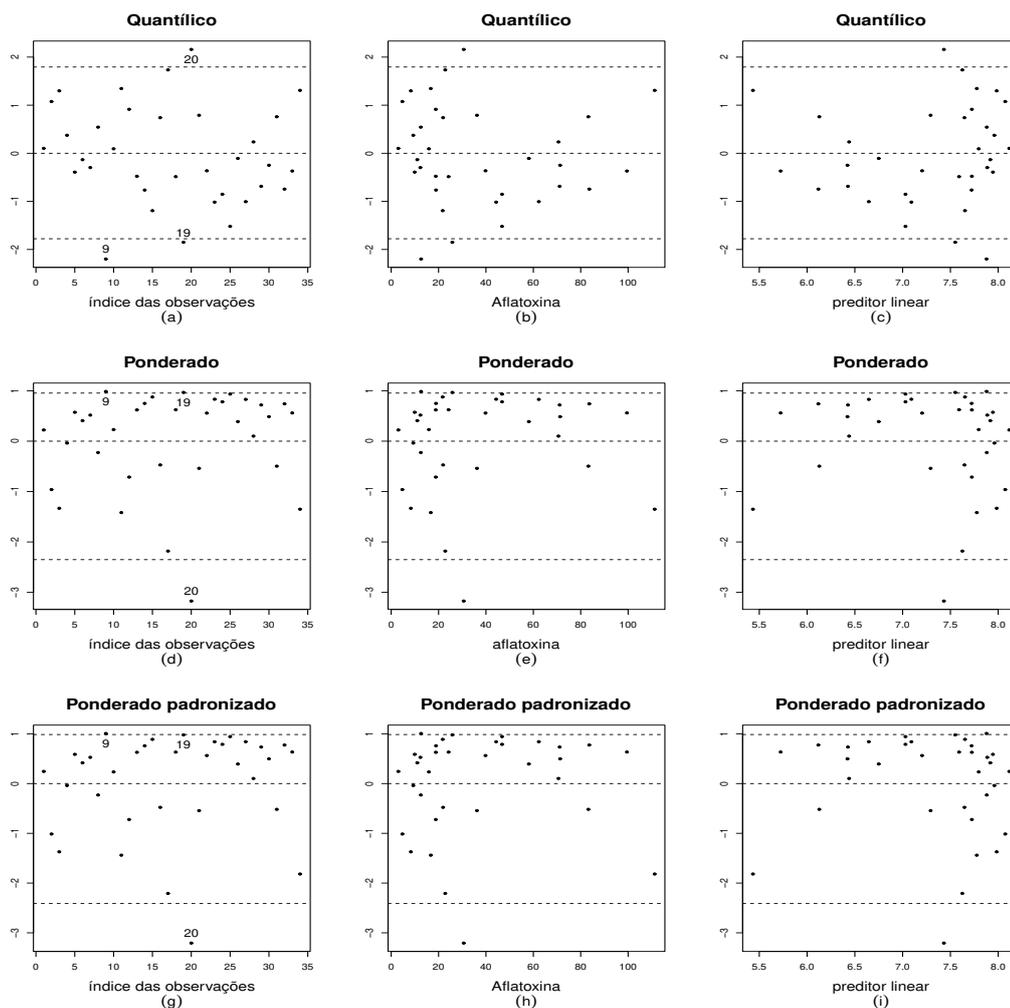
$$M1 : \log\left(\frac{w_t}{1 - w_t}\right) = \beta_1 + \beta_2 x_{t2} \quad \text{e} \quad \log(\phi_t) = \gamma_1 + \gamma_2 \log(x_{t2}),$$

$$t = 1, 2, \dots, 34.$$

Na Figura 10 apresentamos os gráficos dos resíduos quantílico, ponderado e ponderado padronizado contra diferentes elementos que compõem o modelo  $M1$  e na Figura 11 apresentamos os gráficos normais de probabilidade com envelopes simulados. Com base na Figura 11 observamos que, em sua maioria, os pontos estão dentro das bandas do envelope, o que constitui indício de que o modelo está bem ajustado. Os gráficos normais de probabilidade dos resíduos ponderado e ponderado padronizado evidenciam outro fato importante: as distribuições desses resíduos são consideravelmente assimétricas e estão distantes da distribuição normal padrão. Na Figura 10 notamos que os resíduos estão distribuídos de forma aleatória em torno de zero e evidenciamos ainda alguns pontos que merecem nossa atenção; são eles 9, 19 e 20. Este é um bom exemplo em que os usuais

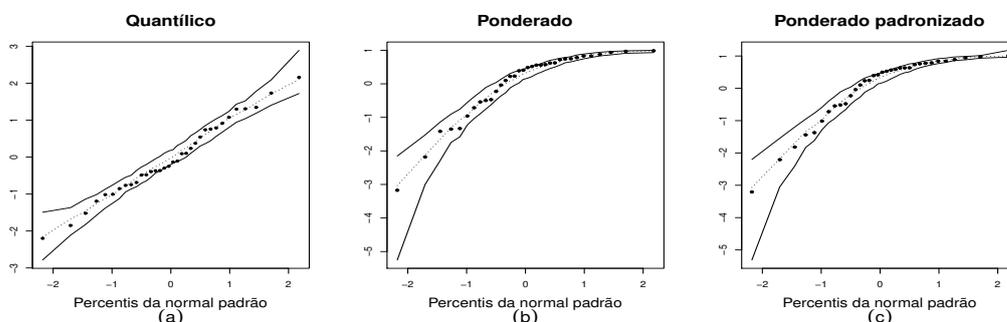
limites de detecção de pontos aberrantes  $(-2, 2)$  não devem ser utilizados em gráficos que utilizam os resíduos ponderado e ponderado padronizado. Esta afirmação é confirmada com base nos gráficos de resíduos contra elementos do modelo. Nestes gráficos, onde são usados os quantis empíricos obtidos com base na simulação do envelope, os limites para a detecção de pontos aberrantes são aproximadamente  $(-2.3, 1)$ , tanto para o resíduo ponderado quanto para o resíduo ponderado padronizado. Nesta aplicação e com base na Figura 10, nota-se uma diferença importante em relação ao desempenho dos resíduos ponderados e do resíduo quantílico aleatorizado. Nota-se que, segundo os três resíduos, o caso 20 é um ponto aberrante. No entanto, os gráficos baseados nos resíduos ponderados destacam com maior ênfase este caso e praticamente não identificam os casos 9 e 19 como aberrantes dado que os mesmos se localizam muito próximos de 1, limite superior empírico para detecção de pontos mal ajustados.

Figura 10 – Gráficos de resíduos do modelo  $\log(w_t/(1 - w_t)) = \beta_1 + \beta_2 x_{t2}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2 \log(x_{t2})$ ,  $t = 1, 2, \dots, 34$ ; Contaminação do Amendoim.



Fonte: O autor (2020).

Figura 11 – Gráficos normais de probabilidade com envelopes simulados para o modelo  $\log(w_t/(1 - w_t)) = \beta_1 + \beta_2 x_{t2}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2 \log(x_{t2})$ ,  $t = 1, 2, \dots, 34$ ; Contaminação do Amendoim.



Fonte: O autor (2020).

Tabela 10 – Estimativas dos parâmetros, erros-padrão e  $p$ -valores do modelo  $\log(w_t/1 - w_t) = \beta_1 + \beta_2 x_{t2}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2 \log(x_{t2})$ ; Contaminação do Amendoim.

Modelo $M1$	Parâmetros	Estimativas	Erros-padrão	$p$ -valores
Dados completo $M1$	$\beta_1$	8.1900	0.1478	< 0.0001
	$\beta_2$	-0.0248	0.0021	< 0.0001
	$\gamma_1$	10.1767	0.9201	< 0.0001
	$\gamma_2$	-0.6374	0.2680	0.0174
Sem a observação 9	$\beta_1$	8.2900	0.1180	< 0.0001
	$\beta_2$	-0.0261	0.0019	< 0.0001
	$\gamma_1$	10.8916	0.7623	< 0.0001
	$\gamma_2$	-0.8229	0.2263	0.0003
Sem a observação 19	$\beta_1$	8.2301	0.1335	< 0.0001
	$\beta_2$	-0.0251	0.0019	< 0.0001
	$\gamma_1$	10.3545	0.8277	< 0.0001
	$\gamma_2$	-0.6727	0.2417	0.0054
Sem as observações 9,19	$\beta_1$	8.3207	0.1087	< 0.0001
	$\beta_2$	-0.0262	0.0017	< 0.0001
	$\gamma_1$	11.0151	0.6930	< 0.0001
	$\gamma_2$	-0.8396	0.2057	< 0.0001
Sem a observação 20	$\beta_1$	8.1136	0.1745	< 0.0001
	$\beta_2$	-0.0238	0.0024	< 0.0001
	$\gamma_1$	9.8520	1.0165	< 0.0001
	$\gamma_2$	-0.5235	0.2935	0.0744

Fonte: O autor (2020).

Na Tabela 10 apresentamos as estimativas de máxima verossimilhança, erros-padrão e  $p$ -valores para o modelo  $M1$ . Como notado anteriormente, devemos dar atenção especial aos pontos destacados pelos gráficos de resíduos. Desta forma, ainda na Tabela 10 apresentamos as estimativas de máxima verossimilhança, erros-padrão e  $p$ -valores para o

modelo  $M1$  obtidos após retirarmos individualmente as observações 9, 19 e 20, e também após a retirada conjunta das observações 9 e 19.

Podemos observar que a retirada individual das observações 9 e 19, assim como a retirada conjunta destas observações, não altera significativamente os  $p$ -valores associados aos testes de significância dos parâmetros do modelo. Por outro lado, a retirada da observação 20 acarreta alteração significativa no  $p$ -valor referente ao parâmetro  $\gamma_2$ , a ponto da modelagem da precisão passar a ser desnecessária, após a exclusão deste caso ao nível de 5%.

Achamos importante destacar uma diferença entre os resíduos quantílico e os resíduos ponderado e ponderado padronizado. Apesar de todos os três resíduos destacarem os mesmos pontos, os resíduos ponderado e ponderado padronizado apresentam melhor desempenho ao destacar o ponto 20 com maior ênfase, uma vez que com base na Tabela 10 este é o ponto que nos leva à conclusão de que o modelo está mal ajustado.

Consideraremos a seguir um modelo com precisão constante e com função de ligação logito para o submodelo da mediana:

$$M2 : \log \left( \frac{w_t}{1 - w_t} \right) = \beta_0 + \beta_1 x_t, \quad t = 1, 2, \dots, 34.$$

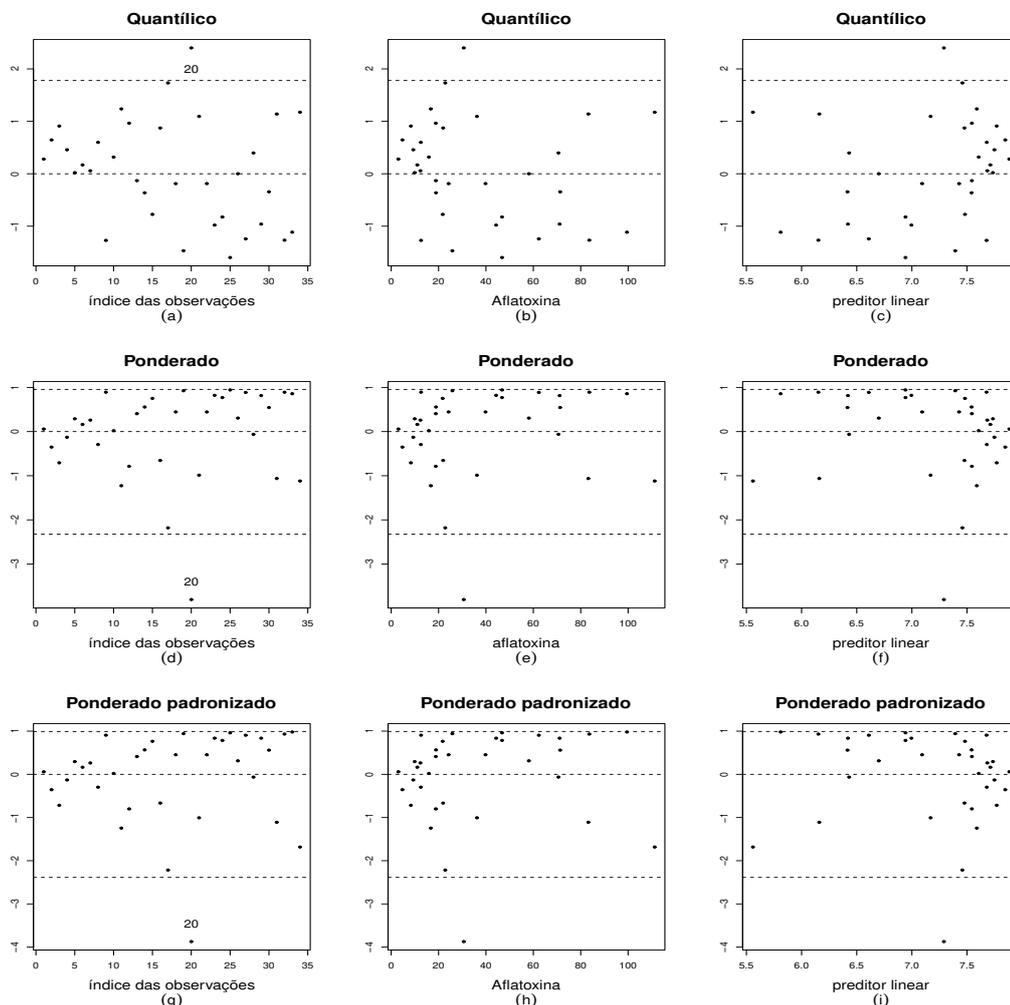
Na Figura 12 apresentamos os gráficos dos resíduos quantílico, ponderado e ponderado padronizado contra diferentes elementos do modelo acima e na Figura 13 apresentamos os gráficos normais de probabilidade com envelopes simulados. Nesta aplicação, os resíduos ponderado e ponderado padronizado apresentam comportamentos muito semelhantes ao do resíduo quantílico. Com base na Figura 13, observamos que todos os pontos estão dentro das bandas do envelope, o que é indício de que o modelo está bem ajustado. Na Figura 12 notamos que os resíduos estão distribuídos de forma aleatória em torno de zero. Evidenciamos ainda que um ponto está fora dos limites de detecção de pontos aberrantes, o ponto 20.

Na Tabela 12 podemos observar os valores das estimativas de máxima verossimilhança, erros-padrão e  $p$ -valores para o modelo  $M2$ . É importante ainda resaltar que, assim como feito no modelo  $M1$ , retiramos o ponto 20 e reestimamos os parâmetros do modelo  $M2$ , ponto esse que foi destacado nos gráficos de resíduos. A partir da Tabela 12, observamos que não houve alteração significativa nos  $p$ -valores dos testes.

No entanto, aqui devemos ressaltar que a não modelagem da precisão implica que o ponto 20 é consideravelmente mal ajustado pelo modelo, o que é evidenciado pelos gráficos construídos com base nos resíduos ponderados.

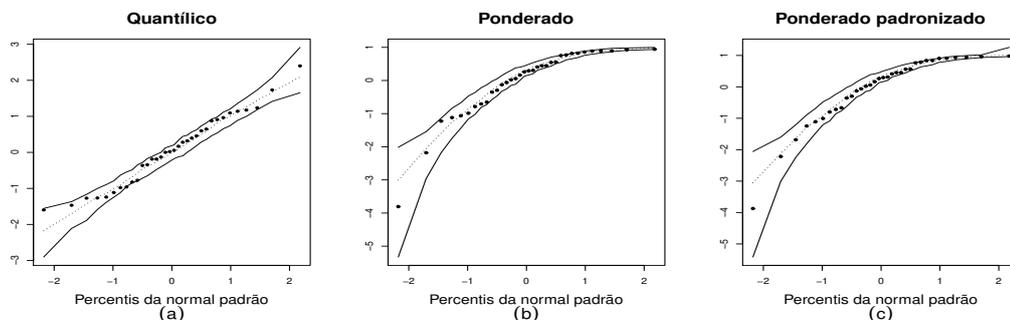
Nota-se que existe relação inversa entre o nível de aflatoxina e a proporção de amendoins não contaminados. Ou seja, quanto maior o nível de aflatoxina menor a

Figura 12 – Gráficos de resíduos do modelo  $\log(w_t/(1 - w_t)) = \beta_0 + \beta_1 x_t$ ,  $t = 1, 2, \dots, 34$ ,  $\phi$  constante; Contaminação do Amendoim.



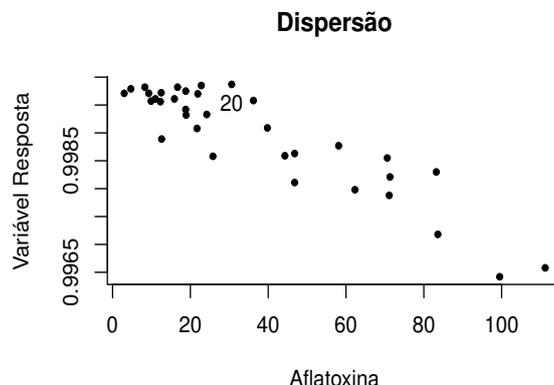
Fonte: O autor (2020).

Figura 13 – Gráficos normais de probabilidade com envelopes simulados para o modelo  $\log(w_t/(1 - w_t)) = \beta_0 + \beta_1 x_t$ ,  $t = 1, 2, \dots, 34$ ,  $\phi$  constante; Contaminação do Amendoim.



Fonte: O autor (2020).

Figura 14 – Gráfico de Dispersão; Contaminação do Amendoim.



Fonte: O autor (2020).

proporção de amendoins não contaminados. O caso 20 contradiz essa relação. Trata-se de uma observação que apresenta nível de aflatoxina igual a 30.6. Com base na Figura 14, a mediana da variável aflatoxina é igual a 23.50. Logo, o caso 20 não apresenta nível de aflatoxina consideravelmente alto. No entanto, a proporção de amendoins não contaminados para esse lote é igual a 0.99987, o maior valor da resposta. Para entender melhor a questão deve-se notar que o segundo maior valor para a resposta (0.99982) apresenta nível de aflatoxina igual a 16.7. Assim, de fato, o caso 20 é consideravelmente atípico.

Tabela 11 – Estimativas dos parâmetros, erros-padrão e  $p$ -valores do modelo  $\log(w_t/(1 - w_t)) = \beta_1 + \beta_2 x_{t2}$  e  $\phi$  constante; Contaminação do Amendoim.

Modelo $M2$	Parâmetros	Estimativas	Erros-padrão	$p$ -valores
Dados completos	$\beta_1$	7.9469	0.1425	< 0.0001
	$\beta_2$	-0.0215	0.0015	< 0.0001
	$\gamma_1$	7.8878	0.1648	< 0.0001
Sem a observação 20	$\beta_1$	7.8800	0.1307	< 0.0001
	$\beta_2$	-0.0208	0.0014	< 0.0001
	$\gamma_1$	7.9782	0.1662	< 0.0001

Fonte: O autor (2020).

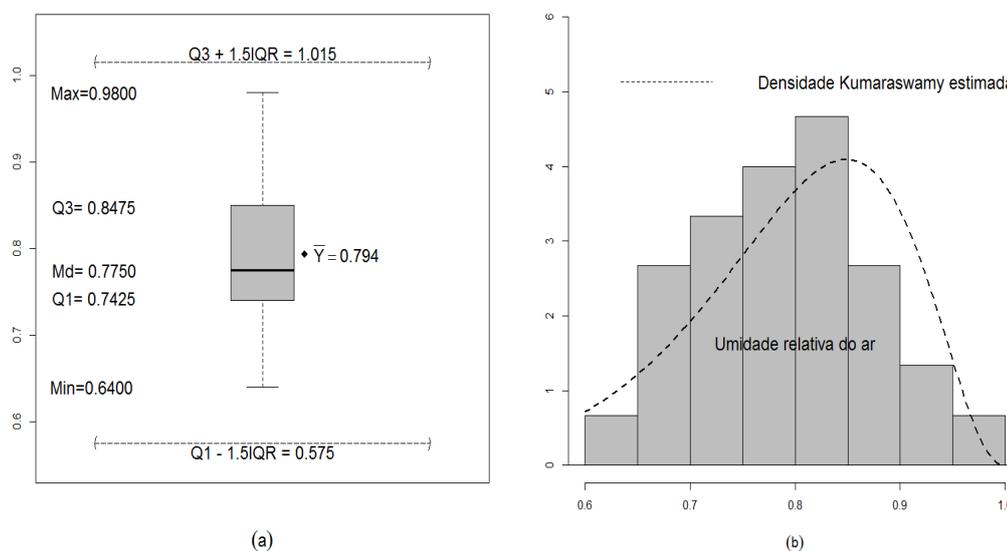
## 2.9 Aplicação II: Umidade relativa do ar

A umidade relativa do ar é uma característica meteorológica importante para a saúde pública, para o planejamento de plantações e em estudos hidrológicos. A Organização Mundial de Saúde (OMS) recomenda que a umidade relativa do ar deve permanecer entre 40% e 70%. Abaixo do limite inferior pode causar ressecamento das mucosas das vias

aéreas, tornando as pessoas mais vulneráveis a problemas respiratórios. Por outro lado, valores acima de 70% deixam o ar praticamente saturado de vapor de água, o que interfere no mecanismo humano de controle de temperatura corporal. Além de ser responsável pelo aumento da precipitação, que pode causar sérias consequências, por exemplo, no caso de plantações de milho, o excesso de umidade pode ser bastante prejudicial variando conforme a fase das lavouras. No período de enchimento dos grãos, por exemplo, a baixa luminosidade implica a redução da taxa de fotossíntese e do metabolismo da planta afetando a translocação de fotoassimilados para os grãos. Nas grandes cidades, como por exemplo o Recife, ela afeta principalmente a drenagem urbana. Do ponto de vista hidrológico, a umidade relativa do ar interfere na pressão de vapor que por sua vez influencia a estimação do grau de evaporação (ALLEN et al., 1998). Compreender e modelar o comportamento da umidade relativa do ar é, assim, de extrema importância.

Os dados que analisaremos representam a medição da umidade relativa do ar e temperatura, feitas às 12 horas de cada dia do mês de setembro de 2019 pela estação Recife(Curado)-PE, com tamanho amostral  $n = 30$ . Os dados estão disponíveis no site do INMET (<http://www.inmet.gov.br/portal/>). Utilizaremos o modelo de regressão Kumaraswamy para avaliar o comportamento da umidade relativa do ar ( $y$ ) através da covariável temperatura ( $x$ ).

Figura 15 – Boxplot(a) e histograma(b) das observações da variável resposta.



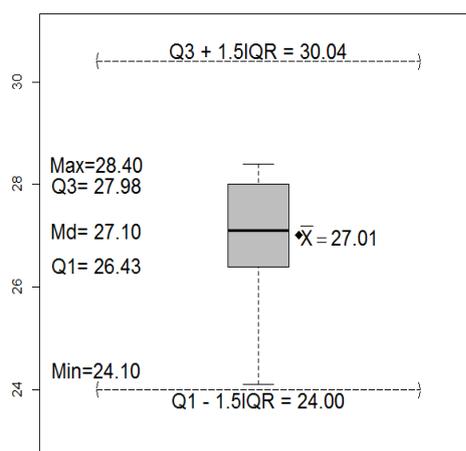
Fonte: O autor (2020).

Iniciamente realizaremos uma análise descritiva dos dados. Construimos os gráficos boxplots das observações da variável resposta e das observações da variável temperatura (Figuras 15 e 16, respectivamente). Os gráficos incluem algumas medidas descritivas. Adicionalmente, na Figura 15 apresentamos o histograma das observações da variável resposta

acompanhado da densidade Kumaraswamy estimada. Os parâmetros foram estimados com base no modelo Kumaraswamy sem covariadas.

Com base no boxplot das observações da variável resposta podemos observar a ausência de outliers. Ainda com base na Figura 15, notamos como a densidade Kumaraswamy se ajusta bem à resposta.

Figura 16 – Boxplot das observações da variável temperatura.



Fonte: O autor (2020).

Para analisar esse conjunto de dados consideramos o modelo de regressão Kumaraswamy com precisão constante e com função de ligação logito para o submodelo da mediana definido por

$$M3 : \log \left( \frac{w_t}{1 - w_t} \right) = \beta_0 + \beta_1 x_t, \quad t = 1, 2, \dots, 30.$$

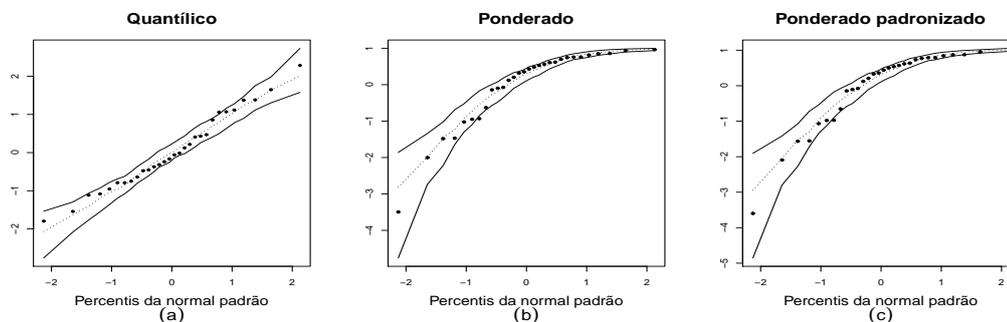
Na Figura 17 apresentamos o gráfico normal de probabilidade com envelope simulado e na Figura 18 apresentamos os gráficos de resíduos quantílico, ponderado e ponderado padronizado contra diferentes elementos do modelo acima.

Com base na Figura 17, observamos que um único ponto se encontra fora das bandas do envelope, o que é indício de que o modelo está bem ajustado. Na Figura 18 notamos que os resíduos estão distribuídos de forma aleatória em torno de zero. Há dois pontos fora dos limites de detecção de pontos aberrantes, os pontos 10 e 30.

A Tabela 12 contém as estimativas de máxima verossimilhança, os erros-padrão e os  $p$ -valores para o modelo  $M3$ .

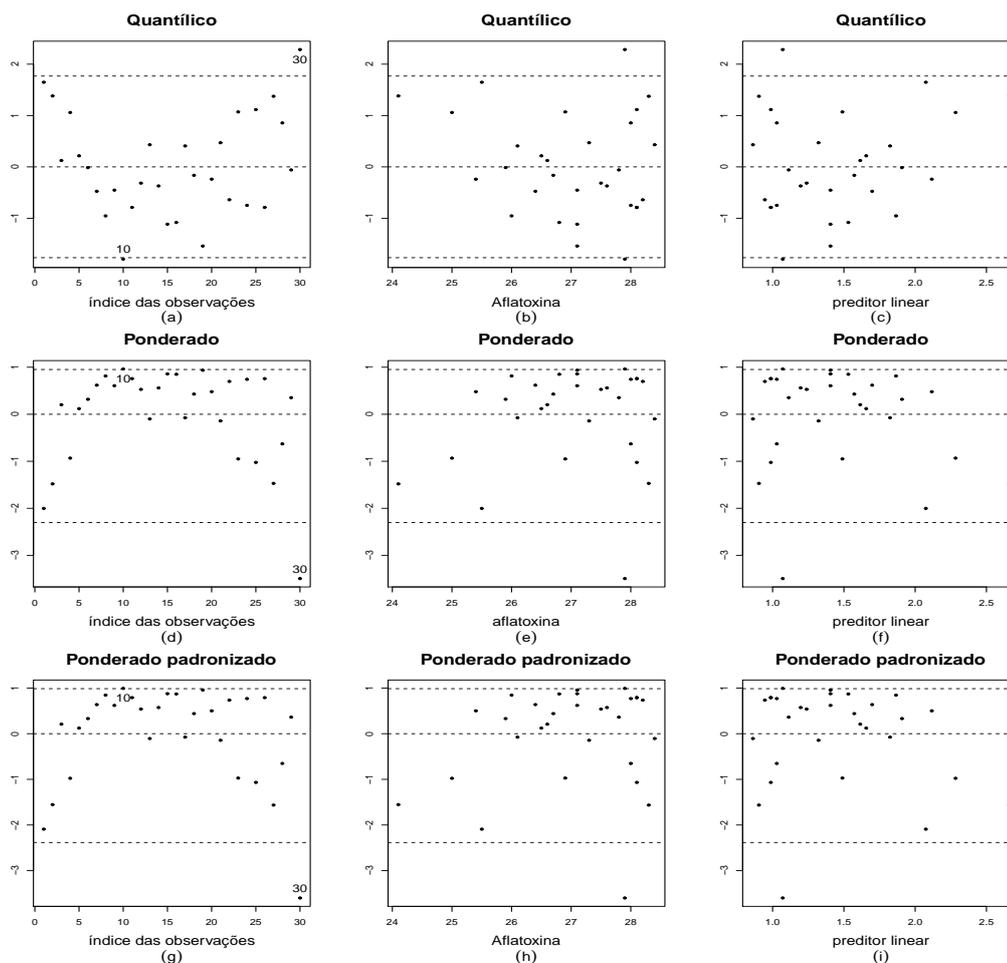
É importante ainda resaltar que retiramos os pontos 10 e 30 separadamente e reestimamos os parâmetros do modelo  $M3$ . Esses pontos foram, como vimos, destacados

Figura 17 – Gráficos normais de probabilidade com envelopes simulados para o modelo  $\log(w_t/(1 - w_t)) = \beta_0 + \beta_1 x_t$ ,  $t = 1, 2, \dots, 30$ ,  $\phi$  constante; Umidade Relativa do Ar.



Fonte: O autor (2020).

Figura 18 – Gráficos de resíduos do modelo  $\log(w_t/(1 - w_t)) = \beta_0 + \beta_1 x_t$ ,  $t = 1, 2, \dots, 30$ ,  $\phi$  constante; Umidade Relativa do Ar.



Fonte: O autor (2020).

nos gráficos de resíduos. A partir da Tabela 12, observamos que não houve alteração significativa nos  $p$ -valores dos testes.

Tabela 12 – Estimativas dos parâmetros, erros-padrão e  $p$ -valores do modelo  $\log(w_t/(1 - w_t)) = \beta_1 + \beta_2 x_{t2}$  e  $\phi$  constante; Umidade Relativa do Ar.

Modelo $M3$	Parâmetros	Estimativas	Erros-padrão	$p$ -valores
Dados completos	$\beta_1$	12.7366	1.4162	< 0.0001
	$\beta_2$	-0.4181	0.0509	< 0.0001
	$\gamma_1$	2.9618	0.1481	< 0.0001
Sem a observação 10	$\beta_1$	12.7081	1.3767	< 0.0001
	$\beta_2$	-0.4164	0.0495	< 0.0001
	$\gamma_1$	3.0245	0.1488	< 0.0001
Sem a observação 30	$\beta_1$	13.0821	1.2334	< 0.0001
	$\beta_2$	-0.4315	0.0442	< 0.0001
	$\gamma_1$	3.0484	0.1517	< 0.0001

Fonte: O autor (2020).

## 2.10 Conclusão

Neste capítulo desenvolvemos o cálculo dos valores iniciais para o método de estimação dos parâmetros do modelo de regressão Kumaraswamy e propusemos dois novos resíduos: resíduo ponderado e ponderado padronizado, que são baseados no processo iterativo Score de Fisher. Mostramos que, para estes resíduos, deve-se utilizar como limites para detecção de pontos aberrante os quantis empíricos, pois ao observar o comportamento destes resíduos, isto é, ao avaliar as distribuições empíricas dos resíduos, notamos que tais distribuições apresentam assimetria à esquerda. No processo de avaliação, utilizamos vários tamanhos amostrais e observamos o referido comportamento em todos os cenários.

Por fim, concluímos que para o conjunto de dados denominado Contaminação do Amendoim os resíduos ponderado e ponderado padronizado apresentam melhor desempenho que o resíduo quantílico aleatorizado, pois ao aplicarmos o modelo de regressão Kumaraswamy a este conjunto de dados pudemos observar que os resíduos propostos quantificaram com maior ênfase a observação 20, observação esta que mostramos ser atípica.

## 3 ANÁLISE DE INFLUÊNCIA

### 3.1 Introdução

Modelos de regressão são comumente utilizados para estudar comportamento de variáveis de interesse que são relacionadas com outras variáveis. Quando esta relação entre a variável dependente (resposta) e as variáveis independentes (regressoras ou covariáveis) é linear, o modelo de regressão frequentemente utilizado é o modelo de regressão normal linear. Tal modelo, contudo, impõe outras restrições, além da linearidade; por exemplo, considera-se que a variável resposta segue distribuição normal. Por esse motivo, o modelo de regressão normal linear não é apropriado em situações em que a variável resposta assume valores no intervalo  $(0, 1)$ , já que a suposição de normalidade pode conduzir a valores ajustados para a variável dependente fora dos limites de tal intervalo.

Em muitas situações práticas, existe a necessidade de modelar dados no intervalo  $(0, 1)$ , por exemplo, taxas e proporções. Vários modelos de regressão foram propostos nas últimas décadas para variáveis dependentes que assumem valores em  $(0, 1)$ . Por exemplo, (BARNDORFF-NIELSEN; JØRGENSEN, 1991) propuseram o modelo de regressão simplex, (FERRARI; CRIBARI-NETO, 2004) propuseram o modelo de regressão beta e (LEMONTE; BAZÁN, 2016) propuseram o modelo de regressão Johnson  $S_B$ . Em particular, (FERRARI; CRIBARI-NETO, 2004) propuseram uma reparametrização para a distribuição beta de modo a permitir a modelagem da média da variável resposta. Tal modelo está implementado no software estatístico **R** distribuído por (<https://www.r-project.org/>) através do pacote **betareg**.

A distribuição beta é uma das mais usadas para modelar experimentos aleatórios que produzem resultados no intervalo  $(0, 1)$ , sendo muito utilizada na modelagem de taxas ou proporções. É uma das famílias de distribuição probabilidade mais flexíveis, fato este devido aos seus dois parâmetros de forma  $p$  e  $q$ , e que facilitam sua aplicação a dados reais.

Segundo (BALAKRISHNAN; NEVZOROV, 2003), a distribuição beta com parâmetros  $p$  e  $q$  possui a propriedade de quase-simetria, isto é, se  $Y \sim \text{Beta}(p, q)$  então  $1 - Y \sim \text{Beta}(q, p)$ . Adicionalmente, se  $Y \sim \text{Beta}(p, q)$ , então  $-\log(Y) \sim \text{Exponencial}(p)$ . Dentre as diversas forma assumidas por sua função de densidade de probabilidade, encontra-se a uniforme, que ocorre quando  $p = q = 1$ . A densidade beta é unimodal quando  $p > 0$  e  $q > 1$ .

Outra opção ainda pouco explorada é o modelo de regressão Kumaraswamy proposto por (MITNIK; BAEK, 2013) em que os autores fazem uma reparametrização da distribuição

Kumaraswamy permitindo a modelagem da mediana da variável resposta. Esta distribuição foi proposta em (KUMARASWAMY, 1976) e reformulada em (KUMARASWAMY, 1980). Assim como a distribuição beta, ela pode ser usada para modelar variáveis dependentes que assumem valores no intervalo  $(0, 1)$ . A distribuição Kumaraswamy é uma distribuição de probabilidade contínua com suporte em  $(a, b)$ ,  $a < b$  com  $a, b \in \mathbb{R}$ . No presente trabalho, consideraremos apenas o intervalo  $(0, 1)$ . Esta distribuição se assemelha em muitos aspectos à distribuição beta. Trata-se de uma distribuição bastante flexível. Sua densidade assume formas muito distintas dependendo dos valores de seus parâmetros, podendo assim ser usada para modelar uma grande variedade de dados.

Diversos autores vêm trabalhando com a distribuição Kumaraswamy. (JONES, 2009) faz um comparativo com a distribuição beta, exibindo algumas vantagens no uso da distribuição Kumaraswamy tais como: função de distribuição e função quantil de forma fechada. (LEMONTE, 2011) desenvolve uma correção de viés para os estimados de máxima verossimilhança dos parâmetros que indexam a distribuição Kumaraswamy. (MITNIK, 2013) desenvolveu muitas propriedades importantes, por exemplo: semelhante ao que ocorre com variáveis que seguem distribuição beta, qualquer transformação linear de uma variável aleatória Kumaraswamy também segue essa distribuição e possui os mesmos parâmetros de forma que a distribuição original. (MITNIK; BAEK, 2013) propuseram o modelo de regressão Kumaraswamy a partir de uma reparametrização dos parâmetros de forma originais.

A análise de diagnóstico teve início com a análise de resíduos e sua evolução foi motivada por ter sido observado que aspectos importantes de um modelo podem ser dominados por uma única observação no processo inferencial. Desta forma, a evolução dos métodos de diagnóstico segue no sentido de detectar observações que exercem efeito desproporcional sobre o ajuste do modelo, podendo tais observações inclusive interferir em resultados inferenciais. Neste contexto, foram propostas a distância de (COOK, 1977) e também as medidas de influência local.

A distância de (COOK, 1977) avalia o impacto individual da retirada de uma observação dos dados sobre as estimativas dos coeficientes de regressão. Esta abordagem de eliminação individual de casos, que é a ideia que se encontra por trás da distância de Cook, é exemplo de análise de influência global.

O método de influência local proposto por (COOK, 1986) é uma importante ferramenta na análise de diagnóstico, sendo recomendado quando há interesse em investigar a sensibilidade do ajuste do modelo sob pequenas perturbações no modelo ou nos dados. Se estas perturbações causarem efeitos desproporcionais nas estimativas, haverá indício de afastamento das suposições impostas ou de mal ajuste do modelo.

Seja  $\delta$  um vetor de perturbações  $m$ -dimensional. Denotamos a função de log-

verossimilhança perturbada por  $\ell(\theta|\delta)$ . Suponha que existe um vetor de não perturbação  $\delta_0$  tal que  $\ell(\theta|\delta_0) = \ell(\theta)$ . A influência de pequenas perturbações no estimador de máxima verossimilhança  $\hat{\theta}$  pode ser avaliada usando o deslocamento da log-verossimilhança:

$$LD_\delta = 2(\ell(\hat{\theta}) - \ell(\hat{\theta}_\delta)), \quad (3.1)$$

em que  $\hat{\theta}_\delta$  denota o estimador de máxima verossimilhança avaliado em  $\ell(\hat{\theta}|\delta)$ .

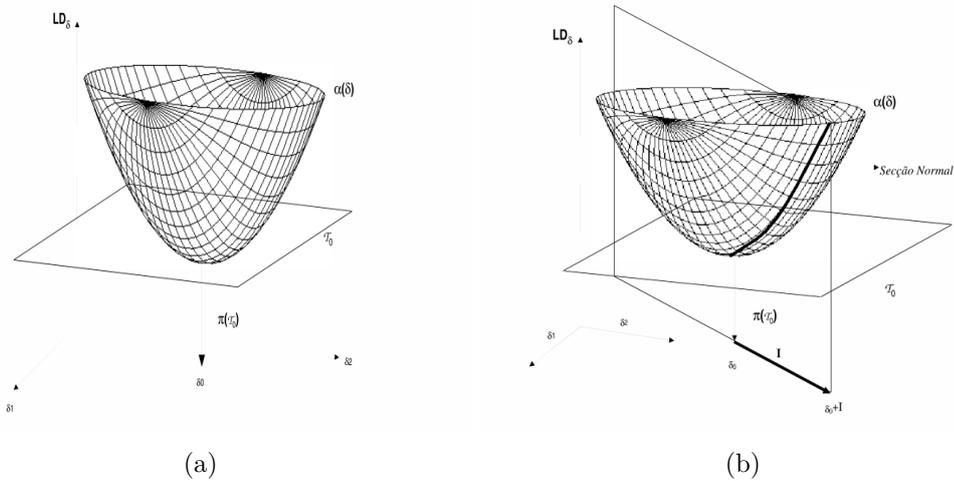
A ideia de (COOK, 1986) foi estudar o comportamento local de  $LD_\delta$ , em uma vizinhança de  $\delta_0$  tal que  $LD_{\delta_0} = 0$ . O procedimento proposto por (COOK, 1986) consiste em avaliar como a superfície  $G(\delta) = (\delta, LD_\delta)^\top$ , se desvia de seu plano tangente em  $\delta_0$  à medida que  $\delta$  se afasta de  $\delta_0$ , ou seja, quando pequenas perturbações são introduzidas no modelo. Tal descrição pode ser obtida pelo estudo das curvaturas das secções normais da superfície  $G(\delta)$  em  $\delta_0$ , que são as chamadas curvaturas normais. Para definir o que é uma secção normal da superfície  $G(\delta)$  em  $\delta_0$  consideramos o plano tangente a  $G(\delta)$  em  $\delta_0$ ,  $\tau_0$ , e o vetor que é ortogonal a  $\tau_0$ , que chamaremos de  $\pi(\tau_0)$ . Na Figura 3.1(a) apresentamos a superfície  $G(\delta) = (\delta, LD_\delta)^\top$ , o plano tangente à superfície em  $\delta_0$  e o vetor  $\pi(\tau_0)$ . Para esta ilustração, admitimos que  $\delta = (\delta_1, \delta_2)$ . Assim,  $\delta_0 = (\delta_{01}, \delta_{02})$ .

Os planos que contêm o vetor normal  $\pi(\tau_0)$  são chamados de planos normais a  $G(\delta)$  em  $\delta_0$ . As secções normais em  $\delta_0$  são definidas como as intersecções destes planos com a superfície  $G(\delta)$ . Consideremos agora um deslocamento ao redor de  $\delta_0$ , ou seja, introduzimos uma pequena perturbação no modelo. Escolhemos uma direção  $a$  tal que  $\|a\| = 1$  e tomamos  $\delta = \delta_0 + a$ . Desta forma, temos o ponto  $LD_{\delta_0+a} \neq 0$ . Por esse ponto passa o plano que contém os vetores  $a$  e  $\pi(\tau_0)$  e a intersecção deste plano com  $G(\delta)$  forma uma secção normal (Figura 3.1(b)). De fato, cada direção  $a$  escolhida associada ao vetor  $\pi(\tau_0)$  especifica uma secção normal e define um valor de curvatura normal. Cada secção normal é equivalente ao gráfico  $LD_{\delta_0+al} \times l$  em que  $l \in \mathbb{R}$ . Esse gráfico é chamado de linha projetada.

Na Figura 3.1 ilustramos o comportamento de duas linhas projetadas,  $LD_{\delta_0+a_1l} \times l$  e  $LD_{\delta_0+a_2l} \times l$ , relacionadas, respectivamente, às direções  $a_1$  e  $a_2$ . Uma vez que  $LD_{\delta_0} = 0$ , segue que  $LD_{\delta_0+al}$  tem mínimo local em  $l = 0$ . Assim, é possível notar que a direção  $a_2$  é a que provoca os maiores deslocamentos na função de verossimilhança à medida que os valores de  $l$  se afastam de zero. A ideia de (COOK, 1986) é encontrar a direção  $a$  tal que a linha projetada resultante apresente a maior das curvaturas normais ( $C_{\max}$ ). Segundo (COOK, 1986), esse vetor, denotado por  $d_{\max}$ , indica os valores que devem ser atribuídos aos componentes de  $\delta$  para obter a maior mudança local no deslocamento pela verossimilhança.

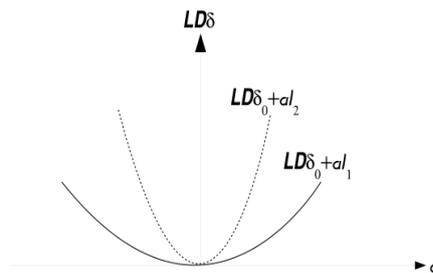
(POON; POON, 1999) chamam atenção para o fato de que a curvatura normal pode assumir qualquer valor real e não é invariante sob uma mudança uniforme de escala, ocasionando perda de objetividade no julgamento da gradeza da curvatura. Com o

Figura 19 – (a) Superfície  $G(\delta)$ , (b) Secção normal.



Fonte: O autor (2020).

Figura 20 – Linhas projetadas  $LD_{\delta_0+a_1l} \times l$  e  $LD_{\delta_0+a_2l} \times l$ .



Fonte: O autor (2020).

objetivo de solucionar esse problema e conseqüentemente aperfeiçoar o método de influência local, (POON; POON, 1999) propõem que façamos uso da curvatura conforme, que está relacionada com a curvatura normal, mas assume valores em um intervalo limitado e é invariante sob uma classe de reparametrizações que definiremos mais adiante. A curvatura normal conforme e suas propriedades possibilitam a difinição de valores de referência que permitem que se avalie a grandeza desta curvatura.

### 3.2 Modelo de regressão Kumaraswamy

Uma variável aleatória  $Y$  segue distribuição Kumaraswamy com parâmetros  $p$  e  $q$ , denotada por  $Y \sim Kum(p, q) = Kum(p, q, 0, 1)$ , se a sua função de densidade é da forma

$$f(y; p, q) = pqy^{p-1}(1 - y^p)^{q-1}, \quad y \in (0, 1), \tag{3.2}$$

com  $p > 0$  e  $q > 0$ . A função de distribuição acumulada é dada por

$$F(y; p, q) = 1 - (1 - y^p)^q. \quad (3.3)$$

A média e a variância de  $Y$  são dadas, respectivamente, por

$$\mathbb{E}(Y) = \frac{q\Gamma(1 + \frac{1}{p})\Gamma(q)}{\Gamma(1 + \frac{1}{p} + q)} \quad (3.4)$$

e

$$\text{Var}(Y) = \mu_2 - \mu_1^2,$$

em que

$$\mu_r = \frac{q\Gamma(1 + \frac{r}{p})\Gamma(q)}{\Gamma(1 + \frac{r}{p} + q)} = qB\left(1 + \frac{r}{p}, q\right), \quad r = 1, 2, \dots,$$

é o  $r$ -ésimo momento, e  $\Gamma(\cdot)$  e  $B(\cdot, \cdot)$  são as funções gama e beta, respectivamente.

Como notado anteriormente, uma vantagem da distribuição Kumaraswamy é possuir função quantil em forma fechada. Tal função é dada por

$$y = F^{-1}(u) = [1 - (1 - u)^{\frac{1}{q}}]^{\frac{1}{p}}, \quad (3.5)$$

em que  $0 < u < 1$ .

É importante destacar algumas características importantes de variáveis aleatórias que seguem distribuição Kumaraswamy. Semelhante ao que ocorre com variáveis aleatórias que seguem distribuição beta, qualquer transformação linear de uma variável aleatória Kumaraswamy também segue distribuição Kumaraswamy e sua distribuição tem os mesmos parâmetros de forma que a distribuição original. Diferentemente de variáveis aleatórias que seguem distribuição beta, variáveis aleatórias que seguem distribuição Kumaraswamy são fechadas sob exponenciação, isto é, para uma variável aleatória que segue distribuição Kumaraswamy, qualquer potência positiva dela também segue distribuição Kumaraswamy. Em suma, temos as propriedades elementares a seguir. Pode-se demonstrar que a distribuição de Kumaraswamy possui algumas propriedades básicas que são características da distribuição beta, a saber:

$$\begin{aligned} p > 1, q > 1 &\Rightarrow \text{unimodal}; & p < 1, q < 1 &\Rightarrow \text{antimodal}; \\ p > 1, q \geq 1 &\Rightarrow \text{crescente}; & p \geq 1, q > 1 &\Rightarrow \text{decrecente}; \\ p = q = 1 &\Rightarrow \text{constante}. \end{aligned}$$

Como notado anteriormente, a distribuição Kumaraswamy é bastante flexível e por este motivo torna-se atrativa para a modelagem de dados. Porém, a expressão para a média dada na Equação (3.4) torna inviável uma reparametrização simples baseada na média da distribuição.

Uma solução foi dada por (MITNIK; BAEK, 2013), que, a partir da Equação (3.5), obtiveram uma expressão para a mediana da distribuição:

$$md(Y) = w = (1 - 0.5^{\frac{1}{q}})^{\frac{1}{p}}. \quad (3.6)$$

Tal resultado fornece uma base para a reparametrização da mediana e da dispersão. Ainda a partir da Equação (3.5) podemos expressar a faixa interquantílica como  $IQR = [(1 - 0.75^{\frac{1}{q}})^{\frac{1}{q}} - (1 - 0.25^{\frac{1}{q}})^{\frac{1}{q}}]$ . (MITNIK, 2013) mostrou que o desvio médio absoluto em torno da mediana é dado por  $\delta_2(y) = \int_0^1 |y - w| f(y) dy$ , que pode também ser expresso por  $\delta_2(y) = [2qB(2^{-\frac{1}{q}}, q, 1 + \frac{1}{p}) - qB(q, 1 + \frac{1}{p})]$ , em que  $B(z, p, q) = \int_0^z S^{p-1}(1 - S)^{q-1} dS$  é a função beta incompleta.

Exibiremos agora as reparametrizações apresentadas por (MITNIK; BAEK, 2013) que decorrem de forma imediata do resultado em (3.6), para os parâmetro de forma  $p = \phi$  e  $q$ . Elas são dadas por

$$q = \frac{\log(0.5)}{\log(1 - w^\phi)} \quad (3.7)$$

e

$$\phi = \frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)}. \quad (3.8)$$

Substituindo as expressões em (3.7) e (3.8) em (3.2), (3.3) e (3.5), obtemos duas reparametrizações da distribuição Kumaraswamy. A primeira reparametrização, denotada por  $K(w, \phi)$ , é dada por

$$f(y; w, \phi) = \frac{\phi \log(0.5)}{\log(1 - w^\phi)} y^{\phi-1} (1 - y^\phi)^{\frac{\log(0.5)}{\log(1 - w^\phi)} - 1}, \quad (3.9)$$

$$F(y) = 1 - (1 - y^\phi)^{\frac{\log(0.5)}{\log(1 - w^\phi)}},$$

$$F^{-1}(u) = y = \left[ 1 - (1 - u)^{\frac{\log(1 - w^\phi)}{\log(0.5)}} \right]^{\frac{1}{\phi}}.$$

A segunda reparametrização, denotada por  $K(w, q)$ , é dada por

$$f(y; w, q) = \frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)} q y^{\frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)} - 1} \left( 1 - y^{\frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)}} \right)^{q-1},$$

$$F(y) = 1 - \left( 1 - y^{\frac{\log(1 - 0.5^{\frac{1}{q}})}{\log(w)}} \right)^q,$$

$$F^{-1}(u) = y = \left[ 1 - (1 - u)^{\frac{1}{q}} \right]^{\frac{\log(w)}{\log(1 - 0.5^{\frac{1}{q}})}}.$$

(MITNIK; BAEK, 2013) mostraram que os parâmetros  $\frac{1}{\phi}$  e  $\frac{1}{q}$  são parâmetros de dispersão. As reparametrizações introduzidas por (MITNIK; BAEK, 2013) possibilitaram o uso da distribuição Kumaraswamy para desenvolver um modelo de regressão que inclui dois submodelos, um para a mediana e outro para a dispersão. Os autores sugerem o uso da primeira reparametrização, por ser mais vantajosa em termo de interpretabilidade dos parâmetros.

Sejam  $Y_1, Y_2, \dots, Y_n$  variáveis aleatórias independentes, em que cada  $Y_t$ ,  $t = 1, 2, \dots, n$ , possui densidade (3.9) com mediana  $w_t$  e parâmetro de precisão  $\phi_t$ . O modelo de regressão Kumaraswamy proposto por (MITNIK; BAEK, 2013) é dado por

$$\begin{aligned} g(w_t) &= \sum_{i=1}^k x_{ti} \beta_i = \eta_{1t}, \\ h(\phi_t) &= \sum_{j=1}^l z_{tj} \gamma_j = \eta_{2t}, \end{aligned} \tag{3.10}$$

em que  $\beta = (\beta_1, \beta_2, \dots, \beta_k)^\top$  e  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_l)^\top$  são vetores de parâmetros desconhecidos ( $\beta \in \mathbb{R}^k$  e  $\gamma \in \mathbb{R}^l$ ),  $x_{t1}, x_{t2}, \dots, x_{tk}$ ,  $z_{t1}, z_{t2}, \dots, z_{tl}$ ,  $t = 1, 2, \dots, n$ , são observações das covariadas fixas e conhecidas ( $k + l < n$ ). As funções  $g(\cdot)$  e  $h(\cdot)$  são funções de ligação, que são estritamente monótonas e duas vezes diferenciáveis,  $g$  tendo domínio no intervalo  $(0,1)$  e imagem no conjunto dos números reais, enquanto  $h$  tem domínio no intervalo  $(0, +\infty)$  e imagem no conjunto dos números reais.

Como exemplos de funções de ligação para o modelo da mediana podemos citar: a logito,  $g(w) = \log(w/(1 - w))$ , a probito,  $g(w) = \Phi^{-1}(w)$ , em que  $\Phi(\cdot)$  é a função de distribuição acumulada normal padrão, a log-log,  $g(w) = -\log(-\log(w))$ , dentre outras. Já para o modelo de precisão citamos  $h(\phi) = \log(\phi)$ .

A função de verossimilhança para o modelo de regressão Kumaraswamy é dada por

$$L(w, \phi) = \prod_{t=1}^n f(y_t; w_t, \phi_t) = \prod_{t=1}^n \frac{\phi_t \log(0.5)}{\log(1 - w_t^{\phi_t})} y_t^{\phi_t - 1} (1 - y_t^{\phi_t})^{\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1}.$$

A função de log-verossimilhança é

$$\ell(\beta, \gamma) = \sum_{t=1}^n \ell_t(w_t, \phi_t), \tag{3.11}$$

em que

$$\begin{aligned} \ell_t(w_t, \phi_t) &= \log(\phi_t) + \log\left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})}\right) + (\phi_t - 1) \log(y_t) \\ &\quad + \left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right) \log(1 - y_t^{\phi_t}). \end{aligned}$$

A seguir obteremos a função escore através da diferenciação da função de log-verossimilhança com relação aos seus parâmetros, representada por  $U(\beta, \gamma) = \left(\frac{\partial \ell}{\partial \beta^\top}, \frac{\partial \ell}{\partial \gamma^\top}\right)^\top$ , que é um vetor de dimensão  $(k + l) \times 1$ . A função escore para o parâmetro  $\beta$  é dada por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial \ell_t(w_t, \phi_t)}{\partial w_t} \frac{dw_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \quad i = 1, 2, \dots, k, \quad (3.12)$$

com

$$\frac{\partial \ell_t(w_t, \phi_t)}{\partial w_t} = \frac{\phi_t w_t^{\phi_t}}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} + \frac{\log(0.5) \phi_t w_t^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})}.$$

A Equação (3.12) reduz-se a

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n (w_t^* + y_t^*) \frac{1}{g'(w_t)} x_{ti},$$

em que

$$w_t^* = \frac{\phi_t w_t^{\phi_t}}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} \text{ e } y_t^* = \frac{\log(0.5) \phi_t w_t^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})}. \quad (3.13)$$

Adicionalmente, verificamos no Apêndice A que

$$\mathbb{E}\left(\frac{\partial \ell_t(w_t, \phi_t)}{\partial w_t}\right) = 0 \Rightarrow \mathbb{E}(w_t^* + y_t^*) = 0 \Rightarrow \mathbb{E}(y_t^*) = -w_t^*.$$

Em forma matricial, a função escore para  $\beta$  é um vetor de dimensão  $k \times 1$  que é dado por

$$U_\beta(\beta, \phi) = X^\top T(w^* + y^*),$$

em que  $X$  é a matriz de covariadas de ordem  $n \times k$ ,  $T = \text{diag}(1/g'(w_1), 1/g'(w_2), \dots, 1/g'(w_n))$ , uma matriz de ordem  $n$ , e  $y^* = (y_1^*, y_2^*, \dots, y_n^*)^\top$  e  $w^* = (w_1^*, w_2^*, \dots, w_n^*)^\top$  são vetores de ordem  $n \times 1$ .

A  $j$ -ésima componente do vetor escore de  $\gamma$  é

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n \frac{\partial \ell_t(w_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j}, \quad j = 1, 2, \dots, l, \quad (3.14)$$

com

$$\begin{aligned} \frac{\partial \ell_t(w_t, \phi_t)}{\partial \phi_t} &= \frac{1}{\phi_t} + \frac{\log(0.5)w_t^{\phi_t} \log(w_t) \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2(1 - w_t^{\phi_t})} + \log(y_t) \\ &+ \frac{w_t^{\phi_t} \log(w_t)}{\log(1 - w_t^{\phi_t})(1 - w_t^{\phi_t})} - \frac{\left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right) y_t^{\phi_t} \log(y_t)}{1 - y_t^{\phi_t}}. \end{aligned} \quad (3.15)$$

A Expressão em (3.15) pode ser reescrita como

$$\frac{\partial \ell_t(w_t, \phi_t)}{\partial \phi_t} = \sum_{t=1}^n [(y_t^* + w_t^*)\alpha_{1t} + \alpha_{2t}],$$

em que

$$\alpha_{1t} = \frac{w_t \log(w_t)}{\phi_t},$$

$$\alpha_{2t} = \frac{1}{\phi_t} + \log(y_t) - \frac{\left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right) y_t^{\phi_t} \log(y_t)}{1 - y_t^{\phi_t}}.$$

Assim, a Equação (3.14) pode ser reescrita como

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n [(y_t^* + w_t^*)\alpha_{1t} + \alpha_{2t}] \frac{1}{h'(\phi_t)} z_{tj}.$$

Em forma matricial, a função escore para  $\gamma$  é um vetor de dimensão  $l \times 1$  que pode ser expresso como

$$U_\gamma(\beta, \gamma) = Z^\top H[(w^* + y^*)\alpha_1 + \alpha_2],$$

em que  $Z$  é a matriz de covariadas de ordem  $n \times l$ ,  $H = \text{diag}(1/h'(\phi_1), 1/h'(\phi_2), \dots, 1/h'(\phi_n))$ ,  $y^* = (y_1^*, y_2^*, \dots, y_n^*)^\top$  e  $w^* = (w_1^*, w_2^*, \dots, w_n^*)^\top$ ,  $\alpha_1 = (\alpha_{11}, \alpha_{12}, \dots, \alpha_{1n})^\top$  e  $\alpha_2 = (\alpha_{21}, \alpha_{22}, \dots, \alpha_{2n})^\top$ .

Os estimadores de máxima verossimilhança para os parâmetros  $\beta$  e  $\gamma$  são obtidos resolvendo o sistema de equações não lineares dado por  $U = 0$ , que não possui solução em forma fechada. Portanto, utilizam-se algoritmos de otimização não-linear, tais como o algoritmos de Newton e quase-Newton, para encontrar as estimativas de máxima verossimilhança, através da maximização numérica da função de log-verossimilhança. Para tanto, há a necessidade de especificar valores iniciais para serem utilizados no esquema iterativo. Seguindo a ideia de (FERRARI; CRIBARI-NETO, 2004), propomos como valores iniciais para  $\beta$  a estimativa de mínimos quadrados ordinários obtidas de uma regressão linear das respostas transformadas pela função de ligação do modelo,  $(g(y_1), g(y_2), \dots, g(y_n))^\top$  sobre  $X$ , isto é,

$$\beta^{(0)} = (X^\top X)^{-1} X^\top g(Y).$$

Por outro lado, de (JONES, 2009) sabemos que o estimador de máxima verossimilhança  $\widehat{\phi}_t$  de  $\phi_t$  é obtido resolvendo a equação

$$f(\phi_t) = \frac{1}{\phi_t} \left( 1 + T_1(\phi_t) + \frac{T_2(\phi_t)}{T_3(\phi_t)} \right) = 0,$$

em que

$$T_1(\phi_t) = \frac{\log(y_t^{\phi_t})}{1 - y_t^{\phi_t}}, \quad T_2(\phi_t) = \frac{y_t^{\phi_t} \log(y_t^{\phi_t})}{1 - y_t^{\phi_t}} \quad \text{e} \quad T_3(\phi_t) = \log(1 - y_t^{\phi_t}).$$

Não é possível expressar a solução da equação acima em forma analítica. Considerando o método iterativo de Newton, temos

$$\phi_t^{(m+1)} = \phi_t^{(m)} - \frac{f(\phi_t^{(m)})}{f'(\phi_t^{(m)})}, \quad m \in \mathbb{N},$$

em que  $m$  indexa as iterações do algoritmo. Tomando  $\phi_t^{(0)} = 1$ , temos

$$\phi_t^{(1)} = 1 - \frac{f(1)}{f'(1)}.$$

Propomos como valores iniciais para  $\gamma$  a estimativa de uma regressão linear de  $\phi_t^{(1)}$  transformada pela função de ligação para o modelo de precisão,  $(h(\phi_1^{(1)}), h(\phi_2^{(1)}), \dots, h(\phi_n^{(1)}))^T$ , sobre  $Z$ , isto é,

$$\gamma^{(0)} = (Z^T Z)^{-1} Z h(\phi^{(1)}).$$

O passo seguinte é a obtenção da expressão da matriz de informação observada. Neste sentido, são necessárias as derivadas de segunda ordem do logaritmo da função de verossimilhança em relação aos parâmetros. A partir da Equação (3.12), a derivada de segunda ordem de  $\ell(\beta, \gamma)$  em relação a  $\beta_i$  e  $\beta_j$ , com  $i, j = 1, 2, \dots, k$ , é dada por

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_j \partial \beta_i} &= \sum_{t=1}^n \frac{\partial}{\partial w_t} \left[ \frac{\partial \ell(w_t, \phi_t)}{\partial w_t} \frac{\partial w_t}{\partial \eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \right] \frac{\partial w_t}{\partial \eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_j} \\ &= \sum_{t=1}^n \frac{\partial^2 \ell(w_t, \phi_t)}{\partial w_t^2} \left( \frac{\partial w_t}{\partial \eta_{1t}} \right)^2 \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i} + \sum_{t=1}^n \frac{\partial \ell(w_t, \phi_t)}{\partial w_t} \left[ \frac{\partial^2 w_t}{\partial \eta_{1t}^2} \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i} \right], \end{aligned}$$

em que

$$\begin{aligned} \frac{\partial^2 \ell(w_t, \phi_t)}{\partial w_t^2} &= \frac{w_t^{\phi_t} \phi_t^2}{w_t^2 \log(1 - w_t^{\phi_t})(1 - w_t^{\phi_t})} + \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1 - w_t^{\phi_t})^2 w_t^2 (1 - w_t^{\phi_t})^2} \\ &\quad - \frac{w_t^{\phi_t} \phi_t}{\log(1 - w_t^{\phi_t}) w_t^2 (1 - w_t^{\phi_t})} + \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1 - w_t^{\phi_t}) w_t^2 (1 - w_t^{\phi_t})^2} \\ &\quad + \frac{\log(0.5) w_t^{\phi_t} \phi_t^2 \log(1 - y_t^{\phi_t})}{w^2 \log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})} + \frac{2 \log(0.5) (w_t^{\phi_t})^2 \phi_t^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^3 w_t^2 (1 - w_t^{\phi_t})^2} \\ &\quad - \frac{\log(0.5) w_t^{\phi_t} \phi_t \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t^2 (1 - w_t^{\phi_t})} + \frac{\log(0.5) (w_t^{\phi_t})^2 \phi_t^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t^2 (1 - w_t^{\phi_t})^2}. \end{aligned}$$

Em forma matricial, temos a seguinte expressão:

$$\ddot{L}_{\beta\beta} = X^\top S_1 X + X^\top S X, \quad (3.16)$$

em que

$$S_1 = \text{diag} \left\{ \left( \frac{\partial^2 \ell(w_1, \phi_1)}{\partial w_1^2} \left( \frac{\partial w_1}{\partial \eta_{11}} \right)^2, \dots, \frac{\partial^2 \ell(w_n, \phi_n)}{\partial w_n^2} \left( \frac{\partial w_n}{\partial \eta_{1n}} \right)^2 \right) \right\}$$

e

$$S = \text{diag} \left\{ \left( \frac{\partial \ell(w_1, \phi_1)}{\partial w_1} \frac{\partial^2 w_1}{\partial \eta_{11}^2}, \dots, \frac{\partial \ell(w_n, \phi_n)}{\partial w_n} \frac{\partial^2 w_n}{\partial \eta_{1n}^2} \right) \right\}$$

A partir da Equação (3.12), a derivada de segunda ordem de  $\ell(\beta, \gamma)$  em relação a  $\beta_i$  e  $\gamma_j$ , com  $i = 1, 2, \dots, k$  e  $j = 1, 2, \dots, l$ , é dada por

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_j \partial \beta_i} &= \sum_{t=1}^n \frac{\partial}{\partial \gamma_j} \left( \frac{\partial \ell(w_t, \phi_t)}{\partial w_t} \frac{\partial w_t}{\partial \eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \right) \\ &= \sum_{t=1}^n \frac{\partial^2 \ell(w_t, \phi_t)}{\partial \phi_t \partial w_t} \frac{\partial \phi_t}{\partial \eta_{2t}} \frac{\partial w_t}{\partial \eta_{1t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} \frac{\partial \eta_{1t}}{\partial \beta_i}, \end{aligned}$$

em que

$$\begin{aligned} \frac{\partial^2 \ell(w_t, \phi_t)}{\partial \phi_t \partial w_t} &= \frac{w_t^{\phi_t} \log(w_t) \phi_t}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} + \frac{w_t^{\phi_t}}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} \\ &+ \frac{(w_t^{\phi_t})^2 \phi_t \log(w_t)}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})^2} + \frac{(w_t^{\phi_t})^2 \phi_t \log(w_t)}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})^2} \\ &+ \frac{\log(0.5) w_t^{\phi_t} \log(w_t) \phi_t \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})} + \frac{\log(0.5) w_t^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})} \\ &- \frac{\log(0.5) w_t^{\phi_t} \phi_t y_t^{\phi_t} \log(y_t)}{(1 - y_t^{\phi_t}) \log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})} \\ &+ \frac{2 \log(0.5) (w_t^{\phi_t})^2 \phi_t \log(1 - y_t^{\phi_t}) \log(w_t)}{\log(1 - w_t^{\phi_t})^3 w_t (1 - w_t^{\phi_t})^2} \\ &+ \frac{\log(0.5) (w_t^{\phi_t})^2 \phi_t \log(1 - y_t^{\phi_t}) \log(w_t)}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})^2}. \end{aligned}$$

Em forma matricial, temos a seguinte expressão:

$$\ddot{L}_{\beta\phi} = X^\top A Z^\top, \quad (3.17)$$

em que  $A = \text{diag} \left\{ \left( \frac{\partial^2 \ell(w_1, \phi_1)}{\partial \phi_1 \partial w_1} \frac{\partial w_1}{\partial \eta_{11}} \frac{\partial \phi_1}{\partial \eta_{21}}, \dots, \frac{\partial^2 \ell(w_n, \phi_n)}{\partial \phi_n \partial w_n} \frac{\partial w_n}{\partial \eta_{1n}} \frac{\partial \phi_n}{\partial \eta_{2n}} \right) \right\}$ .

A partir da Equação (3.14), a derivada de segunda ordem de  $\ell(\beta, \gamma)$  em relação a  $\gamma_i$ ,  $\gamma_j$ , em que  $i, j = 1, 2, \dots, l$ , é da seguinte forma:

$$\frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_j \partial \gamma_i} = \sum_{t=1}^n \frac{\partial}{\partial \gamma_t} \left( \frac{\partial \ell(w_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_i} \right)$$

$$= \sum_{t=1}^n \frac{\partial^2 \ell(w_t, \phi_t)}{\partial \phi_t^2} \left( \frac{\partial \phi_t}{\partial \eta_{2t}} \right)^2 \frac{\partial \eta_{2t}}{\partial \gamma_j} \frac{\partial \eta_{2t}}{\partial \gamma_i} + \sum_{t=1}^n \frac{\partial \ell(w_t, \phi_t)}{\partial \phi_t} \frac{\partial^2 \phi_t}{\partial \eta_{2t}^2} \frac{\partial \eta_{2t}}{\partial \gamma_j} \frac{\partial \eta_{2t}}{\partial \gamma_i},$$

em que

$$\begin{aligned} \frac{\partial^2 \ell(w_t, \phi_t)}{\partial \phi_t^2} &= \frac{w_t^{\phi_t} \log(w_t)^2}{\log(1 - w_t^{\phi_t})(1 - w_t^{\phi_t})} - \frac{2 \log(0.5) w_t^{\phi_t} \log(w_t) y_t^{\phi_t} \log(y_t)}{(1 - y_t^{\phi_t}) \log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})} \\ &+ \frac{(w_t^{\phi_t})^2 \log(w_t)^2}{\log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})^2} + \frac{\log(0.5) w_t^{\phi_t} \log(w_t)^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})} \\ &+ \frac{(w_t^{\phi_t})^2 \log(w_t)^2}{\log(1 - w_t^{\phi_t})(1 - w_t^{\phi_t})^2} + \frac{2 \log(0.5) (w_t^{\phi_t})^2 \log(w_t)^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^3 (1 - w_t^{\phi_t})^2} \\ &+ \frac{\log(0.5) (w_t^{\phi_t})^2 \log(w_t)^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t})^2} - \frac{1}{\phi_t^2} \\ &- \frac{\left[ \frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1 \right] y_t^{\phi_t} \log(y_t)^2}{1 - y_t^{\phi_t}} - \frac{\left[ \frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1 \right] (y_t^{\phi_t})^2 \log(y_t)^2}{(1 - y_t^{\phi_t})^2}. \end{aligned}$$

Portanto,

$$\ddot{L}_{\phi\phi} = Z^\top B_1 Z + Z^\top B Z, \quad (3.18)$$

em que

$$B_1 = \text{diag} \left\{ \left( \frac{\partial^2 \ell(w_1, \phi_1)}{\partial \phi_1^2} \left( \frac{\partial \phi_1}{\partial \eta_{21}} \right)^2, \dots, \frac{\partial^2 \ell(w_n, \phi_n)}{\partial \phi_n^2} \left( \frac{\partial \phi_n}{\partial \eta_{2n}} \right)^2 \right) \right\}$$

e

$$B = \text{diag} \left\{ \left( \frac{\partial \ell(w_1, \phi_1)}{\partial \phi_1} \frac{\partial^2 \phi_1}{\partial \eta_{21}^2}, \dots, \frac{\partial \ell(w_n, \phi_n)}{\partial \phi_n} \frac{\partial^2 \phi_n}{\partial \eta_{2n}^2} \right) \right\}.$$

Finalmente, das expressões em (3.16), (3.17) e (3.18) obtemos a matriz de informação observada, que é dada por

$$\begin{aligned} \ddot{L} &= \begin{bmatrix} -\ddot{L}_{\beta\beta} & -\ddot{L}_{\beta\gamma} \\ -\ddot{L}_{\gamma\beta} & -\ddot{L}_{\gamma\gamma} \end{bmatrix} \\ &= \begin{bmatrix} X^\top S_1 X + X^\top S X & X^\top A Z \\ Z^\top A X & Z^\top B_1 Z + Z^\top B Z \end{bmatrix}, \end{aligned}$$

em que cada uma das matrizes acima já foi definida anteriormente.

### 3.3 Modelo de regressão beta

Nos modelos de regressão beta há a suposição de que a variável resposta segue distribuição beta, que é muito flexível, sua densidade pode apresentar diferentes formas quando os valores dos parâmetros variam. A variável aleatória  $Y$  segue distribuição beta se a sua função de densidade é dada por

$$f(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1} (1-y)^{q-1}, \quad 0 < y < 1,$$

em que  $p, q > 0$  são parâmetros de forma e  $\Gamma(\cdot)$  é a função gama.

A média e a variância de  $y$  são

$$\begin{aligned} E(Y) &= \frac{p}{p+q}, \\ \text{Var}(Y) &= \frac{pq}{(p+q)^2(p+q+1)}. \end{aligned}$$

Nos modelos de regressão clássicos o usual é modelar a média e definir um modelo que contenha o parâmetro de precisão ou dispersão. Neste sentido, o modelo de regressão beta proposto por (FERRARI; CRIBARI-NETO, 2004) utiliza uma parametrização diferente para a distribuição beta, dada por  $\mu = p/(p+q)$  e  $\phi = p+q$ , ou seja,  $p = \mu\phi$  e  $q = (1-\mu)\phi$ . Neste caso a função densidade beta é expressa da seguinte forma:

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1, \quad (3.19)$$

em que  $0 < \mu < 1$  e  $\phi > 0$ . Assim, temos

$$\begin{aligned} E(Y) &= \mu, \\ \text{Var}(Y) &= \frac{V(\mu)}{(1+\phi)}, \end{aligned}$$

em que  $V(\mu) = \mu(1-\mu)$  é a função de variância.

Note que uma vez fixada a média da variável resposta, quanto maior for o valor de  $\phi$ , menor será a variância de  $Y$ . Assim,  $\phi$  pode ser interpretado como um parâmetro de precisão do modelo e  $\phi^{-1}$  é o parâmetro de dispersão.

No modelo de regressão beta proposto por (FERRARI; CRIBARI-NETO, 2004) o parâmetro de dispersão é assumido ser constante ao longo das observações. Similar ao que ocorre em modelos lineares generalizados, o conceito de heteroscedasticidade e dispersão não constante diferem do utilizado nos modelos lineares normais. Neste sentido, (SMITHSON; VERKUILEN, 2006) consideram um modelo de regressão com dispersão variável, isto é, a média e a dispersão são modeladas simultaneamente.

Considerando  $Y_1, Y_2, \dots, Y_n$  variáveis aleatórias independentes, em que cada  $Y_t$ ,  $t = 1, 2, \dots, n$ , segue densidade (3.19), a média e o parâmetro de precisão satisfazem as seguintes relações:

$$g(\mu_t) = \sum_{i=1}^k x_{ti} \beta_i = \eta_{1t}, \quad (3.20)$$

$$h(\phi_t) = \sum_{j=1}^q z_{tj} \gamma_j = \eta_{2t}, \quad (3.21)$$

em que  $\beta = (\beta_1, \beta_2, \dots, \beta_k)^\top$  e  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_q)^\top$  são vetores de parâmetros desconhecidos  $\beta \in \mathbb{R}^k$  e  $\gamma \in \mathbb{R}^q$ ,  $x_{t1}, x_{t2}, \dots, x_{tk}$  e  $z_{t1}, z_{t2}, \dots, z_{tq}$  são observações de  $k$  e  $q$  covariáveis fixas e conhecidas em que  $k < n$  e  $q < n$ ,  $g(\cdot)$  e  $h(\cdot)$  são funções de ligação estritamente monótonas e duas diferenciáveis.

A estimação dos parâmetros pode ser realizada a partir do método de máxima verossimilhança. Assim, baseando-se em uma amostra com  $n$  observações independentes, o logaritmo da função de verossimilhança é dado por

$$\ell(\beta, \gamma) = \sum_{t=1}^n \ell(\mu_t, \phi_t), \quad (3.22)$$

em que

$$\begin{aligned} \ell(\mu_t, \phi_t) &= \log(\Gamma(\phi_t) - \log \Gamma(\mu_t \phi_t) - \log \Gamma((1 - \mu_t) \phi_t) + (\mu_t \phi_t - 1) \log(y_t) \\ &+ [(1 - \mu_t) \phi_t - 1] \log(1 - y_t). \end{aligned}$$

O vetor escore é dado por  $U(\beta, \gamma) = (\frac{\partial \ell}{\partial \beta}, \frac{\partial \ell}{\partial \gamma})^\top$ . Os componentes do vetor escore de  $\beta$  são da forma

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \quad i = 1, 2, \dots, k, \quad (3.23)$$

em que

$$\frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} = \phi_t \left[ \log \left( \frac{y_t}{1 - y_t} \right) - [\psi(\mu_t \phi_t) - \psi((1 - \mu_t) \phi_t)] \right],$$

$\psi(\cdot)$  sendo a função digama. Portanto, a Equação (3.23) reduz-se a

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} x_{ti}, \quad (3.24)$$

em que

$$y_t^* = \log \left( \frac{y_t}{1 - y_t} \right) \quad \text{e} \quad \mu_t^* = \psi(\mu_t \phi_t) - \psi((1 - \mu_t) \phi_t).$$

Em forma matricial a função escore para o parâmetro  $\beta$  é dada pela seguinte expressão:

$$U_\beta(\beta, \gamma) = X^\top \Phi T (y^* - \mu^*), \quad (3.25)$$

em que  $X$  é a matriz de covariadas de ordem  $n \times k$ ,  $\Phi = \text{diag}(\phi_1, \phi_2, \dots, \phi_n)$  e  $T = \text{diag}(1/g'(w_1), 1/g'(w_2), \dots, 1/g'(w_n))$  são matrizes de ordem  $n \times n$ , e  $y^* = (y_1^*, y_2^*, \dots, y_n^*)^\top$  e  $\mu^* = (\mu_1^*, \mu_2^*, \dots, \mu_n^*)^\top$  são vetores de ordem  $n \times 1$ .

A função escore para  $\gamma_j$ ,  $j = 1, 2, \dots, n$  é dada por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j}, \quad j = 1, 2, \dots, l, \quad (3.26)$$

com

$$\begin{aligned} \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} &= \mu_t \left[ \log \left( \frac{y_t}{1 - y_t} \right) - \psi(\mu_t \phi_t) - \psi((1 - \mu_t) \phi_t) \right] \\ &+ \log(1 - y_t) - \psi((1 - \mu_t) \phi_t) + \psi(\phi_t). \end{aligned}$$

Adicionalmente,  $\partial \phi_t / \partial \eta_{2t} = 1/h'(\phi_t)$  e  $\partial \eta_{2t} / \partial \gamma_j = z_{tj}$ .

Utilizando as definições de  $y^*$  e  $\mu^*$ , a função escore para cada um dos parâmetros  $\gamma_j$  é dada por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n [\mu_t(y_t^* - \mu_t^*) + \log(1 + y_t) - \psi((1 - \mu_t) \phi_t) + \psi(\phi_t)] \frac{1}{h'(\phi_t)} z_{tj},$$

que pode ser expressa por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^n a_t \frac{1}{h'(\phi_t)} z_{tj},$$

em que

$$a_t = \mu_t(y_t^* - \mu_t^*) + \log(1 + y_t) - \psi((1 - \mu_t) \phi_t) + \psi(\phi_t). \quad (3.27)$$

Em forma matricial temos

$$U_\gamma(\beta, \gamma) = Z^\top H a,$$

em que  $Z$  é uma matriz de ordem  $n \times q$ ,  $H = \text{diag}(1/h'(\phi_1), 1/h'(\phi_2), \dots, 1/h'(\phi_n))$  é uma matriz de ordem  $n \times n$  e  $a = (a_1, a_2, \dots, a_n)^\top$  é um vetor  $n \times 1$ .

Para obtermos a matriz de informação observada são necessárias as segundas derivadas do logaritmo da função de verossimilhança com relação aos parâmetros  $\beta$  e  $\gamma$ . A partir da Equação (3.23), temos que, para  $i, p = 1, 2, \dots, k$ ,

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_p \partial \beta_i} &= \sum_{t=1}^n \frac{\partial}{\partial \mu_t} \left[ \frac{\partial \ell(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} \right] \frac{\partial \mu_t}{\partial \eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_p} \\ &= \sum_{t=1}^n \frac{\partial^2 \ell(\mu_t, \phi_t)}{\partial \mu_t^2} \left( \frac{\partial \mu_t}{\partial \eta_{1t}} \right)^2 \frac{\partial \eta_{1t}}{\partial \beta_p} \frac{\partial \eta_{1t}}{\partial \beta_i} + \sum_{t=1}^n \frac{\partial \ell(\mu_t, \phi_t)}{\partial \mu_t} \left[ \frac{\partial^2 \mu_t}{\partial \eta_{1t}^2} \frac{\partial \eta_{1t}}{\partial \beta_p} \frac{\partial \eta_{1t}}{\partial \beta_i} \right], \end{aligned}$$

em que

$$\frac{\partial^2 \ell(\mu_t, \phi_t)}{\partial \mu_t^2} = \phi_t^2 [\psi'(\mu_t \phi_t) + \psi'((1 - \mu_t) \phi_t)].$$

Em forma matricial, temos

$$\ddot{L}_{\beta\beta} = X^\top \Phi Q X, \quad (3.28)$$

em que  $Q = \text{diag}(q_1, q_2, \dots, q_n)$ , com

$$q_t = \left[ \phi_t [\psi'(\mu_t \phi_t) + \psi'((1 - \mu_t) \phi_t)] + (y_t^* - \mu_t^*) \frac{g''(\mu_t)}{g'(\mu_t)} \right] \frac{1}{g'(\mu_t)^2}. \quad (3.29)$$

Da Equação (3.24), a segunda derivada de  $\ell(\beta, \gamma)$  em relação a  $\beta_i$  e  $\gamma_j$ , para  $i = 1, 2, \dots, k$  e  $j = 1, 2, \dots, q$ , é dada por

$$\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \gamma_j} = \sum_{t=1}^n \frac{\partial}{\partial \phi_t} \left( \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} x_{ti} \right) \frac{\partial \phi_t}{\partial \eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j},$$

em que

$$\frac{\partial}{\partial \phi_t} \left( \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} x_{ti} \right) = [(y_t^* - \mu_t^*) - \phi_t [\psi'(\mu_t \phi_t) \mu_t - \psi'((1 - \mu_t) \phi_t) (1 - \mu_t)]] \frac{1}{g'(\mu_t)} x_{ti},$$

$\partial \phi_t \partial \eta_{2t} = 1/h'(\phi_t)$  e  $\partial \eta_{2t} / \partial \eta_{2t} = z_{tj}$ .

Em forma matricial, temos a seguinte expressão:

$$\ddot{L}_{\beta\phi} = X^\top F T H Z^\top, \quad (3.30)$$

em que  $F = \text{diag}(f_1, f_2, \dots, f_n)$ , sendo que

$$f_t = (y_t^* - \mu_t^*) - \phi_t [\psi'(\mu_t \phi_t) \mu_t - \psi'((1 - \mu_t) \phi_t) (1 - \mu_t)]. \quad (3.31)$$

A partir da Equação (3.30), a derivada de  $\ell(\beta, \gamma)$  com relação a  $\gamma_j$  e  $\gamma_l$ , para  $j, l = 1, 2, \dots, q$ , é dada por

$$\begin{aligned} \frac{\partial^2 \ell(\beta, \gamma)}{\partial \gamma_j \partial \gamma_l} &= \sum_{t=1}^n \frac{\partial}{\partial \gamma_t} \left( \frac{\partial \ell(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_i} \right) \\ &= \sum_{t=1}^n \frac{\partial^2 \ell(\mu_t, \phi_t)}{\partial \phi_t^2} \left( \frac{\partial \phi_t}{\partial \eta_{2t}} \right)^2 \frac{\partial \eta_{2t}}{\partial \gamma_j} \frac{\partial \eta_{2t}}{\partial \gamma_l} + \sum_{t=1}^n \frac{\partial \ell(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial^2 \phi_t}{\partial \eta_{2t}^2} \frac{\partial \eta_{2t}}{\partial \gamma_j} \frac{\partial \eta_{2t}}{\partial \gamma_l}, \end{aligned}$$

em que

$$\frac{\partial^2 \ell(\mu_t, \phi_t)}{\partial \phi_t^2} = -[\psi'(\mu_t \phi_t) \mu_t^2 + \psi'((1 - \mu_t) \phi_t) (1 - \mu_t)^2 - \psi'(\phi_t)].$$

Em forma matricial, temos

$$\ddot{L}_{\phi\phi} = Z^\top V Z, \quad (3.32)$$

em que  $V = \text{diag}(v_1, v_2, \dots, v_n)$ , sendo que

$$v_t = d_t + a_t \frac{h''(\phi_t)}{(h'(\phi_t))^3}, \quad (3.33)$$

em que  $d_t = [\psi'(\mu_t \phi_t) \mu_t^2 + \psi'((1 - \mu_t) \phi_t) (1 - \mu_t)^2 - \psi'(\phi_t)] \frac{1}{(h'(\phi_t))^2}$ .

Finalmente, de (3.28), (3.30) e (3.32) obtemos a matriz de informação observada, que é dada por

$$\ddot{L} = \begin{bmatrix} -\ddot{L}_{\beta\beta} & -\ddot{L}_{\beta\gamma} \\ -\ddot{L}_{\gamma\beta} & -\ddot{L}_{\gamma\gamma} \end{bmatrix}$$

$$= \begin{bmatrix} X^\top \Phi Q X & X^\top F T H Z \\ Z^\top F T H X & Z^\top V Z \end{bmatrix},$$

em que cada uma das matrizes acima já foi definida anteriormente.

### 3.4 Influência global

Como mencionado na Introdução, o primeiro passo na análise de sensibilidade envolve a avaliação de influências globais, de acordo com a eliminação de casos. A eliminação de caso é uma abordagem comum para estudar o efeito de abandonar o  $i$ -ésimo caso a partir do conjunto de dados.

No que segue, uma quantidade com índice “ $(i)$ ” indica cálculo de quantidades com a  $i$ -ésima observação excluída. Para o modelo anterior, a função de log-verossimilhança é denotada por  $\ell_{(i)}(\theta)$ .

Seja  $\hat{\theta}_{(i)} = (\hat{\beta}_{(i)}, \hat{\gamma}_{(i)})^\top$  o estimador de máxima verossimilhança obtido a partir da maximização  $\ell_{(i)}(\theta)$ . Para avaliar a influência da  $i$ -ésima observação sobre o EMV, comparamos a diferença entre  $\hat{\theta}_{(i)}$  e  $\hat{\theta}$ . Se a eliminação de uma observação influencia acentuadamente uma estimativa, mais atenção deve ser dada a essa observação. Portanto, se  $\hat{\theta}_{(i)}$  está longe de  $\hat{\theta}$ , consideramos que a  $i$ -ésima observação é uma observação influente. Uma medida inicial de influência global é definida como a norma padronizada de  $\hat{\theta}_{(i)}$ , que é conhecida como a distância de Cook generalizada e é dada pela seguinte expressão:

$$GD_i(\theta) = (\hat{\theta}_{(i)} - \hat{\theta})^\top [-\ddot{L}](\hat{\theta}_{(i)} - \hat{\theta}),$$

em que, a matriz  $\ddot{L}$  é avaliada no estimador de máxima verossimilhança.

Podemos calcular separadamente  $GD_i(\beta)$  e  $GD_i(\gamma)$  para avaliar o impacto da  $i$ -ésima observação sobre a estimação de  $\beta$  e  $\gamma$ , respectivamente. Outra medida popular é a da diferença entre função de log-verossimilhanças avaliada em  $\hat{\theta}_{(i)}$  e em  $\hat{\theta}$ :

$$LD_i(\zeta) = 2[l(\hat{\theta}) - l(\hat{\theta}_{(i)})].$$

Adicionalmente, podemos calcular  $\hat{\beta} - \hat{\beta}_{(i)}$  e  $\hat{\gamma} - \hat{\gamma}_{(i)}$ . Outras medidas de influência global podem ser definidas.

### 3.5 Influência local

O método de influência local é recomendado quando há interesse em investigar a sensibilidade do ajuste do modelo a pequenas perturbações no modelo ou nos dados. Seja  $v$  um vetor de perturbações  $m$ -dimensional. Denotamos a função de log-verossimilhança perturbada por  $\ell(\theta|v)$ . Suponha que existe um vetor de não perturbação  $v_0$  tal que  $\ell(\theta|v_0) =$

$\ell(\theta)$ . A influência de pequenas perturbações no estimador de máxima verossimilhança  $\hat{\theta}$  pode ser avaliada usando o deslocamento da log-verossimilhança:

$$LD_v = 2(\ell(\hat{\theta}) - \ell(\hat{\theta}_v)), \quad (3.34)$$

em que  $\hat{\theta}_v$  denota o estimador de máxima verossimilhança de  $\theta$  para o modelo perturbado.

A ideia de (COOK, 1986) para avaliar a influência local é essencialmente analisar o comportamento local de  $LD_v$  em torno de  $v_0$  avaliando a curvatura da curva de  $LD_{v_0+ad}$ , em que  $a \in \mathbb{R}$  e  $d$  é a direção normal unitária. Uma das medidas de particular interesse é a direção  $d_{\max}$  correspondente à maior curvatura  $C_{d_{\max}}$ . O vetor  $d_{\max}$  pode indicar se a observação a ele correspondente exerce influência considerável sobre  $LD_v$  sob perturbações pequenas. Adicionalmente, o gráfico de  $d_{\max}$  em relação aos valores de variáveis independentes pode ser útil para identificar padrões atípicos. (COOK, 1986) mostrou que a curvatura normal na direção de  $d$  é dada pela seguinte expressão:

$$C_d(\theta) = 2|d^\top \Delta^\top (\ddot{L})^{-1} \Delta d|,$$

em que  $\Delta = \frac{\partial^2 \ell(\theta|v)}{\partial \theta \partial v^\top}$ ,  $\Delta$  e  $\ddot{L}$  sendo avaliados em  $\hat{\theta}$  e em  $v_0$ . É possível mostrar que,  $C_{d_{\max}}$  é o maior autovalor de  $C = -\Delta^\top \ddot{L}^{-1} \Delta$  e  $d_{\max}$  é o autovetor normal unitário correspondente. A direção  $d_{\max}$  correspondente à matriz  $C$  mostra-nos como perturbar o modelo ou dados para se obter maior mudança local no deslocamento de log-verossimilhança.

Se o interesse reside em calcular a influência para  $\beta$ , a curvatura normal na direção do vetor  $d$  é

$$C_{d;\beta}(\theta) = -2|d^\top \Delta^\top (\ddot{L}^{-1} - \ddot{L}_{\gamma\gamma}) \Delta d|,$$

em que  $\ddot{L}_{\gamma\gamma} = \begin{pmatrix} 0 & 0 \\ 0 & \ddot{L}_{\gamma\gamma}^{-1} \end{pmatrix}$ ,  $d_{\max;\beta}$  sendo o autovetor normal unitário correspondente ao maior autovalor de  $C_1 = -\Delta^\top (\ddot{L}^{-1} - \ddot{L}_{\gamma\gamma}) \Delta$ .

De maneira análoga, quando o interesse reside no cálculo da influência local para  $\gamma$ , a curvatura normal na direção do vetor  $d$  é dada por

$$C_{d;\gamma}(\theta) = 2|d^\top \Delta^\top (\ddot{L}^{-1} - \ddot{L}_{\beta\beta}) \Delta d|,$$

em que  $\ddot{L}_{\beta\beta} = \begin{pmatrix} \ddot{L}_{\beta\beta}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ ,  $d_{\max;\gamma}$  sendo o autovetor normal unitário correspondente ao autovalor de  $C_2 = -\Delta^\top (\ddot{L}^{-1} - \ddot{L}_{\beta\beta}) \Delta$ .

Outra ferramenta que pode ser usada é a curvatura local total correspondente ao  $i$ -ésimo elemento, tomando  $d_i$  como sendo o vetor  $n \times 1$  de zeros com 1 na  $i$ -ésima coordenada. A curvatura na direção de  $d_i$  assume a forma

$$C_i(\theta) = 2|\Delta_i^\top (\ddot{L}^{-1}) \Delta_i|,$$

em que  $\Delta_i$  denota a  $i$ -ésima linha de  $\Delta$ . Note que  $C_i(\theta)$  é o  $i$ -ésimo elemento diagonal da matriz de influência  $C$  previamente definida. Isto é conhecido como influência local total. É também possível calcular a influência local total do  $i$ -ésimo termo ao estimar um subconjunto de elementos de  $\theta$ . Por exemplo, se o interesse reside em  $\beta$ , podemos usar

$$C_{i;\beta}(\theta) = 2|\Delta_i^\top(\ddot{L}^{-1} - \ddot{L}_{\gamma\gamma})\Delta_i|.$$

Se o interesse reside em  $\gamma$ , podemos usar

$$C_{i;\gamma}(\theta) = 2|\Delta_i^\top(\ddot{L}^{-1} - \ddot{L}_{\beta\beta})\Delta_i|.$$

### 3.6 Influência local conforme

(POON; POON, 1999) definem a curvatura normal conforme ao longo de uma direção  $l$ , em um ponto  $v_0$  pertencente ao gráfico de uma superfície  $\alpha$ , por

$$B_l = \frac{\mathbf{II}(l, l)}{\mathbf{I}(l, l)\text{tr}(\mathbf{II}^2)}, \quad (3.35)$$

em que,

$$\mathbf{I}_{ij} = \delta_{ij} + \frac{\partial LD}{\partial v_i} \frac{\partial LD}{\partial v_j}$$

e

$$\mathbf{II}_{ij} = \frac{1}{(1 + \|\nabla_{LD}\|^2)^{\frac{1}{2}}} \frac{\partial^2 LD}{\partial v_i \partial v_j},$$

com

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j. \end{cases}$$

Em notação matricial,

$$\mathbf{I} = I_n + \nabla_{LD} \nabla_{LD}^\top,$$

$$\mathbf{II} = \frac{1}{(1 + \|\nabla_{LD}\|^2)^{\frac{1}{2}}} H_{LD}, \quad (3.36)$$

em que  $I_n$  é a matriz identidade de ordem  $n$  e  $H_{LD}$  é a matriz hessiana de  $LD_v$ . Por simplicidade, daqui em diante usaremos  $H$  em vez de  $H_{LD}$  e  $\nabla_L$  em vez de  $\nabla_{LD}$ .

É importante observar que a curvatura conforme está bem definida, salve o caso em que a matriz associada a segunda forma fundamental é a matriz nula.

Aqui consideramos a seguinte norma para o espaço das matrizes:

$$\|H\| = \sqrt{\text{tr}(H^2)}.$$

Consequentemente, a partir da Equação (3.36), temos

$$\|H\| = (1 + \|\nabla_L\|^2)^{\frac{1}{2}} \|\mathbf{II}\| = (1 + \|\nabla_L\|^2)^{\frac{1}{2}} (\text{tr}(\mathbf{II}^2))^{\frac{1}{2}}.$$

Segue da definição de curvatura normal conforme em (3.35) que

$$\begin{aligned} B_l &= \frac{l^\top \mathbf{II} l}{l^\top \mathbf{II} [\text{tr}(\mathbf{II}^2)^{\frac{1}{2}}]} \\ &= \frac{l^\top H l}{l^\top [I_n + \nabla_L \nabla_L^\top] l [(1 + \|\nabla_L\|^2)^{\frac{1}{2}}] [\text{tr}(\mathbf{II}^2)^{\frac{1}{2}}]} \\ &= \frac{l^\top H l}{l^\top [I_n + \nabla_L \nabla_L^\top] l \|H\|}, \end{aligned}$$

avaliado em  $v = v_0$ .

É fácil verificar que se  $v_0$  é ponto crítico de  $L$  e o vetor  $l$  é unitário, então

$$B_l = \frac{l^\top H l}{\|H\|}. \quad (3.37)$$

De (COOK, 1986) temos que  $H = -2(\Delta^\top \ddot{L}^{-1} \Delta)$ . Podemos assim, reescrever a Equação (3.37) da seguinte forma:

$$\begin{aligned} B_l &= \frac{l^\top H l}{\|H\|} \\ &= \frac{l^\top [-2(\Delta^\top \ddot{L}^{-1} \Delta)] l}{\| -2(\Delta^\top \ddot{L}^{-1} \Delta) \|} \\ &= -\frac{l^\top [(\Delta^\top \ddot{L}^{-1} \Delta)] l}{\|(\Delta^\top \ddot{L}^{-1} \Delta)\|} \\ &= -\frac{l^\top [(\Delta^\top \ddot{L}^{-1} \Delta)] l}{[\text{tr}([\Delta^\top \ddot{L}^{-1} \Delta]^2)]^{\frac{1}{2}}}, \end{aligned} \quad (3.38)$$

em que as matrizes  $\ddot{L}^{-1}$  e  $\Delta$  estão avaliadas em  $\theta = \hat{\theta}$  e  $v = v_0$  respectivamente.

Podemos observar pela Equação (3.38) que os esforços computacionais de calcular a influência local por meio da curvatura normal ou por meio da curvatura normal conforme são semelhantes. Entretanto, (POON; POON, 1999) mostram que a curvatura normal conforme é detentora de propriedades relevantes quando comparada à curvatura normal.

### 3.7 Propriedades da curvatura normal conforme

Nessa seção apresentaremos algumas propriedades da curvatura normal conforme que foram demonstradas por (POON; POON, 1999). Para melhor compreensão são necessárias algumas definições, que serão apresentadas no texto.

**Definição 1.** *Seja  $H$  uma matriz quadrada de ordem  $n$ . Dizemos que  $H$  é uma matriz conforme se existe um número real positivo  $\alpha$  tal que  $HH^\top = \alpha I_n$ .*

**Proposição 3.** *Se  $H$  é uma matriz conforme, então  $H$  é não singular.*

**Proposição 4.** *Uma condição necessária e suficiente para que um matriz quadrada  $H$  de ordem  $n$  seja conforme é que  $H^\top = \alpha H^{-1}$  para algum  $\alpha \in \mathbb{R}$  positivo.*

**Proposição 5.** *A inversa de uma matriz conforme é uma matriz conforme.*

**Proposição 6.** *Se  $H$  e  $K$  são matrizes conformes de mesma ordem, então  $HK$  é uma matriz conforme.*

**Definição 2.** *Dizemos que uma aplicação diferenciável  $\psi : \Omega \rightarrow \Theta$  é uma reparametrização conforme em  $v_0$  se a matriz jacobiana de  $\psi$  em  $v_0$  é uma matriz conforme.*

**Teorema 1.** *Se  $\psi : \Omega \rightarrow \Theta$  é uma reparametrização conforme em um ponto crítico  $v_0 \in \Omega$  do gráfico de  $LD$ , então a curvatura normal conforme em  $v_0$ , em qualquer direção  $l$ , é invariante sob reparametrização.*

**Teorema 2.** *A curvatura normal conforme  $B_l$  é tal que  $|B_l| \leq 1$ , para qualquer direção  $l$ .*

**Teorema 3.** *Se  $\beta = \{a_i | 1 \leq i \leq n\}$  é uma coleção de autovetores ortonormais de  $\mathbf{II}$ , então a curvatura normal conforme  $B_{a_i}$  é igual ao autovalor normalizado,  $\hat{\lambda}_i$ , associado ao autovetor  $a_i$ . Adicionalmente,  $\sum_{i=1}^n B_{a_i} = 1$ .*

Do teorema anterior temos que  $\sum_{i=1}^n B_{a_i} = 1$ . Logo, se as curvaturas são iguais para todos os autovetores  $a_i$ , para  $i = 1, 2, 3, \dots, n$ , então cada curvatura é igual a  $\frac{1}{\sqrt{n}}$ . E este valor será utilizado na construção de um limiar de referência para avaliação da influência local.

Da Equação (3.38) podemos observar que a curvatura normal conforme e a curvatura normal diferem por um número real positivo. Portanto, o autovetor  $a_i$  que maximiza uma delas também maximiza a outra. Porém, a limitação da curvatura normal conforme facilita em sua interpretação. Com esta finalidade, (POON; POON, 1999) fornecem a definição a seguir:

**Definição 3.** *Um autovetor  $a$  é dito  $q$ -influyente se  $|B_a| \geq \frac{q}{\sqrt{n}}$ .*

### 3.8 Influência de um único autovetor

Seja  $\{e_i | 1 \leq i \leq n\}$  a base canônica de  $\mathbb{R}^n$ . Chamaremos  $e_i$  de  $i$ -ésimo vetor básico de perturbação. A fim de analisar a contribuição de vetores básicos de perturbação à influência de um único vetor  $a$   $q$ -influyente, encontramos os vetores básicos de perturbação

que estão próximos de  $a$ , no sentido de que a distância entre cada um deles e o vetor  $a$  é pequena. Seja  $A = \{a_1, a_2, \dots, a_n\}$  uma coleção de autovetores ortonormais cujos autovalores normalizados correspondentes são  $\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_n$ . Segue que se  $a_i = \sum_{j=1}^n b_{ij} e_j$ , para  $i = 1, 2, \dots, n$ , então  $\sum_{j=1}^n b_{ij}^2 = 1$ . De fato, para  $i = 1, 2, \dots, n$ , temos

$$\sum_{j=1}^n b_{ij}^2 = \langle a_i, a_i \rangle = \|a_i\|^2 = 1,$$

sendo  $\langle \cdot, \cdot \rangle$  o produto interno usual em  $\mathbb{R}^n$  induzido pela norma euclidiana  $\|\cdot\|$ . Para cada  $i = 1, 2, \dots, n$ , se a contribuição de  $b_{ij}$  é a mesma para todos os  $e_j$ ,  $j = 1, 2, \dots, n$ , então  $|b_{ij}| = \frac{1}{\sqrt{n}}$ . Com efeito, dado que a contribuição  $b_{ij}$  é a mesma para todo  $j = 1, 2, \dots, n$ , para cada  $i$  fixo temos que

$$\sum_{j=1}^n b_{ij}^2 = n b_{ij}^2,$$

para qualquer  $j = 1, 2, \dots, n$ . Como  $\sum_{j=1}^n b_{ij}^2 = 1$ , então  $n b_{ij}^2 = 1$ , ou seja,  $|b_{ij}| = \frac{1}{\sqrt{n}}$ .

### 3.9 Contribuição agregada de vetores básicos de perturbação

A ideia apresentada na seção anterior, será agora estendida para analisar a influência de vetores de perturbação básicos para todos os autovetores  $q$ -influentes. Com esse objetivo, definimos  $\mu_i = |\bar{\lambda}_i|$ , para  $i = 1, 2, \dots, n$ . Em seguida, ordenamos

$$\mu_{\max} = \mu_{(1)} \geq \mu_{(2)} \geq \dots \geq \mu_{(k)} \geq \frac{q}{\sqrt{n}} \geq \mu_{(k+1)} \geq \dots \geq \mu_{(n)} \geq 0.$$

(POON; POON, 1999) utilizaram  $j$ -ésima componente que denotamos por  $b_{ij}$  do autovetor normalizado associado a  $\mu_{(i)}$  para definir o seguinte objeto.

**Definição 4.** *A contribuição agregada do  $j$ -ésimo vetor básico de perturbação  $e_j$  para todos os autovetores  $q$ -influentes, denotada por  $m(q)_j$ , é o número real dado por*

$$m(q)_j = \left[ \sum_{i=1}^k \mu_{(i)} a_{ij}^2 \right]^{\frac{1}{2}}.$$

Neste ponto, observamos que se a contribuição de todos os vetores de perturbação é a mesma, então cada uma dela, denotadas uniformemente por  $\bar{m}(q)$ , é dada por

$$\bar{m}(q) = \left[ \frac{1}{n} \sum_{i=1}^k \mu_{(i)} \right]^{\frac{1}{2}}.$$

Devemos tomar  $\bar{m}(q)$  como referência quando tivermos por objetivo determinar a relevância da contribuição de cada um dos vetores básicos de perturbação sobre os autovetores  $q$ -influentes.

Existem dois casos extremos que devem ser levados em consideração ao fazermos uso deste método. O primeiro é permitir  $q$  suficientemente grande de tal forma que consideremos a contribuição de cada um dos vetores básicos de perturbação para  $a_{\max}$  apenas. Sendo assim,

$$m(q)_j = [\mu_{(1)} b_{1j}^2]^{\frac{1}{2}} = [\mu_{\max} b_{1j}^2]^{\frac{1}{2}} = [b_{1j}] [\mu_{\max}]^{\frac{1}{2}}.$$

O outro caso extremo é permitir  $q = 0$ . Desta forma, todos os autovetores são incluídos na análise. A contribuição total, denotada por  $m_j$ , é definida por

$$m_j = m(0)_j = \left[ \sum_{i=1}^n \mu_{(i)} b_{ij}^2 \right]^{\frac{1}{2}}. \quad (3.39)$$

Assim, se a contribuição de todos os vetores básicos de perturbação é a mesma, então cada uma delas, denotada uniformemente por  $\bar{m}$ , é tal que

$$\bar{m} = \left[ \frac{1}{n} \sum_{i=1}^n \mu_{(i)} \right]^{\frac{1}{2}} = \left[ \frac{1}{n} \sum_{i=1}^n |\bar{\lambda}_i| \right]^{\frac{1}{2}} = \left[ \frac{1}{n} \sum_{i=1}^n \left| \frac{\lambda_i}{\sqrt{\sum_{k=1}^n \lambda_k^2}} \right| \right]^{\frac{1}{2}} = \left[ \frac{\sum_{i=1}^n |\lambda_i|}{n \sqrt{\sum_{k=1}^n \lambda_k^2}} \right]^{\frac{1}{2}}. \quad (3.40)$$

**Teorema 4.** *Se todos os autovetores da matriz hessiana  $H_L$  são não negativos, então  $m_j^2 = B_{e_j}$ , para todo  $j = 1, 2, \dots, n$ .*

**Corolário 3.** *Se a contribuição  $m_j$  é a mesma para todo vetor básico de perturbação  $e_j$ , então*

$$B_{e_j} = \bar{m}^2 = \frac{\text{tr}(\mathbf{II})}{n[\text{tr}(\mathbf{II}^2)]^{\frac{1}{2}}}.$$

Inicialmente, obteremos a inversa da matriz de informação observada e, em seguida, consideraremos três esquemas de perturbação. Considere o modelo definido em (3.10) e (3.11) e seja  $\theta = (\beta^\top, \gamma^\top)^\top$ . A matriz de informação observada é dada por

$$\ddot{L} = \begin{bmatrix} \ddot{L}_{\beta\beta} & \ddot{L}_{\beta\gamma} \\ \ddot{L}_{\gamma\beta} & \ddot{L}_{\gamma\gamma} \end{bmatrix}, \quad (3.41)$$

em que  $\ddot{L}_{\beta\beta} = X^\top S_1 X + X^\top S X$ ,  $\ddot{L}_{\beta\gamma} = \ddot{L}_{\gamma\beta}^\top = X^\top A Z$  e  $\ddot{L}_{\gamma\gamma} = Z^\top B_1 Z + Z^\top B Z$ , todas as matrizes foram anteriormente definidas nas Expressões (3.16), (3.17) e (3.18). A partir de (3.41) e usando a fórmula da inversa de uma matriz particionada (RAO, 1973), obtemos

$$\ddot{L}^{-1} = \begin{bmatrix} \ddot{L}^{\beta\beta} & \ddot{L}^{\beta\gamma} \\ \ddot{L}^{\gamma\beta} & \ddot{L}^{\gamma\gamma} \end{bmatrix},$$

em que

$$\ddot{L}^{\beta\beta} = (X^\top S_1 X + X^\top S X - X^\top A Z (Z^\top B_1 Z + Z^\top B Z)^{-1} Z^\top A X)^{-1},$$

$$\begin{aligned}\ddot{L}^{\beta\gamma} &= (\ddot{L}^{\gamma\beta})^\top = -\ddot{L}^{\beta\beta} X^\top AZ (Z^\top B_1 Z + Z^\top BZ)^{-1}, \\ \dot{L}^{\gamma\gamma} &= (Z^\top B_1 Z + Z^\top BZ)^{-1} (I_p + (Z^\top AX) \ddot{L}^{\beta\beta} (X^\top AZ) (Z^\top B_1 Z + Z^\top BZ)^{-1}).\end{aligned}$$

Aqui,  $I_p$  é a matriz identidade de dimensão  $p \times p$ .

Ressaltamos que as quantidades assinaladas com “ $\hat{\cdot}$ ” são avaliadas em  $(\hat{\beta}^\top, \hat{\gamma}^\top)^\top$ . Para mais detalhes sobre a obtenção das quantidades apresentadas nas próximas seções, ver o Apêndice B.

### 3.10 Cálculo das perturbações para o modelo de regressão Kumaraswamy

No que segue, calcularemos, sob três sistemas diferentes de perturbação, a matriz

$$\Delta = \{\Delta_{ri}\}_{(k+q) \times n} = \left\{ \frac{\partial^2 \ell(\theta|v)}{\partial \theta_r \partial v_i} \right\} = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix},$$

em que  $r = 1, 2, \dots, p+q$  e  $i = 1, 2, \dots, n$ . Consideremos o modelo definido pelas Equações (3.10) e sua função de log-verossimilhança dada pela Equação (3.11). No que segue, as quantidades “ $\hat{\cdot}$ ” são avaliadas em  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$ .

### 3.11 Perturbação de ponderação de casos

A perturbação de ponderação de casos é feita através do uso de pesos para as observações na função de log-verossimilhança da seguinte forma:

$$\ell(\theta|v) = \sum_{t=1}^n v_t \ell_t(\theta|v),$$

em que  $v = (v_1, v_2, \dots, v_n)^\top$  é o vetor de peso total, com  $0 \leq v_i \leq 1$ , para  $i = 1, 2, \dots, n$ , e  $v_0 = (1, 1, \dots, 1)^\top$  é o vetor de não perturbação.

As matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são dadas por

$$\begin{aligned}\Delta_\beta &= X^\top T A_1, \\ \Delta_\gamma &= Z^\top H A_2,\end{aligned}$$

em que as matrizes  $X$ ,  $Z$ ,  $T$  e  $H$  foram anteriormente definidas no Capítulo 1 e as matrizes  $A_1$  e  $A_2$  estão definidas no Apêndice B. Ponderação de casos tem sido o esquema de perturbação mais utilizado na análise de influência e pode ser interpretado como uma perturbação na dispersão do  $t$ -ésimo caso, em especial para modelos normais lineares

### 3.12 Perturbação da variável resposta

Consideraremos que cada  $y_t$  é perturbado como  $y_{tv} = y_t + v_t S_y^t$ , onde  $S_y$  é um escalar que pode ser, por exemplo, para  $t$  fixo, igual à distância inter-quantilica estimada das componentes de  $y = (y_1, y_2, \dots, y_n)^\top$ . Neste caso, a perturbação da função log-verossimilhança é dada por

$$\ell(\theta|v) = \sum_{t=1}^n \ell_t(w_t, \alpha_t, y_{tv}),$$

em que

$$\ell_t(w_t, \phi_t, y_{tv}) = \log(\phi_t) + \log\left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})}\right) + (\phi_t - 1) \log(y_{tv}) + \left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right) \log(1 - y_{tv}^{\phi_t}).$$

Aqui,  $v_0 = (0, 0, \dots, 0)^\top$  é o vetor de não perturbação. As matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são dadas por

$$\begin{aligned} \Delta_\beta &= X^\top T M_1, \\ \Delta_\gamma &= Z^\top H M_2. \end{aligned}$$

### 3.13 Perturbação individual de covariáveis

Consideraremos agora uma perturbação aditiva em uma variável explicativa. Modificaremos, a  $p$ -ésima coluna da matriz  $X$ ,  $x_p$ ,  $p = 2, 3, \dots, k$ , fazendo  $x_{tpv} = x_{tp} + v_t S_x^p$ , em que  $S_x$  é um escalar que pode, por exemplo, ser o desvio-padrão estimado de  $x_p$ .

Para os parâmetros de locação e precisão da variável resposta que são modelados simultaneamente, consideraremos três cenários para o esquema de perturbação individual de uma covariável. Primeiro,  $X \neq Z$ . Segundo, a matriz de regressores para o modelo da mediana  $X$  e a matriz de regressores para o modelo da precisão  $Z$  são iguais, isto é,  $X = Z$ . Terceiro,  $X \neq Z$ , no entanto, para algum  $(p, p')$  temos  $z_{tp'} = x_{tp}$  ou  $z_{tp'} = \varphi(x_{tp})$ , sendo  $\varphi$  uma função diferenciável diferente da função identidade. Não consideraremos esquemas de perturbação de covariáveis apenas na modelagem da precisão.

### 3.14 Matriz $Z$ totalmente diferente da matriz $X$

Neste caso, por exemplo, se  $p \neq 2$  ou  $p \neq k$ , temos

$$\eta_{1t}(v) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_p (x_{tp} + v_t S_{x_p}) + \dots + \beta_k x_{tk}. \quad (3.42)$$

Aqui,  $w_t(v)$  é tal que  $g(w_t(v)) = \eta_{1t}(v)$ . Uma vez que os regressores que determinam a mediana não interferem na precisão,  $\phi_t(v) = \phi_t$ . Para esse tipo de perturbação,  $v_0 = (0, 0, \dots, 0)^\top$ . Obtemos

$$\begin{aligned}\Delta_\beta &= S_x^p(\beta_p X^\top T^2 S_2 + \beta_p X^\top A_1(v) T_2 - P_1 A_1(v) T), \\ \Delta_\gamma &= S_x \beta_p (Z^\top H T S_3),\end{aligned}$$

em que  $P_1$  é uma matriz  $k \times n$  de zeros exceto a  $p$ -ésima linha, que é composta por uns.

### 3.15 Matriz $Z$ igual à matriz $X$

Neste caso,  $\eta_{1t}(v)$  é como dado na Equação (3.42) e

$$\eta_{2t}(v) = \gamma_1 + \gamma_2 x_{t2} + \cdots + \gamma_p (x_{tp} + v_t S_{x_p}) + \cdots + \gamma_k x_{tk}. \quad (3.43)$$

Aqui,  $w_t(v)$  é tal que  $g(w_t(v)) = \eta_{1t}(v)$  e  $\phi_t(v)$  é tal que  $h(\phi_t(v)) = \eta_{2t}(v)$ . Para este tipo de perturbação  $v_0 = (0, 0, \dots, 0)^\top$ . Obtemos

$$\begin{aligned}\Delta_\beta &= S_x^p(\gamma_p X^\top S_{10} H T + \beta_p X^\top S_{11} T^2 + \beta_p X^\top S_9 T^2 - P_1 S_9 T), \\ \Delta_\gamma &= S_x^p(\gamma_p X^\top S_{12} H^2 + \beta_p X^\top S_{13} T H + \gamma_p X^\top S_8 H_2 - P_1 S_8 H).\end{aligned}$$

### 3.16 A $p'$ -ésima coluna da matriz $Z$ igual à $p$ -ésima coluna da matriz $X$

Aqui estamos considerando a situação em que algumas covariáveis (não todas) que determinam a mediana também estão envolvidas na modelagem da dispersão e perturbamos tais covariáveis. Ou seja, consideramos que  $z_{tp'} = x_{tp}$ , para algum par  $(p, p')$ , com  $p = 2, \dots, k$  e  $p' = 2, \dots, q$ . Neste caso,  $\eta_{1t}(v)$  é como dado em (3.42) e, por exemplo, se  $p' \neq 2$  ou  $p' \neq q$ . Temos que

$$\eta_{2t}(v) = \gamma_1 + \gamma_2 x_{t2} + \cdots + \gamma_{p'} (x_{tp} + v_t S_{x_p}) + \cdots + \gamma_k x_{tk}. \quad (3.44)$$

Aqui,  $w_t(v)$  é tal que  $g(w_t(v)) = \eta_{1t}(v)$  e  $\phi_t(v)$  é tal que  $h(\phi_t(v)) = \eta_{2t}(v)$ . Para este tipo de perturbação,  $v_0 = (0, 0, \dots, 0)^\top$  e nós obtemos

$$\begin{aligned}\Delta_\beta &= S_x^p(\gamma_{p'} X^\top S_{10} H T + \beta_p X^\top S_{11} T^2 + \beta_p X^\top S_9 T^2 - P_1 S_9 T), \\ \Delta_\gamma &= S_x^p(\gamma_{p'} Z^\top S_{12} H^2 + \beta_p Z^\top S_{13} T H + \gamma_{p'} Z^\top S_8 H_2 - P_2 S_8 H),\end{aligned}$$

em que  $P_2$  é uma matriz  $q \times n$  de zeros exceto pela  $p'$ -ésima linha, que é composta por uns.

### 3.17 A $p'$ -ésima coluna da matriz $Z$ é função da $p$ -ésima coluna da matriz $X$

Aqui estamos considerando a situação em que  $z_{tp'} = G(x_{tp})$ , com  $p = 2, 3, \dots, k$ , e  $p' = 2, 3, \dots, q$ , em que  $G$  é uma função diferenciável (se  $G$  é a função identidade, temos o caso considerado na subseção anterior). Aqui,  $\eta_{1t}(v)$  é o como dado em (3.42) e, por exemplo, para  $p' \neq 2$  e  $p' \neq q$ . Termos

$$\eta_{2t}(v) = \gamma_1 + \gamma_2 z_{t2} + \dots + \gamma_{p'} G(x_{tp} + v_t S_x^p) + \dots + \gamma_q z_{tq},$$

em que  $w_t(v)$  e  $\phi_t(v)$  dependem de  $g(w_t(v)) = \eta_{1t}(v)$  e  $h(\phi_t(v)) = \eta_{2t}(v)$ . Para este tipo de perturbação, tem-se  $v_0 = (0, 0, \dots, 0)^\top$ . As matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são da seguinte forma:

$$\begin{aligned} \Delta_\beta &= S_x^p (\gamma_{p'} X^\top S_{10} H T \dot{G} + \beta_p X^\top S_{11} T^2 + \beta_p X^\top S_9 T^2 - P_1 S_9 T), \\ \Delta_\gamma &= S_x^p (\gamma_{p'} Z^\top S_{12} H^2 \dot{G} + \beta_p Z^\top S_{13} T H + \gamma_{p'} Z^\top S_8 H_2 \dot{G} - P_2 S_8 H). \end{aligned}$$

### 3.18 Aplicações: Análise de Influência

Nosso objetivo a seguir é aplicar análise de influência a modelos de regressão Kumaraswamy ajustados a dados reais. É importante destacar que, para o conjunto de dados, ajustamos o modelo de regressão Kumaraswamy considerando as funções logito e probito para o modelo da mediana e, então, realizamos a análise de diagnóstico. Por não haver diferenças significativas nos ajustes quanto ao uso destas funções de ligação, apresentaremos as análises considerando apenas os modelos de regressão Kumaraswamy que utilizam a função logito para o modelo da mediana. A primeira aplicação refere-se à aplicação denominada Contaminação do Amendoim, sendo os dados fornecidos por (DRAPER; SMITH, 1998) e a segunda aplicação representa a medição da umidade relativa do ar e temperatura disponíveis no site INMET (<http://www.inmet.gov.br/portal/>). Aqui construiremos os gráficos de influência local e influência local conforme considerando os três esquemas de perturbação apresentados na seção anterior. Inicialmente, apresentamos uma figura com os gráficos dos vetores  $d_{\max}$  contra os índices das observações. Nesta figura a primeira coluna,  $d_{\max;\theta}$ , refere-se à influência local relativa ao vetor de parâmetros  $\theta$ . A segunda coluna,  $d_{\max;\beta}$ , refere-se à influência local apenas do vetor  $\beta$ , enquanto a última coluna,  $d_{\max;\gamma}$ , refere-se à influência local apenas do vetor  $\gamma$ . Quanto à composição das linhas da figura, temos, na primeira linha, o esquema de perturbação de ponderação de casos, na segunda linha o esquema de perturbação da variável resposta e na terceira linha o esquema de perturbação da covariável. Em seguida, apresentamos uma figura na qual estão os gráficos de influência local total  $C_t$  contra os índices das observações, que seguem esquema similar ao descrito anteriormente. Por fim, apresentamos figuras com os

gráficos de influência local conforme para os três esquemas de perturbação apresentados anteriormente.

### 3.19 Aplicação I: Contaminação do amendoim, modelo de regressão Kumaraswamy

Nossa ideia aqui é aplicar o método de influência local e influência local conforme ao conjunto de dados Contaminação do Amendoim. Utilizamos o modelo de regressão Kumaraswamy para avaliar o comportamento da proporção de amendoins não contaminados ( $y$ ) através da covariável níveis de aflatoxina ( $x$ ). Consideraremos a seguir um modelo com precisão constante e função de ligação logito para o modelo da mediana:

$$\log\left(\frac{w_t}{1-w_t}\right) = \beta_0 + \beta_1 x_t, \quad t = 1, 2, \dots, 34.$$

Na Tabela 13 apresentamos os resultados inferenciais referentes ao modelo.

Tabela 13 – Estimativas dos parâmetros, erros-padrão e  $p$ -valores do modelo  $\log(\mu_t/(1 - \mu_t)) = \beta_1 + \beta_2 x_{t2}$  e  $\phi$  constante; Contaminação do Amendoim.

Parâmetros	Estimativas	Erros-padrão	$p$ -valores
$\beta_1$	7.9469	0.1425	<0.0001
$\beta_2$	-0.0215	0.0015	<0.0001
$\gamma_1$	7.8878	0.1648	<0.0001

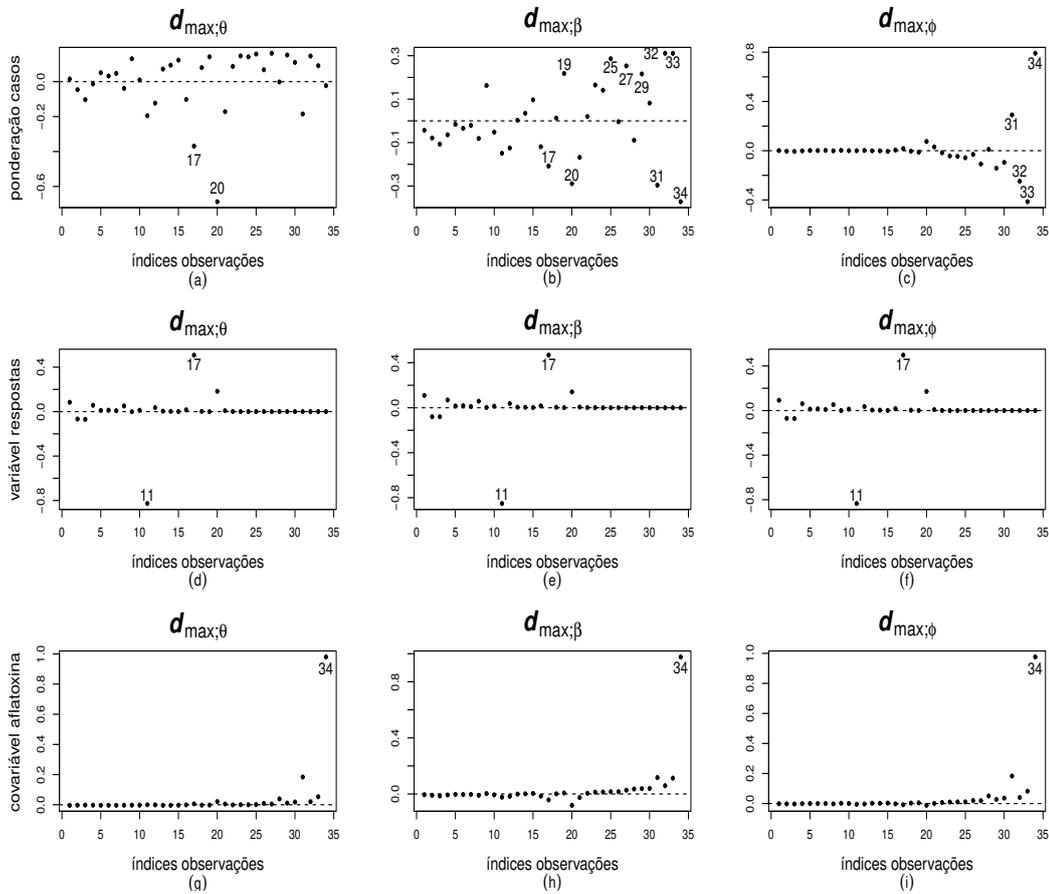
Fonte: O autor (2020).

Iniciando à análise de influência, na Figura 21 estão os gráficos dos valores de  $d_{\max}$  contra os índices das observações e na Figura 22 estão os gráficos de  $C_t$  contra os índices das observações.

Na Figura 21, (gráficos de influência local), notamos a influência individual da observação 34 sobre a estimativa de  $\beta$ , ponto destacado nos gráficos de perturbação covariável aflatoxina. Classificamos como conjuntamente influentes as observações 17 e 20, destacadas no gráfico de perturbação de ponderação de casos  $d_{\max;\theta}$ , e classificamos os pontos 17, 19, 20, 25, 27, 29, 31, 32, 33 e 34 como observações conjuntamente influentes para o modelo com base no gráfico de  $d_{\max;\beta}$ . Por fim, ainda são notadas como conjuntamente influentes as observações 31, 32, 33 e 34 destacadas no gráfico de  $d_{\max;\phi}$  na perturbação de ponderação de casos.

Na Figura 22, (gráfico de influência local total), notamos a influência individual da observação 34, ponto destacado nos gráficos de perturbação da covariável aflatoxina.

Figura 21 – Gráficos influência local; Contaminação do Amendoim.



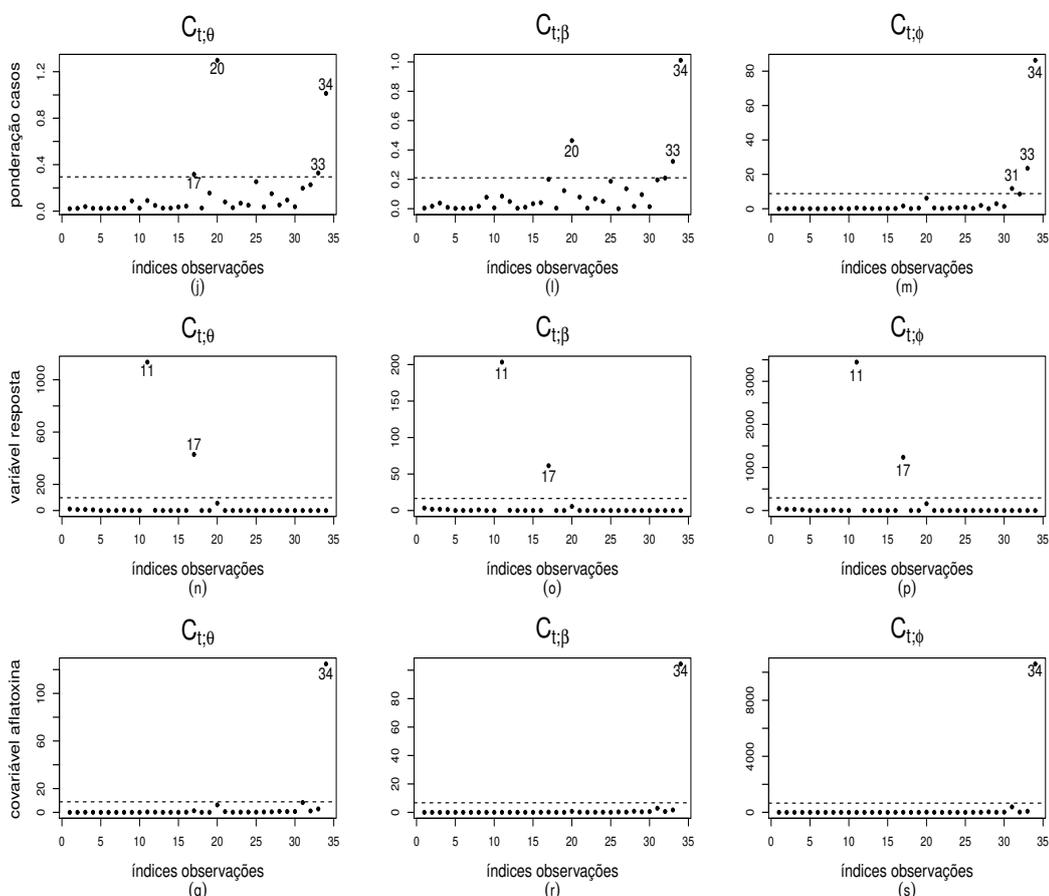
Fonte: O autor (2020).

Observamos como conjuntamente influentes as observações 31, 33 e 34, que são destacadas no gráfico de perturbação de ponderação de casos  $C_{t;\phi}$ .

Na Tabela 14 exibimos as variações nos percentuais nas estimativas dos parâmetros do modelo após a retirada dos pontos influentes. Destacamos que a observação 34 exerce influência principalmente sobre a estimativa do parâmetro  $\beta_2$ , havendo mudança de aproximadamente 13% na estimativa. Os conjuntos  $\{17, 20\}$  e  $\{31, 32, 34\}$  também exercem maior influência sobre a estimativa do parâmetro  $\beta_2$ . Por outro lado, os conjuntos  $\{31, 32, 33, 34\}$  e  $\{17, 19, 20, 25, 27, 29, 31, 32, 33, 34\}$  exercem grande influência tanto sobre as estimativas dos parâmetros do submodelo da mediana quanto sobre a estimativa do parâmetro de precisão.

Continuaremos a análise de diagnóstico considerando a técnica de influência local baseada na curvatura conforme, sob os esquemas de perturbações apresentados anteriormente. Iniciaremos utilizando o esquema de perturbação de ponderação de casos. Consideramos os valores  $q = 1, 2, 3$  e  $4$  a fim de determinar quais são os autovetores (direções) associados às variações máximas da função de distância entre verossimilhanças. Na Figura 23(a)

Figura 22 – Gráficos influência local total; Contaminação do Amendoim.



Fonte: O autor (2020).

Tabela 14 – Variações percentuais nas estimativas dos parâmetros após a retirada de observações influentes; Contaminação do Amendoim.

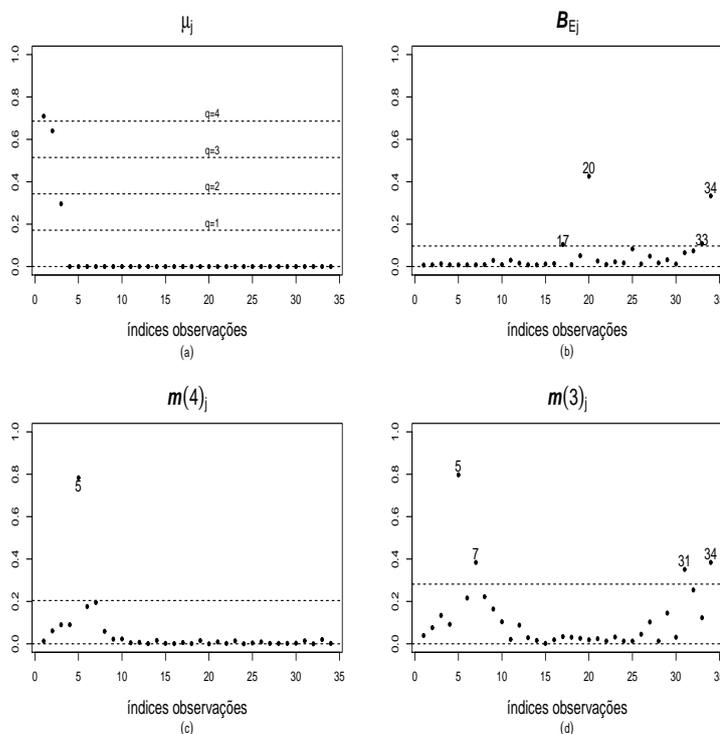
Casos	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\phi}$
34	1.7259	13.1198	0.2923
17, 20	1.4647	5.8350	1.7688
31, 32, 34	2.9837	25.7572	1.5027
31, 32, 33, 34	3.4779	29.9828	29.9828
17, 19, 20, 25, 27, 29, 31, 32, 33, 34	2.3610	15.0749	7.4559

Fonte: O autor (2020).

apresentamos o gráfico do módulo dos autovetores normalizados e indicamos os valores de  $q$ . Temos  $n = 34$  e para  $q = 4$  há apenas um autovalor acima do valor de referência  $q/\sqrt{n}$ , para  $q = 3$  há apenas dois autovalores acima do valor de referência e para  $q = 1$  há 3 autovalores, isto é, existe um único autovetor 4-influente, há dois autovetores 3-influente e existem três vetores 1-influente. Os autovalores normalizados associados a estes autove-

tores são 0.70877, 0.64046 e 0.29573. Desta forma, a curvatura conforme máxima  $B_{l_{\max}}$  é 0.7087, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura 23(b) estão apresentados os valores das contribuições agregadas de todos os autovetores. As observações 17, 20, 33 e 34, que estão acima do valor de referência  $2b$  sugerido no trabalho de (POON; POON, 1999), são consideradas pontencialmente influentes. Na Figura 23(c)(d), temos as contribuições agregadas dos autovetores associados aos maiores autovalores, correspondentes a  $q = 4$  e  $q = 3$ . Neste caso, para  $q = 4$  a observação 5 é detectada como potencialmente influente, pois está acima do valor de referência  $\bar{m}(q)\sqrt{2}$ . Para  $q = 3$ , as observações potencialmente influentes são 5, 7, 31 e 34.

Figura 23 – Gráficos de influência local conforme sob perturbação de ponderação de casos; Contaminação do Amendoim.

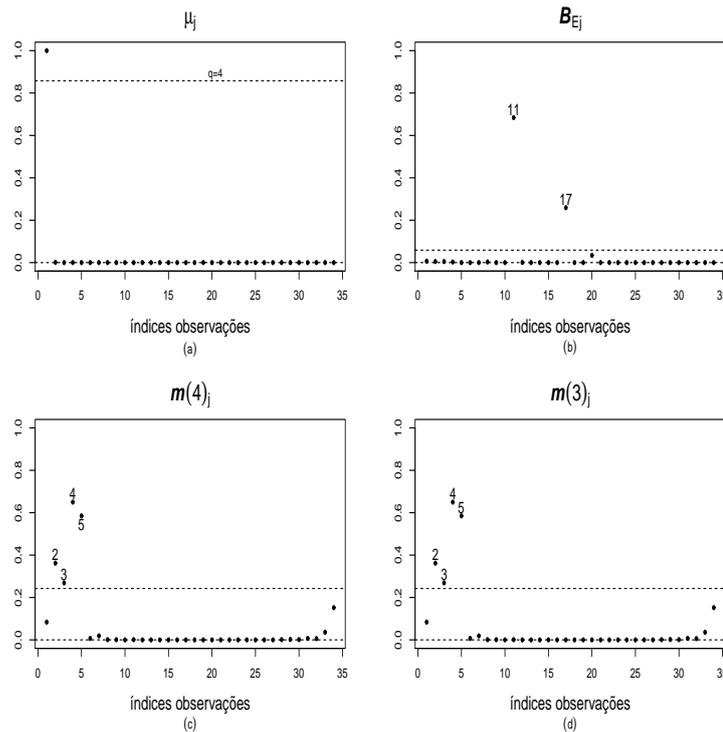


Fonte: O autor (2020).

Consideraremos agora o esquema de perturbação da variável resposta. Na Figura 24(a) apresentamos o gráfico do módulo dos autovetores normalizados. Para  $q = 4$  há apenas um autovalor acima do valor de referência  $q/\sqrt{n}$ , isto é, existe um único autovetor 4-influente. O autovalor normalizado associado a este autovetor é 0.9998. Desta forma, a curvatura conforme máxima  $B_{l_{\max}}$  é 0.9998, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura 24(b) estão apresentados as contribuições agregadas de todos os autovetores. As observações 11 e 17, que estão acima do valor de referência  $2b$  sugerido no trabalho de (POON; POON, 1999), são consideradas pontencialmente influentes. Na

Figura 24(c)(d), temos as contribuições agregadas dos autovetores associados aos maiores autovalores, correspondentes a  $q = 4$  e  $q = 3$ . Em ambos os casos, as observações 2, 3, 4 e 5 são detectadas como potencialmente influentes, pois estão acima do valor de referência  $\bar{m}(q)\sqrt{2}$ .

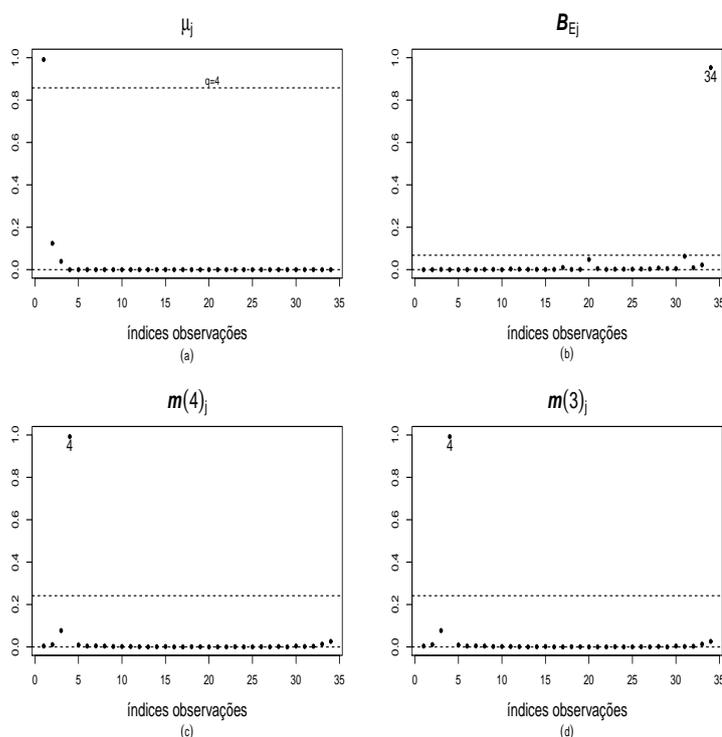
Figura 24 – Gráficos de influência local conforme sob perturbação da variável resposta; Contaminação do Amendoim.



Fonte: O autor (2020).

Por fim, consideraremos o esquema de perturbação da covariada. Na Figura 25(a) apresentamos o gráfico do modulo dos autovetores normalizados para  $q = 4$ , existe apenas um autovalor acima do valor de referência  $q/\sqrt{n}$ , isto é, há um único autovetor 4-influente. O autovalor normalizado associado a este autovetor é 0.9998. Assim, a curvatura conforme máxima  $B_{l_{\max}}$  é 0.9998, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura 25(b) estão apresentados os valores das contribuições agregadas de todos os autovetores. A observação 34, que está acima do valor de referência  $2b$ , é considerada pontencialmente influente. A Figura 25(c)(d), apresenta as contribuições agregadas dos autovetores associados aos maiores autovalores, correspondentes a  $q = 4$  e  $q = 3$ . Em ambos os casos, a observação 4 é detectada como potencialmente influentes, pois esta acima do valor de referência  $\bar{m}(q)\sqrt{2}$ .

Figura 25 – Gráficos de influência local conforme sob perturbação da covariada; Contaminação do Amendoim.



Fonte: O autor (2020).

Tabela 15 – Variações percentuais nas estimativas dos parâmetros após a retirada das observações influentes utilizando o método de influência local conforme; Contaminação do Amendoim.

Casos	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\phi}$
34	1.7259	13.1198	0.2923
5, 7, 34	1.8853	14.1024	0.0770
7, 31, 34	2.3771	20.6087	1.5399
5, 7, 31, 34	2.4349	20.9118	1.3596

Fonte: O autor (2020).

### 3.20 Aplicação II: Umidade relativa do ar, modelo de regressão Kumaraswamy

Nesta segunda aplicação utilizamos o modelo de regressão Kumaraswamy para avaliar o comportamento da umidade relativa do ar ( $y$ ) através da covariável temperatura ( $x$ ), aplicando os métodos de influência local e influência local conforme. Consideraremos a seguir um modelo com precisão constante e função de ligação logito para o modelo da mediana:

$$\log\left(\frac{w_t}{1-w_t}\right) = \beta_0 + \beta_1 x_t, \quad t = 1, 2, \dots, 30.$$

A Tabela 16 contém as estimativas dos parâmetros e os respectivos erros-padrão e  $p$ -valores.

Tabela 16 – Estimativas dos parâmetros, erros-padrão e  $p$ -valores do modelo  $\log(w_t/(1-w_t)) = \beta_1 + \beta_2 x_{t2}$  e  $\phi$  constante; Umidade Relativa do Ar.

Parâmetros	Estimativas	Erros-padrão	$p$ -valores
$\beta_1$	12.7366	1.4162	< 0.0001
$\beta_2$	-0.4181	0.0509	< 0.0001
$\gamma_1$	2.9618	0.1481	< 0.0001

Fonte: O autor (2020).

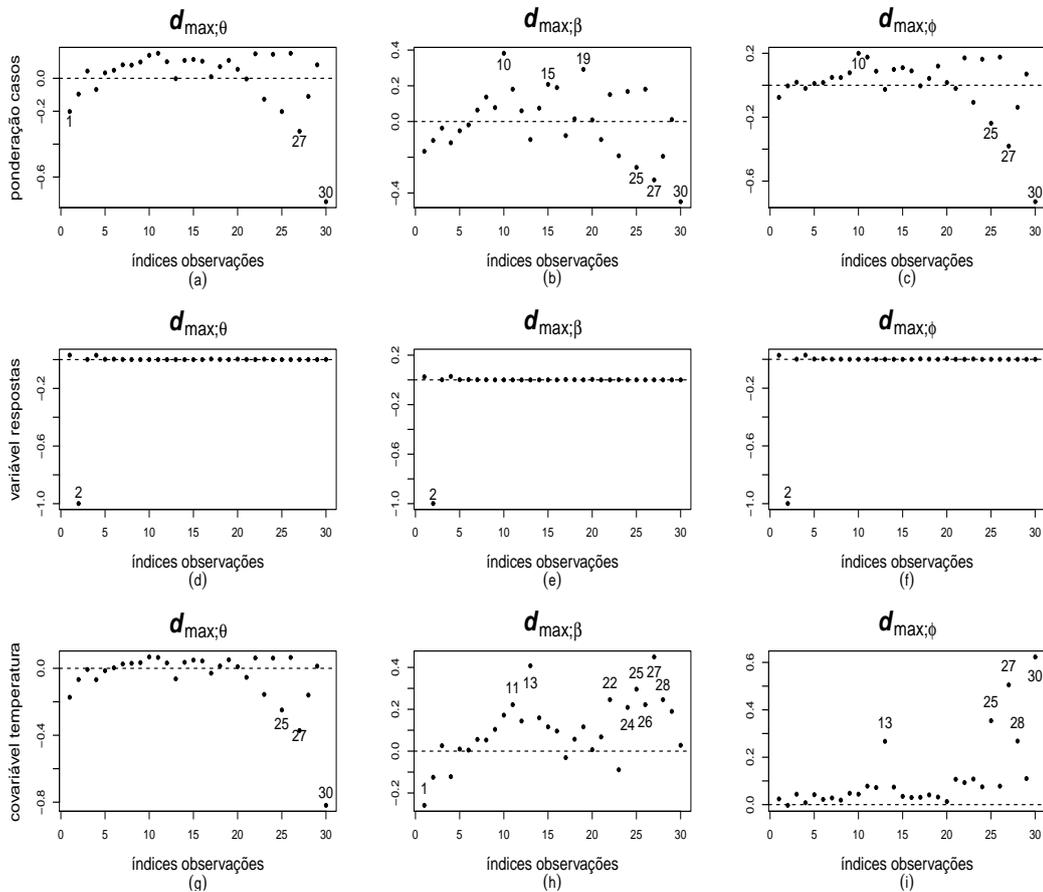
Iniciando a análise de influência, a Figura 26 contém os gráficos dos valores de  $d_{\max}$  contra os índices das observações e a Figura 27 contém os gráficos de  $C_t$  contra os índices das observações.

Na Figura 26 (gráficos de influência local), notamos as influências individuais das observações 1 e 2 sobre as estimativas dos parâmetros  $\beta$ , pontos destacados nos gráficos de perturbação covariável temperatura e da variável resposta. Classificamos como conjuntamente influentes os conjuntos de observações  $\{27, 30\}$ ,  $\{1, 27, 30\}$ ,  $\{10, 15, 19\}$ ,  $\{10, 15, 27\}$ ,  $\{25, 27, 30\}$ ,  $\{10, 25, 27, 30\}$ ,  $\{10, 15, 19, 25, 27\}$  e  $\{10, 15, 19, 25, 27, 30\}$ , que foram destacados no gráfico de perturbação de ponderação de casos, conjuntos de observações  $\{27, 30\}$ ,  $\{13, 25, 27\}$ ,  $\{13, 25, 27, 28\}$ ,  $\{13, 25, 27, 30\}$ , e  $\{13, 25, 27, 28, 30\}$ , segundo os gráficos de perturbação da covariada.

Na Figura 27 (gráfico de influência local total), notamos a influência individual das observações 1 e 2, pontos destacados nos gráficos de perturbação de ponderação de casos. Classificamos como conjuntamente influentes os conjuntos de observações  $\{1, 2\}$ ,  $\{27, 30\}$ ,  $\{1, 2, 10\}$ ,  $\{1, 27, 30\}$ ,  $\{1, 2, 10, 27\}$  e  $\{1, 2, 10, 27, 30\}$ , destacados no gráfico de perturbação de ponderação de casos, e com base no gráfico de perturbação da covariada temperatura classificamos como conjuntamente influentes os conjuntos de observações  $\{1, 13\}$  e  $\{13, 25, 27, 30\}$ .

Na Tabela 17 exibimos as variações nos percentuais nas estimativas dos parâmetros do modelo após a retirada dos pontos influentes. Notamos que as observações 1 e 2 exercem influência principalmente sobre a estimativa do parâmetro  $\beta_2$ . Os conjuntos  $\{1, 2\}$ ,  $\{1, 2, 10\}$  e  $\{25, 27, 30\}$ ,  $\{10, 15, 19, 25, 27, 30\}$  também exercem maior influência sobre a estimativa do parâmetro  $\beta_2$  com variações percentuais de aproximadamente 15%. Adicionalmente, o conjunto  $\{13, 25, 27, 28, 30\}$  exerce grande influência tanto sobre as estimativas dos parâmetros do submodelo da mediana quanto sobre a estimativa do parâmetro de precisão.

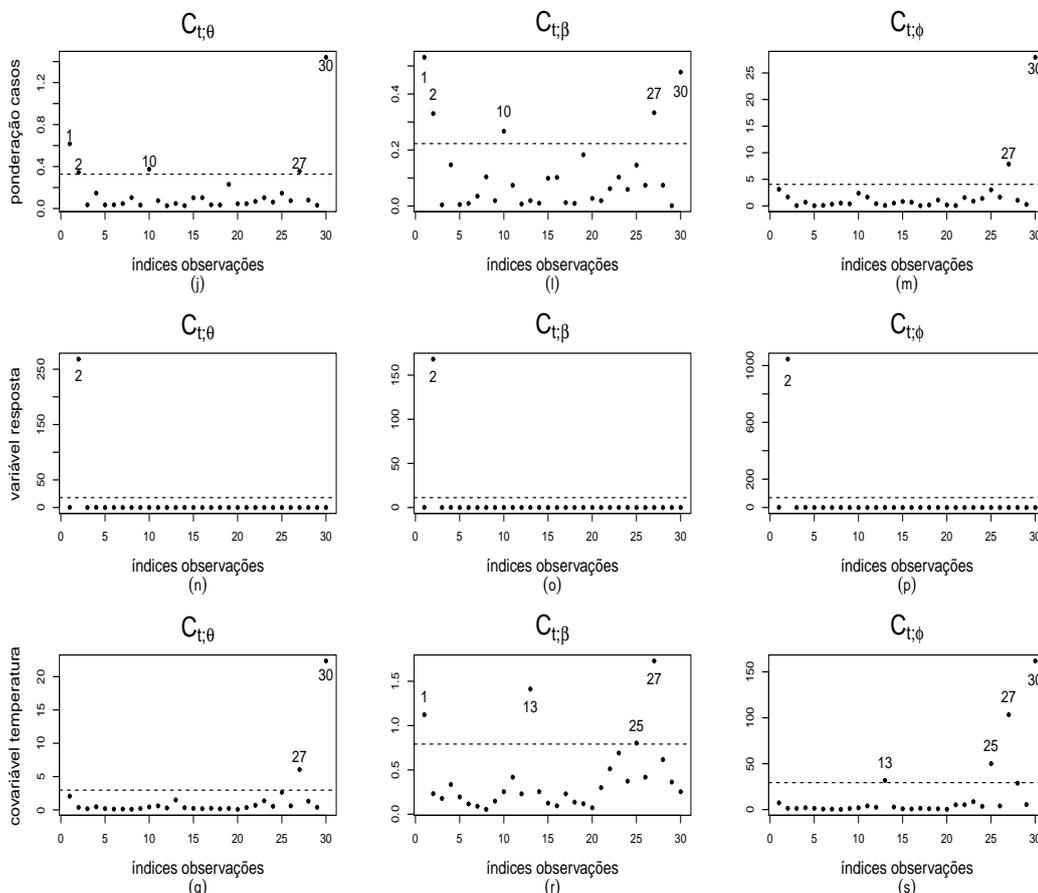
Figura 26 – Gráficos influência local; Umidade Relativa do Ar.



Fonte: O autor (2020).

Continuaremos a análise de diagnóstico considerando a técnica de influência local baseada na curvatura conforme, sob os esquemas de perturbações apresentados anteriormente. Iniciaremos utilizando o esquema de perturbação de ponderação de casos. Consideramos os valores  $q = 2, 3$  e  $4$  a fim de determinar quais são os autovetores (direções) associados às variações máximas da função de distância entre verossimilhanças. Na Figura 28(a) apresentamos o gráfico dos módulos dos autovetores normalizados e indicamos os valores de  $q$ . Temos  $n = 30$  e para  $q = 4$  há apenas um autovalor acima do valor de referência  $q/\sqrt{n}$ , para  $q = 3$  há apenas dois autovalores acima do valor de referência e para  $q = 2$  há 3 autovalores, isto é, existe um único autovetor 4-influente, há dois autovetores 3-influentes e existem três vetores 2-influentes. Os autovalores normalizados associados a estes autovetores são 0.743631, 0.52640 e 0.4122. Desta forma, a curvatura conforme máxima  $B_{l_{\max}}$  é 0.743631, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura 28(b) estão apresentados os valores das contribuições agregadas de todos os autovetores. As observações 1, 2, 10, 27, 30, que estão acima do valor de referência  $2b$  sugerido no

Figura 27 – Gráficos influência local total; Umidade Relativa do Ar.



Fonte: O autor (2020).

trabalho de (POON; POON, 1999), são consideradas potencialmente influentes. Na Figura 28(c)(d), apresentamos as contribuições agregadas dos autovetores associados aos maiores autovalores, correspondentes a  $q = 4$  e  $q = 3$ . Neste caso, para  $q = 4$  a observação 5 é detectada como potencialmente influente, pois está acima do valor de referência  $\bar{m}(q)\sqrt{2}$ . Para  $q = 3$ , as observações potencialmente influentes são 2, 9, 11 e 25.

Consideraremos agora o esquema de perturbação da variável resposta. Na Figura 29(a) apresentamos o gráfico do módulo dos autovetores normalizados. Para  $q = 4$  há apenas um autovalor acima do valor de referência  $q/\sqrt{n}$ , isto é, existe um único autovetor 4-influente. O autovalor normalizado associado a este autovetor é 0.9999. Desta forma, a curvatura conforme máxima  $B_{l_{\max}}$  é 0.9999, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura 29(b) estão apresentadas as contribuições agregadas de todos os autovetores. A observação 2, que está acima do valor de referência  $2b$  sugerido em de (POON; POON, 1999), é considerada potencialmente influente. A Figura 29(c)(d) apresenta as contribuições agregadas dos autovetores associados aos maiores autovalores, correspondentes a  $q = 4$

Tabela 17 – Variação percentual nas estimativas dos parâmetros retirando as observações influentes; Umidade Relativa do Ar.

Casos	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\phi}$
1	6.0307	6.5866	0.7564
2	4.9230	5.3895	0.2057
1, 2	13.6387	14.9524	1.5592
1, 13	6.0354	6.5527	0.0352
27, 30	8.0188	9.2849	4.1327
1, 2, 10	14.1698	15.7197	3.9069
1, 27, 30	2.0600	2.7578	5.2799
10, 15, 19	2.5028	2.3397	5.0219
10, 25, 27	5.0505	5.7326	1.2563
13, 25, 27	7.9778	9.2793	1.6009
25, 27, 30	12.1082	14.0122	5.5036
10, 25, 27, 30	11.9859	13.7109	7.8308
13, 25, 27, 28	10.8774	12.6845	2.2559
13, 25, 27, 30	19.0385	21.9011	5.8917
1, 2, 10, 27, 30	4.3518	4.4742	8.6905
10, 15, 19, 25, 27	7.6167	8.3047	3.9459
13, 25, 27, 28, 30	31.2052	35.6769	11.8577
10, 15, 19, 25, 27, 30	14.8440	16.6166	11.5674

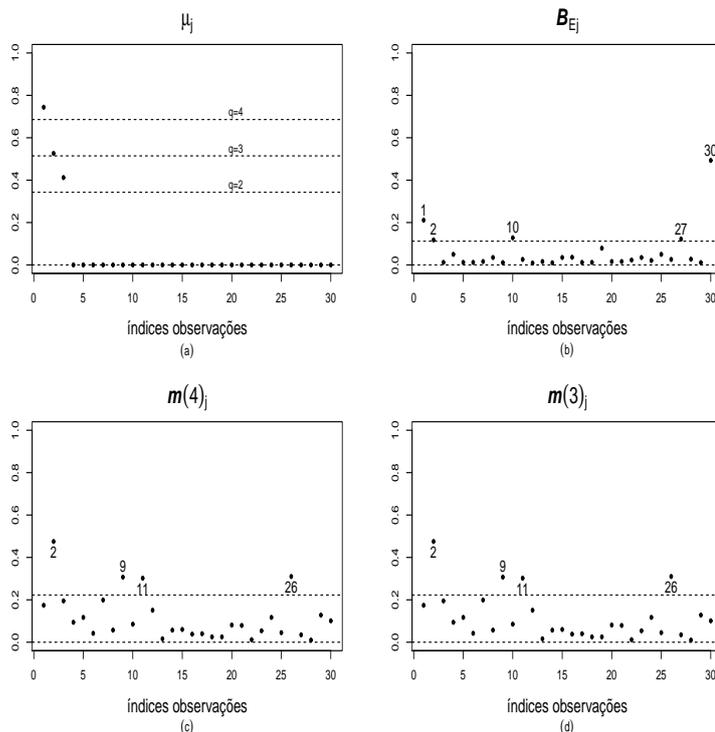
Fonte: O autor (2020).

e  $q = 3$ . Em ambos os casos, as observações 2 e 4 são detectadas como potencialmente influentes, pois estão acima do valor de referência  $\bar{m}(q)\sqrt{2}$ .

Por fim, consideraremos o esquema de perturbação da covariada. Na Figura 30(a) apresentamos o gráfico dos módulos dos autovetores normalizados para  $q = 4$ . Nota-se que existe apenas um autovalor acima do valor de referência  $q/\sqrt{n}$ , isto é, há um único autovetor 4-influente. O autovalor normalizado associado a este autovetor é 0.96651. Assim, a curvatura conforme máxima  $B_{l_{\max}}$  é 0.96651, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura 30(b) estão apresentados os valores das contribuições agregadas de todos os autovetores. As observações 27 e 30, que estão acima do valor de referência  $2b$ , são consideradas pontencialmente influentes. A Figura 30(c)(d) apresenta as contribuições agregadas dos autovetores associados aos maiores autovalores, correspondentes a  $q = 4$  e  $q = 3$ . Em ambos os casos, as observações 2, 3, 11 e 13 são detectadas como potencialmente influentes, pois estão acima do valor de referência  $\bar{m}(q)\sqrt{2}$ .

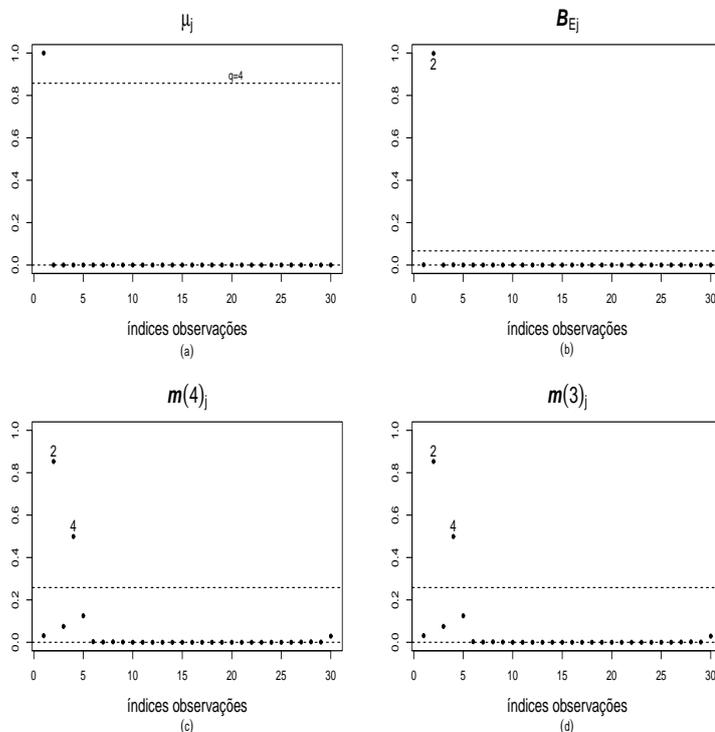
Na Tabela 18 exibimos as variações nos percentuais nas estimativas dos parâmetros do modelo após a retirada dos pontos influentes. Notamos que a observações 1 e 2 exercem influência principalmente sobre a estimativa do parâmetro  $\beta_2$ . O conjunto  $\{2, 4\}$  exerce maior influência sobre a estimativa do parâmetro  $\beta_2$  com variação percentual

Figura 28 – Gráficos de influência local conforme sob perturbação de ponderação de casos; Umidade Relativa do Ar.



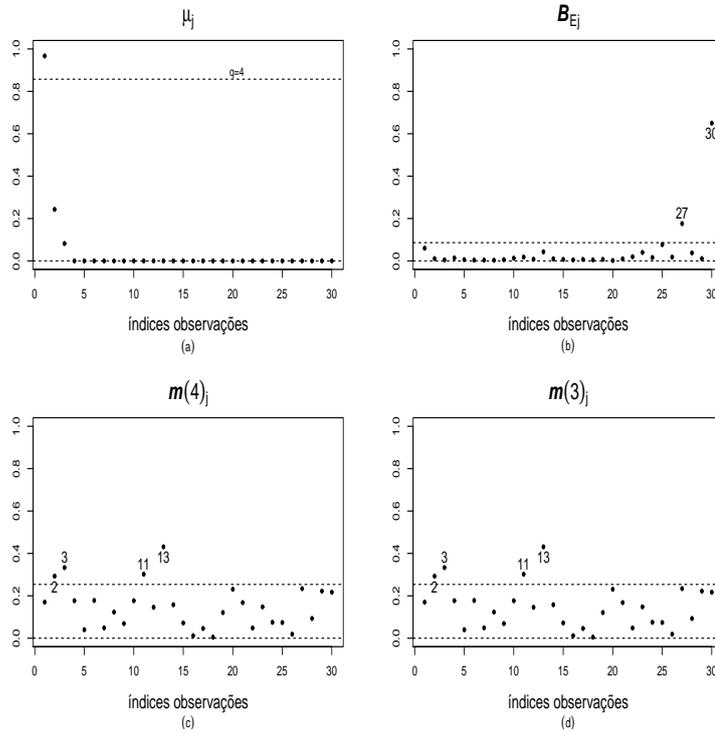
Fonte: O autor (2020).

Figura 29 – Gráficos de influência local conforme sob perturbação da variável resposta; Umidade Relativa do Ar.



Fonte: O autor (2020).

Figura 30 – Gráficos de influência local conforme sob perturbação da covariada; Umidade Relativa do Ar.



Fonte: O autor (2020).

Tabela 18 – Variação percentual nas estimativas dos parâmetros retirando as observações influentes; Umidade Relativa do Ar.

Casos	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\phi}$
1	6.0307	6.5866	0.7564
2	4.9230	5.3895	0.2057
1, 2	13.6387	14.9524	1.5592
2, 3	4.7906	5.2196	0.4734
2, 4	9.4294	10.3110	0.1252
27, 30	8.0188	9.2849	4.1327
1, 2, 10	14.1698	15.7197	3.9069
2, 3, 11	5.8223	6.4484	0.0695
2, 9, 11	5.2710	5.8981	0.4672
1, 27, 30	2.0600	2.7578	5.2799
10, 25, 27	5.0505	5.7326	1.2563
1, 2, 10, 27	11.2806	12.3883	3.2894
2, 9, 11, 26	6.3959	7.2451	1.0296

Fonte: O autor (2020).

de aproximadamente 10%. Adicionalmente, os conjuntos  $\{1, 2\}$ ,  $\{1, 2, 10\}$  também exercem maior influência sobre a estimativa do parâmetro  $\beta_2$  com variações percentuais de aproximadamente 15%.

### 3.21 Cálculo das perturbações para o modelo de regressão beta

A seguir, calcularemos, para o modelo de regressão beta, sob três sistemas diferentes de perturbação, a matriz

$$\Delta = \{\Delta_{ri}\}_{(k+q) \times n} = \left\{ \frac{\partial^2 \ell(\theta|v)}{\partial \theta_r \partial v_i} \right\} = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix},$$

em que  $r = 1, 2, \dots, p+q$  e  $i = 1, 2, \dots, n$ . Consideremos o modelo definido pelas Equações em (3.20) e sua função de log-verossimilhança dada pela equação (3.22). No que segue, as quantidades distinguem-se pela adição de “ $\hat{\sim}$ ” são avaliadas em  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$ .

### 3.22 Perturbação de ponderação de casos

A perturbação de ponderação de casos é feita através do uso de pesos para as observações na função de log-verossimilhança da seguinte forma:

$$\ell(\theta|v) = \sum_{t=1}^n v_t \ell_t(\theta|v),$$

em que  $v = (v_1, v_2, \dots, v_n)^\top$  é o vetor de peso total, com  $0 \leq v_i \leq 1$ , para  $i = 1, 2, \dots, n$ , e  $v_0 = (1, 1, \dots, 1)^\top$  é o vetor de não perturbação.

As matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são dadas por

$$\begin{aligned} \Delta_\beta &= X^\top \Phi T \varepsilon, \\ \Delta_\gamma &= Z^\top H A_3, \end{aligned}$$

em que

$$\varepsilon = \text{diag}(y_1^* - \mu_1^*, \dots, y_n^* - \mu_n^*) \quad (3.45)$$

e

$$A_3 = \text{diag}(a_1, \dots, a_n), \quad (3.46)$$

sendo  $a_t$ , para  $t = 1, 2, \dots, n$ , como definido na Equação (3.27) e sendo a matriz  $T$  como definida após a Equação (3.25).

### 3.23 Perturbação da variável resposta

Consideraremos que cada  $y_t$  é perturbado como  $y_{tv} = y_t + v_t s(y_t)$ , em que  $s(y_t)$  é um escalar que pode ser, por exemplo, para  $t$  fixo igual à distância interquartilica estimada das componentes de  $y = (y_1, y_2, \dots, y_n)^\top$ . Neste caso, a perturbação da função log-verossimilhança é dada por

$$\ell(\theta|v) = \sum_{t=1}^n \ell_t(\mu_t, \alpha_t, y_{tv}).$$

Aqui,  $v_0 = (0, 0, \dots, 0)^\top$  é o vetor de não perturbação. As matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  assumem as formas

$$\begin{aligned}\Delta_\beta &= X^\top \Phi T M S_y, \\ \Delta_\gamma &= Z^\top H B S_y.\end{aligned}$$

em que  $S_y = \text{diag}(s(y_1), s(y_2), \dots, s(y_n))$ ,  $B = \text{diag}(b_1, b_2, \dots, b_n)$  e  $M = \text{diag}(m_1, m_2, \dots, m_n)$ , com  $b_t$  dado por

$$b_t = \frac{(y_t - \mu_t)}{y_t(1 - y_t)}, \quad \text{para } t = 1, 2, \dots, n,$$

e

$$m_t = \frac{1}{y_t(1 - y_t)}, \quad \text{para } t = 1, 2, \dots, n, \quad \text{respectivamente.}$$

### 3.24 Perturbação individual de covariáveis

Consideraremos agora uma perturbação aditiva em uma variável explicativa. Modificaremos, a  $p$ -ésima coluna da matriz  $X$ ,  $x_p$ ,  $p = 2, 3, \dots, k$ , fazendo  $x_{tpv} = x_{tp} + v_t s_{x_p}$ , em que  $s_x$  é um escalar que pode, por exemplo, ser o desvio-padrão estimado de  $x_p$ .

Para os parâmetros de locação e precisão da variável resposta, que são modelados simultaneamente, consideraremos três cenários para o esquema de perturbação individual de uma covariável. Primeiro,  $X \neq Z$ . Segundo, a matriz de regressores para o modelo da mediana  $X$  e a matriz de regressores para o modelo da precisão  $Z$  são iguais, isto é,  $X = Z$ . Terceiro,  $X \neq Z$ , no entanto, para algum  $(p, p')$  temos  $z_{tp'} = x_{tp}$  ou  $z_{tp'} = \varphi(x_{tp})$ , sendo  $\varphi$  uma função diferenciável diferente da função identidade. Não consideraremos esquemas de perturbação de covariáveis apenas na modelagem da precisão.

### 3.25 Matriz $Z$ totalmente diferente da matriz $X$

Neste caso, por exemplo, se  $p \neq 2$  ou  $p \neq k$ , temos

$$\eta_{1t}(v) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_p (x_{tp} + v_t s_{x_p}) + \dots + \beta_k x_{tk}. \quad (3.47)$$

Aqui,  $\mu_t(v)$  é tal que  $g(\mu_t(v)) = \eta_{1t}(v)$ . Uma vez que os regressores que determinam a mediana não interferem na precisão,  $\phi_t(v) = \phi_t$ . Para esse tipo de perturbação,  $v_0 = (0, 0, \dots, 0)^\top$ . Temos que

$$\begin{aligned}\Delta_\beta &= s_{x_p} [\beta_p [X^\top \Phi Q - P \Phi T \varepsilon], \\ \Delta_\gamma &= s_{x_p} \beta_p (Z^\top H T F),\end{aligned}$$

em que  $P$  é uma matriz  $k \times n$  de zeros exceto a  $p$ -ésima linha, que é composta por uns e os componentes das matrizes diagonais  $Q$  e  $F$  estão definidas em (3.29) e (3.31), respectivamente.

### 3.26 Matriz $Z$ igual à matriz $X$

Neste caso,  $\eta_{1t}(v)$  é como dado na Equação (3.47) e

$$\eta_{2t}(v) = \gamma_1 + \gamma_2 x_{t2} + \cdots + \gamma_p (x_{tp} + v_t s_{x_p}) + \cdots + \gamma_k x_{tk}. \quad (3.48)$$

Aqui,  $\mu_t(v)$  é tal que  $g(\mu_t(v)) = \eta_{1t}(v)$  e  $\phi_t(v)$  é tal que  $h(\phi_t(v)) = \eta_{2t}(v)$ . Para este tipo de perturbação,  $v_0 = (0, 0, \dots, 0)^\top$  e obtemos

$$\begin{aligned} \Delta_\beta &= -s_{x_p} (X^\top [\beta_p \Phi Q + FTH\gamma_p] - P\Phi T\varepsilon), \\ \Delta_\gamma &= -s_{x_p} (X^\top [\gamma_p V + FTH\beta_p] - PHA). \end{aligned}$$

### 3.27 A $p'$ -ésima coluna da matriz $Z$ igual à $p$ -ésima coluna da matriz $X$

Aqui estamos considerando a situação em que algumas covariáveis (não todas) que determinam a mediana também estão envolvidas na modelagem da dispersão e perturbamos tais covariáveis. Ou seja, consideramos que  $z_{tp'} = x_{tp}$ , para algum par  $(p, p')$ , com  $p = 2, \dots, k$  e  $p' = 2, \dots, q$ . Neste caso,  $\eta_{1t}(v)$  é como dado em (3.47) e, por exemplo, se  $p' \neq 2$  ou  $p' \neq q$ , temos que

$$\eta_{2t}(v) = \gamma_1 + \gamma_2 x_{t2} + \cdots + \gamma_{p'} (x_{tp} + v_t s_{x_p}) + \cdots + \gamma_k x_{tk}. \quad (3.49)$$

Aqui,  $\mu_t(v)$  é tal que  $g(\mu_t(v)) = \eta_{1t}(v)$  e  $\phi_t(v)$  é tal que  $h(\phi_t(v)) = \eta_{2t}(v)$ . Para este tipo de perturbação,  $v_0 = (0, 0, \dots, 0)^\top$  e nós obtemos

$$\begin{aligned} \Delta_\beta &= -s_{x_p} (X^\top [\beta_p \Phi Q + FTH\gamma_{p'}] - P\Phi T\varepsilon), \\ \Delta_\gamma &= -s_{x_p} (Z^\top [\gamma_{p'} V + FTH\beta_p] - P_1 HA), \end{aligned}$$

em que  $P_1$  é uma matriz  $q \times n$  de zeros exceto pela  $p'$ -ésima linha, que é composta por uns.

### 3.28 A $p'$ -ésima coluna da matriz $Z$ é função da $p$ -ésima coluna da matriz $X$

Aqui consideramos a situação em que  $z_{tp'} = \varphi(x_{tp})$ , com  $p = 2, 3, \dots, k$  e  $p' = 2, 3, \dots, q$ , em que  $\varphi$  é uma função diferenciável (se  $\varphi$  for a função identidade, teremos

o caso considerado na subseção anterior). Aqui,  $\eta_{1t}(v)$  é o como dado em (3.47) e, por exemplo, para  $p' \neq 2$  e  $p' \neq q$ , termos

$$\eta_{2t}(v) = \gamma_1 + \gamma_2 z_{t2} + \dots + \gamma_{p'} \varphi(x_{tp} + v_t S_x^p) + \dots + \gamma_q z_{tq},$$

em que  $\mu_t(v)$  e  $\phi_t(v)$  dependem de  $g(\mu_t(v)) = \eta_{1t}(v)$  e  $h(\phi_t(v)) = \eta_{2t}(v)$ . Para este tipo de perturbação, tem-se  $v_0 = (0, 0, \dots, 0)^\top$ . As matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são da seguinte forma:

$$\begin{aligned} \Delta_\beta &= -s_{x_p} (X^\top [\beta_p \Phi Q + FTH\dot{G}\gamma_{p'}] - P\Phi T\varepsilon), \\ \Delta_\gamma &= -s_{x_p} (Z^\top [\gamma_{p'} \dot{G}V + FTH\beta_p] - P_1 H A \dot{G}), \end{aligned}$$

em que  $\dot{G} = \text{diag}(\varphi'(x_{1p}), \varphi'(x_{2p}), \dots, \varphi'(x_{np}))$ .

### 3.29 Aplicação: contaminação do amendoim, modelo de regressão beta

Aplicaremos o método de influência local conforme ao conjunto de dados Contaminação do Amendoim. Utilizaremos o modelo de regressão beta para avaliar o comportamento da proporção de amendoins não contaminados ( $y$ ) através da covariável níveis de aflatoxina ( $x$ ). Consideraremos a seguir um modelo com precisão constante e com função de ligação logito para o modelo da média:

$$\log\left(\frac{w_t}{1-w_t}\right) = \beta_0 + \beta_1 \sqrt{x_t}, \quad t = 1, 2, \dots, 34.$$

Na Tabela 19 apresentamos os resultados inferenciais referentes ao modelo.

Tabela 19 – Estimativas dos parâmetros, erros-padrão e  $p$ -valores do modelo  $\log(\mu_t/(1 - \mu_t)) = \beta_0 + \beta_1 \sqrt{x_{t2}}$  e  $\phi$  constante; Modelo de regressão beta; Contaminação do Amendoim.

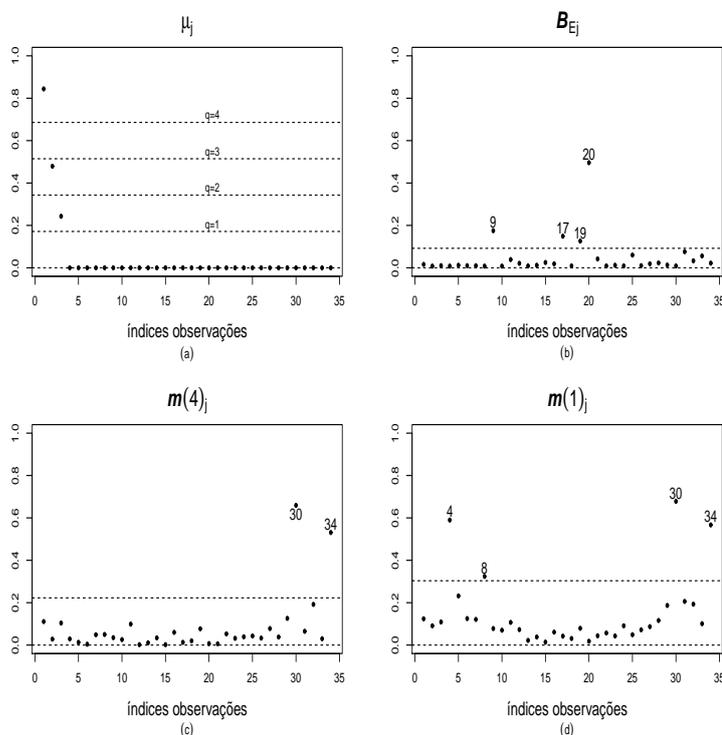
Parâmetros	Estimativas	Erros-padrão	$p$ -valores
$\beta_1$	8.8612	0.2237	<0.0001
$\beta_2$	-0.3098	0.0287	<0.0001
$\gamma_1$	8.6835	0.2521	<0.0001

Fonte: O autor (2020).

Iniciamos análise de influência considerando a técnica de influência local baseada na curvatura conforme, sob os esquemas de perturbações exibidos anteriormente. Iniciaremos utilizando o esquema de perturbação de ponderação de casos. Consideramos os valores  $q = 1, 2, 3$  e  $4$  a fim de determinar quais são os autovetores (direções) associadas às máximas variações da função de distância entre verossimilhanças. Na Figura 31(a) podemos observar

o gráfico do modulo dos autovetores normalizados e os valores de  $q$ . Temos  $n = 34$  observações e para  $q = 4$  há apenas um autovalor acima do valor de referência  $q/\sqrt{n}$ , para  $q = 2$  existem apenas dois autovalores acima do valor de referência e para  $q = 1$  há 3 autovalores acima do valor de referência, isto é, há um único autovetor 4-influente, há dois autovetores 2-influente e existem três vetores 1-influente. Os autovalores normalizados associados a estes autovetores são 0.87126, 0.43478 e 0.22774. Desta forma, a curvatura conforme máxima  $B_{l_{\max}}$  é 0.87126, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura 31(b) estão apresentados os valores das contribuições agregadas de todos os autovetores. As observações 9, 17, 19 e 20 que estão acima do valor de referência  $2b$  sugerido no trabalho de (POON; POON, 1999), são consideradas pontencialmente influentes. Na Figura 31(c)(d), temos a contribuições agregadas dos autovetores associados aos maiores autovalores, correspondentes a  $q = 4$  e  $q = 3$ . Em ambos os casos, as observações 30 e 34 são detectadas como potencialmente influentes, pois estão acima do valor de referência  $\bar{m}(q)\sqrt{2}$ .

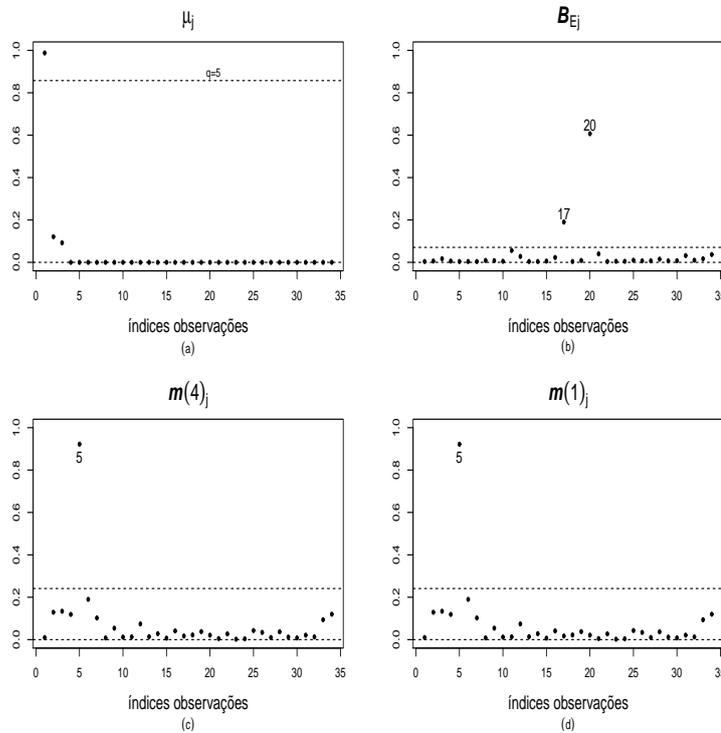
Figura 31 – Gráficos de influência local conforme sob perturbação de ponderação de casos; Modelo de regressão beta; Contaminação do Amendoim.



Fonte: O autor (2020).

Consideraremos agora o esquema de perturbação da variável resposta. Na Figura 32(a) podemos observar o gráfico do modulo dos autovetores normalizados e o valor de  $q = 4$  há apenas um autovalor acima do valor de referência  $q/\sqrt{n}$ , assim, existe um único autovetor 4-influente. O autovalor normalizado associado a este autovetor é 1.0000. Desta

Figura 32 – Gráficos de influência local conforme sob perturbação da variável resposta; Modelo de regressão beta; Contaminação do Amendoim.



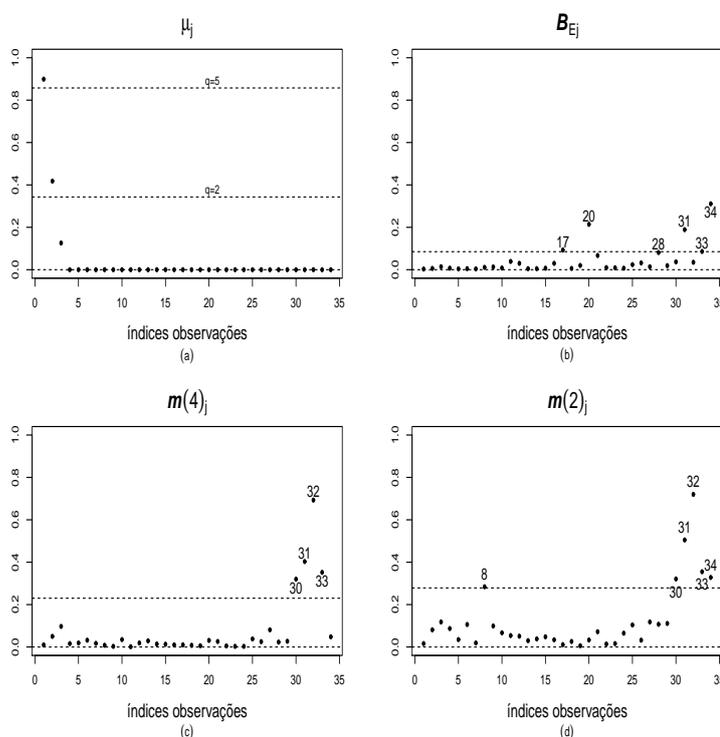
Fonte: O autor (2020).

forma, a curvatura conforme máxima  $B_{l_{\max}}$  é 1.0000, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura 32(b) estão apresentados os valores das contribuições agregadas de todos os autovetores. As observações 17 e 20, que estão acima do valor de referência  $2b$  sugerido no trabalho de (POON; POON, 1999), são consideradas potencialmente influentes. Na Figura 32(c)(d), apresentamos as contribuições agregadas dos autovetores associados aos maiores autovalores, correspondentes a  $q = 4$  e  $q = 3$ . Em ambos os casos, a observação 5 é detectada como potencialmente influentes, pois esta acima do valor de referência  $\bar{m}(q)\sqrt{2}$ .

Por fim vamos agora considerar o esquema de perturbação da covariada. Na Figura 33(a) podemos observar o gráfico dos módulos dos autovetores normalizados e indicamos os valores que  $q$ . Para  $q = 4$  há apenas um autovalor acima do valor de referência  $q/\sqrt{(n)}$ , para  $q = 2$  existem apenas dois autovalores acima do valor de referência, isto é, há um único autovetor 4-influente e dois autovetores 2-influente. Os autovalores normalizados associados a estes autovetores são 0.89947 e 0.41832 desta forma, a curvatura conforme máxima  $B_{l_{\max}}$  é 0.89947, ou seja  $l_{\max}$  é a direção que capta as maiores variações da função de distância entre verossimilhanças quando o modelo é perturbado. Na Figura

33(b) estão apresentados os valores das contribuições agregadas de todos os autovetores. Como notamos, as observações 17, 20, 28,31,33 e 41, estão acima do valor de referência  $2b$  e são consideradas pontencialmente influêntes. Na Figura 33(c), temos as contribuições agregadas dos autovetores associados aos maiores autovalores, correspondente a  $q = 4$ , notamos, as observações 30, 31, 32 e 33 detectadas como potencialmente influentes, e para  $q = 2$ , notamos as observações 8, 30, 31, 32, 33 e 34 são detectada como potêncialmente influêntes, pois estão acima do valor de referência  $\bar{m}(q)\sqrt{2}$ .

Figura 33 – Gráficos de influência local conforme sob perturbação da covariada; Modelo de regressão beta; Contaminação do Amendoim.



Fonte: O autor (2020).

Tabela 20 – Variação percentual nas estimativas dos parâmetros retirando as observações influentes utilizando o método de influência local conforme; Contaminação do Amendoim.

Casos	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\phi}$
9, 19	2.3571	7.0229	2.8319
9, 17, 20	0.3225	0.7025	5.7624
9, 19, 20	1.5689	4.8425	5.8151
17, 19, 20	0.5501	1.6229	5.4476
30, 31, 34	1.0601	6.3546	0.3451
9, 17, 19, 20	0.7321	2.2839	7.7608

Fonte: O autor (2020).

Na Tabela 20 exibimos as variações nos percentuais nas estimativas dos parâmetros do modelo após a retirada dos pontos influentes. utilizando o método de influência local conforme para o modelo de regressão beta.

### 3.30 Conclusão

Neste capítulo desenvolvemos o método de influência local e método de influência local conforme para o modelo de regressão Kumaraswamy e o método de influência local conforme para o modelo de regressão beta.

Utilizando o método de influência local para o modelo de regressão Kumaraswamy aplicado ao conjunto de dados contaminação do amendoim, concluímos que a observação 34 é individualmente influente e que os casos  $\{17, 20\}$ ,  $\{31, 32, 34\}$ ,  $\{31, 32, 33, 34\}$  e  $\{17, 19, 20, 25, 27, 29, 31, 32, 33, 34\}$  são conjuntamente influentes. Utilizando o método de influência local conforme para o modelo de regressão Kumaraswamy concluímos que a observação 34 é individualmente influente e que os casos  $\{5, 7, 34\}$ ,  $\{7, 31, 34\}$  e  $\{5, 7, 31, 34\}$  são conjuntamente influentes.

Utilizando o método de influência local para o modelo de regressão Kumaraswamy aplicado ao conjunto de dados umidade relativa do ar, concluímos que as observações 1 e 2 são individualmente influentes e que os casos  $\{1, 2\}$ ,  $\{1, 13\}$ ,  $\{27, 30\}$ ,  $\{1, 2, 10\}$ ,  $\{1, 27, 30\}$ ,  $\{10, 15, 19\}$ ,  $\{10, 25, 27\}$ ,  $\{13, 25, 27\}$ ,  $\{25, 27, 30\}$ ,  $\{10, 25, 27, 30\}$ ,  $\{13, 25, 27, 28\}$ ,  $\{13, 25, 27, 30\}$ ,  $\{1, 2, 10, 27, 30\}$ ,  $\{10, 15, 19, 25, 27\}$ ,  $\{13, 25, 27, 28, 30\}$  e  $\{10, 15, 19, 25, 27, 29, 30\}$  são conjuntamente influentes. Utilizando o método de influência local conforme para o modelo de regressão Kumaraswamy concluímos que as observações 1 e 2 são individualmente influentes e que os casos  $\{1, 2\}$ ,  $\{2, 3\}$ ,  $\{2, 4\}$ ,  $\{27, 30\}$ ,  $\{1, 2, 10\}$ ,  $\{2, 3, 11\}$ ,  $\{2, 9, 11\}$ ,  $\{1, 27, 30\}$ ,  $\{10, 25, 27\}$ ,  $\{1, 2, 10, 27\}$  e  $\{2, 9, 11, 26\}$  são conjuntamente influentes.

Ressaltamos aqui que, apesar dos gráficos de curvatura normal  $C_t$  e curvatura conforme  $B_E$  destacarem os mesmo casos, as medidas que destacam observações conjuntamente influentes são distintas.

Utilizando o método de influência local conforme para o modelo de regressão beta concluímos que os casos  $\{9, 19\}$ ,  $\{9, 17, 20\}$ ,  $\{9, 19, 20\}$ ,  $\{17, 19, 20\}$ ,  $\{30, 31, 34\}$  e  $\{9, 17, 19, 20\}$  são conjuntamente influentes.

## 4 CONCLUSÃO

No segundo capítulo da tese desenvolvemos o cálculo dos valores iniciais para o método de estimação dos parâmetros do modelo de regressão Kumaraswamy, e propusemos dois novos resíduos (resíduo ponderado e resíduo ponderado padronizado) que são baseados no processo iterativo Score de Fisher. Concluimos que para estes resíduos devemos utilizar como limites para detecção de pontos aberrante quantis empíricos, pois ao observar o comportamento destes resíduos, isto é, ao avaliar a distribuição empírica dos resíduos, notamos que as distribuições dos resíduos propostos apresentam assimetria à esquerda. Neste processo de avaliação utilizamos vários tamanhos amostrais e observamos este comportamento em todos os cenários. Concluimos que para o conjunto de dados denominado Contaminação do amendoim os resíduos ponderado e ponderado padronizado apresentam melhor desempenho que o resíduo quantílico aleatorizado, pois ao aplicarmos o modelo de regressão Kumaraswamy a este conjunto de dados pudemos observar que os resíduos propostos quantificaram com maior ênfase a observação 20, observação esta que mostramos ser atípica nesse conjunto de dados.

No Capítulo 3 desenvolvemos o método de influência local e o método de influência local conforme para o modelo de regressão Kumaraswamy e também o método de influência local conforme para o modelo de regressão beta. Utilizando o método de influência local para o modelo de regressão Kumaraswamy aplicado ao conjunto de dados contaminação do amendoim, concluimos que a observação 34 é individualmente influente e que os casos  $\{17, 20\}$ ,  $\{31, 32, 34\}$ ,  $\{31, 32, 33, 34\}$  e  $\{17, 19, 20, 25, 27, 29, 31, 32, 33, 34\}$  são conjuntamente influentes. Utilizando o método de influência local conforme para o modelo de regressão Kumaraswamy concluimos que a observação 34 é individualmente influente e que os casos  $\{5, 7, 34\}$ ,  $\{7, 31, 34\}$  e  $\{5, 7, 31, 34\}$  são conjuntamente influentes.

Utilizando o método de influência local para o modelo de regressão Kumaraswamy aplicado ao conjunto de dados umidade relativa do ar, concluimos que as observações 1 e 2 são individualmente influentes e que os casos  $\{1, 2\}$ ,  $\{1, 13\}$ ,  $\{27, 30\}$ ,  $\{1, 2, 10\}$ ,  $\{1, 27, 30\}$ ,  $\{10, 15, 19\}$ ,  $\{10, 25, 27\}$ ,  $\{13, 25, 27\}$ ,  $\{25, 27, 30\}$ ,  $\{10, 25, 27, 30\}$ ,  $\{13, 25, 27, 28\}$ ,  $\{13, 25, 27, 30\}$ ,  $\{1, 2, 10, 27, 30\}$ ,  $\{10, 15, 19, 25, 27\}$ ,  $\{13, 25, 27, 28, 30\}$  e  $\{10, 15, 19, 25, 27, 29, 30\}$  são conjuntamente influentes. Utilizando o método de influência local conforme para o modelo de regressão Kumaraswamy concluimos que as observações 1 e 2 são individualmente influentes e que os casos  $\{1, 2\}$ ,  $\{2, 3\}$ ,  $\{2, 4\}$ ,  $\{27, 30\}$ ,  $\{1, 2, 10\}$ ,  $\{2, 3, 11\}$ ,  $\{2, 9, 11\}$ ,  $\{1, 27, 30\}$ ,  $\{10, 25, 27\}$ ,  $\{1, 2, 10, 27\}$  e  $\{2, 9, 11, 26\}$  são conjuntamente influentes.

Ressaltamos aqui que, apesar dos gráficos de curvatura normal  $C_t$  e curvatura conforme  $B_E$  destacarem os mesmo casos, as medidas que destacam observações con-

juntamente influentes são distintas. Utilizando o método de influência local conforme para o modelo de regressão beta concluímos que os casos  $\{9, 19\}$ ,  $\{9, 17, 20\}$ ,  $\{9, 19, 20\}$ ,  $\{17, 19, 20\}$ ,  $\{30, 31, 34\}$  e  $\{9, 17, 19, 20\}$  são conjuntamente influentes.

# REFERÊNCIAS

- ALLEN, R. G. et al. Crop evapotranspiration-guidelines for computing crop water requirements. *fao irrigation and drainage. Fao, Rome*, v. 300, n. 9, p. D05109, 1998. Citado na página [50](#).
- ATKINSON, A. C. Two graphical displays for outlying and influential observations in regression. *Biometrika*, Oxford University Press, v. 68, n. 1, p. 13–20, 1981. Citado 2 vezes nas páginas [14](#) e [17](#).
- BALAKRISHNAN, N.; NEVZOROV, V. A Primer on Statistical Distributions. *Hoboken, New Jersey: A John Wiley & Sons. [S.l.]: Inc*, 2003. Citado 2 vezes nas páginas [13](#) e [54](#).
- BARNDORFF-NIELSEN, O. E.; JØRGENSEN, B. Some parametric models on the simplex. *Journal of Multivariate Analysis*, Academic Press, v. 39, n. 1, p. 106–116, 1991. Citado 2 vezes nas páginas [13](#) e [54](#).
- COOK, R. D. Detection of influential observation in linear regression. *Technometrics*, Taylor & Francis Group, v. 19, n. 1, p. 15–18, 1977. Citado na página [55](#).
- COOK, R. D. Assessment of local influence. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 48, n. 2, p. 133–155, 1986. Citado 6 vezes nas páginas [15](#), [28](#), [55](#), [56](#), [71](#) e [73](#).
- 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. Citado na página [27](#).
- CRIBARI-NETO, F.; ZARKOS, S. G. R. Yet another econometric programming environment. *Journal of Applied Econometrics*, Wiley Online Library, v. 14, n. 3, p. 319–329, 1999. Citado na página [16](#).
- DAVISON, A.; GIGLI, A. Deviance residuals and normal scores plots. *Biometrika*, Oxford University Press, v. 76, n. 2, p. 211–221, 1989. Citado na página [17](#).
- DOORNIK, J. An object-oriented matrix language—ox 4. *Timberlake Consultants Press, London*, 2006. Citado na página [16](#).
- DRAPER, N. R.; SMITH, H. *Applied Regression Analysis*. [S.l.]: New York, John Wiley & Sons, 1998. Citado 2 vezes nas páginas [43](#) e [80](#).
- DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, Taylor & Francis, v. 5, n. 3, p. 236–244, 1996. Citado 3 vezes nas páginas [15](#), [17](#) e [27](#).
- ESPINHEIRA, P. L.; FERRARI, S. L.; CRIBARI-NETO, F. Influence diagnostics in beta regression. *Computational Statistics & Data Analysis*, Elsevier, v. 52, n. 9, p. 4417–4431, 2008. Citado 4 vezes nas páginas [15](#), [17](#), [18](#) e [28](#).

- ESPINHEIRA, P. L.; SANTOS, E. G.; CRIBARI-NETO, F. On nonlinear beta regression residuals. *Biometrical Journal*, Wiley Online Library, v. 59, n. 3, p. 445–461, 2017. Citado na página [15](#).
- ESPINHEIRA, P. L.; SILVA, A. de O. Residual and influence analysis to a general class of simplex regression. *TEST*, Springer, p. 1–30, 2019. Citado 2 vezes nas páginas [15](#) e [17](#).
- FAHRMEIR, L.; TUTZ, G. *Multivariate Statistical Modelling Based on Generalized Linear Models*. [S.l.]: New York, Springer, 1994. Citado na página [17](#).
- FERRARI, S.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, Taylor & Francis, v. 31, n. 7, p. 799–815, 2004. Citado 6 vezes nas páginas [13](#), [24](#), [28](#), [54](#), [62](#) e [66](#).
- JONES, M. Kumaraswamy’s distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, Elsevier, v. 6, n. 1, p. 70–81, 2009. Citado 4 vezes nas páginas [14](#), [24](#), [55](#) e [63](#).
- KUMARASWAMY, P. Sinpower probability density function. *Journal of Hydrology*, Elsevier, v. 31, n. 1-2, p. 181–184, 1976. Citado 3 vezes nas páginas [14](#), [18](#) e [55](#).
- KUMARASWAMY, P. A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, Elsevier, v. 46, n. 1-2, p. 79–88, 1980. Citado 3 vezes nas páginas [14](#), [18](#) e [55](#).
- LAMPORT, L. *LATEX: A Document Preparation System: User’s Guide and Reference Manual*. [S.l.]: Boston, Addison-Wesley, 1994. Citado na página [16](#).
- LAWRANCE, A. J. Regression transformation diagnostics using local influence. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 83, n. 404, p. 1067–1072, 1988. Citado na página [15](#).
- LEMONTE, A. J. Improved point estimation for the kumaraswamy distribution. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 81, n. 12, p. 1971–1982, 2011. Citado 2 vezes nas páginas [14](#) e [55](#).
- LEMONTE, A. J.; BAZÁN, J. L. New class of johnson  $s_b$  distributions and its associated regression model for rates and proportions. *Biometrical Journal*, Wiley Online Library, v. 58, n. 4, p. 727–746, 2016. Citado 4 vezes nas páginas [13](#), [15](#), [17](#) e [54](#).
- LESAFFRE, E.; VERBEKE, G. Local influence in linear mixed models. *Biometrics.*, v. 54, p. 570–582, 1998. Citado na página [15](#).
- MCCULLAGH, P.; NELDER, J. A. *Generalized Linear Models*. [S.l.]: 2ª Edição Londres, Wiley, 1989. Citado na página [17](#).
- MITNIK, P. A. New properties of the kumaraswamy distribution. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 42, n. 5, p. 741–755, 2013. Citado 5 vezes nas páginas [14](#), [20](#), [21](#), [55](#) e [59](#).
- MITNIK, P. A.; BAEK, S. The kumaraswamy distribution: Median-dispersion re-parameterizations for regression modeling and simulation-based estimation. *Statistical Papers*, Springer, v. 54, n. 1, p. 177–192, 2013. Citado 10 vezes nas páginas [5](#), [6](#), [14](#), [15](#), [21](#), [22](#), [54](#), [55](#), [59](#) e [60](#).

- PAULA, G. A. Influence diagnostics in proper dispersion models. *Australian Journal of Statistics*, Wiley Online Library, v. 38, n. 3, p. 307–316, 1996. Citado 2 vezes nas páginas [15](#) e [17](#).
- POON, W.-Y.; POON, Y. S. Conformal normal curvature and assessment of local influence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 61, n. 1, p. 51–61, 1999. Citado 12 vezes nas páginas [15](#), [16](#), [56](#), [57](#), [72](#), [73](#), [74](#), [75](#), [84](#), [89](#), [97](#) e [98](#).
- RAO, C. R. *Linear Statistical Inference and its Applications*. [S.l.]: New York, Wiley, 1973. Citado na página [76](#).
- ROCHA, A. V.; SIMAS, A. B. Influence diagnostics in a general class of beta regression models. *Test*, Springer, v. 20, n. 1, p. 95–119, 2011. Citado 2 vezes nas páginas [15](#) e [17](#).
- SMITHSON, M.; VERKUILEN, J. A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods*, American Psychological Association, v. 11, n. 1, p. 54, 2006. Citado na página [66](#).
- THOMAS, W.; COOK, R. D. Assessing influence on predictions from generalized linear models. *Technometrics*, Taylor & Francis, v. 32, n. 1, p. 59–65, 1990. Citado na página [15](#).
- VENABLES, W. N.; RIPLEY, B. D. In: *Modern Applied Statistics with S*. [S.l.]: 4<sup>a</sup>ed. New York, Springer, 2002. Citado na página [16](#).
- VENABLES, W. N.; SMITH, D. M. *R Development Core Team (2008) An Introduction to R: Notes on R, A Programming Environment for Data Analysis and Graphics*. [S.l.]: Version, 2009. Citado na página [16](#).
- WILLIAMS, D. Residuals in generalized linear models. In: *Proceedings of the 12th International Biometrics Conference, Tokyo*. [S.l.: s.n.], 1984. p. 59–68. Citado na página [17](#).
- WILLIAMS, D. Generalized linear model diagnostics using the deviance and single case deletions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, Wiley Online Library, v. 36, n. 2, p. 181–191, 1987. Citado na página [17](#).

# APÊNDICE A – RESÍDUOS

Nessa seção apresentamos a obtenção do termo  $\kappa_{\beta\beta}$  da matriz informação de Fisher. Esse termo corresponde ao valor esperado da derivada de segunda ordem do logaritmo da função de verossimilhança. Essa matriz é de extrema importância em nosso trabalho, pois é a base para a construção dos resíduos propostos no primeiro capítulo.

Primeiramente, observe que

$$\frac{\partial l(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^n (w_t^* + y_t^*) \frac{1}{g'(w_t)} x_{ti},$$

em que

$$w_t^* = \frac{\phi_t w_t^{\phi_t}}{\log(1 - w_t^{\phi_t}) w_t (1 - w_t^{\phi_t})} \quad \text{e} \quad y_t^* = \frac{\log(0.5) \phi_t w_t^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})}.$$

Portanto,

$$\mathbb{E} \left( \frac{\partial l(\beta, \gamma)}{\partial \beta_i} \right) = \sum_{t=1}^n \left( w_t^* + \frac{\log(0.5) \phi_t w_t^{\phi_t} \mathbb{E}(\log(1 - y_t^{\phi_t}))}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t})} \right) \frac{1}{g'(w_t)} x_{ti}.$$

Seja,  $\zeta = \frac{\log(0.5)}{\log(1 - w^\phi)}$ .

**Lema 1.** *Seja  $Y$  uma variável aleatória tal que  $Y \sim Kum(\phi, \zeta)$ . Então,*

$$\mathbb{E}[\log(1 - Y^\phi)] = -\frac{1}{\zeta}.$$

*Demonstração.* Primeiramente, observe que expandindo  $\log(1 - y^\phi)$  em torno de zero temos

$$\log(1 - y^\phi) = -\sum_{k=1}^{\infty} \frac{y^{k\phi}}{k}.$$

Assim, segue que

$$\mathbb{E}[\log(1 - y^\phi)] = \mathbb{E} \left( -\sum_{k=1}^{\infty} \frac{y^{k\phi}}{k} \right) = -\sum_{k=1}^{\infty} \frac{1}{k} \mathbb{E}(y^{k\phi}).$$

Pela Proposição [2](#) segue que  $Y^{k\phi} \sim Kum(1/k, \zeta)$  e, assim,

$$\mathbb{E}[\log(1 - Y^\phi)] = -\zeta \sum_{k=1}^{\infty} \frac{1}{k} B(k + 1, \zeta).$$

Lembrando que  $\Gamma(k+1) = k\Gamma(k)$  para  $k$  inteiro positivo, temos

$$\frac{1}{k}B(k+1, \zeta) = \frac{1}{k} \frac{\Gamma(k+1)\Gamma(\zeta)}{\Gamma(\zeta+k+1)} = \frac{k!\Gamma(\zeta)}{k\Gamma(\zeta) \prod_{i=0}^k (\zeta+i)} = \frac{(k-1)!}{\prod_{i=0}^k (\zeta+i)} = \frac{1}{\zeta(\zeta+1)} \left( \frac{\zeta+k}{\zeta+1} \right)^{-1}.$$

Portanto,

$$\mathbb{E}[\log(1-y^\phi)] = -\frac{1}{(\zeta+1)} \sum_{k=1}^{\infty} \left( \frac{\zeta+k}{\zeta+1} \right)^{-1} = -\frac{1}{(\zeta+1)} \left[ 1 + \frac{1}{\zeta} \right] = -\frac{1}{\zeta}.$$

□

Assim, pelo Lema [1](#), segue que

$$\mathbb{E} \left( \frac{\partial l(\beta, \gamma)}{\partial \beta_i} \right) = \sum_{t=1}^n \mathbb{E}(y_t^* + w_t^*) \frac{1}{g'(w_t)} x_{ti} = 0.$$

Este fato nos ajudará no nosso objetivo, como veremos a seguir. Como mostrado no Capítulo 2,

$$\frac{\partial^2 l(\beta, \gamma)}{\partial \beta_j \partial \beta_i} = \sum_{t=1}^n \frac{\partial^2 l(w_t, \phi_t)}{\partial w_t^2} \left( \frac{\partial w_t}{\partial \eta_{1t}} \right)^2 \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i} + \sum_{t=1}^n \frac{\partial l(w_t, \phi_t)}{\partial w_t} \left[ \frac{\partial^2 w_t}{\partial \eta_{1t}^2} \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i} \right].$$

Pelo que verificamos anteriormente temos que o valor esperado do segundo termo dessa expressão é nulo, restando calcular apenas o valor esperado da primeira parcela, isto é,

$$\mathbb{E} \left( \frac{\partial^2 l(\beta, \gamma)}{\partial \beta_j \partial \beta_i} \right) = \mathbb{E} \left( \sum_{t=1}^n \frac{\partial^2 l(w_t, \phi_t)}{\partial w_t^2} \left( \frac{\partial w_t}{\partial \eta_{1t}} \right)^2 \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i} \right).$$

Temos

$$\begin{aligned} \mathbb{E} \left( \frac{\partial^2 l(\beta, \gamma)}{\partial \beta_j \partial \beta_i} \right) &= \sum_{t=1}^n \left( \frac{w_t^{\phi_t} \phi_t^2}{w_t^2 \log(1-w_t^{\phi_t})(1-w_t^{\phi_t})} + \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1-w_t^{\phi_t})^2 w_t^2 (1-w_t^{\phi_t})^2} \right. \\ &\quad - \frac{\log(0.5) w_t^{\phi_t} \phi_t \mathbb{E}[\log(1-y_t^{\phi_t})]}{\log(1-w_t^{\phi_t})^2 w_t^2 (1-w_t^{\phi_t})} + \frac{\log(0.5) (w_t^{\phi_t})^2 \phi_t^2 \mathbb{E}[\log(1-y_t^{\phi_t})]}{\log(1-w_t^{\phi_t})^2 w_t^2 (1-w_t^{\phi_t})^2} \\ &\quad + \frac{\log(0.5) w_t^{\phi_t} \phi_t^2 \mathbb{E}[\log(1-y_t^{\phi_t})]}{w_t^2 \log(1-w_t^{\phi_t})^2 (1-w_t^{\phi_t})} + \frac{2 \log(0.5) (w_t^{\phi_t})^2 \phi_t^2 \mathbb{E}[\log(1-y_t^{\phi_t})]}{\log(1-w_t^{\phi_t})^3 w_t^2 (1-w_t^{\phi_t})^2} \\ &\quad \left. - \frac{w_t^{\phi_t} \phi_t}{\log(1-w_t^{\phi_t}) w_t^2 (1-w_t^{\phi_t})} + \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1-w_t^{\phi_t}) w_t^2 (1-w_t^{\phi_t})^2} \right) \left( \frac{\partial w_t}{\partial \eta_{1t}} \right)^2 \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i}. \end{aligned}$$

Pelo Lema [1](#), segue que

$$\mathbb{E} \left( \frac{\partial^2 l(\beta, \gamma)}{\partial \beta_j \partial \beta_i} \right) = - \sum_{t=1}^n \left( \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1-w_t^{\phi_t})^2 w_t^2 (1-w_t^{\phi_t})^2} \right) \left( \frac{\partial w_t}{\partial \eta_{1t}} \right)^2 \frac{\partial \eta_{1t}}{\partial \beta_j} \frac{\partial \eta_{1t}}{\partial \beta_i}.$$

Como  $\kappa_{\beta\beta} = \mathbb{E} \left( -\frac{\partial^2 l(\beta, \gamma)}{\partial \beta_j \partial \beta_i} \right)$ , segue que

$$\kappa_{\beta_i \beta_j} = \sum_{t=1}^n \left( \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1 - w_t^{\phi_t})^2 w_t^2 (1 - w_t^{\phi_t})^2} \right) \left( \frac{\partial w_t}{\partial \eta_{1t}} \right)^2 \frac{\partial \eta_{1t}}{\partial \beta_i} \frac{\partial \eta_{1t}}{\partial \beta_j}.$$

Em forma matricial, tem-se

$$\kappa_{\beta\beta} = X^\top W_{\beta\beta} X,$$

em que  $W_{\beta\beta} = \text{diag}(c_1, c_2, \dots, c_n)$ , com  $c_t = \frac{(w_t^{\phi_t})^2 \phi_t^2}{\log(1 - w_t^{\phi_t})^2 w_t^2 (1 - w_t^{\phi_t})^2} \left( \frac{\partial w_t}{\partial \eta_{1t}} \right)^2$  para  $t = 1, 2, \dots, n$ .

# APÊNDICE B – INFLUÊNCIA LOCAL

## Cálculo das perturbações no modelo de regressão Kumaraswamy

No modelo de regressão Kumaraswamy definido em (3.10), temos que  $\theta = (\beta_1, \beta_2, \dots, \beta_k, \gamma_1, \gamma_2, \dots, \gamma_q)^\top$ . Então,  $\Delta$  é uma matriz  $(k + q) \times n$  dada por

$$\Delta = \begin{bmatrix} \frac{\partial^2 l_\delta(\theta)}{\partial \delta_1 \partial \beta_1} & \dots & \frac{\partial^2 l_\delta(\theta)}{\partial \delta_n \partial \beta_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 l_\delta(\theta)}{\partial \delta_1 \partial \beta_k} & \dots & \frac{\partial^2 l_\delta(\theta)}{\partial \delta_n \partial \beta_k} \\ \frac{\partial^2 l_\delta(\theta)}{\partial \delta_1 \partial \gamma_1} & \dots & \frac{\partial^2 l_\delta(\theta)}{\partial \delta_n \partial \gamma_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 l_\delta(\theta)}{\partial \delta_1 \partial \gamma_q} & \dots & \frac{\partial^2 l_\delta(\theta)}{\partial \delta_n \partial \gamma_q} \end{bmatrix},$$

avaliada em  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$  e em  $\delta_0$ . Podemos particionar a matriz  $\Delta$  como

$$\Delta = \begin{bmatrix} \Delta_\beta \\ \Delta_\gamma \end{bmatrix},$$

em que  $\Delta_\beta$  é uma matriz  $k \times n$  e  $\Delta_\gamma$  é uma matriz  $q \times n$ , sendo tais matrizes dadas por

$$\Delta_\beta = \left. \frac{\partial^2 \ell(\theta|\delta)}{\partial \beta \partial \delta^\top} \right|_{(\theta=\hat{\theta}, \delta=\delta_0)} \quad \text{e} \quad \Delta_\gamma = \left. \frac{\partial^2 \ell(\theta|\delta)}{\partial \gamma \partial \delta^\top} \right|_{(\theta=\hat{\theta}, \delta=\delta_0)},$$

ou seja, ambas são avaliadas em  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$  e em  $\delta_0$ .

### Perturbação ponderados de casos

No esquema de perturbação de casos ponderados, temos que

$$\ell(\theta|\delta) = \sum_{t=1}^n \delta_t \ell_t(w_t, \phi_t).$$

Assim,

$$\frac{\partial \ell(\theta|\delta)}{\partial \delta_t} = \ell_t(w_t, \phi_t). \quad (\text{B.1})$$

Da Equação (B.1) segue que

$$\frac{\partial^2 \ell(\theta|\delta)}{\partial \beta_i \partial \delta_t} = (w_t^* + y_t^*) \frac{1}{g'(w_t)} x_{ti}$$

e também

$$\frac{\partial^2 \ell(\theta|\delta)}{\partial \gamma_j \partial \delta_t} = [(y_t^* + w_t^*)\alpha_{1t} + \alpha_{2t}] \frac{1}{h'(\phi_t)} z_{tj}.$$

Tomando  $A_1 = \text{diag}(w^* + y^*)$  e  $A_2 = \text{diag}([(y^* + w^*)\alpha_1 + \alpha_2])$ , matrizes de dimensão  $n \times n$ , temos que

$$\begin{aligned} \Delta_\beta &= X^\top T A_1, \\ \Delta_\gamma &= Z^\top H A_2. \end{aligned}$$

### Perturbação da variável resposta

Considere a perturbação aditiva dada por  $y_{t\delta} = y_t + \delta_t S_y^t$ , em que  $S_y^t$  é um escalar que pode, por exemplo, para cada  $t = 1, 2, \dots, n$ , ser igual à distância interquartílica estimada em  $y = (y_1, y_2, \dots, y_n)^\top$ . Neste caso, a função log-verossimilhança perturbada é

$$\ell(\theta|\delta) = \sum_{t=1}^n \ell_t(w_t, \phi_t, y_{t\delta}),$$

em que

$$\begin{aligned} \ell_t(w_t, \phi_t, y_{t\delta}) &= \log(\phi_t) + \log\left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})}\right) + (\phi_t - 1) \log(y_{t\delta}) \\ &+ \left(\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right) \log(1 - y_{t\delta}^{\phi_t}). \end{aligned} \quad (\text{B.2})$$

Assim, da Equação (B.2) segue que

$$\frac{\partial \ell(\theta|\delta)}{\partial \delta_t} = \frac{(\phi - 1) S_y^t}{S_y^t \delta_t + y_t} - \frac{\left[\frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1\right] [(S_y^t \delta_t + y_t)^{\phi_t} \phi_t S_y^t]}{(S_y^t \delta_t + y_t)(1 - (S_y^t \delta_t + y_t)^{\phi_t})}. \quad (\text{B.3})$$

Partindo da Equação (B.3), tem-se

$$\frac{\partial^2 \ell(\theta|\delta)}{\partial \beta_i \partial \delta_t} = m_{1t} \frac{1}{g'(w_t)} x_{ti},$$

em que

$$m_{1t} = - \frac{\log(0.5) w_t^{\phi_t} \phi_t^2 (S_y^t \delta_t + y_t)^{\phi_t} S_y^t}{\log(1 - w_t^{\phi_t})^2 w_t (1 - w_t^{\phi_t}) (S_y^t \delta_t + y_t) (1 - (S_y^t \delta_t + y_t)^{\phi_t})},$$

com  $t = 1, 2, \dots, n$ . Então, definimos  $m_1 = (m_{11}, m_{12}, \dots, m_{1n})^\top$  e  $M1 = \text{diag}(m_1)$ .

Ainda da Equação (B.3) obtemos

$$\frac{\partial^2 l_\delta(\theta)}{\partial \gamma_j \partial \delta_t} = m_{2t} \frac{1}{h'(\phi_t)} z_{tj},$$

em que

$$m_{2t} = \frac{S_y^t}{S_y^t \delta_t + y_t} - \frac{\log(0.5) w_t^{\phi_t} \log(w_t) (S_y^t \delta_t + y_t)^{\phi_t} \phi_t S_y^t}{\log(1 - w_t^{\phi_t})^2 (1 - w_t^{\phi_t}) (S_y^t \delta_t + y_t) (1 - (S_y^t \delta_t + y_t)^{\phi_t})} \\ - \frac{\left( \frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1 \right) (S_y^t \delta_t + y_t)^{\phi_t} \log(S_y^t \delta_t + y_t) \phi_t S_y^t}{(S_y^t \delta_t + y_t) (1 - (S_y^t \delta_t + y_t)^{\phi_t})} \\ - \frac{\left( \frac{\ln(0.5)}{\ln(1 - w_t^{\phi_t})} - 1 \right) (S_y^t \delta_t + y_t)^{\phi_t} S_y^t}{(S_y^t \delta_t + y_t) (1 - (S_y^t \delta_t + y_t)^{\phi_t})} - \frac{\left( \frac{\log(0.5)}{\log(1 - w_t^{\phi_t})} - 1 \right) ((S_y^t \delta_t + y_t)^{\phi_t})^2 \phi_t S_y^t \log(S_y^t \delta_t + y_t)}{(S_y^t \delta_t + y_t) (1 - (S_y^t \delta_t + y_t)^{\phi_t})^2},$$

com  $t = 1, 2, \dots, n$ . Definimos  $m_2 = (m_{21}, m_{22}, \dots, m_{2n})^\top$  e  $M_2 = \text{diag}(m_2)$ .

Desta forma, obtemos que as matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são dadas por

$$\begin{aligned} \Delta_\beta &= X^\top T M_1, \\ \Delta_\gamma &= Z^\top H M_2. \end{aligned}$$

### Perturbação individual de covariáveis

A seguir obteremos as matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  com base na perturbação aditiva da  $p$ -ésima covariável  $x_p$ ,  $p = 2, \dots, k$ , isto é,  $x_{tp\delta} = x_{tp} + \delta_t S_x^p$ , considerando o modelo definido em (3.10) e os cenários descritos na Seção 3.13.

### Matriz $Z$ totalmente diferente da matriz $X$

Neste caso, a função de log-verossimilhança do modelo perturbado é dada por

$$\ell(\theta|\delta) = \sum_{t=1}^n \ell_t(w_t(\delta), \phi_t, y_t),$$

em que

$$\begin{aligned} \ell_t(w_t(\delta), \phi_t, y_t) &= \log(\phi_t) + \log\left(\frac{\log(0.5)}{\log(1 - w_t(\delta)^{\phi_t})}\right) \\ &+ (\phi_t - 1) \log(y_t) + \left(\frac{\log(0.5)}{\log(1 - w_t(\delta)^{\phi_t})} - 1\right) \log(1 - y_t^{\phi_t}). \end{aligned} \quad (\text{B.4})$$

Aqui,  $w_t(\delta)$  é tal que  $g(w_t(\delta)) = \eta_{1t}(\delta)$ , com  $\eta_{1t}(\delta)$  é como dado em (3.42). Assim, temos que

$$\frac{\partial \ell(\theta|\delta)}{\partial \delta_t} = (w_t^*(\delta) + y_t^*(\delta)) \frac{1}{g'(w_t(\delta))} \beta_p S_x^p, \quad (\text{B.5})$$

em que

$$w_t^*(\delta) = \frac{\phi_t w_t(\delta)^{\phi_t}}{\log(1 - w_t(\delta)^{\phi_t}) w_t(\delta) (1 - w_t(\delta)^{\phi_t})} \quad \text{e} \quad y_t^*(\delta) = \frac{\log(0.5) \phi_t w_t(\delta)^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t})}.$$

De [\(B.5\)](#) segue que, para  $i = 1, 2, \dots, k$ , com  $i \neq p$ , temos

$$\frac{\partial^2 \ell(\theta|\delta)}{\partial \beta_i \partial \delta_t} = s_{2t} \frac{1}{g'(w_t(\delta))^2} \beta_p S_x^p x_{ti} + (w_t^*(\delta) + y_t^*(\delta)) \frac{g''(w_t(\delta))}{g'(w_t(\delta))} \frac{1}{g'(w_t(\delta))^2} \beta_p S_x^p x_{ti}.$$

Quando  $i = p$ , temos que

$$\begin{aligned} \frac{\partial^2 \ell(\theta|\delta)}{\partial \beta_p \partial \delta_t} &= s_{2t} \frac{1}{g'(w_t(\delta))^2} \beta_p S_x^p (x_{tp} + \delta_t S_x^p) + (w_t^*(\delta) + y_t^*(\delta)) \frac{1}{g'(w_t(\delta))} S_x^p \\ &\quad + (w_t^*(\delta) + y_t^*(\delta)) \frac{g''(w_t(\delta))}{g'(w_t(\delta))} \frac{1}{g'(w_t(\delta))^2} \beta_p S_x^p (x_{tp} + \delta_t S_x^p), \end{aligned}$$

em que

$$\begin{aligned} s_{2t} &= \frac{w_t(\delta)^{\phi_t} \phi_t^2}{w_t(\delta)^2 \log(1 - w_t(\delta)^{\phi_t}) (1 - w_t(\delta)^{\phi_t})} + \frac{(w_t(\delta)^{\phi_t})^2 \phi_t^2}{\log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t})^2} \\ &\quad - \frac{w_t(\delta)^{\phi_t} \phi_t}{\log(1 - w_t(\delta)^{\phi_t}) w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t})} + \frac{(w_t(\delta)^{\phi_t})^2 \phi_t^2}{\log(1 - w_t(\delta)^{\phi_t}) w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t})^2} \\ &\quad + \frac{\log(0.5) w_t(\delta)^{\phi_t} \phi_t^2 \log(1 - y_t^{\phi_t})}{w_t(\delta)^2 \log(1 - w_t(\delta)^{\phi_t})^2 (1 - w_t(\delta)^{\phi_t})} + \frac{2 \log(0.5) (w_t(\delta)^{\phi_t})^2 \phi_t^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t(\delta)^{\phi_t})^3 w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t})^2} \\ &\quad - \frac{\log(0.5) w_t(\delta)^{\phi_t} \phi_t \log(1 - y_t^{\phi_t})}{\log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t})} + \frac{\log(0.5) (w_t(\delta)^{\phi_t})^2 \phi_t^2 \log(1 - y_t^{\phi_t})}{\log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t})^2}, \end{aligned}$$

com  $t = 1, 2, \dots, n$ . Então, definimos  $s_2 = (s_{21}, s_{22}, \dots, s_{2n})$  e  $S_2 = \text{diag}(s_2)$ , uma matriz de dimensão  $n \times n$ . Adicionalmente  $T_2$  é uma matriz diagonal onde o  $t$ -ésimo elemento diagonal é  $T_{2tt} = \frac{g''(w_t(\delta))}{g'(w_t(\delta))} \frac{1}{g'(w_t(\delta))^2}$ .

Segue ainda de [\(B.5\)](#) que, para  $j = 1, 2, \dots, q$ , temos

$$\frac{\partial^2 \ell(\theta|\delta)}{\partial \gamma_j \partial \delta_t} = s_{3t} \frac{1}{g'(w_t(\delta))} \frac{1}{h'(\phi_t)} z_{tj} \beta_p S_x^p,$$

em que

$$\begin{aligned} s_{3t} &= \frac{w_t(\delta)^{\phi_t} \log(w_t(\delta)) \phi_t}{\log(1 - w_t(\delta)^{\phi_t}) w_t(\delta) (1 - w_t(\delta)^{\phi_t})} + \frac{w_t(\delta)^{\phi_t}}{\log(1 - w_t(\delta)^{\phi_t}) w_t(\delta) (1 - w_t(\delta)^{\phi_t})} \\ &\quad + \frac{(w_t(\delta)^{\phi_t})^2 \phi_t \log(w_t(\delta))}{\log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t})^2} + \frac{(w_t(\delta)^{\phi_t})^2 \phi_t \log(w_t(\delta))}{\log(1 - w_t(\delta)^{\phi_t}) w_t(\delta) (1 - w_t(\delta)^{\phi_t})^2} \\ &\quad + \frac{\log(0.5) w_t(\delta)^{\phi_t} \log(w_t(\delta)) \phi_t \log(1 - y_t^{\phi_t})}{\log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t})} + \frac{\log(0.5) w_t(\delta)^{\phi_t} \log(1 - y_t^{\phi_t})}{\log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t})} \\ &\quad - \frac{\log(0.5) w_t(\delta)^{\phi_t} \phi_t y_t^{\phi_t} \log(y_t)}{(1 - y_t^{\phi_t}) \log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t})} + \frac{2 \log(0.5) (w_t(\delta)^{\phi_t})^2 \phi_t \log(1 - y_t^{\phi_t}) \log(w_t(\delta))}{\log(1 - w_t(\delta)^{\phi_t})^3 w_t(\delta) (1 - w_t(\delta)^{\phi_t})^2} \\ &\quad + \frac{\log(0.5) (w_t(\delta)^{\phi_t})^2 \phi_t \log(1 - y_t^{\phi_t}) \log(w_t(\delta))}{\log(1 - w_t(\delta)^{\phi_t})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t})^2}, \end{aligned}$$

com  $t = 1, 2, \dots, n$ . Definimos  $s_3 = (s_{31}, s_{32}, \dots, s_{3n})$  e  $S_3 = \text{diag}(s_3)$  uma matriz de dimensão  $n \times n$ .

Avaliando as expressões anteriores em  $\delta_0 = (0, 0, \dots, 0)^\top$  e em  $\hat{\theta}$ , obtemos que as matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são da seguinte forma:

$$\begin{aligned}\Delta_\beta &= S_x^p(\beta_p X^\top T^2 S_2 + \beta_p X^\top A_1(\delta) T_2 - P_1 A_1(\delta) T), \\ \Delta_\gamma &= S_x \beta_p (Z^\top H T S_3).\end{aligned}$$

### Matriz $Z$ igual à matriz $X$

Neste caso, a função de log-verossimilhança do modelo perturbado é

$$\ell(\theta|\delta) = \sum_{t=1}^n \ell_t(w_t(\delta), \phi_t(\delta), y_{t\delta}),$$

em que

$$\begin{aligned}\ell_t(w_t(\delta), \phi_t(\delta), y_t) &= \log(\phi_t(\delta)) + \log\left(\frac{\log(0.5)}{\log(1 - w_t(\delta)^{\phi_t(\delta)})}\right) + (\phi_t(\delta) - 1) \log(y_t) \\ &+ \left(\frac{\log(0.5)}{\log(1 - w_t(\delta)^{\phi_t(\delta)})} - 1\right) \log(1 - y_t^{\phi_t(\delta)}).\end{aligned}\tag{B.6}$$

Aqui,  $w_t(\delta)$  e  $\phi_t(\delta)$  dependem de  $g(w_t(\delta)) = \eta_{1t}(\delta)$  e  $h(\phi_t(\delta)) = \eta_{2t}(\delta)$ , que estão dados [\(3.42\)](#) e [\(3.43\)](#), respectivamente. Assim, temos que

$$\frac{\partial \ell(\theta|\delta)}{\partial \delta_t} = (w_t^*(\delta) + y_t^*(\delta)) \frac{1}{g'(w_t(\delta))} \beta_p S_x^p + [(y_t^*(\delta) + w_t^*(\delta)) \alpha_{1t}(\delta) + \alpha_{2t}(\delta)] \frac{1}{h'(\phi_t(\delta))} \gamma_p S_x^p.\tag{B.7}$$

De [\(B.7\)](#) segue que, para  $i = 1, 2, \dots, k$ , com  $i \neq p$ , temos

$$\frac{\partial^2 \ell(\theta|\delta)}{\partial \beta_i \partial \delta_t} = s_{10t}(\delta) \frac{1}{h'(\phi_t(\delta))} \frac{1}{g'(w_t(\delta))} \gamma_p S_x^p x_{ti} + s_{11t} \frac{1}{g'(w_t(\delta))^2} \beta_p S_x^p x_{ti} + s_{9t} \frac{1}{g'(w_t(\delta))^2} \beta_p S_x^p x_{ti}.$$

Quando  $i = p$ ,

$$\begin{aligned}\frac{\partial^2 \ell(\theta|\delta)}{\partial \beta_i \partial \delta_t} &= s_{10t}(\delta) \frac{1}{h'(\phi_t(\delta))} \frac{1}{g'(w_t(\delta))} \gamma_p S_x^p (x_{ti} + \delta_t S_x^p) + s_{11t} \frac{1}{g'(w_t(\delta))^2} \beta_p S_x^p (x_{ti} + \delta_t S_x^p) \\ &+ s_{9t} \frac{1}{g'(w_t(\delta))^2} \beta_p S_x^p (x_{ti} + \delta_t S_x^p) + s_{9t} \frac{1}{g'(w_t(\delta))} S_x^p.\end{aligned}$$

em que  $s_{9t} = (w_t^*(\delta) + y_t^*(\delta))$ , com  $t = 1, 2, \dots, n$ . Definimos o vetor  $s_9 = (s_{91}, s_{92}, \dots, s_{9n})$  e  $S_9 = \text{diag}(s_9)$  uma matriz de dimensão  $n \times n$ .

Seja

$$\begin{aligned}
 s_{10t} = & \frac{w_t(\delta)^{\phi_t(\delta)} \log(w_t(\delta)) \phi_t(\delta)}{\log(1 - w_t(\delta)^{\phi_t(\delta)}) w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})} + \frac{w_t(\delta)^{\phi_t(\delta)}}{\log(1 - w_t(\delta)^{\phi_t(\delta)}) w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})} \\
 & + \frac{(w_t(\delta)^{\phi_t(\delta)})^2 \phi_t(\delta) \log(w_t(\delta))}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})^2} + \frac{(w_t(\delta)^{\phi_t(\delta)})^2 \phi_t(\delta) \log(w_t(\delta))}{\log(1 - w_t(\delta)^{\phi_t(\delta)}) w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})^2} \\
 & + \frac{\log(0.5) w_t(\delta)^{\phi_t(\delta)} \log(w_t(\delta)) \phi_t(\delta) \log(1 - y_t^{\phi_t(\delta)})}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})} + \frac{\log(0.5) w_t(\delta)^{\phi_t(\delta)} \log(1 - y_t^{\phi_t(\delta)})}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})} \\
 & - \frac{\log(0.5) w_t(\delta)^{\phi_t(\delta)} \phi_t y_t^{\phi_t(\delta)} \log(y_t)}{(1 - y_t^{\phi_t(\delta)}) \log(1 - w_t(\delta)^{\phi_t(\delta)})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})} \\
 & + \frac{2 \log(0.5) (w_t(\delta)^{\phi_t(\delta)})^2 \phi_t(\delta) \log(1 - y_t^{\phi_t(\delta)}) \log(w_t(\delta))}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^3 w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})^2} \\
 & + \frac{\log(0.5) (w_t(\delta)^{\phi_t(\delta)})^2 \phi_t(\delta) \log(1 - y_t^{\phi_t(\delta)}) \log(w_t(\delta))}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 w_t(\delta) (1 - w_t(\delta)^{\phi_t(\delta)})^2},
 \end{aligned}$$

com  $t = 1, 2, \dots, n$ . Definimos o vetor  $s_{10} = (s_{101}, s_{102}, \dots, s_{10n})$  e  $S_{10} = \text{diag}(s_{10})$ , uma matriz de dimensão  $n \times n$ .

Seja

$$\begin{aligned}
 s_{11t} = & \frac{w_t(\delta)^{\phi_t(\delta)} \phi_t(\delta)^2}{w_t(\delta)^2 \log(1 - w_t(\delta)^{\phi_t(\delta)}) (1 - w_t(\delta)^{\phi_t(\delta)})} + \frac{(w_t(\delta)^{\phi_t(\delta)})^2 \phi_t(\delta)^2}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t(\delta)})^2} \\
 & - \frac{w_t(\delta)^{\phi_t(\delta)} \phi_t(\delta)}{\log(1 - w_t(\delta)^{\phi_t(\delta)}) w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t(\delta)})} + \frac{(w_t(\delta)^{\phi_t(\delta)})^2 \phi_t(\delta)^2}{\log(1 - w_t(\delta)^{\phi_t(\delta)}) w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t(\delta)})^2} \\
 & + \frac{\log(0.5) w_t(\delta)^{\phi_t(\delta)} \phi_t(\delta)^2 \log(1 - y_t^{\phi_t(\delta)})}{w_t(\delta)^2 \log(1 - w_t(\delta)^{\phi_t(\delta)})^2 (1 - w_t(\delta)^{\phi_t(\delta)})} + \frac{2 \log(0.5) (w_t(\delta)^{\phi_t(\delta)})^2 \phi_t(\delta)^2 \log(1 - y_t^{\phi_t(\delta)})}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^3 w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t(\delta)})^2} \\
 & - \frac{\log(0.5) w_t(\delta)^{\phi_t(\delta)} \phi_t(\delta) \log(1 - y_t^{\phi_t(\delta)})}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t(\delta)})} + \frac{\log(0.5) (w_t(\delta)^{\phi_t(\delta)})^2 \phi_t(\delta)^2 \log(1 - y_t^{\phi_t(\delta)})}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 w_t(\delta)^2 (1 - w_t(\delta)^{\phi_t(\delta)})^2},
 \end{aligned}$$

com  $t = 1, 2, \dots, n$ . Definimos o vetor  $s_{11} = (s_{111}, s_{112}, \dots, s_{11n})$  e  $S_{11} = \text{diag}(s_{11})$ , uma matriz de dimensão  $n \times n$ .

Neste caso, para calcular as segundas derivadas com relação ao vetor  $\gamma$  também teremos que considerar  $j \neq p$  e  $j = p$ . Assim, com base em [\[B.7\]](#) e para  $j = 1, 2, \dots, q$ , com  $j \neq p$ , temos que

$$\frac{\partial^2 \ell(\theta|\delta)}{\partial \gamma_j \partial \delta_t} = s_{12t} \frac{1}{h'(\phi_t(\delta))^2} \gamma_p S_x^p x_{tj} + s_{13t} \frac{1}{h'(\phi_t(\delta))} \frac{1}{g'(w_t(\delta))} \beta_p S_x^p x_{tj} + s_{8t} \frac{h''(\phi_t(\delta))}{h'(\phi_t(\delta))} \frac{1}{h'(\phi_t(\delta))} x_{tj} \gamma_p S_x^p.$$

Quando  $j = p$ , temos

$$\begin{aligned}
 \frac{\partial^2 \ell(\theta|\delta)}{\partial \gamma_j \partial \delta_t} = & s_{12t} \frac{1}{h'(\phi_t(\delta))^2} \gamma_p S_x^p (x_{tj} + \delta_t S_x^p) + s_{13t} \frac{1}{h'(\phi_t(\delta))} \frac{1}{g'(w_t(\delta))} \beta_p S_x^p (x_{tj} + \delta_t S_x^p) \\
 & + s_{8t} \frac{h''(\phi_t(\delta))}{h'(\phi_t(\delta))^2} \frac{1}{h'(\phi_t(\delta))} (x_{tj} + \delta_t S_x^p) \gamma_p S_x^p + s_{8t} \frac{1}{h'(\phi_t(\delta))} S_x^p,
 \end{aligned}$$

em que  $s_{8t} = [(y_t^*(\delta) + w_t^*(\delta)) \alpha_{1t}(\delta) + \alpha_{2t}(\delta)]$ , com  $t = 1, 2, \dots, n$ . Definimos o vetor  $s_8 = (s_{81}, s_{82}, \dots, s_{8n})$  e  $S_8 = \text{diag}(s_8)$  uma matriz de dimensão  $n \times n$ .

Seja

$$\begin{aligned}
 s_{12t} = & \frac{w_t(\delta)^{\phi_t(\delta)} \log(w_t(\delta))^2}{\log(1 - w_t(\delta)^{\phi_t(\delta)})(1 - w_t(\delta)^{\phi_t(\delta)})} - \frac{2 \log(0.5) w_t(\delta)^{\phi_t(\delta)} \log(w_t(\delta)) y_t^{\phi_t(\delta)} \log(y_t)}{(1 - y_t^{\phi_t(\delta)}) \log(1 - w_t(\delta)^{\phi_t(\delta)})^2 (1 - w_t(\delta)^{\phi_t(\delta)})} \\
 & + \frac{(w_t(\delta)^{\phi_t(\delta)})^2 \log(w_t(\delta))^2}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 (1 - w_t(\delta)^{\phi_t(\delta)})^2} + \frac{\log(0.5) w_t(\delta)^{\phi_t(\delta)} \log(w_t(\delta))^2 \log(1 - y_t^{\phi_t(\delta)})}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 (1 - w_t(\delta)^{\phi_t(\delta)})} \\
 & + \frac{(w_t(\delta)^{\phi_t(\delta)})^2 \log(w_t(\delta))^2}{\log(1 - w_t(\delta)^{\phi_t(\delta)})(1 - w_t(\delta)^{\phi_t(\delta)})^2} + \frac{2 \log(0.5) (w_t(\delta)^{\phi_t(\delta)})^2 \log(w_t(\delta))^2 \log(1 - y_t^{\phi_t(\delta)})}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^3 (1 - w_t(\delta)^{\phi_t(\delta)})^2} \\
 & + \frac{\log(0.5) (w_t(\delta)^{\phi_t(\delta)})^2 \log(w_t(\delta))^2 \log(1 - y_t^{\phi_t(\delta)})}{\log(1 - w_t(\delta)^{\phi_t(\delta)})^2 (1 - w_t(\delta)^{\phi_t(\delta)})^2} - \frac{1}{\phi_t(\delta)^2} \\
 & - \frac{\left[ \frac{\log(0.5)}{\log(1 - w_t(\delta)^{\phi_t(\delta)})} - 1 \right] y_t^{\phi_t(\delta)} \log(y_t)^2}{1 - y_t^{\phi_t(\delta)}} - \frac{\left[ \frac{\log(0.5)}{\log(1 - w_t(\delta)^{\phi_t(\delta)})} - 1 \right] (y_t^{\phi_t(\delta)})^2 \log(y_t)^2}{(1 - y_t^{\phi_t(\delta)})^2},
 \end{aligned}$$

com  $t = 1, 2, \dots, n$ . Definimos o vetor  $s_{12} = (s_{121}, s_{122}, \dots, s_{12n})$ , a matriz  $S_{12} = \text{diag}(s_{12})$ , a matriz  $S_{13} = S_{10}$  e  $H_2$  é uma matriz cujo  $t$ -ésimo elemento diagonal  $H_{2tt} = \frac{h''(\phi_t(\delta))}{h'(\phi_t(\delta))^2} \frac{1}{h'(\phi_t(\delta))}$ .

Avaliando as expressões anteriores em  $\delta_0 = (0, 0, \dots, 0)^\top$  e em  $\hat{\theta}$  obtemos que as matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são da seguinte forma:

$$\begin{aligned}
 \Delta_\beta &= S_x^p (\gamma_p X^\top S_{10} H T + \beta_p X^\top S_{11} T^2 + \beta_p X^\top S_9 T^2 - P_1 S_9 T), \\
 \Delta_\gamma &= S_x^p (\gamma_p X^\top S_{12} H^2 + \beta_p X^\top S_{13} T H + \gamma_p X^\top S_8 H_2 - P_1 S_8 H).
 \end{aligned}$$

**A  $p'$ -ésima coluna da matriz  $Z$  é igual à  $p$ -ésima coluna da matriz  $X$**

Neste caso, a função de log-verossimilhança do modelo perturbado é

$$\ell(\theta|\delta) = \sum_{t=1}^n \ell_t(w_t(\delta), \phi_t(\delta), y_{t\delta}),$$

em que

$$\begin{aligned}
 \ell_t(w_t(\delta), \phi_t(\delta), y_t) &= \log(\phi_t(\delta)) + \log\left(\frac{\log(0.5)}{\log(1 - w_t(\delta)^{\phi_t(\delta)})}\right) + (\phi_t(\delta) - 1) \log(y_t) \\
 &+ \left(\frac{\log(0.5)}{\log(1 - w_t(\delta)^{\phi_t(\delta)})} - 1\right) \log(1 - y_t^{\phi_t(\delta)}). \tag{B.8}
 \end{aligned}$$

Aqui,  $w_t(\delta)$  e  $\phi_t(\delta)$  dependem de  $g(w_t(\delta)) = \eta_{1t}(\delta)$  e  $h(\phi_t(\delta)) = \eta_{2t}(\delta)$ , que estão dados em (3.42) e (3.43), respectivamente. Assim, temos que

$$\frac{\partial \ell(\theta|\delta)}{\partial \delta_t} = (w_t^*(\delta) + y_t^*(\delta)) \frac{1}{g'(w_t(\delta))} \beta_p S_x^p + [(y_t^*(\delta) + w_t^*(\delta)) \alpha_{1t}(\delta) + \alpha_{2t}(\delta)] \frac{1}{h'(\phi_t(\delta))} \gamma_p' S_x^p. \tag{B.9}$$

Note que a única diferença entre as expressões em (B.7) e (B.9) é que na primeira aparece  $\gamma_p$  ao passo que na segunda aparece  $\gamma_p'$ . Ocorre o mesmo com a expressão de  $\Delta_\beta$  e a matriz equivalente referente ao esquema de perturbação atual.

Quanto às segundas derivadas com relação ao vetor  $\gamma$ , temos que, para  $1, 2, \dots, q$ , com  $j \neq p$ ,

$$\frac{\partial^2 \ell(\theta|\delta)}{\partial \gamma_j \partial \delta_t} = s_{12t} \frac{1}{h'(\phi_t(\delta))^2} \gamma_{p'} S_x^p z_{tj} + s_{13t} \frac{1}{h'(\phi_t(\delta))} \frac{1}{g'(w_t(\delta))} \beta_p S_x^p z_{tj} + s_{8t} \frac{h''(\phi_t(\delta))}{h'(\phi_t(\delta))} \frac{1}{h'(\phi_t(\delta))} x_{tj} \gamma_{p'} S_x^p.$$

Quando  $j = p$ , temos que

$$\begin{aligned} \frac{\partial^2 \ell(\theta|\delta)}{\partial \gamma_j \partial \delta_t} &= s_{12t} \frac{1}{h'(\phi_t(\delta))^2} \gamma_{p'} S_x^p (x_{tj} + \delta_t S_x^p) + s_{13t} \frac{1}{h'(\phi_t(\delta))} \frac{1}{g'(w_t(\delta))} \beta_p S_x^p (x_{tj} + \delta_t S_x^p) \\ &+ s_{8t} \frac{h''(\phi_t(\delta))}{h'(\phi_t(\delta))^2} \frac{1}{h'(\phi_t(\delta))} (x_{tj} + \delta_t S_x^p) \gamma_{p'} S_x^p + s_{8t} \frac{1}{h'(\phi_t(\delta))} S_x^p. \end{aligned}$$

Avaliando as expressões anteriores em  $\delta_0 = (0, 0, \dots, 0)^\top$  e em  $\hat{\theta}$ , obtemos que as matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são da seguinte forma:

$$\begin{aligned} \Delta_\beta &= S_x^p (\gamma_{p'} X^\top S_{10} H T + \beta_p X^\top S_{11} T^2 + \beta_p X^\top S_9 T^2 - P_1 S_9 T), \\ \Delta_\gamma &= S_x^p (\gamma_{p'} Z^\top S_{12} H^2 + \beta_p Z^\top S_{13} T H + \gamma_{p'} X^\top S_8 H_2 - P_2 S_8 H). \end{aligned}$$

**A  $p'$ -ésima coluna da matriz  $Z$  é função da  $p$ -ésima coluna da matriz  $X$**

Aqui estamos considerando a situação em que  $z_{tp'} = \varphi(x_{tp})$ , com  $p = 2, 3, \dots, k$  e  $p' = 2, 3, \dots, q$ , em que  $\varphi$  é uma função diferenciável (se  $\varphi$  é a função identidade, temos que o caso considerado na seção anterior). Aqui  $\eta_{1t}(\delta)$  é como dado em (3.42) e, para  $p' \neq 2$  e  $p' \neq q$ ,

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 z_{t2} + \dots + \gamma_{p'} \varphi(x_{tp} + \delta_t S_x^p) + \dots + \gamma_q z_{tq},$$

em que  $w_t(\delta)$  e  $\phi_t(\delta)$  dependem de  $g(w_t(\delta)) = \eta_{1t}(\delta)$  e  $h(\phi_t(\delta)) = \eta_{2t}(\delta)$ . Para esse tipo de perturbação, tem-se  $\delta_0 = (0, 0, \dots, 0)^\top$  e as matrizes  $\Delta_\beta$  e  $\Delta_\gamma$  são da seguinte forma:

$$\begin{aligned} \Delta_\beta &= S_x^p (\gamma_{p'} X^\top S_{10} H T \dot{G} + \beta_p X^\top S_{11} T^2 + \beta_p X^\top S_9 T^2 - P_1 S_9 T), \\ \Delta_\gamma &= S_x^p (\gamma_{p'} Z^\top S_{12} H^2 \dot{G} + \beta_p Z^\top S_{13} T H + \gamma_{p'} X^\top S_8 H_2 \dot{G} - P_2 S_8 H). \end{aligned}$$