



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

SUELENA DE SOUZA ROCHA

DIAGNÓSTICO EM MODELOS DE REGRESSÃO GAMA UNITÁRIA

Recife  
2020

SUELENA DE SOUZA ROCHA

DIAGNÓSTICO EM MODELOS DE REGRESSÃO GAMA UNITÁRIA

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

**Área de Concentração:** Estatística Matemática.

Orientadora: **Prof<sup>a</sup>. Dr<sup>a</sup>. Patrícia Leone Espinheira Ospina**

Co-orientador: **Prof. Dr. Francisco Cribari Neto**

Recife  
2020

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

R672d Rocha, Suelena de Souza  
Diagnóstico em modelos de regressão gama unitária / Suelena de Souza  
Rocha. – 2020.  
147 f.: il., fig., tab.

Orientadora: Patrícia Leone Espinheira Ospina.  
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN,  
Estatística, Recife, 2020.  
Inclui referências e apêndices.

1. Estatística matemática. 2. Regressão gama unitária. I. Ospina, Patrícia  
Leone Espinheira (orientadora). II. Título.

519.5

CDD (23. ed.)

UFPE- CCEN 2020 - 50

**SUELENA DE SOUZA ROCHA**

**DIAGNÓSTICO EM MODELOS DE REGRESSÃO GAMA UNITÁRIA**

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

Aprovada em: 19 de fevereiro de 2020.

**BANCA EXAMINADORA**

---

Prof.<sup>(a)</sup> Patrícia Leone Espinheira Ospina  
UFPE

---

Prof.<sup>(o)</sup> Abraão David Costa do Nascimento  
UFPE

---

Prof.<sup>(a)</sup> Aline Barbosa Tsuyuguchi  
UFPE

---

Prof.<sup>(o)</sup> Arthur José Lemante  
UFRN

---

Prof.<sup>(o)</sup> Juvêncio Santos Nobre  
UFC

# AGRADECIMENTOS

À Deus e Nossa Senhora, por todas as bênçãos a mim concedidas.

À professora Patrícia Ospina pela excelente orientação nesta tese, pelo carinho e dedicação durante esses quatro anos.

A toda minha família, Suely(mãe), Sebastião(pai) e Suelen(irmã), pelo amor e por todos os seus ensinamentos valiosos.

Aos professores da Universidade Federal de Pernambuco que contribuíram para minha formação, em especial aos professores Alex Dias Ramos, Francisco Cribari Neto, Maria do Carmo Soares de Lima, Klaus Leite Pinto Vasconcellos, Gauss Cordeiro e Patrícia Leone Espinheira Ospina. Obrigada por todo ensinamento compartilhado.

Aos amigos Cássio Anderson, Cristiany Moura, Leon Tarquino, Luana Cecília, Mariana Lima e Victor José, por todos os momentos compartilhados.

A todos os funcionários do Departamento de Estatística da UFPE, em especial à Valéria Bittencourt pela dedicação e carinho.

Aos demais amigos da Pós-graduação em Estatística da UFPE.

Aos participantes da banca examinadora, pelas valiosas sugestões.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

# RESUMO

Nesta tese tratamos da questão de diagnóstico e qualidade de ajuste do modelo de regressão gama unitária. O modelo de regressão gama unitária foi proposto por (MOUSA; EL-SHEIKH; ABDEL-FATTAH, 2016) com o objetivo de modelar variáveis contínuas duplamente limitadas, considerando simultaneamente a modelagem da média e da precisão. Objetivando desenvolver ferramentas de diagnóstico obtemos as expressões dos resíduos ponderado, ponderado padronizado ((ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008)) e combinado ((ESPINHEIRA; SANTOS; CRIBARI-NETO, 2017)) para o modelo de regressão gama unitária. Em seguida consideramos quatro métodos de perturbação, a saber: ponderação de casos, perturbação da variável resposta, perturbação individual de covariadas e perturbação conjunta de covariadas, a partir daí, obtivemos as quantidades associadas ao método de diagnóstico via influência local ((COOK, 1986)). Além disso, consideramos o problema de investigar a qualidade do modelo tanto do ponto de vista de predição quanto do ponto de vista de variabilidade. Neste sentido desenvolvemos para o modelo de regressão gama unitária as expressões das estatísticas PRESS e  $P^2$  ((ESPINHEIRA; OLIVEIRA SILVA, 2019)) e investigamos o comportamento destas medidas conjuntamente com versões do coeficiente de determinação  $R^2$ . Por fim, realizamos simulações de Monte Carlo e aplicações à dados reais embasam as conclusões a serem apresentadas.

**Palavras-chave:** Regressão gama unitária. Resíduos. Distribuição do resíduo. Influência local. Medidas de predição.

# ABSTRACT

We address the diagnostical and model selection issue on unit gamma regression, which models doubly limited continuous variables. We derived the weighted, standardized (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) and combined ((ESPINHEIRA; SANTOS; CRIBARI-NETO, 2017)) residual expressions for the unit gamma regression model. We then consider four perturbation methods, namely: case weighting, response variable perturbation, individual covariate perturbation, and joint covariate perturbation, from which we obtained the quantities associated with the local influence diagnostic method ((COOK, 1986)). In addition, we consider the problem of investigating model quality from both the prediction and variability points of view. In this sense we developed for the unit gamma regression model the expressions of the PRESS and  $P^2$  ((ESPINHEIRA; OLIVEIRA SILVA, 2019)) statistics and investigated the behavior of these measures together with versions of the coefficient of determination  $R^2$ . Finally, we perform Monte Carlo simulations and a real data application support the conclusions to be presented.

**Keywords:** Regression unit gama. Waste. Waste distribution. Local influence. Prediction measures.

# LISTA DE FIGURAS

|            |  |    |
|------------|--|----|
| Figura 1–  | Gráfico $\text{Var}(y) \times \phi_s$ , $s = 1, 2, 3$ . Densidade beta. . . . .  | 18 |
| Figura 2–  | Gráfico $\text{Var}(y) \times \phi_s$ , $s = 1, 2, 3$ . Densidade gama unitária. . . . .   | 18 |
| Figura 3–  | Densidades da gama unitária para diversos valores de $\mu$ e $\phi$ . . . . .  | 20 |
| Figura 4–  | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, \dots, 40$ , $\phi = 1.5$ . . . . .   | 41 |
| Figura 5–  | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, \dots, 40$ , $\phi = 10$ . . . . .  | 42 |
| Figura 6–  | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, \dots, 40$ , $\phi = 30$ . . . . .  | 43 |
| Figura 7–  | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, \dots, 120$ , $\phi = 1.5$ . . . . .  | 44 |
| Figura 8–  | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, \dots, 120$ , $\phi = 10$ . . . . .   | 45 |
| Figura 9–  | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ , $t = 1, \dots, 120$ , $\phi = 30$ . . . . .   | 46 |
| Figura 10– | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$ , $t = 1, \dots, 40$ , $\lambda = 30$ . . . . .   | 54 |
| Figura 11– | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$ , $t = 1, \dots, 40$ , $\lambda = 100$ . . . . .  | 55 |
| Figura 12– | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$ , $t = 1, \dots, 40$ , $\lambda = 530$ . . . . .  | 56 |
| Figura 13– | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$ , $t = 1, \dots, 120$ , $\lambda = 30$ . . . . .  | 57 |
| Figura 14– | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$ , $t = 1, \dots, 120$ , $\lambda = 100$ . . . . . | 58 |

|            |  |     |
|------------|--|-----|
| Figura 15– | Gráficos normais de probabilidade; Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 120$ , $\lambda = 530$ . . . . . | 59  |
| Figura 16– | A curva $\alpha(\delta) = (\delta, LD_\delta)^\top$ , o plano tangente $\mathcal{T}_0$ e o vetor $\pi(\mathcal{T}_0)$ . . . . .  | 61  |
| Figura 17– | Seção normal. . . . .  | 61  |
| Figura 18– | Linhas projetadas. . . . .   | 62  |
| Figura 19– | Gráficos de histograma e boxplot da resposta. Modelos gama unitária com precisão variável. Dados da Linha de Pobreza. . . . .  | 99  |
| Figura 20– | Gráficos de boxplot e dispersão das covariadas. Dados da Linha de Pobreza. . . . .   | 99  |
| Figura 21– | Gráficos dos resíduos, covariada e envelopes. Modelo gama unitária com precisão fixa. Dados da Linha de Pobreza. . . . .   | 101 |
| Figura 22– | Gráficos dos resíduos e envelopes. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza. . . . .  | 104 |
| Figura 23– | Gráficos dos resíduos e envelopes. Modelo beta com dispersão variável. Dados da Linha de Pobreza. . . . .  | 105 |
| Figura 24– | Gráficos de $I_{\max}$ dos índices das observações versus a ponderação de casos, resposta e Persistência no Primário. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza. . . . .   | 106 |
| Figura 25– | Gráficos de $I_{\max}$ dos índices das observações versus as covariadas usadas no submodelo da média. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza. . . . .   | 107 |
| Figura 26– | Gráficos de $C_t$ dos índices das observações versus a ponderação de casos, resposta e Persistência no Primário. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza. . . . .  | 108 |
| Figura 27– | Gráficos de $C_t$ dos índices das observações versus as covariadas usadas no submodelo da média. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza. . . . .  | 109 |
| Figura 28– | Gráficos de $I_{\max}$ dos índices das observações versus a ponderação de casos, resposta e Persistência no Primário. Modelo beta com precisão variável. Dados da Linha de Pobreza. . . . .  | 110 |
| Figura 29– | Gráficos de $I_{\max}$ dos índices das observações versus as covariadas usadas no submodelo da média. Modelo beta com precisão variável. Dados da Linha de Pobreza. . . . .  | 111 |
| Figura 30– | Gráficos de $C_t$ dos índices das observações versus a ponderação de casos, resposta e Persistência no Primário. Modelo beta com precisão variável. Dados da Linha de Pobreza. . . . .   | 112 |
| Figura 31– | Gráficos de $C_t$ dos índices das observações versus as covariadas usadas no submodelo da média. Modelo beta com precisão variável. Dados da Linha de Pobreza. . . . .   | 113 |

|            |   |     |
|------------|---|-----|
| Figura 32– | Gráficos dos resíduos versus. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza. . . . .   | 118 |
| Figura 33– | Gráficos dos resíduos. Função de ligação Cloglog. Modelo beta com dispersão variável. Dados da Linha de Pobreza. . . . .  | 119 |
| Figura 34– | Gráficos de $I_{\max}$ dos índices das observações versus a ponderação de casos, resposta e a covariada usada no submodelo da precisão. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza. . . . . | 120 |
| Figura 35– | Gráficos de $I_{\max}$ dos índices das observações versus as covariadas usadas no submodelo da média. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza . . . . .                                  | 121 |
| Figura 36– | Gráficos de $C_t$ dos índices das observações versus a ponderação de casos, resposta e a covariada usada no submodelo da precisão. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza. . . . .      | 122 |
| Figura 37– | Gráficos de $C_t$ dos índices das observações versus as covariadas usadas no submodelo da média. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza . . . . .                                       | 123 |
| Figura 38– | Gráficos de $I_{\max}$ dos índices das observações versus as covariadas usadas no submodelo da média e precisão. Perturbação conjunta. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza . . . . . | 124 |
| Figura 39– | Gráficos de $I_{\max}$ dos índices das observações versus a ponderação de casos, resposta e a covariada usada no submodelo da precisão. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha de Pobreza. . . . .          | 125 |
| Figura 40– | Gráficos de $I_{\max}$ dos índices das observações versus as covariadas usadas no submodelo da média. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha de Pobreza . . . . .   | 126 |
| Figura 41– | Gráficos de $C_t$ dos índices das observações versus a ponderação de casos, resposta e a covariada usada no submodelo da precisão. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha de Pobreza. . . . .               | 127 |
| Figura 42– | Gráficos de $C_t$ dos índices das observações versus as covariadas usadas no submodelo da média. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha de Pobreza . . . . .  | 128 |

# LISTA DE TABELAS

|           |  |
|-----------|--|
| Tabela 1– | Formas da densidade da gama unitária para diversos valores de $\mu$ e $\phi$ . 19  |
| Tabela 2– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ).<br>Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 1.5$ e $\mu \in (0.0107; 0.3723)$ . . . . . 36 |
| Tabela 3– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ).<br>Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 1.5$ e $\mu \in (0.3709; 0.8507)$ . . . . . 36 |
| Tabela 4– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ).<br>Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 1.5$ e $\mu \in (0.8212; 0.9903)$ . . . . . 37 |
| Tabela 5– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ).<br>Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 10$ e $\mu \in (0.0107; 0.3723)$ . . . . . 37  |
| Tabela 6– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ).<br>Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 10$ e $\mu \in (0.3709; 0.8507)$ . . . . . 38  |
| Tabela 7– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ).<br>Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 10$ e $\mu \in (0.8212; 0.9903)$ . . . . . 38  |
| Tabela 8– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ).<br>Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 30$ e $\mu \in (0.0107; 0.3723)$ . . . . . 39  |

|            |   |    |
|------------|---|----|
| Tabela 9–  | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 30$ e $\mu \in (0.3709; 0.8507)$ . . . . .   | 39 |
| Tabela 10– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ , $t = 1, \dots, 20$ , $\phi = 30$ e $\mu \in (0.8212; 0.9903)$ . . . . .   | 40 |
| Tabela 11– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 30$ e $\mu \in (0.0167; 0.2055)$ . . . . .  | 47 |
| Tabela 12– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 30$ e $\mu \in (0.2048; 0.8712)$ . . . . .  | 48 |
| Tabela 13– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 30$ e $\mu \in (0.8172; 0.9908)$ . . . . .  | 49 |
| Tabela 14– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 100$ e $\mu \in (0.0167; 0.2055)$ . . . . . | 49 |
| Tabela 15– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 100$ e $\mu \in (0.2048; 0.8712)$ . . . . . | 50 |
| Tabela 16– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 100$ e $\mu \in (0.8172; 0.9908)$ . . . . . | 50 |

|            |   |    |
|------------|---|----|
| Tabela 17– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 530$ e $\mu \in (0.0167; 0.2055)$ . . . . . | 52 |
| Tabela 18– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 530$ e $\mu \in (0.2048; 0.8712)$ . . . . . | 52 |
| Tabela 19– | Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo: $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ e $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ , $t = 1, \dots, 20$ , $\lambda = 530$ e $\mu \in (0.8172; 0.9908)$ . . . . . | 53 |
| Tabela 20– | Valores médios das medidas de predição; Funções de ligação; Modelo: $g(\mu_t) = \beta_1 + \beta_2x_{t2}$ , $\phi_t$ constante e $\mu \approx 0.0$ . Modelo corretamente especificado . . . . .  | 79 |
| Tabela 21– | Valores médios das medidas de predição; Funções de ligação; Modelo: $g(\mu_t) = \beta_1 + \beta_2x_{t2}$ , $\phi_t$ constante e $\mu \approx 0.5$ . Modelo corretamente especificado . . . . .  | 80 |
| Tabela 22– | Valores médios das medidas de predição; Funções de ligação; Modelo: $g(\mu_t) = \beta_1 + \beta_2x_{t2}$ , $\phi_t$ constante e $\mu \approx 1$ . Modelo corretamente especificado . . . . .  | 81 |
| Tabela 23– | Valores médios das medidas de predição; Função de ligação Log-Log; Modelo verdadeiro: $g(\mu_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t3} + \beta_5x_{t5}$ , $x_{ti} \sim U(0, 1)$ , $i = 2, 3, 4, 5$ , $t = 1, \dots, n$ , $\phi$ constante e $\mu \approx 0$ . Modelo mal especificado: omissão de covariadas (Cenários 1, 2 e 3). . . . .  | 83 |
| Tabela 24– | Valores médios das medidas de predição; Função de ligação Probit; Modelo verdadeiro: $g(\mu_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t3} + \beta_5x_{t5}$ , $x_{ti} \sim U(0, 1)$ , $i = 2, 3, 4, 5$ , $t = 1, \dots, n$ , $\phi$ constante e $\mu \approx 0.5$ . Modelo mal especificado: omissão de covariadas (Cenários 1, 2 e 3). . . . .   | 84 |
| Tabela 25– | Valores médios das medidas de predição; Função de ligação Clog-Log; Modelo verdadeiro: $g(\mu_t) = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t3} + \beta_5x_{t5}$ , $x_{ti} \sim U(0, 1)$ , $i = 2, 3, 4, 5$ , $t = 1, \dots, n$ , $\phi$ constante e $\mu \approx 1$ . Modelo mal especificado: omissão de covariadas (Cenários 1, 2 e 3). . . . .   | 85 |
| Tabela 26– | Valores médios das estatísticas; Função de ligação Log-Log; Modelos corretamente especificado com $\phi_t$ constante e $\mu \approx 0$ . . . . .  | 88 |
| Tabela 27– | Valores médios das estatísticas; Função de ligação Probit; Modelos corretamente especificado com $\phi_t$ constante e $\mu \approx 0.5$ . . . . .   | 89 |

|            |   |     |
|------------|---|-----|
| Tabela 28– | Valores médios das estatísticas; Função de ligação Clog-Log; Modelos corretamente especificado com $\phi_t$ constante e $\mu \approx 1$ . . . . .               | 90  |
| Tabela 29– | Valores médios das estatísticas; Função de ligação Log-Log; Modelos corretamente especificado com $\phi_t$ variável e $\mu \approx 0$ . . . . .                 | 91  |
| Tabela 30– | Valores médios das estatísticas; Função de ligação Probit; Modelos corretamente especificado com $\phi_t$ variável e $\mu \approx 0.5$ . . . . .                | 92  |
| Tabela 31– | Valores médios das estatísticas; Função de ligação Clog-Log; Modelos corretamente especificado com $\phi_t$ variável e $\mu \approx 1$ . . . . .                | 93  |
| Tabela 32– | Valores médios das estatísticas; Função de ligação Log-Log; Modelos gerados com $\phi_t$ variável e estimado com $\phi_t$ constante onde $\mu \approx 0$ . . .  | 94  |
| Tabela 33– | Valores médios das estatísticas; Função de ligação Probit; Modelos gerados com $\phi_t$ variável e estimado com $\phi_t$ constante onde $\mu \approx 0.5$ . .   | 95  |
| Tabela 34– | Valores médios das estatísticas; Função de ligação Clog-Log; Modelos gerados com $\phi_t$ variável e estimado com $\phi_t$ constante onde $\mu \approx 1$ . . . | 96  |
| Tabela 35– | Estimativas, erros-padrão e $p$ -valores. Modelo gama unitária com precisão fixa. Dados da Linha de Pobreza. . . . .  | 100 |
| Tabela 36– | Estimativas, erros-padrão e $p$ -valores. Modelos beta e gama unitária com precisão variável. Dados da Linha Pobreza. . . . .                                   | 103 |
| Tabela 37– | Estimativas, erro-padrão e $p$ -valor. Observações destacadas no resíduo e influência local. Modelo beta. Dados da Linha de Pobreza. . .                        | 114 |
| Tabela 38– | Estimativas, erros-padrão e $p$ -valores. Função de ligação cloglog. Modelos beta e gama unitária com precisão variável. Dados da Linha Pobreza. . . . .        | 116 |
| Tabela 39– | Observações aberrantes e influentes. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha Pobreza. . . . .                              | 129 |
| Tabela 40– | Medidas de predição. Função de ligação Cloglog. Modelos beta e gama unitária com precisão variável. Dados da Linha Pobreza. . . . .                             | 129 |
| Tabela 41– | Observações influentes. Função de ligação Cloglog. Modelos beta com precisão variável. Dados da Linha Pobreza. . . . .  | 130 |

# SUMÁRIO

|           |   |           |
|-----------|---|-----------|
| <b>1</b>  | <b>INTRODUÇÃO</b>   | <b>16</b> |
| 1.1       | ORGANIZAÇÃO DA TESE . . . . .   | 22        |
| 1.2       | SUPORTE COMPUTACIONAL . . . . .   | 23        |
| <b>2</b>  | <b>MODELO DE REGRESSÃO GAMA UNITÁRIA</b>  | <b>24</b> |
| 2.1       | MODELO DE REGRESSÃO GAMA UNITÁRIA COM PRECISÃO FIXA   | 24        |
| 2.2       | RESÍDUOS . . . . .  | 26        |
| 2.2.1     | <b>Resíduo ponderado</b> . . . . .  | 26        |
| 2.2.2     | <b>Resíduo ponderado padronizado</b> . . . . .  | 27        |
| 2.3       | MODELO DE REGRESSÃO GAMA UNITÁRIA COM PRECISÃO VA-<br>RIÁVEL . . . . .  | 28        |
| 2.4       | RESÍDUOS . . . . .  | 31        |
| 2.4.1     | <b>Resíduo ponderado</b> . . . . .  | 31        |
| 2.4.2     | <b>Resíduo ponderado padronizado</b> . . . . .  | 32        |
| 2.4.3     | <b>Resíduo combinado</b> . . . . .  | 33        |
| <b>3</b>  | <b>ESTUDO DE SIMULAÇÃO DOS RESÍDUOS</b>   | <b>35</b> |
| <b>4</b>  | <b>INFLUÊNCIA LOCAL</b>   | <b>60</b> |
| 4.1       | INFLUÊNCIA LOCAL COM PRECISÃO FIXA . . . . .  | 65        |
| 4.1.1     | <b>Esquema de perturbação</b> . . . . .   | 65        |
| 4.1.1.1   | Ponderação de casos . . . . .   | 65        |
| 4.1.1.2   | Perturbação da variável resposta . . . . .  | 66        |
| 4.1.1.3   | Perturbação individual de covariáveis . . . . .   | 66        |
| 4.2       | INFLUÊNCIA LOCAL COM PRECISÃO VARIÁVEL . . . . .  | 67        |
| 4.2.1     | <b>Esquemas de perturbação</b> . . . . .  | 68        |
| 4.2.1.1   | Ponderação de casos . . . . .   | 68        |
| 4.2.1.2   | Perturbação da variável resposta . . . . .  | 68        |
| 4.2.1.3   | Perturbação individual de covariáveis . . . . .   | 69        |
| 4.2.1.3.1 | <i>A matriz <math>Z</math> totalmente diferente da matriz <math>X</math></i> . . . . .  | 69        |
| 4.2.1.3.2 | <i>A matriz <math>Z</math> é igual à matriz <math>X</math></i> . . . . .  | 70        |
| 4.2.1.3.3 | <i>A <math>k'</math>-ésima coluna da matriz <math>Z</math> é igual à <math>k</math>-ésima coluna da matriz <math>X</math></i> . . . . . | 71        |

|           |   |            |
|-----------|---|------------|
| 4.2.1.3.4 | <i>A <math>k'</math>-ésima coluna da matriz <math>Z</math> é uma função da <math>k</math>-ésima coluna da matriz <math>X</math></i> | 71         |
| 4.2.1.4   | Perturbação conjunta de covariáveis . . . . .   | 72         |
| <b>5</b>  | <b>MEDIDAS DE QUALIDADE DO MODELO ESTIMADO</b>  | <b>73</b>  |
| 5.1       | MEDIDAS DE PREDIÇÃO: DEFINIÇÕES E PROPRIEDADES . . . . .  | 74         |
| 5.2       | SIMULAÇÃO DAS MEDIDAS DE PREDIÇÃO . . . . .   | 77         |
| 5.2.1     | Precisão constante: análise das funções de ligação . . . . .  | 77         |
| 5.2.2     | Precisão constante: omissão de covariadas no submodelo da média   | 82         |
| 5.2.3     | Precisão constante e variável: inclusão de covariadas . . . . .   | 86         |
| 5.2.4     | Precisão variável: Omissão da modelagem da precisão . . . . .   | 87         |
| <b>6</b>  | <b>APLICAÇÕES</b>   | <b>97</b>  |
| 6.1       | APRESENTAÇÃO DOS DADOS . . . . .  | 97         |
| 6.1.1     | Aplicação I: usando a função logito para o submodelo da média   | 98         |
| 6.1.2     | Aplicação II: usando a função cloglog para o submodelo da média   | 115        |
|           | <b>REFERÊNCIAS</b>  | <b>131</b> |
|           | <b>APÊNDICE A - DERIVADAS SEGUNDAS</b>  | <b>133</b> |
|           | <b>APÊNDICE B - INFLUÊNCIA LOCAL</b>  | <b>135</b> |
|           | <b>APÊNDICE C - PROPRIEDADE</b>   | <b>146</b> |

# 1 INTRODUÇÃO

A distribuição gama possui a seguinte função densidade:

$$g(w; \alpha, \phi) = \frac{\alpha^\phi}{\Gamma(\phi)} \exp(-\alpha w) w^{\phi-1}, \quad (1.1)$$

com  $\alpha, \phi > 0$  e  $0 < w < \infty$ . Temos que  $E(W) = \phi/\alpha$  e  $\text{Var}(W) = \phi/\alpha^2$ . Alguns casos particulares da densidade gama; São  $\phi = 1$ , então  $W \sim \text{Exponencial}(\alpha)$ , se  $\phi$  é inteiro, então  $W \sim \text{Erlang}(\phi, \alpha)$  e se  $\phi = n/2$  e  $\alpha = 1/2$ , então  $W \sim \chi^2(n)$ .

Também podemos verificar que várias distribuições são derivadas da distribuição gama com densidade (1.1). Por exemplo, fazendo  $w = \log(1/y)$ , como em (OLSHEN, 1938), temos que  $0 < y < 1$  e obtemos uma nova distribuição, denominada gama unitária.

As variáveis que pertencem ao intervalo unitário  $(0, 1)$ , podem ser interpretadas como por exemplo taxas ou proporções. Diversas distribuições têm sido desenvolvidas para esse tipo de dados, a principal delas é a distribuição beta.

A distribuição beta é bastante flexível para modelar variáveis que assumem valores continuamente no intervalo  $(0, 1)$ , uma vez que sua função densidade pode assumir diversas formas dependendo dos valores dos dois parâmetros que indexam a distribuição. Uma variável aleatória  $Y$  segue distribuição beta com parâmetros  $p$  e  $q$ , denotada por  $Y \sim \mathcal{B}(p, q)$ , se sua função de densidade é da forma

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

com  $0 < y < 1$ , em que  $p > 0$ ,  $q > 0$  e  $\Gamma(\cdot)$  é a função gama. A média e a variância de  $y$  são dadas, respectivamente, por  $E(Y) = p/(p+q)$  e  $\text{Var}(Y) = pq/(p+q)^2(p+q+1)$ . Se  $p = q$ , a função densidade é simétrica. Em particular, quando  $p = q = 1$ , tem-se a distribuição uniforme padrão. Quando  $q < p$ , há assimetria à esquerda e, de forma análoga, quando  $q > p$ , existe assimetria à direita.

Quando  $W \sim \text{Gama}(\alpha, \phi)$  e utilizando as transformações  $w = \log(1/y)$  ou  $w = \log(1/(1-y))$ , em que  $w \in (0, +\infty)$  temos que  $y \in (0, 1)$ . Este fato tem sido pouco explorado na literatura. A distribuição de  $Y$ , a gama unitária, apresenta suporte no intervalo unitário e muitas das características apresentadas pela distribuição beta. (GRIFFITHS; CSIRO, 1981) estudaram as semelhanças das distribuições gama unitária e beta, respectivamente.

A densidade e diversas propriedades da distribuição gama unitária são semelhantes às da distribuição beta, o que pode ser visto em (MOUSA; EL-SHEIKH; ABDEL-FATTAH, 2016), (TADIKAMALLA, 1981) e (RATNAPARKHL; MOSIMANN, 1990), sendo assim, a distribuição gama unitária é uma candidata para modelagem de variáveis contínuas

limitadas.

Ao fazermos a mudança de variável  $w = \log(1/y)$  e usando o teorema da mudança de variável, temos a seguinte expressão para o módulo do determinante do jacobiano  $|-1/y|$ , dessa forma, obtemos a expressão

$$ug(y; \alpha, \phi) = \frac{\alpha^\phi}{\Gamma(\phi)} y^{\alpha-1} [-\log(y)]^{\phi-1}, \quad (1.2)$$

com  $0 < y < 1$ , a qual é denominada como a densidade da variável aleatória com distribuição gama unitária, em que  $\alpha, \phi > 0$  são parâmetros de forma,  $\Gamma(\cdot)$  é a função gama e denotaremos esta distribuição por  $ug(\alpha, \phi)$ ; para mais detalhes, recomendamos (GRASSIA, 1977). Desta forma,  $Y$  segue distribuição gama unitária com parâmetros  $\alpha$  e  $\phi$ , ou seja,  $Y \sim ug(\alpha, \phi)$ .

A média, a variância e o  $n$ -ésimo momento não central de  $Y$ , são respectivamente,

$$E(Y) = \left(\frac{\alpha}{\alpha+1}\right)^\phi, \text{Var}(Y) = \left(\frac{\alpha}{\alpha+2}\right)^\phi - \left(\frac{\alpha}{\alpha+1}\right)^{2\phi} \text{ e } E(Y^n) = \left(\frac{\alpha}{n+\alpha}\right)^\phi. \quad (1.3)$$

É de grande interesse verificar se duas ou mais variáveis estão relacionadas de alguma forma. Para expressar esta relação postulamos um modelo estocástico. Adicionalmente, ao considerarmos a aleatoriedade das variáveis envolvidas nesta relação temos o contexto dos modelos de regressão. Esta relação pode ser analisada como um processo. Neste processo, as quantidades  $x_1, x_2, \dots, x_k$  são chamadas de variáveis de entrada ou regressoras (inputs). Enquanto que  $Y$  é chamada de variável de saída ou resposta (output). A análise de regressão possibilita encontrar uma relação entre as variáveis de entrada e saída, por meio de interações empíricas. A utilização desta abordagem necessita de coleta de dados e do uso de métodos estatísticos.

Se estamos interessados na relação de apenas uma variável de entrada com a variável resposta temos o caso de regressão simples. Mas, se queremos relacionar a variável resposta com mais de uma variável regressora, a regressão múltipla é utilizada. Como exemplo da análise de regressão, podemos citar a regressão beta.

O modelo de regressão beta é uma ferramenta bastante útil para a modelagem de variáveis que assumem valores continuamente no intervalo  $(0, 1)$ . Diversos autores definiram modelos de regressão baseados na distribuição beta, tais como, (PAOLINO, 2001), (KIESCHNICK; MCCULLOUGH, 2003), (FERRARI; CRIBARI-NETO, 2004), (SMITHSON; VERKUILEN, 2006), (SIMAS; BARRETO-SOUZA; ROCHA, 2010) e (ROCHA; SIMAS, 2011).

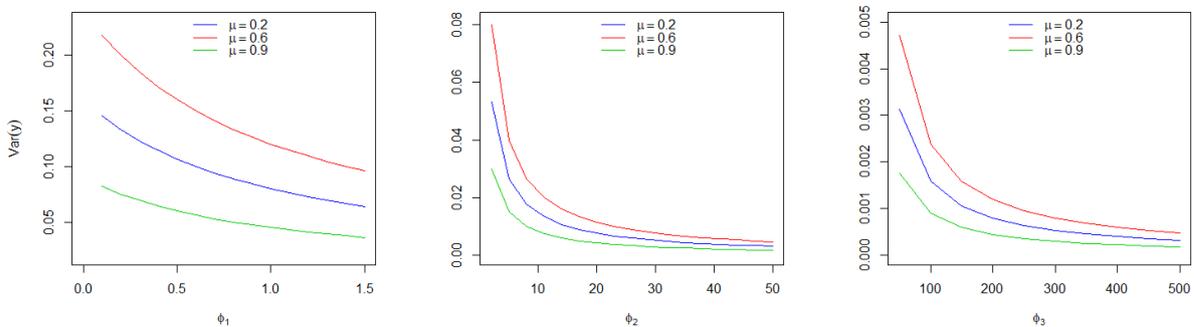
Tais modelos são baseados na suposição de que a variável resposta segue distribuição beta e é relacionada com outras variáveis explicativas através de uma estrutura de regressão. No entanto, o modelo proposto por (FERRARI; CRIBARI-NETO, 2004) propõe uma parametrização permitindo que a densidade da distribuição beta possa ser expressa em função da média da variável resposta, denotada por  $\mu$  e por um parâmetro de precisão denotado por  $\phi$ , onde a variância é dada por  $\text{Var}(Y) = \mu(1-\mu)/(1+\phi)$ . O modelo

beta proposto por (FERRARI; CRIBARI-NETO, 2004) foi generalizado por (SIMAS; BARRETO-SOUZA; ROCHA, 2010), que introduziram estruturas não lineares tanto para o parâmetro  $\mu$  quanto para o parâmetro  $\phi$ .

A distribuição gama unitária, como já mencionado, é usada para situações em que a variável resposta se distribui continuamente no intervalo  $(0, 1)$ . Note que esta variável pode ser explicada por outras variáveis, através de uma estrutura de regressão. Podemos encontrar um modelo de regressão gama unitária no artigo (MOUSA; EL-SHEIKH; ABDEL-FATTAH, 2016). Os autores propuseram uma parametrização diferente para a distribuição gama unitária, parametrização esta que permite a modelagem da média da resposta envolvendo também o parâmetro de precisão.

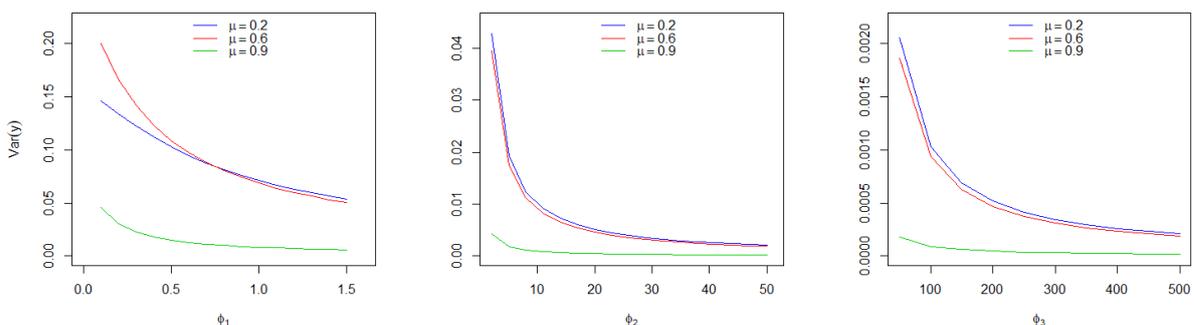
Considere a seguinte parametrização  $\alpha = \mu^{1/\phi}/(1 - \mu^{1/\phi})$ . Assim, temos que  $E(Y) = \mu$  e  $\text{Var}(Y) = \mu \{1/[(2 - \mu^{1/\phi})^\phi] - \mu\}$ . Note que a variância e a média de  $Y$  estão relacionadas. As Figuras 1-2 foram geradas considerando  $\mu = 0.2, 0.6, 0.9$  e os parâmetros  $\phi_1 = (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5)$ ,  $\phi_2 = (2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47, 50)$  e  $\phi_3 = (50, 100, 150, 200, 250, 300, 350, 400, 450, 500)$ . Note que uma vez fixada a média da variável resposta  $\mu$ , quanto maior for o valor que  $\phi$  assume, menor será a variância de  $Y$ . Neste sentido  $\phi$  pode ser interpretado como um parâmetro de precisão tanto do modelo beta quanto para o modelo gama unitária.

Figura 1 – Gráfico  $\text{Var}(y) \times \phi_s$ ,  $s = 1, 2, 3$ . Densidade beta.



Fonte: O autor (2020).

Figura 2 – Gráfico  $\text{Var}(y) \times \phi_s$ ,  $s = 1, 2, 3$ . Densidade gama unitária.



Fonte: O autor (2020).

Na Figura 1, note que a variância diminui ao aumentarmos o valor do parâmetro de

precisão independentemente do valor da média, e esta mesma característica também ocorre na Figura 2, isto decorre dos seguintes fatos:

$$\lim_{\phi \rightarrow \infty} \frac{\mu(\mu - 1)}{1 + \phi} = 0 \quad \text{e} \quad \lim_{\phi \rightarrow \infty} \frac{\mu}{(2 - \mu^{1/\phi})^\phi} - \mu^2 = 0,$$

ou seja, quando o parâmetro de precisão aumenta a variância tende a zero, para ambos os modelos beta e gama unitária.

Observe ainda nas Figuras 1 e 2 quando a média da variável resposta  $\mu$  está próxima de um, o modelo gama unitária possui menor variância que o modelo beta, o que nos mostra forte indícios que o modelo gama unitária pode modelar melhor dados com  $\mu \approx 1$  que o modelo beta.

Sob esta parametrização, a função de densidade de probabilidade de  $Y$  é dada por

$$ug(y; \mu, \phi) = \frac{\alpha^\phi}{\Gamma(\phi)} y^{\alpha-1} [-\log(y)]^{\phi-1}, \quad (1.4)$$

com  $0 < y < 1$ ,  $0 < \mu < 1$  e  $\phi > 0$ , onde  $0 < \alpha = \mu^{1/\phi} / (1 - \mu^{1/\phi})$ .

Detalharemos agora as possíveis formas da densidade da variável  $Y$  com distribuição gama unitária parametrizada para diferentes valores de  $\mu$  e  $\phi$ , observando a Tabela (1) e a Figura (3).

Tabela 1 – Formas da densidade da gama unitária para diversos valores de  $\mu$  e  $\phi$ .

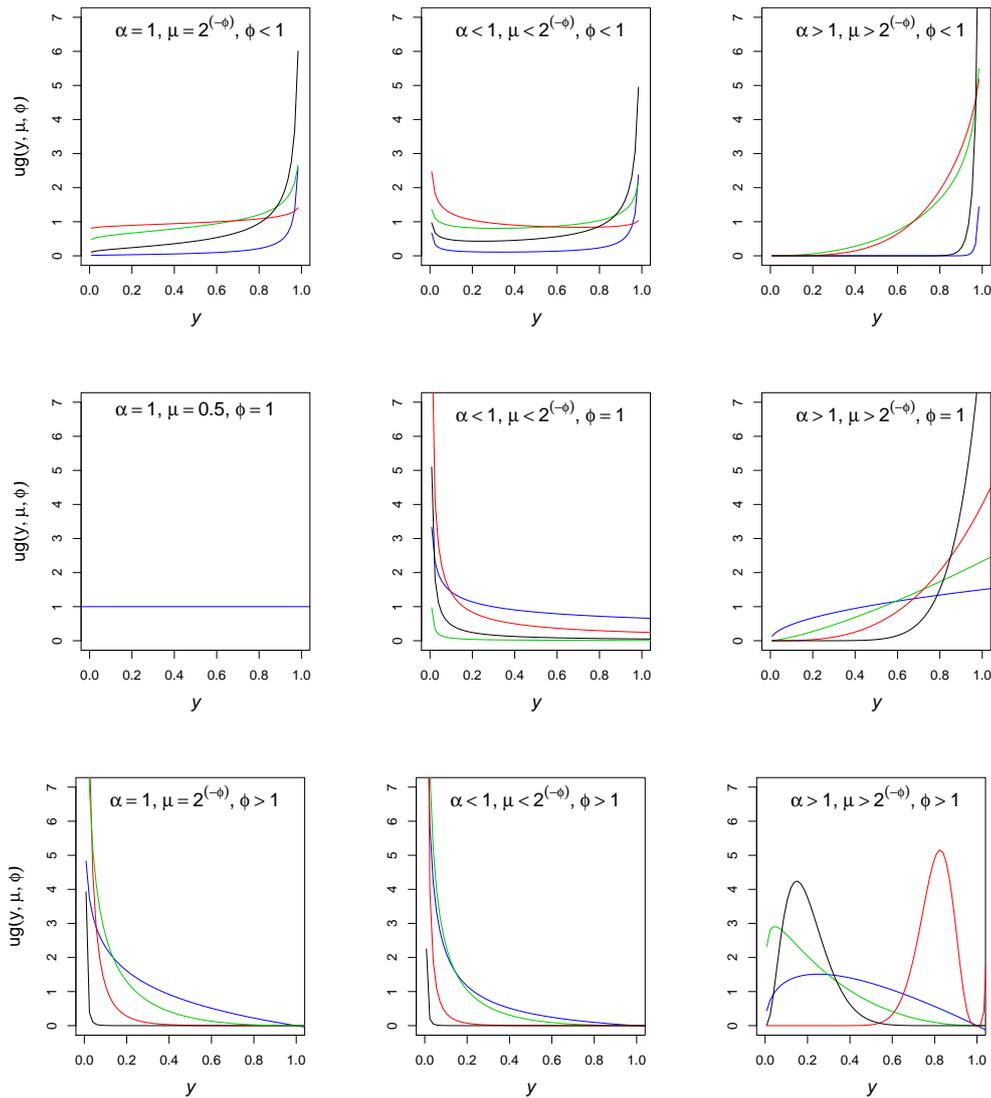
| Valor do parâmetro de precisão | Valor de $\alpha$ e $\mu$        | Gráficos da densidade   |
|--------------------------------|----------------------------------|---|
| $\phi < 1$                     | $\alpha < 1$ e $\mu < 2^{-\phi}$ | o gráfico tem forma de $U$  |
|                                | $\alpha = 1$ e $\mu = 2^{-\phi}$ | o gráfico tem forma de $J$  |
|                                | $\alpha > 1$ e $\mu > 2^{-\phi}$ | o gráfico tem forma de $J$  |
| $\phi = 1$                     | $\alpha < 1$ e $\mu < 0.5$       | o gráfico tem forma de $J$ invertido  |
|                                | $\alpha = 1$ e $\mu = 0.5$       | o gráfico tem a forma da densidade uniforme no $(0, 1)$                     |
|                                | $\alpha > 1$ e $\mu > 0.5$       | o gráfico começa na origem e vai crescendo até o infinito                   |
| $\phi > 1$                     | $\alpha < 1$ e $\mu < 2^{-\phi}$ | o gráfico tem a forma de $J$ invertido                                      |
|                                | $\alpha = 1$ e $\mu = 2^{-\phi}$ | o gráfico tem a forma de $J$ invertido                                      |
|                                | $\alpha > 1$ e $\mu > 2^{-\phi}$ | começa na origem vai crescendo até atingir o valor máximo e depois decresce |

**Fonte:** O autor (2020).

Sabemos que todos os modelos são inevitavelmente simplificações, aproximações da realidade. Assim, uma etapa imprescindível da análise de regressão é a validação do modelo, no sentido de avaliar a qualidade desta aproximação. É sabido que mais de um modelo pode descrever um mesmo fenômeno, haja vista que não há um receita a ser seguida, tendo cada pesquisador a liberdade de modelar o fenômeno seguindo a metodologia que julgar mais adequada. Desse modo, ao se deparar com dois (ou mais modelos) é natural

questionar: “Dentre estes modelos qual deles é o mais adequado?” O conceito de melhor modelo é controverso, mas um bom modelo deve conseguir equilibrar a qualidade do ajuste e a complexidade, sendo esta, em geral, medida pelo número de parâmetros presentes no modelo; quanto mais parâmetros, mais complexo o modelo, sendo então mais difícil interpretar e estimar o modelo. A seleção do “melhor” modelo torna-se então evidente.

Figura 3 – Densidades da gama unitária para diversos valores de  $\mu$  e  $\phi$ .



Fonte: O autor (2020).

Diversas são as metodologias utilizadas para selecionar modelos tais como:  $C_p$  de Mallows, Regressão Stepwise, Critério de Informação Geeralizado (GIC), Critérios de Informação de Akaike (AIC), Critério de Informação Baysiano (BIC), dentre outros. Nesta tese, focaremos nosso estudo nos Critérios AIC e BIC. O modelo que apresentar o menor valor de AIC e BIC será considerado o “melhor” modelo.

Esses critérios essencialmente penalizam a função de verossimilhança utilizando o número de parâmetros do modelo e, eventualmente, o tamanho da amostra. Esta penalização é feita subtraindo-se do valor da verossimilhança uma determinada quantidade, que

depende do quão complexo é o modelo (quanto mais parâmetros, mais complexo).

(AKAIKE, 1974) propôs utilizar a divergência de Kullback-Leibler para seleção de modelos. Ele estabeleceu uma relação entre a máxima verossimilhança e a divergência de Kullback-Leibler desenvolvendo então um critério para estimar a divergência de Kullback-Leibler, que foi posteriormente chamado de Critério de Divergência de Akaike (AIC), dada pela seguinte expressão:  $AIC = -2 \log(\hat{\ell}) + 2k$ , onde  $k$  é o número de parâmetros do modelo e  $\hat{\ell}$  é o valor máximo da log-verossimilhança. O outro critério de seleção de modelos foi proposto por (SCHWARZ, 1978) é conhecido como critério de divergência bayesiano (BIC), dado por:  $BIC = -2 \log(\hat{\ell}) + k \log(n)$ . Este critério é bem similar ao AIC, sendo que neste há uma penalização maior aos modelos que têm um maior número de parâmetros.

Em relação ao contexto da análise de diagnóstico, existem duas direções. Numa direção, o interesse recai em avaliar possíveis afastamentos das suposições admitidas para o modelo, entre as quais está a distribuição de probabilidades dos dados. Em outra direção, o interesse recai em investigar o quanto o modelo é sensível a pequenas perturbações, no sentido de avaliar a estabilidade dos resultados inferências. O modelo é considerado sensível a pequenas perturbações se em sua constituição original implicar em resultados inferenciais significativamente distintos.

A análise de diagnóstico teve início com a análise de resíduos objetivando detectar pontos mal ajustados ou aberrantes e avaliar indícios de afastamentos das suposições sobre o modelo, entre estas, a adequações da distribuição proposta para a variável resposta. A análise de resíduos pode se basear nos resíduos ordinários e suas possíveis padronizações. As técnicas gráficas utilizando resíduos são frequentemente adotadas para a análise de diagnóstico. O uso de envelopes simulados, por exemplo, conforme proposto por (ATKINSON, 1981) inicialmente para o modelo de regressão com erros normais, permite avaliar a qualidade do ajuste do modelo postulado.

Quando se adota um modelo de regressão este é ajustado a um conjunto de dados. A avaliação desse ajuste pode ser realizada por meio da análise dos resíduos. A partir dessa análise identificam-se observações discrepantes, além de avaliar a existência de afastamentos sérios das suposições inerentes ao modelo. Se o modelo for adequado, gráficos de resíduos versus ordens das observações ou valores preditos devem apresentar comportamento aleatório em torno de zero.

Define-se como resíduo uma medida cujo objetivo é identificar discrepância entre o modelo ajustado e os dados. Assim, a maioria das formulações dos resíduos baseia-se na diferença  $y_t - \widehat{E}(y_t)$ . No entanto, respeitando o formato da distribuição de probabilidade da variável resposta, é mais interessante pensar como uma função  $r(y_t, \widehat{E}(y_t))$ , definição geral de resíduos proposta por (COX; SNELL, 1968). Sob esse ponto de vista, (FERRARI; CRIBARI-NETO, 2004). (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) propõem resíduos para o modelo de regressão beta. Dessa forma, adaptamos o resíduo ponderado padronizado proposto por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) ao modelo de regressão gama unitária.

Para o modelo de regressão gama unitária, iremos observar que em muitas situações as distribuições dos resíduos não se aproximam da distribuição normal padrão. Esse fato implica que os usuais limites  $-2$  e  $2$  usados para a detecção de pontos aberrantes nos gráficos de resíduos contra elementos do modelo (índice das observações, valores preditos, valores das covariadas) não são adequados. Neste sentido, vamos utilizar como limites de detecção de pontos aberrantes os quantis empíricos dos resíduos gerados com base em suas distribuições estimadas por processo de reamostragem para a construção das bandas do envelope dos gráficos normais de probabilidade; para mais detalhes, recomendamos (ESPINHEIRA; SANTOS; CRIBARI-NETO, 2017).

Um dos métodos mais modernos de diagnóstico foi proposto por (COOK, 1986). A ideia básica consiste em estudar o comportamento de alguma medida particular de influência segundo pequenas perturbações (influência local) nos dados ou no modelo. Isto é, deseja-se identificar pontos que, sob modificações modestas no modelo, causam variações desproporcionais nos resultados inferenciais. O interesse é estudarmos a influência local das observações no ajuste. A sugestão de Cook é perturbarmos as covariáveis ou a variável resposta e utilizarmos alguma medida adequada para quantificarmos a influência das observações.

A análise de influência local, que visa avaliar conjuntamente o impacto das observações sob pequenas perturbações no modelo ou nos dados, representou um grande avanço para a análise de diagnóstico. Se modificações infinitesimais na formulação inicial do modelo causam efeitos desproporcionais nos resultados inferenciais, existem fortes indícios de falta de qualidade no ajuste ou violação das suposições iniciais, sugerindo a escolha de um modelo mais adequado aos dados.

Nesta tese agregamos diversas contribuições ao modelo de regressão proposto por (MOUSA; EL-SHEIKH; ABDEL-FATTAH, 2016). Uma contribuição foi a demonstração que se  $Y$  pertence a família exponencial de distribuições gama unitária, então a variável  $1 - Y$  também pertence à mesma família (ver Apêndice C). Ainda apresentamos um estudo comparativo das variâncias de variáveis aleatórias gama unitária e beta em função de diferentes valores para os parâmetros  $\mu$  e  $\phi$  que indexam ambas distribuições e que representam respectivamente a média da variável resposta  $Y$  e o parâmetro de precisão do modelo.

Adicionalmente, com relação ao modelo de regressão gama unitária propomos três resíduos baseados nos processos iterativos score de Fisher, com base em estudos de simulações avaliamos as distribuições empíricas destes resíduos. E por fim, propomos duas novas perturbações que podem ser estudadas na análise de influencia local.

## 1.1 ORGANIZAÇÃO DA TESE

A tese está dividida em seis capítulos da seguinte forma; no segundo capítulo, na primeira seção, exibiremos o vetor score e a matriz de informação de Fisher para o modelo de regressão gama unitária com o parâmetro de precisão fixo. Na segunda seção

propomos dois resíduos: ponderado e ponderado padronizado. Na terceira seção estudamos o modelo de regressão gama unitária com o parâmetro de precisão variável. Na quarta seção propomos três resíduos, a saber: ponderado, ponderado padronizado e combinado.

No terceiro capítulo analisamos as distribuições empíricas dos três resíduos propostos no capítulo anterior através das simulações de Monte Carlo. No quarto capítulo realizamos a análise de influência local, baseado em quatro tipos de perturbações: ponderação de casos, perturbação da variável resposta, perturbação individual de covariadas e perturbação conjunta das covariadas. Estudamos o modelo de regressão gama unitária com o parâmetro de precisão fixo e variável, na primeira e segunda seção, respectivamente.

No quinto capítulo, estudamos as medidas de predição no modelo de regressão gama unitária, através de duas seções, na primeira seção, definimos as medidas de predição  $P^2$ ,  $P_c^2$ ,  $R_{FC}^2$ ,  $R_{FC_c}^2$ ,  $R_{LR}^2$ ,  $R_{LR_c}^2$  e apresentamos algumas propriedades das mesmas.

Na segunda seção analisamos as simulações das medidas de predição usando diversas funções de ligações para o submodelo da média, para que desta forma possamos identificar qual função de ligação é mais apropriada quando a média da variável resposta se encontra próxima de zero, próxima de meio e próxima de um. Em seguida, com base nesta informação, estudamos três tipos de cenários, a saber: omissão de covariadas no submodelo da média, comparação dos modelos de regressão gama unitária usando o parâmetro de precisão fixa e variável e o estudo da geração dos dados no modelo de regressão gama unitária ao estimamos com o parâmetro de precisão fixa.

No sexto capítulo estudamos uma aplicação a dados reais (Dados da Linha de pobreza). Na primeira seção estudamos esta aplicação usando a função de ligação logito para o submodelo de média, inicialmente realizamos uma análise descritiva dos dados, em seguida, o estudo das medidas de AIC e BIC, análise dos resíduos e influência local. Na segunda seção utilizamos para a modelagem do submodelo da média a função de ligação Clog-log. E, a partir daí, realizamos todo o estudo da seção anterior incluindo as medidas de predição. No final de cada seção, destacaremos as descobertas realizadas ao analisamos os dados da Linha de Pobreza comparando os modelos de regressão beta e gama unitária.

## 1.2 SUPORTE COMPUTACIONAL

A linguagem de programação utilizada na tese foi, a linguagem de programação matricial Ox. Criada por Jurgen Doornik em 1994. Detalhes sobre esta linguagem de programação podem ser encontrados em Doornik (2001). As análises gráficas foram produzidas utilizando o programa R em sua versão 3.3.2 que se encontra disponível gratuitamente no endereço <http://www.r-project.org>.

## 2 MODELO DE REGRESSÃO GAMA UNITÁRIA

### 2.1 MODELO DE REGRESSÃO GAMA UNITÁRIA COM PRECISÃO FIXA

Sejam  $y_1, y_2, \dots, y_n$  variáveis aleatórias independentes, em que cada  $y_t$ , com  $t = 1, \dots, n$ , segue a densidade dada em (1.4), com média  $\mu_t$  e precisão desconhecida  $\phi$ . Considere o seguinte modelo de regressão:

$$g_1(\mu_t) = \sum_{i=1}^p x_{ti}\beta_i = \eta_t, \quad (2.1)$$

em que  $\beta = (\beta_1, \dots, \beta_p)^\top$  é um vetor de parâmetros da regressão desconhecido ( $\beta \in \mathbb{R}^p$ ),  $x_{t1}, \dots, x_{tp}$  são observações de  $p$  covariadas ( $p < n$ ), assumidas fixas e conhecidas e a função  $g_1(\cdot)$ , que é estritamente monótona e duas vezes diferenciável, tem domínio no intervalo  $(0, 1)$  e imagem que compreende todos os reais. Como exemplo de funções de ligação podemos citar: a ligação logito,  $g_1(\mu) = \log(\mu/(1 - \mu))$ , a probito,  $g_1(\mu) = \Phi^{-1}(\mu)$ , em que  $\Phi(\cdot)$  é a função de distribuição acumulada de uma variável aleatória com distribuição normal padrão, a log - log,  $g_1(\cdot) = -\log(-\log(\mu))$  entre outras. Assim, de acordo com este modelo, temos que

$$\mu_t = g_1^{-1}(\eta_t) \text{ e } \text{Var}(y_t) = g_1^{-1}(\eta_t) \left[ \frac{1}{\{2 - [g_1^{-1}(\eta_t)]^{\frac{1}{\phi}}\}^\phi} - g_1^{-1}(\eta_t) \right]. \quad (2.2)$$

Note que, a variância da resposta depende de  $\mu$ , mesmo quando parâmetro de precisão sendo constante para todas as observações, as variâncias não são constantes.

Considerando uma amostra de  $n$  observações independentes de uma população gama unitária, o logaritmo da função verossimilhança é dado por

$$\ell(\beta, \phi) = \sum_{t=1}^n \ell_t(\mu_t, \phi),$$

em que

$$\ell_t(\mu_t, \phi) = \log(ug(y_t; \mu_t, \phi)) = \phi \log(\alpha_t) - \log(\Gamma(\phi)) + (\alpha_t - 1) \log(y_t) + (\phi - 1) \log(-\log(y_t)),$$

com  $\alpha_t = \mu_t^{1/\phi}/(1 - \mu_t^{1/\phi})$  e  $\mu_t = g_1^{-1}(\eta_t)$ .

Observe inicialmente, que a densidade gama unitária pertence a família exponencial

biparamétrica (bidimensional):

$$ug(y_t; \mu_t, \phi) = \exp\{\tau_1 T_1 + \tau_2 T_2 - A(\tau)\} \left[ \frac{-1}{y_t \log(y_t)} \right],$$

em que  $\tau = (\tau_1, \tau_2)^\top = (\alpha_t, \phi)^\top$  onde  $(T_1, T_2) = (\log(y_t), \log(-\log(y_t)))$  e  $A(\tau) = \log(\Gamma(\phi)) - \phi \log(\alpha_t) = \log(\Gamma(\tau_2)) - \tau_2 \log(\tau_1)$ . Utilizando este fato, podemos afirmar que são válidas as condições de regularidades associadas ao processo de estimação por máxima verossimilhança que serão utilizadas neste trabalho, para mais detalhes, recomendamos (LEHMANN; CASELLA, 1998). Adicionalmente, segue que  $\text{Var}(T_1) = \phi/\alpha_t^2$  e  $\text{E}(T_1) = -\phi/\alpha_t$ . Faremos as estimações dos nossos parâmetros, através do método de máxima verossimilhança.

A função escore para o parâmetro  $\beta$  é um vetor de dimensão  $p$ , que é dado por  $U_\beta(\beta, \phi) = X^\top T s$ , em que  $X$  é a matriz de covariadas de ordem  $n \times p$  cuja  $t$ -ésima linha é  $x_t^\top$ ,  $s = (\mu_1^* + y_1^*, \dots, \mu_n^* + y_n^*)^\top$  e  $T = \text{diag}\{1/g_1'(\mu_1), \dots, 1/g_1'(\mu_n)\}$ , em que

$$\mu_t^* = \frac{\alpha_t}{\mu_t^{1/\phi+1}} \text{ e } y_t^* = \frac{\alpha_t^2 \log(y_t)}{\phi \mu_t^{1/\phi+1}}. \quad (2.3)$$

A função escore para o parâmetro de precisão  $\phi$  é um escalar, que é dado por  $U_\phi(\beta, \phi) = \sum_{t=1}^n u_t$ , onde

$$u_t = \log(-\log(y_t)) - \left[ \frac{1}{\phi} \alpha_t \log(\mu_t) \right] \left[ 1 + \frac{\alpha_t \log(y_t)}{\phi \mu_t^{1/\phi}} \right] - \log\left(\frac{\mu_t^{1/\phi}}{\alpha_t}\right) - \psi(\phi), \quad (2.4)$$

sendo  $\psi(\cdot)$  a função digama, isto é,  $\psi(\phi) = d \log(\Gamma(\phi))/d\phi$ .

A matriz de informação de Fisher é dada por

$$K = K(\beta, \phi) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\phi} \\ K_{\phi\beta} & K_{\phi\phi} \end{pmatrix}, \quad (2.5)$$

em que  $K_{\beta\beta} = X^\top W_{\beta\beta} X$ ,  $K_{\beta\phi} = K_{\phi\beta} = X^\top W_{\beta\phi} \mathbf{1}$  e  $K_{\phi\phi} = \mathbf{1}^\top W_{\phi\phi} \mathbf{1}$ , onde  $\mathbf{1}$  é um vetor com todos os elementos iguais a 1, de ordem  $n \times 1$ ,  $W_{\beta\beta} = \text{diag}\{(a_t^2/\phi) [1/g_1'(\mu_t)]^2\}$ ,  $W_{\beta\phi} = \text{diag}\{(b_t/\phi) (c_t/\phi + 1) [1/g_1'(\mu_t)]\}$  e  $W_{\phi\phi} = \text{diag}[(2c_t/\phi^2) + (c_t^2/\phi^3) + \psi'(\phi)]$  são matrizes diagonais de ordem  $n \times n$ , sendo  $a_t = \alpha_t/(\mu_t^{1/\phi+1})$ ,  $b_t = -a_t$ ,  $c_t = [\alpha_t \log(\mu_t)]/(\mu_t^{1/\phi})$  e  $\psi'(\cdot)$  é a função trigama. Note que podemos escrever a matriz de informação de Fisher de ordem  $(p+1) \times (p+1)$  da seguinte forma:  $K = K(\beta, \phi) = \tilde{X}^\top \tilde{W} \tilde{X}$ , onde

$$\tilde{X}_{2n \times (p+1)} = \begin{pmatrix} X_{n \times p} & \mathbf{0}_{n \times 1} \\ \mathbf{0}_{n \times p} & \mathbf{1}_{n \times 1} \end{pmatrix} \text{ e } \tilde{W}_{2n \times 2n} = \begin{pmatrix} W_{\beta\beta} & W_{\beta\phi} \\ W_{\phi\beta} & W_{\phi\phi} \end{pmatrix}.$$

Sob certas condições de regularidades e em grandes amostras,

$$\begin{pmatrix} \hat{\beta} \\ \hat{\phi} \end{pmatrix} \sim N_{p+1} \left( \begin{pmatrix} \beta \\ \phi \end{pmatrix}, K^{-1} \right),$$

aproximadamente, em que  $\hat{\beta}$  e  $\hat{\phi}$  são os estimadores de máxima verossimilhança de  $\beta$  e  $\phi$  respectivamente. Diferentemente do que ocorre em modelos lineares generalizados, aqui os parâmetros de  $\beta$  e  $\phi$  não são ortogonais, portanto,  $\hat{\beta}$  e  $\hat{\phi}$  não são assintoticamente independentes. Estes estimadores são obtidos como solução do sistema

$$\begin{cases} U_{\beta}(\beta, \phi) = 0 \\ U_{\phi}(\beta, \phi) = 0. \end{cases}$$

Como eles não possuem forma fechada, devemos obtê-los numericamente utilizando algum algoritmo de estimação não linear, como por exemplo, Newton-Raphson, Escore de Fisher, BHHH, BFGS. Para mais detalhes, sobre o vetor Escore e a matriz de informação de Fisher, recomendamos (MOUSA; EL-SHEIKH; ABDEL-FATTAH, 2016).

A proposta de chute inicial para o modelo de regressão gama unitária com o parâmetro de precisão fixo, baseia-se na estimação dos parâmetros de uma regressão de  $X$  contra  $g(y)$  a partir do método dos mínimos quadrados ordinários. Desta forma temos que  $\hat{\beta} = (X^{\top}X)^{-1}X^{\top}g(y)$ . Quanto ao chute inicial de  $\phi$  é importante relacionar a variância da resposta com o parâmetro  $\phi$ . Com base na Figura 2 temos que a medida que  $\phi$  aumenta a variância diminui, onde observamos que  $\phi$  é um parâmetro de precisão.

Considerando  $\sigma^2 = \text{Var}(y)$  temos que  $\sigma^2 \approx \frac{1}{\phi}$  de onde surge a proposta de chute inicial para  $\phi$  dada por

$$\phi^{(0)} = \frac{1}{\sum_{i=1}^n \check{r}_i^2},$$

em que  $\check{r}_i$  é o resíduo de mínimos quadrados ordinário da regressão de  $X$  contra  $g(y)$ , tal que  $\hat{\beta} = (X^{\top}X)^{-1}X^{\top}g(y)$  e  $\check{r} = g(y) - X\hat{\beta}$ .

## 2.2 RESÍDUOS

### 2.2.1 Resíduo ponderado

Utilizaremos a ideia proposta por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) para obtenção de resíduos para o modelo de regressão beta com dispersão constante, que é baseada no processo iterativo Scoring de Fisher para o vetor  $\beta$ .

Para o modelo de regressão gama unitária, com precisão constante e usando o processo iterativo Scoring de Fisher para  $\beta$ , temos que  $\beta^{m+1} = \beta^m + [K_{\beta\beta}^{(m)}]^{-1}U_{\beta}^{(m)}(\beta, \phi)$ , em que  $m$  indexa as iterações até convergência do processo, o qual acontece quando a distância

entre  $\beta^{(m+1)}$  e  $\beta^{(m)}$  torna-se menor que uma constante pequena pré-estabelecida. Obtemos assim a seguinte expressão:

$$\beta^{m+1} = \beta^m + [X^T W_{\beta\beta}^{(m)} X]^{-1} X^T T^{(m)} s^{(m)}.$$

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

$$\beta^{m+1} = [X^T W_{\beta\beta}^{(m)} X]^{-1} X^T W_{\beta\beta} a^{(m)}, \quad (2.6)$$

em que  $a^{(m)} = \eta^{(m)} + W_{\beta\beta}^{(m)-1} T^{(m)} s^{(m)}$  com  $\eta = (\eta_1, \dots, \eta_n)^\top = X\beta$ . Após convergência, temos que  $\hat{\beta} = [X^T \widehat{W}_{\beta\beta}^{(m)} X]^{-1} X^T \widehat{W}_{\beta\beta} a$ , com  $a = \hat{\eta} + \widehat{W}_{\beta\beta}^{-1} \widehat{T} \hat{s}$ , que pode ser visto como um estimador de mínimos quadrados considerando a regressão linear de  $W_{\beta\beta}^{1/2} a$  em  $W_{\beta\beta}^{1/2} X$ .

O resíduo de mínimos quadrados dessa regressão é dado por  $r^\beta = \widehat{W}_{\beta\beta}^{1/2} (a - \hat{\eta}) = \widehat{W}_{\beta\beta}^{-1/2} \widehat{T} \hat{s}$ . Assim, o resíduo da  $t$ -ésima observação é dado por  $r_t^\beta = (1/\hat{w}_{\beta\beta}^{1/2}) [1/g'_1(\hat{\mu}_t)] \hat{s}$ , o qual denominamos de resíduo ponderado. Com base nas definições de  $W_{\beta\beta}$  em (2.5), ele pode ser expresso como

$$r_t^\beta = \frac{1}{\hat{w}_{\beta\beta}^{1/2}} \frac{1}{g'_1(\hat{\mu})} \hat{s} = \frac{\hat{s}}{\left[ \frac{(\hat{\mu}_t^*)^2}{\hat{\phi}} \right]^{1/2}}, \quad (2.7)$$

onde  $s = (\mu_1^* + y_1^*, \dots, \mu_n^* + y_n^*)^\top$  é um vetor com  $\mu_t^*$  e  $y_t^*$  definidos em (2.3).

## 2.2.2 Resíduo ponderado padronizado

Uma possível padronização de (2.7), também proposta por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008), baseia-se na matriz de variâncias e covariâncias de  $a$ . Para isso, a equação (2.6) pode ser escrita como

$$(X^T \widehat{W}_{\beta\beta} X) \hat{\beta} = X^T \widehat{W}_{\beta\beta} a. \quad (2.8)$$

Assuma que  $\text{Cov}(\hat{\beta}) \approx (X^T W_{\beta\beta} X)^{-1}$  e  $\widehat{W}_{\beta\beta} \approx W_{\beta\beta}$ . Note que, ao consideramos isto, implicitamente, sempre estamos considerando também que  $\hat{\phi} \approx \phi$ , já que a expressão  $W_{\beta\beta}$  depende de  $\phi$ . Substituindo essas aproximações e usando a propriedade  $\text{Cov}(Ay) = A \text{Cov}(y) A^\top$  na Expressão (2.8), temos  $\text{Cov}[(X^T W_{\beta\beta} X) \hat{\beta}] \approx \text{Cov}(X^T W_{\beta\beta} a)$ . Então  $\text{Cov}(a) \approx W_{\beta\beta}^{-1}$ .

Com base em (2.8), o resíduo dado em (2.7) pode ser expresso como

$$r^\beta = W_{\beta\beta}^{1/2} a - W_{\beta\beta}^{1/2} X (X^T W_{\beta\beta} X)^{-1} X^T W_{\beta\beta} a = [W_{\beta\beta}^{1/2} - W_{\beta\beta}^{1/2} X (X^T W_{\beta\beta} X)^{-1} X^T W_{\beta\beta}] a.$$

Assim, mais uma vez considerando  $\widehat{W}_{\beta\beta} \approx W_{\beta\beta}$ , segue que

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

Logo,  $\widehat{\text{Cov}}(r^\beta) \approx (I - \widehat{H})$ , onde  $H = W_{\beta\beta}^{1/2}X(X^\top W_{\beta\beta}X)^{-1}X^\top W_{\beta\beta}^{1/2}$  é uma matriz simétrica e idempotente.

Portanto, a padronização do resíduo é dado por

$$r_{pt}^\beta = \frac{r_t^\beta}{\sqrt{\widehat{\text{Cov}}(r_t^\beta)}} = \frac{\hat{s}}{\sqrt{\frac{\alpha_t^2}{\phi(\mu_t^{1/\phi+1})^2}(1 - \hat{h}_{tt})}} = \frac{\hat{s}}{\sqrt{\frac{(\hat{\mu}_t^*)^2}{\hat{\phi}}(1 - \hat{h}_{tt})}}. \quad (2.9)$$

onde  $h_{tt}$  é o  $t$ -ésimo elemento diagonal de  $H$ ,  $s = (\mu_1^* + y_1^*, \dots, \mu_n^* + y_n^*)^\top$  com  $\mu_t^*$  e  $y^*$  definido em (2.3). O resíduo em (2.9) é nomeado como resíduo ponderado padronizado.

## 2.3 MODELO DE REGRESSÃO GAMA UNITÁRIA COM PRECISÃO VARIÁVEL

Consideremos agora que  $y_1, \dots, y_n$  são variáveis aleatórias independentes, onde cada  $y_t$  com  $t = 1, 2, \dots, n$ , segue a densidade (1.4). A média e o parâmetro de precisão satisfazem às seguintes relações funcionais:

$$g(\mu_t) = \sum_{i=1}^p x_{ti}\beta_i = \eta_{1t} \quad \text{e} \quad h(\phi_t) = \sum_{j=1}^q z_{tj}\gamma_j = \eta_{2t}, \quad (2.10)$$

em que  $\beta = (\beta_1, \dots, \beta_p)^\top$  e  $\gamma = (\gamma_1, \dots, \gamma_q)^\top$  são vetores de parâmetros de regressão desconhecidos com  $\beta \in \mathbb{R}^p$  e  $\gamma \in \mathbb{R}^q$ ,  $x_{t1}, \dots, x_{tp}$  e  $z_{t1}, \dots, z_{tq}$  são observações de  $p$  e  $q$  covaridas ( $p < n$  e  $q < n$ ), assumidas fixas e conhecidas, as matrizes de covariadas  $X$  e  $Z$  possuam posto completo e as funções  $g(\cdot)$ ,  $h(\cdot)$  são estritamente monótonas e duas vezes diferenciáveis, com domínios no intervalo  $(0, 1)$  e em  $(0, +\infty)$ , respectivamente, e imagem que compreende todos os reais.

Considerando uma amostra de  $n$  observações independentes, o logaritmo da função verossimilhança é dado por

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

em que

$$\begin{aligned}\ell_t(\mu_t, \phi_t) &= \log(ug(y_t; \mu_t, \phi_t)) \\ &= \phi_t \log(\alpha_t) - \log(\Gamma(\phi_t)) + (\alpha_t - 1) \log(y_t) + (\phi_t - 1) \log(-\log(y_t)),\end{aligned}\quad (2.11)$$

onde  $\alpha_t = \mu_t^{1/\phi_t}/(1 - \mu_t^{1/\phi_t})$ ,  $\mu_t = g^{-1}(\eta_{1t})$  e  $\phi_t = h^{-1}(\eta_{2t})$ .

A função escore para  $\beta$  é um vetor de dimensão  $p$ , que é dado por  $U_\beta^*(\beta, \phi) = X^\top T b$ , em que  $X$  é a matriz de covariadas de ordem  $n \times p$  cuja  $t$ -ésima linha é  $x_t^\top$ ,  $b = (\mu_1^* + y_1^*, \dots, \mu_n^* + y_n^*)^\top$  e  $T = \text{diag}\{1/g'(\mu_1), \dots, 1/g'(\mu_n)\}$ , com

$$\mu_t^* = \frac{\alpha_t}{\mu_t^{1/\phi_t+1}} \quad \text{e} \quad y_t^* = \frac{\alpha_t^2 \log(y_t)}{\phi_t \mu_t^{1/\phi_t+1}}. \quad (2.12)$$

A função escore para  $\gamma$  é um vetor de dimensão  $q$ , que é dado por  $U_\gamma^*(\beta, \gamma) = Z^\top H e$ , em que  $Z$  é a matriz de covariadas de ordem  $n \times q$  cuja  $t$ -ésima linha é  $z_t^\top$ ,  $e = (e_1, \dots, e_n)^\top$  e  $H = \text{diag}\{1/h'(\phi_1), \dots, 1/h'(\phi_n)\}$  com

$$e_t = \log(-\log(y_t)) - \left[ \frac{1}{\phi_t} \alpha_t \log(\mu_t) \right] \left[ 1 + \frac{\alpha_t \log(y_t)}{\phi_t \mu_t^{1/\phi_t}} \right] - \log\left(\frac{\mu_t^{1/\phi_t}}{\alpha_t}\right) - \psi(\phi_t) \quad (2.13)$$

Observe que a função de densidade gama unitária pode ser escrita na forma exponencial biparamétrica como

$$ug(y_t; \mu_t, \phi) = \exp\{\tau_1 T_1 + \tau_2 T_2 - A(\tau)\} \left[ \frac{-1}{y_t \log(y_t)} \right],$$

em que  $f(y_t) = (-1/(y_t \log(y_t)))$ ,  $\tau = (\tau_1, \tau_2) = (\alpha_t, \phi_t)$ ,  $(T_1, T_2) = (\log(y_t), \log(-\log(y_t)))$  e  $A(\tau) = (\log(\Gamma(\phi_t)) - \phi_t \log(\alpha_t)) = (\log(\Gamma(\tau_2)) - \tau_2 \log(\tau_1))$ .

Utilizando este fato, podemos afirmar que são válidas as condições de regularidades associadas a estimação por máxima verossimilhança que serão utilizadas neste trabalho, para mais detalhes, recomendamos (LEHMANN; CASELLA, 1998). Adicionalmente, segue que

$$E(T_1) = \frac{\partial A(\tau)}{\partial \tau_1} = \frac{-\tau_2}{\tau_1} = \frac{-\phi}{\alpha_t}, \quad (2.14)$$

$$\text{Var}(T_1) = \frac{\partial^2 A(\tau)}{\partial \tau_1^2} = -\left(\frac{0 - \tau_2}{\tau_1^2}\right) = \frac{\tau_2}{\tau_1^2} = \frac{\phi}{\alpha_t^2}, \quad (2.15)$$

$$E(T_2) = \frac{\partial A(\tau)}{\partial \tau_2} = \psi(\tau_2) - \log(\tau_1) = \psi(\phi) - \log(\alpha_t), \quad (2.16)$$

$$\text{Var}(T_2) = \frac{\partial^2 A(\tau)}{\partial \tau_2^2} = \psi'(\tau_2) = \psi'(\phi), \quad (2.17)$$

$$\text{Cov}(T_1, T_2) = \frac{\partial^2 A(\tau)}{\partial \tau_1 \partial \tau_2} = \frac{\partial}{\partial \tau_2} \left( \frac{-\tau_2}{\tau_1} \right) = \frac{-1}{\tau_1} = \frac{-1}{\alpha_t}. \quad (2.18)$$

A matriz de informação de Fisher é dada por

$$K^* = K^*(\beta, \gamma) = \begin{pmatrix} K_{\beta\beta}^* & K_{\beta\gamma}^* \\ K_{\gamma\beta}^* & K_{\gamma\gamma}^* \end{pmatrix}, \quad (2.19)$$

onde  $K_{\beta\beta}^* = X^\top V_{\beta\beta} X$ ,  $K_{\beta\gamma}^* = K_{\gamma\beta}^* = X^\top V_{\beta\phi} Z$  e  $K_{\gamma\gamma}^* = Z^\top V_{\phi\phi} Z$ , em que  $V_{\beta\beta} = \text{diag} \left\{ (\tilde{a}_t^2 / \phi_t) [1/g'(\mu_t)]^2 \right\}$ ,  $V_{\beta\gamma} = \text{diag} \left\{ [(\tilde{b}_t / \phi_t) (\tilde{c}_t / \phi_t + 1)] (1/g'(\mu_t)) (1/h'(\phi_t)) \right\}$  e  $V_{\gamma\gamma} = \text{diag} \left\{ [(2\tilde{c}_t / \phi_t^2) + (\tilde{c}_t^2 / \phi_t^3) + \psi'(\phi_t)] (1/h'(\phi_t))^2 \right\}$  são matrizes diagonais de ordem  $n \times n$ , sendo

$$\tilde{a}_t = \alpha_t / \mu_t^{1/\phi_t + 1}, \tilde{b}_t = -\tilde{a}_t \text{ e } \tilde{c}_t = [\alpha_t \log(\mu_t)] / \mu_t^{1/\phi_t}. \quad (2.20)$$

Note que podemos escrever a matriz de informação de Fisher de ordem  $(p+q) \times (p+q)$ , da seguinte forma:  $K^* = K^*(\beta, \gamma) = \tilde{R}^\top \tilde{V} \tilde{R}$ , onde

$$\tilde{R}_{2n \times (p+q)} = \begin{pmatrix} X_{n \times p} & 0_{n \times q} \\ 0_{n \times p} & Z_{n \times q} \end{pmatrix} \text{ e } \tilde{V}_{2n \times 2n} = \begin{pmatrix} V_{\beta\beta} & V_{\beta\gamma} \\ V_{\gamma\beta} & V_{\gamma\gamma} \end{pmatrix}.$$

Sob certas condições de regularidades e em grandes amostras,

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \sim N_{p+q} \left( \begin{pmatrix} \beta \\ \gamma \end{pmatrix}, K^{*-1} \right),$$

aproximadamente, em que  $\hat{\beta}$  e  $\hat{\gamma}$  são os estimadores de máxima verossimilhança de  $\beta$  e  $\gamma$ , respectivamente. Diferentemente do que ocorre em modelos lineares generalizados, aqui os parâmetros de  $\beta$  e  $\gamma$  não são ortogonais, portanto,  $\hat{\beta}$  e  $\hat{\gamma}$  não são assintoticamente independentes. Estes estimadores são obtidos através do sistema

$$\begin{cases} U_\beta^*(\beta, \gamma) = 0 \\ U_\gamma^*(\beta, \gamma) = 0. \end{cases}$$

Eles não podem ser expressos de forma fechada. Eles devem ser obtidos numericamente utilizando algum algoritmo de estimação não linear, como por exemplo, Newton-Raphson, Escore de Fisher, BHHH, BFGS. Para mais detalhes, sobre o vetor escore e a matriz de informação de Fisher, recomendamos (MOUSA; EL-SHEIKH; ABDEL-FATTAH, 2016).

A proposta de chute inicial para o modelo de regressão gama unitária com o parâmetro de precisão variável, baseia-se na estimação dos parâmetros de uma regressão de  $X$  contra  $g(y)$  e de  $Z$  contra  $h(\phi)$  a partir do método dos mínimos quadrados ordinários. Desta forma temos que  $\hat{\beta}^{(0)} = (X^\top X)^{-1} X^\top g(y)$  e  $\gamma = (Z^\top Z)^{-1} Z^\top h(\phi^{(0)})$ , respectivamente. Quanto ao chute inicial de  $\phi_t$  é importante encontrarmos uma relação entre a variância da resposta com o parâmetro  $\phi$ . Com base na Figura 2 temos que a medida que  $\phi$  aumenta a variância diminui, observando que  $\phi$  é um parâmetro de precisão.

Considerando  $\sigma^2 = \text{Var}(y)$  temos que  $\sigma^2 \approx 1/\phi$  de onde surge a proposta de chute inicial para  $\phi_t$  dada por

$$\phi_t^{(0)} = \frac{1}{\check{r}_t^2},$$

em que  $\check{r}_t$  é o resíduo de mínimos quadrados ordinário da regressão de  $X$  contra  $g(y)$ , tal que  $\hat{\beta}^{(0)} = (X^\top X)^{-1} X^\top g(y)$  e  $\check{r} = g(y) - X\hat{\beta}^{(0)}$ .

## 2.4 RESÍDUOS

### 2.4.1 Resíduo ponderado

Para calcularmos os resíduos do modelo gama unitária, com parâmetro de precisão variável, iremos utilizar a ideia proposta por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008) para modelos de regressão beta com dispersão constante, a qual foi adaptada para os casos em que a dispersão pode ser variável ao longo das observações.

Utilizando o modelo de regressão gama unitária, com precisão variável e o processo iterativo Scoring de Fisher para estimar  $\beta$ , temos  $\beta^{m+1} = \beta^m + [K_{\beta\beta}^{*(m)}]^{-1} U_{\beta}^{*(m)}(\beta, \phi)$ , em que  $m$  indexa as iterações até convergência, o qual acontece quando a distância entre  $\beta^{(m+1)}$  e  $\beta^{(m)}$  torna-se menor que uma constante pequena pré-estabelecida. Obtemos a seguinte expressão para o método iterativo:

$$\beta^{m+1} = \beta^m + \left[ X^\top V_{\beta\beta}^{(m)} X \right]^{-1} X^\top T^{(m)} b^{(m)}.$$

O processo acima pode ser representado como um processo iterativo de mínimos quadrados reponderados, ou seja,  $\beta^{m+1} = [X^\top V_{\beta\beta}^{(m)} X]^{-1} X^\top V_{\beta\beta} t_1^{(m)}$ , em que  $t_1^{(m)} = \eta_1^{(m)} + V_{\beta\beta}^{(m)-1} T^{(m)} b^{(m)}$  com  $\eta_1 = (\eta_{11}, \dots, \eta_{1n})^\top = X\beta$ , o qual. Após a convergência, temos que

$$\hat{\beta} = [X^\top \hat{V}_{\beta\beta}^{(m)} X]^{-1} X^\top \hat{V}_{\beta\beta} t_1, \quad (2.21)$$

com  $t_1 = \hat{\eta}_1 + \hat{V}_{\beta\beta}^{-1} \hat{T} \hat{b}$ , que pode ser visto como um estimador de mínimos quadrados considerando a regressão linear de  $V_{\beta\beta}^{1/2} t_1$  em  $V_{\beta\beta}^{1/2} X$ .

O resíduo de mínimos quadrados dessa regressão é dado por

$$r^\beta = \hat{V}_{\beta\beta}^{1/2} (t_1 - \hat{\eta}_1) = \hat{V}_{\beta\beta}^{-1/2} \hat{T} \hat{b}.$$

Assim, o resíduo da  $t$ -ésima observação é

$$r_t^\beta = \frac{1}{\hat{v}_{\beta\beta}^{1/2}} \frac{1}{g'(\hat{\mu}_t)} \hat{b}_t,$$

o qual denominamos de resíduo ponderado. Com base nas definições de  $V_{\beta\beta}$  em (2.19), ele

pode ser expresso como

$$r_t^\beta = \frac{1}{\hat{v}_{\beta\beta}^{1/2}} \frac{1}{g'(\hat{\mu}_t)} \hat{b}_t = \frac{\hat{b}_t}{\left[ \frac{(\hat{\mu}_t^*)^2}{\hat{\phi}_t} \right]^{1/2}}, \quad (2.22)$$

onde  $b = (\mu_1^* + y_1^*, \dots, \mu_n^* + y_n^*)^\top$  é um vetor com  $\mu_t^*$  e  $y_t^*$  definido em (2.12).

## 2.4.2 Resíduo ponderado padronizado

Uma possível padronização de (2.22), também considerada por (ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008), baseia-se na matriz de variâncias e covariâncias de  $t_1$ . Para isso, a Equação (2.21) pode ser escrita como

$$(X^\top \hat{V}_{\beta\beta} X) \hat{\beta} = X^\top \hat{V}_{\beta\beta} t_1.$$

Dado que  $\text{Cov}(\hat{\beta}) \approx (X^\top V_{\beta\beta} X)^{-1}$  e  $\hat{V}_{\beta\beta} \approx V_{\beta\beta}$ . Note que ao considerarmos isto, implicitamente, estamos considerando também que  $\hat{\phi}_t \approx \phi_t$ , já que a expressão de  $V_{\beta\beta}$  depende de  $\phi_t$ . Substituindo essas igualdades e usando a propriedade  $\text{Cov}(Ay) = A\text{Cov}(y)A^\top$  na expressão acima, temos  $\text{Cov}[(X^\top \hat{V}_{\beta\beta} X) \hat{\beta}] \approx \text{Cov}(X^\top \hat{V}_{\beta\beta} t_1)$ . Assim,  $\text{Cov}(t_1) \approx V_{\beta\beta}^{-1}$ .

Logo, o resíduo anterior pode ser expresso como

$$r^\beta = V_{\beta\beta}^{1/2} t_1 - V_{\beta\beta}^{1/2} X (X^\top V_{\beta\beta} X)^{-1} X^\top V_{\beta\beta} t_1 = \left[ V_{\beta\beta}^{1/2} - V_{\beta\beta}^{1/2} X (X^\top V_{\beta\beta} X)^{-1} X^\top V_{\beta\beta} \right] t_1.$$

Assim, mais uma vez considerando  $\widehat{W}_{\beta\beta} \approx W_{\beta\beta}$ , segue que

$$\begin{aligned} \text{Cov}(r^\beta) &\approx \text{Cov} \left\{ \left[ V_{\beta\beta}^{1/2} - V_{\beta\beta}^{1/2} X (X^\top V_{\beta\beta} X)^{-1} X^\top V_{\beta\beta} \right] t_1 \right\} \\ &\approx \left[ I - V_{\beta\beta}^{1/2} X (X^\top V_{\beta\beta} X)^{-1} X^\top V_{\beta\beta}^{1/2} \right] \\ &\approx (I - H^*). \end{aligned}$$

Logo,  $\widehat{\text{Cov}}(r^\beta) \approx (I - \widehat{H}^*)$ , onde  $H^* = V_{\beta\beta}^{1/2} X (X^\top V_{\beta\beta} X)^{-1} X^\top V_{\beta\beta}^{1/2}$  é uma matriz de projeção de  $V_{\beta\beta}^{1/2} t_1$  contra  $V_{\beta\beta}^{1/2} X$ , e além disso, é uma matriz simétrica e idempotente.

Portanto, o resíduo padronizado é dado por

$$r_{pt}^\beta = \frac{r_t^\beta}{\sqrt{\widehat{\text{Cov}}(r_t^\beta)}} = \frac{\hat{b}_t}{\sqrt{\frac{\hat{\alpha}_t^2}{\hat{\phi}_t (\hat{\mu}_t^{1/\hat{\phi}_t+1})^2} (1 - \hat{h}_{tt})}} = \frac{\hat{b}_t}{\sqrt{\frac{(\hat{\mu}_t^*)^2}{\hat{\phi}_t} (1 - \hat{h}_{tt}^*)}}, \quad (2.23)$$

onde  $b = (\mu_1^* + y_1^*, \dots, \mu_n^* + y_n^*)^\top$  é um vetor com  $\mu_t^*$  e  $y_t^*$  definidos em (2.12) e  $h_{tt}^*$  é o  $t$ -ésimo elemento diagonal de  $H^*$ . O resíduo acima é nomeado como resíduo ponderado padronizado.

### 2.4.3 Resíduo combinado

A construção do resíduo combinado é em parte similar à construção do resíduo ponderado, que é baseado no processo iterativo Scoring de Fisher para estimar  $\beta$ . No entanto, para a construção do resíduo combinado, considera-se adicionalmente o processo iterativo para  $\gamma$  que é dado por

$$\gamma^{m+1} = \gamma^m + [K_{\gamma\gamma}^{*(m)}]^{-1} U_{\gamma}^{*(m)}(\beta, \gamma),$$

em que  $m$  indexa as iterações até convergência. Portanto, obtemos a seguinte expressão para o método iterativo:  $\gamma^{m+1} = \gamma^m + [Z^\top V_{\gamma\gamma}^{(m)} Z]^{-1} Z^\top H^{(m)} e^{(m)}$ , em que  $e = (e_1, \dots, e_n)^\top$  com  $e_t$  definido em (2.13).

O processo acima pode ser representado como um processo iterativo de mínimos quadrados reponderados, ou seja,  $\gamma^{m+1} = [Z^\top V_{\gamma\gamma}^{(m)} Z]^{-1} Z^\top V_{\gamma\gamma} t_2^{(m)}$ , em que  $t_2^{(m)} = \eta_2^{(m)} + V_{\gamma\gamma}^{(m)-1} H^{(m)} e^{(m)}$  com  $\eta_2 = (\eta_{21}, \dots, \eta_{2n})^\top = Z\gamma$ . Após convergência, temos que  $\hat{\gamma} = [Z^\top \hat{V}_{\gamma\gamma}^{(m)} Z]^{-1} Z^\top \hat{V}_{\gamma\gamma} t_2$ , com  $t_2 = \hat{\eta}_2 + \hat{V}_{\gamma\gamma}^{-1} \hat{H} \hat{e}$ , que pode ser visto como um estimador de mínimos quadrados considerando a regressão linear de  $V_{\gamma\gamma}^{1/2} t_2$  em  $V_{\gamma\gamma}^{1/2} Z$ .

O resíduo de mínimos quadrados dessa regressão é dado por

$$r^\beta = \hat{V}_{\gamma\gamma}^{1/2} (t_2 - \hat{\eta}_2) = \hat{V}_{\gamma\gamma}^{-1/2} \hat{H} \hat{e}. \quad (2.24)$$

Assim, o resíduo da  $t$ -ésima observação é dado por

$$r_t^\gamma = \frac{1}{\hat{v}_{\gamma\gamma}^{1/2}} \frac{1}{h'(\hat{\phi}_t)} \hat{e}_t.$$

Logo o resíduo ponderado é dado por

$$r_t^\gamma = \frac{h'(\hat{\phi}_t)}{\left\{ \left[ \left( \frac{2\hat{c}_t}{\hat{\phi}_t^2} \right) + \left( \frac{\hat{c}_t^2}{\hat{\phi}_t^3} \right) + \psi'(\hat{\phi}_t) \right] \right\}^{1/2}} \frac{1}{h'(\hat{\phi}_t)} \hat{e}_t = \frac{\hat{e}_t}{\left\{ \left[ \left( \frac{2\hat{c}_t}{\hat{\phi}_t^2} \right) + \left( \frac{\hat{c}_t^2}{\hat{\phi}_t^3} \right) + \psi'(\hat{\phi}_t) \right] \right\}^{1/2}}, \quad (2.25)$$

onde  $\tilde{c}_t$  e  $e_t$  estão definido em (2.20) e (2.13), respectivamente.

Vale a pena lembrar que o resíduo para  $t$ -ésima observação pode ser definido como

$$R_t = r_t(y_t, \hat{\mu}_t, \hat{\phi}_t).$$

Ele mede a discrepância entre o valor observado  $y_t$  e o valor ajustado  $\hat{\mu}_t$  da  $t$ -ésima observação ((COX; SNELL, 1968)).

O resíduo combinado é construído baseando-se nos processos iterativos Scoring de Fisher de  $\beta$  e  $\gamma$ , e assim contém a informação sobre todos os parâmetros do modelo. Assim baseando-se na definição de resíduo acima, o resíduo combinado, que tem componentes dos resíduos (2.22) e (2.25) é dado por

$$r_t^{\beta\gamma} = \hat{b}_t + \hat{e}_t.$$

Calcularemos agora, a variância de  $b_t + e_t$ . Temos que

$$\begin{aligned} \text{Var}(b_t + e_t) &= \text{Var} \left\{ \frac{\alpha_t^2 \log(y_t)}{\phi_t \mu_t^{\frac{1}{\phi_t} + 1}} + \log[-\log(y_t)] - \frac{\alpha_t^2 \log(\mu_t) \log(y_t)}{\phi_t^2 \mu_t^{\frac{1}{\phi_t}}} \right\} \\ &= \frac{\alpha_t^2}{\phi_t \left( \mu_t^{1/\phi_t + 1} \right)^2} + \psi'(\phi_t) + \frac{\alpha_t^2 [\log(\mu_t)]^2}{\phi_t^3 (\mu_t^{1/\phi_t})^2} - 2 \left[ \frac{\alpha_t}{\phi_t (\mu_t^{1/\phi_t + 1})} + \frac{\alpha_t^2 \log(\mu_t)}{\phi_t^2 (\mu_t^{2/\phi_t + 1})} - \frac{\alpha_t \log(\mu_t)}{\phi_t (\mu_t^{1/\phi_t})} \right] \\ &= d_t, \end{aligned}$$

$$\text{em que } d_t = \psi'(\phi_t) + \frac{\alpha_t}{\phi_t} \left\{ \frac{\alpha_t}{(\mu_t^{1/\phi_t + 1})^2} + \frac{\alpha_t [\log(\mu_t)]^2}{\phi_t^2 (\mu_t^{1/\phi_t})^2} - 2 \left[ \frac{1}{(\mu_t^{1/\phi_t + 1})} + \frac{\alpha_t \log(\mu_t)}{\phi_t (\mu_t^{2/\phi_t + 1})} - \frac{\log(\mu_t)}{(\mu_t^{1/\phi_t})} \right] \right\}.$$

Assim, o resíduo combinado padronizado é definido como

$$r_t^{\beta\gamma} = \frac{\hat{b}_t + \hat{e}_t}{\sqrt{\hat{d}_t}}, \quad t = 1, \dots, n, \quad (2.26)$$

em que  $b = (\mu_1^* + y_1^*, \dots, \mu_n^* + y_n^*)^\top$  é um vetor com  $\mu_t^*$  e  $y_t^*$  definidos em (2.12),  $e_t$  está definido em (2.13) e  $\hat{d}_t$  é obtido avaliando  $d_t$  em  $\hat{\mu}_t$  e  $\hat{\phi}_t$ . Para o modelo de precisão constante, considera-se  $\phi_1 = \phi_2 = \dots = \phi_n = \phi$ .

Neste capítulo estudamos a densidade gama unitária, verificamos que esta pertence a família exponencial, estudamos também o modelo de regressão gama unitária com o parâmetro de precisão fixo e variável, exibimos o vetor escore, matriz de informação de Fisher e o estudo do chute inicial para o parâmetro de precisão e por fim propomos três resíduos, a saber: ponderado, ponderado padronizado e combinado.

### 3 ESTUDO DE SIMULAÇÃO DOS RESÍDUOS

Para avaliar a distribuição empírica dos resíduos ponderado, ponderado padronizado e combinado, definidos na Seção 2.3, realizamos simulações de Monte Carlo com 5.000 réplicas sob diferente cenários. Os resultados das simulações foram obtidos através do programa de linguagem matricial Ox, para mais detalhes ver <http://www.doornik.com>. Inicialmente, consideramos o modelo de regressão gama unitária com precisão constante dado por

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

Os valores das covariadas foram obtidos de forma independente através da distribuição uniforme padrão,  $x_i \sim U(0, 1)$ ,  $i = 2, \dots, 5$ , e mantidos fixos para cada réplica. Tomamos as covariadas geradas da distribuição uniforme, pois queremos analisar as distribuições dos resíduos considerando os cenários mais simples possíveis, porém se gerarmos as covariadas de outras distribuições, estas podem gerar outliers, dificultando assim as nossas interpretações. Esse procedimento é tipicamente realizado para avaliação de distribuições de resíduos ((ESPINHEIRA; FERRARI; CRIBARI-NETO, 2008)).

Iremos exibir os resultados em relação a média, desvio-padrão, assimetria e curtose das 5.000 réplicas relativo aos resíduos ponderado, ponderado padronizado e combinado, definidos nas Seções 2.2.1, 2.2.2 e 2.4.3, respectivamente, considerando  $n = 20$  observações. Assim, é possível avaliar o comportamento das distribuições dos resíduos quando a amostra é consideravelmente pequena e adicionalmente facilita a apresentação dos resultados tabelados.

Como não existe na literatura nenhum estudo relativo ao parâmetro de precisão da distribuição gama unitária, com base no estudo que fizemos consideramos inicialmente  $\phi = 1.5$  como um valor pequeno, pois valores menores que estes acarretavam problemas de convergência e estimação gerando assim impecílios para o estudo com  $\phi < 1.5$ . Em seguida, consideramos  $\phi = 10$  como um valor intermediário e por fim consideramos  $\phi = 30$  como um valor grande para o parâmetro de precisão fixo. É importante ressaltar que para valores de  $\phi > 30$  não há alterações importantes nas distribuições dos resíduos.

Nas Tabelas 2,3 e 4 a média da variável resposta pertence aos intervalos (0.0107; 0.3723), (0.3709; 0.8507) e (0.8212; 0.9903), respectivamente, e  $\phi = 1.5$ . Os valores verdadeiros dos parâmetros são  $\beta_1 = -2.5, \beta_2 = -1.8, \beta_3 = 2.2, \beta_4 = 3.0$  e  $\beta_5 = -2.5$  quando  $\mu \in (0.0107; 0.3723)$ ,  $\beta_1 = -1.9, \beta_2 = 1.2, \beta_3 = 1.3, \beta_4 = 2.0$  e  $\beta_5 = 1.3$  quando  $\mu \in (0.3709; 0.8507)$  e  $\beta_1 = 2.0, \beta_2 = 1.4, \beta_3 = 1.8, \beta_4 = -1.3$  e  $\beta_5 = 1.4$  quando  $\mu \in (0.8212; 0.9903)$ .

Observa-se nas Tabelas 2,3 e 4 que para os três resíduos, a média está próxima de zero e o desvio-padrão está próximo de um. Além disso, observa-se que há assimetria negativa (próxima de  $-1$ ), para os três cenários da média da variável resposta e a curtose também está distante de três, valor de referência da distribuição normal padrão.

Tabela 2 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 1.5$  e  $\mu \in (0.0107; 0.3723)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | 0.001     | 0.001       | 0.000            | 0.962          | 1.034       | 0.961            | -1.075     | -1.074      | -1.113           | 4.539     | 4.536       | 4.636            |
| 2  | 0.015     | 0.018       | 0.016            | 0.858          | 1.012       | 0.855            | -0.804     | -0.794      | -0.883           | 3.757     | 3.682       | 3.904            |
| 3  | 0.014     | 0.015       | 0.013            | 0.910          | 1.021       | 0.909            | -0.901     | -0.907      | -0.927           | 3.878     | 3.885       | 3.940            |
| 4  | -0.005    | -0.006      | -0.006           | 0.913          | 1.029       | 0.916            | -0.877     | -0.871      | -1.000           | 3.684     | 3.671       | 3.945            |
| 5  | 0.002     | 0.003       | 0.002            | 0.908          | 1.014       | 0.907            | -0.898     | -0.899      | -0.964           | 3.812     | 3.807       | 3.946            |
| 6  | 0.025     | 0.038       | 0.038            | 0.677          | 1.017       | 0.675            | -0.436     | -0.388      | -0.603           | 3.060     | 3.065       | 3.233            |
| 7  | 0.008     | 0.010       | 0.008            | 0.894          | 1.011       | 0.893            | -0.878     | -0.877      | -0.893           | 3.859     | 3.834       | 3.889            |
| 8  | 0.016     | 0.018       | 0.016            | 0.890          | 0.983       | 0.889            | -0.895     | -0.897      | -0.909           | 3.922     | 3.913       | 3.957            |
| 9  | -0.021    | -0.022      | -0.027           | 1.014          | 1.064       | 1.025            | -1.259     | -1.259      | -1.434           | 5.213     | 5.211       | 5.803            |
| 10 | -0.032    | -0.035      | -0.035           | 0.942          | 1.053       | 0.951            | -0.887     | -0.887      | -1.063           | 3.682     | 3.685       | 4.054            |
| 11 | 0.014     | 0.018       | 0.016            | 0.807          | 1.004       | 0.804            | -0.618     | -0.615      | -0.698           | 3.220     | 3.212       | 3.329            |
| 12 | 0.010     | 0.017       | 0.026            | 0.616          | 0.999       | 0.613            | -0.288     | -0.216      | -0.448           | 2.983     | 2.924       | 3.087            |
| 13 | 0.007     | 0.007       | 0.007            | 0.876          | 0.997       | 0.875            | -0.883     | -0.887      | -0.922           | 3.924     | 3.906       | 4.003            |
| 14 | -0.003    | -0.003      | -0.003           | 0.937          | 1.037       | 0.936            | -1.031     | -1.030      | -1.078           | 4.303     | 4.281       | 4.415            |
| 15 | -0.016    | -0.017      | -0.017           | 0.992          | 1.056       | 0.991            | -1.097     | -1.099      | -1.1405          | 4.458     | 4.461       | 4.577            |
| 16 | -0.023    | -0.025      | -0.023           | 0.965          | 1.048       | 0.965            | -0.971     | -0.974      | -0.993           | 3.964     | 3.967       | 4.014            |
| 17 | 0.025     | 0.036       | 0.034            | 0.740          | 1.016       | 0.740            | -0.600     | -0.567      | -0.763           | 3.267     | 3.233       | 3.503            |
| 18 | -0.012    | -0.013      | -0.015           | 0.938          | 1.041       | 0.941            | -0.948     | -0.947      | -1.085           | 3.906     | 3.903       | 4.230            |
| 19 | -0.022    | -0.023      | -0.027           | 0.988          | 1.055       | 0.994            | -1.058     | -1.059      | -1.199           | 4.198     | 4.197       | 4.546            |
| 20 | -0.006    | -0.007      | -0.005           | 0.855          | 1.016       | 0.854            | -0.726     | -0.732      | -0.755           | 3.401     | 3.405       | 3.450            |

Fonte: O autor (2020).

Tabela 3 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 1.5$  e  $\mu \in (0.3709; 0.8507)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | -0.011    | -0.012      | -0.016           | 1.014          | 1.0915      | 1.026            | -0.991     | -0.991      | -1.195           | 3.888     | 3.887       | 4.436            |
| 2  | 0.006     | 0.007       | 0.009            | 0.882          | 1.070       | 0.878            | -0.640     | -0.637      | -0.782           | 2.946     | 2.941       | 3.172            |
| 3  | 0.042     | 0.048       | 0.042            | 0.934          | 1.070       | 0.938            | -0.849     | -0.852      | -1.042           | 3.372     | 3.379       | 3.797            |
| 4  | -0.013    | -0.015      | -0.015           | 0.984          | 1.091       | 0.984            | -0.865     | -0.864      | -0.986           | 3.341     | 3.341       | 3.586            |
| 5  | 0.008     | 0.010       | 0.008            | 0.945          | 1.065       | 0.944            | -0.825     | -0.825      | -0.966           | 3.283     | 3.283       | 3.569            |
| 6  | 0.059     | 0.081       | 0.065            | 0.771          | 1.069       | 0.764            | -0.521     | -0.520      | -0.621           | 2.745     | 2.756       | 2.858            |
| 7  | 0.002     | 0.002       | 0.002            | 0.937          | 1.096       | 0.939            | -0.803     | -0.803      | -0.964           | 3.267     | 3.266       | 3.594            |
| 8  | -0.001    | -0.001      | -0.000           | 0.935          | 1.083       | 0.935            | -0.838     | -0.837      | -0.990           | 3.449     | 3.447       | 3.772            |
| 9  | -0.047    | -0.049      | -0.055           | 1.078          | 1.135       | 1.092            | -1.125     | -1.125      | -1.317           | 4.342     | 4.340       | 4.942            |
| 10 | -0.028    | -0.031      | -0.032           | 0.997          | 1.108       | 1.002            | -0.839     | -0.840      | -1.008           | 3.326     | 3.329       | 3.684            |
| 11 | 0.038     | 0.049       | 0.041            | 0.824          | 1.054       | 0.817            | -0.571     | -0.572      | -0.662           | 2.767     | 2.775       | 2.879            |
| 12 | 0.019     | 0.024       | 0.029            | 0.814          | 1.078       | 0.814            | -0.570     | -0.573      | -0.749           | 2.843     | 2.823       | 3.106            |
| 13 | 0.005     | 0.005       | 0.007            | 0.915          | 1.055       | 0.919            | -0.761     | -0.766      | -0.950           | 3.232     | 3.233       | 3.599            |
| 14 | 0.011     | 0.013       | 0.009            | 0.952          | 1.080       | 0.960            | -0.890     | -0.891      | -1.091           | 3.516     | 3.515       | 3.976            |
| 15 | -0.017    | -0.019      | -0.023           | 1.031          | 1.110       | 1.038            | -0.993     | -0.993      | -1.166           | 3.799     | 3.800       | 4.249            |
| 16 | -0.019    | -0.021      | -0.022           | 0.997          | 1.107       | 1.001            | -0.867     | -0.866      | -1.015           | 3.347     | 3.347       | 3.660            |
| 17 | 0.023     | 0.029       | 0.030            | 0.858          | 1.076       | 0.859            | -0.660     | -0.661      | -0.834           | 2.983     | 2.979       | 3.281            |
| 18 | -0.026    | -0.029      | -0.029           | 0.996          | 1.102       | 0.996            | -0.867     | -0.866      | -0.991           | 3.405     | 3.403       | 3.674            |
| 19 | -0.042    | -0.045      | -0.047           | 1.036          | 1.116       | 1.038            | -0.907     | -0.907      | -1.032           | 3.408     | 3.409       | 3.675            |
| 20 | 0.005     | 0.005       | 0.008            | 0.884          | 1.070       | 0.883            | -0.647     | -0.649      | -0.810           | 2.878     | 2.879       | 3.140            |

Fonte: O autor (2020).

Tabela 4 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 1.5$  e  $\mu \in (0.8212; 0.9903)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | -0.020    | -0.022      | -0.020           | 1.011          | 1.092       | 1.011            | -0.970     | -0.970      | -0.983           | 3.811     | 3.811       | 3.842            |
| 2  | 0.001     | 0.002       | 0.002            | 0.894          | 1.071       | 0.891            | -0.664     | -0.664      | -0.709           | 2.984     | 2.983       | 3.050            |
| 3  | 0.043     | 0.050       | 0.043            | 0.920          | 1.067       | 0.919            | -0.827     | -0.827      | -0.838           | 3.326     | 3.326       | 3.346            |
| 4  | -0.011    | -0.012      | -0.011           | 0.984          | 1.088       | 0.982            | -0.877     | -0.877      | -0.902           | 3.378     | 3.378       | 3.426            |
| 5  | 0.005     | 0.005       | 0.005            | 0.944          | 1.064       | 0.943            | -0.821     | -0.821      | -0.847           | 3.275     | 3.275       | 3.323            |
| 6  | 0.070     | 0.094       | 0.071            | 0.796          | 1.064       | 0.794            | -0.597     | -0.597      | -0.609           | 2.852     | 2.853       | 2.866            |
| 7  | 0.001     | 0.001       | 0.001            | 0.938          | 1.094       | 0.937            | -0.806     | -0.806      | -0.816           | 3.272     | 3.272       | 3.289            |
| 8  | -0.005    | -0.006      | -0.005           | 0.942          | 1.085       | 0.942            | -0.848     | -0.848      | -0.856           | 3.461     | 3.460       | 3.476            |
| 9  | -0.062    | -0.065      | -0.063           | 1.085          | 1.140       | 1.085            | -1.111     | -1.111      | -1.145           | 4.282     | 4.281       | 4.380            |
| 10 | -0.034    | -0.038      | -0.034           | 0.993          | 1.105       | 0.992            | -0.823     | -0.823      | -0.845           | 3.288     | 3.288       | 3.328            |
| 11 | 0.050     | 0.063       | 0.051            | 0.835          | 1.048       | 0.831            | -0.625     | -0.626      | -0.662           | 2.867     | 2.868       | 2.914            |
| 12 | 0.008     | 0.010       | 0.013            | 0.814          | 1.082       | 0.809            | -0.549     | -0.551      | -0.653           | 2.782     | 2.776       | 2.918            |
| 13 | 0.025     | 0.030       | 0.026            | 0.883          | 1.047       | 0.880            | -0.738     | -0.738      | -0.772           | 3.184     | 3.185       | 3.240            |
| 14 | 0.009     | 0.010       | 0.008            | 0.939          | 1.075       | 0.937            | -0.865     | -0.865      | -0.909           | 3.449     | 3.450       | 3.537            |
| 15 | -0.030    | -0.032      | -0.030           | 1.037          | 1.114       | 1.036            | -0.989     | -0.989      | -1.001           | 3.784     | 3.784       | 3.811            |
| 16 | -0.023    | -0.025      | -0.023           | 1.002          | 1.108       | 1.002            | -0.875     | -0.875      | -0.882           | 3.364     | 3.364       | 3.378            |
| 17 | 0.027     | 0.034       | 0.027            | 0.841          | 1.072       | 0.840            | -0.625     | -0.625      | -0.639           | 2.902     | 2.902       | 2.921            |
| 18 | -0.023    | -0.025      | -0.024           | 1.000          | 1.101       | 0.998            | -0.892     | -0.892      | -0.934           | 3.487     | 3.487       | 3.575            |
| 19 | -0.038    | -0.041      | -0.039           | 1.038          | 1.115       | 1.038            | -0.929     | -0.929      | -0.945           | 3.473     | 3.473       | 3.505            |
| 20 | 0.009     | 0.012       | 0.010            | 0.871          | 1.065       | 0.870            | -0.633     | -0.633      | -0.650           | 2.851     | 2.851       | 2.872            |

Fonte: O autor (2020).

Nas Tabelas 5, 6 e 7 são apresentados as medidas baseadas nos momentos da distribuição empírica dos resíduos quando o parâmetro de precisão é igual a  $\phi = 10$ . É possível notar que nos três cenários as médias dos resíduos são próximas do valor zero, os desvios-padrão estão próximos de um e os valores da assimetria e curtose estão consideravelmente próximos dos valores relativos à distribuição normal padrão.

Tabela 5 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 10$  e  $\mu \in (0.0107; 0.3723)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | -0.014    | -0.015      | -0.014           | 1.059          | 1.140       | 1.059            | -0.445     | -0.444      | -0.450           | 2.879     | 2.879       | 2.885            |
| 2  | 0.050     | 0.058       | 0.050            | 0.976          | 1.147       | 0.978            | -0.426     | -0.428      | -0.437           | 2.903     | 2.904       | 2.914            |
| 3  | 0.003     | 0.004       | 0.003            | 1.040          | 1.136       | 1.040            | -0.379     | -0.380      | -0.382           | 2.798     | 2.803       | 2.802            |
| 4  | -0.022    | -0.026      | -0.022           | 0.983          | 1.129       | 0.988            | -0.350     | -0.351      | -0.364           | 2.772     | 2.772       | 2.782            |
| 5  | 0.022     | 0.024       | 0.022            | 1.028          | 1.142       | 1.030            | -0.398     | -0.397      | -0.407           | 2.798     | 2.797       | 2.805            |
| 6  | -0.015    | -0.027      | -0.013           | 0.686          | 1.134       | 0.691            | -0.090     | -0.089      | -0.104           | 2.748     | 2.688       | 2.750            |
| 7  | -0.004    | -0.004      | -0.004           | 1.047          | 1.148       | 1.047            | -0.359     | -0.360      | -0.361           | 2.762     | 2.763       | 2.764            |
| 8  | -0.039    | -0.042      | -0.039           | 1.078          | 1.154       | 1.078            | -0.353     | -0.355      | -0.356           | 2.688     | 2.689       | 2.690            |
| 9  | -0.005    | -0.005      | -0.005           | 1.097          | 1.151       | 1.104            | -0.482     | -0.482      | -0.499           | 2.891     | 2.892       | 2.909            |
| 10 | -0.004    | -0.005      | -0.004           | 1.006          | 1.129       | 1.014            | -0.348     | -0.348      | -0.367           | 2.822     | 2.823       | 2.839            |
| 11 | -0.007    | -0.009      | -0.007           | 0.921          | 1.131       | 0.923            | -0.268     | -0.267      | -0.277           | 2.675     | 2.672       | 2.679            |
| 12 | -0.018    | -0.035      | -0.016           | 0.592          | 1.137       | 0.597            | -0.024     | -0.006      | -0.038           | 2.729     | 2.663       | 2.728            |
| 13 | -0.006    | -0.006      | -0.006           | 1.029          | 1.145       | 1.030            | -0.382     | -0.379      | -0.387           | 2.736     | 2.734       | 2.741            |
| 14 | 0.005     | 0.006       | 0.005            | 1.047          | 1.136       | 1.048            | -0.428     | -0.427      | -0.435           | 2.938     | 2.933       | 2.946            |
| 15 | 0.003     | 0.003       | 0.003            | 1.089          | 1.154       | 1.090            | -0.409     | -0.409      | -0.416           | 2.916     | 2.917       | 2.924            |
| 16 | 0.003     | 0.003       | 0.003            | 1.077          | 1.156       | 1.077            | -0.414     | -0.415      | -0.418           | 2.861     | 2.862       | 2.866            |
| 17 | 0.021     | 0.032       | 0.022            | 0.774          | 1.140       | 0.779            | -0.221     | -0.202      | -0.236           | 2.788     | 2.734       | 2.795            |
| 18 | 0.003     | 0.003       | 0.003            | 1.019          | 1.138       | 1.024            | -0.346     | -0.346      | -0.360           | 2.733     | 2.733       | 2.743            |
| 19 | 0.027     | 0.029       | 0.027            | 1.078          | 1.151       | 1.083            | -0.433     | -0.433      | -0.449           | 2.847     | 2.848       | 2.865            |
| 20 | -0.002    | -0.002      | -0.002           | 0.977          | 1.134       | 0.978            | -0.357     | -0.357      | -0.361           | 2.777     | 2.772       | 2.780            |

Fonte: O autor (2020).

Tabela 6 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 10$  e  $\mu \in (0.3709; 0.8507)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | -0.038    | -0.041      | -0.039           | 1.066          | 1.144       | 1.074            | -0.431     | -0.431      | -0.451           | 2.872     | 2.872       | 2.893            |
| 2  | 0.028     | 0.034       | 0.028            | 0.940          | 1.151       | 0.943            | -0.337     | -0.337      | -0.349           | 2.741     | 2.741       | 2.749            |
| 3  | 0.023     | 0.026       | 0.023            | 0.994          | 1.130       | 1.001            | -0.353     | -0.353      | -0.371           | 2.766     | 2.767       | 2.783            |
| 4  | -0.032    | -0.035      | -0.032           | 1.016          | 1.128       | 1.018            | -0.390     | -0.390      | -0.400           | 2.815     | 2.815       | 2.823            |
| 5  | 0.022     | 0.025       | 0.022            | 1.012          | 1.142       | 1.015            | -0.389     | -0.389      | -0.400           | 2.792     | 2.792       | 2.801            |
| 6  | 0.002     | 0.002       | 0.002            | 0.795          | 1.138       | 0.796            | -0.183     | -0.182      | -0.189           | 2.695     | 2.689       | 2.698            |
| 7  | 0.018     | 0.021       | 0.018            | 0.977          | 1.146       | 0.981            | -0.315     | -0.315      | -0.328           | 2.715     | 2.715       | 2.725            |
| 8  | 0.002     | 0.002       | 0.002            | 0.980          | 1.144       | 0.984            | -0.319     | -0.319      | -0.331           | 2.668     | 2.669       | 2.677            |
| 9  | -0.015    | -0.016      | -0.016           | 1.093          | 1.152       | 1.098            | -0.473     | -0.473      | -0.488           | 2.880     | 2.880       | 2.897            |
| 10 | 0.000     | 0.000       | 0.000            | 1.017          | 1.128       | 1.022            | -0.370     | -0.370      | -0.387           | 2.867     | 2.868       | 2.883            |
| 11 | 0.000     | 0.000       | 0.000            | 0.872          | 1.131       | 0.873            | -0.226     | -0.227      | -0.233           | 2.655     | 2.657       | 2.657            |
| 12 | 0.010     | 0.014       | 0.011            | 0.872          | 1.138       | 0.878            | -0.266     | -0.265      | -0.283           | 2.743     | 2.742       | 2.753            |
| 13 | -0.017    | -0.019      | -0.017           | 1.005          | 1.139       | 1.014            | -0.331     | -0.330      | -0.350           | 2.700     | 2.698       | 2.715            |
| 14 | 0.003     | 0.004       | 0.003            | 1.008          | 1.135       | 1.016            | -0.369     | -0.368      | -0.388           | 2.823     | 2.821       | 2.843            |
| 15 | -0.004    | -0.005      | -0.005           | 1.071          | 1.155       | 1.076            | -0.384     | -0.384      | -0.400           | 2.857     | 2.857       | 2.873            |
| 16 | -0.001    | -0.001      | -0.001           | 1.039          | 1.157       | 1.043            | -0.363     | -0.363      | -0.376           | 2.760     | 2.760       | 2.773            |
| 17 | 0.023     | 0.029       | 0.024            | 0.923          | 1.145       | 0.928            | -0.298     | -0.298      | -0.315           | 2.774     | 2.770       | 2.786            |
| 18 | -0.014    | -0.016      | -0.014           | 1.029          | 1.141       | 1.031            | -0.342     | -0.343      | -0.352           | 2.722     | 2.723       | 2.730            |
| 19 | 0.003     | 0.003       | 0.002            | 1.067          | 1.152       | 1.070            | -0.390     | -0.390      | -0.400           | 2.766     | 2.766       | 2.776            |
| 20 | -0.014    | -0.017      | -0.013           | 0.944          | 1.135       | 0.949            | -0.299     | -0.299      | -0.314           | 2.685     | 2.685       | 2.695            |

Fonte: O autor (2020).

Tabela 7 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 10$  e  $\mu \in (0.8212; 0.9903)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | -0.044    | -0.048      | -0.044           | 1.059          | 1.144       | 1.059            | -0.420     | -0.420      | -0.421           | 2.853     | 2.853       | 2.854            |
| 2  | 0.026     | 0.032       | 0.026            | 0.962          | 1.152       | 0.963            | -0.363     | -0.363      | -0.366           | 2.777     | 2.778       | 2.780            |
| 3  | 0.025     | 0.029       | 0.025            | 0.974          | 1.131       | 0.974            | -0.339     | -0.339      | -0.340           | 2.759     | 2.760       | 2.760            |
| 4  | -0.032    | -0.036      | -0.032           | 1.020          | 1.129       | 1.020            | -0.394     | -0.394      | -0.396           | 2.824     | 2.824       | 2.825            |
| 5  | 0.019     | 0.021       | 0.019            | 1.014          | 1.142       | 1.014            | -0.384     | -0.384      | -0.386           | 2.789     | 2.789       | 2.791            |
| 6  | 0.008     | 0.011       | 0.008            | 0.848          | 1.136       | 0.848            | -0.236     | -0.236      | -0.237           | 2.715     | 2.714       | 2.715            |
| 7  | 0.018     | 0.021       | 0.018            | 0.981          | 1.146       | 0.981            | -0.321     | -0.321      | -0.321           | 2.726     | 2.726       | 2.727            |
| 8  | 0.000     | 0.000       | 0.000            | 0.991          | 1.144       | 0.991            | -0.329     | -0.329      | -0.330           | 2.678     | 2.678       | 2.678            |
| 9  | -0.025    | -0.026      | -0.025           | 1.097          | 1.154       | 1.097            | -0.469     | -0.469      | -0.471           | 2.877     | 2.878       | 2.880            |
| 10 | -0.004    | -0.004      | -0.004           | 1.013          | 1.128       | 1.013            | -0.359     | -0.359      | -0.360           | 2.852     | 2.852       | 2.853            |
| 11 | 0.005     | 0.006       | 0.005            | 0.901          | 1.130       | 0.901            | -0.258     | -0.258      | -0.260           | 2.672     | 2.673       | 2.673            |
| 12 | 0.006     | 0.008       | 0.007            | 0.868          | 1.139       | 0.869            | -0.252     | -0.253      | -0.260           | 2.729     | 2.729       | 2.732            |
| 13 | 0.000     | 0.000       | 0.000            | 0.955          | 1.133       | 0.955            | -0.297     | -0.297      | -0.299           | 2.660     | 2.659       | 2.661            |
| 14 | 0.003     | 0.004       | 0.003            | 0.991          | 1.134       | 0.992            | -0.343     | -0.343      | -0.346           | 2.787     | 2.787       | 2.790            |
| 15 | -0.012    | -0.013      | -0.012           | 1.075          | 1.156       | 1.075            | -0.381     | -0.381      | -0.382           | 2.853     | 2.853       | 2.854            |
| 16 | -0.003    | -0.003      | -0.003           | 1.046          | 1.158       | 1.046            | -0.369     | -0.369      | -0.370           | 2.768     | 2.768       | 2.769            |
| 17 | 0.027     | 0.034       | 0.027            | 0.896          | 1.143       | 0.896            | -0.270     | -0.270      | -0.271           | 2.740     | 2.740       | 2.740            |
| 18 | -0.012    | -0.013      | -0.012           | 1.036          | 1.141       | 1.036            | -0.353     | -0.353      | -0.356           | 2.733     | 2.733       | 2.735            |
| 19 | 0.004     | 0.004       | 0.004            | 1.072          | 1.152       | 1.072            | -0.400     | -0.400      | -0.401           | 2.789     | 2.789       | 2.790            |
| 20 | -0.009    | -0.012      | -0.009           | 0.927          | 1.134       | 0.927            | -0.282     | -0.282      | -0.284           | 2.672     | 2.673       | 2.673            |

Fonte: O autor (2020).

As Tabelas 8, 9 e 10 apresentam os valores das estatísticas descritivas para os três resíduos considerando  $\phi = 30$  sob os três cenários apresentados anteriormente. Nota-se que a assimetria, curtose, média e desvios-padrão estão próximos dos valores de referência da distribuição normal padrão, de forma mais acentuada que os dois cenários anteriores.

Tabela 8 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 30$  e  $\mu \in (0.0107; 0.3723)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | 0.003     | 0.003       | 0.003            | 1.076          | 1.1595      | 1.076            | -0.241     | -0.241      | -0.242           | 2.717     | 2.717       | 2.717            |
| 2  | 0.000     | 0.000       | 0.000            | 0.976          | 1.146       | 0.977            | -0.216     | -0.216      | -0.219           | 2.684     | 2.685       | 2.685            |
| 3  | 0.007     | 0.008       | 0.007            | 1.044          | 1.137       | 1.044            | -0.261     | -0.260      | -0.262           | 2.686     | 2.685       | 2.686            |
| 4  | -0.000    | 0.000       | -0.000           | 0.995          | 1.145       | 0.996            | -0.188     | -0.186      | -0.191           | 2.736     | 2.738       | 2.737            |
| 5  | -0.008    | -0.009      | -0.008           | 1.031          | 1.144       | 1.031            | -0.196     | -0.196      | -0.197           | 2.712     | 2.711       | 2.712            |
| 6  | 0.022     | 0.036       | 0.022            | 0.687          | 1.153       | 0.689            | -0.030     | -0.041      | -0.032           | 2.660     | 2.647       | 2.660            |
| 7  | 0.001     | 0.002       | 0.001            | 1.049          | 1.146       | 1.050            | -0.191     | -0.191      | -0.191           | 2.660     | 2.661       | 2.661            |
| 8  | 0.003     | 0.003       | 0.003            | 1.078          | 1.149       | 1.078            | -0.260     | -0.260      | -0.261           | 2.786     | 2.788       | 2.787            |
| 9  | -0.020    | -0.021      | -0.020           | 1.092          | 1.146       | 1.094            | -0.287     | -0.287      | -0.290           | 2.872     | 2.872       | 2.875            |
| 10 | -0.002    | -0.002      | -0.002           | 1.034          | 1.159       | 1.037            | -0.172     | -0.172      | -0.176           | 2.656     | 2.656       | 2.658            |
| 11 | -0.001    | -0.002      | -0.001           | 0.931          | 1.140       | 0.931            | -0.160     | -0.159      | -0.162           | 2.682     | 2.676       | 2.682            |
| 12 | 0.002     | 0.004       | 0.002            | 0.580          | 1.141       | 0.582            | -0.019     | -0.023      | -0.021           | 2.622     | 2.600       | 2.621            |
| 13 | 0.002     | 0.003       | 0.002            | 1.032          | 1.145       | 1.032            | -0.202     | -0.201      | -0.203           | 2.688     | 2.686       | 2.689            |
| 14 | 0.035     | 0.038       | 0.035            | 1.052          | 1.139       | 1.052            | -0.238     | -0.238      | -0.239           | 2.727     | 2.727       | 2.728            |
| 15 | 0.009     | 0.010       | 0.009            | 1.107          | 1.172       | 1.107            | -0.253     | -0.252      | -0.254           | 2.669     | 2.668       | 2.670            |
| 16 | -0.016    | -0.017      | -0.016           | 1.069          | 1.146       | 1.069            | -0.253     | -0.253      | -0.254           | 2.770     | 2.770       | 2.770            |
| 17 | -0.012    | -0.018      | -0.011           | 0.763          | 1.138       | 0.765            | -0.081     | -0.079      | -0.083           | 2.678     | 2.658       | 2.678            |
| 18 | -0.003    | -0.003      | -0.003           | 1.036          | 1.159       | 1.038            | -0.252     | -0.252      | -0.255           | 2.822     | 2.821       | 2.824            |
| 19 | 0.000     | 0.000       | 0.000            | 1.084          | 1.158       | 1.087            | -0.226     | -0.226      | -0.229           | 2.654     | 2.654       | 2.655            |
| 20 | -0.021    | -0.025      | -0.021           | 1.014          | 1.173       | 1.014            | -0.194     | -0.194      | -0.195           | 2.632     | 2.632       | 2.633            |

Fonte: O autor (2020).

Tabela 9 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 30$  e  $\mu \in (0.3709; 0.8507)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | -0.010    | -0.011      | -0.010           | 1.081          | 1.160       | 1.084            | -0.236     | -0.236      | -0.240           | 2.706     | 2.706       | 2.707            |
| 2  | -0.004    | -0.005      | -0.003           | 0.939          | 1.150       | 0.940            | -0.159     | -0.159      | -0.161           | 2.625     | 2.625       | 2.625            |
| 3  | 0.013     | 0.015       | 0.013            | 0.993          | 1.129       | 0.996            | -0.226     | -0.226      | -0.230           | 2.689     | 2.688       | 2.691            |
| 4  | -0.007    | -0.007      | -0.007           | 1.031          | 1.146       | 1.032            | -0.211     | -0.211      | -0.213           | 2.756     | 2.756       | 2.756            |
| 5  | -0.011    | -0.013      | -0.011           | 1.015          | 1.145       | 1.016            | -0.189     | -0.189      | -0.192           | 2.691     | 2.691       | 2.693            |
| 6  | 0.036     | 0.051       | 0.036            | 0.802          | 1.151       | 0.802            | -0.087     | -0.090      | -0.089           | 2.657     | 2.656       | 2.657            |
| 7  | 0.014     | 0.017       | 0.014            | 0.975          | 1.144       | 0.976            | -0.161     | -0.162      | -0.164           | 2.640     | 2.640       | 2.640            |
| 8  | 0.019     | 0.023       | 0.020            | 0.980          | 1.145       | 0.982            | -0.226     | -0.225      | -0.228           | 2.778     | 2.777       | 2.779            |
| 9  | -0.026    | -0.027      | -0.026           | 1.087          | 1.146       | 1.089            | -0.273     | -0.273      | -0.276           | 2.860     | 2.859       | 2.862            |
| 10 | 0.002     | 0.002       | 0.002            | 1.048          | 1.162       | 1.050            | -0.185     | -0.184      | -0.188           | 2.673     | 2.673       | 2.675            |
| 11 | 0.005     | 0.007       | 0.005            | 0.878          | 1.140       | 0.878            | -0.132     | -0.133      | -0.134           | 2.686     | 2.685       | 2.686            |
| 12 | 0.018     | 0.024       | 0.019            | 0.876          | 1.140       | 0.878            | -0.158     | -0.157      | -0.161           | 2.610     | 2.606       | 2.610            |
| 13 | -0.007    | -0.008      | -0.007           | 1.013          | 1.145       | 1.016            | -0.165     | -0.164      | -0.169           | 2.667     | 2.667       | 2.669            |
| 14 | 0.033     | 0.037       | 0.033            | 1.009          | 1.136       | 1.012            | -0.200     | -0.200      | -0.204           | 2.734     | 2.734       | 2.737            |
| 15 | 0.000     | 0.000       | 0.000            | 1.084          | 1.169       | 1.086            | -0.243     | -0.243      | -0.246           | 2.651     | 2.651       | 2.654            |
| 16 | -0.020    | -0.022      | -0.020           | 1.031          | 1.148       | 1.033            | -0.218     | -0.218      | -0.221           | 2.750     | 2.750       | 2.751            |
| 17 | -0.011    | -0.014      | -0.011           | 0.917          | 1.137       | 0.919            | -0.132     | -0.131      | -0.135           | 2.686     | 2.684       | 2.687            |
| 18 | -0.008    | -0.009      | -0.008           | 1.044          | 1.158       | 1.044            | -0.242     | -0.242      | -0.244           | 2.798     | 2.798       | 2.800            |
| 19 | -0.013    | -0.014      | -0.013           | 1.076          | 1.161       | 1.076            | -0.210     | -0.210      | -0.212           | 2.670     | 2.670       | 2.671            |
| 20 | -0.021    | -0.025      | -0.021           | 0.977          | 1.174       | 0.979            | -0.157     | -0.157      | -0.160           | 2.596     | 2.595       | 2.597            |

Fonte: O autor (2020).

Tabela 10 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 20$ ,  $\phi = 30$  e  $\mu \in (0.8212; 0.9903)$ .

| t  | Média     |             |                  | Desvios-Padrão |             |                  | Assimetria |             |                  | Curtose   |             |                  |
|----|-----------|-------------|------------------|----------------|-------------|------------------|------------|-------------|------------------|-----------|-------------|------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\phi}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\phi}$ |
| 1  | -0.014    | -0.015      | -0.014           | 1.074          | 1.161       | 1.074            | -0.230     | -0.230      | -0.230           | 2.701     | 2.701       | 2.701            |
| 2  | -0.006    | -0.007      | -0.006           | 0.961          | 1.149       | 0.961            | -0.176     | -0.176      | -0.177           | 2.640     | 2.640       | 2.641            |
| 3  | 0.012     | 0.014       | 0.012            | 0.972          | 1.128       | 0.972            | -0.209     | -0.210      | -0.210           | 2.673     | 2.673       | 2.673            |
| 4  | -0.007    | -0.007      | -0.007           | 1.036          | 1.146       | 1.036            | -0.214     | -0.214      | -0.215           | 2.761     | 2.761       | 2.761            |
| 5  | -0.013    | -0.014      | -0.013           | 1.017          | 1.146       | 1.017            | -0.190     | -0.190      | -0.190           | 2.687     | 2.687       | 2.687            |
| 6  | 0.041     | 0.055       | 0.041            | 0.857          | 1.148       | 0.857            | -0.126     | -0.126      | -0.126           | 2.654     | 2.654       | 2.654            |
| 7  | 0.015     | 0.017       | 0.015            | 0.979          | 1.144       | 0.979            | -0.164     | -0.164      | -0.164           | 2.635     | 2.635       | 2.635            |
| 8  | 0.019     | 0.022       | 0.019            | 0.993          | 1.146       | 0.993            | -0.229     | -0.229      | -0.229           | 2.776     | 2.775       | 2.776            |
| 9  | -0.032    | -0.033      | -0.032           | 1.090          | 1.147       | 1.090            | -0.271     | -0.271      | -0.272           | 2.859     | 2.859       | 2.860            |
| 10 | 0.000     | 0.000       | 0.000            | 1.043          | 1.161       | 1.043            | -0.179     | -0.179      | -0.179           | 2.668     | 2.668       | 2.668            |
| 11 | 0.009     | 0.011       | 0.009            | 0.908          | 1.139       | 0.908            | -0.156     | -0.156      | -0.156           | 2.698     | 2.698       | 2.699            |
| 12 | 0.015     | 0.020       | 0.015            | 0.870          | 1.141       | 0.871            | -0.154     | -0.154      | -0.155           | 2.608     | 2.607       | 2.608            |
| 13 | 0.000     | 0.000       | 0.000            | 0.965          | 1.146       | 0.965            | -0.138     | -0.138      | -0.138           | 2.645     | 2.645       | 2.645            |
| 14 | 0.031     | 0.036       | 0.031            | 0.993          | 1.136       | 0.993            | -0.191     | -0.191      | -0.191           | 2.741     | 2.741       | 2.742            |
| 15 | -0.004    | -0.005      | -0.004           | 1.087          | 1.169       | 1.087            | -0.243     | -0.243      | -0.243           | 2.647     | 2.647       | 2.647            |
| 16 | -0.020    | -0.023      | -0.020           | 1.037          | 1.148       | 1.037            | -0.223     | -0.223      | -0.223           | 2.755     | 2.755       | 2.755            |
| 17 | -0.007    | -0.009      | -0.007           | 0.890          | 1.136       | 0.890            | -0.120     | -0.120      | -0.120           | 2.675     | 2.675       | 2.675            |
| 18 | -0.008    | -0.009      | -0.008           | 1.051          | 1.158       | 1.052            | -0.250     | -0.250      | -0.250           | 2.807     | 2.807       | 2.807            |
| 19 | -0.011    | -0.012      | -0.011           | 1.081          | 1.161       | 1.081            | -0.219     | -0.219      | -0.219           | 2.673     | 2.673       | 2.673            |
| 20 | -0.018    | -0.022      | -0.018           | 0.958          | 1.173       | 0.958            | -0.148     | -0.148      | -0.148           | 2.587     | 2.587       | 2.587            |

Fonte: O autor (2020).

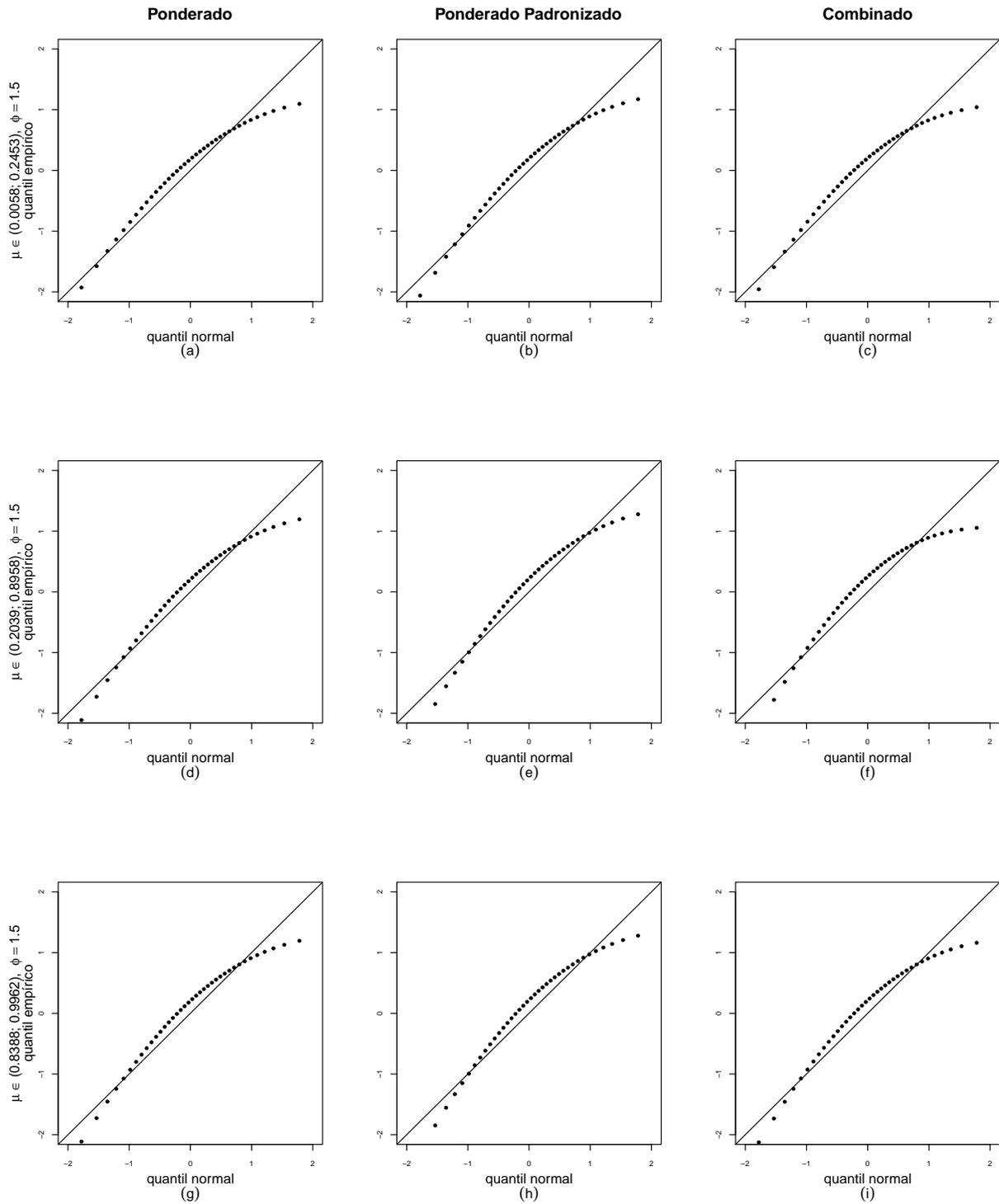
Portanto, a distribuição dos resíduos é bem aproximada pela distribuição normal padrão quando o valor do parâmetro de precisão é grande.

Com o objetivo de comparar os quantis empíricos dos resíduos com os quantis da distribuição normal padrão, construímos gráficos normais de probabilidade, que são apresentados nas Figuras 4-9. Para tanto, consideramos dois tamanhos amostrais ( $n = 40$  e  $n = 120$ ) e três valores para o parâmetro de precisão ( $\phi = 1.5, 10, 30$ ). Foram considerados três cenários para a média da variável resposta como definidos anteriormente.

Com base na Figura 4 é possível notar que os gráficos dos resíduos ponderado, ponderado padronizado e combinado se comportam de maneira semelhantes para os três cenários relativos à média da variável resposta, embora observamos uma má aderência da distribuição dos três resíduos a distribuição normal padrão. No entanto, na Figura 5 os gráficos dos resíduos ponderado, ponderado padronizado e combinado continuam sendo semelhantes entre si, porém mostram uma maior aderência a distribuição normal padrão, independentemente do valor da média da variável resposta esteja localizada. Observe o seguinte: à medida que o valor da precisão aumenta à aderência da distribuição dos três resíduos a distribuição normal padrão se torna melhor, tanto para  $n = 40$  quanto para  $n = 120$ .

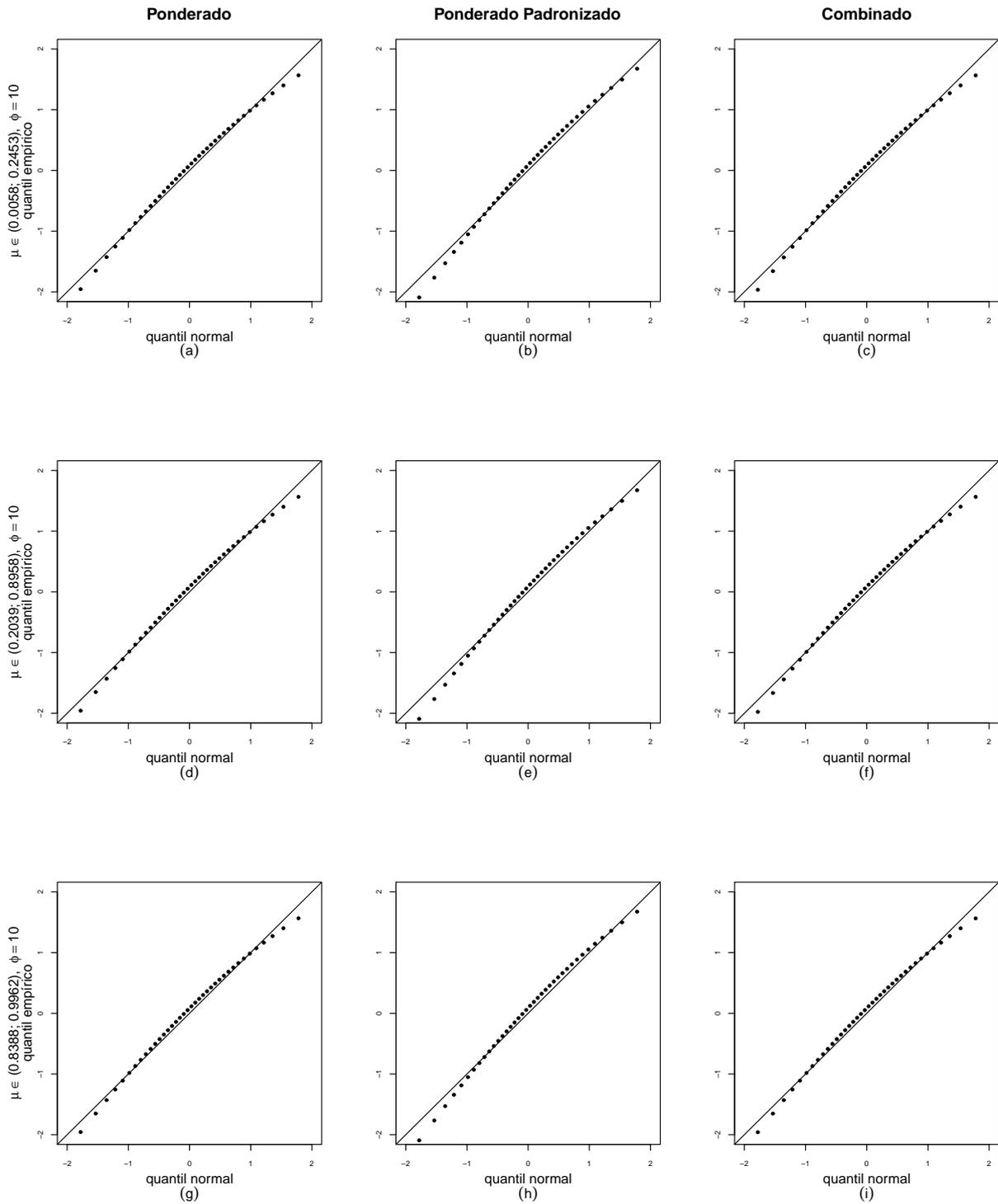
De fato, o tamanho da amostra não parece interferir na distribuição dos resíduos diferentemente do valor da precisão de acordo com as Figuras 4-9. Quando  $\phi = 10$  e  $n = 40$ , Figura (5), a concordância dos quantis da distribuição normal padrão com os quantis empíricos já pode ser considerada boa.

Figura 4 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 40$ ,  $\phi = 1.5$ .



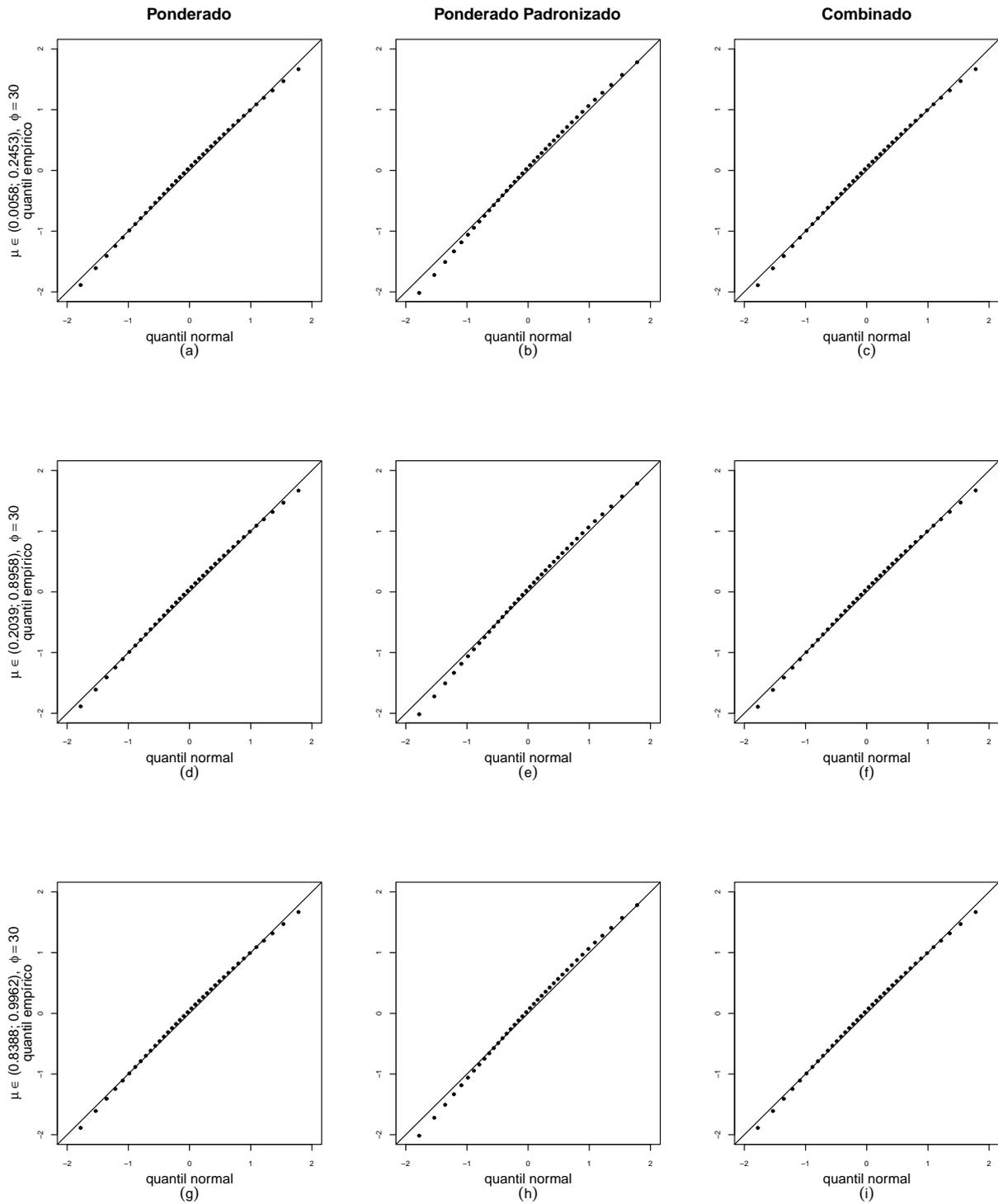
Fonte: O autor (2020).

Figura 5 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 40$ ,  $\phi = 10$ .



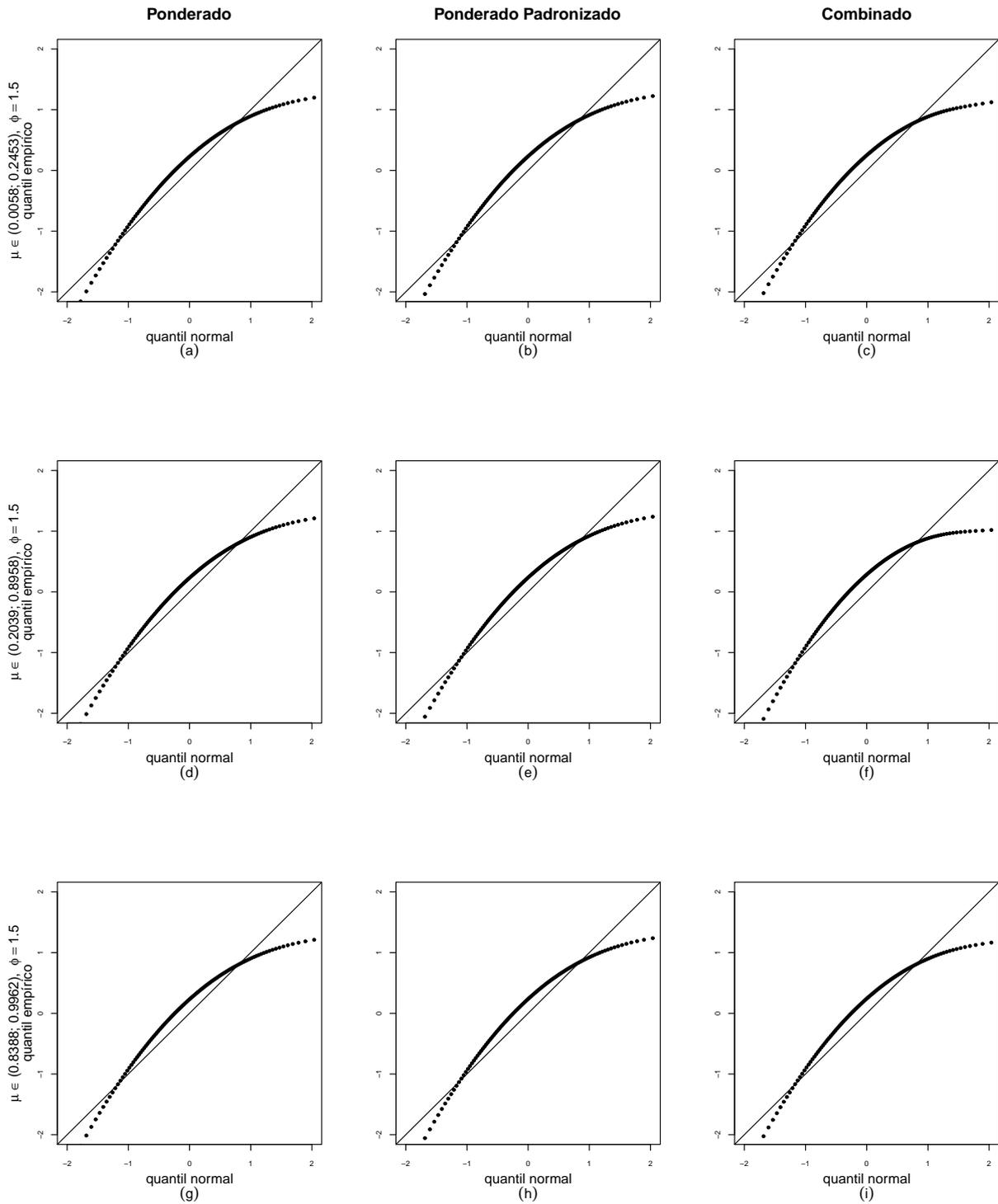
Fonte: O autor (2020).

Figura 6 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 40$ ,  $\phi = 30$ .



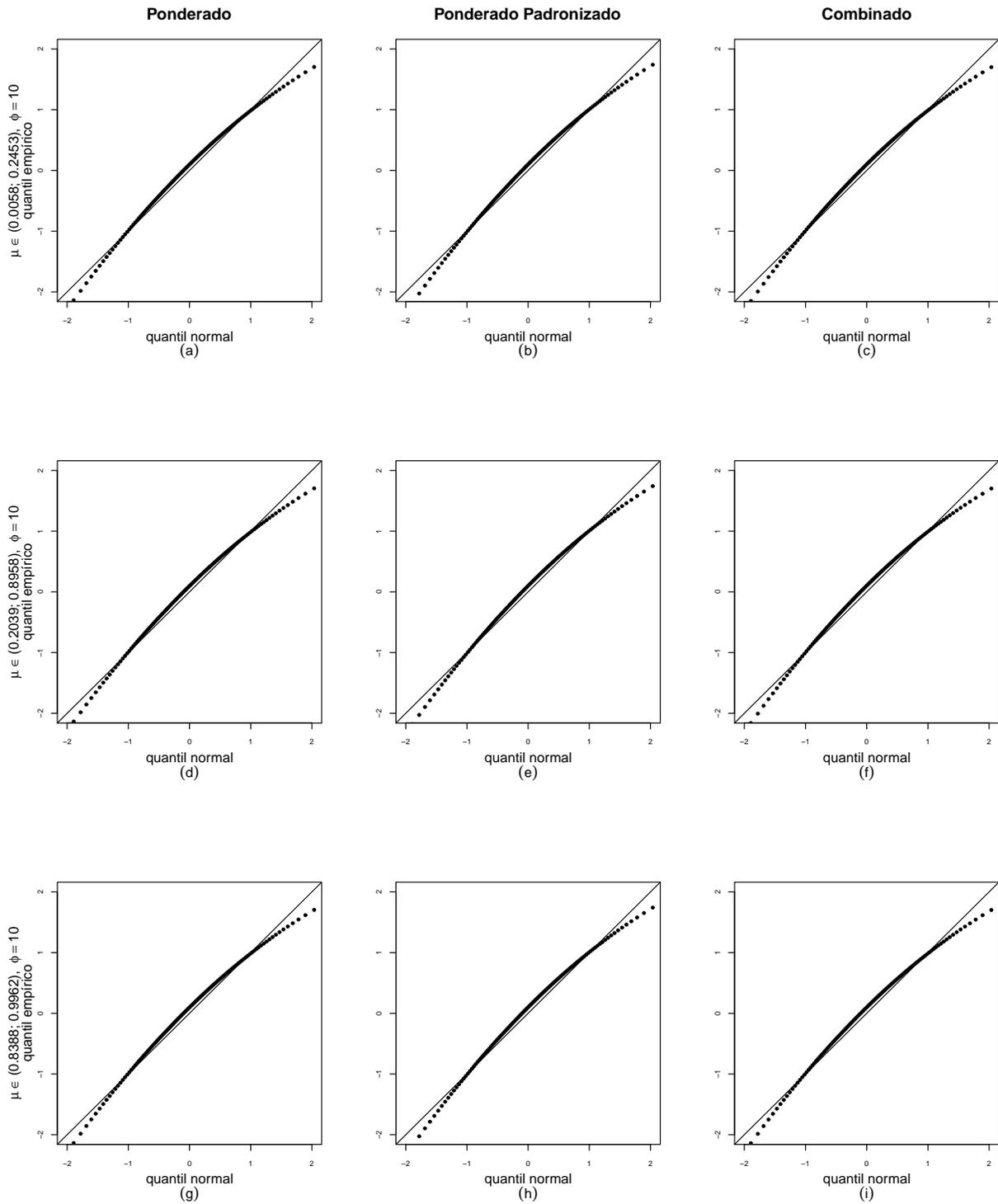
Fonte: O autor (2020).

Figura 7 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 120$ ,  $\phi = 1.5$ .



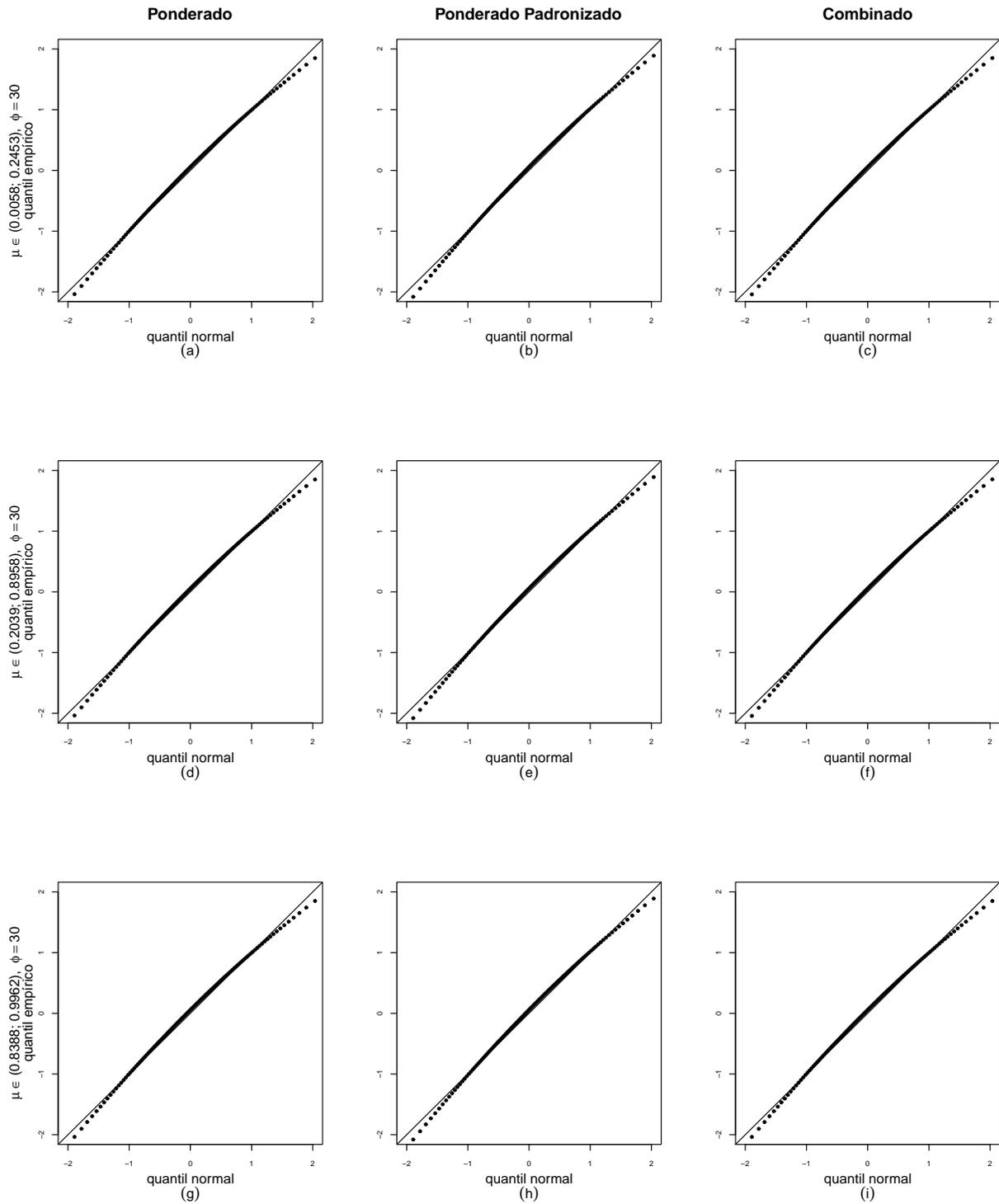
Fonte: O autor (2020).

Figura 8 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 120$ ,  $\phi = 10$ .



Fonte: O autor (2020).

Figura 9 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$ ,  $t = 1, \dots, 120$ ,  $\phi = 30$ .



Fonte: O autor (2020).

Para avaliar o comportamento das distribuições dos resíduos sob a modelagem do parâmetro de precisão, foram realizadas simulações de Monte Carlo com 5.000 réplicas. Para tanto, consideramos o modelo de regressão gama unitária, dado por

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5},$$

$$\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5} \quad \text{com } t = 1, \dots, n.$$

Inicialmente, consideramos o tamanho da amostra  $n = 20$  e três cenários para a média da variável resposta:  $\mu \in (0.0167; 0.2055)$  onde  $\beta = (-2.5, 0.8, -0.8, 1.5, -2)^\top$ ,  $\mu \in (0.2048; 0.8712)$  onde  $\beta = (-3, 1.4, 2.3, 1.3, 2.5)^\top$  e  $\mu \in (0.8172; 0.9908)$  onde  $\beta = (1.9, 1.4, 1, -1.3, 2)^\top$ . Os valores das covariadas foram obtidas como realizações independente das seguintes distribuições:  $x_t \sim U(0, 1)$  e  $z_t \sim U(0, 1)$  e foram mantidos fixas para cada réplica. Além disso, consideramos três valores para o grau de precisão não constante,  $\lambda = 30$  onde  $\gamma = (-1.4, 2.1, 1.7, 1.7, 1)^\top$ ,  $\lambda = 100$  onde  $\gamma = (-2.1, 2.2, 2, 2.1, 2.14)^\top$  e  $\lambda = 530$  onde  $\gamma = (2.5062, 2.7, -2.7, -2.2, 2.4)^\top$ , em que  $\lambda = \max(\phi_t)/\min(\phi_t)$ , com  $t = 1, \dots, n$ . Consideramos estes valores para o parâmetros de precisão plausíveis de acordo com aplicações realizadas.

As Tabelas 11, 12 e 13 mostram as médias, desvios-padrão, assimetrias e curtoses dos resíduos ( $r^\beta$ ), ( $r_p^\beta$ ) e ( $r_p^{\beta\gamma}$ ), considerando os três cenários da média da variável resposta e  $\lambda = 30$ .

Tabela 11 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1-\mu_t)] = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$ ,  $t = 1, \dots, 20$ ,  $\lambda = 30$  e  $\mu \in (0.0167; 0.2055)$ .

| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | 0.044     | 0.046       | 0.044              | 0.917          | 0.945       | 0.916              | -0.117     | -0.121      | -0.127             | 2.257     | 2.259       | 2.264              |
| 2  | -0.002    | -0.003      | 0.002              | 0.803          | 0.874       | 0.796              | -0.205     | -0.206      | -0.259             | 2.319     | 2.325       | 2.384              |
| 3  | -0.026    | -0.030      | -0.027             | 1.096          | 1.195       | 1.099              | -0.172     | -0.162      | -0.189             | 2.325     | 2.302       | 2.339              |
| 4  | 0.029     | 0.047       | 0.029              | 1.071          | 1.490       | 1.073              | -0.042     | -0.029      | -0.044             | 1.807     | 1.679       | 1.811              |
| 5  | 0.033     | 0.053       | 0.034              | 0.919          | 1.472       | 0.920              | -0.121     | -0.115      | -0.124             | 2.075     | 1.709       | 2.078              |
| 6  | 0.045     | 0.062       | 0.045              | 0.851          | 1.289       | 0.851              | 0.035      | -0.045      | 0.034              | 2.746     | 2.170       | 2.748              |
| 7  | -0.001    | 0.014       | -0.001             | 0.928          | 1.187       | 0.928              | 0.029      | 0.031       | 0.026              | 2.628     | 2.284       | 2.630              |
| 8  | 0.001     | 0.005       | 0.000              | 0.962          | 1.121       | 0.963              | -0.094     | -0.064      | -0.103             | 2.488     | 2.290       | 2.496              |
| 9  | -0.027    | -0.028      | -0.027             | 1.022          | 1.051       | 1.023              | -0.172     | -0.168      | -0.199             | 2.077     | 2.078       | 2.090              |
| 10 | -0.044    | -0.045      | -0.045             | 1.110          | 1.153       | 1.112              | -0.271     | -0.263      | -0.286             | 2.503     | 2.474       | 2.518              |
| 11 | -0.176    | -0.184      | -0.183             | 1.173          | 1.261       | 1.189              | -0.446     | -0.400      | -0.514             | 2.456     | 2.349       | 2.550              |
| 12 | -0.044    | -0.051      | -0.036             | 0.877          | 1.052       | 0.879              | -0.552     | -0.405      | -0.644             | 3.350     | 2.889       | 3.590              |
| 13 | 0.074     | 0.146       | 0.074              | 0.750          | 1.383       | 0.751              | -0.083     | -0.114      | -0.086             | 3.058     | 2.147       | 3.064              |
| 14 | -0.089    | -0.100      | -0.090             | 1.290          | 1.452       | 1.293              | -0.026     | -0.011      | -0.030             | 2.176     | 2.029       | 2.176              |
| 15 | -0.029    | -0.031      | -0.030             | 1.111          | 1.161       | 1.113              | -0.220     | -0.213      | -0.233             | 2.218     | 2.187       | 2.229              |
| 16 | -0.150    | -0.155      | -0.151             | 0.998          | 1.027       | 0.998              | -0.071     | -0.062      | -0.085             | 2.412     | 2.390       | 2.420              |
| 17 | -0.091    | -0.109      | -0.092             | 0.961          | 1.244       | 0.964              | -0.129     | -0.058      | -0.145             | 2.802     | 2.363       | 2.823              |
| 18 | 0.021     | 0.022       | 0.026              | 0.830          | 0.862       | 0.818              | -0.470     | -0.460      | -0.571             | 2.467     | 2.453       | 2.616              |
| 19 | 0.129     | 0.131       | 0.129              | 0.894          | 0.921       | 0.892              | -0.430     | -0.428      | -0.454             | 2.585     | 2.601       | 2.616              |
| 20 | 0.031     | 0.033       | 0.031              | 0.911          | 0.977       | 0.911              | -0.080     | -0.083      | -0.088             | 2.368     | 2.383       | 2.374              |

Fonte: O autor (2020).

Tabela 12 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 20$ ,  $\lambda = 30$  e  $\mu \in (0.2048; 0.8712)$ .

| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | 0.038     | 0.042       | 0.058              | 0.919          | 0.988       | 0.931              | -0.518     | -0.444      | -0.609             | 2.794     | 2.740       | 2.868              |
| 2  | 0.026     | 0.028       | 0.037              | 0.931          | 1.001       | 0.923              | -0.931     | -0.850      | -1.090             | 3.602     | 3.423       | 4.000              |
| 3  | 0.016     | 0.015       | 0.019              | 1.070          | 1.181       | 1.079              | -0.382     | -0.339      | -0.427             | 2.580     | 2.503       | 2.635              |
| 4  | 0.020     | 0.026       | 0.020              | 1.179          | 1.573       | 1.180              | -0.096     | -0.070      | -0.098             | 1.658     | 1.539       | 1.660              |
| 5  | 0.004     | -0.001      | 0.008              | 1.042          | 1.550       | 1.045              | -0.109     | -0.053      | -0.121             | 1.954     | 1.552       | 1.952              |
| 6  | -0.021    | -0.039      | -0.021             | 0.704          | 1.362       | 0.705              | -0.275     | -0.021      | -0.281             | 4.329     | 2.245       | 4.341              |
| 7  | -0.009    | 0.005       | 0.000              | 0.743          | 1.147       | 0.748              | -0.545     | -0.134      | -0.585             | 4.061     | 2.715       | 4.090              |
| 8  | 0.009     | 0.030       | 0.027              | 0.812          | 1.068       | 0.819              | -0.682     | -0.286      | -0.747             | 3.854     | 2.790       | 3.918              |
| 9  | 0.011     | 0.009       | 0.018              | 1.033          | 1.077       | 1.042              | -0.424     | -0.397      | -0.487             | 2.305     | 2.289       | 2.370              |
| 10 | 0.023     | 0.023       | 0.026              | 1.102          | 1.148       | 1.110              | -0.489     | -0.467      | -0.527             | 2.783     | 2.719       | 2.830              |
| 11 | -0.024    | -0.023      | -0.022             | 1.077          | 1.160       | 1.084              | -0.595     | -0.509      | -0.662             | 2.655     | 2.477       | 2.786              |
| 12 | 0.043     | 0.044       | 0.048              | 0.863          | 0.976       | 0.862              | -0.463     | -0.404      | -0.532             | 2.663     | 2.703       | 2.773              |
| 13 | 0.005     | 0.017       | 0.005              | 0.773          | 1.279       | 0.775              | -0.087     | 0.008       | -0.098             | 3.489     | 2.363       | 3.504              |
| 14 | -0.047    | -0.052      | -0.047             | 1.267          | 1.375       | 1.271              | -0.159     | -0.130      | -0.166             | 2.230     | 2.070       | 2.233              |
| 15 | -0.010    | -0.012      | -0.005             | 1.116          | 1.177       | 1.124              | -0.378     | -0.347      | -0.413             | 2.381     | 2.325       | 2.405              |
| 16 | -0.054    | -0.056      | -0.036             | 1.130          | 1.165       | 1.137              | -0.606     | -0.580      | -0.682             | 2.859     | 2.801       | 2.947              |
| 17 | 0.009     | 0.020       | 0.009              | 0.931          | 1.202       | 0.938              | -0.460     | -0.268      | -0.497             | 3.328     | 2.617       | 3.405              |
| 18 | 0.015     | 0.014       | 0.038              | 0.948          | 0.981       | 0.939              | -0.836     | -0.795      | -1.001             | 3.138     | 3.084       | 3.497              |
| 19 | 0.017     | 0.018       | 0.046              | 0.950          | 0.994       | 0.959              | -0.584     | -0.538      | -0.683             | 3.140     | 3.055       | 3.200              |
| 20 | -0.002    | 0.000       | 0.008              | 0.913          | 1.036       | 0.922              | -0.649     | -0.474      | -0.725             | 3.408     | 3.136       | 3.553              |

Fonte: O autor (2020).

É possível notar que as médias para cada uma das vinte observações dos três resíduos estão bem próximas de zero nos três cenários, nos desvios-padrão temos uma maior variação de valores principalmente no resíduo ponderado padronizado quando comparamos com o valor de referência da normal padrão que é igual a um.

Com relação aos valores de assimetria na Tabela 11 temos assimetria negativa exceto para a observação 6, no resíduo ponderado e combinado e para observação 7 com relação aos três resíduos; na Tabela 12, temos assimetria negativa, exceto para a observação 13, no resíduo ponderado padronizado, e na Tabela 13, temos assimetria negativa para todos os resíduos, nas vinte observações.

Nos valores da curtose, na Tabela 11, temos em geral valores relativamente próximos de três, embora ainda existam observações com valores superiores a este valor de referência. O mesmo fato também pode ser observado principalmente nas Tabelas 12 e 13.

As tabelas 14, 15 e 16 contêm os valores das estatísticas referentes aos três resíduos quando  $\lambda = 100$ . Os valores das médias nos três cenários para as vinte observações estão relativamente próximas de zero, os valores dos desvios-padrão estão relativamente próximos de um.

Com relação à assimetria de modo geral, podemos afirmar que é negativa, embora existam algumas observações com assimetria positiva e é possível notarmos ainda que no primeiro cenário esta medida esta próxima de zero e nos demais cenários isso não ocorre. Sobre os valores da curtose, na Tabela 14 os valores estão razoavelmente próximos de três, embora existam observações com valores acima deste valor, como a observação 12.

Tabela 13 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}, t = 1, \dots, 20$ ,  $\lambda = 30$  e  $\mu \in (0.8172; 0.9908)$ .

| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | 0.003     | 0.001       | 0.015              | 0.912          | 0.984       | 0.915              | -0.592     | -0.461      | -0.625             | 3.177     | 3.026       | 3.184              |
| 2  | 0.010     | 0.011       | 0.019              | 0.935          | 0.983       | 0.935              | -1.142     | -1.058      | -1.206             | 4.494     | 4.235       | 4.679              |
| 3  | 0.019     | 0.018       | 0.019              | 1.026          | 1.153       | 1.026              | -0.526     | -0.436      | -0.530             | 3.115     | 2.882       | 3.120              |
| 4  | -0.001    | 0.001       | -0.001             | 1.185          | 1.573       | 1.186              | -0.104     | -0.064      | -0.105             | 1.736     | 1.586       | 1.736              |
| 5  | -0.005    | -0.015      | 0.006              | 1.015          | 1.538       | 1.019              | -0.120     | -0.031      | -0.145             | 2.134     | 1.629       | 2.109              |
| 6  | -0.032    | -0.039      | -0.032             | 0.783          | 1.350       | 0.783              | -0.358     | -0.047      | -0.359             | 3.924     | 2.295       | 3.926              |
| 7  | -0.013    | 0.005       | -0.009             | 0.722          | 1.123       | 0.723              | -0.561     | -0.144      | -0.570             | 4.244     | 2.899       | 4.236              |
| 8  | -0.008    | 0.000       | 0.000              | 0.782          | 1.029       | 0.785              | -0.670     | -0.289      | -0.691             | 4.027     | 2.983       | 4.010              |
| 9  | -0.018    | -0.018      | -0.009             | 1.034          | 1.081       | 1.035              | -0.466     | -0.422      | -0.495             | 2.568     | 2.550       | 2.591              |
| 10 | -0.001    | 0.000       | 0.000              | 1.088          | 1.138       | 1.089              | -0.498     | -0.466      | -0.505             | 2.926     | 2.874       | 2.935              |
| 11 | -0.000    | 0.000       | 0.005              | 1.051          | 1.112       | 1.053              | -0.715     | -0.638      | -0.742             | 2.975     | 2.807       | 3.013              |
| 12 | 0.024     | 0.028       | 0.060              | 0.837          | 0.983       | 0.843              | -0.798     | -0.556      | -0.926             | 3.830     | 3.375       | 3.993              |
| 13 | -0.025    | -0.030      | -0.025             | 0.689          | 1.268       | 0.689              | -0.344     | 0.080       | -0.346             | 4.457     | 2.490       | 4.459              |
| 14 | -0.011    | -0.013      | -0.010             | 1.201          | 1.328       | 1.201              | -0.285     | -0.234      | -0.288             | 2.509     | 2.242       | 2.510              |
| 15 | -0.011    | -0.013      | -0.008             | 1.111          | 1.165       | 1.112              | -0.426     | -0.396      | -0.433             | 2.495     | 2.413       | 2.497              |
| 16 | -0.031    | -0.033      | -0.026             | 1.109          | 1.137       | 1.110              | -0.632     | -0.620      | -0.646             | 3.030     | 2.992       | 3.040              |
| 17 | 0.000     | 0.010       | 0.000              | 0.874          | 1.159       | 0.874              | -0.634     | -0.301      | -0.641             | 3.923     | 2.859       | 3.934              |
| 18 | 0.000     | 0.000       | 0.027              | 0.940          | 0.962       | 0.943              | -1.064     | -1.025      | -1.161             | 4.110     | 4.013       | 4.314              |
| 19 | -0.013    | -0.014      | -0.001             | 0.936          | 0.991       | 0.938              | -0.568     | -0.482      | -0.601             | 3.323     | 3.316       | 3.332              |
| 20 | -0.040    | -0.042      | -0.038             | 0.898          | 1.011       | 0.898              | -0.700     | -0.494      | -0.711             | 3.827     | 3.424       | 3.846              |

Fonte: O autor (2020).

Tabela 14 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}, t = 1, \dots, 20$ ,  $\lambda = 100$  e  $\mu \in (0.0167; 0.2055)$ .

| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | 0.028     | 0.030       | 0.027              | 0.950          | 0.975       | 0.949              | -0.112     | -0.109      | -0.119             | 2.319     | 2.314       | 2.323              |
| 2  | -0.076    | -0.082      | -0.069             | 0.793          | 0.854       | 0.786              | -0.239     | -0.237      | -0.298             | 2.339     | 2.319       | 2.378              |
| 3  | -0.061    | -0.070      | -0.062             | 1.107          | 1.243       | 1.109              | -0.166     | -0.151      | -0.174             | 2.423     | 2.386       | 2.429              |
| 4  | 0.005     | 0.014       | 0.005              | 0.995          | 1.524       | 0.995              | 0.004      | 0.014       | 0.003              | 1.777     | 1.626       | 1.778              |
| 5  | 0.046     | 0.081       | 0.046              | 1.047          | 1.524       | 1.048              | -0.143     | -0.130      | -0.144             | 1.874     | 1.622       | 1.876              |
| 6  | 0.061     | 0.117       | 0.061              | 0.677          | 1.274       | 0.677              | 0.068      | -0.050      | 0.068              | 3.442     | 2.251       | 3.442              |
| 7  | -0.095    | -0.101      | -0.095             | 0.996          | 1.151       | 0.997              | 0.049      | 0.066       | 0.045              | 2.486     | 2.284       | 2.487              |
| 8  | 0.026     | 0.039       | 0.026              | 0.910          | 1.195       | 0.911              | -0.016     | -0.003      | -0.022             | 2.764     | 2.269       | 2.773              |
| 9  | 0.002     | 0.002       | 0.002              | 0.943          | 0.954       | 0.942              | -0.234     | -0.230      | -0.277             | 2.211     | 2.210       | 2.241              |
| 10 | -0.107    | -0.111      | -0.108             | 1.166          | 1.211       | 1.168              | -0.222     | -0.211      | -0.228             | 2.400     | 2.364       | 2.405              |
| 11 | -0.241    | -0.251      | -0.250             | 1.205          | 1.274       | 1.224              | -0.389     | -0.355      | -0.459             | 2.386     | 2.305       | 2.481              |
| 12 | 0.024     | 0.028       | 0.037              | 0.812          | 0.899       | 0.804              | -0.964     | -0.839      | -1.105             | 6.429     | 5.277       | 6.921              |
| 13 | 0.057     | 0.142       | 0.057              | 0.593          | 1.393       | 0.593              | 0.018      | -0.081      | 0.016              | 3.627     | 2.171       | 3.630              |
| 14 | -0.096    | -0.106      | -0.096             | 1.329          | 1.476       | 1.330              | -0.001     | 0.011       | -0.003             | 2.172     | 2.025       | 2.172              |
| 15 | -0.043    | -0.044      | -0.044             | 1.138          | 1.185       | 1.141              | -0.160     | -0.154      | -0.170             | 2.223     | 2.198       | 2.23               |
| 16 | -0.002    | -0.003      | -0.003             | 0.856          | 0.872       | 0.853              | -0.338     | -0.332      | -0.363             | 2.752     | 2.732       | 2.786              |
| 17 | -0.147    | -0.174      | -0.148             | 1.015          | 1.217       | 1.018              | -0.114     | -0.079      | -0.130             | 2.699     | 2.450       | 2.714              |
| 18 | -0.017    | -0.017      | -0.012             | 0.813          | 0.838       | 0.797              | -0.516     | -0.511      | -0.634             | 2.546     | 2.528       | 2.730              |
| 19 | 0.082     | 0.082       | 0.083              | 0.922          | 0.942       | 0.923              | -0.383     | -0.377      | -0.399             | 2.602     | 2.591       | 2.614              |
| 20 | 0.031     | 0.032       | 0.031              | 0.917          | 0.988       | 0.917              | -0.172     | -0.178      | -0.178             | 2.430     | 2.473       | 2.435              |

Fonte: O autor (2020).

Tabela 15 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 20$ ,  $\lambda = 100$  e  $\mu \in (0.2048; 0.8712)$ .

| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | 0.017     | 0.014       | 0.031              | 0.944          | 1.019       | 0.955              | -0.418     | -0.364      | -0.480             | 2.610     | 2.596       | 2.639              |
| 2  | -0.006    | -0.005      | 0.006              | 0.950          | 1.001       | 0.942              | -0.938     | -0.870      | -1.104             | 3.437     | 3.295       | 3.848              |
| 3  | -0.001    | -0.002      | 0.000              | 1.091          | 1.243       | 1.098              | -0.288     | -0.241      | -0.309             | 2.539     | 2.403       | 2.555              |
| 4  | -0.012    | -0.023      | -0.012             | 1.088          | 1.597       | 1.088              | -0.046     | -0.007      | -0.048             | 1.742     | 1.528       | 1.745              |
| 5  | 0.013     | 0.023       | 0.017              | 1.176          | 1.582       | 1.178              | -0.107     | -0.069      | -0.117             | 1.699     | 1.463       | 1.696              |
| 6  | 0.001     | 0.008       | 0.001              | 0.566          | 1.360       | 0.566              | -0.288     | 0.002       | -0.291             | 6.399     | 2.264       | 6.411              |
| 7  | -0.022    | -0.026      | -0.009             | 0.788          | 1.050       | 0.795              | -0.529     | -0.264      | -0.574             | 3.674     | 2.872       | 3.681              |
| 8  | 0.011     | 0.038       | 0.020              | 0.728          | 1.101       | 0.733              | -0.513     | -0.129      | -0.546             | 3.997     | 2.513       | 3.998              |
| 9  | 0.034     | 0.034       | 0.043              | 0.976          | 0.994       | 0.985              | -0.575     | -0.557      | -0.666             | 2.581     | 2.588       | 2.702              |
| 10 | -0.011    | -0.012      | -0.009             | 1.191          | 1.242       | 1.197              | -0.364     | -0.346      | -0.381             | 2.45      | 2.420       | 2.468              |
| 11 | -0.056    | -0.056      | -0.056             | 1.121          | 1.182       | 1.129              | -0.547     | -0.493      | -0.612             | 2.438     | 2.326       | 2.543              |
| 12 | 0.111     | 0.114       | 0.117              | 0.776          | 0.831       | 0.769              | -0.777     | -0.699      | -0.896             | 3.410     | 3.516       | 3.723              |
| 13 | 0.005     | 0.003       | 0.005              | 0.650          | 1.293       | 0.651              | 0.046      | 0.042       | 0.039              | 4.523     | 2.314       | 4.535              |
| 14 | -0.013    | -0.014      | -0.013             | 1.300          | 1.407       | 1.302              | -0.127     | -0.107      | -0.129             | 2.152     | 1.984       | 2.154              |
| 15 | -0.010    | -0.012      | -0.006             | 1.154          | 1.209       | 1.161              | -0.344     | -0.326      | -0.370             | 2.307     | 2.257       | 2.325              |
| 16 | 0.009     | 0.009       | 0.038              | 1.019          | 1.036       | 1.023              | -0.790     | -0.778      | -0.919             | 3.401     | 3.371       | 3.614              |
| 17 | 0.005     | 0.009       | 0.006              | 0.986          | 1.174       | 0.993              | -0.414     | -0.317      | -0.448             | 3.029     | 2.669       | 3.089              |
| 18 | 0.009     | 0.008       | 0.032              | 0.927          | 0.950       | 0.914              | -0.884     | -0.852      | -1.070             | 3.238     | 3.194       | 3.679              |
| 19 | 0.006     | 0.009       | 0.026              | 0.986          | 1.029       | 0.994              | -0.502     | -0.447      | -0.567             | 3.072     | 2.997       | 3.086              |
| 20 | -0.017    | -0.008      | -0.006             | 0.888          | 1.026       | 0.896              | -0.529     | -0.375      | -0.584             | 3.106     | 2.968       | 3.170              |

Fonte: O autor (2020).

Tabela 16 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 20$ ,  $\lambda = 100$  e  $\mu \in (0.8172; 0.9908)$ .

| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | -0.021    | -0.027      | -0.014             | 0.941          | 1.015       | 0.943              | -0.472     | -0.405      | -0.489             | 2.842     | 2.76        | 2.841              |
| 2  | 0.004     | 0.005       | 0.013              | 0.943          | 0.980       | 0.942              | -1.149     | -1.081      | -1.211             | 4.248     | 4.068       | 4.414              |
| 3  | 0.009     | 0.010       | 0.009              | 1.020          | 1.208       | 1.020              | -0.400     | -0.308      | -0.402             | 2.979     | 2.687       | 2.981              |
| 4  | -0.022    | -0.031      | -0.022             | 1.087          | 1.593       | 1.087              | -0.026     | 0.010       | -0.026             | 1.765     | 1.582       | 1.765              |
| 5  | 0.002     | 0.008       | 0.015              | 1.173          | 1.581       | 1.177              | -0.109     | -0.058      | -0.136             | 1.803     | 1.515       | 1.787              |
| 6  | -0.015    | -0.010      | -0.015             | 0.635          | 1.356       | 0.635              | -0.447     | 0.005       | -0.448             | 6.006     | 2.338       | 6.008              |
| 7  | -0.014    | -0.017      | -0.010             | 0.743          | 1.002       | 0.744              | -0.528     | -0.243      | -0.541             | 3.920     | 3.180       | 3.916              |
| 8  | -0.004    | 0.003       | -0.001             | 0.704          | 1.082       | 0.706              | -0.526     | -0.129      | -0.536             | 4.300     | 2.686       | 4.286              |
| 9  | 0.009     | 0.011       | 0.021              | 0.949          | 0.971       | 0.950              | -0.634     | -0.599      | -0.675             | 2.900     | 2.887       | 2.947              |
| 10 | -0.022    | -0.022      | -0.022             | 1.190          | 1.243       | 1.190              | -0.376     | -0.351      | -0.378             | 2.535     | 2.494       | 2.537              |
| 11 | -0.020    | -0.021      | -0.015             | 1.091          | 1.136       | 1.092              | -0.668     | -0.626      | -0.692             | 2.798     | 2.678       | 2.825              |
| 12 | 0.056     | 0.060       | 0.121              | 0.771          | 0.848       | 0.776              | -1.005     | -0.802      | -1.215             | 4.380     | 4.241       | 4.719              |
| 13 | -0.024    | -0.038      | -0.024             | 0.575          | 1.300       | 0.575              | -0.389     | 0.099       | -0.391             | 5.812     | 2.356       | 5.812              |
| 14 | 0.011     | 0.013       | 0.011              | 1.233          | 1.362       | 1.234              | -0.234     | -0.197      | -0.235             | 2.406     | 2.166       | 2.406              |
| 15 | -0.020    | -0.023      | -0.019             | 1.133          | 1.179       | 1.133              | -0.383     | -0.366      | -0.387             | 2.431     | 2.376       | 2.433              |
| 16 | 0.014     | 0.015       | 0.024              | 0.993          | 1.010       | 0.993              | -0.859     | -0.832      | -0.887             | 3.916     | 3.889       | 3.945              |
| 17 | 0.007     | 0.013       | 0.008              | 0.927          | 1.115       | 0.927              | -0.571     | -0.396      | -0.579             | 3.474     | 2.931       | 3.483              |
| 18 | 0.005     | 0.004       | 0.034              | 0.932          | 0.948       | 0.932              | -1.089     | -1.058      | -1.191             | 4.002     | 3.931       | 4.221              |
| 19 | -0.022    | -0.020      | -0.014             | 0.980          | 1.024       | 0.981              | -0.482     | -0.393      | -0.504             | 3.191     | 3.120       | 3.190              |
| 20 | -0.024    | -0.019      | -0.022             | 0.872          | 1.001       | 0.872              | -0.553     | -0.382      | -0.561             | 3.328     | 3.147       | 3.332              |

Fonte: O autor (2020).

Nas Tabelas 15 e 16 há valores próximos de três, porém existem observações, como por exemplo a observação 6 em ambas as tabelas, que possuem valores superiores a seis, sendo estes valores muito superiores ao valor de referência da distribuição normal padrão. Este fato ocorre para os resíduos ponderado e combinado.

A seguir, apresentamos as Tabelas 17, 18 e 19, que mostram o comportamento dos resíduos ponderado, ponderado padronizado e combinado para os três cenários da média da variável resposta com  $\lambda = 530$ . Os valores das médias nos três cenários para as vinte observações são próximas de zero, os valores dos desvios-padrão estão relativamente próximos de um. Em geral, há assimetria negativa, embora existam algumas observações com assimetria positiva, esta medida estando razoavelmente próxima de zero.

Quanto à medida de curtose, na Tabela 17 se destacam as observações 11, 19 e 20, apresentando valores superior a três, para os resíduos ponderado e combinado. Também em relação a esses dois resíduos nas Tabelas 18 e 19, temos observações cujas curtoses excedem muito o valor três, como por exemplo as observações 19 e 20, com valores superiores a oito, ou seja, os valores das medidas de curtose estão em geral distantes de três e ainda existem observações que excedem esse valor.

De forma geral, quando a precisão do modelo não é constante à medida que o grau de heterogeneidade da precisão aumenta as distribuições dos resíduos ponderado e combinado apresentam progressivamente pior aderência à distribuição normal padrão, em especial com relação à curtose. É interessante notar que este fato é muito mais acentuado quando as médias da variável resposta estão próximas do extremo superior do intervalo unitário, ou seja, quando  $\mu \approx 1$ .

Para comparar os quantis empíricos dos resíduos com os quantis da distribuição normal padrão, foram construídos os gráficos normais de probabilidade apresentados nas Figuras 10-15. Consideramos os tamanhos amostrais  $n = 40$  e  $n = 120$ , os três cenários para a média da variável resposta e  $\lambda = (30, 100, 530)^T$ . Com base nessas figuras notamos que as distribuições dos resíduos ponderado, ponderado padronizado e combinado se comportam de maneira semelhante considerando todos os valores de  $\lambda$  e tamanhos amostrais, apresentando em muitos casos certa falta de aderência à distribuição normal padrão.

Este fato, para o modelo de regressão gama unitária, implica que os limites usuais  $-2$  e  $2$  usados para detecção de pontos aberrantes nos gráficos de resíduos contra elementos do modelo (índice das observações, valores preditos, valores das covariadas) não são adequados. Neste sentido, propomos aqui utilizar como limites de detecção de pontos aberrantes, os quantis empíricos dos resíduos gerados com base em suas distribuições estimadas por processo de reamostragem para a construção das bandas do envelope dos gráficos normais de probabilidade, para mais detalhes, recomendamos (ESPINHEIRA; SANTOS; CRIBARI-NETO, 2017).

Tabela 17 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 20$ ,  $\lambda = 530$  e  $\mu \in (0.0167; 0.2055)$ .

| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | -0.046    | -0.049      | -0.047             | 1.159          | 1.217       | 1.159              | 0.059      | 0.066       | 0.058              | 1.803     | 1.801       | 1.804              |
| 2  | -0.038    | -0.037      | -0.037             | 0.934          | 0.955       | 0.934              | -0.054     | -0.049      | -0.070             | 2.177     | 2.173       | 2.179              |
| 3  | 0.025     | 0.031       | 0.025              | 1.162          | 1.405       | 1.163              | -0.061     | -0.059      | -0.062             | 2.250     | 2.095       | 2.250              |
| 4  | -0.000    | -0.001      | 0.000              | 0.959          | 0.988       | 0.966              | -0.396     | -0.375      | -0.444             | 2.517     | 2.532       | 2.562              |
| 5  | 0.140     | 0.143       | 0.145              | 0.669          | 0.700       | 0.651              | -0.619     | -0.611      | -0.731             | 3.229     | 3.452       | 3.551              |
| 6  | -0.127    | -0.129      | -0.127             | 1.126          | 1.168       | 1.126              | 0.011      | 0.028       | 0.009              | 2.308     | 2.274       | 2.308              |
| 7  | -0.166    | -0.168      | -0.166             | 0.978          | 0.993       | 0.977              | 0.195      | 0.200       | 0.188              | 2.317     | 2.321       | 2.316              |
| 8  | -0.030    | -0.031      | -0.030             | 1.105          | 1.200       | 1.105              | -0.034     | -0.020      | -0.034             | 2.328     | 2.221       | 2.328              |
| 9  | -0.041    | -0.050      | -0.041             | 1.223          | 1.343       | 1.223              | 0.009      | 0.016       | 0.008              | 1.829     | 1.795       | 1.828              |
| 10 | 0.092     | 0.149       | 0.092              | 0.864          | 1.305       | 0.864              | -0.102     | -0.102      | -0.102             | 2.910     | 2.291       | 2.910              |
| 11 | 0.152     | 0.157       | 0.167              | 0.631          | 0.6574      | 0.598              | -0.856     | -0.824      | -1.074             | 3.650     | 3.671       | 4.446              |
| 12 | -0.012    | -0.016      | -0.013             | 0.812          | 1.289       | 0.814              | -0.106     | -0.005      | -0.119             | 2.986     | 2.303       | 3.017              |
| 13 | -0.019    | -0.019      | -0.019             | 1.060          | 1.122       | 1.062              | -0.157     | -0.132      | -0.163             | 2.332     | 2.310       | 2.335              |
| 14 | -0.053    | -0.063      | -0.054             | 1.046          | 1.233       | 1.048              | -0.051     | -0.032      | -0.055             | 2.496     | 2.284       | 2.499              |
| 15 | 0.034     | 0.052       | 0.034              | 0.940          | 1.345       | 0.940              | -0.047     | -0.047      | -0.047             | 2.385     | 2.059       | 2.386              |
| 16 | -0.240    | -0.241      | -0.240             | 1.171          | 1.180       | 1.172              | -0.017     | -0.014      | -0.023             | 2.252     | 2.248       | 2.252              |
| 17 | -0.037    | -0.041      | -0.037             | 1.043          | 1.246       | 1.044              | -0.096     | -0.075      | -0.099             | 2.452     | 2.262       | 2.453              |
| 18 | -0.061    | -0.063      | -0.061             | 0.997          | 1.022       | 1.001              | -0.173     | -0.166      | -0.198             | 2.245     | 2.224       | 2.259              |
| 19 | 0.007     | -0.007      | 0.007              | 0.599          | 1.206       | 0.600              | 0.059      | -0.020      | 0.058              | 4.418     | 2.376       | 4.420              |
| 20 | 0.058     | 0.169       | 0.058              | 0.552          | 1.192       | 0.552              | -0.116     | -0.106      | -0.118             | 4.786     | 2.560       | 4.791              |

Fonte: O autor (2020).

Tabela 18 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 20$ ,  $\lambda = 530$  e  $\mu \in (0.2048; 0.8712)$ .

| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | -0.017    | -0.012      | -0.010             | 1.180          | 1.269       | 1.185              | -0.170     | -0.143      | -0.189             | 1.882     | 1.886       | 1.877              |
| 2  | 0.011     | 0.015       | 0.014              | 0.902          | 0.940       | 0.910              | -0.493     | -0.440      | -0.544             | 2.704     | 2.696       | 2.779              |
| 3  | -0.001    | 0.000       | -0.001             | 1.188          | 1.404       | 1.189              | -0.121     | -0.099      | -0.122             | 2.350     | 2.148       | 2.350              |
| 4  | 0.025     | 0.026       | 0.035              | 0.887          | 0.916       | 0.894              | -0.556     | -0.506      | -0.628             | 3.023     | 3.054       | 3.117              |
| 5  | 0.104     | 0.105       | 0.160              | 0.720          | 0.755       | 0.715              | -0.961     | -0.882      | -1.207             | 4.277     | 4.345       | 4.801              |
| 6  | -0.039    | -0.035      | -0.039             | 1.041          | 1.136       | 1.042              | -0.352     | -0.275      | -0.356             | 2.881     | 2.700       | 2.885              |
| 7  | 0.007     | 0.006       | 0.030              | 1.018          | 1.047       | 1.027              | -0.558     | -0.519      | -0.638             | 2.863     | 2.828       | 2.912              |
| 8  | -0.010    | -0.013      | -0.008             | 0.994          | 1.145       | 0.995              | -0.175     | -0.129      | -0.181             | 2.563     | 2.263       | 2.562              |
| 9  | -0.014    | -0.015      | -0.014             | 1.130          | 1.305       | 1.132              | -0.044     | -0.033      | -0.046             | 1.990     | 1.950       | 1.990              |
| 10 | 0.023     | 0.046       | 0.023              | 0.812          | 1.306       | 0.812              | -0.115     | -0.045      | -0.116             | 3.730     | 2.464       | 3.731              |
| 11 | 0.068     | 0.070       | 0.149              | 0.717          | 0.740       | 0.705              | -1.192     | -1.104      | -1.530             | 4.798     | 4.705       | 5.829              |
| 12 | 0.006     | 0.003       | 0.006              | 0.927          | 1.239       | 0.929              | -0.068     | -0.042      | -0.074             | 2.356     | 2.243       | 2.360              |
| 13 | -0.018    | -0.014      | -0.017             | 0.996          | 1.044       | 1.001              | -0.181     | -0.158      | -0.192             | 2.436     | 2.433       | 2.439              |
| 14 | -0.036    | -0.034      | -0.036             | 1.087          | 1.223       | 1.090              | -0.105     | -0.076      | -0.110             | 2.441     | 2.319       | 2.442              |
| 15 | 0.006     | 0.016       | 0.007              | 0.894          | 1.309       | 0.894              | -0.070     | -0.023      | -0.071             | 2.706     | 2.201       | 2.707              |
| 16 | 0.008     | 0.008       | 0.012              | 1.101          | 1.113       | 1.105              | -0.528     | -0.519      | -0.562             | 2.858     | 2.843       | 2.894              |
| 17 | -0.013    | -0.022      | -0.012             | 1.035          | 1.207       | 1.038              | -0.201     | -0.165      | -0.207             | 2.729     | 2.520       | 2.731              |
| 18 | -0.026    | -0.030      | -0.021             | 0.941          | 0.985       | 0.949              | -0.403     | -0.374      | -0.443             | 2.494     | 2.462       | 2.533              |
| 19 | -0.014    | -0.028      | -0.014             | 0.548          | 1.218       | 0.549              | -0.254     | -0.026      | -0.255             | 6.625     | 2.602       | 6.626              |
| 20 | -0.007    | -0.004      | -0.007             | 0.506          | 1.189       | 0.507              | -0.526     | 0.033       | -0.541             | 8.362     | 2.858       | 8.449              |

Fonte: O autor (2020).

Tabela 19 – Médias, desvios-padrão, assimetrias e curtoses empíricas dos resíduos ponderado ( $r^\beta$ ), ponderado padronizado ( $r_p^\beta$ ) e combinado ( $r_p^{\beta\gamma}$ ). Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 20$ ,  $\lambda = 530$  e  $\mu \in (0.8172; 0.9908)$ .

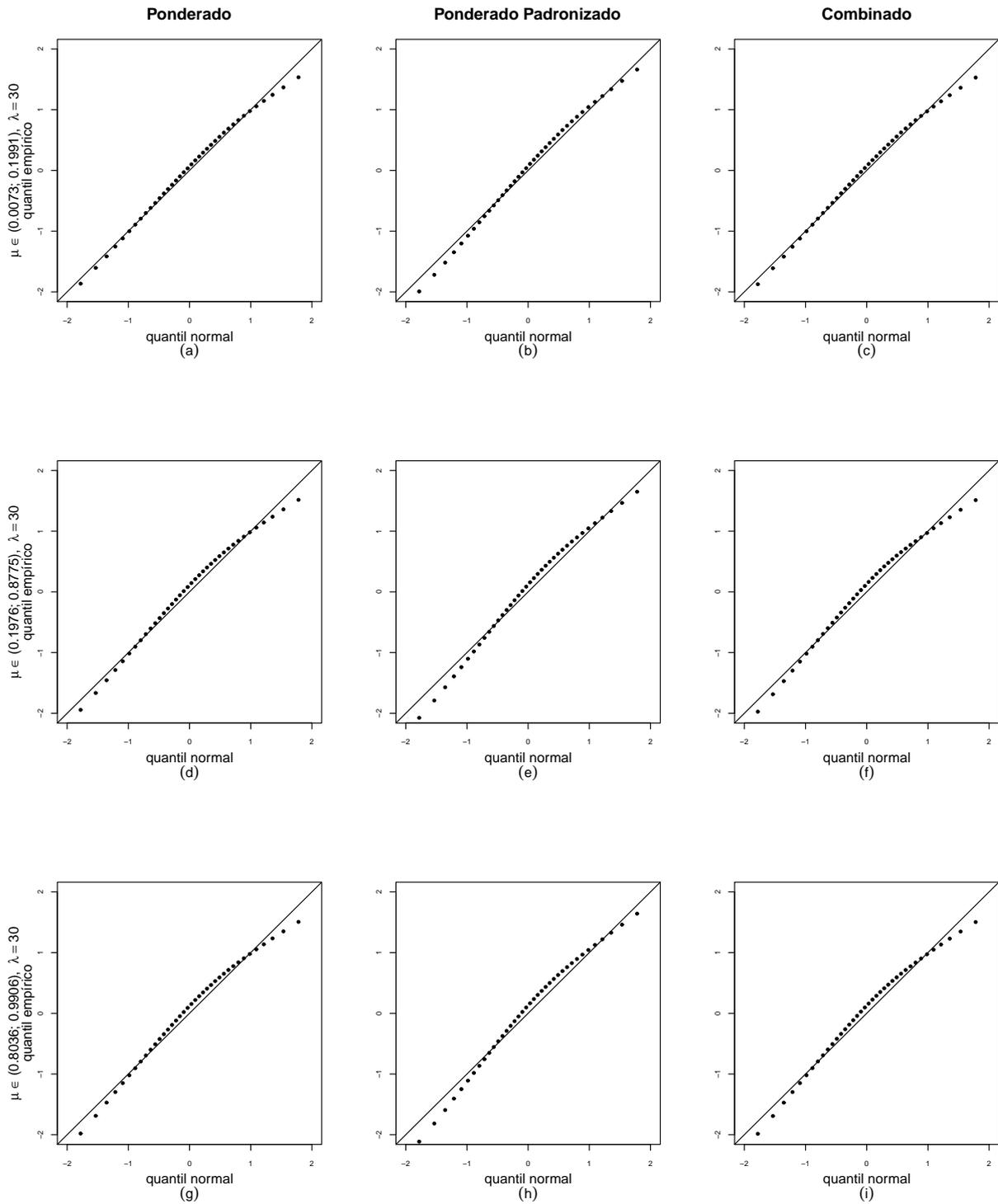
| t  | Média     |             |                    | Desvios-Padrão |             |                    | Assimetria |             |                    | Curtose   |             |                    |
|----|-----------|-------------|--------------------|----------------|-------------|--------------------|------------|-------------|--------------------|-----------|-------------|--------------------|
|    | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$      | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$  | $r_p^\beta$ | $r^{\beta,\gamma}$ | $r^\beta$ | $r_p^\beta$ | $r^{\beta,\gamma}$ |
| 1  | -0.017    | -0.015      | -0.011             | 1.201          | 1.277       | 1.203              | -0.208     | -0.179      | -0.221             | 1.949     | 1.913       | 1.944              |
| 2  | 0.014     | 0.015       | 0.015              | 0.856          | 0.879       | 0.856              | -0.604     | -0.573      | -0.614             | 3.138     | 3.105       | 3.155              |
| 3  | 0.004     | 0.010       | 0.004              | 1.100          | 1.342       | 1.100              | -0.154     | -0.124      | -0.154             | 2.712     | 2.415       | 2.712              |
| 4  | -0.001    | -0.002      | 0.002              | 0.839          | 0.863       | 0.839              | -0.634     | -0.605      | -0.651             | 3.457     | 3.451       | 3.480              |
| 5  | 0.033     | 0.035       | 0.137              | 0.751          | 0.778       | 0.776              | -1.063     | -0.960      | -1.258             | 4.986     | 4.892       | 5.025              |
| 6  | -0.013    | -0.009      | -0.013             | 1.011          | 1.077       | 1.011              | -0.361     | -0.305      | -0.363             | 3.016     | 2.907       | 3.016              |
| 7  | -0.004    | -0.004      | 0.008              | 0.977          | 0.997       | 0.981              | -0.616     | -0.590      | -0.642             | 3.099     | 3.050       | 3.089              |
| 8  | -0.006    | -0.007      | -0.006             | 0.956          | 1.072       | 0.956              | -0.204     | -0.159      | -0.205             | 2.718     | 2.472       | 2.718              |
| 9  | -0.032    | -0.039      | -0.032             | 1.140          | 1.274       | 1.140              | -0.050     | -0.039      | -0.050             | 2.104     | 2.075       | 2.104              |
| 10 | 0.006     | 0.008       | 0.006              | 0.757          | 1.237       | 0.757              | -0.034     | -0.020      | -0.034             | 4.132     | 2.793       | 4.132              |
| 11 | 0.000     | 0.000       | 0.073              | 0.729          | 0.746       | 0.738              | -1.254     | -1.168      | -1.419             | 5.315     | 5.270       | 5.602              |
| 12 | 0.007     | 0.012       | 0.008              | 0.811          | 1.190       | 0.812              | -0.141     | -0.068      | -0.146             | 2.947     | 2.600       | 2.947              |
| 13 | -0.011    | -0.008      | -0.011             | 0.925          | 0.985       | 0.925              | -0.365     | -0.306      | -0.368             | 2.900     | 2.826       | 2.901              |
| 14 | -0.005    | 0.001       | -0.004             | 0.992          | 1.160       | 0.992              | -0.183     | -0.108      | -0.185             | 2.851     | 2.654       | 2.851              |
| 15 | -0.023    | -0.022      | -0.023             | 0.895          | 1.270       | 0.895              | -0.079     | 0.009       | -0.079             | 2.825     | 2.395       | 2.825              |
| 16 | 0.005     | 0.006       | 0.006              | 1.027          | 1.036       | 1.027              | -0.581     | -0.572      | -0.584             | 3.190     | 3.178       | 3.193              |
| 17 | 0.002     | 0.003       | 0.002              | 0.979          | 1.134       | 0.979              | -0.257     | -0.178      | -0.258             | 3.126     | 2.810       | 3.126              |
| 18 | -0.030    | -0.032      | -0.029             | 0.886          | 0.916       | 0.887              | -0.485     | -0.456      | -0.496             | 2.803     | 2.773       | 2.814              |
| 19 | 0.006     | 0.016       | 0.006              | 0.551          | 1.168       | 0.551              | -0.148     | -0.067      | -0.148             | 6.921     | 2.869       | 6.921              |
| 20 | -0.011    | -0.015      | -0.011             | 0.454          | 1.143       | 0.454              | -0.638     | 0.034       | -0.639             | 8.846     | 3.104       | 8.848              |

Fonte: O autor (2020).

De forma geral, podemos afirmar com base nas Figuras 10-15 que o tamanho da amostra ( $n = 40$  e  $n = 120$ ), os valores atribuídos ao grau de precisão ( $\lambda = 30$ ,  $\lambda = 100$  e  $\lambda = 530$ ) e considerando os três cenários relativos a média da variável resposta, notamos que não ocorre grandes alterações ao compararmos as distribuições dos resíduos ponderado, ponderado padronizado e combinado.

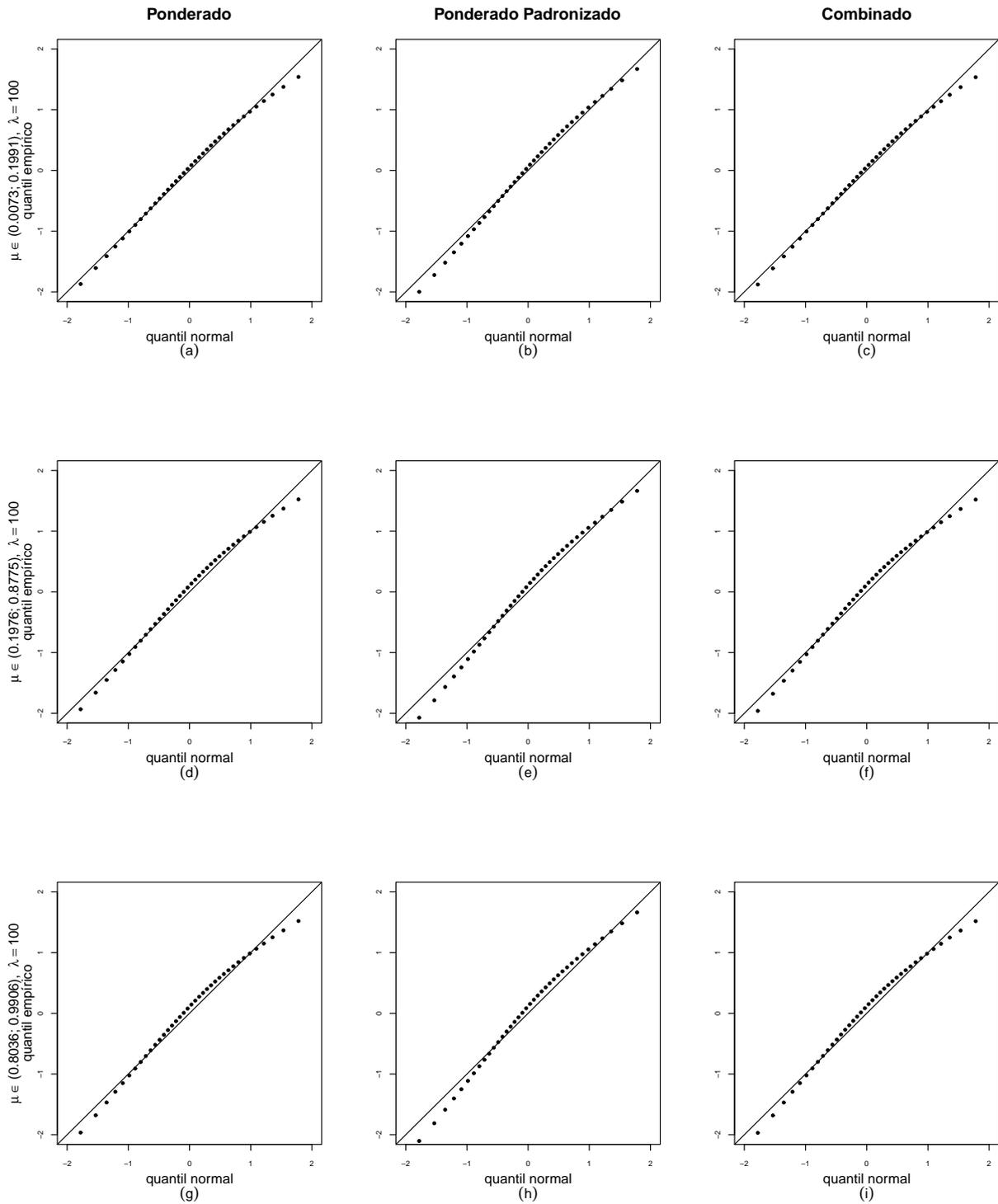
Inicialmente, neste capítulo avaliamos o comportamento dos resíduos estudados no Capítulo 2 do modelo de regressão gama unitária com precisão fixo. É notável que para valores entre 1.5 e 10 as distribuições dos resíduos apresentam comportamentos eventualmente diferentes, com alguma assimetria. No entanto, a partir de  $\phi = 30$  as distribuições empíricas dos resíduos são similares e se aproximam melhor da distribuição normal padrão. No geral, quando o parâmetro de precisão é variável, independentemente do valor atribuído a ele, temos uma melhor adequação da distribuição empírica dos resíduos a distribuição normal padrão. Contudo, é mais adequado utilizarmos os quantis empíricos gerados a partir da construção do envelope, para a construção dos gráficos de resíduos contra elementos do modelo ((ESPINHEIRA; SANTOS; CRIBARI-NETO, 2017)), ou seja, não usaremos os tradicionais limites  $-2$  e  $2$  para detecção de pontos aberrantes nas nossas aplicações.

Figura 10 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 40$ ,  $\lambda = 30$ .



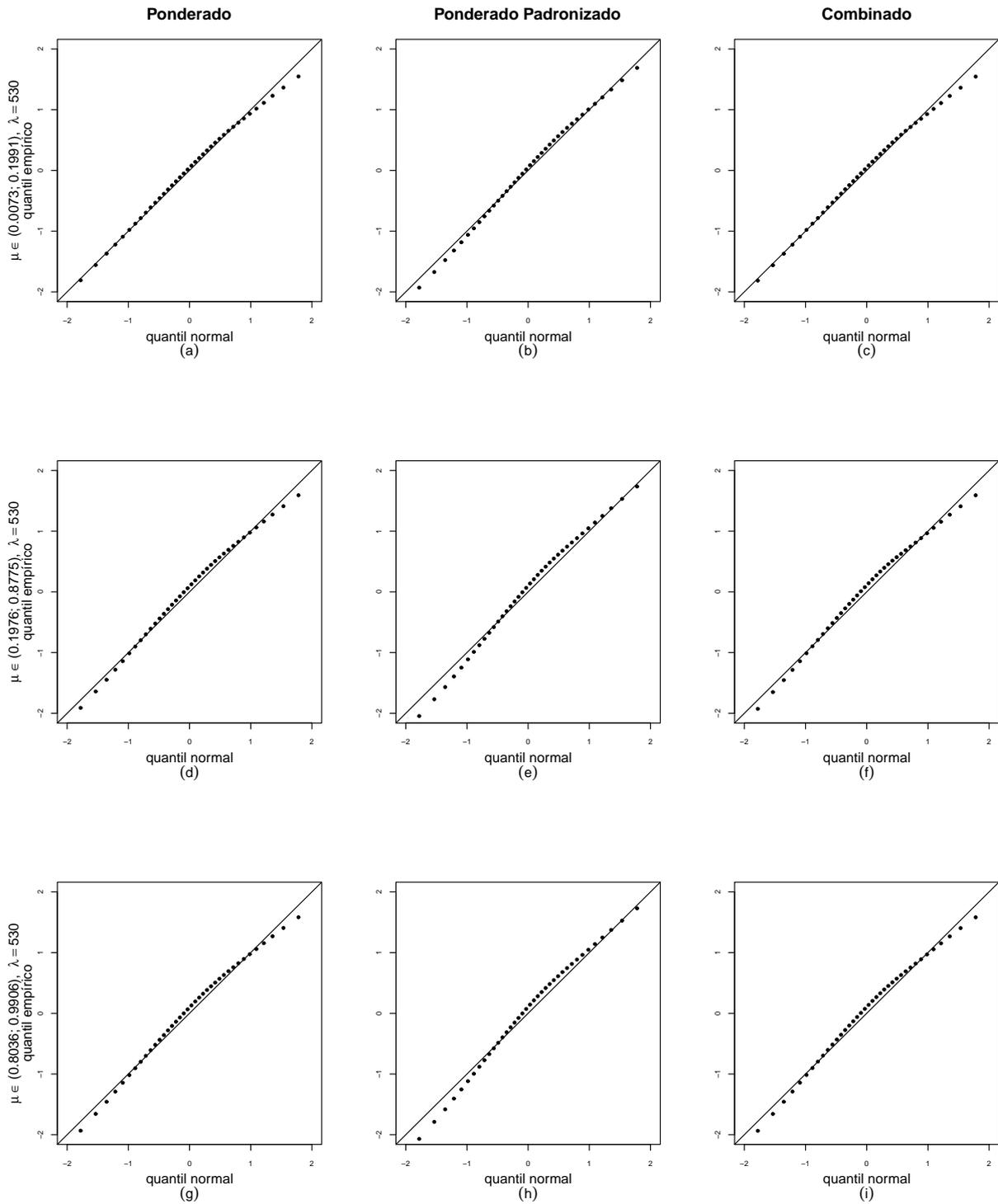
Fonte: O autor (2020).

Figura 11 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 40$ ,  $\lambda = 100$ .



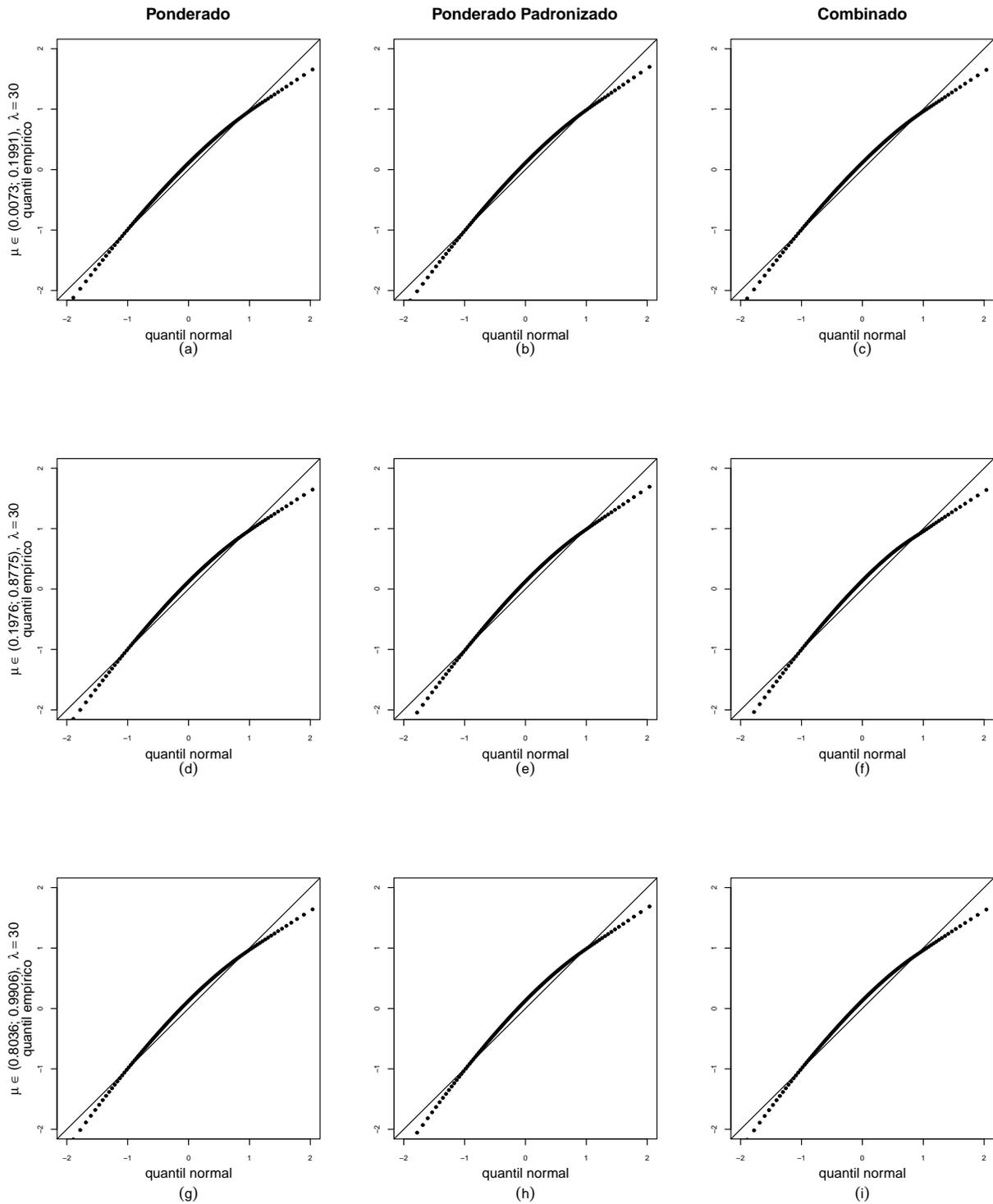
Fonte: O autor (2020).

Figura 12 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 40$ ,  $\lambda = 530$ .



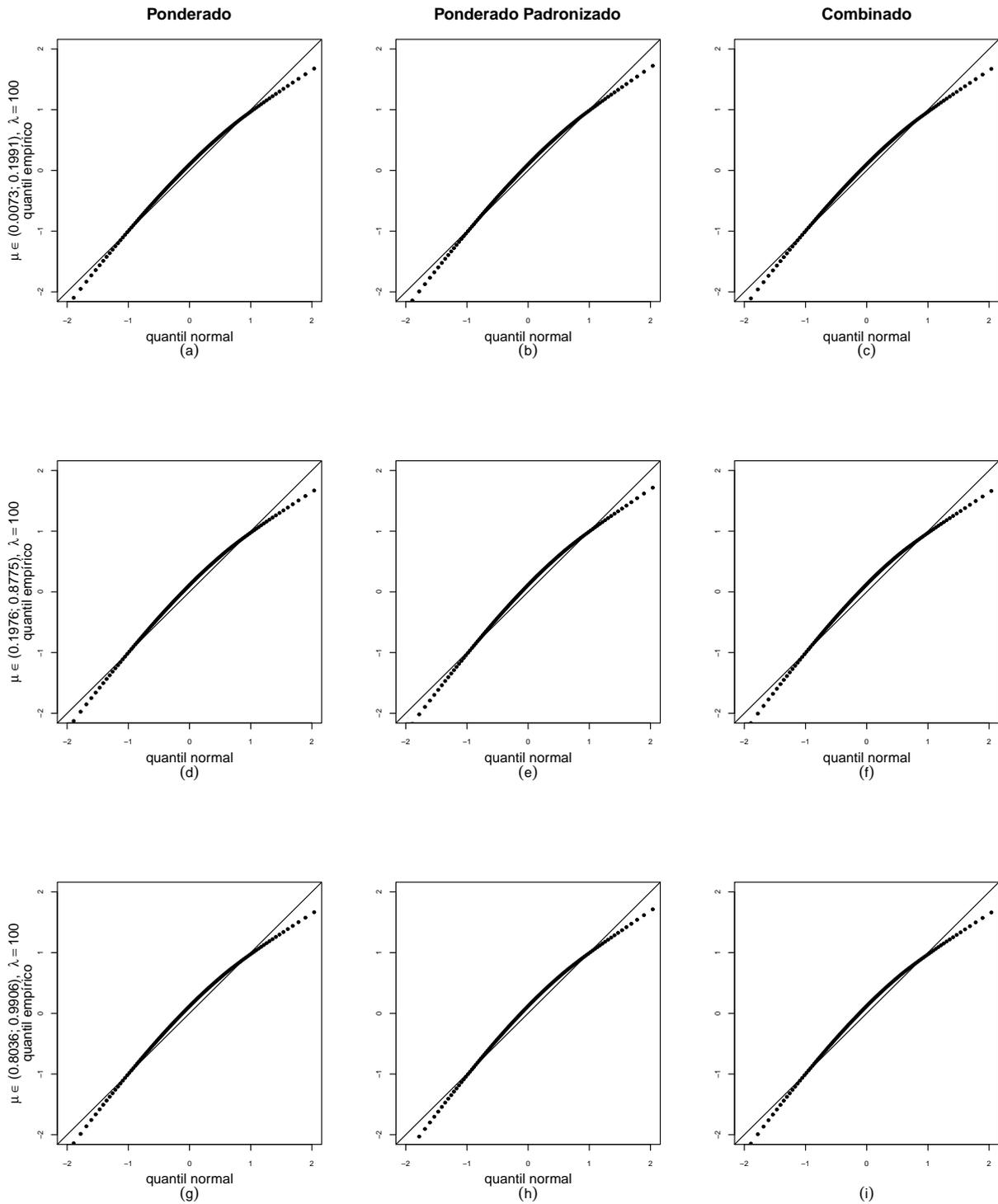
Fonte: O autor (2020).

Figura 13 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 120$ ,  $\lambda = 30$ .



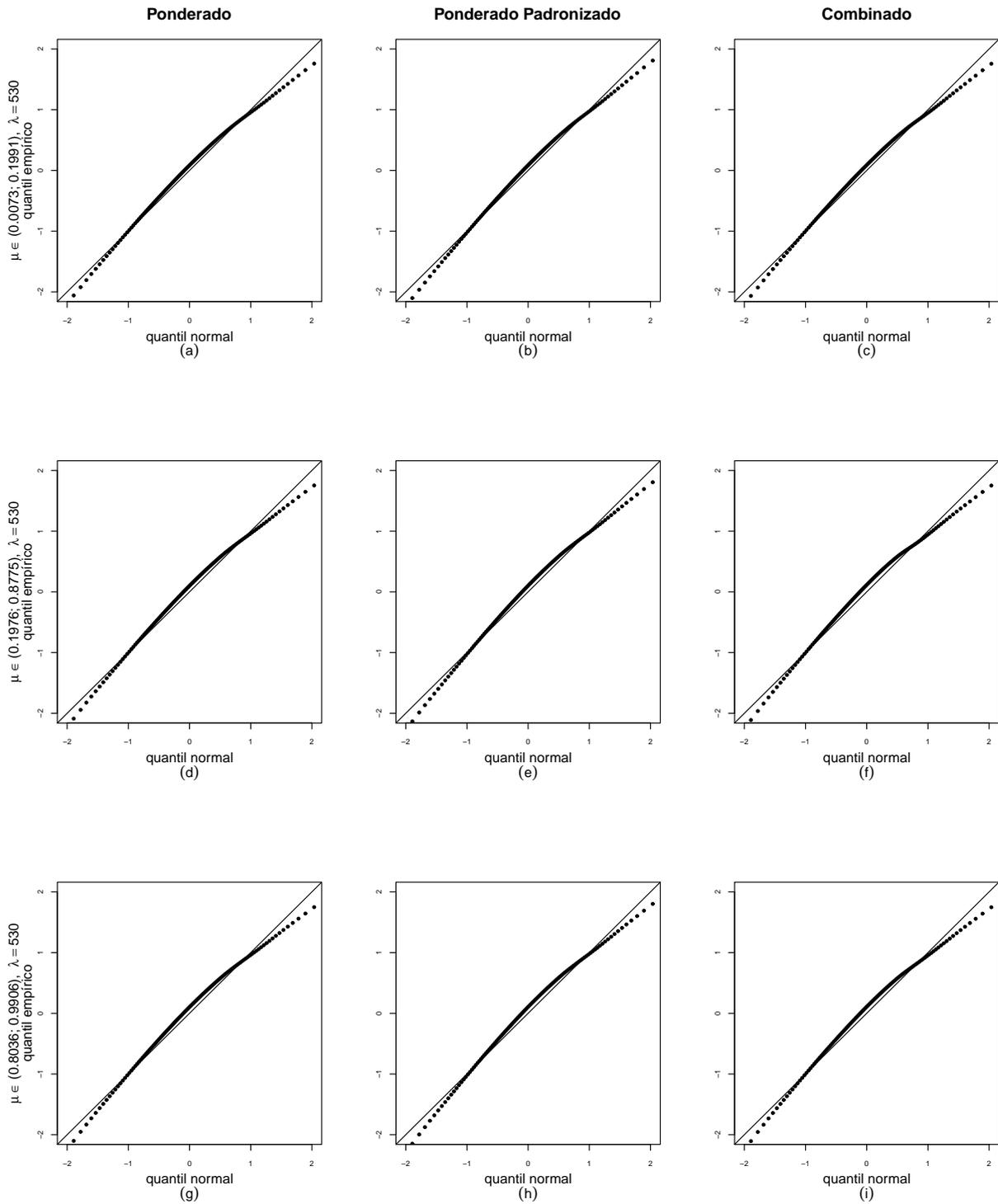
Fonte: O autor (2020).

Figura 14 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 120$ ,  $\lambda = 100$ .



Fonte: O autor (2020).

Figura 15 – Gráficos normais de probabilidade; Modelo:  $\log[\mu_t/(1 - \mu_t)] = \beta_1 + \beta_2x_{t2} + \beta_3x_{t3} + \beta_4x_{t4} + \beta_5x_{t5}$  e  $\log(\phi_t) = \gamma_1 + \gamma_2z_{t2} + \gamma_3z_{t3} + \gamma_4z_{t4} + \gamma_5z_{t5}$ ,  $t = 1, \dots, 120$ ,  $\lambda = 530$ .



Fonte: O autor (2020).

## 4 INFLUÊNCIA LOCAL

Para um certo conjunto de dados observados, seja  $\ell(\theta)$  o logaritmo da função de verossimilhança correspondente ao modelo postulado, em que  $\theta$  é um vetor  $s \times 1$  desconhecido; Considere agora uma perturbação neste modelo postulado, realizada através de um vetor  $\delta$ , em geral de dimensão  $n \times 1$ , o qual é restrito a algum subconjunto aberto  $\mathcal{D}$  de  $\mathbb{R}^n$ . Seja  $\ell_\delta(\theta)$  o logaritmo da função de verossimilhança do modelo perturbado para um dado  $\delta \in \mathcal{D}$ .

Sejam  $\hat{\theta}$  e  $\hat{\theta}_\delta$  os estimadores de máxima verossimilhança de  $\theta$  para o modelo postulado e para o modelo perturbado, respectivamente. O deslocamento pela verossimilhança, que neste caso mais geral é definido por

$$LD_\delta = 2\{\ell(\hat{\theta}) - \ell(\hat{\theta}_\delta)\}, \quad (4.1)$$

pode ser utilizado como uma medida para avaliar a influência sobre a estimativa de  $\theta$  ao se variar  $\delta$  através de  $\mathcal{D}$ .

No entanto, avaliar o comportamento de  $LD_\delta$  para todo  $\delta \in \mathcal{D}$  pode ser inviável. Neste sentido, (COOK, 1986) propôs estudar o comportamento local de  $LD_\delta$  em uma vizinhança de  $\delta_0$ , de modo que devemos escolher  $\delta$  pertencente a vizinhança de  $\delta_0$ , onde  $\delta_0$  representa o vetor de ausência de perturbação do modelo postulado, de tal forma que  $LD_{\delta_0} = 0$ . Pela forma como definimos (4.1), temos que  $LD_\delta \geq 0$ .

Uma vez que tomamos qualquer vetor  $\delta$  que pertença à vizinhança de  $\delta_0$ , isso irá representar um incremento em  $LD_{\delta_0}$ , este procedimento descreve a sensibilidade de  $\ell(\hat{\theta})$  com respeito a uma leve perturbação introduzida em  $\ell(\theta)$ . Para estudar esta sensibilidade, (COOK, 1986) propôs trabalhar com curvaturas normais.

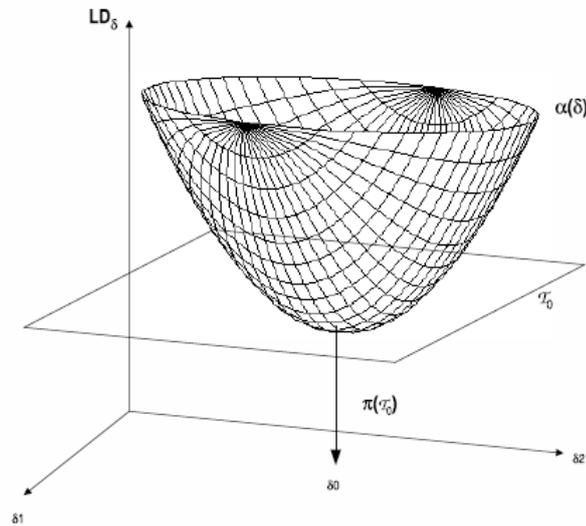
O procedimento proposto por Cook, consiste em avaliar como a superfície geométrica  $\alpha(\delta) = (\delta, LD_\delta)^\top$  desvia-se de seu plano tangente em  $\delta_0$  à medida que  $\delta$  se afasta levemente de  $\delta_0$ , ou seja, quando pequenas perturbações são introduzidas no modelo. Tal descrição pode ser obtida pelo estudo das curvaturas das seções normais da superfície  $\alpha(\delta)$  em  $\delta_0$ , que são chamadas de curvaturas normais.

Para definir o que é uma seção normal da superfície  $\alpha(\delta)$  em  $\delta_0$ , consideramos o plano tangente a  $\alpha(\delta)$  em  $\delta_0$ , denotado por  $\mathcal{T}_0$ , e o vetor que é ortogonal a  $\mathcal{T}_0$ , que denotamos por  $\pi(\mathcal{T}_0)$ .

Na Figura 16 apresentamos a curva  $\alpha(\delta) = (\delta, LD_\delta)^\top$ , o plano tangente a esta curva em  $\delta_0$  e o vetor  $\pi(\mathcal{T}_0)$ . Para esta ilustração, estamos admitindo que  $\delta = (\delta_1, \delta_2)^\top$ , assim  $\delta_0 = (\delta_{01}, \delta_{02})^\top$ .

Os planos que contêm o vetor normal  $\pi(\mathcal{T}_0)$  são chamados de planos normais a  $\alpha(\delta)$  em  $\delta_0$ . As seções normais em  $\delta_0$  são definidas como as intersecções desses planos normais com a superfície  $\alpha(\delta)$ .

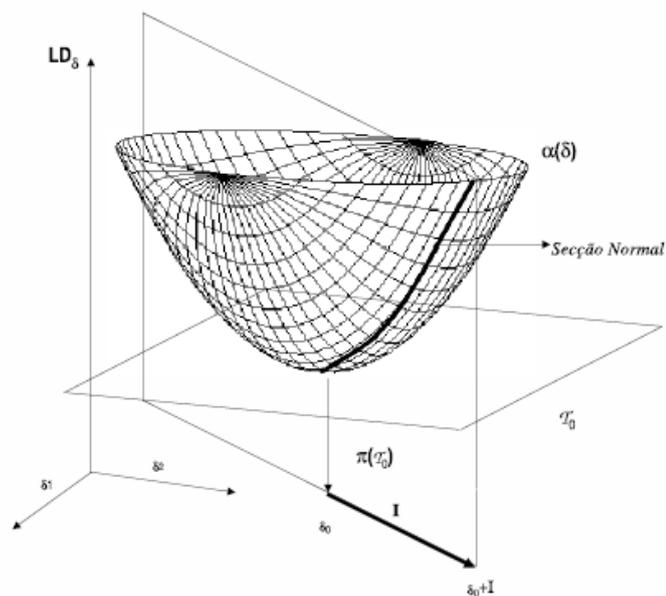
Figura 16 – A curva  $\alpha(\delta) = (\delta, LD_\delta)^\top$ , o plano tangente  $\mathcal{T}_0$  e o vetor  $\pi(\mathcal{T}_0)$ .



**Fonte:** O autor (2020).

Consideremos agora um deslocamento ao redor de  $\delta_0$ , ou seja, introduziremos uma pequena perturbação no modelo. Escolhemos uma direção  $I$ , de modo que  $\|I\| = 1$  e tomamos  $\delta = \delta_0 + I$ . Dessa forma, teremos o ponto  $LD_{\delta_0+I} > 0$ . Por esse ponto passa o plano que contém os vetores  $I$  e  $\pi(\mathcal{T}_0)$  e a intersecção desse plano com  $\alpha(\delta)$  forma uma seção normal, como podemos observar na Figura 17.

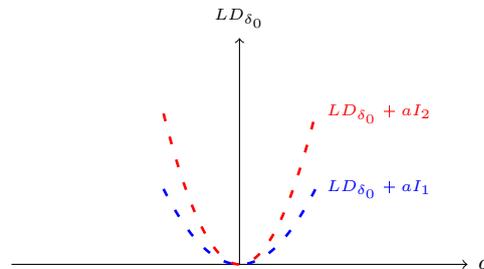
Figura 17 – Seção normal.



**Fonte:** O autor (2020).

De fato, cada direção  $I$  escolhida, de modo que  $\|I\| = 1$ , associada ao vetor  $\pi(\mathcal{T}_0)$ , especifica uma seção normal e define um valor de curvatura normal. Cada seção normal é equivalente ao gráfico  $LD_{\delta_0+aI} \times a$ , que  $a \in \mathbb{R}$ . Esse gráfico é chamado de linha projetada.

Figura 18 – Linhas projetadas.



Fonte: O autor (2020).

Na Figura 18 ilustramos o comportamento de duas linhas projetadas  $LD_{\delta_0+aI_1} \times a$  e  $LD_{\delta_0+aI_2} \times a$ , relacionadas, respectivamente, às direções  $I_1$  e  $I_2$ . Uma vez que  $LD_{\delta_0} = 0$ , segue que  $LD_{\delta_0+aI}$  tem mínimo local em  $a = 0$ . É possível notar que a direção  $I_2$  é a que provoca os maiores deslocamentos pela verossimilhança, à medida que os valores de  $a$  se afastam de zero.

A ideia de (COOK, 1986) é analisar a direção  $I$ , tal que a linha projetada resultante apresente a maior curvatura normal ( $C_{\max}$ ). Segundo (COOK, 1986), este vetor, denotado por  $I_{\max}$ , indica que valores atribuir aos componentes de  $\delta$  para obter a maior mudança local no deslocamento pela verossimilhança. Ele mostra ainda que a curvatura normal na direção  $I$ , em que estamos interessados, fica resumida ao módulo da segunda derivada de  $LD_{\delta_0+aI}$  com respeito a  $a$ , que avaliada em  $a = 0$  é dada por

$$C_I = |I^\top \ddot{L}D_\delta I|,$$

em que

$$\ddot{L}D_\delta = \frac{\partial^2 LD_\delta}{\partial \delta \partial \delta^\top}$$

é avaliada em  $\delta = \delta_0$  e  $\theta = \hat{\theta}$ . Note que  $\ddot{L}D_\delta(\cdot)$  é uma matriz simétrica e positiva definida. Pela definição do deslocamento pela verossimilhança, segue que

$$\ddot{L}D_\delta = -2\Delta^\top \ddot{\ell}^{-1} \Delta,$$

em que

$$\ddot{\ell} = \frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^\top} \Big|_{\theta=\hat{\theta}} \quad \text{e} \quad \Delta = \frac{\partial^2 \ell_\delta(\theta)}{\partial \theta \partial \delta^\top} \Big|_{\theta=\hat{\theta}},$$

é avaliada em  $\delta = \delta_0$ . Ressaltamos que estamos considerando  $\delta = (\delta_1, \dots, \delta_n)^\top$  e, dado que  $\theta$  é um vetor  $s \times 1$ , temos que  $\Delta$  é uma matriz de ordem  $s \times n$ .

Finalmente, chegamos à expressão da curvatura normal:

$$C_I(\theta) = 2|I^\top \Delta^\top \ddot{\ell}^{-1} \Delta I|,$$

que também pode ser expressa como

$$C_I = 2I^\top (-\ddot{\mathcal{F}})I,$$

em que  $\ddot{\mathcal{F}} = \Delta^\top \ddot{\ell}^{-1} \Delta$ . É possível também avaliar a influência local apenas para parte do vetor de parâmetros. Suponha que particionamos o vetor de parâmetros como  $\theta = (\theta_1^\top, \theta_2^\top)^\top$ . Seja

$$\ddot{\ell} = \begin{pmatrix} \ddot{\ell}_{\theta_1\theta_1} & \ddot{\ell}_{\theta_1\theta_2} \\ \ddot{\ell}_{\theta_2\theta_1} & \ddot{\ell}_{\theta_2\theta_2} \end{pmatrix},$$

em que  $\ddot{\ell}_{\theta_1\theta_1} = \partial^2 \ell(\theta) / \partial \theta_1 \partial \theta_1^\top$ ,  $\ddot{\ell}_{\theta_1\theta_2} = \partial^2 \ell(\theta) / \partial \theta_1 \partial \theta_2^\top$ ,  $\ddot{\ell}_{\theta_2\theta_1} = \partial^2 \ell(\theta) / \partial \theta_2 \partial \theta_1^\top$  e  $\ddot{\ell}_{\theta_2\theta_2} = \partial^2 \ell(\theta) / \partial \theta_2 \partial \theta_2^\top$ .

Segundo (COOK, 1986), se o interesse é calcular a influência local apenas para  $\theta_1$ , o afastamento pela verossimilhança é dado por

$$LD_{\delta:\theta_1} = 2 \left\{ \ell(\hat{\theta}_1, \hat{\theta}_2) - \ell(\hat{\theta}_{1\delta}, \mathcal{J}(\theta_{1\delta})) \right\},$$

em que  $\mathcal{J}(\theta_{1\delta})$  é uma função que maximiza  $\ell(\theta_1, \theta_2)$  para cada  $\theta_1$  fixo,  $\hat{\theta}_{1\delta}$  é obtido da partição  $(\hat{\theta}_{1\delta}, \hat{\theta}_{2\delta})^\top$  e  $\ell(\theta_{1\delta}, \mathcal{J}(\theta_{1\delta}))$  é o logaritmo da função de verossimilhança perfilada para  $\theta_1$ . Neste caso, a curvatura normal na direção do vetor  $I$  é dada por

$$C_{I:\theta_1} = |I^\top \Delta^\top (\ddot{\ell}^{-1} - \ddot{\ell}_{22}) \Delta I|,$$

em que

$$\ddot{\ell}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & \ddot{\ell}_{\theta_2\theta_2}^{-1} \end{pmatrix},$$

e  $I_{\max:\theta_1}$  é o autovetor de norma igual a um correspondente ao maior autovalor da matriz  $-\Delta^\top (\ddot{\ell}^{-1} - \ddot{\ell}_{22}) \Delta$ . O mesmo procedimento pode ser utilizado para avaliar a influência local apenas para  $\theta_2$ .

Uma vez que o vetor  $I_{\max}$  mostra como perturbar o modelo postulado (o modelo para  $\delta = \delta_0$ ) no sentido de obter a maior mudança local no deslocamento pela verossimilhança, (COOK, 1986) propõe inspecionar  $I_{\max}$  independentemente da magnitude de  $C_{\max}$ , porque isto pode revelar observações que são simultaneamente influentes.

No entanto, podem existir casos individualmente influentes e que são revelados através da análise dos componentes de  $I_{\max}$ . Neste sentido, os autores (LESAFFRE; VERBEKE, 1998) sugerem avaliar a curvatura na direção do  $i$ -ésimo indivíduo, isto é, o vetor em que o  $i$ -ésimo componente assume o valor um e os demais o valor zero. Neste caso, a curvatura

normal é dada por

$$C_i = 2|\Delta_i^\top \ddot{\ell}^{-1} \Delta_i|,$$

em que  $\Delta_i$  é a  $i$ -ésima coluna de matriz  $\Delta$ .

Uma vez que atribuímos a  $C_i$  o maior valor (o valor total) possível da  $i$ -ésima coordenada de  $\|I\| = 1$ , estes autores denotam tal curvatura por influência local do  $i$ -ésimo indivíduo. Também é possível calcular a influência local total do  $i$ -ésimo indivíduo na estimação de parte do vetor  $\theta$ . Por exemplo, se o interesse recai em  $\theta_1$  temos que

$$C_{i:\theta_1} = 2|\Delta_i^\top (\ddot{\ell}^{-1} - \ddot{\ell}_{22}) \Delta_i|,$$

(LESAFFRE; VERBEKE, 1998) argumentam que os componentes de  $I_{\max}$  e as medidas de influência  $C_i$  com  $I = 1, \dots, n$ , contêm informações diferentes. Os autores mostram que podemos expressar  $C_i$  como

$$C_i = 2 \sum_{p=1}^s \zeta_p v_{pi}^2,$$

em que  $\zeta_p$  com  $p = 1, \dots, s$ , são os autovalores diferentes de zero da matriz  $-\Delta^\top \ell^{-1} \Delta$ , tais que  $\zeta_1 \geq \zeta_2 \geq \zeta_3 \geq \dots \geq \zeta_s \geq \zeta_{s+1} = \dots = \zeta_n = 0$  e  $v_p = (v_{p1}, \dots, v_{pn})^\top$ ,  $p = 1, \dots, s$ , correspondentes autovetores ortonormais, tal que uma base ortonormal da imagem de  $-\Delta^\top \ell^{-1} \Delta$  é dada pelas colunas da matriz

$$\begin{bmatrix} v_{11}, \dots, v_{s1} \\ \vdots \\ v_{1n}, \dots, v_{sn} \end{bmatrix}$$

Assim,  $C_{\max}/2 = \zeta_1$  e  $I_{\max} = (v_{11}, v_{12}, \dots, v_{1t}, \dots, v_{1n})^\top$ .

Note que podem existir casos individuais com curvatura total  $C_i$  grande sem ter o  $i$ -ésimo componente na direção  $I_{\max}$  de máxima curvatura. Isto ocorrerá com casos com um valor expressivo em qualquer autovetor  $v_p$ , com  $p \neq 1$ , correspondente a um autovalor relativamente grande, digamos  $\zeta_2$  ou  $\zeta_3$ .

Assim, parece razoável considerar em uma análise de influência o estudo dos componentes de  $I_{\max}$  e adicionalmente a medida de influência total referente à  $i$ -ésima observação,  $C_i$ . Gráficos dos componentes de  $I_{\max}$  contra os índices das observações podem sugerir quais observações são conjuntamente influentes, enquanto os gráficos de  $C_i$  contra os índices das observações destacam casos individualmente influentes.

Ressaltamos ainda que (LESAFFRE; VERBEKE, 1998) sugerem utilizar como ponto de corte nos gráficos de  $C_i$ , duas vezes o valor médio desta medida, ou seja, observações com  $C_i$  maior que  $2 \sum_{i=1}^n C_i/n$  podem ser classificadas como influentes.

O método de influência local fica bem definido a partir da obtenção da matriz  $\Delta$ , a qual depende do esquema de perturbação adotado e da matriz de informação observada.

## 4.1 INFLUÊNCIA LOCAL COM PRECISÃO FIXA

Considere o modelo de regressão gama unitária definido em (2.1) e sejam  $s = p + 1$  e  $\theta = (\beta^\top, \phi)^\top$ . Temos que

$$\ddot{\ell} = \ddot{\ell}(\beta, \phi) = \begin{pmatrix} \ddot{\ell}_{\beta\beta} & \ddot{\ell}_{\beta\phi} \\ \ddot{\ell}_{\phi\beta} & \ddot{\ell}_{\phi\phi} \end{pmatrix}, \quad (4.2)$$

em que  $\ddot{\ell}_{\beta\beta} = X^\top QX$ ,  $\ddot{\ell}_{\beta\phi} = (\ddot{\ell}_{\phi\beta})^\top = X^\top T f$  e  $\ddot{\ell}_{\phi\phi} = \text{tr}(G)$ , onde

$$Q = \text{diag}\{q_1, \dots, q_n\} \text{ em que } q_t = \{-v_t + (\mu_t^* + y_t^*) [g''(\mu_t)/g'(\mu_t)]\} [1/g'(\mu_t)]^2, \quad (4.3)$$

$$f = (-c_1, \dots, -c_n)^\top, \quad (4.4)$$

$$T = \text{diag}[g'^{-1}(\mu_1), \dots, g'^{-1}(\mu_n)], \quad (4.5)$$

$$G = \text{diag}(-g_1, \dots, -g_n), \quad (4.6)$$

onde  $\text{tr}(G)$  indica o traço da matriz  $G$ . As expressões de  $\mu_t^*$  e  $y_t^*$  estão apresentadas em (2.3) e  $v_t, c_t$  e  $g_t$ , estão definidas no Apêndice B.

A partir de (4.2) e usando a fórmula da inversa de uma matriz particionada, (RAO, 1973), tem-se

$$\ddot{\ell}^{-1} = \ddot{\ell}^{-1}(\beta, \phi) = \begin{pmatrix} \ddot{\ell}^{\beta\beta} & \ddot{\ell}^{\beta\phi} \\ \ddot{\ell}^{\phi\beta} & \ddot{\ell}^{\phi\phi} \end{pmatrix},$$

em que  $(\ddot{\ell}^{\phi\phi})^{-1} = \text{tr}(G) - f^\top T^\top X (X^\top QX)^{-1} X^\top T f$ ,  $\ddot{\ell}^{\beta\phi} = (\ddot{\ell}^{\phi\beta})^\top = (X^\top QX)^{-1} X^\top T f \ddot{\ell}^{\phi\phi}$  e  $\ddot{\ell}^{\beta\beta} = (X^\top QX)^{-1} [1_k + X^\top T f f^\top T^\top X (X^\top QX)^{-1} \ddot{\ell}^{\phi\phi}]$  e  $1_k$  é a matriz identidade de ordem  $k \times k$ .

### 4.1.1 Esquema de perturbação

#### 4.1.1.1 Ponderação de casos

Considere o seguinte esquema de perturbação:

$$\ell_\delta(\beta, \phi) = \sum_{t=1}^n \delta_t \ell_t(\mu_t, \phi),$$

com  $0 \leq \delta_t \leq 1$ . Para o método de influência local discutido anteriormente, temos que  $\delta_0 = (1, 1, \dots, 1)^\top$  e  $\Delta_t = \partial \ell_t(\hat{\theta}) / \partial \theta$ , com  $t = 1, \dots, n$ . Então, usando o Apêndice B, temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\phi \end{pmatrix} = \begin{pmatrix} X^\top \hat{T} \hat{\mathcal{E}} \\ \hat{u}^\top \end{pmatrix},$$

em que

$$\mathcal{E} = \text{diag}[(\mu_t^* + y_t^*), \dots, (\mu_n^* + y_n^*)] \quad (4.7)$$

e  $T$ ,  $\mu_t^*$ ,  $y_t^*$  e  $u_t$  estão definidos, respectivamente, em (4.5), (2.3) e (2.4).

#### 4.1.1.2 Perturbação da variável resposta

No esquema aditivo de perturbação da resposta em que  $y = (y_1, \dots, y_n)^\top$  é alterado através da adição de um vetor  $\delta$  de pequenas perturbações. Em situações em que cada  $y_t$  apresentam uma variância diferente é comum utilizar um fator de escala para padronizar os componentes de  $\delta$ , por exemplo, a estimativa do desvio padrão de  $y_t$ , de forma que

$$y_t(\delta) = y_t + \delta_t s(y_t),$$

em que  $s(y_t) = \sqrt{\hat{\mu}_t(1/[(2 - \hat{\mu}_t^{1/\hat{\phi}})^{\hat{\phi}}] - \hat{\mu}_t)}$ . Para esse tipo de perturbação,  $\delta_0 = (0, 0, \dots, 0)^\top$  e, do Apêndice B, temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\phi \end{pmatrix} = \begin{pmatrix} X^\top \hat{T} M S_y \\ \hat{b}^\top S_y \end{pmatrix},$$

em que  $M = \text{diag}(m_1, \dots, m_n)$  com

$$m_t = \frac{\alpha_t^2}{y_t \mu_t^{1/\phi + 1}}, \quad (4.8)$$

$S_y = \text{diag}[s(y_1), \dots, s(y_n)]$  e  $\hat{b} = (\hat{b}_1, \dots, \hat{b}_n)$  com

$$b_t = \left[ \frac{-\alpha_t^2 \log(\mu_t)}{y_t \phi^2 \mu_t^{1/\phi}} + \frac{1}{\log(y_t) y_t} \right], \quad (4.9)$$

e lembremos que  $\alpha_t = \mu_t^{1/\phi} / (1 - \mu_t^{1/\phi})$ .

#### 4.1.1.3 Perturbação individual de covariáveis

(THOMAS; COOK, 1989) sugerem modificar a  $k$ -ésima coluna da matriz  $X$ , denotada por  $x_k$ , com  $k = 2, \dots, p$ , adicionando um vetor  $\delta$  de pequenas perturbações ponderado

por um fator de escala  $s_{x_k}$ , desvio-padrão da coluna modificada  $x_k$ , de forma que

$$x_{tk}(\delta) = x_{tk} + \delta_t s_{x_k}.$$

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

$$\eta_t(\delta) = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k (x_{tk} + \delta_t s_{x_k}) + \cdots + \beta_p x_{tp}, \quad (4.10)$$

e  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_t(\delta)$ . Para esse tipo de perturbação, temos  $\delta_0 = (0, 0, \dots, 0)^\top$  e do Apêndice B temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\phi \end{pmatrix} = \begin{bmatrix} -s_{x_k} \left( \hat{\beta}_k X^\top \hat{Q} - P \hat{T} \mathcal{E} \right) \\ -\hat{\beta}_k s_{x_k} \hat{f}^\top \hat{T} \end{bmatrix},$$

em que  $P$  é uma matriz  $p \times n$  de zeros exceto a  $k$ -ésima linha, que é composta por uns e os componentes das matrizes diagonais  $Q$  e  $\mathcal{E}$  e do vetor  $f$  estão definidos em (4.3), (4.7) e (4.4), respectivamente.

## 4.2 INFLUÊNCIA LOCAL COM PRECISÃO VARIÁVEL

Considere o modelo de regressão gama unitária definido em (2.10) e sejam  $s = p + q$  e  $\theta = (\beta^\top, \gamma^\top)^\top$ . Temos que

$$\ddot{\ell} = \ddot{\ell}(\beta, \gamma) = \begin{pmatrix} \ddot{\ell}_{*\beta\beta} & \ddot{\ell}_{*\beta\gamma} \\ \ddot{\ell}_{*\gamma\beta} & \ddot{\ell}_{*\gamma\gamma} \end{pmatrix}, \quad (4.11)$$

em que  $\ddot{\ell}_{*\beta\beta} = X^\top Q^* X$ ,  $\ddot{\ell}_{*\beta\gamma} = (\ddot{\ell}_{*\gamma\beta})^\top = X^\top F^* T^* H^* Z$  e  $\ddot{\ell}_{*\gamma\gamma} = Z^\top \mathcal{V} Z$ , onde

$$Q^* = \text{diag}\{q_1^*, \dots, q_n^*\} \text{ com } q_t^* = \{-v_t^* + (\mu_t^* + y_t^*) [g''(\mu_t)/g'(\mu_t)]\} [1/g'(\mu_t)]^2, \quad (4.12)$$

$$F^* = (-c_1^*, \dots, -c_n^*)^\top, \quad (4.13)$$

$$T^* = \text{diag} [g'(\mu_1)^{-1}, \dots, g'(\mu_n)^{-1}], \quad (4.14)$$

$$H^* = \text{diag} [h'(\phi_1)^{-1}, \dots, h'(\phi_n)^{-1}], \quad (4.15)$$

$$V^* = \text{diag} (v_1^*, \dots, v_n^*) \text{ em que } v_t^* = \{-e_t^* + e_t [h''(\phi_t)/h'(\phi_t)]\} [1/h'(\phi_t)]^2, \quad (4.16)$$

com  $\mu_t^*$ ,  $y_t^*$  e  $e_t$  definidos em (2.12) e (2.13), respectivamente, e  $v_t^*$ ,  $c_t^*$  e  $e_t^*$  definidos no Apêndice A.

A partir de (4.11) e usando a fórmula da inversa de uma matriz particionada, (RAO,

1973), obtém-se

$$\ddot{\ell}_*^{-1} = \ddot{\ell}_*^{-1}(\beta, \gamma) = \begin{pmatrix} \ddot{\ell}_*^{\beta\beta} & \ddot{\ell}_*^{\beta\gamma} \\ \ddot{\ell}_*^{\gamma\beta} & \ddot{\ell}_*^{\gamma\gamma} \end{pmatrix},$$

em que

$$\begin{aligned} (\ddot{\ell}_*^{\beta\beta})^{-1} &= \left[ X^\top Q^* X - X^\top F^* T^* H^* Z (Z^\top V^* Z)^{-1} Z^\top H^* T^* F^{*T} X \right]^{-1}, \\ \ddot{\ell}_*^{\beta\gamma} &= (\ddot{\ell}_*^{\gamma\beta})^\top = -\ddot{\ell}_*^{\beta\beta} X^\top F^* T^* H^* Z (Z^\top V^* Z)^{-1}, \end{aligned}$$

e

$$\ddot{\ell}_*^{\gamma\gamma} = (Z^\top V^* Z)^{-1} \left[ I_q + (Z^\top H^* T^* F^* X) \ddot{\ell}_*^{\beta\beta} X^\top F^* T^* H^* Z (Z^\top V^* Z)^{-1} \right]$$

onde  $I_q$  é a matriz identidade de ordem  $q \times q$ .

## 4.2.1 Esquemas de perturbação

### 4.2.1.1 Ponderação de casos

Considere o seguinte esquema de perturbação:

$$\ell_\delta(\beta, \gamma) = \sum_{t=1}^n \delta_t \ell_t(\mu_t, \phi_t),$$

com  $0 \leq \delta_t \leq 1$ . Para o método de influência local discutido anteriormente, temos que  $\delta_0 = (1, 1, \dots, 1)^\top$  e  $\Delta_t = \partial \ell_t(\hat{\theta}) / \partial \theta$ , com  $t = 1, \dots, n$ . Então, do Apêndice B, temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{pmatrix} X^\top \hat{T}^* \hat{\mathcal{E}}^* \\ Z^\top \hat{H}^* \hat{E}^* \end{pmatrix},$$

em que

$$\mathcal{E}^* = \text{diag} [(\mu_1^* + y_1^*), \dots, (\mu_n^* + y_n^*)] \quad (4.17)$$

e

$$E^* = \text{diag} (e_1, \dots, e_n), \quad (4.18)$$

sendo  $\mu_t^*$ ,  $y_t^*$  e  $e_t$  definidos em (2.12) e (2.13), respectivamente.

### 4.2.1.2 Perturbação da variável resposta

Consideremos o esquema aditivo de perturbação da resposta em que

$$y_t(\delta) = y_t + \delta_t s(y_t), \quad (4.19)$$

com  $s(y_t) = \sqrt{\hat{\mu}_t(1/[(2 - \hat{\mu}_t^{1/\hat{\phi}_t})^{\hat{\phi}_t}] - \hat{\mu}_t)}$ . Para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$  e usando o Apêndice B, temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{pmatrix} X^\top \hat{T}^* \hat{M}^* S_y^* \\ Z^\top \hat{H}^* \hat{B}^* S_y^* \end{pmatrix},$$

$T^*$  e  $H^*$  definidos em (4.14) e (4.15), respectivamente,  $M^* = \text{diag}(m_1^*, \dots, m_n^*)$ , com

$$m_t^* = \frac{\alpha_t^2}{y_t \phi_t \mu^{1/\phi_t + 1}}, \quad (4.20)$$

$S_y^* = \text{diag}[s^*(y_1), \dots, s^*(y_n)]$  e  $\hat{B}^* = (\hat{b}_1^*, \dots, \hat{b}_n^*)$  com

$$b_t^* = \left[ -\frac{\alpha_t^2 \log(\mu_t)}{y_t \mu_t^{1/\phi_t} \phi_t^2} + \frac{1}{\log(y_t) y_t} \right]. \quad (4.21)$$

Lembremos que  $\alpha_t = \mu_t^{1/\phi_t} / (1 - \mu_t^{1/\phi_t})$ .

#### 4.2.1.3 Perturbação individual de covariáveis

Pertubaremos a  $k$ -ésima coluna da matriz  $X$ , denotada por  $x_k$ , com  $k = 2, \dots, p$ , de forma que

$$x_{tk}(\delta) = x_{tk} + \delta_t s_{x_k}. \quad (4.22)$$

Neste caso, em que o parâmetro de precisão e a média da variável resposta são modelados simultaneamente, consideramos três cenários para o esquema de perturbação individual de uma variável.

Primeiro:  $X \neq Z$ .

Segundo:  $X = Z$ , ou seja, a matriz de regressores para o modelo da média  $X$  e a matriz de regressores para o modelo de precisão  $Z$  são iguais.

Terceiro:  $Z \neq X$ , no entanto, para algum par  $(k, k')$ , temos  $z_{tp'} = x_{tp}$  ou  $z_{tp'} = \mathcal{G}(x_{tp})$ , sendo  $\mathcal{G}$  uma função diferenciável diferente da função identidade. Aqui, não consideramos esquemas de perturbação de covariáveis em que estejam envolvidas apenas na modelagem da precisão.

##### 4.2.1.3.1 A matriz $Z$ totalmente diferente da matriz $X$

Iremos analisar duas situações, à saber: Perturbação na matriz  $X$  e Perturbação na matriz  $Z$ .

**Situação 1:** Perturbação na matriz  $X$

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

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k (x_{tk} + \delta_t s_{xk}) + \cdots + \beta_p x_{tp}$$

e  $\mu_t(\delta)$  é tal que  $g(\mu_t(\delta)) = \eta_{1t}(\delta)$ . Uma vez que os regressores que determinam a média não interferem na precisão,  $\phi_t(\delta) = \phi_t$ . Para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$  e do Apêndice B temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{bmatrix} -s_{xk} \left( \hat{\beta}_k X^\top \hat{Q}^* - P \hat{T}^* \hat{\mathcal{E}}^* \right) \\ -\hat{\beta}_k s_{xk} Z^\top \hat{F}^* \hat{T}^* \hat{H}^* \end{bmatrix}, \quad (4.23)$$

em que  $P$  é uma matriz  $p \times n$  de zeros exceto a  $k$ -ésima linha, que é composta por uns e os componentes das matrizes diagonais  $Q^*$  e  $F^*$  estão definidos em (4.12) e (4.13), respectivamente, e o vetor  $\mathcal{E}^*$  está definido em (4.17).

### Situação 2: Perturbação na matriz $Z$

Consideremos o esquema aditivo de perturbação na  $k$ -ésima coluna da matriz  $Z$ , denotada por  $z_k$ , com  $k = 1, 2, \dots, q$ , à saber:

$$z_{tk}(\delta) = z_{tk} + \delta_t s_{z_k}.$$

Neste caso, suponha  $k \neq 2$  ou  $k \neq q$ , temos

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 z_{t2} + \cdots + \gamma_k (z_{tk} + \delta_t s_{z_k}) + \cdots + \gamma_p z_{tq}$$

e  $\phi_t(\delta)$  é tal que  $h(\phi_t(\delta)) = \eta_{2t}(\delta)$ . Para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$  e do Apêndice B temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{bmatrix} -\hat{\gamma}_k s_{z_k} X^\top \hat{F}^* \hat{T}^* \hat{H}^* \\ -s_{z_k} \left( \hat{\gamma}_k Z^\top \hat{Q}^* - P \hat{H}^* \hat{\mathcal{E}}^* \right) \end{bmatrix}, \quad (4.24)$$

em que  $P$  é uma matriz  $q \times n$  de zeros exceto a  $k$ -ésima linha, que é composta por uns e as expressões  $F^*$ ,  $T^*$ ,  $H^*$ ,  $Q^*$  e  $E^*$  estão definidos em (4.13), (4.14), (4.15), (4.12) e (4.18), respectivamente.

#### 4.2.1.3.2 A matriz $Z$ é igual à matriz $X$

Neste caso,

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k (x_{tk} + \delta_t s_{xk}) + \cdots + \beta_p x_{tp},$$

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 x_{t2} + \cdots + \gamma_k (x_{tk} + \delta_t s_{xk}) + \cdots + \gamma_q x_{tq},$$

onde  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_{1t}(\delta)$  e  $\phi_t(\delta)$  é tal que  $h[\phi_t(\delta)] = \eta_{2t}(\delta)$ . Para esse tipo de perturbação,  $\delta_0 = (0, 0, \dots, 0)^\top$ . Do Apêndice B, temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \left\{ \begin{array}{l} -s_{x_k} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\gamma}_k \right) - P \hat{T}^* \hat{\mathcal{E}}^* \right] \\ -s_{x_k} \left[ X^\top \left( \hat{\gamma}_k \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P \hat{H}^* \hat{E}^* \right] \end{array} \right\}, \quad (4.25)$$

em que  $P$  é uma matriz  $p \times n$  de zeros exceto a  $k$ -ésima linha que é composta por uns e  $Q^*, F^*, T^*, H^*, \mathcal{E}^*, V^*$  e  $E^*$  estão definidos, respectivamente, em (4.12), (4.13), (4.14), (4.15), (4.17), (4.16) e (4.18).

#### 4.2.1.3.3 A $k'$ -ésima coluna da matriz $Z$ é igual à $k$ -ésima coluna da matriz $X$

Aqui estamos considerando a situação em que algumas covariáveis (não todas) que determinam a média também estão envolvidas na modelagem da precisão e pretendemos perturbar tais covariáveis. Ou seja, consideramos que  $z_{tk'} = x_{tk}$ , para algum par  $(k, k')$ , com  $k = 2, \dots, p$  e  $k' = 2, \dots, q$ . Neste caso,

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k (x_{tk} + \delta_t s_{xk}) + \dots + \beta_p x_{tp}.$$

Suponha que  $k' \neq 2$  ou  $k' \neq q$ ,

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 x_{t2} + \dots + \gamma_{k'} (x_{tk} + \delta_t s_{xk}) + \dots + \gamma_q z_{tq},$$

onde  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_{1t}(\delta)$  e  $\phi_t(\delta)$  é tal que  $h[\phi_t(\delta)] = \eta_{2t}(\delta)$ , para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$  e do Apêndice B, temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \left\{ \begin{array}{l} -s_{x_k} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\gamma}_{k'} \right) - P \hat{T}^* \hat{\mathcal{E}}^* \right] \\ -s_{x_k} \left[ Z^\top \left( \hat{\gamma}_{k'} \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P \hat{H}^* \hat{E}^* \right] \end{array} \right\}, \quad (4.26)$$

em que  $P$  é uma matriz  $q \times n$  de zeros exceto a  $k$ -ésima linha que é composta por uns e  $Q^*, F^*, T^*, H^*, \mathcal{E}^*, V^*$  e  $E^*$  estão definidos, respectivamente, em (4.12), (4.13), (4.14), (4.15), (4.17), (4.16) e (4.18).

#### 4.2.1.3.4 A $k'$ -ésima coluna da matriz $Z$ é uma função da $k$ -ésima coluna da matriz $X$

Aqui estamos considerando a situação em que  $z_{tk'} = \mathcal{G}(x_{tk})$ , com  $k = 2, \dots, p$  e  $k' = 2, \dots, q$ , em que  $\mathcal{G}$  é uma função contínua e diferenciável; se  $\mathcal{G}$  é a função identidade, temos a situação tratada na seção anterior. Aqui,

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k (x_{tk} + \delta_t s_{xk}) + \dots + \beta_p x_{tp}.$$

Suponha que  $k' \neq 2$  e  $k' \neq q$ , temos que

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 x_{t2} + \cdots + \gamma_{k'} \mathcal{G}(x_{tk} + \delta_t s_{xk}) + \cdots + \gamma_q z_{tq},$$

onde  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_{1t}(\delta)$  e  $\phi_t(\delta)$  é tal que  $h[\phi_t(\delta)] = \eta_{2t}(\delta)$ , para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$  e do Apêndice B, temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \left\{ \begin{array}{l} -s_{x_k} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \dot{G} \hat{\gamma}_{k'} \right) - P \hat{T}^* \hat{\mathcal{E}}^* \right] \\ -s_{x_k} \left[ Z^\top \left( \hat{\gamma}_{k'} \dot{G} \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P \hat{H}^* \hat{E}^* \dot{G} \right] \end{array} \right\},$$

em que  $P$  é uma matriz  $q \times n$  de zeros exceto a  $k$ -ésima linha que é composta por uns e  $Q^*, F^*, T^*, H^*, \mathcal{E}^*, V^*$  e  $E^*$  estão definidos, respectivamente, em (4.12), (4.13), (4.14), (4.15), (4.17), (4.16), (4.18) e  $\dot{G} = \text{diag}[\mathcal{G}'(x_{1k}), \dots, \mathcal{G}'(x_{nk})]$  com  $\mathcal{G}'(x_{tk}) = d\mathcal{G}(x_{tk})/dx_{tk}$ .

#### 4.2.1.4 Perturbação conjunta de covariáveis

Neste caso, estamos considerando as matrizes  $X$  e  $Z$  distintas, pois o caso destas matrizes serem iguais já esta sendo analisado na Seção 4.2.1.3.2 . Temos

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k (x_{tk} + \delta_t s_{xk}) + \cdots + \beta_p x_{tp},$$

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 z_{t2} + \cdots + \gamma_k (z_{tk} + \delta_t s_{zk}) + \cdots + \gamma_q z_{tq},$$

onde  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_{1t}(\delta)$  e  $\phi_t(\delta)$  é tal que  $h[\phi_t(\delta)] = \eta_{2t}(\delta)$ . Para esse tipo de perturbação,  $\delta_0 = (0, 0, \dots, 0)^\top$ . Do Apêndice B, temos

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \left\{ \begin{array}{l} -s_{x_k} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\gamma}_k \right) - P_m \hat{T}^* \hat{\mathcal{E}}^* \right] \\ -s_{z_k} \left[ Z^\top \left( \hat{\gamma}_k \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P_p \hat{H}^* \hat{E}^* \right] \end{array} \right\}, \quad (4.27)$$

em que  $P_m$  e  $P_p$  são matrizes  $p \times n$  e  $q \times n$ , respectivamente, de zeros exceto a  $k$ -ésima linha que é composta por uns e  $Q^*, F^*, T^*, H^*, \mathcal{E}^*, V^*$  e  $E^*$  estão definidos, respectivamente, em (4.12), (4.13), (4.14), (4.15), (4.17), (4.16) e (4.18).

Resumidamente, neste capítulo fizemos o estudo da influência local, no modelo de regressão gama unitária considerando o parâmetro de precisão fixo e variável, no qual podemos verificar mais detalhes sobre os cálculos e as manipulações algébricas no Apêndice B.

## 5 MEDIDAS DE QUALIDADE DO MODELO ESTIMADO

Na literatura existem diversas medidas para a seleção de modelos e variáveis. Dentre os principais podemos citar soma de quadrados de resíduos(SQR) e várias funções do SQR tais como  $R^2$  e o  $R^2$ -ajustado ((BAYER; CRIBARI-NETO, 2017)). É importante destacar que critérios de tipo  $R^2$  objetivam selecionar modelos em que a variabilidade da variável resposta seja explicada o máximo possível pelo modelo de regressão. No entanto, tais medidas não oferecem nenhuma visão da qualidade dos valores preditos. Essa informação está muito mais relacionada com o desempenho do modelo de regressão em explicar a média da variável resposta. Assim, todos os elementos do submodelo da média, como por exemplo, a função de ligação, preditor linear e não linear são melhor avaliados a partir de medidas de seleção de modelos com foco em predição.

(ALLEN, 1971) propôs utilizar a estatística PRESS, definida como a soma dos quadrados preditos, como critério para selecionar o melhor modelo preditivo em modelos lineares. A PRESS é uma medida que seleciona modelos na perspectiva de predição, ou seja, identifica o modelo com a melhor habilidade para predizer. O cálculo da PRESS consiste em ajustar o modelo, repetidamente, excluindo uma observação de cada vez. Em cada repetição o modelo é utilizado para predizer a observação que ficou de fora. O melhor modelo para predição terá o valor da PRESS relativamente pequeno. No entanto, não é possível utilizar a PRESS para comparar modelos competidores. Neste sentido (MEDIIVILLA; LANDRAM; SHAH, 2008) propôs a estatística  $P^2$ , definida como o complementar do quociente entre a PRESS e a soma de quadrados totais sem a  $t$ -ésima observação, a estatística  $P^2$  mede a qualidade de predição do modelo.

Assim, é importante ressaltar que medidas do tipo  $R^2$  e  $P^2$  contém informações diferentes sobre o modelo ajustado. Por exemplo, medidas do tipo  $R^2$  podem selecionar modelos que explicam bem a variabilidade da variável resposta para um específico conjuntos de dados. Isso não garante uma boa qualidade de ajuste para novos conjuntos de dados associados ao mesmo problema de interesse. Quando o ajuste a um novo conjunto de dados não se mostra satisfatório, inclusive apresentando valores reduzidos de critérios do tipo  $R^2$ , dizemos que ocorreu overfitting.

Overfitting é identificável quando um modelo ajustado apresenta altos valores para os critérios do tipo  $R^2$  e baixos valores para critérios do tipo  $P^2$ . Quando este fato ocorre devemos dar uma atenção especial à especificação do submodelo da média, o que inclui uma atenção maior a função de ligação e a ausência de covariadas importantes.

Outro aspecto importante é que os critérios do tipo  $R^2$  são muito favorecidos quando os dados apresentam pouca dispersão, independentemente da modelagem da média, o que não ocorre com os critérios do tipo  $P^2$ . Dessa forma, uma abordagem de seleção de

modelos deve considerar que critérios do tipo  $P^2$  e  $R^2$  devem apresentar valores similares, demonstrando a importância da proposta do critério de seleção  $P^2$ .

## 5.1 MEDIDAS DE PREDIÇÃO: DEFINIÇÕES E PROPRIEDADES

Considere o modelo linear,  $Y = X\beta + \varepsilon$ , em que  $Y$  é um vetor  $n \times 1$  de respostas,  $X$  é uma matriz conhecida de covariadas de dimensão  $n \times p$ ,  $\beta$  é o vetor de parâmetros de dimensão  $p \times 1$  e  $\varepsilon$  é um vetor  $n \times 1$  de erros com distribuição  $N_n(0; \sigma^2 I_n)$ . Seja  $\beta = (X^\top X)^{-1} X^\top y$ ,  $\varepsilon_t = y_t - x_t^\top \hat{\beta}$ ,  $\hat{y} = x_t^\top \hat{\beta}$ , em que  $\hat{\beta}_{(t)}$  é o estimador de  $\beta$  sem a  $t$ -ésima observação e  $\hat{y}_{(t)} = x_t^\top \hat{\beta}_{(t)}$  o valor predito do caso deletado quando a covariável assume o valor  $x_t$ . Assim, a PRESS para o modelo de regressão múltipla é dada por

$$\text{PRESS} = \sum_{t=1}^n [y_t - \hat{y}_{(t)}]^2, \quad (5.1)$$

que pode ser reescrita como  $\text{PRESS} = \sum_{t=1}^n (y_t - \hat{y}_t)^2 / (1 - \hat{h}_{tt})^2$ , onde  $h_{tt}$  é o  $t$ -ésimo elemento da diagonal da matriz  $X(X^\top X)^{-1}X^\top$ . Para mais detalhes recomendamos (ALLEN, 1971), (HOCKING, 1972) e (GEISSER; EDDY, 1979), por exemplo.

Para o modelo de regressão gama unitária, temos que  $\hat{\beta}$  em (2.21), pode ser visto como a solução de mínimos quadrados ordinários da regressão

$$\check{y} = \hat{V}_{\beta\beta}^{1/2} t_1 \quad \text{contra} \quad \check{X} = \hat{V}_{\beta\beta}^{1/2} X,$$

em que  $V_{\beta\beta}$  está definido em (2.19). Logo, o erro de predição é  $\check{y} - \hat{y}_{(t)} = \hat{v}_{\beta\beta,t}^{1/2} t_{1,t} - \hat{v}_{\beta\beta,t}^{1/2} x_t^\top \hat{\beta}_{(t)}$ . Adicionalmente, segundo (PREGIBON et al., 1981), temos que

$$\hat{\beta}_{(t)} = \hat{\beta} - \frac{(X^\top V_{\beta\beta} X)^{-1} x_t \hat{v}_{\beta\beta,t}^{1/2} r_t^\beta}{1 - \hat{h}_{tt}},$$

no qual,  $r_t^\beta$  é dado em (2.22) e neste caso  $h_{tt}$  é o  $t$ -ésimo elemento da diagonal principal de

$$H = (V_{\beta\beta})^{1/2} X (X^\top V_{\beta\beta} X)^{-1} X^\top (V_{\beta\beta})^{1/2}.$$

Podemos escrever o erro predito como

$$\begin{aligned} \check{y} - \hat{y}_{(t)} &= \hat{v}_{\beta\beta,t}^{1/2} t_{1,t} - \hat{v}_{\beta\beta,t}^{1/2} x_t^\top \hat{\beta}_{(t)} = \hat{v}_{\beta\beta,t}^{1/2} t_{1,t} - \hat{v}_{\beta\beta,t}^{1/2} x_t^\top \left[ \hat{\beta} - \frac{(X^\top V_{\beta\beta} X)^{-1} x_t \hat{v}_{\beta\beta,t}^{1/2} r_t^\beta}{1 - \hat{h}_{tt}} \right] \\ &= \frac{r_t^\beta - r_t \hat{h}_{tt} + \hat{h}_{tt} r_t^\beta}{1 - \hat{h}_{tt}} = \frac{r_t^\beta}{1 - \hat{h}_{tt}}. \end{aligned}$$

Assim, a estatística PRESS para o modelo de regressão gama unitária é dada por

$$\text{PRESS} = \sum_{t=1}^n \left[ \check{y}_t - \hat{y}_{(t)} \right]^2 = \sum_{t=1}^n \left( \frac{r_t^\beta}{1 - \hat{h}_{tt}} \right)^2, \quad (5.2)$$

sendo  $r_t^\beta$  o resíduo ponderado dado em (2.22).

Podemos calcular a estatística  $P^2$  baseada na PRESS, que é dada por

$$P^2 = 1 - \frac{\text{PRESS}}{\text{SST}_{(t)}}, \quad (5.3)$$

em que  $\text{SST}_{(t)} = \sum_{t=1}^n [y_t - \bar{y}_{(t)}]^2$ , sendo  $\bar{y}_{(t)}$  a média aritmética de  $y$  sem a  $t$ -ésima observação. Pode-se mostrar que  $\text{SST}_{(t)} = [n/(n-r)]^2 \text{SST}$ , em que  $r$  é o número de parâmetros do modelo. Para o modelo gama unitária  $\text{SST} = \sum_{t=1}^n (\check{y}_t - \bar{\check{y}})^2$ ,  $\bar{\check{y}}$  é a média aritmética de  $\check{y} = V_{\beta\beta}^{1/2} t_{1,t}$ ,  $t = 1, \dots, n$ ,  $r = p + q$ , onde  $p$  e  $q$  são os números respectivamente de covariadas no submodelo da média e no submodelo da precisão e  $V_{\beta\beta}$  está definido em (2.19).

Para avaliar a qualidade de ajuste do modelo consideramos a medida de adequação  $R_{FC}^2$  proposto por (FERRARI; CRIBARI-NETO, 2004) que é definido como o quadrado da correlação entre  $g(y)$  e  $\hat{\eta}_1$  e o  $R_{LR}^2$  que é baseado na razão de verossimilhanças definida por (BAYER; CRIBARI-NETO, 2017), da seguinte forma

$$R_{LR}^2 = 1 - \left( \frac{L_{null}}{L_{fit}} \right)^{2/n}, \quad (5.4)$$

em que  $L_{null}$  é a função de verossimilhança maximizada do modelo sem regressores e  $L_{fit}$  é a função de verossimilhança maximizada do modelo ajustado. Consideramos ainda as versões corrigidas definidas para o modelo de regressão beta propostos por (BAYER; CRIBARI-NETO, 2017), que são dados por

$$R_{FC_c}^2 = 1 - (1 - R_{FC}^2) \frac{(n-1)}{[n - (p+q)]}, \quad (5.5)$$

$$R_{LR_c}^2 = 1 - (1 - R_{LR}^2) \left[ \frac{n-1}{n - (1+\alpha)p - (1-\alpha)q} \right], \quad (5.6)$$

em que  $\alpha \in [0, 1]$  e  $\delta > 0$ . (BAYER; CRIBARI-NETO, 2017) sugerem  $\alpha = 0.4$  e  $\delta = 1$ . Similarmente a  $R_{FC_c}^2$  definimos a versão corrigida de  $P^2$  dada por

$$P_c^2 = 1 - (1 - P^2)(n-1)/[n - (p+q)]. \quad (5.7)$$

As expressões (5.3) e (5.7) podem ser positivas ou negativas, dependendo do conjunto de valores assumidos por  $\{\text{PRESS e } \text{SST}_{(t)}\}$  e  $\{P^2, n, p \text{ e } q\}$ , respectivamente. De fato,

pela forma como foi definida a expressão da PRESS na equação (5.2), temos que o seu valor será sempre positivo e o mesmo acontecerá para o valor de  $SST_{(t)}$ . Então a fração  $PRESS/SST_{(t)}$  será sempre positiva, logo o valor de  $P^2$  será positivo, se a fração mencionada anteriormente for menor que um, caso contrário, será negativo.

Note que

$$\text{Se } \left( \frac{PRESS}{SST_{(t)}} \right)_{40} < \left( \frac{PRESS}{SST_{(t)}} \right)_{80} < \left( \frac{PRESS}{SST_{(t)}} \right)_{120} \text{ então } P_{40}^2 > P_{80}^2 > P_{120}^2,$$

onde o índice 40, 80 e 120 representa o tamanho da amostra. No nosso caso tomamos  $n = 40, 80, 120$ , pois será onde iremos realizar o nosso estudo.

E as relações a seguir das estatísticas  $P^2$ ,  $P_c^2$ ,  $R_{FC}^2$  e  $R_{LR}^2$  com o número de observações são obtidas diretamente da definição.

$$\begin{aligned} P_{c40}^2 > P_{c80}^2 > P_{c120}^2 & \text{ então } P_{40}^2 > P_{80}^2 > P_{120}^2, \\ P_{40}^2 < P_{80}^2 < P_{120}^2 & \text{ então } P_{c40}^2 < P_{c80}^2 < P_{c120}^2, \\ R_{FC40}^2 > R_{FC80}^2 > R_{FC120}^2, \\ R_{LR40}^2 > R_{LR80}^2 > R_{LR120}^2. \end{aligned}$$

Analisando agora, os possíveis valores de  $P_c^2$  observando a expressão (5.7), notamos que o denominador será sempre positivo e que a expressão  $(1 - P^2)(n - 1)$ , pode assumir valores positivo ou negativo devido ao valor de  $P^2$ , logo o valor de  $P_c^2$  será positivo se a expressão anterior for menor que um, e será negativo caso contrário.

Ambas medidas  $P^2$  e  $P_c^2$  medem a predição do modelo, o qual analisamos da seguinte forma: quanto mais próximo de um estiverem essas medidas, melhor será o poder preditivo do modelo e quanto mais distante de um (podendo até ser negativo como visto anteriormente), menor será o valor preditivo do modelo.

Com relação as medidas de  $R_{FC}^2$  e  $R_{LR}^2$  sempre serão positivas pelas suas definições e os valores de  $R_{FCc}^2$  e  $R_{LRc}^2$  podem ser positivos ou negativos dependendo do conjunto de valores recebidos  $\{R_{FC}^2, n, p \text{ e } q\}$  e  $\{R_{LR}^2, n, p \text{ e } q\}$ , respectivamente.

As relações que podemos encontrar destas estatísticas analisando o número de observações são:

$$\begin{aligned} R_{FCc40}^2 > R_{FCc80}^2 > R_{FCc120}^2 & \text{ então } R_{FC40}^2 > R_{FC80}^2 > R_{FC120}^2, \\ R_{FC40}^2 < R_{FC80}^2 < R_{FC120}^2 & \text{ então } R_{FCc40}^2 < R_{FCc80}^2 < R_{FCc120}^2, \\ R_{LRc40}^2 > R_{LRc80}^2 > R_{LRc120}^2 & \text{ então } R_{LR40}^2 > R_{LR80}^2 > R_{LR120}^2, \\ R_{LR40}^2 < R_{LR80}^2 < R_{LR120}^2 & \text{ então } R_{LRc40}^2 < R_{LRc80}^2 < R_{LRc120}^2, \end{aligned}$$

As medidas  $R_{FC}^2$ ,  $R_{LR}^2$ ,  $R_{FCc}^2$  e  $R_{LRc}^2$  medem a qualidade do ajuste do modelo e analisaremos da seguinte forma: quando o seu valor se encontra próximo de um, melhor será a qualidade de ajuste do modelo e quanto mais distante de um (podendo assumir

valores negativos, como visto anteriormente), menor será a qualidade de ajuste do modelo.

Resumindo, as medidas  $P^2$  e  $P_c^2$  assumem valores positivos ou negativos,  $R_{FC}^2$  e  $R_{LR}^2$  serão sempre positivos e os valores de  $R_{FC_c}^2$  e  $R_{LR_c}^2$  podem ser positivos ou negativos.

Em todas as tabelas que serão usadas a seguir, a notação  $P_\beta^2$  e  $P_{\beta_c}^2$  correspondem as estatística de  $P^2$  e  $P_c^2$ , respectivamente, com relação ao resíduo ponderado. De maneira análoga temos  $P_{\beta p}^2$  e  $P_{\beta p_c}^2$  representando o resíduo ponderado padronizado e  $P_{\beta \gamma}^2$  e  $P_{\beta \gamma_c}^2$  representando o resíduo combinado.

## 5.2 SIMULAÇÃO DAS MEDIDAS DE PREDIÇÃO

Nesta seção começamos o nosso estudo no modelo de regressão gama unitária com precisão fixa. Estudaremos este modelo utilizando as funções de ligação Logito, Log-Log, Complementar Log-Log e Probit, para modelarmos a média da variável resposta e a partir deste estudo analisando as medidas de predição iremos obter qual a função de ligação mais apropriada, quando a média da variável resposta se encontra nos seguintes cenários: próximo de zero, próximo de meio e próximo de um. Aqui não iremos utilizar a função de ligação Cauchy, pois encontramos grandes problemas de convergência do algoritmo de maximização scoring de Fisher ao utilizar essa função, impossibilitando a sua utilização neste estudo. A função de ligação utilizada na precisão foi a Logaritmo e esta não foi alterada em nenhuma simulação.

Iremos estudar as medidas de predição em diversos situações e cenários para a média da variável resposta, para que dessa forma possamos entender melhor os seus comportamento no modelo de regressão gama unitária. Analisamos quatro situações. As duas primeiras situações consideramos o modelo com precisão constante; na primeira situação o modelo está corretamente especificado e investigamos a influência das funções de ligação na qualidade do modelo estimado e na segunda situação avaliamos o desempenho das medidas de predição diante do problema de omissão de covariadas.

Na terceira situação investigamos o comportamento das medidas de predição nos modelos com precisão constante e variável à medida que são introduzidas novas covariadas importantes nos dois submodelos. Finalmente, na quarta situação avaliamos os desempenhos das medidas de qualidade do modelo estimado quando o erro de especificação é a omissão de covariadas. Para isso os dados são gerados com precisão variável e estimados com precisão constante.

### 5.2.1 Precisão constante: análise das funções de ligação

Por definição temos as seguintes expressões para as funções de ligação, Logito:  $\log(\mu/(1-\mu))$ , Probit:  $\Phi^{-1}$ , que significa a inversa da função de distribuição acumulada da distribuição normal padrão, Complementar Log-Log:  $\log(-\log(1-\mu))$  e Log-Log:  $-\log(-\log(\mu))$ .

Para encontrarmos qual a função de ligação apropriada para a média da variável resposta, dependendo de onde está esteja localizada (próxima de zero, próxima de meio

e próxima de um), iremos analisar as medidas de predição da seção anterior no modelo de regressão gama unitária com o parâmetro de precisão fixo. O resultado deste estudo será apresentado nas Tabelas 20, 21 e 22 onde informamos a função de ligação, o modelo de regressão, os valores atribuídos ao parâmetro de precisão e os cenários no qual se encontra a média da variável resposta. Vale a pena destacarmos também que, foi feito o estudo análogo, para os modelos com 2, 3 e 4 parâmetros, os quais mostraram o mesmo desempenho observados nas Tabelas 20, 21 e 22

Para este estudo, realizamos simulações de Monte Carlo com 10.000 réplicas. As observações das covariadas foram gerados de forma independente através da distribuição  $x_t \sim \mathcal{U}(0, 1)$  com  $t = 1, \dots, n$  e mantiveram-se fixas em todas as réplicas. Foram considerados os seguintes valores para o parâmetro de precisão e tamanho da amostra,  $\phi = (5.0, 15.0, 30.0)$  e  $n = (40, 80, 120)$ , respectivamente.

Analisando a Tabela 20 observamos que a função de ligação Log-Log é a mais adequada quando a média da variável resposta está próxima de zero, pois esta apresentou os melhores valores nas medidas de predição.

Na Tabela 21, as funções de ligação que forneceram os melhores valores das medidas predição foram Probit e Log-Log. Quando comparamos os valores de  $R_{FC}^2$ ,  $R_{LR}^2$  e suas correções na função de ligação Probit comparando com a Log-Log, não encontramos grande alteração de valores. No entanto, quando comparamos os valores de  $P^2$  relacionados a cada resíduo e as suas versões corrigidas notamos que a função de ligação Probit mostra melhor desempenho quanto ao poder preditivo do modelo. Logo, para dados modelados com base na regressão gama unitária a função de ligação mais adequada, quando a média da variável resposta está próximo de meio é a Probit.

Ao fazermos uma avaliação semelhante da Tabela 21 para a Tabela 22, usando as funções de ligação Log-Log e Complementar Log-Log, concluímos que a função de ligação mais adequada quando a média da variável resposta está próximo de um é a Complementar Log-Log.

Ressaltamos que as medidas de qualidade de ajuste e predição do modelo são muito altas quando  $\mu \approx 1$  independentemente da função de ligação utilizada enfatizando assim, o bom desempenho do modelo de regressão gama unitária quando a média da variável resposta se encontra próximo de um.

Com base nas conclusões relativas as funções de ligação, nas simulações subsequentes, para cada faixa de  $\mu$  utilizaremos a função de ligação que se mostrou mais adequada. Desta forma os valores das medidas irão refletir se existe algum tipo de má especificação do modelo que não seja a função de ligação.

Tabela 20 – Valores médios das medidas de ligação; Funções de ligação; Modelo:  $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ ,  $\phi_t$  constante e  $\mu \approx 0.0$ . Modelo corretamente especificado

| n                      | Modelo                 | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ |         |         |                            |         |         |                            |         |         |                            |         |         |
|------------------------|------------------------|---------------------------------------|---------|---------|----------------------------|---------|---------|----------------------------|---------|---------|----------------------------|---------|---------|
|                        |                        | Logito                                |         |         | Log-log                    |         |         | Comp.-loglog               |         |         | Probito                    |         |         |
|                        |                        | $\mu \in (0.0201; 0.3284)$            |         |         | $\mu \in (0.0202; 0.3285)$ |         |         | $\mu \in (0.0200; 0.3285)$ |         |         | $\mu \in (0.0204; 0.3286)$ |         |         |
| 40                     | $\mu$                  | 5.0                                   | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    |
|                        | $\phi$                 | 0.28991                               | 0.37076 | 0.48340 | 0.60328                    | 0.73662 | 0.83164 | 0.20348                    | 0.22725 | 0.29613 | 0.46259                    | 0.58171 | 0.70066 |
|                        | $P_2^\beta$            | 0.25153                               | 0.33675 | 0.45547 | 0.58183                    | 0.72238 | 0.82254 | 0.16043                    | 0.18548 | 0.25808 | 0.43354                    | 0.55910 | 0.68448 |
|                        | $P_2^{\beta_c}$        | 0.28016                               | 0.35926 | 0.47081 | 0.59132                    | 0.72657 | 0.82427 | 0.19499                    | 0.21777 | 0.28487 | 0.45018                    | 0.56917 | 0.68965 |
|                        | $P_2^{\beta_p}$        | 0.24125                               | 0.32463 | 0.44221 | 0.56923                    | 0.71179 | 0.81477 | 0.15148                    | 0.17549 | 0.24622 | 0.42046                    | 0.54588 | 0.67287 |
|                        | $P_2^{\beta_{pc}}$     | 0.28739                               | 0.36932 | 0.48254 | 0.60002                    | 0.73530 | 0.83111 | 0.20193                    | 0.22637 | 0.29554 | 0.45966                    | 0.58026 | 0.69996 |
|                        | $P_2^{\beta_\gamma}$   | 0.24887                               | 0.33522 | 0.45457 | 0.57840                    | 0.72099 | 0.82199 | 0.15880                    | 0.18456 | 0.25746 | 0.43045                    | 0.55757 | 0.68374 |
|                        | $R_2^{\beta_\gamma c}$ | 0.47181                               | 0.67254 | 0.79217 | 0.48334                    | 0.69268 | 0.81165 | 0.41679                    | 0.62693 | 0.76079 | 0.47576                    | 0.67899 | 0.80029 |
|                        | $R_2^{FC}$             | 0.44326                               | 0.65484 | 0.78094 | 0.45541                    | 0.67607 | 0.80147 | 0.38526                    | 0.60676 | 0.74786 | 0.44743                    | 0.66164 | 0.78950 |
|                        | $R_2^{C_c}$            | 0.50552                               | 0.69480 | 0.80841 | 0.50167                    | 0.69407 | 0.80977 | 0.49219                    | 0.68483 | 0.80277 | 0.49941                    | 0.69167 | 0.80791 |
|                        | $R_2^{LR}$             | 0.47309                               | 0.67478 | 0.79584 | 0.46899                    | 0.67401 | 0.79730 | 0.45889                    | 0.66416 | 0.78984 | 0.46659                    | 0.67145 | 0.79532 |
|                        | $R_2^{LR_c}$           | 0.26500                               | 0.34600 | 0.46116 | 0.58690                    | 0.72580 | 0.82331 | 0.17639                    | 0.20084 | 0.26916 | 0.44079                    | 0.56428 | 0.68605 |
|                        | $P_2^\beta$            | 0.24591                               | 0.32901 | 0.44716 | 0.57617                    | 0.71868 | 0.81872 | 0.15499                    | 0.18008 | 0.25018 | 0.42627                    | 0.55296 | 0.67790 |
|                        | $P_2^{\beta_c}$        | 0.26028                               | 0.34040 | 0.45493 | 0.58086                    | 0.72074 | 0.81958 | 0.17271                    | 0.19670 | 0.26409 | 0.43466                    | 0.55805 | 0.68054 |
|                        | $P_2^{\beta_p}$        | 0.24107                               | 0.32327 | 0.44077 | 0.56997                    | 0.71348 | 0.81489 | 0.15123                    | 0.17584 | 0.24498 | 0.41997                    | 0.54657 | 0.67225 |
|                        | $P_2^{\beta_{pc}}$     | 0.26222                               | 0.34448 | 0.46024 | 0.58322                    | 0.72439 | 0.82274 | 0.17476                    | 0.19997 | 0.26858 | 0.43748                    | 0.56273 | 0.68530 |
| $P_2^{\beta_\gamma}$   | 0.24305                | 0.32745                               | 0.44622 | 0.57239 | 0.71723                    | 0.81814 | 0.15332 | 0.17919                    | 0.24958 | 0.42287 | 0.55137                    | 0.67713 |         |
| $R_2^{\beta_\gamma c}$ | 0.46269                | 0.66549                               | 0.78764 | 0.47607 | 0.68879                    | 0.80770 | 0.40731 | 0.62113                    | 0.75541 | 0.46883 | 0.67490                    | 0.79615 |         |
| $R_2^{FC}$             | 0.44873                | 0.65680                               | 0.78212 | 0.46246 | 0.68070                    | 0.80271 | 0.39192 | 0.61129                    | 0.74905 | 0.45503 | 0.66646                    | 0.79086 |         |
| $R_2^{C_c}$            | 0.50096                | 0.69025                               | 0.80513 | 0.49554 | 0.69021                    | 0.80586 | 0.48509 | 0.68058                    | 0.79860 | 0.49286 | 0.68765                    | 0.80391 |         |
| $R_2^{LR}$             | 0.48532                | 0.68054                               | 0.79902 | 0.47974 | 0.68050                    | 0.79978 | 0.46896 | 0.67057                    | 0.79229 | 0.47697 | 0.67786                    | 0.79776 |         |
| $R_2^{LR_c}$           | 0.25646                | 0.33810                               | 0.45378 | 0.58190 | 0.72177                    | 0.82048 | 0.16766 | 0.19185                    | 0.26013 | 0.43395 | 0.55803                    | 0.68111 |         |
| $P_2^\beta$            | 0.24375                | 0.32678                               | 0.44444 | 0.57475 | 0.71702                    | 0.81741 | 0.15344 | 0.17803                    | 0.24748 | 0.42427 | 0.55047                    | 0.67565 |         |
| $P_2^{\beta_c}$        | 0.25335                | 0.33439                               | 0.44964 | 0.57786 | 0.71839                    | 0.81798 | 0.16533 | 0.18922                    | 0.25686 | 0.42987 | 0.55388                    | 0.67743 |         |
| $P_2^{\beta_p}$        | 0.24058                | 0.32302                               | 0.44023 | 0.57064 | 0.71357                    | 0.81487 | 0.15106 | 0.17536                    | 0.24416 | 0.42012 | 0.54626                    | 0.67192 |         |
| $P_2^{\beta_{pc}}$     | 0.25361                | 0.33654                               | 0.45285 | 0.57806 | 0.72033                    | 0.81991 | 0.16600 | 0.19099                    | 0.25955 | 0.43050 | 0.55644                    | 0.68035 |         |
| $P_2^{\beta_\gamma}$   | 0.24085                | 0.32520                               | 0.44349 | 0.57085 | 0.71555                    | 0.81683 | 0.15175 | 0.17716                    | 0.24689 | 0.42076 | 0.54886                    | 0.67488 |         |
| $R_2^{\beta_\gamma c}$ | 0.45872                | 0.66330                               | 0.78616 | 0.47412 | 0.68697                    | 0.80627 | 0.40488 | 0.61865                    | 0.75343 | 0.46711 | 0.67303                    | 0.79462 |         |
| $R_2^{FC}$             | 0.44947                | 0.65755                               | 0.78251 | 0.46513 | 0.68162                    | 0.80296 | 0.39470 | 0.61213                    | 0.74921 | 0.45800 | 0.66744                    | 0.79111 |         |
| $R_2^{C_c}$            | 0.49895                | 0.68890                               | 0.80397 | 0.48390 | 0.68845                    | 0.80449 | 0.48310 | 0.67867                    | 0.79713 | 0.49106 | 0.68582                    | 0.80250 |         |
| $R_2^{LR}$             | 0.48863                | 0.68250                               | 0.79994 | 0.48348 | 0.68203                    | 0.80047 | 0.47246 | 0.67206                    | 0.79296 | 0.48058 | 0.67936                    | 0.79844 |         |

Fonte: O autor (2020).

Tabela 21 – Valores médios das medidas de ligação; Funções de ligação; Modelo:  $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ ,  $\phi_t$  constante e  $\mu \approx 0.5$ . Modelo corretamente especificado

| n            | Modelo               | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ |         |         |                            |         |         |                            |         |         |                            |         |         |         |
|--------------|----------------------|---------------------------------------|---------|---------|----------------------------|---------|---------|----------------------------|---------|---------|----------------------------|---------|---------|---------|
|              |                      | Logito                                |         |         | Log-log                    |         |         | Comp.-loglog               |         |         | Probito                    |         |         |         |
|              |                      | $\mu \in (0.3182; 0.7971)$            |         |         | $\mu \in (0.3181; 0.7971)$ |         |         | $\mu \in (0.3181; 0.7971)$ |         |         | $\mu \in (0.3182; 0.7971)$ |         |         |         |
| 40           | $\mu$                | 5.0                                   | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    |         |
|              | $\phi$               | 0.80820                               | 0.88244 | 0.87906 | 0.78711                    | 0.87906 | 0.86907 | 0.77276                    | 0.86907 | 0.85172 | 0.75727                    | 0.84842 | 0.83983 |         |
|              | $P_2^\beta$          | 0.58639                               | 0.79783 | 0.88663 | 0.54690                    | 0.77561 | 0.87252 | 0.54121                    | 0.76048 | 0.86199 | 0.60181                    | 0.80787 | 0.89293 |         |
|              | $P_2^{\beta_c}$      | 0.59556                               | 0.80004 | 0.88731 | 0.55786                    | 0.77834 | 0.87341 | 0.55204                    | 0.76335 | 0.86287 | 0.61027                    | 0.80983 | 0.89351 |         |
|              | $P_2^{\beta_p}$      | 0.57370                               | 0.78923 | 0.88122 | 0.53396                    | 0.76636 | 0.86657 | 0.52782                    | 0.75056 | 0.85546 | 0.58921                    | 0.79955 | 0.88775 |         |
|              | $P_2^{\beta_{pc}}$   | 0.60180                               | 0.80644 | 0.89186 | 0.56484                    | 0.78545 | 0.87849 | 0.55871                    | 0.77072 | 0.86835 | 0.61646                    | 0.81602 | 0.89786 |         |
|              | $P_2^{\beta_\gamma}$ | 0.58028                               | 0.79598 | 0.88601 | 0.54132                    | 0.77385 | 0.87192 | 0.53486                    | 0.75833 | 0.86123 | 0.59573                    | 0.80608 | 0.89234 |         |
|              | $R_2^{\beta_{7c}}$   | 0.52665                               | 0.76219 | 0.86515 | 0.53385                    | 0.77256 | 0.87151 | 0.48929                    | 0.72947 | 0.84282 | 0.52505                    | 0.75970 | 0.86334 |         |
|              | $R_2^{FC}$           | 0.50106                               | 0.74933 | 0.85786 | 0.50866                    | 0.76027 | 0.86457 | 0.46168                    | 0.71485 | 0.83432 | 0.49938                    | 0.74671 | 0.85595 |         |
|              | $R_2^{C_c}$          | 0.54877                               | 0.77136 | 0.86971 | 0.57201                    | 0.78913 | 0.88043 | 0.53883                    | 0.76377 | 0.86473 | 0.54496                    | 0.76857 | 0.86790 |         |
|              | $R_2^{LR}$           | 0.51918                               | 0.75637 | 0.86116 | 0.54395                    | 0.77531 | 0.87259 | 0.50859                    | 0.74828 | 0.85586 | 0.51512                    | 0.75339 | 0.85924 |         |
|              | 80                   | $P_2^\beta$                           | 0.58884 | 0.79962 | 0.88684                    | 0.55144 | 0.77719 | 0.87317                    | 0.54321 | 0.76261 | 0.86227                    | 0.60395 | 0.80964 | 0.89313 |
|              |                      | $P_2^{\beta_c}$                       | 0.57816 | 0.79442 | 0.88390                    | 0.53979 | 0.77141 | 0.86988                    | 0.53135 | 0.75644 | 0.85870                    | 0.59367 | 0.80470 | 0.89035 |
|              |                      | $P_2^{\beta_p}$                       | 0.58278 | 0.79552 | 0.88424                    | 0.54527 | 0.77277 | 0.87032                    | 0.53694 | 0.75794 | 0.85918                    | 0.59795 | 0.80568 | 0.89066 |
|              |                      | $P_2^{\beta_{pc}}$                    | 0.57194 | 0.79021 | 0.88124                    | 0.53346 | 0.76687 | 0.86695                    | 0.52491 | 0.75165 | 0.85553                    | 0.58750 | 0.80063 | 0.88782 |
|              |                      | $P_2^{\beta_\gamma}$                  | 0.58228 | 0.79774 | 0.88621                    | 0.54557 | 0.77542 | 0.87257                    | 0.53633 | 0.76041 | 0.86150                    | 0.59743 | 0.80782 | 0.89253 |
|              |                      | $P_2^{\beta_{7c}}$                    | 0.57143 | 0.79249 | 0.88325                    | 0.53377 | 0.76959 | 0.86926                    | 0.52429 | 0.75418 | 0.85791                    | 0.58697 | 0.80282 | 0.88974 |
|              |                      | $R_2^{FC}$                            | 0.51979 | 0.75871 | 0.86205                    | 0.52757 | 0.76856 | 0.86896                    | 0.48145 | 0.72526 | 0.83917                    | 0.51851 | 0.75633 | 0.86030 |
|              |                      | $R_2^{C_c}$                           | 0.50732 | 0.75244 | 0.85847                    | 0.51530 | 0.76255 | 0.86555                    | 0.46799 | 0.71812 | 0.83499                    | 0.50601 | 0.75000 | 0.85667 |
|              |                      | $R_2^{LR}$                            | 0.54253 | 0.76802 | 0.86678                    | 0.56857 | 0.78585 | 0.87821                    | 0.53211 | 0.76024 | 0.86167                    | 0.53864 | 0.76519 | 0.86494 |
| $R_2^{LR_c}$ |                      | 0.52819                               | 0.76075 | 0.86261 | 0.55506                    | 0.77914 | 0.87439 | 0.51745                    | 0.75272 | 0.85734 | 0.52419                    | 0.75783 | 0.86071 |         |
| 120          |                      | $P_2^\beta$                           | 0.58289 | 0.79644 | 0.88493                    | 0.54480 | 0.77394 | 0.87117                    | 0.53630 | 0.75887 | 0.85996                    | 0.59816 | 0.80663 | 0.89133 |
|              |                      | $P_2^{\beta_c}$                       | 0.57576 | 0.79296 | 0.88297                    | 0.53702 | 0.77007 | 0.86897                    | 0.52837 | 0.75474 | 0.85757                    | 0.59130 | 0.80332 | 0.88948 |
|              |                      | $P_2^{\beta_p}$                       | 0.57884 | 0.79369 | 0.88320                    | 0.54067 | 0.77098 | 0.86926                    | 0.53213 | 0.75575 | 0.85790                    | 0.59415 | 0.80398 | 0.88968 |
|              |                      | $P_2^{\beta_{pc}}$                    | 0.57164 | 0.79017 | 0.88120                    | 0.53282 | 0.76706 | 0.86703                    | 0.52413 | 0.75158 | 0.85547                    | 0.58721 | 0.80063 | 0.88780 |
|              |                      | $P_2^{\beta_\gamma}$                  | 0.57608 | 0.79451 | 0.88429                    | 0.53873 | 0.77213 | 0.87056                    | 0.52913 | 0.75660 | 0.85917                    | 0.59138 | 0.80476 | 0.89072 |
|              |                      | $P_2^{\beta_{7c}}$                    | 0.56884 | 0.79100 | 0.88231                    | 0.53084 | 0.76823 | 0.86834                    | 0.52108 | 0.75244 | 0.85676                    | 0.58440 | 0.80142 | 0.88885 |
|              |                      | $R_2^{FC}$                            | 0.51805 | 0.75714 | 0.86091                    | 0.52485 | 0.76725 | 0.86806                    | 0.47951 | 0.72342 | 0.83781                    | 0.51687 | 0.75480 | 0.85917 |
|              |                      | $R_2^{C_c}$                           | 0.50982 | 0.75299 | 0.85854                    | 0.51673 | 0.76327 | 0.86380                    | 0.47061 | 0.71869 | 0.83504                    | 0.50861 | 0.75060 | 0.85676 |
|              |                      | $R_2^{LR}$                            | 0.54080 | 0.76653 | 0.86576                    | 0.56698 | 0.78481 | 0.87742                    | 0.53021 | 0.75868 | 0.86059                    | 0.53689 | 0.76368 | 0.86390 |
|              |                      | $R_2^{LR_c}$                          | 0.53135 | 0.76173 | 0.86299                    | 0.55807 | 0.78038 | 0.87490                    | 0.52054 | 0.75371 | 0.85772                    | 0.52736 | 0.75882 | 0.86110 |

Fonte: O autor (2020).

Tabela 22 – Valores médios das medidas de ligação; Funções de ligação; Modelo:  $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ ,  $\phi_t$  constante e  $\mu \approx 1$ . Modelo corretamente especificado

| Modelo            |                        | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ |         |         |                            |         |         |                            |         |         |                            |         |         |         |
|-------------------|------------------------|---------------------------------------|---------|---------|----------------------------|---------|---------|----------------------------|---------|---------|----------------------------|---------|---------|---------|
| Função de ligação |                        | Logito                                |         |         | Log-log                    |         |         | Comp.-loglog               |         |         | Probit                     |         |         |         |
|                   |                        | $\mu \in (0.7938; 0.9956)$            |         |         | $\mu \in (0.7938; 0.9956)$ |         |         | $\mu \in (0.7938; 0.9957)$ |         |         | $\mu \in (0.7938; 0.9954)$ |         |         |         |
|                   |                        | 5.0                                   | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    | 5.0                        | 15.0    | 30.0    |         |
| 40                | $\mu$                  | 0.89003                               | 0.96018 | 0.97977 | 0.87731                    | 0.95518 | 0.97718 | 0.97101                    | 0.99007 | 0.99503 | 0.94997                    | 0.98265 | 0.99129 |         |
|                   | $\phi$                 | 0.88408                               | 0.95802 | 0.97868 | 0.87068                    | 0.95276 | 0.97594 | 0.96944                    | 0.98953 | 0.99476 | 0.94726                    | 0.98171 | 0.99082 |         |
|                   | $P^2_{\beta}$          | 0.88485                               | 0.95812 | 0.97870 | 0.87162                    | 0.95288 | 0.97597 | 0.96946                    | 0.98952 | 0.99475 | 0.94740                    | 0.98172 | 0.99082 |         |
|                   | $P^2_{\beta^c}$        | 0.87862                               | 0.95586 | 0.97755 | 0.86468                    | 0.95033 | 0.97468 | 0.96781                    | 0.98896 | 0.99447 | 0.94455                    | 0.98073 | 0.99032 |         |
|                   | $P^2_{\beta^p}$        | 0.88989                               | 0.96015 | 0.97976 | 0.87717                    | 0.95515 | 0.97717 | 0.97095                    | 0.99006 | 0.99503 | 0.94988                    | 0.98264 | 0.99129 |         |
|                   | $P^2_{\beta^{\gamma}}$ | 0.88394                               | 0.95800 | 0.97867 | 0.87053                    | 0.95273 | 0.97593 | 0.96938                    | 0.98952 | 0.99476 | 0.94717                    | 0.98170 | 0.99082 |         |
|                   | $R^2_{FC}$             | 0.85961                               | 0.95100 | 0.97538 | 0.86012                    | 0.95138 | 0.97559 | 0.82114                    | 0.93339 | 0.96602 | 0.84433                    | 0.94394 | 0.97164 |         |
|                   | $R^2_{FC_c}$           | 0.85202                               | 0.94835 | 0.97405 | 0.85256                    | 0.94875 | 0.97427 | 0.81147                    | 0.92979 | 0.96418 | 0.83591                    | 0.94091 | 0.97011 |         |
|                   | $R^2_{LR}$             | 0.86540                               | 0.95080 | 0.97494 | 0.86573                    | 0.95095 | 0.97502 | 0.85146                    | 0.94494 | 0.97184 | 0.85546                    | 0.94666 | 0.97275 |         |
|                   | $R^2_{LR_c}$           | 0.85658                               | 0.94757 | 0.97329 | 0.85692                    | 0.94773 | 0.97338 | 0.84172                    | 0.94132 | 0.96999 | 0.84598                    | 0.94316 | 0.97096 |         |
|                   | 80                     | $P^2_{\beta}$                         | 0.88379 | 0.95818 | 0.97866                    | 0.87035 | 0.95293 | 0.97592                    | 0.96937 | 0.98957 | 0.99476                    | 0.94708 | 0.98178 | 0.99081 |
|                   |                        | $P^2_{\beta^c}$                       | 0.88077 | 0.95710 | 0.97810                    | 0.86699 | 0.95171 | 0.97530                    | 0.96852 | 0.98930 | 0.99462                    | 0.94571 | 0.98131 | 0.99057 |
|                   |                        | $P^2_{\beta^p}$                       | 0.88115 | 0.95715 | 0.97812                    | 0.86746 | 0.95177 | 0.97531                    | 0.96854 | 0.98930 | 0.99462                    | 0.94578 | 0.98132 | 0.99057 |
|                   |                        | $P^2_{\beta^{\gamma}}$                | 0.87806 | 0.95603 | 0.97755                    | 0.86402 | 0.95052 | 0.97467                    | 0.96772 | 0.98902 | 0.99448                    | 0.94437 | 0.98083 | 0.99033 |
|                   |                        | $P^2_{\beta^{\gamma c}}$              | 0.88362 | 0.95816 | 0.97865                    | 0.87018 | 0.95291 | 0.97591                    | 0.96925 | 0.98956 | 0.99475                    | 0.94698 | 0.98177 | 0.99081 |
|                   |                        | $P^2_{\beta^{\gamma p}}$              | 0.88060 | 0.95707 | 0.97810                    | 0.86681 | 0.95168 | 0.97529                    | 0.96845 | 0.98929 | 0.99462                    | 0.94560 | 0.98130 | 0.99057 |
| $R^2_{FC}$        |                        | 0.85664                               | 0.95003 | 0.97475 | 0.85715                    | 0.95042 | 0.97497 | 0.81745                    | 0.93211 | 0.96513 | 0.84126                    | 0.94289 | 0.97092 |         |
| $R^2_{FC_c}$      |                        | 0.85291                               | 0.94873 | 0.97409 | 0.85344                    | 0.94913 | 0.97432 | 0.81271                    | 0.93035 | 0.96423 | 0.83713                    | 0.94140 | 0.97017 |         |
| $R^2_{LR}$        |                        | 0.86290                               | 0.94987 | 0.97431 | 0.86326                    | 0.95002 | 0.97439 | 0.84843                    | 0.94386 | 0.97113 | 0.85263                    | 0.94563 | 0.97207 |         |
| $R^2_{LR_c}$      |                        | 0.85861                               | 0.94830 | 0.97351 | 0.85898                    | 0.94846 | 0.97359 | 0.84368                    | 0.94210 | 0.97022 | 0.84802                    | 0.94393 | 0.97119 |         |
| 120               |                        | $P^2_{\beta}$                         | 0.88179 | 0.95745 | 0.97828                    | 0.86812 | 0.95211 | 0.97549                    | 0.96877 | 0.98939 | 0.99466                    | 0.94616 | 0.98146 | 0.99065 |
|                   |                        | $P^2_{\beta^c}$                       | 0.87977 | 0.95673 | 0.97791                    | 0.86586 | 0.95130 | 0.97508                    | 0.96824 | 0.98921 | 0.99457                    | 0.94523 | 0.98115 | 0.99049 |
|                   |                        | $P^2_{\beta^p}$                       | 0.88002 | 0.95676 | 0.97792                    | 0.86618 | 0.95134 | 0.97509                    | 0.96826 | 0.98921 | 0.99457                    | 0.94529 | 0.98115 | 0.99049 |
|                   |                        | $P^2_{\beta^{\gamma}}$                | 0.87797 | 0.95602 | 0.97754                    | 0.86389 | 0.95050 | 0.97466                    | 0.96771 | 0.98902 | 0.99448                    | 0.94435 | 0.98083 | 0.99033 |
|                   |                        | $P^2_{\beta^{\gamma c}}$              | 0.88161 | 0.95743 | 0.97827                    | 0.86794 | 0.95209 | 0.97549                    | 0.96870 | 0.98938 | 0.99466                    | 0.94605 | 0.98145 | 0.99064 |
|                   |                        | $P^2_{\beta^{\gamma p}}$              | 0.87959 | 0.95670 | 0.97790                    | 0.86568 | 0.95127 | 0.97507                    | 0.96817 | 0.98920 | 0.99457                    | 0.94513 | 0.98113 | 0.99048 |
|                   | $R^2_{FC}$             | 0.85570                               | 0.94962 | 0.97452 | 0.85620                    | 0.95002 | 0.97475 | 0.81645                    | 0.93157 | 0.96479 | 0.84037                    | 0.94244 | 0.97065 |         |
|                   | $R^2_{FC_c}$           | 0.85323                               | 0.94876 | 0.97408 | 0.85374                    | 0.94916 | 0.97431 | 0.81331                    | 0.93040 | 0.96419 | 0.83764                    | 0.94146 | 0.97015 |         |
|                   | $R^2_{LR}$             | 0.86213                               | 0.94948 | 0.97409 | 0.86251                    | 0.94964 | 0.97418 | 0.84749                    | 0.94342 | 0.97088 | 0.85176                    | 0.94520 | 0.97183 |         |
|                   | $R^2_{LR_c}$           | 0.85929                               | 0.94844 | 0.97356 | 0.85968                    | 0.94860 | 0.97365 | 0.84435                    | 0.94225 | 0.97028 | 0.84871                    | 0.94408 | 0.97125 |         |

Fonte: O autor (2020).

### 5.2.2 Precisão constante: omissão de covariadas no submodelo da média

Nesta seção realizamos simulações Monte Carlos com 10.000 réplicas para verificar o comportamento das medidas de predição sob omissão de covariadas, consideramos os cenários 1, 2 e 3, na qual são omitidas 3, 2 e 1 covariadas, respectivamente. No cenário 4, o modelo está corretamente especificado, ou seja, é gerado e estimado corretamente. Os resultados desta simulação são apresentados nas Tabelas 23, 24 e 25 em que a média da variável resposta  $\mu$ , se encontra próximo de zero, próximo de meio e próximo de um, respectivamente.

Observando as Tabelas 23, 24 e 25 é possível notar que o valor das medidas de predição aumentam a medida que inserimos as covariadas (fato esperado), até obtermos o modelo gerado e estimado corretamente. Além disso, quando o valor de  $\phi$  aumenta obtemos uma melhora nas medidas de predição, pelo fato de  $\phi$  ser um parâmetro de precisão, ou seja, quando maior o valor do parâmetro menor é a variância, logo, mais preciso será os valores das medidas de predição.

É possível notar ainda que os valores dos coeficientes de predição são maiores em  $\mu \approx 1$  quando comparados com  $\mu \approx 0$  e  $\mu \approx 0.5$ . Notamos uma maior dificuldade de predição do modelo quando  $\mu$  assume valores na extremidade esquerda e central do intervalo  $(0, 1)$ . Encontramos em  $\mu \approx 0$  a menor predição do modelo corretamente especificado e  $\mu \approx 0.5$  detectou mais fortemente as omissões das covariadas, mas esta não ultrapassou o valor de predição do modelo gerado e estimado corretamente em  $\mu \approx 1.0$ . Dessa forma, podemos ver neste intervalo a maior adequação do modelo de distribuição gama unitária, a ponto de não sentir drasticamente a omissão de covariadas e também, os valores das medidas de predição encontram-se relativamente mais próximos uns dos outros, quando comparados com  $\mu \approx 0$  e  $\mu \approx 0.5$ , o que é um forte indício de melhor adequação do modelo.

Observe que, quando a média da variável resposta  $\mu$  se aproxima de zero a qualidade do ajuste não é boa, mesmo no cenário 3, com o parâmetros de precisão 30; quando  $\mu$  se aproxima de meio, os valores das medidas de predição possui o pior valor no cenário 1, comparando com todos os cenários de  $\mu \approx 0$  e  $\mu \approx 0.5$ . Porém, quando  $\mu \approx 1$ , temos os melhores valores das medidas de qualidade de ajuste, em todos os cenários incluindo no cenário 5, cujo o modelo é gerado e estimado corretamente, dessa forma, comportando-se semelhante aos valores das medidas de predição, donde, podemos observar nas Tabelas(23), (24) e (25). Além disso, notamos que o valor da qualidade de ajuste é razoável, mesmo no cenário 2, mostrando assim a melhor adequação e não sensibilidade do modelo de regressão gama unitária a omissão de covariadas, quando a média da variável resposta se aproxima de um.

Na análise das tabelas de omissão de covariadas, quando  $\mu \approx 0$ ,  $\mu \approx 0.5$  e  $\mu \approx 1$  foi possível observamos que a regressão gama unitária mostrou-se mais adequada quando  $\mu \approx 1$ , pois os valores das medidas de predição se assemelham e estão muito próximos de um, que é o valor desejado.

Tabela 23 – Valores médios das medidas de predição; Função de ligação Log-Log; Modelo verdadeiro:  $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5} + \beta_5 x_{t5}$ ,  $x_{ti} \sim U(0, 1)$ ,  $i = 2, 3, 4, 5$ ,  $t = 1, \dots, n$ ,  $\phi$  constante e  $\mu \approx 0$ . Modelo mal especificado: omissão de covariadas (Cenários 1, 2 e 3).

| Cenários        |                        | Cenário 1                             |          |          |         | Cenário 2  |         |         |         | Cenário 3   |         |         |         | Cenário 4  |      |      |  |  |
|-----------------|------------------------|---------------------------------------|----------|----------|---------|--|---------|---------|---------|---|---------|---------|---------|--|------|------|--|--|
| Modelo estimado |                        | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ |          |          |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$ |         |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ |         |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ |      |      |  |  |
| $\mu$           |                        | $\mu \in (0.0202; 0.3285)$            |          |          |         |  |         |         |         |   |         |         |         |  |      |      |  |  |
| n               |                        | 5.0                                   | 15.0     | 30.0     | 5.0     | 15.0   | 30.0    | 5.0     | 15.0    | 30.0  | 5.0     | 15.0    | 30.0    | 5.0  | 15.0 | 30.0 |  |  |
| 40              | $\phi$                 | 0.10163                               | 0.12867  | 0.15479  | 0.38832 | 0.47230  | 0.53919 | 0.50297 | 0.59838 | 0.67421   | 0.62271 | 0.73827 | 0.82802 |  |      |      |  |  |
|                 | $P_{\beta}^2$          | 0.053075                              | 0.081568 | 0.10910  | 0.33735 | 0.42833  | 0.50079 | 0.44617 | 0.55248 | 0.63698   | 0.56722 | 0.69978 | 0.80273 |  |      |      |  |  |
|                 | $P_{\beta^c}^2$        | 0.067960                              | 0.087764 | 0.10873  | 0.34794 | 0.42914  | 0.49538 | 0.46625 | 0.56165 | 0.63948   | 0.59104 | 0.71124 | 0.80767 |  |      |      |  |  |
|                 | $P_{\beta^p}^2$        | 0.017579                              | 0.038454 | 0.060556 | 0.29360 | 0.38156  | 0.45333 | 0.40525 | 0.51155 | 0.59827   | 0.53090 | 0.66878 | 0.77938 |  |      |      |  |  |
|                 | $P_{\beta^{pc}}^2$     | 0.096080                              | 0.12399  | 0.15060  | 0.38458 | 0.46986  | 0.53747 | 0.49928 | 0.59627 | 0.67289   | 0.61997 | 0.73713 | 0.82755 |  |      |      |  |  |
|                 | $P_{\beta^{\gamma}}^2$ | 0.047219                              | 0.076640 | 0.10468  | 0.33330 | 0.42568  | 0.49892 | 0.44206 | 0.55013 | 0.63551   | 0.56409 | 0.69847 | 0.80219 |  |      |      |  |  |
|                 | $R_{\beta^c}^2$        | 0.16367                               | 0.22875  | 0.26676  | 0.32900 | 0.46517  | 0.54500 | 0.40683 | 0.57019 | 0.66581   | 0.49503 | 0.69179 | 0.80620 |  |      |      |  |  |
|                 | $R_{FC}^2$             | 0.11846                               | 0.18707  | 0.22712  | 0.27308 | 0.42060  | 0.50708 | 0.33904 | 0.52107 | 0.62762   | 0.42076 | 0.64646 | 0.77770 |  |      |      |  |  |
|                 | $R_{FC}^2$             | 0.17041                               | 0.22619  | 0.26165  | 0.34438 | 0.46766  | 0.54437 | 0.42576 | 0.57281 | 0.66430   | 0.51929 | 0.69752 | 0.80760 |  |      |      |  |  |
|                 | $R_{LR}^2$             | 0.11601                               | 0.17545  | 0.21323  | 0.27360 | 0.41019  | 0.49518 | 0.33742 | 0.50709 | 0.61266   | 0.42136 | 0.63590 | 0.76841 |  |      |      |  |  |
|                 | 80                     | $P_{\beta}^2$                         | 0.15428  | 0.18984  | 0.22067 | 0.39685  | 0.48560 | 0.55485 | 0.48676 | 0.59043   | 0.67060 | 0.58725 | 0.71432 | 0.81186  |      |      |  |  |
|                 |                        | $P_{\beta^c}^2$                       | 0.13232  | 0.16879  | 0.20043 | 0.37304  | 0.46530 | 0.53728 | 0.45939 | 0.56859   | 0.65304 | 0.55936 | 0.69501 | 0.79915  |      |      |  |  |
|                 |                        | $P_{\beta^p}^2$                       | 0.14051  | 0.17313  | 0.20184 | 0.37889  | 0.46652 | 0.53560 | 0.46968 | 0.57345   | 0.65466 | 0.57166 | 0.70091 | 0.80171  |      |      |  |  |
|                 |                        | $P_{\beta^{pc}}^2$                    | 0.11818  | 0.15165  | 0.18111 | 0.35437  | 0.44546 | 0.51727 | 0.44140 | 0.55070   | 0.63624 | 0.54272 | 0.68070 | 0.78831  |      |      |  |  |
|                 |                        | $P_{\beta^{\gamma}}^2$                | 0.14852  | 0.18540  | 0.21677 | 0.39274  | 0.48312 | 0.55313 | 0.48256 | 0.58819   | 0.66923 | 0.58394 | 0.71300 | 0.81132  |      |      |  |  |
|                 |                        | $R_{\beta^c}^2$                       | 0.12640  | 0.16424  | 0.19642 | 0.36877  | 0.46272 | 0.53549 | 0.45497 | 0.56623   | 0.65159 | 0.55583 | 0.69360 | 0.79858  |      |      |  |  |
| $R_{FC}^2$      |                        | 0.15659                               | 0.22366  | 0.26366  | 0.31509 | 0.45583  | 0.53895 | 0.38644 | 0.55679 | 0.65674   | 0.46771 | 0.67359 | 0.79418 |  |      |      |  |  |
| $R_{FC}^2$      |                        | 0.13468                               | 0.20349  | 0.24453  | 0.28805 | 0.43435  | 0.52075 | 0.35372 | 0.53316 | 0.63843   | 0.43175 | 0.65154 | 0.78028 |  |      |      |  |  |
| $R_{LR}^2$      |                        | 0.16327                               | 0.22092  | 0.25838  | 0.33008 | 0.45783  | 0.53803 | 0.40547 | 0.55925 | 0.65513   | 0.49289 | 0.67959 | 0.79568 |  |      |      |  |  |
| $R_{LR}^2$      |                        | 0.13706                               | 0.19651  | 0.23515  | 0.29622 | 0.43043  | 0.51468 | 0.36358 | 0.52820 | 0.63083   | 0.44666 | 0.65038 | 0.77705 |  |      |      |  |  |
| 120             |                        | $P_{\beta}^2$                         | 0.16846  | 0.20723  | 0.23912 | 0.39874  | 0.48955 | 0.55922 | 0.48093 | 0.58744   | 0.66953 | 0.57529 | 0.70630 | 0.80649  |      |      |  |  |
|                 |                        | $P_{\beta^c}^2$                       | 0.15425  | 0.19368  | 0.22612 | 0.38319  | 0.47635 | 0.54782 | 0.46287 | 0.57309   | 0.65804 | 0.55666 | 0.69342 | 0.79800  |      |      |  |  |
|                 |                        | $P_{\beta^p}^2$                       | 0.15981  | 0.19670  | 0.22729 | 0.38716  | 0.47727 | 0.54687 | 0.46977 | 0.57636   | 0.65916 | 0.56493 | 0.69737 | 0.79972  |      |      |  |  |
|                 |                        | $P_{\beta^{pc}}^2$                    | 0.14544  | 0.18297  | 0.21408 | 0.37131  | 0.46375 | 0.53515 | 0.45132 | 0.56163   | 0.64731 | 0.54585 | 0.68410 | 0.79094  |      |      |  |  |
|                 |                        | $P_{\beta^{\gamma}}^2$                | 0.16264  | 0.20284  | 0.23530 | 0.39451  | 0.48705 | 0.55751 | 0.47656 | 0.58516   | 0.66815 | 0.57177 | 0.70492 | 0.80593  |      |      |  |  |
|                 |                        | $R_{\beta^c}^2$                       | 0.14833  | 0.18922  | 0.22223 | 0.37885  | 0.47378 | 0.54606 | 0.45835 | 0.57073   | 0.65661 | 0.55298 | 0.69198 | 0.79742  |      |      |  |  |
|                 | $R_{FC}^2$             | 0.15365                               | 0.22232  | 0.26277  | 0.31009 | 0.45327  | 0.53700 | 0.37944 | 0.55238 | 0.65408   | 0.45857 | 0.66750 | 0.79024 |  |      |      |  |  |
|                 | $R_{LR}^2$             | 0.13919                               | 0.20902  | 0.25017  | 0.29225 | 0.43913  | 0.52502 | 0.35785 | 0.53681 | 0.64204   | 0.43482 | 0.65291 | 0.78104 |  |      |      |  |  |
|                 | $R_{LR}^2$             | 0.16056                               | 0.21962  | 0.25739  | 0.32533 | 0.45522  | 0.53599 | 0.39875 | 0.55488 | 0.65247   | 0.48425 | 0.67366 | 0.79174 |  |      |      |  |  |
|                 | $R_{LR}^2$             | 0.14328                               | 0.20356  | 0.24211  | 0.30307 | 0.43725  | 0.52069 | 0.37128 | 0.53454 | 0.63659   | 0.45396 | 0.65450 | 0.77951 |  |      |      |  |  |

Fonte: O autor (2020).

Tabela 24 – Valores médios das medidas de predição; Função de ligação Probit; Modelo verdadeiro:  $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t3} + \beta_5 x_{t3} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, 3, 4, 5, t = 1, \dots, n, \phi$  constante e  $\mu \approx 0.5$ . Modelo mal especificado: omissão de covariadas (Cenários 1, 2 e 3).

| $\mu$                      | Modelo estimado        | Cenário 1                             |            |            |            |         | Cenário 2  |         |         |         |         | Cenário 3   |         |         |         |         | Cenário 4  |         |         |         |         |
|----------------------------|------------------------|---------------------------------------|------------|------------|------------|---------|--|---------|---------|---------|---------|---|---------|---------|---------|---------|--|---------|---------|---------|---------|
|                            |                        | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ |            |            |            |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$ |         |         |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ |         |         |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ |         |         |         |         |
| $\mu \in (0.3181; 0.7972)$ |                        |                                       |            |            |            |         |  |         |         |         |         |   |         |         |         |         |  |         |         |         |         |
| n                          | $\phi$                 | 5.0                                   | 15.0       | 30.0       | 5.0        | 15.0    | 30.0   | 5.0     | 15.0    | 30.0    | 5.0     | 15.0  | 30.0    | 5.0     | 15.0    | 30.0    | 5.0  | 15.0    | 30.0    |         |         |
| 40                         | $P^2_\beta$            | -0.097909                             | -0.10247   | -0.10636   | 0.18225    | 0.26262 | 0.29763  | 0.41723 | 0.55569 | 0.61621 | 0.64021 | 0.61621   | 0.55569 | 0.41723 | 0.55569 | 0.61621 | 0.64021  | 0.61621 | 0.55569 | 0.41723 |         |
|                            | $P^2_{\beta_c}$        | -0.15726                              | -0.16206   | -0.16617   | 0.11411    | 0.20117 | 0.23910  | 0.35063 | 0.50491 | 0.57235 | 0.58730 | 0.57235   | 0.50491 | 0.35063 | 0.50491 | 0.57235 | 0.58730  | 0.57235 | 0.50491 | 0.35063 |         |
|                            | $P^2_{\beta_p}$        | -0.11137                              | -0.11873   | -0.12383   | 0.15208    | 0.22556 | 0.25780  | 0.37513 | 0.50924 | 0.56955 | 0.60828 | 0.56955   | 0.50924 | 0.37513 | 0.50924 | 0.56955 | 0.60828  | 0.56955 | 0.50924 | 0.37513 |         |
|                            | $P^2_{\beta_{pc}}$     | -0.17145                              | -0.17920   | -0.18457   | 0.081415   | 0.16102 | 0.19595  | 0.30372 | 0.45315 | 0.52035 | 0.55068 | 0.52035   | 0.45315 | 0.30372 | 0.45315 | 0.52035 | 0.55068  | 0.52035 | 0.45315 | 0.30372 |         |
|                            | $P^2_{\beta_\gamma}$   | -0.10086                              | -0.10566   | -0.10961   | 0.17634    | 0.25715 | 0.29261  | 0.41127 | 0.55187 | 0.61342 | 0.63616 | 0.61342   | 0.55187 | 0.41127 | 0.55187 | 0.61342 | 0.63616  | 0.61342 | 0.55187 | 0.41127 |         |
|                            | $P^2_{\beta_\gamma c}$ | -0.16036                              | -0.16543   | -0.16959   | 0.10771    | 0.19524 | 0.23366  | 0.34399 | 0.50066 | 0.56924 | 0.58265 | 0.56924   | 0.50066 | 0.34399 | 0.50066 | 0.56924 | 0.58265  | 0.56924 | 0.50066 | 0.34399 |         |
|                            | $R^2_{FC}$             | 0.046730                              | 0.057941   | 0.062191   | 0.20500    | 0.28117 | 0.31331  | 0.38287 | 0.53356 | 0.59851 | 0.55256 | 0.53356   | 0.59851 | 0.38287 | 0.53356 | 0.59851 | 0.55256  | 0.53356 | 0.38287 | 0.53356 |         |
|                            | $R^2_{FCc}$            | -0.0047985                            | 0.0070186  | 0.011499   | 0.13875    | 0.22127 | 0.25608  | 0.31235 | 0.48025 | 0.52627 | 0.48675 | 0.48025   | 0.52627 | 0.31235 | 0.48025 | 0.52627 | 0.48675  | 0.48025 | 0.31235 | 0.48025 |         |
|                            | $R^2_{LR}$             | 0.050640                              | 0.059393   | 0.062540   | 0.22935    | 0.30230 | 0.33354  | 0.40362 | 0.54314 | 0.60387 | 0.57533 | 0.54314   | 0.60387 | 0.40362 | 0.54314 | 0.60387 | 0.57533  | 0.54314 | 0.40362 | 0.54314 |         |
|                            | $R^2_{LRc}$            | -0.011613                             | -0.0022860 | 0.0010674  | 0.14615    | 0.22698 | 0.26159  | 0.31187 | 0.47286 | 0.54293 | 0.48883 | 0.47286   | 0.54293 | 0.31187 | 0.47286 | 0.54293 | 0.48883  | 0.47286 | 0.31187 | 0.47286 |         |
|                            | 80                     | $P^2_\beta$                           | -0.024414  | -0.015131  | -0.011172  | 0.19579 | 0.28650  | 0.32743 | 0.39746 | 0.55213 | 0.62020 | 0.60164   | 0.62020 | 0.55213 | 0.39746 | 0.55213 | 0.62020  | 0.60164 | 0.62020 | 0.55213 | 0.39746 |
|                            |                        | $P^2_{\beta_c}$                       | -0.051022  | -0.041498  | -0.037436  | 0.16404 | 0.25834  | 0.30089 | 0.36533 | 0.52824 | 0.59995 | 0.57472   | 0.59995 | 0.52824 | 0.36533 | 0.52824 | 0.59995  | 0.57472 | 0.59995 | 0.52824 | 0.36533 |
|                            |                        | $P^2_{\beta_p}$                       | -0.028635  | -0.020646  | -0.017302  | 0.18325 | 0.27048  | 0.31010 | 0.37910 | 0.53194 | 0.60025 | 0.58605   | 0.60025 | 0.53194 | 0.37910 | 0.53194 | 0.60025  | 0.58605 | 0.60025 | 0.53194 | 0.37910 |
|                            |                        | $P^2_{\beta_{pc}}$                    | -0.055353  | -0.047156  | -0.043725  | 0.15101 | 0.24168  | 0.28286 | 0.34599 | 0.50698 | 0.57893 | 0.55808   | 0.57893 | 0.50698 | 0.34599 | 0.50698 | 0.57893  | 0.55808 | 0.57893 | 0.50698 | 0.34599 |
|                            |                        | $P^2_{\beta_\gamma}$                  | -0.027156  | -0.017961  | -0.014003  | 0.18927 | 0.28086  | 0.32240 | 0.39047 | 0.54796 | 0.61725 | 0.59661   | 0.61725 | 0.54796 | 0.39047 | 0.54796 | 0.61725  | 0.59661 | 0.61725 | 0.54796 | 0.39047 |
|                            |                        | $P^2_{\beta_\gamma c}$                | -0.053836  | -0.044402  | -0.040341  | 0.15727 | 0.25247  | 0.29565 | 0.35797 | 0.52385 | 0.59684 | 0.56935   | 0.59684 | 0.52385 | 0.35797 | 0.52385 | 0.59684  | 0.56935 | 0.59684 | 0.52385 | 0.35797 |
| $R^2_{FC}$                 |                        | 0.040413                              | 0.054332   | 0.060459   | 0.19302    | 0.27446 | 0.31010  | 0.36492 | 0.52338 | 0.59283 | 0.52913 | 0.59283   | 0.52338 | 0.36492 | 0.52338 | 0.59283 | 0.52913  | 0.59283 | 0.52338 | 0.36492 |         |
| $R^2_{FCc}$                |                        | 0.015489                              | 0.029769   | 0.036056   | 0.16117    | 0.24582 | 0.28287  | 0.33105 | 0.49796 | 0.57112 | 0.49732 | 0.57112   | 0.49796 | 0.33105 | 0.49796 | 0.57112 | 0.49732  | 0.57112 | 0.49796 | 0.33105 |         |
| $R^2_{LR}$                 |                        | 0.042936                              | 0.054962   | 0.060452   | 0.21588    | 0.29488 | 0.32998  | 0.38503 | 0.53279 | 0.59819 | 0.55245 | 0.59819   | 0.53279 | 0.38503 | 0.53279 | 0.59819 | 0.55245  | 0.59819 | 0.53279 | 0.38503 |         |
| $R^2_{LRc}$                |                        | 0.012950                              | 0.025353   | 0.031015   | 0.17626    | 0.25925 | 0.29612  | 0.34170 | 0.49987 | 0.56988 | 0.51165 | 0.56988   | 0.49987 | 0.34170 | 0.49987 | 0.56988 | 0.51165  | 0.56988 | 0.49987 | 0.34170 |         |
| 120                        |                        | $P^2_\beta$                           | -0.0032722 | 0.0097573  | 0.015605   | 0.19897 | 0.29317  | 0.33564 | 0.39020 | 0.55043 | 0.62032 | 0.58813   | 0.62032 | 0.55043 | 0.39020 | 0.55043 | 0.62032  | 0.58813 | 0.62032 | 0.55043 | 0.39020 |
|                            |                        | $P^2_{\beta_c}$                       | -0.020422  | -0.0071699 | -0.0012219 | 0.17825 | 0.27489  | 0.31845 | 0.36899 | 0.53479 | 0.60712 | 0.57006   | 0.60712 | 0.53479 | 0.36899 | 0.53479 | 0.60712  | 0.57006 | 0.60712 | 0.53479 | 0.36899 |
|                            |                        | $P^2_{\beta_p}$                       | -0.0056409 | 0.0065182  | 0.011955   | 0.19106 | 0.28293  | 0.32453 | 0.37841 | 0.53746 | 0.60755 | 0.57779   | 0.60755 | 0.53746 | 0.37841 | 0.53746 | 0.60755  | 0.57779 | 0.60755 | 0.53746 | 0.37841 |
|                            |                        | $P^2_{\beta_{pc}}$                    | -0.022831  | -0.010464  | -0.0049345 | 0.17014 | 0.26438  | 0.30706 | 0.35678 | 0.52137 | 0.59390 | 0.55927   | 0.59390 | 0.52137 | 0.35678 | 0.52137 | 0.59390  | 0.55927 | 0.59390 | 0.52137 | 0.35678 |
|                            |                        | $P^2_{\beta_\gamma}$                  | -0.0058791 | 0.0070409  | 0.012897   | 0.19228 | 0.28747  | 0.33059 | 0.38288 | 0.54614 | 0.61732 | 0.58274   | 0.61732 | 0.54614 | 0.38288 | 0.54614 | 0.61732  | 0.58274 | 0.61732 | 0.54614 | 0.38288 |
|                            |                        | $P^2_{\beta_\gamma c}$                | -0.023074  | -0.0099328 | -0.0039765 | 0.17139 | 0.26904  | 0.31328 | 0.36142 | 0.53035 | 0.60401 | 0.56443   | 0.60401 | 0.53035 | 0.36142 | 0.53035 | 0.60401  | 0.56443 | 0.60401 | 0.53035 | 0.36142 |
|                            | $R^2_{FC}$             | 0.038192                              | 0.053294   | 0.059927   | 0.18918    | 0.27248 | 0.30901  | 0.35887 | 0.52035 | 0.59062 | 0.52110 | 0.59062   | 0.52035 | 0.35887 | 0.52035 | 0.59062 | 0.52110  | 0.59062 | 0.52035 | 0.35887 |         |
|                            | $R^2_{FCc}$            | 0.021751                              | 0.037111   | 0.043858   | 0.16821    | 0.25366 | 0.29114  | 0.33657 | 0.50367 | 0.57638 | 0.50009 | 0.57638   | 0.50367 | 0.33657 | 0.50367 | 0.57638 | 0.50009  | 0.57638 | 0.50367 | 0.33657 |         |
|                            | $R^2_{LR}$             | 0.040389                              | 0.053723   | 0.059760   | 0.21153    | 0.29258 | 0.32865  | 0.37890 | 0.52961 | 0.59589 | 0.54467 | 0.59589   | 0.52961 | 0.37890 | 0.52961 | 0.59589 | 0.54467  | 0.59589 | 0.52961 | 0.37890 |         |
|                            | $R^2_{LRc}$            | 0.020637                              | 0.034246   | 0.040407   | 0.18552    | 0.26925 | 0.30650  | 0.35052 | 0.50811 | 0.57742 | 0.51794 | 0.57742   | 0.50811 | 0.35052 | 0.50811 | 0.57742 | 0.51794  | 0.57742 | 0.50811 | 0.35052 |         |

Fonte: O autor (2020).

Tabela 25 – Valores médios das medidas de predição; Função de ligação Clog-Log; Modelo verdadeiro:  $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t3} + \beta_5 x_{t5} + \beta_5 x_{t5}, x_{ti} \sim U(0, 1), i = 2, 3, 4, 5, t = 1, \dots, n, \phi$  constante e  $\mu \approx 1$ . Modelo mal especificado: omissão de covariadas (Cenários 1, 2 e 3).

|       | Cenário 1                             |         |         |         |         | Cenário 2  |         |         |         |         | Cenário 3   |         |         |                 |         | Cenário 4  |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |
|-------|---------------------------------------|---------|---------|---------|---------|--|---------|---------|---------|---------|---|---------|---------|-----------------|---------|--|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
|       | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ |         |         |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$ |         |         |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ |         |         |                 |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |
| $\mu$ | $\mu \in (0.7938; 0.99566)$           |         |         |         |         |  |         |         |         |         |   |         |         |                 |         |  |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |
| n     | 5.0                                   | 15.0    | 30.0    | 5.0     | 15.0    | 30.0   | 5.0     | 15.0    | 30.0    | 5.0     | 15.0  | 30.0    | 5.0     | 15.0            | 30.0    | 5.0  | 15.0    | 30.0    | 5.0     | 15.0    | 30.0    |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |
| 40    | $\phi$                                | 0.52878 | 0.56961 | 0.58001 | 0.78423 | 0.82198  | 0.83185 | 0.90846 | 0.93332 | 0.93980 | 0.97641   | 0.99180 | 0.99588 | $P_2^\beta$     | 0.50331 | 0.54635  | 0.55730 | 0.76625 | 0.80715 | 0.81783 | 0.89800 | 0.92570 | 0.93292 | 0.97294 | 0.99060 | 0.99527 | $P_2^{\beta_c}$    | 0.46151 | 0.49726 | 0.50603 | 0.75449 | 0.79476 | 0.80536 | 0.89121 | 0.91925 | 0.92660 | 0.97272 | 0.99048 | 0.99521 | $P_2^{\beta_{pp}}$ | 0.43240 | 0.47009 | 0.47933 | 0.73403 | 0.77766 | 0.78914 | 0.87877 | 0.91002 | 0.91821 | 0.96870 | 0.98908 | 0.99451 | $P_2^{\beta_{pc}}$   | 0.52413 | 0.56595 | 0.57663 | 0.78296 | 0.82126 | 0.83127 | 0.90821 | 0.93323 | 0.93975 | 0.97639 | 0.99180 | 0.99588 | $P_2^{\beta_\gamma}$ | 0.49840 | 0.54249 | 0.55375 | 0.76488 | 0.80637 | 0.81721 | 0.89772 | 0.92560 | 0.93286 | 0.97292 | 0.99059 | 0.99527 | $R_2^{FC}$   | 0.39291 | 0.43918 | 0.45138 | 0.52013 | 0.58000 | 0.59593 | 0.68314 | 0.76060 | 0.78163 | 0.85091 | 0.94591 | 0.97231 | $R_2^{FC_c}$ | 0.36009 | 0.40887 | 0.42173 | 0.48014 | 0.54500 | 0.56225 | 0.64693 | 0.73324 | 0.75667 | 0.82899 | 0.93796 | 0.96824 | $R_2^{LR}$   | 0.35361 | 0.38544 | 0.39359 | 0.50787 | 0.55295 | 0.56488 | 0.68790 | 0.74884 | 0.76534 | 0.87730 | 0.95542 | 0.97727 | $R_2^{LR_c}$ | 0.31122 | 0.34515 | 0.35383 | 0.45474 | 0.50469 | 0.51790 | 0.63989 | 0.71020 | 0.72924 | 0.85230 | 0.94634 | 0.97264 |
| 80    | $P_2^\beta$                           | 0.57620 | 0.62519 | 0.63884 | 0.78895 | 0.83104  | 0.84243 | 0.90667 | 0.93515 | 0.94263 | 0.97367   | 0.99098 | 0.99545 | $P_2^{\beta_c}$ | 0.56519 | 0.61546  | 0.62946 | 0.78062 | 0.82437 | 0.83621 | 0.90169 | 0.93170 | 0.93957 | 0.97189 | 0.99037 | 0.99515 | $P_2^{\beta_{pp}}$ | 0.55000 | 0.59848 | 0.61204 | 0.77590 | 0.81957 | 0.83143 | 0.89954 | 0.92971 | 0.93765 | 0.97189 | 0.99035 | 0.99514 | $P_2^{\beta_{pc}}$ | 0.53831 | 0.58805 | 0.60197 | 0.76705 | 0.81245 | 0.82478 | 0.89418 | 0.92596 | 0.93433 | 0.96999 | 0.98970 | 0.99481 | $P_2^{\beta_\gamma}$ | 0.57182 | 0.62198 | 0.63595 | 0.78752 | 0.83029 | 0.84185 | 0.90634 | 0.93505 | 0.94257 | 0.97365 | 0.99097 | 0.99545 | $R_2^{FC}$           | 0.56069 | 0.61216 | 0.62649 | 0.77913 | 0.82360 | 0.83561 | 0.90135 | 0.93158 | 0.93950 | 0.97187 | 0.99036 | 0.99515 | $R_2^{FC_c}$ | 0.39015 | 0.43769 | 0.45091 | 0.51599 | 0.57808 | 0.59539 | 0.67695 | 0.75800 | 0.78052 | 0.84249 | 0.94272 | 0.97059 | $R_2^{LR}$   | 0.37431 | 0.42309 | 0.43665 | 0.49689 | 0.56142 | 0.57942 | 0.65972 | 0.74510 | 0.76882 | 0.83185 | 0.93885 | 0.96861 | $R_2^{LR_c}$ | 0.34979 | 0.38345 | 0.39299 | 0.50268 | 0.55069 | 0.56421 | 0.68118 | 0.74608 | 0.76424 | 0.86982 | 0.95253 | 0.97569 |              |         |         |         |         |         |         |         |         |         |         |         |         |
| 120   | $P_2^\beta$                           | 0.58853 | 0.63965 | 0.65411 | 0.78980 | 0.83342  | 0.84531 | 0.90575 | 0.93541 | 0.94317 | 0.97269   | 0.99069 | 0.99531 | $P_2^{\beta_c}$ | 0.58149 | 0.63349  | 0.64820 | 0.78436 | 0.82911 | 0.84131 | 0.90247 | 0.93317 | 0.94119 | 0.97149 | 0.99028 | 0.99510 | $P_2^{\beta_{pp}}$ | 0.57222 | 0.62331 | 0.63782 | 0.78140 | 0.82613 | 0.83836 | 0.90119 | 0.93200 | 0.94007 | 0.97150 | 0.99027 | 0.99510 | $P_2^{\beta_{pc}}$ | 0.56490 | 0.61688 | 0.63163 | 0.77574 | 0.82164 | 0.83418 | 0.89775 | 0.92963 | 0.93798 | 0.97025 | 0.98984 | 0.99488 | $P_2^{\beta_\gamma}$ | 0.58421 | 0.63655 | 0.65134 | 0.78831 | 0.83266 | 0.84473 | 0.90539 | 0.93530 | 0.94310 | 0.97266 | 0.99068 | 0.99531 | $R_2^{FC}$           | 0.57711 | 0.63034 | 0.64538 | 0.78284 | 0.82834 | 0.84071 | 0.90210 | 0.93305 | 0.94112 | 0.97146 | 0.99027 | 0.99510 | $R_2^{FC_c}$ | 0.38894 | 0.43722 | 0.45083 | 0.51451 | 0.57753 | 0.59523 | 0.67467 | 0.75719 | 0.78011 | 0.83948 | 0.94164 | 0.97001 | $R_2^{LR}$   | 0.37849 | 0.42760 | 0.44145 | 0.50195 | 0.56661 | 0.58477 | 0.66335 | 0.74874 | 0.77247 | 0.83244 | 0.93908 | 0.96870 | $R_2^{LR_c}$ | 0.34859 | 0.38296 | 0.39288 | 0.50099 | 0.54997 | 0.56396 | 0.67898 | 0.74522 | 0.76372 | 0.86726 | 0.95156 | 0.97515 |              |         |         |         |         |         |         |         |         |         |         |         |         |
|       | 0.33518                               | 0.37026 | 0.38038 | 0.48453 | 0.53512 | 0.54958  | 0.66432 | 0.73357 | 0.75292 | 0.85947 | 0.94871   | 0.97369 |         |                 |         |  |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                    |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |                      |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |              |         |         |         |         |         |         |         |         |         |         |         |         |

Fonte: O autor (2020).

### 5.2.3 Precisão constante e variável: inclusão de covariadas

Nesta seção vamos avaliar o desempenho das medidas de predição tanto no modelo com precisão constante quanto variável. Nos dois modelos serão considerados quatro cenários, a saber: uma covariada, duas covariadas, três covariadas e quatro covariadas. Salientamos ainda que, para  $\mu \approx 0$  utilizamos a função de ligação log-log, para  $\mu \approx 0.5$  utilizamos a função de ligação probito e para  $\mu \approx 1$  utilizamos a função de ligação complementar log-log.

Assim, para os modelos com precisão constante temos que  $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$  com  $\beta_1 = -1.369, \beta_2 = 1.266, \mu \in (0.0202; 0.3285), \beta_1 = -0.480, \beta_2 = 1.316, \mu \in (0.3182; 0.7971), \beta_1 = 0.450, \beta_2 = 1.250, \mu \in (0.7938; 0.9957); g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$  com  $\beta_1 = -0.733, \beta_2 = 0.720, \beta_3 = -0.900, \mu \in (0.0205; 0.3285), \beta_1 = -0.611, \beta_2 = 0.780, \beta_3 = 0.787, \mu \in (0.3181; 0.7972), \beta_1 = 1.224, \beta_2 = 0.550, \beta_3 = -1.000, \mu \in (0.7936; 0.9957); g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4}$  com  $\beta_1 = -1.630, \beta_2 = 0.696, \beta_3 = 0.600, \beta_4 = 0.600, \mu \in (0.0203; 0.3285), \beta_1 = -0.778, \beta_2 = 0.600, \beta_3 = 0.649, \beta_4 = 0.800, \mu \in (0.3181; 0.7970), \beta_1 = 0.799, \beta_2 = 0.600, \beta_3 = -0.720, \beta_4 = 0.576, \mu \in (0.7938; 0.9953)$  e  $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$  com  $\beta_1 = -0.400, \beta_2 = -0.493, \beta_3 = 0.752, \beta_4 = -0.490, \beta_5 = -0.500, \mu \in (0.0202; 0.3282), \beta_1 = -0.498, \beta_2 = -0.500, \beta_3 = 0.750, \beta_4 = 0.740, \beta_5 = 0.690, \mu \in (0.3181; 0.7973)$  e  $\beta_1 = 0.900, \beta_2 = -0.900, \beta_3 = 0.500, \beta_4 = 0.530, \beta_5 = 0.500, \mu \in (0.7935; 0.9973)$ .

Para o modelo com precisão variável os valores dos betas são os mesmo descritos acima, e em adição a um modelo para  $\mu$  consideramos os seguintes modelos para  $\phi$ :  $h(\phi_t) = \gamma_1 + \gamma_2 z_{t2}$  com  $\gamma_1 = 3.000, \gamma_2 = -2.723, \phi \in (19.9879; 1.3324), \gamma_1 = 3.800, \gamma_2 = -3.420, \phi \in (44.4283; 1.4808), \gamma_1 = 4.400, \gamma_2 = -4.120, \phi \in (80.8523; 1.3433); h(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3}$  com  $\gamma_1 = 1.690, \gamma_2 = -1.600, \gamma_3 = 1.464, \phi \in (19.8831; 1.3184), \gamma_1 = 2.149, \gamma_2 = -2.000, \gamma_3 = 1.849, \phi \in (44.0290; 1.4657), \gamma_1 = 2.508, \gamma_2 = -2.500, \gamma_3 = 2.143, \phi \in (80.8631; 1.3443); h(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4}$  com  $\gamma_1 = 1.260, \gamma_2 = 1.387, \gamma_3 = 1.300, \gamma_4 = -1.418, \phi \in (19.8555; 1.3233), \gamma_1 = 1.893, \gamma_2 = 1.700, \gamma_3 = 1.300, \gamma_4 = -2.03, \phi \in (44.0348; 1.4660), \gamma_1 = 1.530, \gamma_2 = 2.100, \gamma_3 = 2.300, \gamma_4 = -1.923, \phi \in (80.8465; 1.3455)$  e  $h(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}$  com  $\gamma_1 = 3.495, \gamma_2 = -1.300, \gamma_3 = 1.300, \gamma_4 = -1.550, \gamma_5 = -2.091, \phi \in (19.8428; 1.3193), \gamma_1 = 4.232, \gamma_2 = -1.400, \gamma_3 = 1.400, \gamma_4 = -1.500, \gamma_5 = -2.834, \phi \in (44.0400; 1.4656)$  e  $\gamma_1 = 4.771, \gamma_2 = -1.700, \gamma_3 = 1.700, \gamma_4 = -1.550, \gamma_5 = -3.416, \phi \in (80.8549; 1.3420)$ .

Nas Tabelas (26, 27 e 28) e (29, 30 e 31) apresentamos os resultados das simulações para os modelos com precisão constante e corretamente especificado, para o modelo gama unitária com precisão fixa e variável, respectivamente. Nestas tabelas iremos analisar os cenários onde a média da variável resposta encontra-se concentrada próximo de zero, próximo de meio e próximo de um e o grau de variabilidade da precisão definido por  $\lambda = \max(\phi_t) / \min(\phi_t)$ ,  $t = 1, \dots, n$  é igual a 15, 30 e 60 sendo o valor do  $\max(\phi_t)$  e do  $\min(\phi_t)$ , (19.8555, 1.3233), (44.0348, 1.4660) e (80.8465, 1.3455), respectivamente.

Nas Tabelas 27 e 30, observamos que as medidas de  $P^2$  com relação aos três resíduos e

as suas versões corrigidas, obtiveram valores semelhantes. É notável que o desempenho do modelo quanto a predição é melhor para o modelo com precisão variável. O mesmo não ocorre com as medidas de qualidade de ajuste.

Dessa forma, podemos afirmar com relação as medidas de predição, que o modelo de regressão gama unitária possui bons resultados para a modelagem de dados tanto para o parâmetro de precisão fixa como o parâmetro de precisão variável, ou seja, irá depender dos dados a serem analisados para escolhermos qual modelo é mais adequado. Vale a pena ressaltar que em ambos os modelos, os valores da medida de predição são mais adequados quando a média da variável resposta encontra-se próxima de um, o que é coerente com o estudo feito na Figura 2.

#### 5.2.4 Precisão variável: Omissão da modelagem da precisão

Nesta seção, o modelo é dito corretamente especificado, quando gerarmos e estimarmos o modelo gama unitária com o parâmetro de precisão variável. No entanto, causamos um erro na estimação do modelo proposto considerando que a precisão é constante ao longo das observações, ou seja, a má especificação é omissão da modelagem da precisão.

Nas Tabelas 32, 33 e 34 iremos analisar respectivamente, os cenários onde a média da variável resposta encontra-se concentrada próximo de zero, próximo de meio, próximo de um e em cada tabela informamos os submodelos da média e precisão usados para gerarmos os dados, usando graus de variabilidade da precisão iguais a 15, 30 e 60, no qual consideramos o  $\max(\phi_t)$  e o  $\min(\phi_t)$ , (50.1551; 3.3348), (40.2003; 1.3381) e (30.0428; 0.4991), respectivamente. Em seguida exibimos os submodelos da média da variável resposta que usamos para estimação, faremos o nosso estudo para 40, 80 e 120 observações e em cada tamanho de amostra estudado exibimos o valor estimado para o parâmetro de precisão.

Observamos nas Tabelas 32, 33 e 34 que a medida que aumentamos o valor de  $\lambda$ , os valores das medidas de predição diminuem, este fato nos mostra que essas medidas são capazes de detectar o erro que estamos causando na simulação, ao gerarmos os dados com precisão variável e estimarmos com precisão fixa. Quando o grau de variabilidade aumenta, os valores estimados para o parâmetro  $\phi$  diminuem, e como  $\phi$  é um parâmetro de precisão, quanto menor o valor estimado deste parâmetro maior a variabilidade dos dados e isso ocorre exatamente para acomodar os casos que se tornam atípicos decorrente do fato das observações terem sido geradas com precisão variável e estarem sendo estimadas supondo precisão constante. Importante destacar o bom desempenho das medidas de predição que mesmo neste cenário de dispersão superestimada, conseguem identificar a má especificação do modelo.

Tabela 26 – Valores médios das estatísticas; Função de ligação Log-Log; Modelos corretamente especificado com  $\phi_t$  constante e  $\mu \approx 0$ .

| $n$                        | Modelo da média            | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$<br>(0.0202; 0.3285) |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$<br>(0.0205; 0.3285) |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4}$<br>(0.0203; 0.3285) |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$<br>(0.0202; 0.3282) |         |         |         |
|----------------------------|----------------------------|---|---------|---------|--|---------|---------|---|---------|---------|--|---------|---------|---------|
|                            |                            | 5   | 15      | 30      | 5  | 15      | 30      | 5   | 15      | 30      | 5  | 15      | 30      |         |
| 40                         | $\phi$                     | 0.60328   | 0.73662 | 0.83164 | 0.57204  | 0.69887 | 0.80107 | 0.52819   | 0.66570 | 0.77600 | 0.62271  | 0.73827 | 0.82802 |         |
|                            | $P_{\beta}^2$              | 0.58183   | 0.72238 | 0.82254 | 0.53637  | 0.67377 | 0.78449 | 0.47427   | 0.62749 | 0.75040 | 0.56722  | 0.69978 | 0.80273 |         |
|                            | $P_{\beta^c}^2$            | 0.59132   | 0.72657 | 0.82427 | 0.55305  | 0.68199 | 0.78800 | 0.50271   | 0.64190 | 0.75690 | 0.59104  | 0.71124 | 0.80767 |         |
|                            | $P_{\beta^p}^2$            | 0.56923   | 0.71179 | 0.81477 | 0.51580  | 0.65549 | 0.77033 | 0.44588   | 0.60097 | 0.72912 | 0.53090  | 0.66878 | 0.77938 |         |
|                            | $P_{\beta^e}^2$            | 0.60002   | 0.73530 | 0.83111 | 0.56899  | 0.69754 | 0.80050 | 0.52480   | 0.66414 | 0.77532 | 0.61997  | 0.73713 | 0.82755 |         |
|                            | $P_{\beta^{\gamma}}^2$     | 0.57840   | 0.72099 | 0.82199 | 0.53308  | 0.67233 | 0.78388 | 0.47049   | 0.62576 | 0.74964 | 0.56409  | 0.69847 | 0.80219 |         |
|                            | $R_{\beta^{\gamma c}}^2$   | 0.48334   | 0.69268 | 0.81165 | 0.44099  | 0.64757 | 0.77667 | 0.40971   | 0.61612 | 0.75096 | 0.49503  | 0.69179 | 0.80620 |         |
|                            | $R_{\beta^{\gamma c}^c}^2$ | 0.45541   | 0.67607 | 0.80147 | 0.39441  | 0.61820 | 0.75805 | 0.34225   | 0.57225 | 0.72249 | 0.42076  | 0.64646 | 0.77770 |         |
|                            | $R_{\beta^{\gamma c}^c}^2$ | 0.50177   | 0.69423 | 0.80979 | 0.47042  | 0.65777 | 0.78093 | 0.43374   | 0.62327 | 0.75319 | 0.51994  | 0.69797 | 0.80793 |         |
|                            | $R_{\beta^{\gamma c}^c}^2$ | 0.46910   | 0.67418 | 0.79732 | 0.41325  | 0.62082 | 0.75728 | 0.34663   | 0.56531 | 0.71521 | 0.42215  | 0.63645 | 0.76880 |         |
|                            | 80                         | $P_{\beta}^2$   | 0.58690 | 0.72580 | 0.82331  | 0.54878 | 0.68269 | 0.78849   | 0.49196 | 0.64165 | 0.75839  | 0.58725 | 0.71432 | 0.81186 |
|                            |                            | $P_{\beta^c}^2$   | 0.57617 | 0.71868 | 0.81872  | 0.53097 | 0.67016 | 0.78015   | 0.46486 | 0.62254 | 0.74551  | 0.55936 | 0.69501 | 0.79915 |
|                            |                            | $P_{\beta^p}^2$   | 0.58086 | 0.72074 | 0.81958  | 0.53939 | 0.67432 | 0.78197   | 0.47943 | 0.62981 | 0.74883  | 0.57166 | 0.70091 | 0.80171 |
|                            |                            | $P_{\beta^e}^2$   | 0.56997 | 0.71348 | 0.81489  | 0.52121 | 0.66146 | 0.77337   | 0.45166 | 0.61007 | 0.73544  | 0.54272 | 0.68070 | 0.78831 |
|                            |                            | $P_{\beta^{\gamma}}^2$                                    | 0.58322 | 0.72439 | 0.82274  | 0.54524 | 0.68123 | 0.78788   | 0.48796 | 0.63990 | 0.75763  | 0.58394 | 0.71300 | 0.81132 |
|                            |                            | $P_{\beta^{\gamma c}}^2$                                  | 0.57239 | 0.71723 | 0.81814  | 0.52729 | 0.66865 | 0.77950   | 0.46066 | 0.62070 | 0.74470  | 0.55583 | 0.69360 | 0.79858 |
| $R_{\beta^{\gamma c}^c}^2$ |                            | 0.47607   | 0.68879 | 0.80770 | 0.42703  | 0.63872 | 0.76899 | 0.38597   | 0.60068 | 0.73912 | 0.46771  | 0.67359 | 0.79418 |         |
| $R_{\beta^{\gamma c}^c}^2$ |                            | 0.46246   | 0.68070 | 0.80271 | 0.40442  | 0.62446 | 0.75987 | 0.35323   | 0.57938 | 0.72521 | 0.43175  | 0.65154 | 0.78028 |         |
| $R_{\beta^{\gamma c}^c}^2$ |                            | 0.49553   | 0.69027 | 0.80591 | 0.45770  | 0.64959 | 0.77353 | 0.41121   | 0.60826 | 0.74150 | 0.49291  | 0.67974 | 0.79584 |         |
| $R_{\beta^{\gamma c}^c}^2$ |                            | 0.47972   | 0.68057 | 0.79983 | 0.43030  | 0.63188 | 0.76208 | 0.36973   | 0.58065 | 0.72329 | 0.44669  | 0.65055 | 0.77723 |         |
| 120                        |                            | $P_{\beta}^2$   | 0.58190 | 0.72177 | 0.82048  | 0.54122 | 0.67662 | 0.78423   | 0.47975 | 0.63349 | 0.75237  | 0.57529 | 0.70630 | 0.80649 |
|                            |                            | $P_{\beta^c}^2$   | 0.57475 | 0.71702 | 0.81741  | 0.52936 | 0.66826 | 0.77865   | 0.46166 | 0.62074 | 0.74375  | 0.55666 | 0.69342 | 0.79800 |
|                            |                            | $P_{\beta^p}^2$   | 0.57786 | 0.71839 | 0.81798  | 0.53497 | 0.67104 | 0.77988   | 0.47144 | 0.62560 | 0.74598  | 0.56493 | 0.69737 | 0.79972 |
|                            |                            | $P_{\beta^e}^2$   | 0.57064 | 0.71357 | 0.81487  | 0.52294 | 0.66254 | 0.77419   | 0.45305 | 0.61258 | 0.73714  | 0.54585 | 0.68410 | 0.79094 |
|                            |                            | $P_{\beta^{\gamma}}^2$                                    | 0.57806 | 0.72033 | 0.81991  | 0.53750 | 0.67511 | 0.78360   | 0.47554 | 0.63167 | 0.75157  | 0.57177 | 0.70492 | 0.80593 |
|                            |                            | $P_{\beta^{\gamma c}}^2$                                  | 0.57085 | 0.71555 | 0.81683  | 0.52554 | 0.66671 | 0.77800   | 0.45729 | 0.61886 | 0.74293  | 0.55298 | 0.69198 | 0.79742 |
|                            | $R_{\beta^{\gamma c}^c}^2$ | 0.47412   | 0.68697 | 0.80627 | 0.42250  | 0.63506 | 0.76634 | 0.37816   | 0.59539 | 0.73503 | 0.45857  | 0.66750 | 0.79024 |         |
|                            | $R_{\beta^{\gamma c}^c}^2$ | 0.46513   | 0.68162 | 0.80296 | 0.40756  | 0.62562 | 0.76029 | 0.35653   | 0.58132 | 0.72581 | 0.43482  | 0.65291 | 0.78104 |         |
|                            | $R_{\beta^{\gamma c}^c}^2$ | 0.49390   | 0.68850 | 0.80449 | 0.45366  | 0.64564 | 0.77124 | 0.40364   | 0.60300 | 0.73744 | 0.48425  | 0.67365 | 0.79191 |         |
|                            | $R_{\beta^{\gamma c}^c}^2$ | 0.48348   | 0.68209 | 0.80047 | 0.43564  | 0.63395 | 0.76369 | 0.37639   | 0.58485 | 0.72545 | 0.45397  | 0.65448 | 0.77969 |         |

Fonte: O autor (2020).

Tabela 27 – Valores médios das estatísticas; Função de ligação Probit; Modelos corretamente especificado com  $\phi_t$  constante e  $\mu \approx 0.5$ .

| $n$                    | Modelo da média        | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$<br>(0.3181; 0.7971) |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$<br>(0.3181; 0.7971) |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4}$<br>(0.3181; 0.7972) |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$<br>(0.3182; 0.7970) |         |         |         |
|------------------------|------------------------|---|---------|---------|--|---------|---------|---|---------|---------|--|---------|---------|---------|
|                        |                        | 5   | 15      | 30      | 5  | 15      | 30      | 5   | 15      | 30      | 5  | 15      | 30      |         |
| 40                     | $\phi$                 | 0.62223   | 0.81772 | 0.89842 | 0.51607  | 0.73274 | 0.84152 | 0.58819   | 0.78755 | 0.87812 | 0.64021  | 0.82176 | 0.89929 |         |
|                        | $P_{\beta}^2$          | 0.60181   | 0.80787 | 0.89293 | 0.47574  | 0.71047 | 0.82832 | 0.54112   | 0.76327 | 0.86419 | 0.58730  | 0.79555 | 0.88448 |         |
|                        | $P_{\beta^c}^2$        | 0.61027   | 0.80983 | 0.89351 | 0.49621  | 0.71649 | 0.83025 | 0.56212   | 0.76861 | 0.86578 | 0.60828  | 0.80035 | 0.88583 |         |
|                        | $P_{\beta^p}^2$        | 0.58921   | 0.79955 | 0.88775 | 0.45422  | 0.69286 | 0.81611 | 0.51208   | 0.74216 | 0.85044 | 0.55068  | 0.77099 | 0.86904 |         |
|                        | $P_{\beta^e}^2$        | 0.61646   | 0.81602 | 0.89786 | 0.51027  | 0.73054 | 0.84070 | 0.58319   | 0.78589 | 0.87755 | 0.63616  | 0.82053 | 0.89887 |         |
|                        | $P_{\beta^{\gamma}}^2$ | 0.59573   | 0.80608 | 0.89234 | 0.46946  | 0.70808 | 0.82743 | 0.53556   | 0.76142 | 0.86355 | 0.58265  | 0.79413 | 0.88400 |         |
|                        | $R_{\beta^c}^2$        | 0.52505   | 0.75970 | 0.86334 | 0.42135  | 0.66735 | 0.79852 | 0.48958   | 0.72586 | 0.83939 | 0.55256  | 0.77371 | 0.87053 |         |
|                        | $R_{\beta^p}^2$        | 0.49938   | 0.74671 | 0.85595 | 0.37312  | 0.63962 | 0.78173 | 0.43124   | 0.69453 | 0.82104 | 0.48675  | 0.74044 | 0.85149 |         |
|                        | $R_{\beta^e}^2$        | 0.54496   | 0.76857 | 0.86790 | 0.43843  | 0.67594 | 0.80303 | 0.50901   | 0.73487 | 0.84432 | 0.57533  | 0.78426 | 0.87660 |         |
|                        | $R_{\beta^{\gamma}}^2$ | 0.51512   | 0.75339 | 0.85924 | 0.37781  | 0.64096 | 0.78176 | 0.43347   | 0.69408 | 0.82036 | 0.48883  | 0.74031 | 0.85146 |         |
|                        | 80                     | $P_{\beta}^2$   | 0.60395 | 0.80964 | 0.89313  | 0.48443 | 0.71668 | 0.83127   | 0.55286 | 0.77090 | 0.86794  | 0.60164 | 0.80396 | 0.88890 |
|                        |                        | $P_{\beta^c}^2$   | 0.59367 | 0.80470 | 0.89035  | 0.46408 | 0.70549 | 0.82461   | 0.52901 | 0.75868 | 0.86090  | 0.57472 | 0.79071 | 0.88139 |
|                        |                        | $P_{\beta^p}^2$   | 0.59795 | 0.80568 | 0.89066  | 0.47489 | 0.70870 | 0.82572   | 0.54018 | 0.76157 | 0.86183  | 0.58605 | 0.79340 | 0.88223 |
|                        |                        | $P_{\beta^e}^2$   | 0.58750 | 0.80063 | 0.88782  | 0.45416 | 0.69721 | 0.81884   | 0.51566 | 0.74886 | 0.85447  | 0.55808 | 0.77944 | 0.87427 |
|                        |                        | $P_{\beta^{\gamma}}^2$                                    | 0.59743 | 0.80782 | 0.89253  | 0.47765 | 0.71423 | 0.83037   | 0.54687 | 0.76902 | 0.86729  | 0.59661 | 0.80250 | 0.88841 |
|                        |                        | $R_{\beta^c}^2$   | 0.58697 | 0.80282 | 0.88974  | 0.45703 | 0.70295 | 0.82367   | 0.52271 | 0.75670 | 0.86022  | 0.56935 | 0.78916 | 0.88088 |
| $R_{\beta^p}^2$        |                        | 0.51851   | 0.75633 | 0.86030 | 0.40657  | 0.65883 | 0.79198 | 0.46944   | 0.71463 | 0.83159 | 0.52913  | 0.76075 | 0.86234 |         |
| $R_{\beta^e}^2$        |                        | 0.50601   | 0.75000 | 0.85667 | 0.38314  | 0.64537 | 0.78377 | 0.44115   | 0.69941 | 0.82261 | 0.49732  | 0.74458 | 0.85304 |         |
| $R_{\beta^{\gamma}}^2$ |                        | 0.53864   | 0.76519 | 0.86494 | 0.42332  | 0.66703 | 0.79643 | 0.48873   | 0.72361 | 0.83655 | 0.55245  | 0.77159 | 0.86853 |         |
| $R_{L,R}^2$            |                        | 0.52419   | 0.75783 | 0.86071 | 0.39418  | 0.65020 | 0.78614 | 0.45271   | 0.70414 | 0.82504 | 0.51165  | 0.75077 | 0.85655 |         |
| 120                    |                        | $P_{\beta}^2$   | 0.59816 | 0.80663 | 0.89133  | 0.47390 | 0.71117 | 0.82778   | 0.54077 | 0.76521 | 0.86443  | 0.58813 | 0.79804 | 0.88541 |
|                        |                        | $P_{\beta^c}^2$   | 0.59130 | 0.80332 | 0.88948  | 0.46029 | 0.70371 | 0.82333   | 0.52480 | 0.75705 | 0.85971  | 0.57006 | 0.78918 | 0.88038 |
|                        |                        | $P_{\beta^p}^2$   | 0.59415 | 0.80398 | 0.88968  | 0.46760 | 0.70589 | 0.82409   | 0.53237 | 0.75901 | 0.86036  | 0.57779 | 0.79101 | 0.88096 |
|                        |                        | $P_{\beta^e}^2$   | 0.58721 | 0.80063 | 0.88780  | 0.45383 | 0.69828 | 0.81954   | 0.51611 | 0.75062 | 0.85550  | 0.55927 | 0.78184 | 0.87574 |
|                        |                        | $P_{\beta^{\gamma}}^2$                                    | 0.59138 | 0.80476 | 0.89072  | 0.46677 | 0.70864 | 0.82685   | 0.53444 | 0.76325 | 0.86375  | 0.58274 | 0.79651 | 0.88490 |
|                        |                        | $R_{\beta^c}^2$   | 0.58440 | 0.80142 | 0.88885  | 0.45298 | 0.70111 | 0.82237   | 0.51825 | 0.75502 | 0.85901  | 0.56443 | 0.78758 | 0.87985 |
|                        | $R_{\beta^p}^2$        | 0.51687   | 0.75480 | 0.85917 | 0.40202  | 0.65570 | 0.78972 | 0.46270   | 0.71080 | 0.82887 | 0.52110  | 0.75663 | 0.85958 |         |
|                        | $R_{\beta^e}^2$        | 0.50861   | 0.75060 | 0.85676 | 0.38656  | 0.64680 | 0.78428 | 0.44402   | 0.70075 | 0.82292 | 0.50009  | 0.74595 | 0.85342 |         |
|                        | $R_{\beta^{\gamma}}^2$ | 0.53689   | 0.76368 | 0.86390 | 0.41845  | 0.66384 | 0.79417 | 0.48185   | 0.71975 | 0.83385 | 0.54467  | 0.76752 | 0.86582 |         |
|                        | $R_{L,R}^2$            | 0.52736   | 0.75882 | 0.86110 | 0.39926  | 0.65275 | 0.78738 | 0.45818   | 0.70695 | 0.82626 | 0.51794  | 0.75387 | 0.85795 |         |

Fonte: O autor (2020).



Tabela 29 – Valores médios das estatísticas; Função de ligação Log-Log; Modelos corretamente especificado com  $\phi_t$  variável e  $\mu \approx 0$ .

| n       | Modelo da média    | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$      |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$ |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ |         |         |  |         |         |         |         |
|---------|--------------------|--|---------|---------|--|---------|---------|--|---------|---------|--|---------|---------|---------|---------|
|         |                    | 15   | 30      | 60      | 15   | 30      | 60      | 15   | 30      | 60      |  |         |         |         |         |
| 40      | Modelo da precisão | $h(\gamma_t) = \gamma_1 + \gamma_2 x_{t2}$ |         |         |  |         |         |  |         |         | $h(\gamma_t) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t3} + \gamma_4 x_{t4} + \gamma_5 x_{t5}$ |         |         |         |         |
|         |                    | $(0.0204; 0.3284)$                         |         |         |  |         |         |  |         |         | $(0.0202; 0.3284)$   |         |         |         |         |
|         | $\mu$              | $\lambda$                                  | 0.73369 | 0.83277 | 0.89326  | 0.72384 | 0.81886 | 0.88127  | 0.77038 | 0.85463 | 0.90814  | 0.76899 | 0.82687 | 0.87501 |         |
|         |                    | $P_2^2$                                    | 0.71149 | 0.81883 | 0.88436  | 0.68323 | 0.79222 | 0.86381  | 0.72015 | 0.82283 | 0.88805  | 0.69969 | 0.77493 | 0.83751 |         |
|         |                    | $P_1^2$                                    | 0.72297 | 0.82462 | 0.88739  | 0.70648 | 0.80424 | 0.86968  | 0.74904 | 0.83172 | 0.89478  | 0.74122 | 0.80155 | 0.85242 |         |
|         |                    | $P_2^1$                                    | 0.69989 | 0.81001 | 0.87800  | 0.66331 | 0.77546 | 0.85051  | 0.69415 | 0.80149 | 0.87176  | 0.66358 | 0.74202 | 0.80815 |         |
|         |                    | $P_2^1$                                    | 0.73206 | 0.83185 | 0.89275  | 0.72122 | 0.81724 | 0.88031  | 0.76888 | 0.85369 | 0.90761  | 0.76750 | 0.82572 | 0.87423 |         |
|         |                    | $P_2^2$                                    | 0.70973 | 0.81783 | 0.88382  | 0.68022 | 0.79036 | 0.86271  | 0.71832 | 0.82169 | 0.88739  | 0.69776 | 0.77344 | 0.83650 |         |
|         |                    | $R_2^1$                                    | 0.45229 | 0.51841 | 0.53926  | 0.38891 | 0.45263 | 0.47541  | 0.37560 | 0.43054 | 0.45009  | 0.33127 | 0.37211 | 0.37607 |         |
|         |                    | $R_2^1$                                    | 0.40664 | 0.47828 | 0.50086  | 0.29904 | 0.37213 | 0.39826  | 0.23901 | 0.30598 | 0.32980  | 0.13065 | 0.18374 | 0.18889 |         |
|         |                    | $R_2^1$                                    | 0.65645 | 0.73693 | 0.79476  | 0.61331 | 0.69304 | 0.75894  | 0.58776 | 0.65973 | 0.72237  | 0.58532 | 0.64804 | 0.69226 |         |
|         |                    | $R_2^1$                                    | 0.62783 | 0.71501 | 0.77766  | 0.55644 | 0.64790 | 0.72349  | 0.49758 | 0.58529 | 0.66164  | 0.46092 | 0.54245 | 0.59994 |         |
|         |                    | 80   | $P_2^2$ | 0.70917 | 0.81594  | 0.88298 | 0.67660 | 0.78596  | 0.85898 | 0.70029 | 0.80526  | 0.87585 | 0.66753 | 0.73786 | 0.80718 |
|         |                    |  |         | 0.69769 | 0.80868  | 0.87836 | 0.68475 | 0.77150  | 0.84945 | 0.67115 | 0.78633  | 0.86378 | 0.62478 | 0.70416 | 0.78239 |
|         |                    |  | $P_2^2$ | 0.70380 | 0.81190  | 0.88013 | 0.66795 | 0.77895  | 0.85367 | 0.68921 | 0.79641  | 0.86933 | 0.65313 | 0.72462 | 0.79583 |
|         |                    |  | $P_2^1$ | 0.69210 | 0.80447  | 0.87540 | 0.64552 | 0.76402  | 0.84378 | 0.65899 | 0.77661  | 0.85663 | 0.60853 | 0.68921 | 0.76957 |
| $P_2^1$ | 0.70711            |  | 0.81476 | 0.88233 | 0.67318  | 0.78385 | 0.85770 | 0.69790  | 0.80374 | 0.87495 | 0.66483  | 0.73571 | 0.80561 |         |         |
| $P_2^2$ | 0.69555            |  | 0.80745 | 0.87768 | 0.65110  | 0.76925 | 0.84809 | 0.66853  | 0.78466 | 0.86279 | 0.62174  | 0.70173 | 0.78061 |         |         |
| $R_2^1$ | 0.44235            |  | 0.51003 | 0.52786 | 0.37608  | 0.43966 | 0.46193 | 0.35759  | 0.41598 | 0.43553 | 0.30914  | 0.35263 | 0.35811 |         |         |
| $R_2^1$ | 0.42033            |  | 0.49068 | 0.50922 | 0.33392  | 0.40180 | 0.42557 | 0.29514  | 0.35920 | 0.38065 | 0.22032  | 0.26939 | 0.27558 |         |         |
| $R_2^1$ | 0.65136            |  | 0.73241 | 0.79219 | 0.60113  | 0.68399 | 0.75576 | 0.55191  | 0.63026 | 0.69786 | 0.53812  | 0.60575 | 0.65560 |         |         |
| $R_2^1$ | 0.63760            |  | 0.72185 | 0.78398 | 0.57418  | 0.66264 | 0.73926 | 0.50835  | 0.59432 | 0.66849 | 0.47873  | 0.55507 | 0.61132 |         |         |
| 120     | $P_2^2$            |  | 0.70023 | 0.81021 | 0.87912  | 0.66044 | 0.77427 | 0.85097  | 0.67548 | 0.78705 | 0.86386  | 0.63138 | 0.70676 | 0.78218 |         |
|         |                    |  | 0.69248 | 0.80530 | 0.87600  | 0.64554 | 0.76437 | 0.84443  | 0.65520 | 0.77374 | 0.85535  | 0.60123 | 0.68276 | 0.76436 |         |
|         | $P_2^2$            |  | 0.69664 | 0.80752 | 0.87723  | 0.65464 | 0.76959 | 0.84747  | 0.66796 | 0.78104 | 0.85946  | 0.62156 | 0.69768 | 0.77440 |         |
|         | $P_2^1$            |  | 0.68880 | 0.80255 | 0.87406  | 0.63949 | 0.75949 | 0.84077  | 0.64720 | 0.76735 | 0.85068  | 0.59059 | 0.67294 | 0.75594 |         |
|         | $P_2^1$            | 0.69800                                    | 0.80892 | 0.87841 | 0.65669  | 0.77196 | 0.84956 | 0.67271  | 0.78527 | 0.86278 | 0.62812  | 0.70414 | 0.78023 |         |         |
|         | $P_2^2$            | 0.69019                                    | 0.80398 | 0.87526 | 0.64163  | 0.76196 | 0.84296 | 0.65226  | 0.77185 | 0.85421 | 0.59769  | 0.67993 | 0.76225 |         |         |
|         | $R_2^1$            | 0.43907                                    | 0.50587 | 0.52295 | 0.37162  | 0.43539 | 0.45756 | 0.35123  | 0.41005 | 0.42939 | 0.30061  | 0.34467 | 0.35034 |         |         |
|         | $R_2^1$            | 0.42456                                    | 0.49309 | 0.51061 | 0.34406  | 0.41062 | 0.43377 | 0.31068  | 0.37318 | 0.39372 | 0.24338  | 0.29105 | 0.29718 |         |         |
|         | $R_2^1$            | 0.64966                                    | 0.73101 | 0.79102 | 0.59786  | 0.68121 | 0.75531 | 0.54044  | 0.62044 | 0.69002 | 0.52279  | 0.59343 | 0.64469 |         |         |
|         | $R_2^1$            | 0.64060                                    | 0.72405 | 0.78561 | 0.58022  | 0.66723 | 0.74458 | 0.51172  | 0.59672 | 0.67064 | 0.48374  | 0.56016 | 0.61562 |         |         |

Fonte: O autor (2020).



Tabela 31 – Valores médios das estatísticas; Função de ligação Clog-Log; Modelos corretamente especificado com  $\phi_t$  variável e  $\mu \approx 1$ .

| n  | Modelo da média | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ |         |         |         |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$ |         |         |         |         |         | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ |         |         |         |         |         |         |         |         |         |         |         |
|----|-----------------|---------------------------------------|---------|---------|---------|---------|---------|--|---------|---------|---------|---------|---------|--|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
|    |                 | 15                                    | 30      | 60      | 15      | 30      | 60      | 15   | 30      | 60      | 15      | 30      | 60      | 15   | 30      | 60      |         |         |         |         |         |         |         |         |         |
| 40 | $P_2^2$         | 0.98881                               | 0.99459 | 0.99687 | 0.96930 | 0.98277 | 0.98914 | 0.99067  | 0.99540 | 0.99758 | 0.98960 | 0.99324 | 0.99564 | 0.98788  | 0.99414 | 0.99661 | 0.96479 | 0.98023 | 0.98754 | 0.98863 | 0.99439 | 0.99704 | 0.98648 | 0.99121 | 0.99433 |
|    | $P_2^2$         | 0.98815                               | 0.99425 | 0.99667 | 0.96617 | 0.98088 | 0.98784 | 0.98889  | 0.99435 | 0.99694 | 0.98694 | 0.99147 | 0.99432 | 0.98716  | 0.99378 | 0.99639 | 0.96117 | 0.97807 | 0.98606 | 0.98646 | 0.99312 | 0.99627 | 0.98303 | 0.98891 | 0.99261 |
|    | $P_2^2$         | 0.98879                               | 0.99458 | 0.99687 | 0.96926 | 0.98275 | 0.98913 | 0.99066  | 0.99539 | 0.99757 | 0.98959 | 0.99323 | 0.99564 | 0.98786  | 0.99413 | 0.99661 | 0.96474 | 0.98021 | 0.98753 | 0.98862 | 0.99439 | 0.99704 | 0.98647 | 0.99120 | 0.99433 |
|    | $P_2^2$         | 0.83126                               | 0.87882 | 0.89453 | 0.73884 | 0.80430 | 0.82885 | 0.79800  | 0.85466 | 0.87152 | 0.69280 | 0.74182 | 0.75184 | 0.81720  | 0.86872 | 0.88575 | 0.70043 | 0.77532 | 0.80368 | 0.75382 | 0.82287 | 0.84341 | 0.66064 | 0.66436 | 0.67739 |
|    | $R_2^{FC}$      | 0.89489                               | 0.93481 | 0.95244 | 0.81252 | 0.87227 | 0.90263 | 0.86735  | 0.91227 | 0.93288 | 0.81934 | 0.86251 | 0.88450 | 0.89489  | 0.93481 | 0.95244 | 0.81252 | 0.87227 | 0.90263 | 0.86735 | 0.91227 | 0.93288 | 0.81934 | 0.86251 | 0.88450 |
|    | $R_2^{LR}$      | 0.88613                               | 0.92938 | 0.94848 | 0.78495 | 0.85349 | 0.88831 | 0.83833  | 0.89308 | 0.91820 | 0.76515 | 0.82126 | 0.84985 | 0.89489  | 0.93481 | 0.95244 | 0.81252 | 0.87227 | 0.90263 | 0.86735 | 0.91227 | 0.93288 | 0.81934 | 0.86251 | 0.88450 |
|    | $P_2^2$         | 0.98772                               | 0.99404 | 0.99657 | 0.96142 | 0.97849 | 0.98643 | 0.98673  | 0.99354 | 0.99659 | 0.98252 | 0.98821 | 0.99244 | 0.98723  | 0.99380 | 0.99643 | 0.96881 | 0.97704 | 0.98552 | 0.98544 | 0.99291 | 0.99626 | 0.98027 | 0.98669 | 0.99146 |
|    | $P_2^2$         | 0.98738                               | 0.99388 | 0.99647 | 0.95982 | 0.97757 | 0.98583 | 0.98592  | 0.99312 | 0.99636 | 0.98120 | 0.98729 | 0.99183 | 0.98689  | 0.99363 | 0.99633 | 0.96710 | 0.97605 | 0.98487 | 0.98455 | 0.99246 | 0.99601 | 0.97878 | 0.98566 | 0.99078 |
|    | $P_2^2$         | 0.98769                               | 0.99403 | 0.99656 | 0.95136 | 0.97847 | 0.98642 | 0.98671  | 0.99353 | 0.99659 | 0.98250 | 0.98819 | 0.99243 | 0.98721  | 0.99380 | 0.99643 | 0.96875 | 0.97701 | 0.98550 | 0.98542 | 0.99290 | 0.99626 | 0.98025 | 0.98668 | 0.99146 |
|    | $R_2^{FC}$      | 0.82711                               | 0.87665 | 0.89187 | 0.73284 | 0.79987 | 0.82460 | 0.79202  | 0.85069 | 0.86831 | 0.68367 | 0.73425 | 0.74499 | 0.82028  | 0.87178 | 0.88760 | 0.71479 | 0.78635 | 0.81275 | 0.77180 | 0.83617 | 0.85550 | 0.64300 | 0.70008 | 0.71220 |
|    | $R_2^{LR}$      | 0.88708                               | 0.92992 | 0.94882 | 0.78595 | 0.85457 | 0.88870 | 0.83911  | 0.89340 | 0.91830 | 0.76590 | 0.82087 | 0.84960 | 0.88729  | 0.93385 | 0.95645 | 0.95870 | 0.97695 | 0.98544 | 0.98528 | 0.99284 | 0.99622 | 0.97986 | 0.98633 | 0.99122 |
|    | $P_2^2$         | 0.98696                               | 0.99369 | 0.99636 | 0.95689 | 0.97594 | 0.98480 | 0.98437  | 0.99239 | 0.99599 | 0.97821 | 0.98521 | 0.99050 | 0.98707  | 0.99374 | 0.99639 | 0.95762 | 0.97633 | 0.98504 | 0.98474 | 0.99256 | 0.99607 | 0.97893 | 0.98569 | 0.99079 |
|    | $P_2^2$         | 0.98674                               | 0.99358 | 0.99630 | 0.95576 | 0.97529 | 0.98438 | 0.98378  | 0.99210 | 0.99582 | 0.97720 | 0.98452 | 0.99004 | 0.98727  | 0.99384 | 0.99645 | 0.95864 | 0.97693 | 0.98526 | 0.98326 | 0.99283 | 0.99622 | 0.97983 | 0.98632 | 0.99121 |
|    | $P_2^2$         | 0.98694                               | 0.99369 | 0.99636 | 0.95682 | 0.97592 | 0.98479 | 0.98434  | 0.99238 | 0.99598 | 0.97818 | 0.98520 | 0.99049 | 0.982575   | 0.87545 | 0.89088 | 0.73073 | 0.79834 | 0.82336 | 0.78969 | 0.84903 | 0.86679 | 0.67934 | 0.73061 | 0.74126 |
|    | $R_2^{FC}$      | 0.82124                               | 0.87223 | 0.88805 | 0.71892 | 0.78950 | 0.81561 | 0.77654  | 0.83960 | 0.85846 | 0.65310 | 0.70857 | 0.72009 | 0.89003  | 0.93174 | 0.95008 | 0.79546 | 0.86094 | 0.84879 | 0.89973 | 0.92314 | 0.78416 | 0.83474 | 0.86116 |         |
|    | $R_2^{LR}$      | 0.88718                               | 0.92997 | 0.94879 | 0.78648 | 0.85484 | 0.88879 | 0.83934  | 0.89346 | 0.91834 | 0.76650 | 0.82122 | 0.84980 | 0.88718  | 0.92997 | 0.94879 | 0.78648 | 0.85484 | 0.88879 | 0.83934 | 0.89346 | 0.91834 | 0.76650 | 0.82122 |         |

Fonte: O autor (2020).

Tabela 32 – Valores médios das estatísticas; Função de ligação Log-Log; Modelos gerados com  $\phi_t$  variável e estimado com  $\phi_t$  constante onde  $\mu \approx 0$ .

| n            | $\lambda$                | Modelo gerado da média                |  |  | Modelo gerado da precisão                  |  |  | Modelo estimado da média |         |         |         |         |         |         |
|--------------|--------------------------|---------------------------------------|--|--|--|--|--|--------------------------|---------|---------|---------|---------|---------|---------|
|              |                          | $g(\mu_t) = \beta_1 + \beta_2 x_{t2}$ | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$ | $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$ | $h(\gamma_t) = \gamma_1 + \gamma_2 x_{t2}$ | $h(\gamma_t) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t3}$ | $h(\gamma_t) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t3} + \gamma_4 x_{t4} + \gamma_5 x_{t5}$ |                          |         |         |         |         |         |         |
| 40           | $P_{\beta}^2$            | 0.68589                               | 0.52731  | 0.30955  | 0.63815                                    | 0.50519  | 0.31093  | 0.62875                  | 0.50218 | 0.38285 | 0.62383 | 0.53375 | 0.52099 |         |
|              | $P_{\beta_c}^2$          | 0.66891                               | 0.50176  | 0.27223  | 0.60800                                    | 0.46396  | 0.25350  | 0.58632                  | 0.44529 | 0.31232 | 0.56851 | 0.46519 | 0.45055 |         |
|              | $P_{\beta^p}^2$          | 0.67582                               | 0.51667  | 0.30155  | 0.62119                                    | 0.48770  | 0.29923  | 0.60419                  | 0.47903 | 0.36334 | 0.58962 | 0.49712 | 0.47845 |         |
|              | $P_{\beta^{pc}}^2$       | 0.65830                               | 0.49054  | 0.26379  | 0.58962                                    | 0.44500  | 0.24083  | 0.55896                  | 0.41949 | 0.29058 | 0.52927 | 0.42317 | 0.40175 |         |
|              | $P_{\beta^{\gamma}}^2$   | 0.68376                               | 0.52200  | 0.30090  | 0.63585                                    | 0.50109  | 0.30299  | 0.62652                  | 0.49766 | 0.37643 | 0.62152 | 0.52974 | 0.51560 |         |
|              | $P_{\beta^{\gamma c}}^2$ | 0.66667                               | 0.49617  | 0.26311  | 0.60551                                    | 0.45951  | 0.24490  | 0.58384                  | 0.44025 | 0.30516 | 0.56587 | 0.46058 | 0.44437 |         |
|              | $R_{FC}^2$               | 0.65288                               | 0.49283  | 0.29678  | 0.58867                                    | 0.43533  | 0.23902  | 0.57611                  | 0.42216 | 0.29310 | 0.56619 | 0.43306 | 0.32283 |         |
|              | $R_{FC_c}^2$             | 0.63412                               | 0.46542  | 0.25877  | 0.55439                                    | 0.38828  | 0.17561  | 0.52767                  | 0.35612 | 0.21231 | 0.50240 | 0.34969 | 0.22324 |         |
|              | $R_{LR}^2$               | 0.64668                               | 0.47762  | 0.26600  | 0.60010                                    | 0.45584  | 0.23319  | 0.58765                  | 0.42600 | 0.28984 | 0.59823 | 0.48965 | 0.47773 |         |
|              | $R_{LR_c}^2$             | 0.62351                               | 0.44337  | 0.21787  | 0.55693                                    | 0.39710  | 0.15041  | 0.52421                  | 0.33770 | 0.18058 | 0.51639 | 0.38569 | 0.37134 |         |
|              | $\hat{\phi}$             | 13.939                                | 6.436  | 2.0582   | 12.63                                      | 5.6612   | 1.7328   | 12.358                   | 6.1527  | 3.1082  | 13.239  | 6.1502  | 2.3101  |         |
|              | 80                       | $P_{\beta}^2$                         | 0.66754  | 0.49052  | 0.22743                                    | 0.61075  | 0.45375  | 0.21559                  | 0.59566 | 0.44938 | 0.28850 | 0.58696 | 0.48296 | 0.44446 |
|              |                          | $P_{\beta_c}^2$                       | 0.65891  | 0.47729  | 0.20736                                    | 0.59539  | 0.43219  | 0.18463                  | 0.57410 | 0.42002 | 0.25055 | 0.55905 | 0.44802 | 0.40692 |
|              |                          | $P_{\beta^p}^2$                       | 0.66240  | 0.48517  | 0.22412                                    | 0.60238  | 0.44546  | 0.21129                  | 0.58344 | 0.43829 | 0.28065 | 0.57008 | 0.46526 | 0.42511 |
|              |                          | $P_{\beta^{pc}}^2$                    | 0.65363  | 0.47180  | 0.20397                                    | 0.58668  | 0.42357  | 0.18016                  | 0.56122 | 0.40833 | 0.24229 | 0.54103 | 0.42913 | 0.38626 |
|              |                          | $P_{\beta^{\gamma}}^2$                | 0.66514  | 0.48382  | 0.21596                                    | 0.60810  | 0.44851  | 0.20548                  | 0.59296 | 0.44329 | 0.27870 | 0.58410 | 0.47722 | 0.43410 |
|              |                          | $P_{\beta^{\gamma c}}^2$              | 0.65644  | 0.47041  | 0.19560                                    | 0.59263  | 0.42674  | 0.17412                  | 0.57125 | 0.41360 | 0.24023 | 0.55599 | 0.44190 | 0.39587 |
|              |                          | $R_{FC}^2$                            | 0.64642  | 0.48343  | 0.28377                                    | 0.57695  | 0.42173  | 0.22117                  | 0.55724 | 0.39769 | 0.26800 | 0.53824 | 0.39831 | 0.27897 |
|              |                          | $R_{FC_c}^2$                          | 0.63724  | 0.47001  | 0.26517                                    | 0.56025  | 0.39890  | 0.19042                  | 0.53362 | 0.36556 | 0.22896 | 0.50704 | 0.35765 | 0.23025 |
|              |                          | $R_{LR}^2$                            | 0.64030  | 0.46777  | 0.23838                                    | 0.58826  | 0.44168  | 0.20373                  | 0.56842 | 0.39877 | 0.24553 | 0.57291 | 0.45812 | 0.44141 |
| $R_{LR_c}^2$ |                          | 0.62903                               | 0.45109  | 0.21452  | 0.56746                                    | 0.41347  | 0.16349  | 0.53801                  | 0.35641 | 0.19237 | 0.53397 | 0.40872 | 0.39049 |         |
| $\hat{\phi}$ |                          | 13.055                                | 5.9324   | 1.8031   | 11.715                                     | 5.153  | 1.5207   | 11.243                   | 5.5882  | 2.7359  | 11.841  | 5.4565  | 2.0053  |         |
| 120          |                          | $P_{\beta}^2$                         | 0.66047  | 0.47487  | 0.19341                                    | 0.60039  | 0.43187  | 0.17645                  | 0.58369 | 0.42886 | 0.24486 | 0.57415 | 0.46367 | 0.40792 |
|              |                          | $P_{\beta_c}^2$                       | 0.65466  | 0.46589  | 0.17963                                    | 0.59006  | 0.41718  | 0.15515                  | 0.56921 | 0.40899 | 0.21860 | 0.55548 | 0.44014 | 0.38195 |
|              |                          | $P_{\beta^p}^2$                       | 0.65702  | 0.47130  | 0.19145                                    | 0.59483  | 0.42647  | 0.17403                  | 0.57554 | 0.42158 | 0.24027 | 0.56293 | 0.45198 | 0.39540 |
|              |                          | $P_{\beta^{pc}}^2$                    | 0.65116  | 0.46226  | 0.17763                                    | 0.58435  | 0.41163  | 0.15267                  | 0.56078 | 0.40146 | 0.21385 | 0.54376 | 0.42794 | 0.36888 |
|              |                          | $P_{\beta^{\gamma}}^2$                | 0.65795  | 0.46748  | 0.18092                                    | 0.59760  | 0.42613  | 0.16590                  | 0.58080 | 0.42205 | 0.23340 | 0.57106 | 0.45717 | 0.39495 |
|              |                          | $P_{\beta^{\gamma c}}^2$              | 0.65210  | 0.45837  | 0.16692                                    | 0.58720  | 0.41129  | 0.14433                  | 0.56622 | 0.40195 | 0.20673 | 0.55224 | 0.43336 | 0.36842 |
|              |                          | $R_{FC}^2$                            | 0.64372  | 0.48014  | 0.27921                                    | 0.57211  | 0.41656  | 0.21456                  | 0.55039 | 0.38886 | 0.26056 | 0.52899 | 0.38599 | 0.26296 |
|              |                          | $R_{FC_c}^2$                          | 0.63763  | 0.47126  | 0.26689                                    | 0.56104  | 0.40147  | 0.19425                  | 0.53476 | 0.36761 | 0.23484 | 0.50833 | 0.35906 | 0.23064 |
|              |                          | $R_{LR}^2$                            | 0.63762  | 0.46380  | 0.22872                                    | 0.58352  | 0.43647  | 0.19307                  | 0.56170 | 0.38853 | 0.22753 | 0.56461 | 0.44690 | 0.42633 |
|              | $R_{LR_c}^2$             | 0.63016                               | 0.45277  | 0.21284  | 0.56978                                    | 0.41788  | 0.16646  | 0.54167                  | 0.36059 | 0.19223 | 0.53905 | 0.41442 | 0.39265 |         |
|              | $\hat{\phi}$             | 12.739                                | 5.7706   | 1.7171   | 11.411                                     | 4.9734   | 1.4462   | 10.892                   | 5.3952  | 2.6035  | 11.408  | 5.2456  | 1.904   |         |

Fonte: O autor (2020).

Tabela 33 – Valores médios das estatísticas; Função de ligação Probit; Modelos gerados com  $\phi_t$  variável e estimado com  $\phi_t$  constante onde  $\mu \approx 0.5$ .

| n            | $\lambda$        | Modelo gerado da média |         |         |         | Modelo gerado da precisão |         |         |         | Modelo estimado da média |         |         |         |         |
|--------------|------------------|------------------------|---------|---------|---------|---------------------------|---------|---------|---------|--------------------------|---------|---------|---------|---------|
|              |                  | 15                     | 30      | 60      | 150     | 15                        | 30      | 60      | 150     | 15                       | 30      | 60      | 150     |         |
| 40           | $P_2^2$          | 0.79577                | 0.67404 | 0.50326 | 0.68213 | 0.54218                   | 0.38640 | 0.69425 | 0.56968 | 0.46969                  | 0.69244 | 0.56121 | 0.44169 |         |
|              | $P_2^c$          | 0.78474                | 0.65642 | 0.47641 | 0.65564 | 0.50402                   | 0.33527 | 0.65931 | 0.52050 | 0.40909                  | 0.64721 | 0.49668 | 0.35959 |         |
|              | $P_2^p$          | 0.78787                | 0.66397 | 0.49289 | 0.66558 | 0.52448                   | 0.37071 | 0.67173 | 0.54567 | 0.44646                  | 0.65927 | 0.52440 | 0.40637 |         |
|              | $P_2^{pc}$       | 0.77640                | 0.64581 | 0.46548 | 0.63771 | 0.48486                   | 0.31827 | 0.63422 | 0.49375 | 0.38320                  | 0.60916 | 0.45445 | 0.31907 |         |
|              | $P_2^{\gamma}$   | 0.79356                | 0.66898 | 0.49394 | 0.67900 | 0.53632                   | 0.38050 | 0.69135 | 0.56476 | 0.46326                  | 0.68970 | 0.55635 | 0.43597 |         |
|              | $P_2^{\gamma c}$ | 0.78240                | 0.65109 | 0.46659 | 0.65225 | 0.49768                   | 0.32888 | 0.65608 | 0.51502 | 0.40191                  | 0.64407 | 0.49111 | 0.35302 |         |
|              | $R_{FC}^2$       | 0.73549                | 0.59330 | 0.40771 | 0.61550 | 0.46021                   | 0.26327 | 0.63233 | 0.47960 | 0.38179                  | 0.63361 | 0.48339 | 0.34880 |         |
|              | $R_{FC}^c$       | 0.72119                | 0.57132 | 0.37570 | 0.58345 | 0.41523                   | 0.20188 | 0.59031 | 0.42012 | 0.31114                  | 0.57973 | 0.40742 | 0.25303 |         |
|              | $R_{LR}^2$       | 0.74214                | 0.60178 | 0.41835 | 0.61599 | 0.45949                   | 0.31167 | 0.64948 | 0.51132 | 0.40345                  | 0.64283 | 0.49554 | 0.36480 |         |
|              | $R_{LR}^c$       | 0.72524                | 0.57566 | 0.38021 | 0.57453 | 0.40114                   | 0.23736 | 0.59555 | 0.43614 | 0.31168                  | 0.57008 | 0.39278 | 0.23540 |         |
|              | $\hat{\phi}$     | 14.736                 | 7.6322  | 3.4411  | 13.355  | 6.8754                    | 2.7462  | 13.065  | 6.6574  | 3.9904                   | 13.745  | 6.7979  | 3.0693  |         |
|              | 80               | $P_2^2$                | 0.78427 | 0.65221 | 0.45682 | 0.65946                   | 0.50543 | 0.32500 | 0.66838 | 0.52801                  | 0.41189 | 0.66084 | 0.51181 | 0.36781 |
|              |                  | $P_2^c$                | 0.77866 | 0.64318 | 0.44271 | 0.64601                   | 0.48591 | 0.29835 | 0.65069 | 0.50284                  | 0.38053 | 0.63793 | 0.47883 | 0.32509 |
|              |                  | $P_2^p$                | 0.78026 | 0.64711 | 0.45171 | 0.65137                   | 0.49697 | 0.31800 | 0.65707 | 0.51612                  | 0.40089 | 0.64459 | 0.49410 | 0.35180 |
|              |                  | $P_2^{pc}$             | 0.77455 | 0.63795 | 0.43747 | 0.63761                   | 0.47711 | 0.29108 | 0.63878 | 0.49032                  | 0.36894 | 0.62058 | 0.45992 | 0.30800 |
|              |                  | $P_2^{\gamma}$         | 0.78184 | 0.64615 | 0.44341 | 0.65586                   | 0.49819 | 0.31479 | 0.66493 | 0.52160                  | 0.40265 | 0.65750 | 0.50529 | 0.35743 |
|              |                  | $P_2^{\gamma c}$       | 0.77617 | 0.63696 | 0.42896 | 0.64227                   | 0.47838 | 0.28775 | 0.64706 | 0.49609                  | 0.37079 | 0.63436 | 0.47186 | 0.31401 |
|              |                  | $R_{FC}^2$             | 0.73046 | 0.58518 | 0.39586 | 0.60466                   | 0.44614 | 0.24192 | 0.61577 | 0.45788                  | 0.35549 | 0.61155 | 0.45372 | 0.30884 |
|              |                  | $R_{FC}^c$             | 0.72346 | 0.57440 | 0.38016 | 0.58905                   | 0.42428 | 0.21200 | 0.59527 | 0.42897                  | 0.32111 | 0.58531 | 0.41681 | 0.26214 |
|              |                  | $R_{LR}^2$             | 0.73711 | 0.59398 | 0.40618 | 0.60411                   | 0.44356 | 0.29166 | 0.63456 | 0.49155                  | 0.37861 | 0.61896 | 0.46325 | 0.32345 |
| $R_{LR}^c$   |                  | 0.72887                | 0.58126 | 0.38757 | 0.58410 | 0.41545                   | 0.25587 | 0.60881 | 0.45572 | 0.33483                  | 0.58422 | 0.41432 | 0.26177 |         |
| $\hat{\phi}$ |                  | 13.801                 | 7.0603  | 3.1592  | 12.396  | 6.3702                    | 2.4962  | 11.889  | 6.0625  | 3.5947                   | 12.269  | 6.0246  | 2.7074  |         |
| 120          |                  | $P_2^2$                | 0.77984 | 0.64316 | 0.43878 | 0.65118                   | 0.49052 | 0.30049 | 0.65904 | 0.51191                  | 0.38752 | 0.64951 | 0.49317 | 0.33673 |
|              |                  | $P_2^c$                | 0.77608 | 0.63706 | 0.42919 | 0.64216                   | 0.47734 | 0.28240 | 0.64718 | 0.49493                  | 0.36621 | 0.63414 | 0.47094 | 0.30764 |
|              |                  | $P_2^p$                | 0.77715 | 0.63974 | 0.43540 | 0.64581                   | 0.48495 | 0.29605 | 0.65147 | 0.50400                  | 0.38038 | 0.63869 | 0.48143 | 0.32634 |
|              |                  | $P_2^{pc}$             | 0.77334 | 0.63358 | 0.42575 | 0.63665                   | 0.47163 | 0.27784 | 0.63935 | 0.48674                  | 0.35883 | 0.62284 | 0.45869 | 0.29679 |
|              |                  | $P_2^{\gamma}$         | 0.77732 | 0.63664 | 0.42356 | 0.64741                   | 0.48268 | 0.28836 | 0.65538 | 0.50486                  | 0.37679 | 0.64594 | 0.48599 | 0.32434 |
|              |                  | $P_2^{\gamma c}$       | 0.77351 | 0.63043 | 0.41370 | 0.63830                   | 0.46930 | 0.26996 | 0.64340 | 0.48764                  | 0.35511 | 0.63041 | 0.46344 | 0.29470 |
|              |                  | $R_{FC}^2$             | 0.72835 | 0.58241 | 0.39205 | 0.60050                   | 0.44029 | 0.23486 | 0.61013 | 0.44961                  | 0.34648 | 0.60394 | 0.44317 | 0.29492 |
|              |                  | $R_{FC}^c$             | 0.72370 | 0.57527 | 0.38165 | 0.59017                   | 0.42581 | 0.21507 | 0.59657 | 0.43047                  | 0.32375 | 0.58656 | 0.41875 | 0.26400 |
|              |                  | $R_{LR}^2$             | 0.73497 | 0.59109 | 0.40234 | 0.59955                   | 0.43709 | 0.28498 | 0.62932 | 0.48412                  | 0.36955 | 0.61070 | 0.45205 | 0.30793 |
|              | $R_{LR}^c$       | 0.72951                | 0.58267 | 0.39004 | 0.58634 | 0.41852                   | 0.26139 | 0.61238 | 0.46055 | 0.34075                  | 0.58784 | 0.41987 | 0.26730 |         |
|              | $\hat{\phi}$     | 13.47                  | 6.8833  | 3.064   | 12.077  | 6.1913                    | 2.4132  | 11.52   | 5.861   | 3.4656                   | 11.818  | 5.7953  | 2.5924  |         |

Fonte: O autor (2020).

Tabela 34 – Valores médios das estatísticas; Função de ligação Clog-Log; Modelos gerados com  $\phi_t$  variável e estimado com  $\phi_t$  constante onde  $\mu \approx 1$ .

| n            | $\lambda$                | Modelo gerado da média   |         |         | Modelo gerado da precisão |         |         | Modelo estimado da média |         |         |         |         |         |         |
|--------------|--------------------------|--------------------------|---------|---------|---------------------------|---------|---------|--------------------------|---------|---------|---------|---------|---------|---------|
|              |                          | 15                       | 30      | 60      | 15                        | 30      | 60      | 15                       | 30      | 60      |         |         |         |         |
| 40           | $P_{\beta}^2$            | 0.98875                  | 0.97889 | 0.95720 | 0.97746                   | 0.95819 | 0.91247 | 0.98213                  | 0.96699 | 0.94907 | 0.97873 | 0.95967 | 0.92405 |         |
|              | $P_{\beta_c}^2$          | 0.98814                  | 0.97775 | 0.95488 | 0.97558                   | 0.95471 | 0.90518 | 0.98009                  | 0.96322 | 0.94325 | 0.97560 | 0.95374 | 0.91288 |         |
|              | $P_{\beta^p}^2$          | 0.98818                  | 0.97784 | 0.95522 | 0.97549                   | 0.95471 | 0.90619 | 0.97984                  | 0.96333 | 0.94276 | 0.97521 | 0.95321 | 0.91233 |         |
|              | $P_{\beta^{pc}}^2$       | 0.98754                  | 0.97664 | 0.95280 | 0.97345                   | 0.95094 | 0.89837 | 0.97754                  | 0.95914 | 0.93622 | 0.97156 | 0.94633 | 0.89944 |         |
|              | $P_{\beta^{\gamma}}^2$   | 0.98874                  | 0.97886 | 0.95710 | 0.97744                   | 0.95814 | 0.91256 | 0.98212                  | 0.96695 | 0.94900 | 0.97872 | 0.95967 | 0.92431 |         |
|              | $P_{\beta^{\gamma c}}^2$ | 0.92664                  | 0.86788 | 0.75203 | 0.88027                   | 0.79271 | 0.56838 | 0.90569                  | 0.83951 | 0.76087 | 0.87803 | 0.78021 | 0.61974 |         |
|              | $R_{FC}^2$               | 0.92268                  | 0.86074 | 0.73863 | 0.87030                   | 0.77543 | 0.53242 | 0.89491                  | 0.82117 | 0.73354 | 0.86009 | 0.74788 | 0.56382 |         |
|              | $R_{FC_c}^2$             | 0.93712                  | 0.88439 | 0.77819 | 0.88737                   | 0.80260 | 0.62919 | 0.91627                  | 0.84863 | 0.77412 | 0.88580 | 0.79513 | 0.64830 |         |
|              | $R_{LR}^2$               | 0.93299                  | 0.87681 | 0.76365 | 0.87521                   | 0.78129 | 0.58916 | 0.90339                  | 0.82535 | 0.73937 | 0.86254 | 0.75339 | 0.57666 |         |
|              | $\hat{\phi}$             | 14.546                   | 7.4528  | 3.3035  | 13.268                    | 6.8171  | 2.7221  | 13.088                   | 6.67    | 4.046   | 13.733  | 6.7895  | 3.0677  |         |
|              | 80                       | $P_{\beta}^2$            | 0.98811 | 0.97747 | 0.95303                   | 0.97579 | 0.95486 | 0.90195                  | 0.98051 | 0.96355 | 0.94302 | 0.97652 | 0.95475 | 0.91188 |
|              |                          | $P_{\beta_c}^2$          | 0.98780 | 0.97688 | 0.95181                   | 0.97483 | 0.95308 | 0.89808                  | 0.97947 | 0.96161 | 0.93998 | 0.97494 | 0.95169 | 0.90592 |
|              |                          | $P_{\beta^p}^2$          | 0.98782 | 0.97694 | 0.95200                   | 0.97487 | 0.95321 | 0.89882                  | 0.97942 | 0.96178 | 0.94002 | 0.97481 | 0.95155 | 0.90587 |
|              |                          | $P_{\beta^{pc}}^2$       | 0.98750 | 0.97634 | 0.95076                   | 0.97387 | 0.95137 | 0.89482                  | 0.97832 | 0.95974 | 0.93682 | 0.97311 | 0.94828 | 0.89951 |
|              |                          | $P_{\beta^{\gamma}}^2$   | 0.98809 | 0.97743 | 0.95287                   | 0.97577 | 0.95479 | 0.90176                  | 0.98049 | 0.96350 | 0.94290 | 0.97651 | 0.95471 | 0.91191 |
|              |                          | $P_{\beta^{\gamma c}}^2$ | 0.98778 | 0.97684 | 0.95165                   | 0.97481 | 0.95301 | 0.89788                  | 0.97945 | 0.96155 | 0.93985 | 0.97492 | 0.95165 | 0.90596 |
| $R_{FC}^2$   |                          | 0.92503                  | 0.86465 | 0.74409 | 0.87641                   | 0.78663 | 0.55214 | 0.90134                  | 0.83203 | 0.75005 | 0.86937 | 0.76424 | 0.59194 |         |
| $R_{FC_c}^2$ |                          | 0.92308                  | 0.86113 | 0.73745 | 0.87154                   | 0.77821 | 0.53446 | 0.89608                  | 0.82307 | 0.73672 | 0.86054 | 0.74831 | 0.56437 |         |
| $R_{LR}^2$   |                          | 0.93570                  | 0.88162 | 0.77108 | 0.88319                   | 0.79595 | 0.61622 | 0.91197                  | 0.84176 | 0.76175 | 0.87753 | 0.78043 | 0.62324 |         |
| $R_{LR_c}^2$ |                          | 0.93369                  | 0.87791 | 0.76391 | 0.87729                   | 0.78564 | 0.59683 | 0.90576                  | 0.83061 | 0.74496 | 0.86637 | 0.76041 | 0.58890 |         |
| $\hat{\phi}$ |                          | 13.676                   | 6.8416  | 3.0741  | 12.387                    | 6.3484  | 2.5307  | 11.911                   | 6.121   | 3.6688  | 12.271  | 6.0435  | 2.7387  |         |
| 120          |                          | $P_{\beta}^2$            | 0.98782 | 0.97684 | 0.95165                   | 0.97519 | 0.95339 | 0.89737                  | 0.97990 | 0.96222 | 0.94058 | 0.97570 | 0.95283 | 0.90653 |
|              |                          | $P_{\beta_c}^2$          | 0.98761 | 0.97644 | 0.95083                   | 0.97454 | 0.95218 | 0.89472                  | 0.97920 | 0.96090 | 0.93851 | 0.97463 | 0.95076 | 0.90243 |
|              |                          | $P_{\beta^p}^2$          | 0.98762 | 0.97648 | 0.95096                   | 0.97458 | 0.95229 | 0.89526                  | 0.97918 | 0.96103 | 0.93857 | 0.97456 | 0.95068 | 0.90241 |
|              |                          | $P_{\beta^{pc}}^2$       | 0.98741 | 0.97608 | 0.95012                   | 0.97392 | 0.95106 | 0.89255                  | 0.97846 | 0.95968 | 0.93643 | 0.97344 | 0.94852 | 0.89813 |
|              |                          | $P_{\beta^{\gamma}}^2$   | 0.98781 | 0.97679 | 0.95147                   | 0.97516 | 0.95331 | 0.89705                  | 0.97988 | 0.96216 | 0.94044 | 0.97568 | 0.95278 | 0.90643 |
|              | $P_{\beta^{\gamma c}}^2$ | 0.98760                  | 0.97640 | 0.95064 | 0.97452                   | 0.95210 | 0.89439 | 0.97918                  | 0.96084 | 0.93836 | 0.97461 | 0.95071 | 0.90233 |         |
|              | $R_{FC}^2$               | 0.92422                  | 0.86328 | 0.74261 | 0.87492                   | 0.78383 | 0.54676 | 0.89966                  | 0.82938 | 0.74658 | 0.86640 | 0.75897 | 0.58273 |         |
|              | $R_{FC_c}^2$             | 0.92292                  | 0.86095 | 0.73821 | 0.87168                   | 0.77824 | 0.53503 | 0.89617                  | 0.82345 | 0.73777 | 0.86054 | 0.74840 | 0.56442 |         |
|              | $R_{LR}^2$               | 0.93493                  | 0.88045 | 0.76994 | 0.88157                   | 0.79294 | 0.61195 | 0.91035                  | 0.83908 | 0.75766 | 0.87464 | 0.77535 | 0.61453 |         |
|              | $R_{LR_c}^2$             | 0.93359                  | 0.87799 | 0.76520 | 0.87767                   | 0.78611 | 0.59915 | 0.90626                  | 0.83172 | 0.74659 | 0.86728 | 0.76216 | 0.59189 |         |
|              | $\hat{\phi}$             | 13.466                   | 6.8836  | 2.9278  | 12.118                    | 6.2483  | 2.4217  | 11.573                   | 5.9266  | 3.5091  | 11.838  | 5.8176  | 2.6167  |         |

Fonte: O autor (2020).

## 6 APLICAÇÕES

Neste capítulo realizamos uma aplicação a dados reais (Dados da Linha de Pobreza). Para avaliar a qualidade do modelo de regressão gama unitária realizamos o seguinte estudo no modelo, a saber: AIC e BIC, análise de resíduos, influência local, medidas de predição e medidas de qualidade de ajuste. Usando o chute inicial realizado no Capítulo 2 para obtenção dos parâmetros da distribuição através do método dos mínimos quadrados ordinários obtemos as estimativas para os parâmetros da regressão através do método de máxima verossimilhança (BFGS).

Nas Seções 6.1.1 e 6.1.2 estudamos os dados da Linha de Pobreza, onde usamos para o submodelo da média as funções de ligação logito e clog-log, respectivamente.

Nesta aplicação decidimos usar inicialmente a função de ligação logito, pois os dados se encontram concentrados próximos do limite superior do intervalo unitário e neste cenário, vimos com base nas simulações, que o modelo de regressão gama unitária apresenta um desempenho muito favorável independentemente da função de ligação. De fato, vimos através dos resultados de simulação que ao utilizarmos a função de ligação logito as medidas de predição apresentam valores muito próximos o que é favorável para a qualidade de ajuste do modelo

Adicionalmente, uma vez que a função de ligação logito apresenta a facilidade da interpretação de seus parâmetros, esse fato nos levou a escolher inicialmente esta função de ligação para ajustar os dados da linha de pobreza. No entanto, uma vez que também com base nos resultados de simulações, vimos que, a função de ligação Complementar Log-Log apresentava valores de  $P^2$  consideravelmente mais altos do que os demais valores relacionados as funções de ligação, o seguinte passo em relação ao ajuste de dados é considerar a função de ligação Complementar Log-Log e com base na análise de resíduos, gráficos de influência local e análise confirmatória, verificarmos qual das duas funções de ligação de fato é a mais adequada para esse conjunto de dados

### 6.1 APRESENTAÇÃO DOS DADOS

Nas Seções 6.1.1 e 6.1.2 apresentamos aplicações referentes aos dados sobre o percentual de pessoas fora da linha de pobreza no cenário mundial. Para o estudo desta variável coletamos os dados dos países do World Bank. O World Bank é um grupo formado por cinco instituições que tem como objetivo reduzir a extrema pobreza e promover a prosperidade. Neste sentido decidimos investigar como variáveis relacionadas com a educação explicam o percentual de pessoas fora da linha de pobreza de um país.

A variável resposta é o percentual de pessoas fora da linha de pobreza, complementar da variável percentual de pessoas abaixo da linha de pobreza disponibilizada pelo World Bank. O cálculo desta variável é feita a partir da quantidade de pessoas cuja a renda é inferior a U\$1.90 por dia. Segundo o World Bank esta medida reflete a profundidade da pobreza.

Consideramos o complementar desta taxa uma vez que queremos identificar fatores relacionados às políticas de educação que contribuem para que as pessoas estejam fora da linha de pobreza. Neste mesmo sentido consideramos diversas covariadas relacionadas à educação como um todo, inclusive investimentos em alta tecnologia, recurso cada vez mais presente no sistema educacional. As covariadas finais desta investigação são todas percentuais, a saber:

- $x_{t2}$  : Alta Tecnologia (Al.Tec.);
- $x_{t3}$  : Investimento em Educação Básica (Ed.Bas.);
- $x_{t5}$  : Percentual de alunos que terminaram o primário e se matricularam no secundário (Mat.Sec.),
- $x_{t6}$  : Percentual de alunos que terminaram o jardim da infância e permaneceram cursando o ensino primário (Per.Pri.).

### 6.1.1 Aplicação I: usando a função logito para o submodelo da média

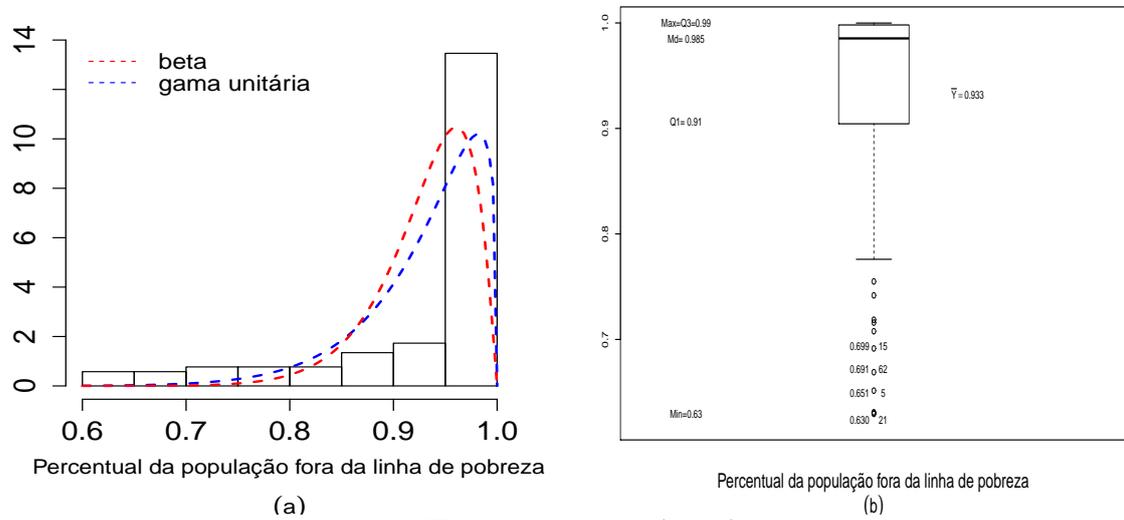
Agora iremos apresentar a análise descritiva dos dados da linha de pobreza. Na Figura 19 apresentamos o histograma e o boxplot das observações da resposta, sendo que o histograma considera os valores verdadeiros da resposta e as curvas consideram os valores ajustados com base nos modelos gama unitária e beta.

A partir desta figura notamos que a distribuição gama unitária se ajusta melhor a este conjunto de dados do que a distribuição beta. Essa assertiva é mais evidente quando  $y \in [0.7; 0.8]$  e  $y \in (0.9; 1.0)$ , uma vez que a curva ajustada gama unitária apresenta maior área de cobertura relativa aos dados verdadeiros que a curva ajustada com base na distribuição beta; especialmente onde ocorre a maior concentração de dados, a saber: o intervalo (0.9; 1.0).

O boxplot, Figura 19 (b), enfatiza a forte assimetria negativa da variável resposta, já evidenciada pelo histograma. Adicionalmente, é possível identificarmos diversas observações outliers, todas inferiores a 0.75 revelando que essa variável resposta se caracteriza por apresentar valores próximos do limite superior do intervalo unitário, de fato, tem-se que a mediana da resposta é igual 0.985 e a média é igual 0.933.

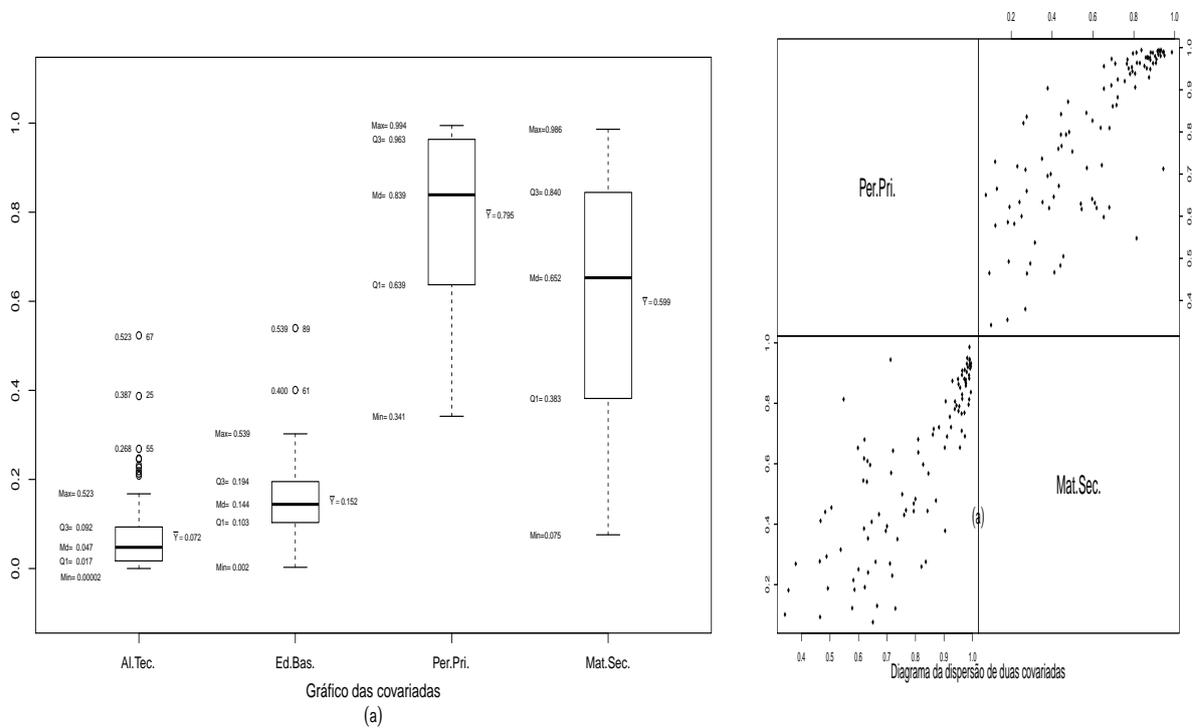
Na Figura 20 (a) apresentamos os boxplots dos valores das covariadas. Com base nessas figuras notamos que os percentuais de Investimento em Alta Tecnologia e em Educação Básica são tipicamente pequenos, tal que, percentuais próximos a 50% já são considerados

Figura 19 – Gráficos de histograma e boxplot da resposta. Modelos gama unitária com precisão variável. Dados da Linha de Pobreza.



Fonte: O autor (2020).

Figura 20 – Gráficos de boxplot e dispersão das covariadas. Dados da Linha de Pobreza.



Fonte: O autor (2020).

outliers. Enquanto que os valores dos percentuais das covariadas Persistência no Primário e Matrícula no Secundário são altos, ocorrendo inclusive percentuais acima de 90%.

Na Figura 20 (b) apresentamos o gráfico de dispersão entre as covariadas Persistência no Primário e Matrícula no Secundário. Esse gráfico evidencia como essas duas candidatas a covariadas são correlacionadas e desta forma apenas uma poderia ser considerada em cada um dos submodelos, o da média e o da precisão.

Investigamos extensivamente diversos candidatos a preditores lineares para o submodelo da média. Escolhemos entre os candidatos aquele que apresentou a análise de resíduos, de influência e as mediadas de seleção de modelos AIC e BIC mais adequadas. Salientamos que nesta primeira aplicação não vamos aplicar as medidas de predição o que será realizado na segunda aplicação. Assim, o modelo final a ser postulado como verdadeiro foi o já mencionado na Expressão 6.1.

Inicialmente ajustamos um modelo de regressão gama unitária com precisão fixa usando esses dados, em que a variável resposta é o percentual da população fora da linha de pobreza ( $y$ ). Usamos o seguinte modelo:

$$M1 : \log\left(\frac{\mu_t}{1 - \mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t6} \quad \text{e} \quad \log(\phi_t) = \gamma_1,$$

com  $t = 1, 2, \dots, 104$ . Na Tabela 35 constam as estimativas, erros-padrão e os  $p$ -valores associados aos teste de hipóteses do modelo  $M1$  para os dados da linha de pobreza.

Tabela 35 – Estimativas, erros-padrão e  $p$ -valores. Modelo gama unitária com precisão fixa. Dados da Linha de Pobreza.

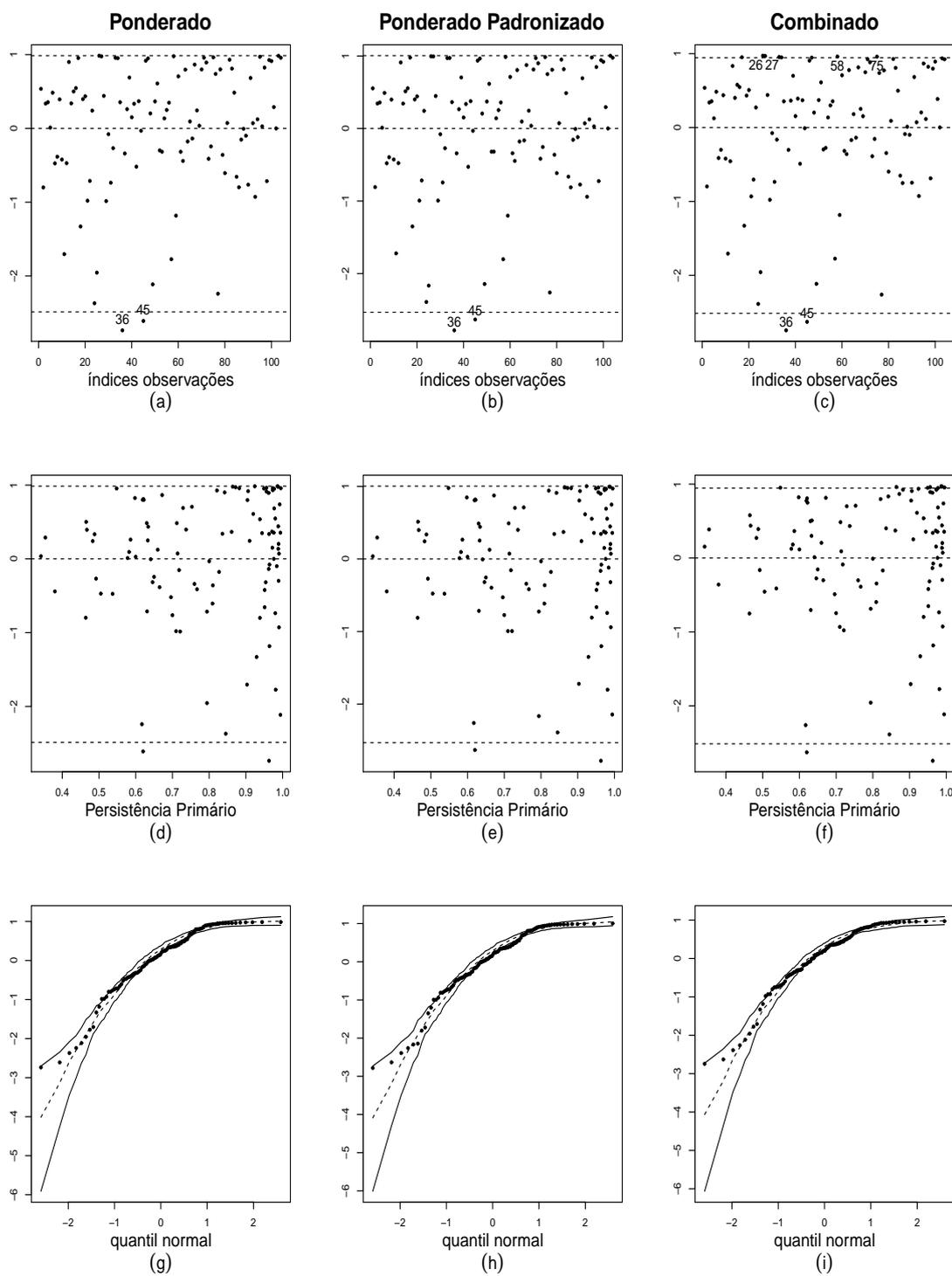
| Modelo       | Gama Unitária |           |           |           |            |
|--------------|---------------|-----------|-----------|-----------|------------|
|              | $\beta_1$     | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\gamma_1$ |
| Descrição    | Const         | Al.Tec.   | Ed.Bas.   | Mat.Sec.  | Const.     |
| Estimativa   | -0.5690       | 4.9790    | 5.3928    | 5.4238    | 0.0074     |
| Erros-Padrão | 0.2638        | 1.2452    | 1.5072    | 0.4463    | 0.1221     |
| $p$ -valor   | 0.0310        | 0.0001    | 0.0003    | < 0.0000  |            |

**Fonte:** O autor (2020).

Podemos notar com base na Tabela 35 que todos os parâmetros são significativos ao nível 5%. Em seguida realizaremos a análise de resíduos do modelo estimado apresentado na Tabela 35.

Na Figura 21 apresentaremos os gráficos que se mostraram mais relevantes, a saber: os gráficos dos resíduos contra os índices das observações e contra os valores da covariada Persistência no Primário e os gráficos normais de probabilidade com envelope simulados. Nesta figura foram utilizados o resíduo ponderado, o resíduo ponderado padronizado e o resíduo combinado. A análise da Figura 21 revela que o comportamento dos três resíduos é muito semelhante. Adicionalmente, nota-se com base nos gráficos normais de probabilidades com envelopes simulados, Figura 21 (g), (h) e (i), que os resíduos se distribuem com forte assimetria negativa, resultado compatível com os estudos de simulação, uma vez que  $\hat{\phi} = 1.007$  e no cenário em que  $\phi = 1.5$  esse tipo de simetria foi evidenciada.

Figura 21 – Gráficos dos resíduos, covariada e envelopes. Modelo gama unitária com precisão fixa. Dados da Linha de Pobreza.



Fonte: O autor (2020).

Outra importante constatação evidenciada a partir da Figura 21 é que os resíduos não se encontram aleatoriamente distribuídos dentro das bandas dos envelopes e o padrão apresentado sugere a necessidade da modelagem da precisão e ainda, nos painéis (a), (b) e (c), da Figura 21, destacamos as observações aberrantes (mal ajustadas) e notamos que nem todas as observações estão aleatoriamente distribuídas em torno do zero.

Essa assertiva é composta pelo padrão apresentado nos gráficos dos resíduos contra os valores da covariada Persistência no Primário, Figura 21 (d), (e) e (f). Com relação a esta covariada nota-se que entre os valores 0.4 e 0.6 aproximadamente a dispersão dos resíduos é consideravelmente menor que a dispersão dos resíduos entre 0.6 e 1.0. Esses gráficos sugerem que a precisão dos dados apresenta forte relação com essa covariada e, portanto, a mesma deve ser considerada no modelo de precisão. Ressaltamos que vários outros gráficos foram analisados, no entanto, como citado anteriormente, apenas os mais relevantes foram apresentados. Observemos ainda, que as medidas de AIC e BIC obtidas do modelo  $M1$  foram  $-601.56$  e  $-588.33$ , respectivamente.

Diante das evidências de precisão não constante decidimos investigar diversos modelos de regressão gama com precisão variável. Utilizamos os critérios AIC e BIC, análise de resíduos e análise de influencia e o modelo selecionado foi:

$$M2 : \log\left(\frac{\mu_t}{1 - \mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t4} + \beta_4 x_{t6}$$

$$\log(\phi_t) = \gamma_1 + \gamma_2 x_{t5}, \quad t = 1, 2, \dots, 104. \quad (6.1)$$

A partir de agora, iremos comparar os modelos de regressão gama unitária e beta com precisão variável. Em ambas regressões usaremos o modelo  $M2$ . Ao compararmos os modelos beta e gama unitária com relação as medidas de AIC e BIC temos indícios que o modelo gama unitária é mais adequado para modelarmos os dados da linha de pobreza, pois este assume os seguintes valores de AIC e BIC,  $-611.29$  e  $-595.42$ , respectivamente, enquanto que o modelo beta possui os seguintes valores de AIC e BIC,  $-570.97$  e  $-555.11$ , respectivamente.

Na Tabela 36 temos os valores das estimativas, erros-padrão e  $p$ -valores associados ao processo de inferência do modelo  $M2$ , para os modelos gama unitária e beta com o parâmetro de precisão variável. Esta tabela nos revela que para o modelo gama unitária todos os parâmetros são significativos a um nível superior a 1%, enquanto que na regressão beta apenas a covariada Matrícula no Secundário é considerada importante para a explicação da média da variável resposta.

Com base nos modelos gama unitária e beta estimados (Tabela 36) vamos realizar as análises de resíduos. Nas Figuras 22 e 23 temos a análise dos resíduos dos modelos gama unitária e beta, respectivamente. Observamos que o modelo gama unitária destacou menos pontos aberrantes e estes estão mais próximos dos limites usados nos gráficos dos resíduos. Os resíduos da gama unitária estão aleatoriamente distribuídos nos gráficos dos valores preditos e seus gráficos de envelope se adequaram melhor a todas as observações, mostrando assim uma boa qualidade de ajuste do modelo, ao compararmos com o modelo

Tabela 36 – Estimativas, erros-padrão e  $p$ -valores. Modelos beta e gama unitária com precisão variável. Dados da Linha Pobreza.

| Modelos      | Beta      |           |           |           |            |            | Gama Unitária |           |           |           |            |            |
|--------------|-----------|-----------|-----------|-----------|------------|------------|---------------|-----------|-----------|-----------|------------|------------|
|              | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\gamma_1$ | $\gamma_2$ | $\beta_1$     | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\gamma_1$ | $\gamma_2$ |
| Descrição    | Const     | Al.Tec.   | Ed.Bas.   | Mat.Sec.  | Const.     | Per.Pri.   | Const         | Alt.Tec.  | Ed.Bas.   | Mat.Sec.  | Const.     | Per.Pri    |
| Estimativa   | 0.295     | 0.417     | 1.710     | 4.740     | 2.184      | 1.497      | -0.703        | 5.005     | 5.712     | 5.504     | 2.218      | -2.644     |
| Erros-Padrão | 0.185     | 1.058     | 1.248     | 0.369     | 0.596      | 0.764      | 0.223         | 1.270     | 1.599     | 0.424     | 0.572      | 0.691      |
| $p$ -valor   | 0.111     | 0.693     | 0.170     | < 0.000   | < 0.000    | 0.050      | 0.001         | < 0.000   | < 0.000   | < 0.000   | < 0.000    | < 0.000    |

**Fonte:** O autor (2020).

beta.

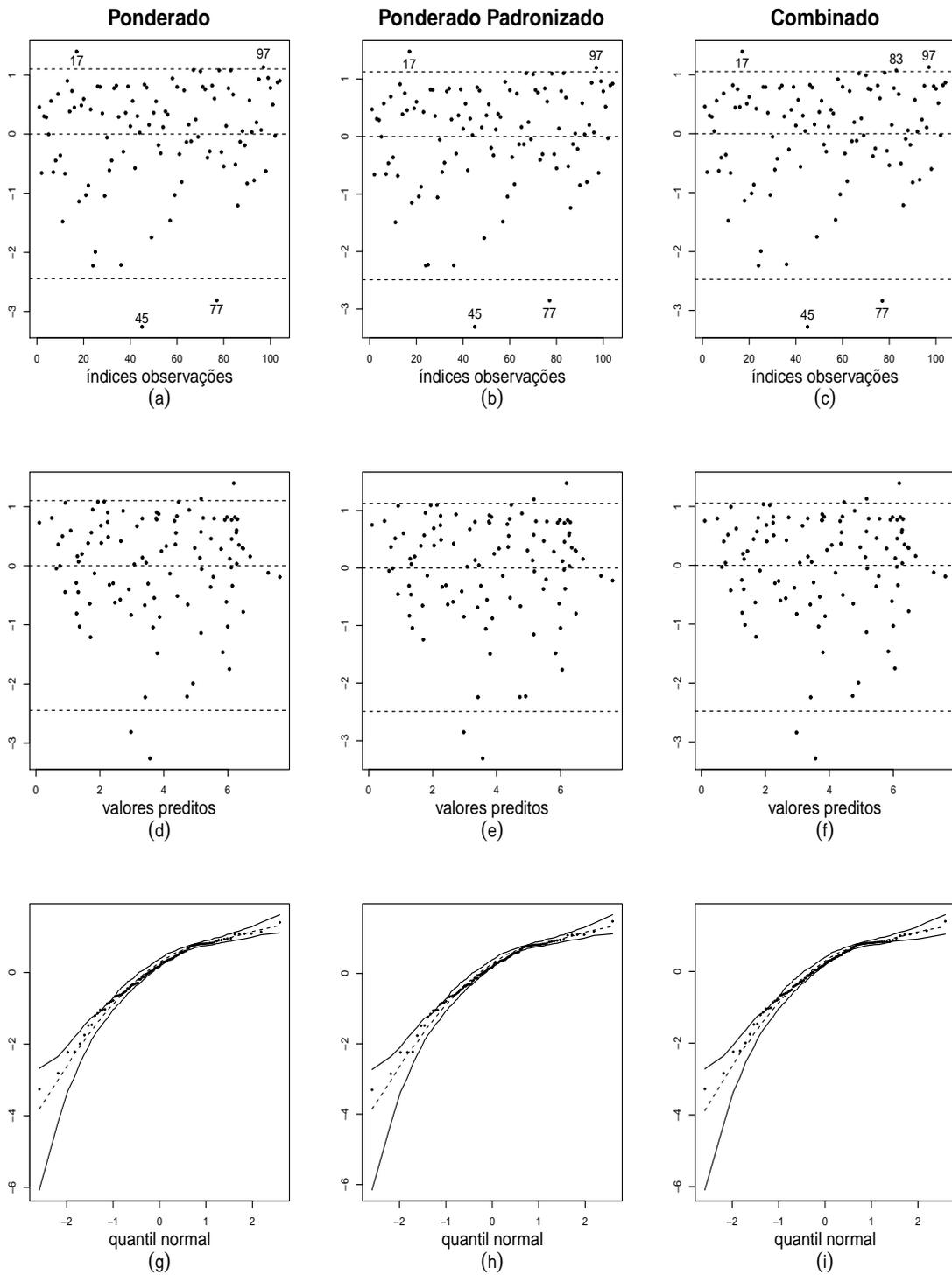
De fato, no ajuste beta percebe-se mais pontos aberrantes os quais estão mais distantes dos limites usados nos gráficos de resíduos. Adicionalmente, os gráficos dos resíduos versus os valores preditos apresenta uma tendência notável de acúmulo de observações na margem direita inferior, sugerindo falta de aleatoriedade da distribuição dos resíduos. Finalmente, os gráficos de envelope destacam muitas observações mal ajustadas, mostrando assim a má qualidade de ajuste do modelo, ou seja, a falta de ajuste da distribuição beta aos dados.

Os gráficos de envelopes das Figuras 22 e 23 evidenciam a superioridade do ajuste do modelo de regressão gama unitária em relação ao modelo de regressão beta a este conjunto de dados, em particular. As Figuras 23 (a) e (c) revelam em que particular duas observações são muito mal ajustadas a partir da regressão beta, a saber: os casos 95 e 99. Adicionalmente, no gráfico normal de probabilidades com envelopes simulados do modelo gama unitária os resíduos encontram-se todos dentro das bandas do envelope enquanto claramente para o ajuste beta os resíduos entre os respectivos quantis da distribuição normal 0 e 1 estão abaixo da banda inferior do envelope, enquanto entre os quantis  $-1$  e 0 além de alguns resíduos estarem fora do envelope nota-se uma tendência linear em sua distribuição. Desta forma, a análise de resíduos sugere fortemente que a distribuição gama unitária é mais adequada para a modelagem destes dados que a distribuição beta. Adicionalmente, as observações 13, 17, 21, 32, 45, 77, 83, 95, 97 e 99 são considerados pontos mal ajustados e devem ser investigados cuidadosamente. O que faremos após realizarmos as análises de influência local.

Nas Figuras (24 e 25) e (26 e 27) temos o estudo da análise de influência local do modelo gama unitária,  $I_{\max}$  e  $C_t$ , respectivamente. Destacamos os pontos conjuntamente e individualmente influentes, respectivamente. De modo análogo, estudamos o modelo beta, nas Figuras (28 e 29) e (30 e 31).

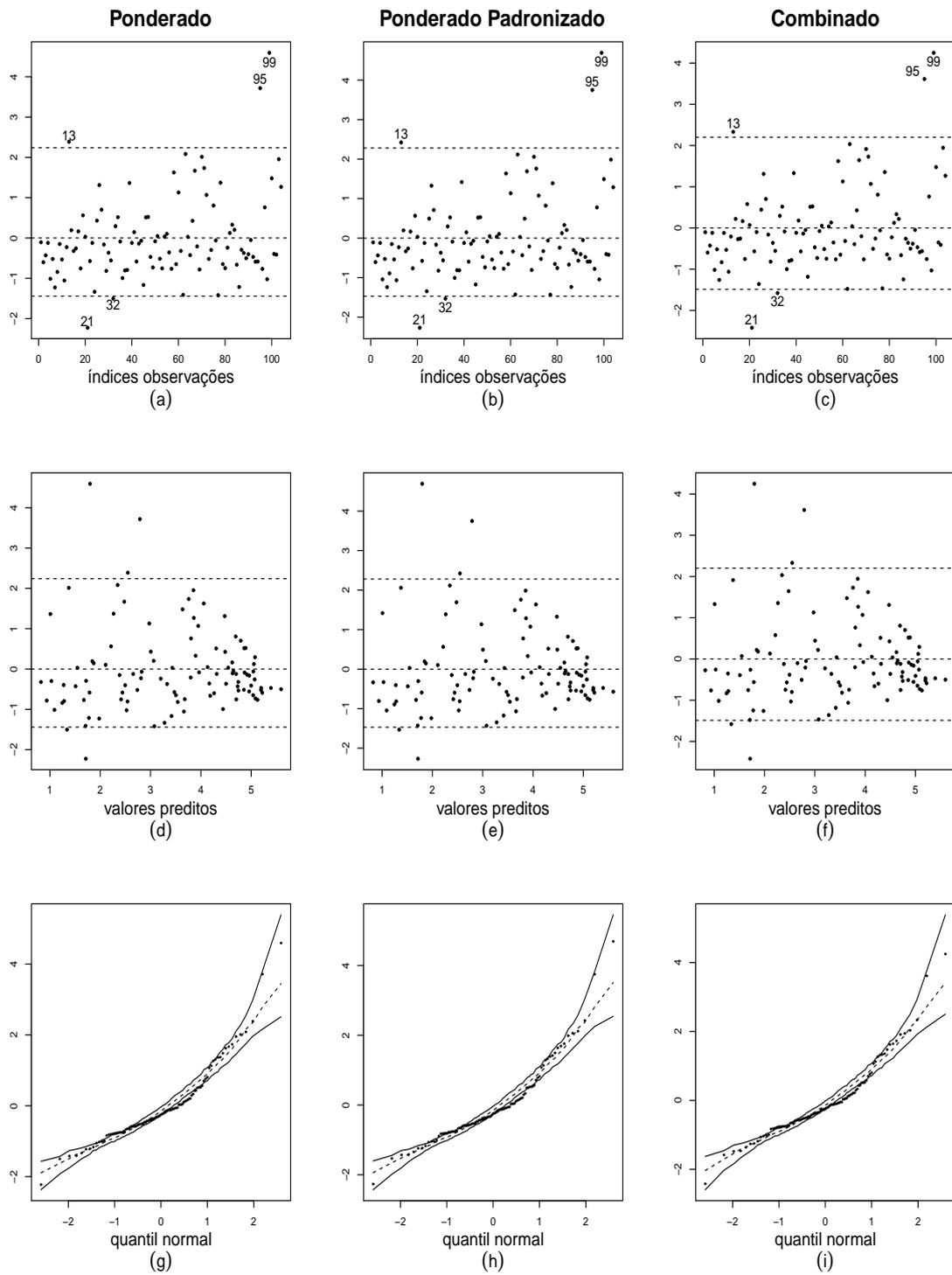
Nesta aplicação todas as covariadas podem ser perturbadas sendo que três covariadas pertencem apenas ao modelo da média: Alta Tecnologia, Investimento em Educação Básica e Matrícula no Secundário e a outra covariada pertence apenas ao modelo da precisão: Permanência no Primário. Para todos os esquemas de perturbação foram construídos gráficos para avaliar a influência das observações sobre as EMV (estimativas de máxima verossimilhança) do vetor  $\theta = (\beta^\top, \gamma^\top)^\top$  e das partições  $\beta$  e  $\gamma$ .

Figura 22 – Gráficos dos resíduos e envelopes. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza.



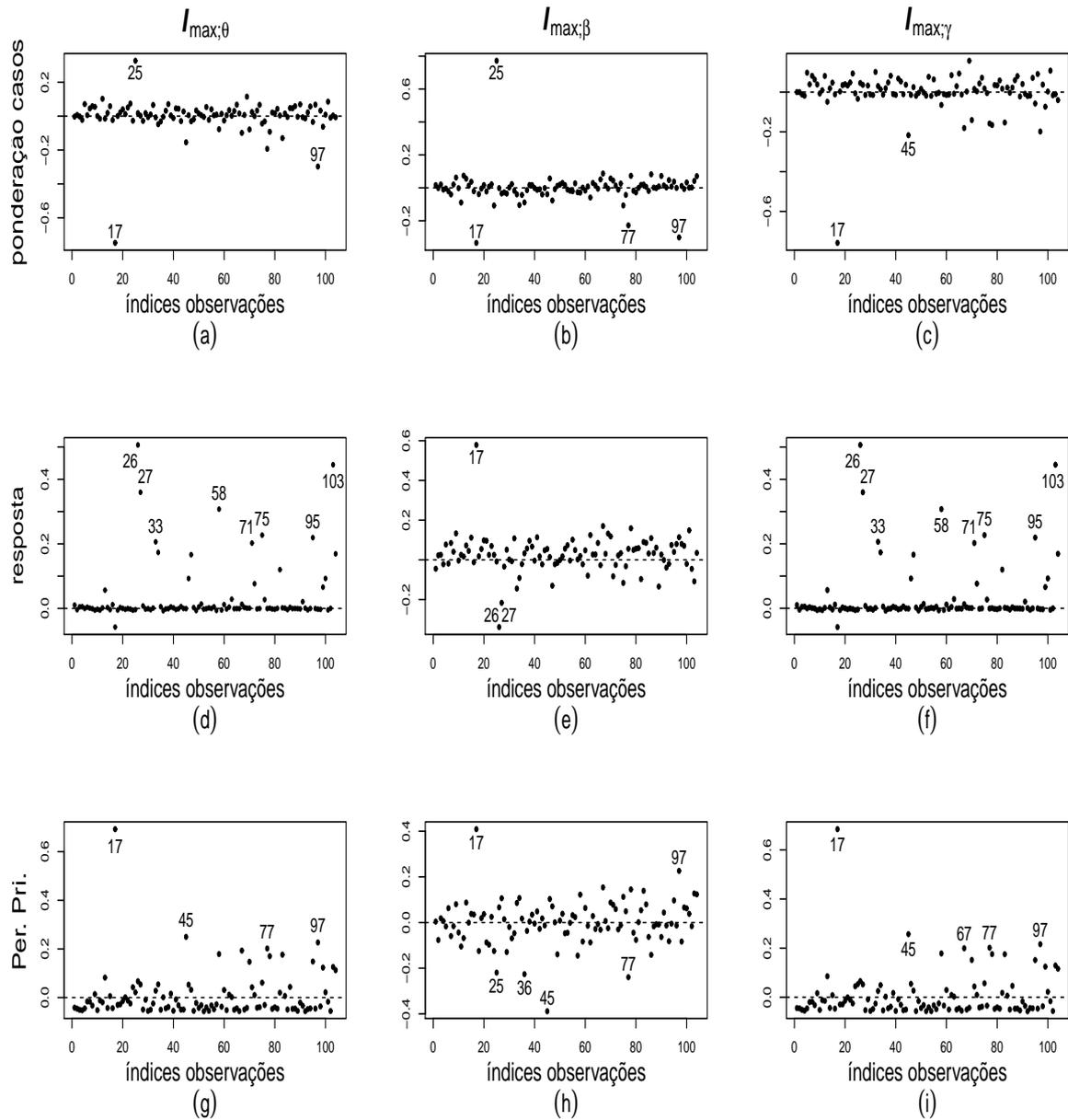
Fonte: O autor (2020).

Figura 23 – Gráficos dos resíduos e envelopes. Modelo beta com dispersão variável. Dados da Linha de Pobreza.



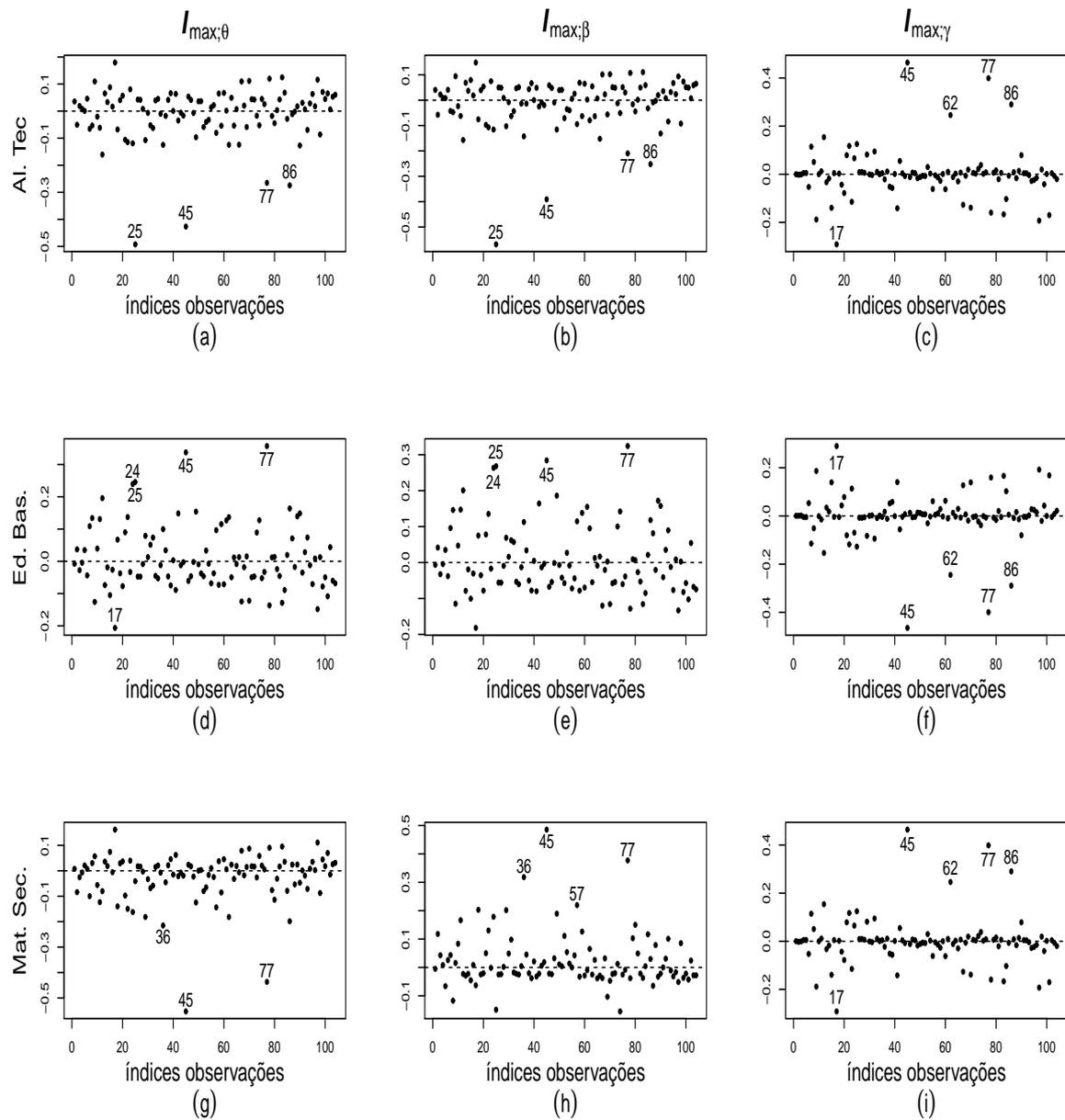
Fonte: O autor (2020).

Figura 24 – Gráficos de  $I_{\max}$  dos índices das observações versus a ponderação de casos, resposta e Persistência no Primário. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza.



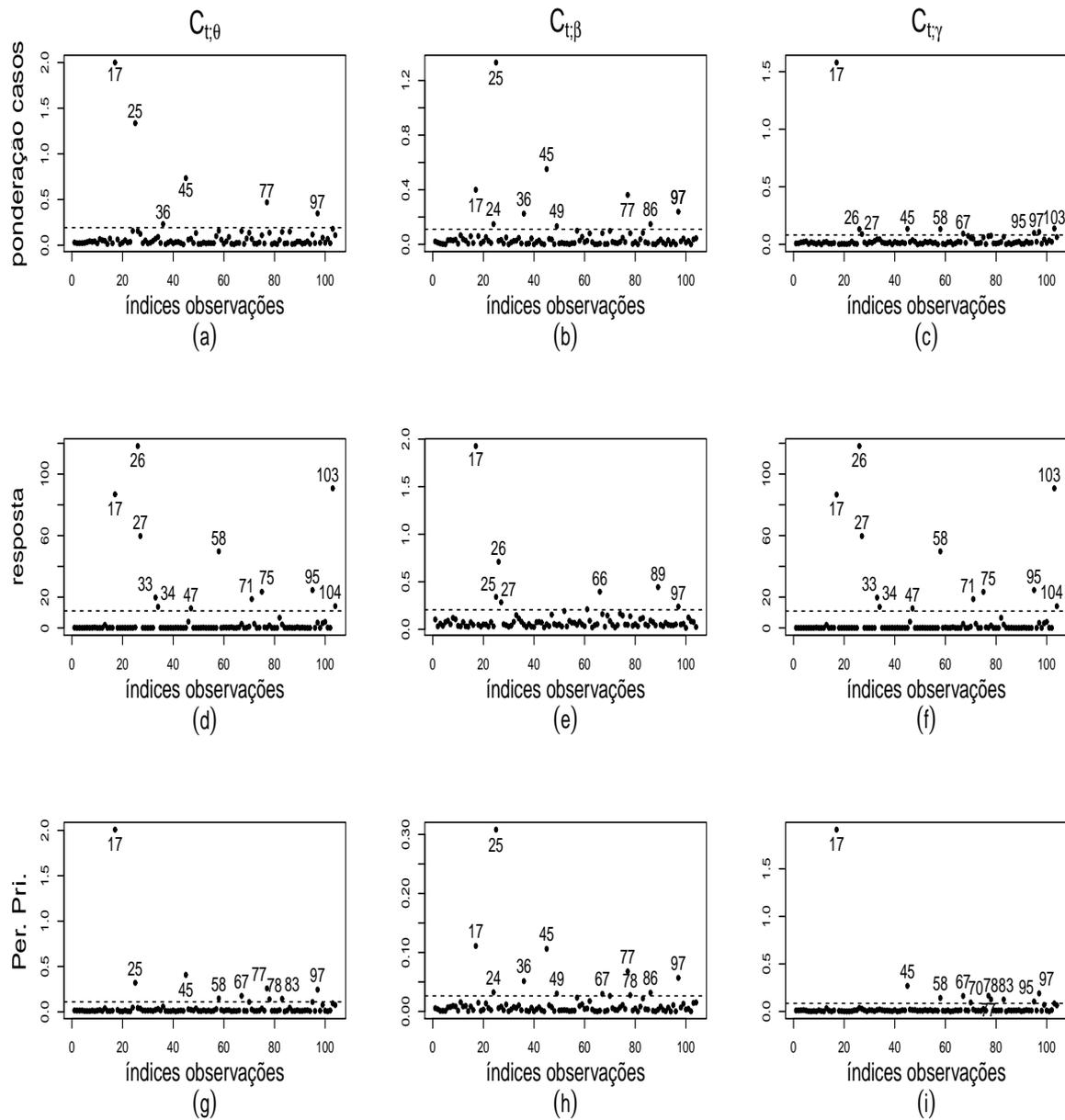
Fonte: O autor (2020).

Figura 25 – Gráficos de  $I_{\max}$  dos índices das observações versus as covariadas usadas no submodelo da média. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza.



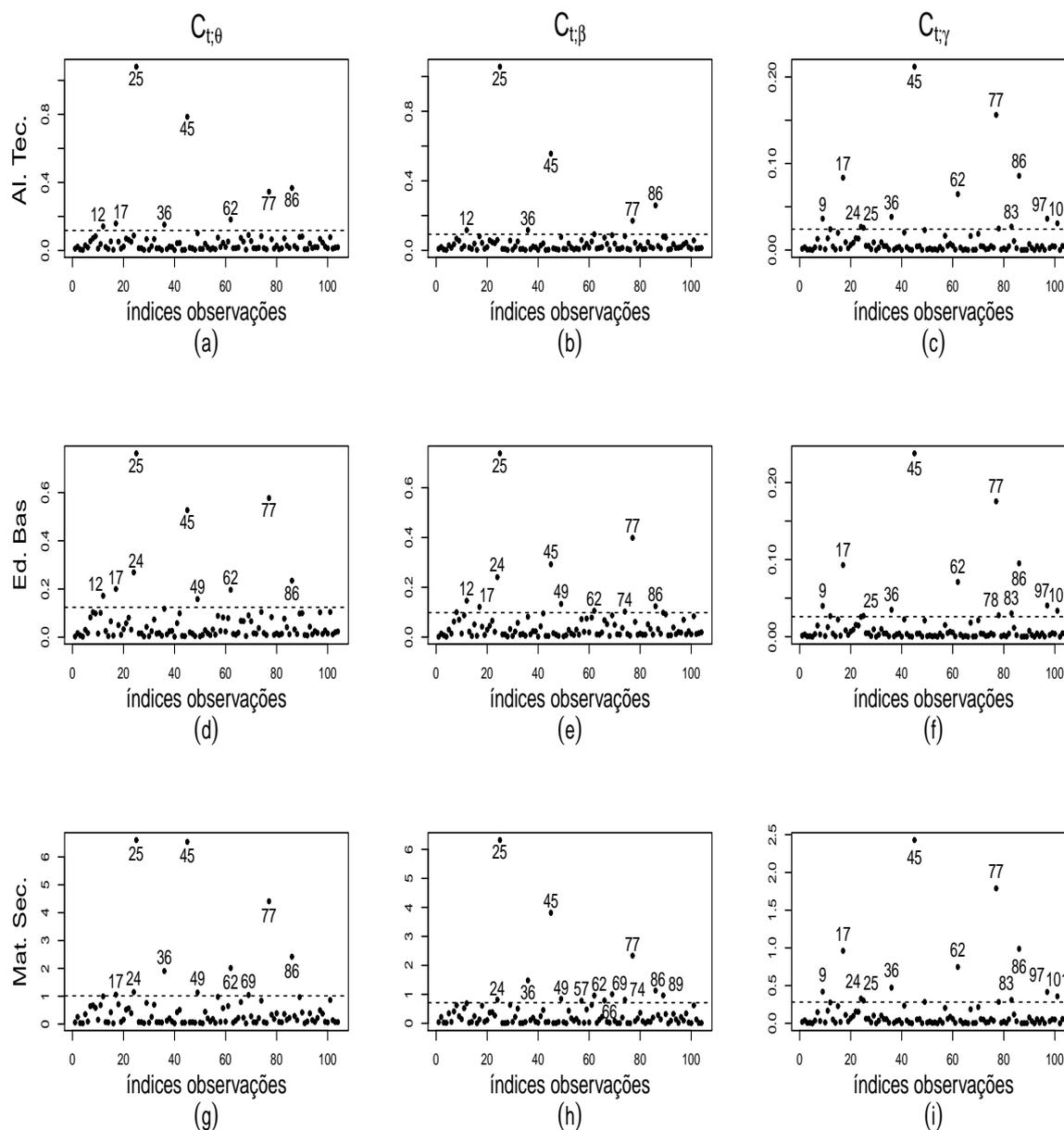
Fonte: O autor (2020).

Figura 26 – Gráficos de  $C_t$  dos índices das observações versus a ponderação de casos, resposta e Persistência no Primário. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza.



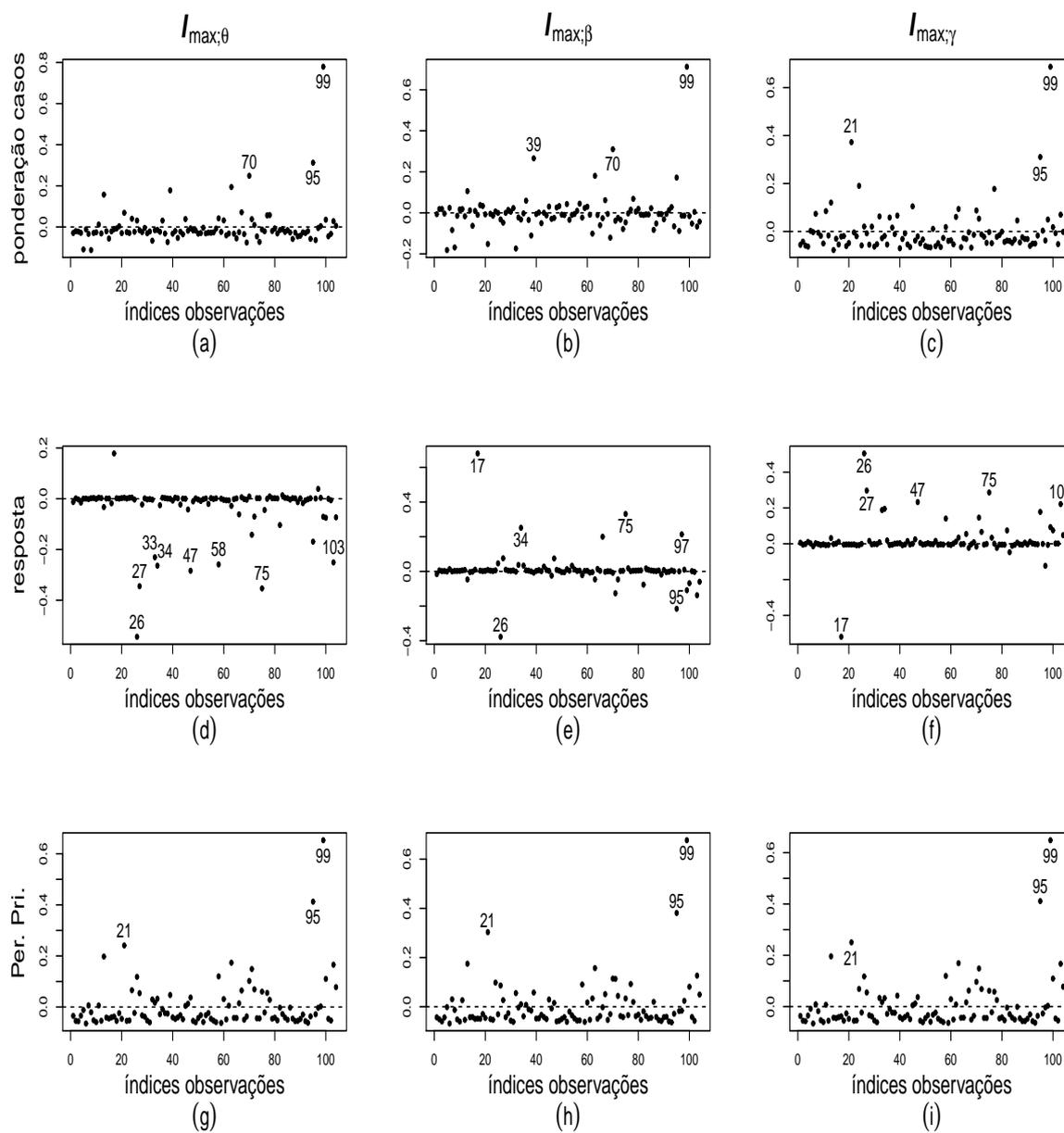
Fonte: O autor (2020).

Figura 27 – Gráficos de  $C_t$  dos índices das observações versus as covariadas usadas no submodelo da média. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza.



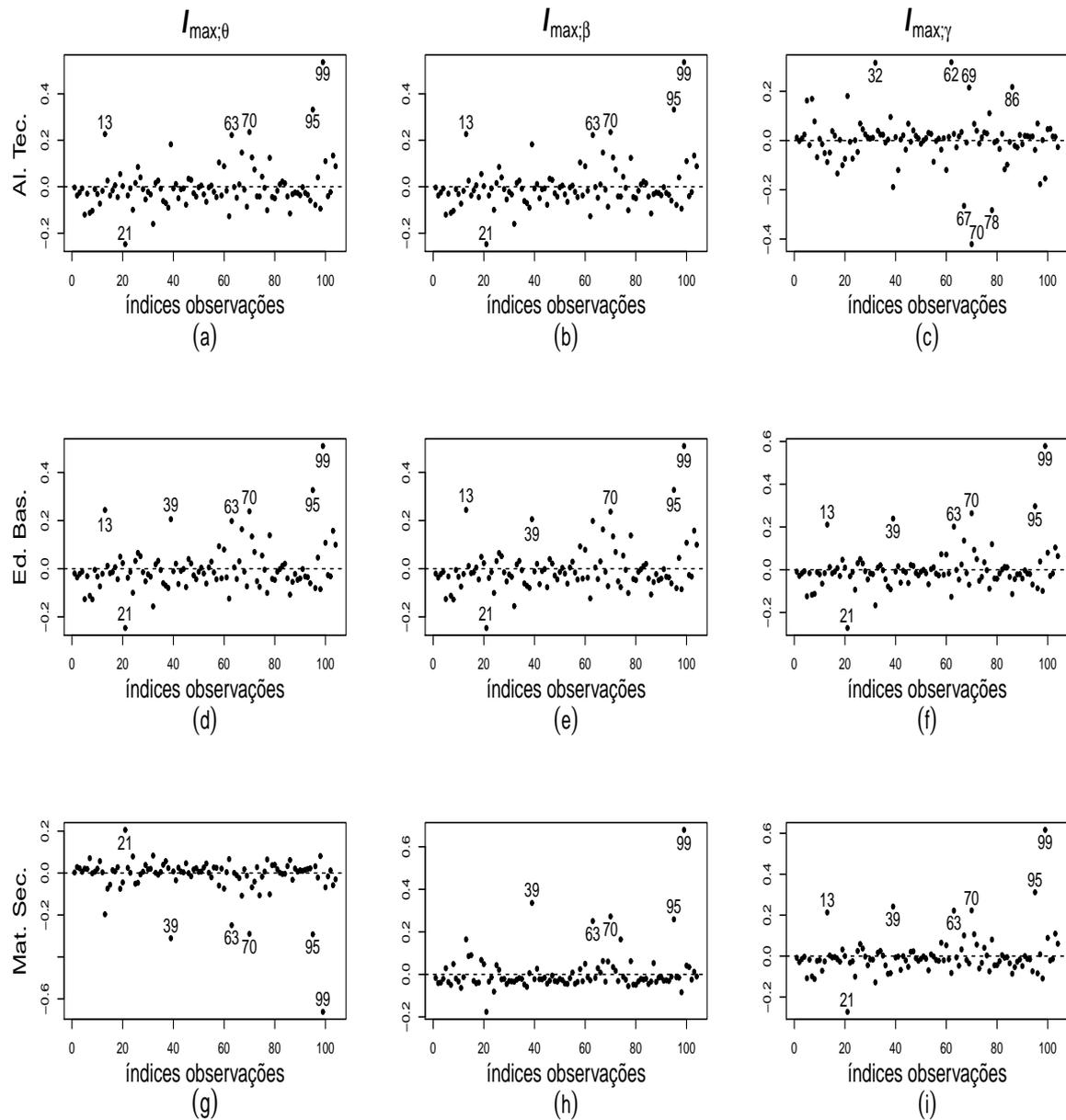
Fonte: O autor (2020).

Figura 28 – Gráficos de  $I_{\max}$  dos índices das observações versus a ponderação de casos, resposta e Persistência no Primário. Modelo beta com precisão variável. Dados da Linha de Pobreza.



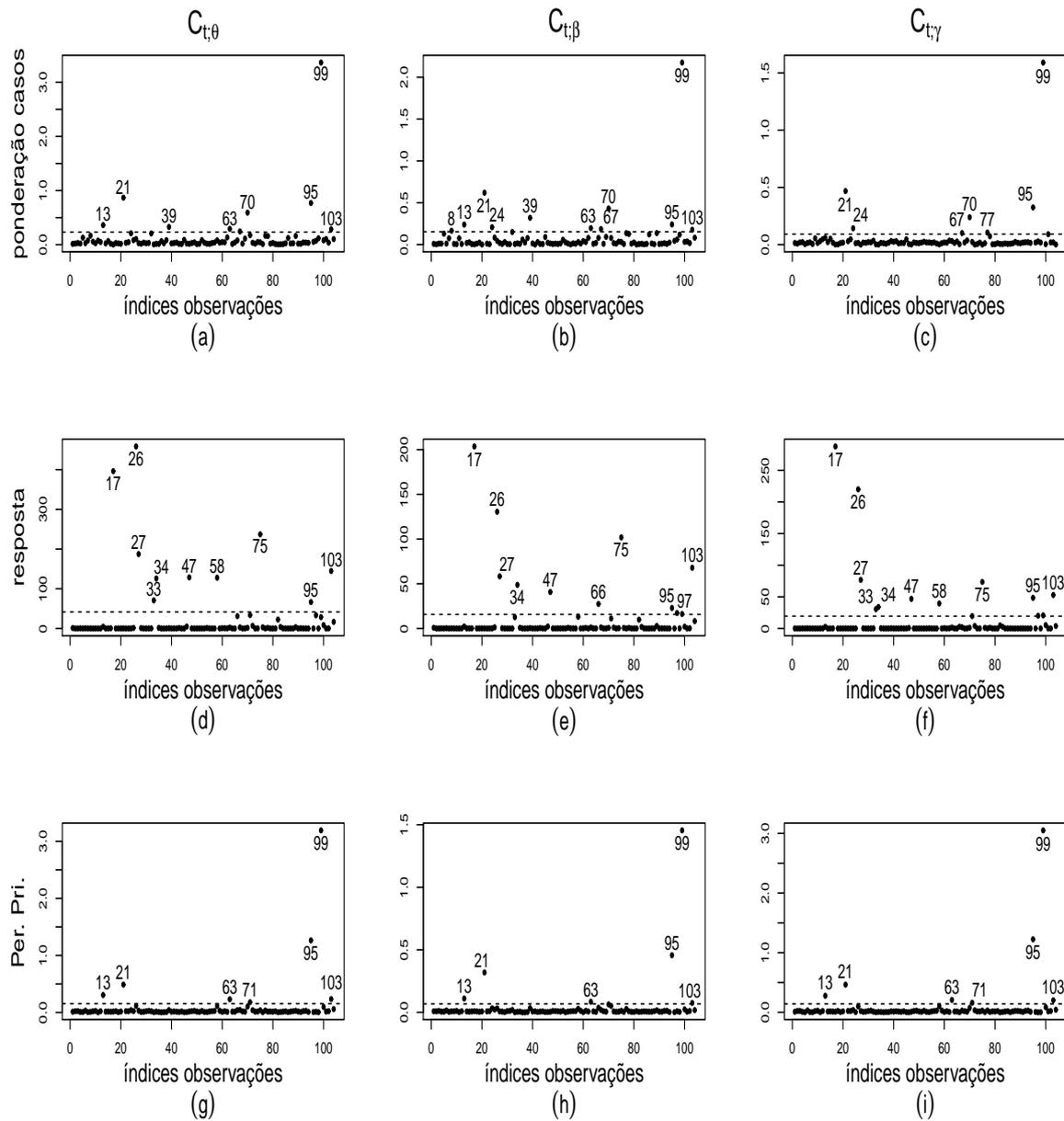
Fonte: O autor (2020).

Figura 29 – Gráficos de  $I_{\max}$  dos índices das observações versus as covariadas usadas no submodelo da média. Modelo beta com precisão variável. Dados da Linha de Pobreza.



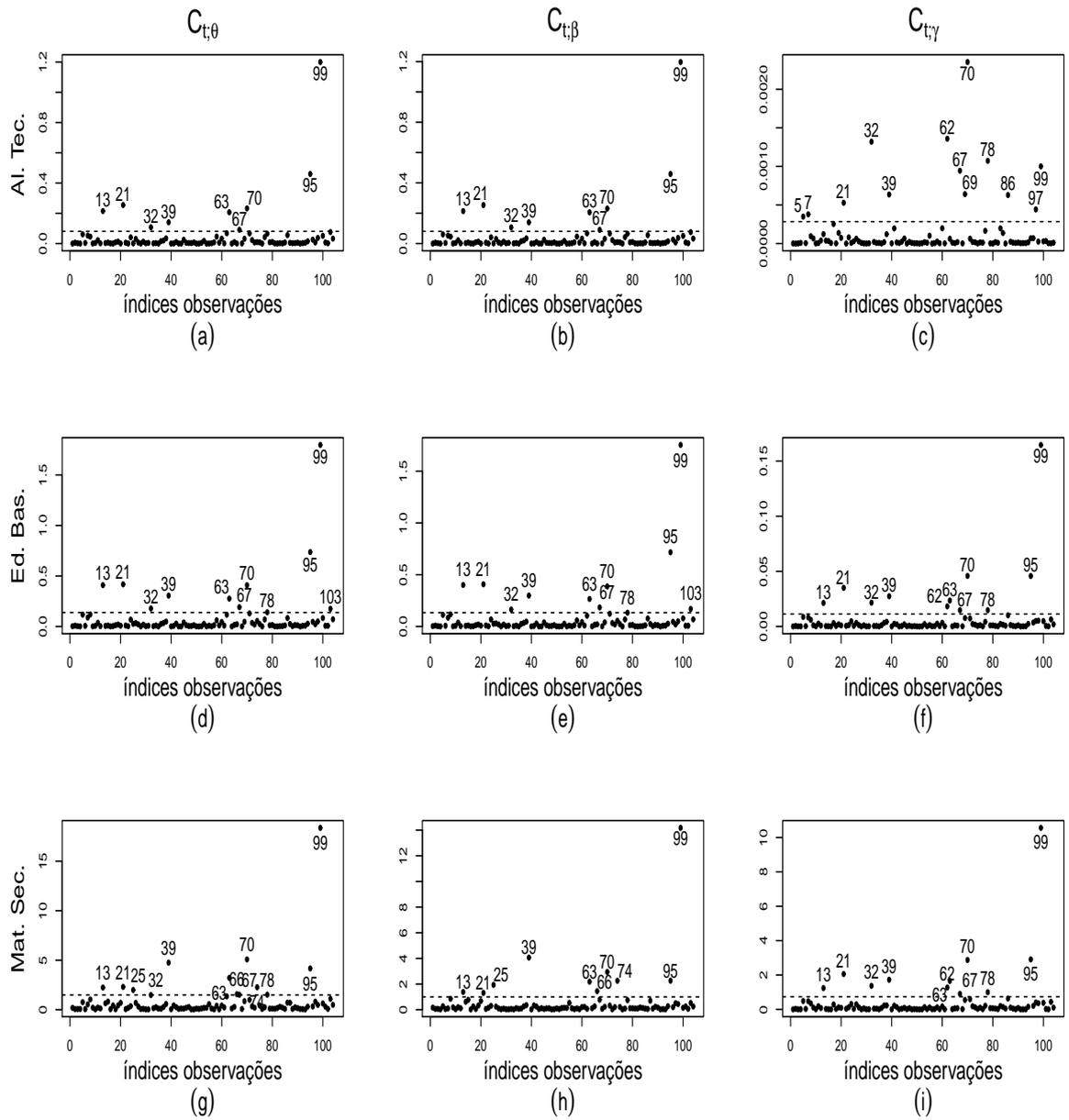
Fonte: O autor (2020).

Figura 30 – Gráficos de  $C_t$  dos índices das observações versus a ponderação de casos, resposta e Persistência no Primário. Modelo beta com precisão variável. Dados da Linha de Pobreza.



Fonte: O autor (2020).

Figura 31 – Gráficos de  $C_t$  dos índices das observações versus as covariadas usadas no submodelo da média. Modelo beta com precisão variável. Dados da Linha de Pobreza.



Fonte: O autor (2020).

Tabela 37 – Estimativas, erro-padrão e  $p$ -valor. Observações destacadas no resíduo e influência local. Modelo beta. Dados da Linha de Pobreza.

| Modelo   |                      | Beta         |              |              |              |              |              |
|--|----------------------|--------------|--------------|--------------|--------------|--------------|--------------|
| Parâmetro  |                      | $\beta_1$    | $\beta_2$    | $\beta_3$    | $\beta_4$    | $\gamma_1$   | $\gamma_2$   |
| Descrição  |                      | Const        | Al.Tec       | Ed.Básica    | Mat.Sec.     | Const.       | Per.Pri.     |
| Observação<br>21<br>Deletada                       | mudança estimativa   | 21.401       | 7.066        | -16.368      | 1.423        | -0.532       | 8.911        |
|  | mudança erros-padrão | 0.255        | -0.913       | -0.680       | -1.484       | 0.153        | -0.137       |
|  | $p$ -valor           | <b>0.054</b> | <b>0.670</b> | <b>0.248</b> | <b>0.000</b> | <b>0.000</b> | <b>0.032</b> |
| Observações<br>21, 95, 99<br>Deletada              | mudança estimativa   | -29.613      | 74.509       | -46.412      | 11.811       | -4.790       | 41.702       |
|  | mudança erros-padrão | -7.162       | -5.171       | -6.399       | -8.922       | -0.036       | -0.864       |
|  | $p$ -valor           | <b>0.227</b> | <b>0.467</b> | <b>0.432</b> | <b>0.000</b> | <b>0.000</b> | <b>0.005</b> |
| Observações<br>39, 70, 99<br>Deletada              | mudança estimativa   | -105.098     | 24.113       | -12.202      | 14.713       | 16.684       | -5.398       |
|  | mudança erros-padrão | -6.278       | -5.170       | -6.825       | -6.089       | 0.983        | 0.194        |
|  | $p$ -valor           | <b>0.931</b> | <b>0.605</b> | <b>0.196</b> | <b>0.000</b> | <b>0.000</b> | <b>0.064</b> |
| Observações<br>103<br>Deletada                     | mudança estimativa   | -8.871       | 16.171       | 23.207       | -1.539       | -2.166       | 5.541        |
|  | mudança erros-padrão | 0.067        | 0.044        | 0.786        | -0.058       | 0.131        | 0.220        |
|  | $p$ -valor           | <b>0.147</b> | <b>0.646</b> | <b>0.093</b> | <b>0.000</b> | <b>0.000</b> | <b>0.039</b> |
| Observações<br>13, 63, 70, 95, 99<br>Deletada      | mudança estimativa   | -114.533     | 248.056      | -22.000      | 14.674       | 3.178        | 33.864       |
|  | mudança erros-padrão | -12.122      | -4.080       | -9.686       | -12.842      | 0.520        | -0.497       |
|  | $p$ -valor           | <b>0.792</b> | <b>0.152</b> | <b>0.236</b> | <b>0.000</b> | <b>0.000</b> | <b>0.008</b> |
| Observações<br>25, 36, 45 e 77<br>Deletada         | mudança estimativa   | 0.9506       | -100.9928    | -14.4946     | 4.7080       | -2.8660      | 11.3434      |
|  | mudança erros-padrão | -0.5702      | 12.5067      | -0.0635      | 2.5708       | 1.8545       | 1.7576       |
|  | $p$ -valor           | <b>0.106</b> | <b>0.997</b> | <b>0.241</b> | <b>0.000</b> | <b>0.000</b> | <b>0.032</b> |
| Observações<br>13, 39, 63, 70, 95 e 99<br>Deletada | mudança estimativa   | -169.4770    | 252.8880     | -0.7817      | 19.6879      | 5.7660       | 38.4918      |
|  | mudança erros-padrão | -11.5358     | -5.9223      | -11.7570     | -14.2733     | 0.8997       | -0.2934      |
|  | $p$ -valor           | <b>0.211</b> | <b>0.139</b> | <b>0.123</b> | <b>0.000</b> | <b>0.000</b> | <b>0.006</b> |

Fonte: O autor (2020).

Desta forma é possível avaliar o impacto das observações na estimação conjunto dos dois submodelo: média e precisão, (primeira coluna), apenas na estimação do submodelo da média (segunda coluna) e apenas no submodelo da precisão (terceira coluna). Vale a pena ressaltar que a exclusão dos pontos destacados como aberrantes e/ou influentes no modelo gama unitária, não afetou os  $p$ -valores relacionados aos teste sobre os parâmetros, diferentemente do que ocorre no modelo beta, o qual mostrou-se sensível aos pontos aberrantes e influentes.

Na Tabela 37 exibiremos algumas das observações que foram detectadas no estudo de influência local e resíduos no modelo beta. Ao analisarmos essa tabela observamos que ao retirarmos a observação 21, destacada tanto nos gráficos dos resíduos quanto no estudo de influência local realizado no modelo beta, o parâmetro  $\beta_1$  passa a ser diferente de zero a um nível próximo de 5%, enquanto que o parâmetro  $\beta_3$  passa a ser não significativo ao um nível superior a 24%. Assim, a exclusão do caso 21 implica considerar a covariada Investimento em educação básica não importante para a explicação do percentual de pessoas fora da linha da pobreza. Quanto ao submodelo da precisão quando o caso 21 não está no conjunto de dados o parâmetro  $\gamma_2$  passa a ser significativo ( $p$ -valor = 0.032). A exclusão do conjunto {21, 95 e 99} destacado nos gráficos de ponderação de casos no modelo beta causa um efeito ainda mais impressionante, a não significância de todos os parâmetros do submodelo da média.

Ao tirarmos o conjunto  $\{13, 63, 70, 95 \text{ e } 99\}$  destacado pelo gráfico de influência local sob a perturbação individual da Alta Tecnologia, observamos uma grande redução no  $p$ -valor relacionado ao parâmetro  $\beta_2$  que passa de 0.693 para 0.139. Por outro lado, a exclusão de  $\{25, 36, 45 \text{ e } 77\}$ , destacadas no modelo gama unitária ao perturbamos o parâmetro de precisão, aumenta o  $p$ -valor para 0.997. Estes fatos analisados na Tabela 37 nos mostra a grande sensibilidade do modelo beta aos pontos aberrantes e influentes.

Em um estudo mais minucioso sobre os resíduos e influência local observamos que ao retirarmos os pontos destacados como aberrantes e influentes no modelo beta, na gama unitária, esta continuou sem sofrer alterações significativas nos valores do  $p$ -valor relacionados aos seus parâmetros, mostrando assim, um modelo capaz de suportar as perturbações dos pontos aberrantes e influentes de ambos os modelos, Diferentemente do modelo beta, pois este também foi sensível aos pontos aberrantes e influentes de ambos os modelos.

Em resumo, sobre a amostra da linha de pobreza, iniciamos o nosso estudo analisando o modelo gama unitária com precisão fixo, onde estudamos as medidas de AIC e BIC, análise dos resíduos e gráficos de envelopes; observamos que o modelo mais adequado para a amostra da linha de pobreza foi o modelo gama unitária com o parâmetro de precisão variável, pois este nos forneceu os melhores valores de AIC e BIC, gráficos de resíduos, gráficos de envelopes e não foi sensível aos pontos aberrantes nem influentes.

A partir daí, começamos o nosso estudo da comparação dos modelos gama unitária e beta com precisão variável, onde analisamos as medidas de AIC e BIC, gráficos dos resíduos, gráficos dos envelopes e por fim, estudamos a sensibilidade de cada modelo através da análise de influência local, foi notável a melhor adequação do modelo gama unitária ao modelo beta, pois o modelo gama unitária se mostrou melhor em todos os estudos realizados.

Assim concluímos que o modelo gama unitária com precisão variável é o mais adequado que o modelo beta com dispersão variável, no cenário onde a média da variável resposta se encontra concentrada próxima de um. Nesta aplicação utilizamos a função de ligação logito. Na próxima aplicação vamos investigar o uso da função de ligação complementar log-log, uma vez que essa ligação apresentou nos estudos de simulação tanto as medidas de predição quanto as de qualidade de ajuste mais altas quando  $\mu \approx 1$ .

### 6.1.2 Aplicação II: usando a função cloglog para o submodelo da média

Nessa aplicação vamos investigar modelos de regressão considerando transformações para as covariadas e a função de ligação complementar log-log para o submodelo da média. Iniciamos o nosso estudo ajustando um modelo de regressão gama unitária com precisão variável usando esses dados, em que a variável resposta é o percentual da população fora da linha de pobreza ( $y$ ). Para escolhermos o modelo mais adequado para esses dados fizemos uso das medidas  $P^2$ ,  $P_c^2$ ,  $R_{FC}^2$ ,  $R_{FC_c}^2$ ,  $R_{LR}^2$  e  $R_{LR_c}^2$  apresentadas na Tabela 40 e o

modelo selecionado foi:

$$M3 : \log(-\log(1 - \mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 \log(x_{t3}) + \beta_4 \log(x_{t4}) \quad e$$

$$\log(\phi_t) = \gamma_1 + \gamma_2 \exp(x_{t5}^2),$$

com  $t = 1, 2, \dots, 104$ . Para que possamos comparar os modelos de regressão beta e gama unitária utilizamos em ambos os modelos o modelo  $M3$ . Os resultados da Tabela 40 consideram os modelos logito apresentados na seção anterior e os modelos complementar log-log definidos a partir dos preditores lineares  $M3$ .

Nossa proposta nesta aplicação é comprovar através da análise de resíduos e de influência local que os modelos selecionados com base nas medidas  $P^2$ ,  $P_c^2$ ,  $R_{FC}^2$ ,  $R_{FC_c}^2$ ,  $R_{LR}^2$  e  $R_{LR_c}^2$  de fato apresenta um bom desempenho, em especial quanto a distribuição gama unitária.

Na Tabela 38 constam as estimativas, erros-padrão e os  $p$ -valores associados aos teste de hipóteses do modelo  $M3$  para os dados da linha de pobreza do modelo de regressão gama unitária e beta, respectivamente. Note que a nível de significância de 1% todos os parâmetros são significativos do modelo gama unitária.

Tabela 38 – Estimativas, erros-padrão e  $p$ -valores. Função de ligação cloglog. Modelos beta e gama unitária com precisão variável. Dados da Linha Pobreza.

| Modelos      | gama unitária |           |                |                |            |                  | beta      |           |                |                |            |                  |
|--------------|---------------|-----------|----------------|----------------|------------|------------------|-----------|-----------|----------------|----------------|------------|------------------|
|              | $\beta_1$     | $\beta_2$ | $\beta_3$      | $\beta_4$      | $\gamma_1$ | $\gamma_2$       | $\beta_1$ | $\beta_2$ | $\beta_3$      | $\beta_4$      | $\gamma_1$ | $\gamma_2$       |
| Parâmetros   | const.        | $x_{t2}$  | $\log(x_{t3})$ | $\log(x_{t4})$ | const.     | $\exp(z_{t2}^2)$ | const.    | $x_{t2}$  | $\log(x_{t3})$ | $\log(x_{t4})$ | const.     | $\exp(z_{t2}^2)$ |
| Estimativa   | 2.142         | 1.029     | 0.238          | 0.795          | 1.870      | -0.885           | 1.718     | 0.304     | 0.061          | 0.668          | 2.131      | 0.629            |
| Erros-Padrão | 0.104         | 0.228     | 0.057          | 0.066          | 0.505      | 0.239            | 0.085     | 0.256     | 0.037          | 0.047          | 0.527      | 0.266            |
| $p$ -valor   | < 0.000       | < 0.000   | < 0.000        | < 0.000        | < 0.000    | < 0.000          | < 0.000   | 0.234     | 0.098          | < 0.000        | < 0.000    | 0.018            |

**Fonte:** O autor (2020).

Observe na Tabela 38 o parâmetro  $\beta_2$  só poderia ser considerado significativo ao nível de significância de 24%. Para entendermos melhor os valores fornecidos nesta tabela, com relação aos parâmetros  $\beta_2$  e  $\beta_3$  iremos estudar a análise de resíduos e influência local.

Nas Figuras 32 e 33 notamos que os gráficos de resíduos e envelopes se comportam de maneira semelhantes para o modelo beta e o modelo gama unitária. Nas letras (a), (b) e (c) das Figuras 32 e 33 notamos uma semelhança das observações na aleatoriedade em torno do zero e em ambas as figuras são destacadas cinco observações como aberrantes, porém na Figura 33 é possível notamos um maior distanciamento dos pontos aberrantes dos limitantes dos resíduos, indicando uma menor qualidade de ajuste do modelo.

Nos painéis (d), (e) e (f) das Figuras 32 e 33 também observamos aleatoriedade nas distribuições das observações em torno do zero, porém na Figura 33 os casos destacados como aberrantes estão mais distantes dos limitantes dos resíduos. Observando os painéis (g), (h) e (i) da Figura 32 notamos uma melhor qualidade de ajuste, pois nenhuma observação sai completamente do envelope e uma melhor aleatoriedade das observações em torno da mediana com relação as observações da Figura 33, mostrando assim uma melhor adequação da distribuição ao modelo.

A seguir relatamos o estudo de influência feito nos modelos de regressão gama unitária e beta.

Da Figura 34 à Figura 38 apresentamos os gráficos de influência em relação ao ajuste gama unitária. Enquanto, da Figura 39 à Figura 42 apresentamos os gráficos de influência em relação ao ajuste beta.

Apresentamos inicialmente os gráficos de valores de  $I_{\max}$  e  $C_t$  contra os índices das observações considerando os esquemas de ponderação de casos, perturbação da variável resposta e perturbação individual das covariadas da média e da precisão Figuras 34, 35, 36 e 37, respectivamente. Em seguida, na Figura 38 geramos os gráficos de  $I_{\max}$  obtido da perturbação conjunta de covariadas. Todos relativos com citado acima ao ajuste gama unitária. Esse mesmo script foi considerado para a construção dos gráficos do ajuste beta, Figuras 39, 40, 41 e 42.

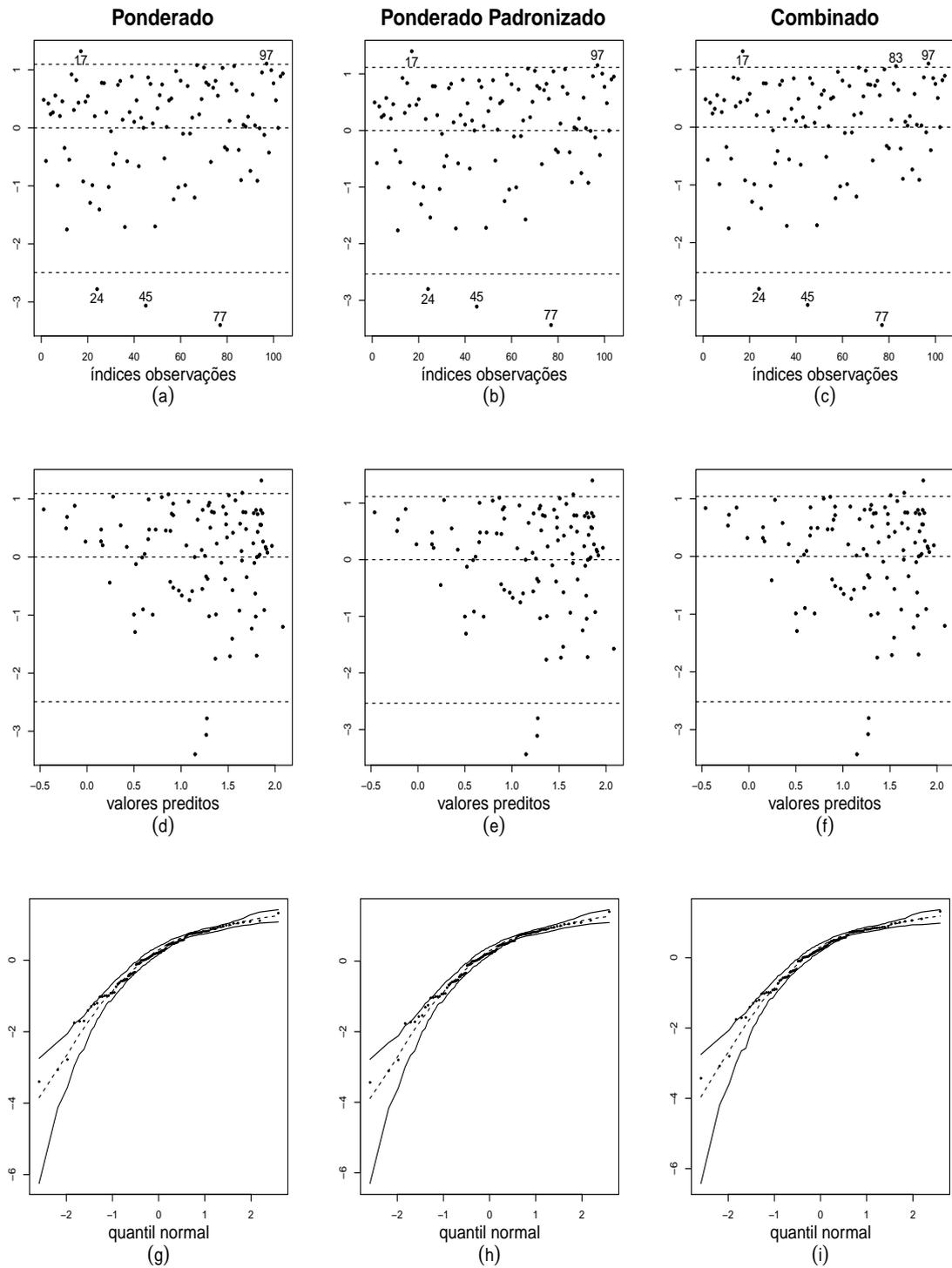
Em uma análise mais detalhada ao tirarmos os pontos destacados como aberrantes e influentes, no modelo gama unitária não foi possível observamos nenhuma alteração nos  $p$ -valores, mostrando que este ajuste não apresentou instabilidade no processo de estimação na presença dos pontos aberrantes nem possivelmente influentes, diferentemente do que ocorre no modelo beta. Na Tabela 39 exibimos algumas das alterações sofridas pelo modelo beta ao tirarmos as observações destacadas como aberrantes e possivelmente influentes.

Na Tabela 39 é possível notarmos alterações no aumento ou na redução do valor assumido pelo  $p$ -valor, ao tirarmos as observações aberrantes e influentes. Por exemplo, o parâmetro  $\gamma_2$  que tornou-se não significativo a nível de 5%, ao tirarmos conjuntamente as observações {21, 32 e 62} que foram destacadas na análise de resíduos.

Observe também que ao tirarmos conjuntamente as observações {26, 27, 71, 95, 99, 100 e 103} notamos grande mudança assumida nos  $p$ -valores associadas aos parâmetros  $\beta_2$ ,  $\beta_3$  e  $\gamma_2$ , pois obtemos todos os parâmetros significativos a nível de 4%. Dessa forma, notamos que o processo de estimação do modelo beta é sensível a pontos aberrantes e influentes.

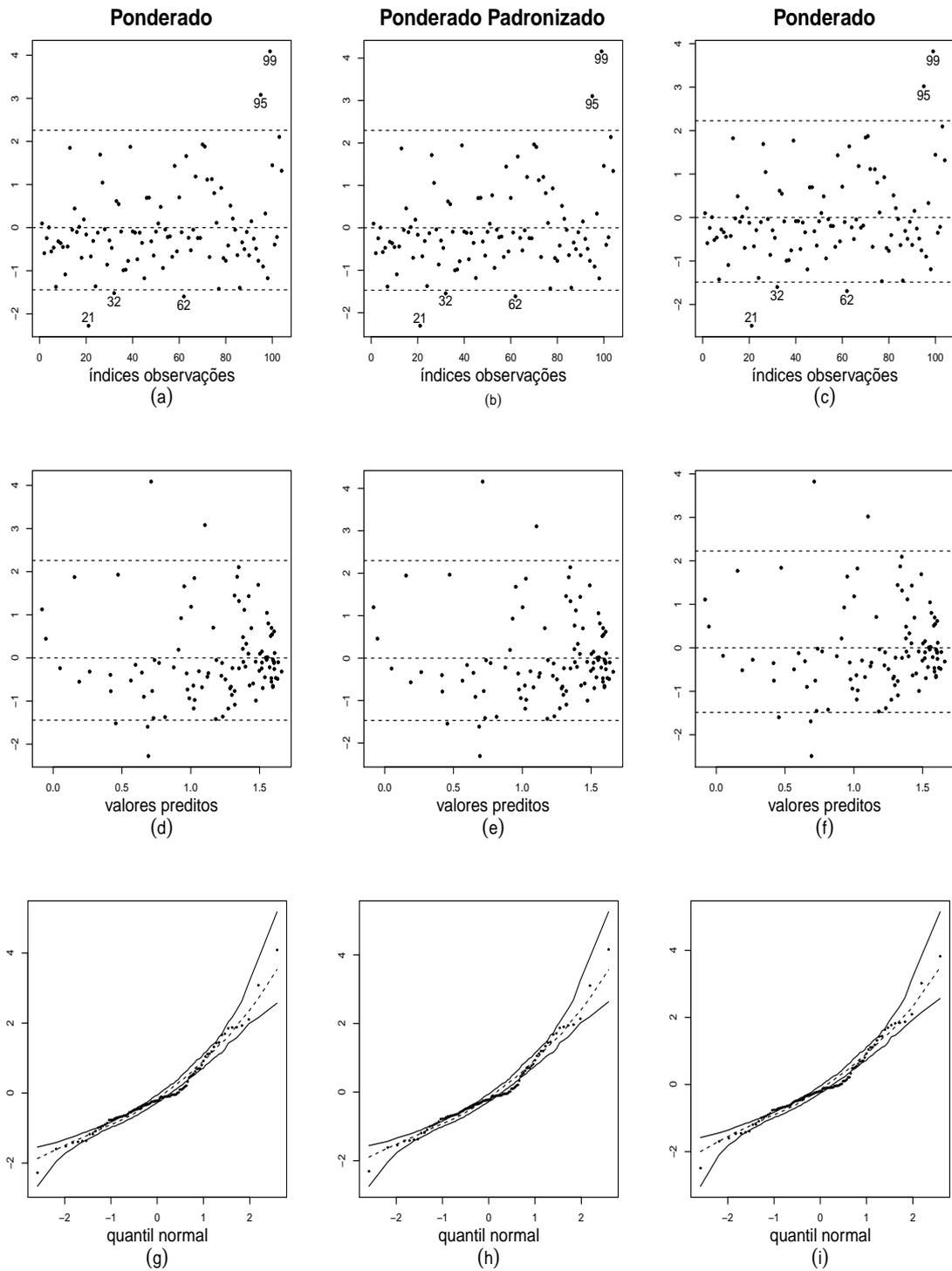
Com relação as medidas de AIC e BIC obtivemos os seguintes valores para o modelo gama unitária  $-606.86$  e  $-590.99$ , respectivamente e os seguintes valores para o modelo beta  $-574.53$  e  $-558.67$ , respectivamente. Assim, o modelo gama unitária com relação a esta medida é mais adequado para modelarmos os dados da Linha de Pobreza que o modelo beta.

Figura 32 – Gráficos dos resíduos versus. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza.



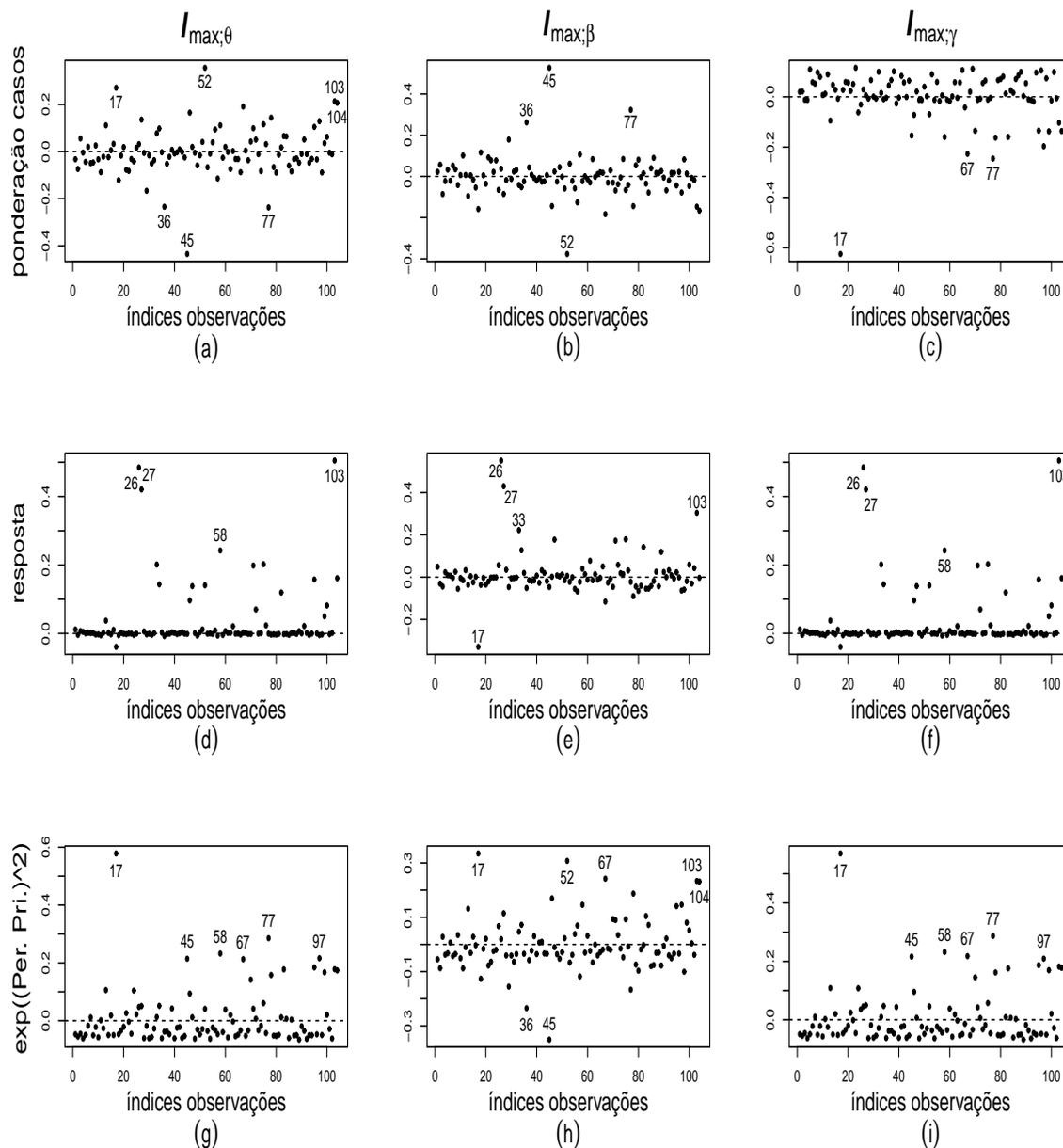
Fonte: O autor (2020).

Figura 33 – Gráficos dos resíduos. Função de ligação Cloglog. Modelo beta com dispersão variável. Dados da Linha de Pobreza.



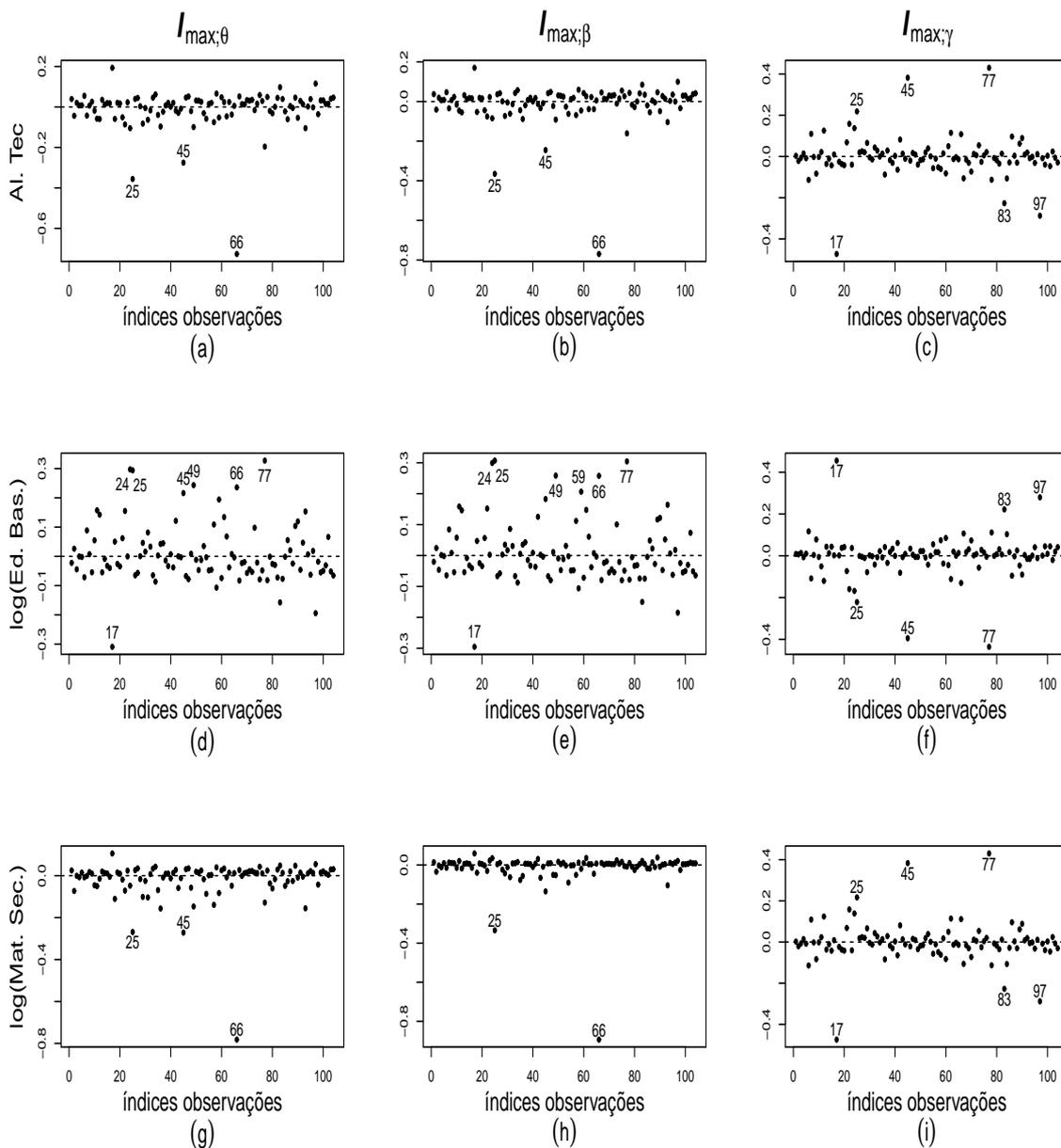
Fonte: O autor (2020).

Figura 34 – Gráficos de  $I_{\max}$  dos índices das observações versus a ponderação de casos, resposta e a covariada usada no submodelo da precisão. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza.



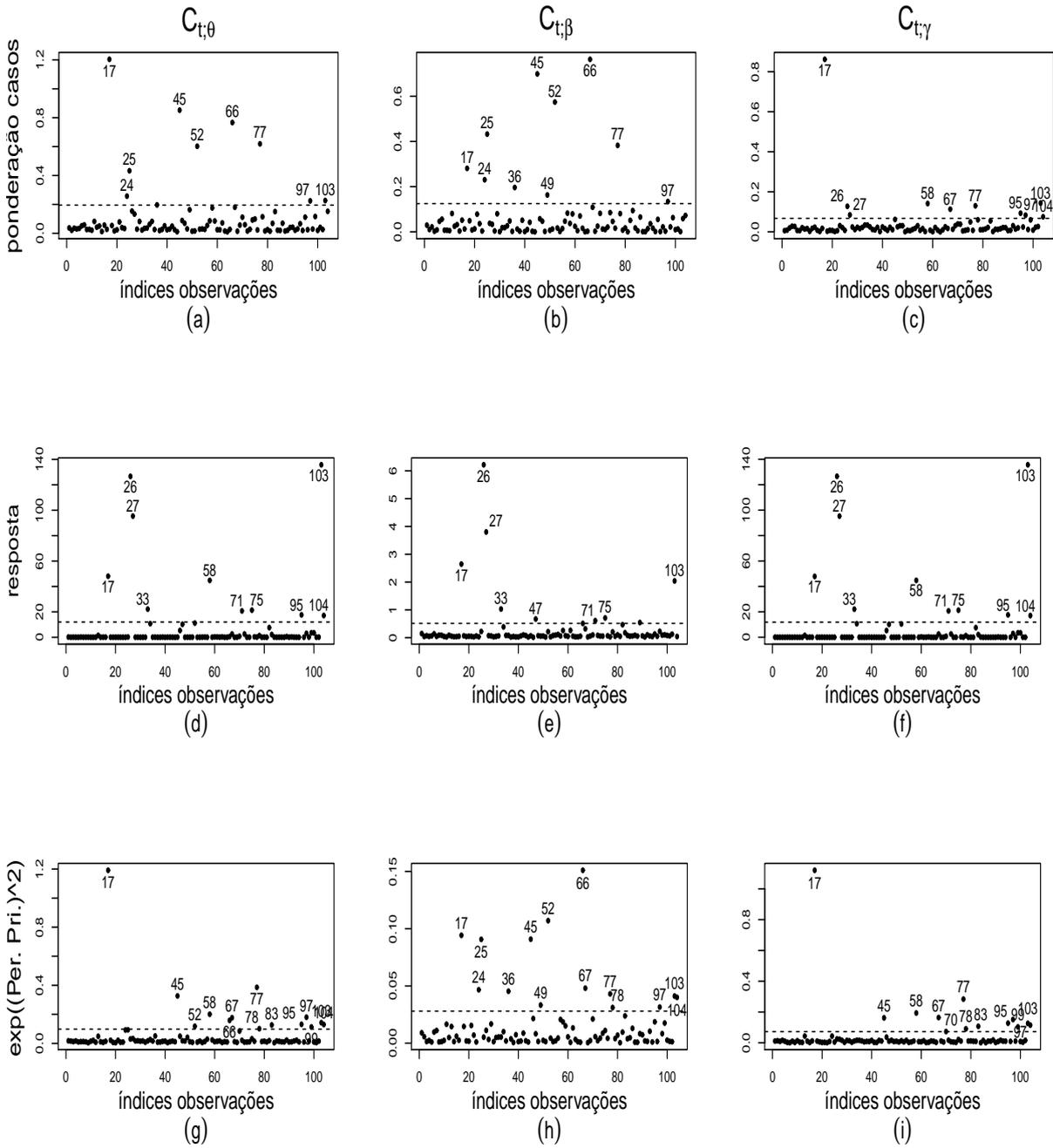
Fonte: O autor (2020).

Figura 35 – Gráficos de  $I_{\max}$  dos índices das observações versus as covariadas usadas no submodelo da média. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza



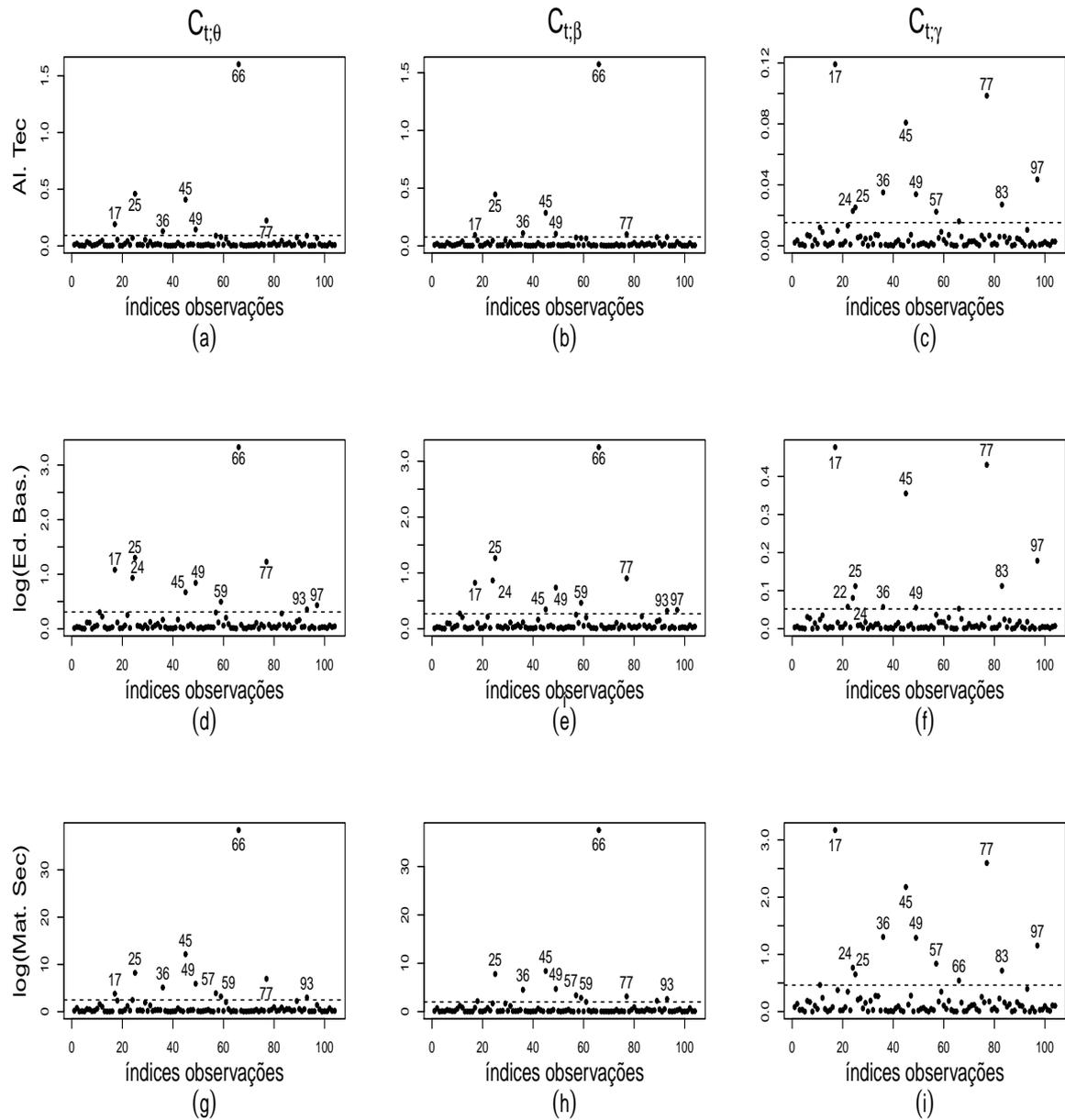
Fonte: O autor (2020).

Figura 36 – Gráficos de  $C_t$  dos índices das observações versus a ponderação de casos, resposta e a covariada usada no submodelo da precisão. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza.



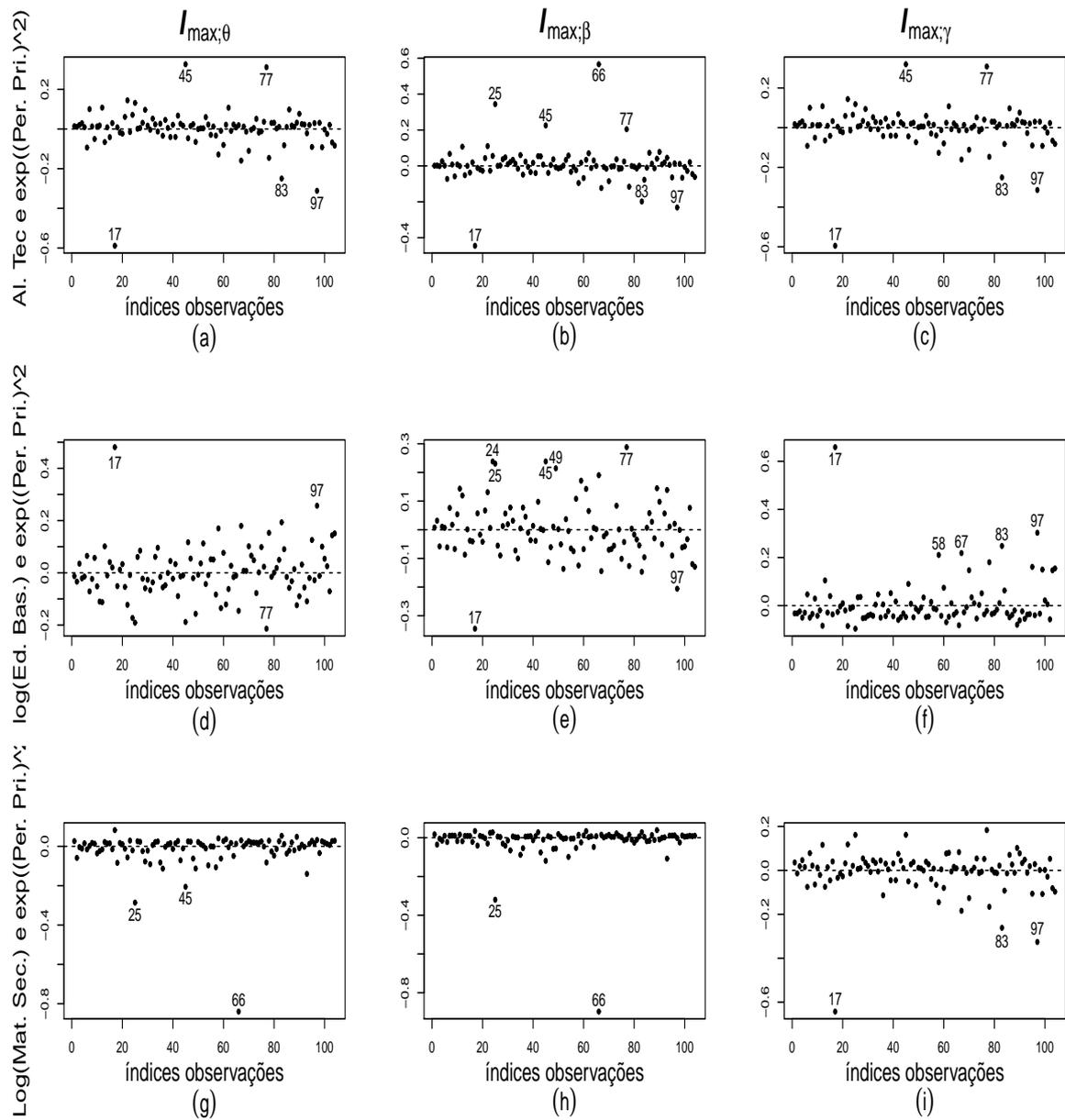
Fonte: O autor (2020).

Figura 37 – Gráficos de  $C_t$  dos índices das observações versus as covariadas usadas no submodelo da média. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza



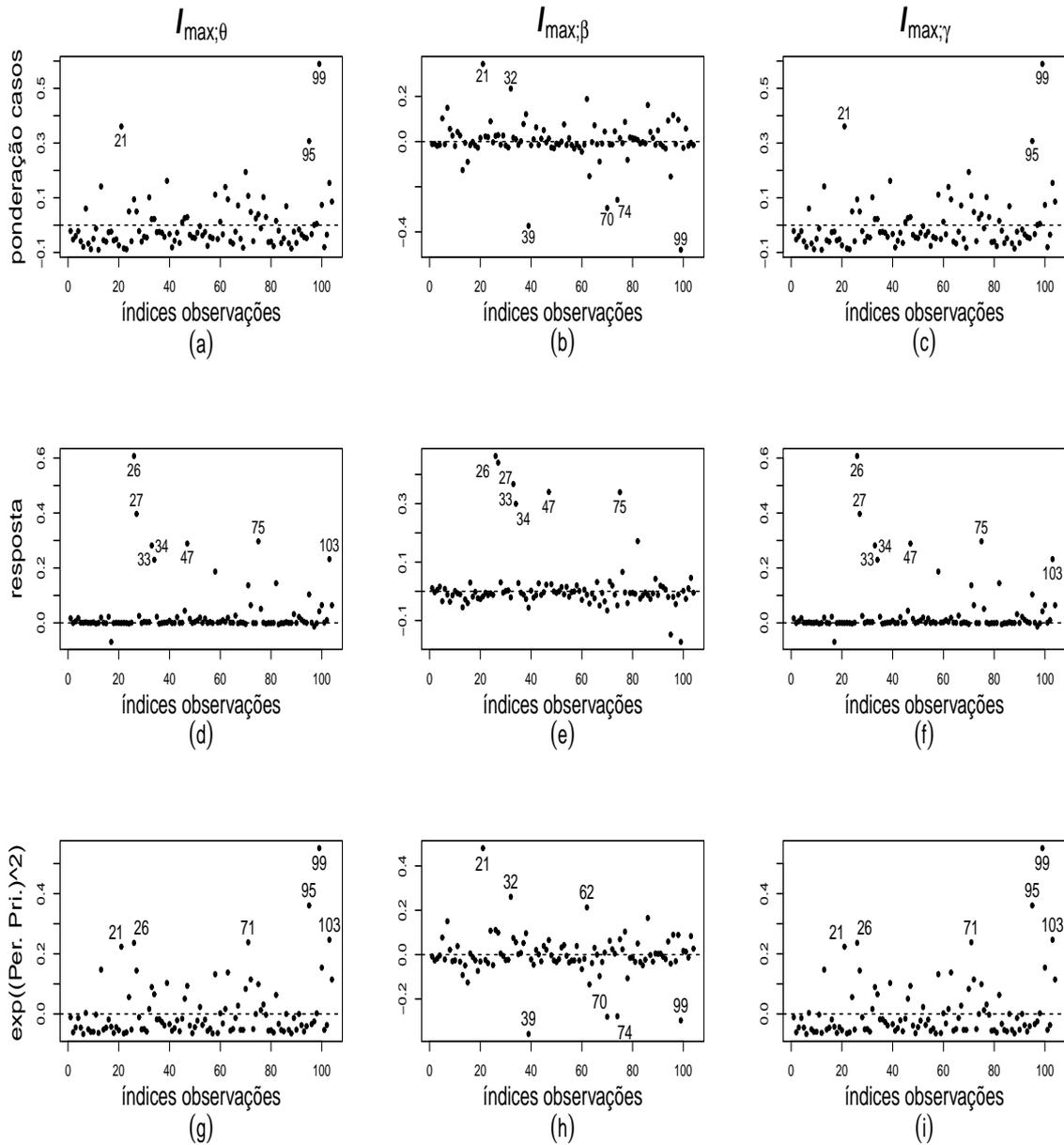
Fonte: O autor (2020).

Figura 38 – Gráficos de  $I_{\max}$  dos índices das observações versus as covariadas usadas no submodelo da média e precisão. Perturbação conjunta. Função de ligação Cloglog. Modelo gama unitária com precisão variável. Dados da Linha de Pobreza



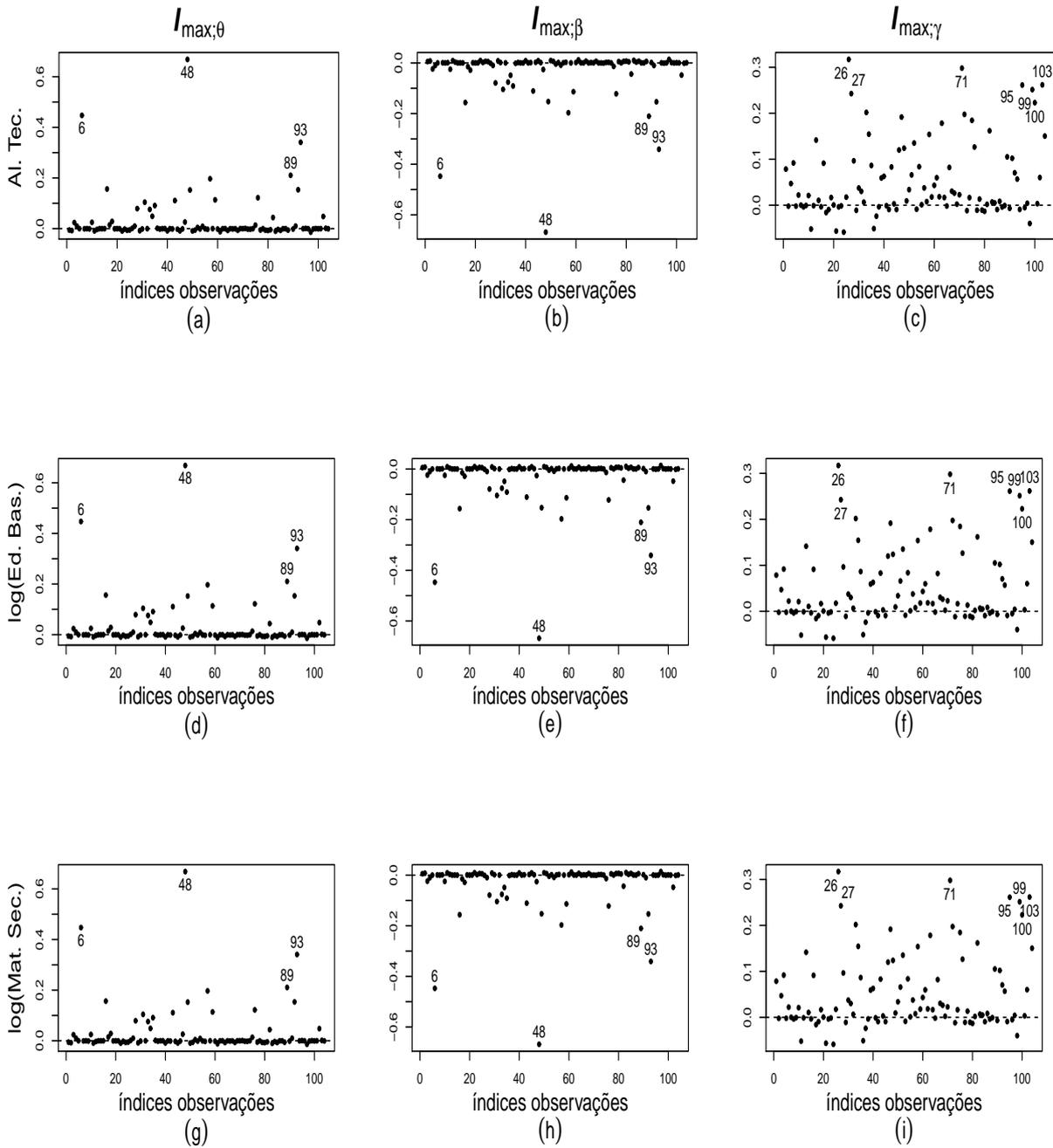
Fonte: O autor (2020).

Figura 39 – Gráficos de  $I_{\max}$  dos índices das observações versus a ponderação de casos, resposta e a covariada usada no submodelo da precisão. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha de Pobreza.



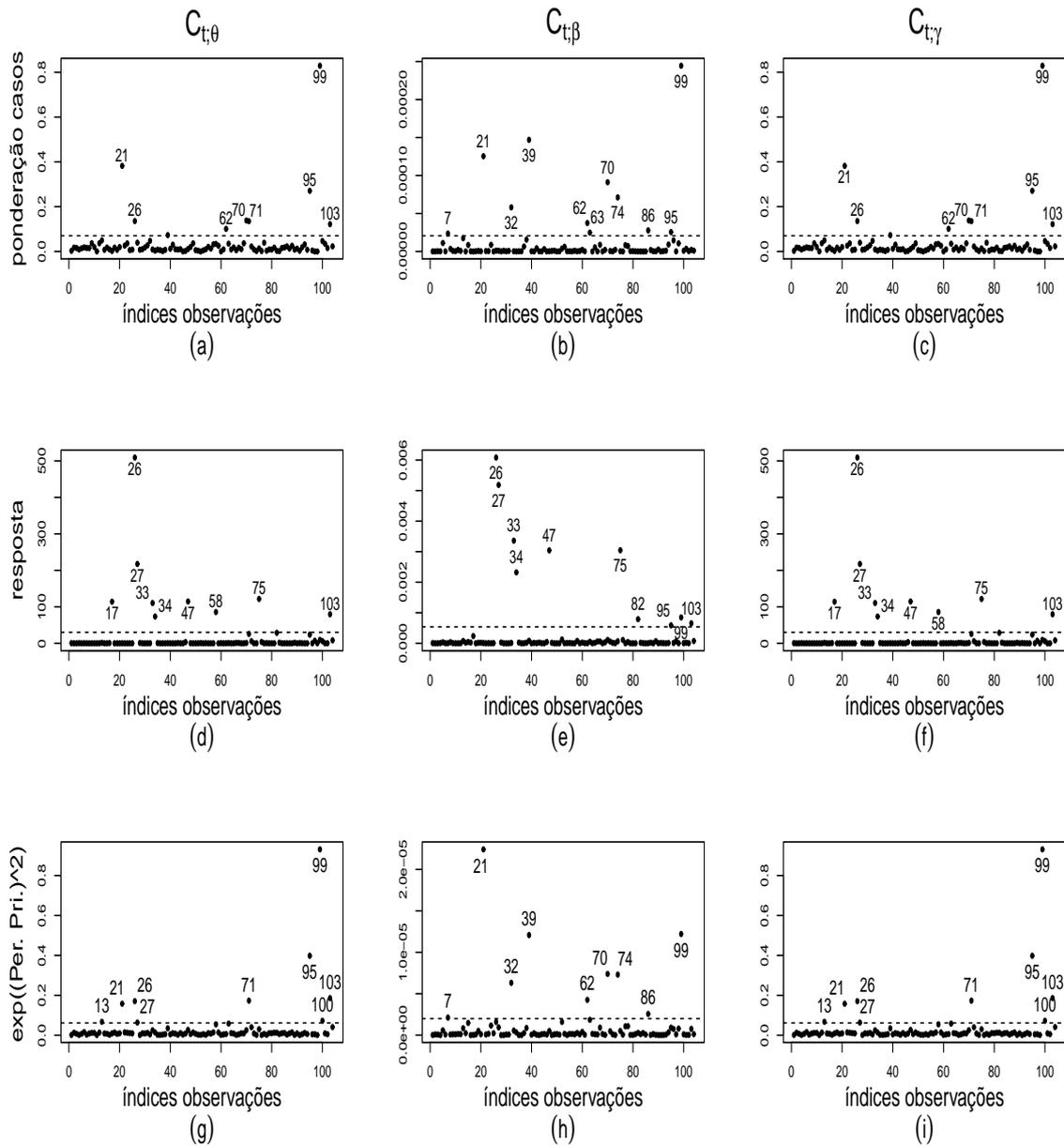
Fonte: O autor (2020).

Figura 40 – Gráficos de  $I_{\max}$  dos índices das observações versus as covariadas usadas no submodelo da média. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha de Pobreza



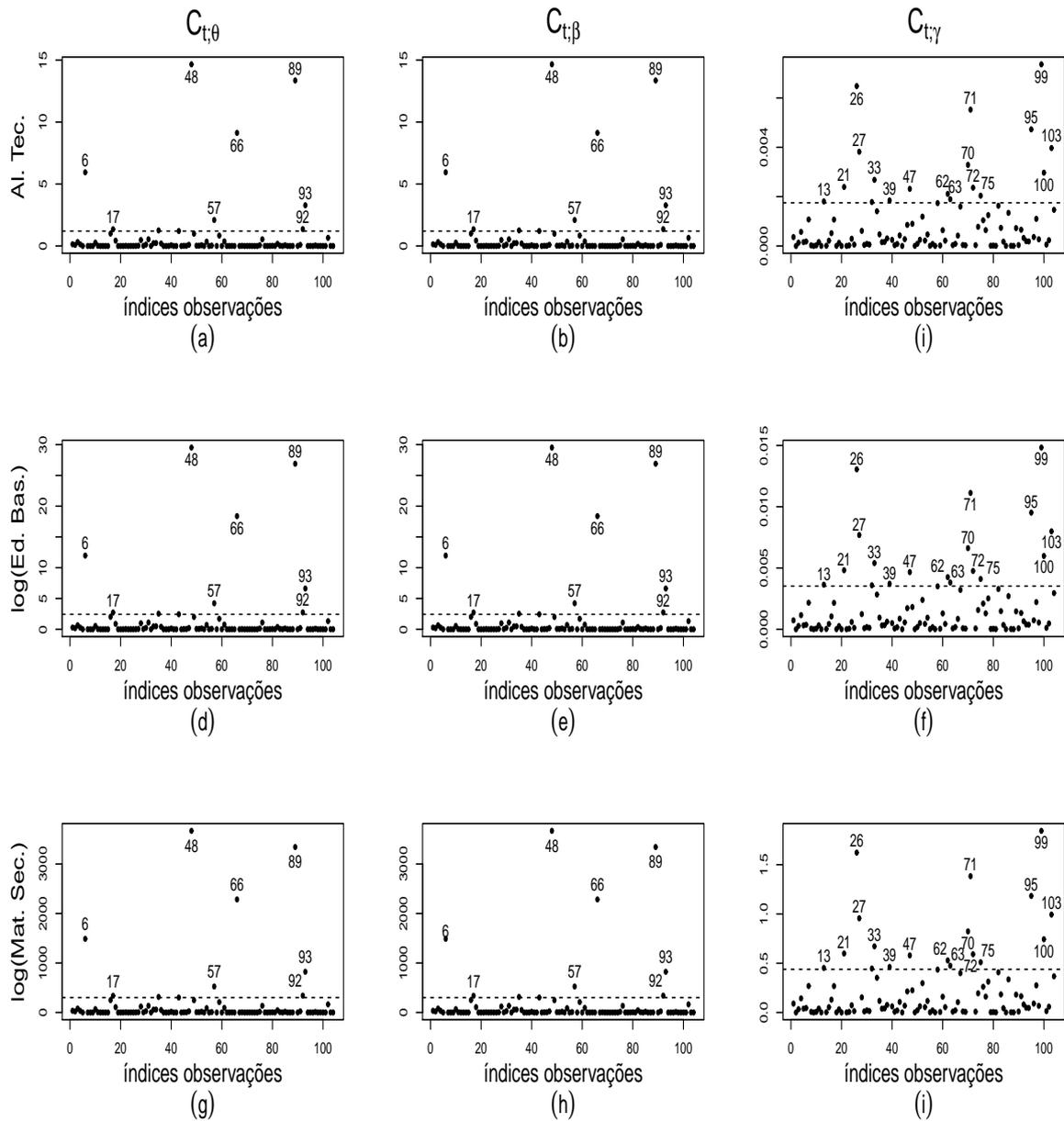
Fonte: O autor (2020).

Figura 41 – Gráficos de  $C_t$  dos índices das observações versus a ponderação de casos, resposta e a covariada usada no submodelo da precisão. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha de Pobreza.



Fonte: O autor (2020).

Figura 42 – Gráficos de  $C_t$  dos índices das observações versus as covariadas usadas no submodelo da média. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha de Pobreza



Fonte: O autor (2020).

Tabela 39 – Observações aberrantes e influentes. Função de ligação Cloglog. Modelo beta com precisão variável. Dados da Linha Pobreza.

| Modelos   |                      | beta          |               |                |                |               |                  |
|---|----------------------|---------------|---------------|----------------|----------------|---------------|------------------|
| Parâmetros  |                      | $\beta_1$     | $\beta_2$     | $\beta_3$      | $\beta_4$      | $\gamma_1$    | $\gamma_2$       |
| Descrição   |                      | const.        | $x_{t2}$      | $\log(x_{t3})$ | $\log(x_{t4})$ | const.        | $\exp(z_{t2}^2)$ |
| Observações<br>21, 32 e 62<br>Deletadas                   | Mudança Estimada     | -0.2574       | -2.5760       | -10.0631       | -3.7644        | 19.6281       | -21.6525         |
|   | Mudança Erros-padrão | -2.3321       | -2.2023       | -2.1590        | -5.7017        | 3.2266        | 1.7121           |
|   | Mudança $p$ -valor   | <b>0.0000</b> | <b>0.2367</b> | <b>0.1290</b>  | <b>0.0000</b>  | <b>0.0000</b> | <b>0.0684</b>    |
| Observações<br>6, 48, 89 e 93<br>Deletadas                | Mudança Estimada     | 0.0127        | 2.2286        | 6.3344         | -0.2853        | 5.6544        | -13.2596         |
|   | Mudança Erros-padrão | 7.5266        | 1.8165        | 5.4199         | 2.5257         | 2.3460        | 3.1828           |
|   | Mudança $p$ -valor   | <b>0.0000</b> | <b>0.2330</b> | <b>0.0958</b>  | <b>0.0000</b>  | <b>0.0000</b> | <b>0.0467</b>    |
| Observações<br>26, 27, 33, 34, 47 e 75<br>Deletadas       | Mudança Estimada     | -1.0275       | -25.9047      | -10.4964       | -1.2336        | -10.0905      | 19.3250          |
|   | Mudança Erros-padrão | -0.2870       | 3.6869        | -0.9241        | 0.8359         | 2.0325        | 3.1222           |
|   | Mudança $p$ -valor   | <b>0.0000</b> | <b>0.3960</b> | <b>0.1358</b>  | <b>0.0000</b>  | <b>0.0004</b> | <b>0.0062</b>    |
| Observações<br>26, 27, 71, 95, 99, 100 e 103<br>Deletadas | Mudança Estimada     | 2.1451        | 85.9886       | 11.7353        | 6.5553         | -32.2191      | 85.2217          |
|   | Mudança Erros-padrão | -13.5199      | -9.1217       | -13.4467       | -7.3467        | 1.7910        | 2.2746           |
|   | Mudança $p$ -valor   | <b>0.0000</b> | <b>0.0151</b> | <b>0.0330</b>  | <b>0.0000</b>  | <b>0.0071</b> | <b>0.0000</b>    |
| Observações<br>39, 70, 74 e 99<br>Deletadas               | Mudança Estimada     | 3.4224        | -19.1955      | -20.2589       | 18.9820        | 10.5279       | 5.4492           |
|   | Mudança Erros-padrão | -6.9876       | -5.1386       | -6.2176        | 3.3666         | 2.5599        | 1.6289           |
|   | Mudança $v$ -valor   | <b>0.0000</b> | <b>0.3116</b> | <b>0.1603</b>  | <b>0.0000</b>  | <b>0.0000</b> | <b>0.0141</b>    |

Fonte: O autor (2020).

Na Tabela 40 informamos os valores obtidos das medidas  $P^2$ ,  $P_c^2$ ,  $R_{FC}^2$ ,  $R_{FC_c}^2$ ,  $R_{LR}^2$  e  $R_{LR_c}^2$  com relação aos dados da Linha de pobreza. Os resultados consideram os modelos logito apresentados na seção anterior e os modelos complementar log-log.

Notamos um melhor valor assumido destas medidas na distribuição gama unitária em comparação a distribuição beta.

Tabela 40 – Medidas de predição. Função de ligação Cloglog. Modelos beta e gama unitária com precisão variável. Dados da Linha Pobreza.

| Medidas                | Modelo Clog-Log |         | Modelo Logito |         |
|------------------------|-----------------|---------|---------------|---------|
|                        | Gama Unitária   | Beta    | Gama Unitária | Beta    |
| $P_\beta^2$            | 0.93950         | 0.86533 | 0.65406       | 0.46619 |
| $P_{\beta_c}^2$        | 0.93641         | 0.85846 | 0.63641       | 0.43896 |
| $P_{\beta p}^2$        | 0.93514         | 0.85924 | 0.62928       | 0.42983 |
| $P_{\beta p_c}^2$      | 0.93183         | 0.85205 | 0.61036       | 0.40074 |
| $P_{\beta \gamma}^2$   | 0.93999         | 0.86776 | 0.65751       | 0.47322 |
| $P_{\beta \gamma_c}^2$ | 0.93693         | 0.86101 | 0.64004       | 0.44635 |
| $R_{FC}^2$             | 0.76512         | 0.76394 | 0.66484       | 0.66698 |
| $R_{FC_c}^2$           | 0.75314         | 0.75189 | 0.64774       | 0.64999 |
| $R_{LR}^2$             | 0.78245         | 0.70312 | 0.71845       | 0.66894 |
| $R_{LR_c}^2$           | 0.76463         | 0.67880 | 0.69539       | 0.64181 |

Fonte: O autor (2020).

Na Tabela 40 é possível observarmos uma maior diferença nos valores de  $P^2$  e  $P_c^2$  em relação aos valores de  $R_{FC}^2$ ,  $R_{FC_c}^2$ ,  $R_{LR}^2$  e  $R_{LR_c}^2$  em comparação a distribuição gama

unitária e beta, mostrando que a distribuição gama unitária possui uma melhor qualidade de predição e ajuste em relação a distribuição beta. Observe que a análise das medidas de predição são coerentes com todo o estudo feito da análise de resíduos e influência local.

Em um estudo mais aprofundado na análise de influência local, ao tirarmos os pontos que foram destacados com influentes na distribuição beta na distribuição gama unitária, esta não foi sensível a estes pontos influentes, porém ao tirarmos os pontos influentes da gama unitária na beta, esta foi sensível a estes pontos influentes, como é possível observamos na Tabela 41.

Em resumo, sobre o estudo feito nos dados da Linha de Pobreza, iniciamos modelando estes dados usando o modelo  $M3$  e através deste modelo comparamos o modelo gama unitária e beta, a parti daí, observamos o melhor gráficos dos resíduos, uma melhor adequação das observações aos gráficos de envelopes, uma não sensibilidade aos pontos aberrantes e influentes, os melhores valores obtidos para as medidas AIC, BIC,  $P^2$ ,  $P_c^2$ ,  $R_{FC}^2$ ,  $R_{FC_c}^2$ ,  $R_{LR}^2$  e  $R_{LR_c}^2$  no modelo gama unitária em comparação ao modelo beta. Assim, podemos afirmar que o modelo gama unitária com precisão variável é mais adequado para modelarmos dados onde a média da variável resposta se encontra próxima de um, do que o modelo beta com precisão variável.

Tabela 41 – Observações influentes. Função de ligação Cloglog. Modelos beta com precisão variável. Dados da Linha Pobreza.

| Modelos                 |             | beta                   |               |                |                |               |                  |               |
|-------------------------|-------------|------------------------|---------------|----------------|----------------|---------------|------------------|---------------|
| Parâmetros              |             | $\beta_1$              | $\beta_2$     | $\beta_3$      | $\beta_4$      | $\gamma_1$    | $\gamma_2$       |               |
| Descrição               |             | const.                 | $x_{t2}$      | $\log(x_{t3})$ | $\log(x_{t4})$ | const.        | $\exp(z_{t2}^2)$ |               |
| 24, 25, 49, 59, 66 e 77 | Observações | Mudança Estimada       | 1.9346        | -24.2057       | 6.4605         | 1.4391        | 0.9248           | 1.8779        |
|                         |             | Mudança Desvios-padrão | 2.7316        | 35.4178        | -0.5294        | 2.3897        | 1.9272           | 2.5222        |
|                         | Deletadas   | Mudança $p$ -valor     | <b>0.0000</b> | <b>0.5061</b>  | <b>0.0772</b>  | <b>0.0000</b> | <b>0.0001</b>    | <b>0.0187</b> |
| 24, 45 e 77             | Observações | Mudança Estimada       | 2.5212        | -20.4641       | -2.0193        | 2.6596        | -10.1298         | 27.5271       |
|                         |             | Mudança Desvios-padrão | -4.0342       | -2.5619        | -4.6524        | -2.6792       | 1.8866           | 1.4169        |
|                         | Deletadas   | Mudança $p$ -valor     | <b>0.0000</b> | <b>0.3322</b>  | <b>0.0897</b>  | <b>0.0000</b> | <b>0.0004</b>    | <b>0.0029</b> |
| 26, 27, 58 e 103        | Observações | Mudança Estimada       | -0.4481       | 20.1903        | 6.7164         | -1.7669       | -6.4109          | 15.3715       |
|                         |             | Mudança Desvios-padrão | -1.0517       | -0.9439        | -0.8888        | -0.8096       | 1.0565           | 1.4853        |
|                         | Deletadas   | Mudança $p$ -valor     | <b>0.0000</b> | <b>0.1495</b>  | <b>0.0754</b>  | <b>0.0000</b> | <b>0.0002</b>    | <b>0.0072</b> |
| 36, 45 e 77             | Observações | Mudança Estimada       | 1.6910        | -17.1684       | -7.6880        | 2.6589        | -7.7156          | 20.6392       |
|                         |             | Mudança Desvios-padrão | -3.1178       | -1.7860        | -3.2267        | -1.2712       | 2.1439           | 1.9702        |
|                         | Deletadas   | Mudança $p$ -valor     | <b>0.0000</b> | <b>0.3164</b>  | <b>0.1152</b>  | <b>0.0000</b> | <b>0.0003</b>    | <b>0.0051</b> |
| 17, 58, 67, 83 e 97     | Observações | Mudança Estimada       | -1.5467       | 0.2300         | 0.4970         | -1.6364       | 13.3252          | -25.9910      |
|                         |             | Mudança Desvios-padrão | 4.2021        | 6.1131         | 3.7612         | 0.5138        | 6.9047           | 6.1675        |
|                         | Deletadas   | Mudança $p$ -valor     | <b>0.0000</b> | <b>0.2619</b>  | <b>0.1097</b>  | <b>0.0000</b> | <b>0.0000</b>    | <b>0.0991</b> |
| 17, 52, 67, 103 e 104   | Observações | Mudança Estimada       | 8.1970        | 5.6146         | 155.9751       | -6.9888       | 10.1665          | -14.9124      |
|                         |             | Mudança Desvios-padrão | 39.3712       | 4.1049         | 61.2016        | 5.0022        | 4.8628           | 5.2626        |
|                         | Deletadas   | Mudança $p$ -valor     | <b>0.0000</b> | <b>0.2282</b>  | <b>0.0087</b>  | <b>0.0000</b> | <b>0.0000</b>    | <b>0.0558</b> |

Fonte: O autor (2020).

## REFERÊNCIAS

- AKAIKE, Hirotugu. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Ieee, v. 19(6), 716–723, 1974.
- ALLEN, David M. **The prediction sum of squares as a criterion for selecting predictor variables**. [S.l.]: University of Kentucky, 1971.
- ATKINSON, Anthony C. Two graphical displays for outlying and influential observations in regression. **Biometrika**, Oxford University Press, v. 68(1), 13–20, 1981.
- BAYER, Fábio M; CRIBARI-NETO, Francisco. Model selection criteria in beta regression with varying dispersion. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 46(1), 729–746, 2017.
- COOK, R Dennis. Assessment of local influence. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 48(2), 133–155, 1986.
- COX, David R; SNELL, E Joyce. A general definition of residuals. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 30(2), 248–265, 1968.
- ESPINHEIRA, Patrícia L; FERRARI, Silvia LP; CRIBARI-NETO, Francisco. On beta regression residuals. **Journal of Applied Statistics**, Taylor & Francis, v. 35(4), 407–419, 2008.
- ESPINHEIRA, Patrícia L; OLIVEIRA SILVA, Alisson de. Residual and influence analysis to a general class of simplex regression. **TEST**, Springer, 1–30, 2019.
- ESPINHEIRA, Patrícia L; SANTOS, Evelyne G; CRIBARI-NETO, Francisco. On nonlinear beta regression residuals. **Biometrical Journal**, Wiley Online Library, v. 59(3), 445–461, 2017.
- FERRARI, Silvia; CRIBARI-NETO, Francisco. Beta regression for modelling rates and proportions. **Journal of Applied Statistics**, Taylor & Francis, v. 31(7), 799–815, 2004.
- GEISSER, Seymour; EDDY, William F. A predictive approach to model selection. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 74(365), 153–160, 1979.
- GRASSIA, A. On a family of distributions with argument between 0 and 1 obtained by transformation of the gamma and derived compound distributions. **Australian and New Zealand Journal of Statistics**, Wiley Online Library, v. 19(2), 108–114, 1977.
- GRIFFITHS, David; CSIRO, Christine Schafer. Closeness of Grassia's transformed Gammas and the Beta distribution. **Australian Journal and New Zealand of Statistics**, Wiley Online Library, v. 23(2), 240–246, 1981.
- HOCKING, RR. Criteria for selection of a subset regression: which one should be used? **Technometrics**, Taylor & Francis Group, v. 14(4), 967–976, 1972.
- KIESCHNICK, Robert; MCCULLOUGH, Bruce D. Regression analysis of variates observed on (0, 1): percentages, proportions and fractions. **Statistical modelling**, Sage Publications Sage CA: Thousand Oaks, CA, v. 3(3), 193–213, 2003.
- LEHMANN, Erich L.; CASELLA, George. **Theory of Point Estimation**. Second. New York, NY, USA: Springer-Verlag, 1998.

- LESAFFRE, Emmanuel; VERBEKE, Geert. Local influence in linear mixed models. **Biometrics**, JSTOR, 570–582, 1998.
- MEDIAVILLA, Francis A Méndez; LANDRAM, Frank; SHAH, Vivek. A comparison of the coefficient of predictive power, the coefficient of determination and AIC for linear regression. **The Journal of Applied Business and Economics**, North American Business Press, v. 8(4), 44, 2008.
- MOUSA, Amany M; EL-SHEIKH, Ahmed A; ABDEL-FATTAH, Mahmoud A. A Gamma Regression for Bounded Continuous Variables. **Advances and Applications in Statistics**, v. 49(4), 305–326, 2016.
- OLSHEN, Abraham Charles. Transformations of the pearson type III distribution. **The Annals of Mathematical Statistics**, JSTOR, v. 9(3), 176–200, 1938.
- PAOLINO, Philip. Maximum likelihood estimation of models with beta-distributed dependent variables. **Political Analysis**, Cambridge University Press, v. 9(4), 325–346, 2001.
- PREGIBON, Daryl et al. Logistic regression diagnostics. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 9(4), 705–724, 1981.
- RAO, C. R. **Linear Statistical Inference and its Applications**. Wiley New York: [s.n.], 1973.
- RATNAPARKHL, Makarand V; MOSIMANN, James E. On the normality of transformed beta and unit-gamma random variables. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 19(10), 3833–3854, 1990.
- ROCHA, Andréa V; SIMAS, Alexandre B. Influence diagnostics in a general class of beta regression models. **Test**, Springer, v. 20(1), 95–119, 2011.
- SCHWARZ, Gideon. Estimating the dimension of a model. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 6(2), 461–464, 1978.
- SIMAS, Alexandre B; BARRETO-SOUZA, Wagner; ROCHA, Andréa V. Improved estimators for a general class of beta regression models. **Computational Statistics & Data Analysis**, Elsevier, v. 54(2), 348–366, 2010.
- SMITHSON, Michael; VERKUILEN, Jay. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. **Psychological Methods**, American Psychological Association, v. 11(1), 54, 2006.
- TADIKAMALLA, Pandu R. On a family of distributions obtained by the transformation of the gamma distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 13(3-4), 209–214, 1981.
- THOMAS, William; COOK, R Dennis. Assessing influence on regression coefficients in generalized linear models. **Biometrika**, Oxford University Press, v. 76(4), 741–749, 1989.

## APÊNDICE A – DERIVADAS SEGUNDAS

As derivadas segundas da expressão (2.11) com relação aos parâmetros da distribuição gama unitária são

$$\begin{aligned}
 \frac{\partial^2 \ell_t(\mu_t, \phi_t)}{\partial \mu_t^2} &= \frac{\partial}{\partial \mu_t} \left[ \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} \right] \\
 &= \frac{\partial}{\partial \mu_t} \left\{ \frac{\alpha_t}{\mu_t^{1/\phi_t+1}} \left[ 1 + \frac{1}{\phi_t} \alpha_t \log(y_t) \right] \right\} \\
 &= \frac{\log(y_t) \mu_t^{1/\phi_t} \left[ \phi_t (\mu_t^{1/\phi_t} - 1) + \mu_t^{1/\phi_t} + 1 \right] + \phi_t \mu_t^{1/\phi_t} \left( 1 - \mu_t^{1/\phi_t} \phi_t + 2\phi_t - \mu_t^{1/\phi_t} \right) - \phi_t^2}{\mu_t^2 (1 - \mu_t^{1/\phi_t})^3 \phi_t^2} \\
 &= v_t^*,
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial^2 \ell_t(\mu_t, \phi_t)}{\partial \mu_t \partial \phi_t} &= \frac{\partial}{\partial \phi_t} \left[ \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} \right] \\
 &= \frac{\partial}{\partial \phi_t} \left\{ \frac{\alpha_t}{\mu_t^{1/\phi_t+1}} \left[ 1 + \frac{1}{\phi_t} \alpha_t \log(y_t) \right] \right\} \\
 &= - \frac{\log(y_t) \left[ \log(\mu_t) \mu_t^{1/\phi_t-1} \left( 1 + \mu_t^{1/\phi_t} \right) + \mu_t^{1/\phi_t-1} \phi_t \left( 1 - \mu_t^{1/\phi_t} \right) \right]}{\phi_t^3 (1 - \mu_t^{1/\phi_t})^3} \\
 &\quad + \frac{\phi_t \log(\mu_t) \mu_t^{1/\phi_t-1} \left( 1 - \mu_t^{1/\phi_t} \right)}{\phi_t^3 (1 - \mu_t^{1/\phi_t})^3} \\
 &= c_t^*,
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial^2 \ell_t(\mu_t, \phi_t)}{\partial \phi_t^2} &= \frac{\partial}{\partial \phi_t} \left[ \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} \right] \\
 &= \frac{\partial}{\partial \phi_t} \left\{ \log(-\log(y_t)) - \left[ \frac{1}{\phi_t} \alpha_t \log(\mu_t) \right] \left[ 1 + \frac{\alpha_t \log(y_t)}{\phi_t \mu_t^{1/\phi_t}} \right] - \log \left( \frac{\mu_t^{1/\phi_t}}{\alpha_t} \right) - \psi(\phi_t) \right\} \\
 &= \log(\mu_t) \left( \frac{\mu_t^{\frac{2}{\phi_t}} \{ \log(y_t) [\log(\mu_t) - 2\phi_t] - \log(\mu_t) \phi_t \} + \mu_t^{\frac{1}{\phi_t}} \{ \log(y_t) [\log(\mu_t) + 2\phi_t] \}}{\phi_t^4 (1 - \mu_t^{1/\phi_t})^3} \right) \\
 &\quad + \frac{(\log(\mu_t))^2 \phi_t}{\phi_t^4 (1 - \mu_t^{1/\phi_t})^3} - \psi'(\phi_t) \\
 &= e_t^*.
 \end{aligned}$$

Considerando nas expressões acima  $\phi_t = \phi$ , ou seja, tornando o parâmetro de precisão fixo, temos que  $v_t^* = v_t$ ,  $c_t^* = c_t$  e  $e_t^* = g_t$ .

## APÊNDICE B – INFLUÊNCIA LOCAL

Agora, estudaremos em detalhes a influência local com precisão variável; o caso de precisão fixa é análogo.

### ESQUEMA DE PERTURBAÇÃO

#### Ponderação de casos

Considere o seguinte esquema de perturbação:

$$\ell_\delta(\beta, \gamma) = \sum_{t=1}^n \delta_t \ell_t(\mu_t, \phi_t),$$

com  $0 \leq \delta_t \leq 1$ . Para o método de influência local discutido anteriormente, temos que  $\delta_0 = (1, 1, \dots, 1)^\top$  e, assim, a  $t$ -ésima coluna de  $\Delta$ ,  $t = 1, \dots, n$  é dada por

$$\Delta_t = \frac{\partial^2 \ell_\delta(\theta)}{\partial \theta \partial \delta_t} = \frac{\partial}{\partial \theta} \left[ \frac{\partial \ell_\delta(\theta)}{\partial \delta_t} \right] = \frac{\partial \ell_t(\theta)}{\partial \theta} = \left[ \frac{\partial \ell_t(\theta)}{\partial \beta_1}, \dots, \frac{\partial \ell_t(\theta)}{\partial \beta_p}, \frac{\partial \ell_t(\theta)}{\partial \gamma_1}, \dots, \frac{\partial \ell_t(\theta)}{\partial \gamma_q} \right]^\top \Bigg|_{\theta=\hat{\theta}}.$$

Com base em (2.11), para  $i = 1, \dots, p$  e  $j = 1, \dots, q$ , obtemos

$$\frac{\partial \ell_t(\theta)}{\partial \beta_i} = (\mu_t^* + y_t^*) \frac{1}{g'(\mu_t)} x_{ti} \quad \text{e} \quad \frac{\partial \ell_t(\theta)}{\partial \gamma_j} = e_t \frac{1}{h'(\phi_t)} z_{tj},$$

em que, os  $t$ -ésimos elementos de  $\mu_t^*$ ,  $y_t^*$  e  $e_t$  estão definidos em (2.12) e (2.13), respectivamente. Então, temos que a  $i$ -ésima linha  $\Delta_\beta$ ,  $i = 1, \dots, p$  é dada por

$$\Delta_{\beta_i} = \left[ (\mu_1^* + y_1^*) \frac{1}{g'(\mu_1)} x_{1i}, \dots, (\mu_n^* + y_n^*) \frac{1}{g'(\mu_n)} x_{ni} \right] \Bigg|_{\theta=\hat{\theta}},$$

e a  $j$ -ésima linha de  $\Delta_\gamma$ , com  $j = 1, \dots, q$ , é dada por

$$\Delta_{\gamma_j} = \left[ e_1 \frac{1}{h'(\phi_1)} z_{1j}, \dots, e_n \frac{1}{h'(\phi_n)} z_{nj} \right] \Bigg|_{\theta=\hat{\theta}}.$$

Dessa forma, podemos escrever  $\Delta_\beta = X^\top \hat{T}^* \hat{\mathcal{E}}^*$  e  $\Delta_\gamma = Z^\top \hat{H}^* \hat{E}^*$ . Logo,

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{pmatrix} X^\top \hat{T}^* \hat{\mathcal{E}}^* \\ Z^\top \hat{H}^* \hat{E}^* \end{pmatrix},$$

onde  $\mathcal{E}^*$  e  $E^*$  estão definidos em (4.17) e (4.18), respectivamente.

## Perturbação da variável resposta

Consideremos o esquema aditivo de perturbação da resposta em que

$$y_t(\delta) = y_t + \delta_t s(y_t), \quad (\text{B.2})$$

com  $s(y_t) = \sqrt{\hat{\mu}_t(1/[(2 - \hat{\mu}_t^{1/\hat{\phi}})\hat{\phi}] - \hat{\mu}_t)}$ . Para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$  e com base na definição da variável resposta sob perturbação aditiva apresentada (B.2) e considerando (2.11), temos que o logaritmo da função de verossimilhança do modelo perturbado é

$$\ell_\delta(\beta, \gamma) = \sum_{t=1}^n \phi_t \log(\alpha_t) - \log(\Gamma(\phi_t)) + (\alpha_t - 1) \log(y_t + \delta_t s(y_t)) + (\phi_t - 1) \log(-\log(y_t + \delta_t s(y_t))).$$

Então,

$$\begin{aligned} \frac{\partial \ell_\delta(\beta, \gamma)}{\partial \delta_t} &= \frac{(\alpha_t - 1)s(y_t)}{y_t + \delta_t s(y_t)} + (\phi_t - 1) \frac{s(y_t)}{\log(y_t + \delta_t s(y_t)) [y_t + \delta_t s(y_t)]} \\ &= \frac{\alpha_t s(y_t)}{y_t + \delta_t s(y_t)} - \frac{s(y_t)}{y_t + \delta_t s(y_t)} + \frac{\phi_t s(y_t)}{\log(y_t + \delta_t s(y_t)) [y_t + \delta_t s(y_t)]} \\ &\quad - \frac{s(y_t)}{\log(y_t + \delta_t s(y_t)) [y_t + \delta_t s(y_t)]}. \end{aligned}$$

Assim, para  $i = 1, 2, \dots, p$ , temos

$$\begin{aligned} \frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} &= \frac{\partial}{\partial \beta_i} \left[ \frac{\partial \ell_\delta(\beta, \gamma)}{\partial \delta_t} \right] = \frac{\partial}{\partial \mu_t} \left[ \frac{\partial \ell_\delta(\beta, \gamma)}{\partial \delta_t} \right] \frac{d\mu_t}{d\eta_t} \frac{\partial \eta_t}{\partial \beta_i} \\ &= \frac{s(y_t)}{y_t + \delta_t s(y_t)} \frac{\mu_t^{1/\phi_t - 1}}{(1 - \mu_t^{1/\phi_t})^2 \phi_t} \frac{1}{g'(\mu_t)} x_{ti} = \frac{s(y_t)}{y_t + \delta_t s(y_t)} \frac{\alpha_t^2}{\phi_t \mu_t^{1/\phi_t + 1}} \frac{1}{g'(\mu_t)} x_{ti}. \end{aligned}$$

e

$$\begin{aligned} \frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} &= \frac{\partial}{\partial \gamma_j} \left[ \frac{\partial \ell_\delta(\beta, \gamma)}{\partial \delta_t} \right] = \frac{\partial}{\partial \phi_t} \left[ \frac{\partial \ell_\delta(\beta, \gamma)}{\partial \delta_t} \right] \frac{d\phi_t}{dn_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} \\ &= \left\{ \frac{s(y_t)}{\log(y_t + \delta_t s(y_t)) [y_t + \delta_t s(y_t)]} + \frac{s(y_t)}{y_t + \delta_t s(y_t)} \left[ \frac{-\mu_t^{1/\phi_t} \log(\mu_t)}{(1 - \mu_t^{1/\phi_t})^2 \phi_t^2} \right] \right\} \frac{d\phi_t}{dn_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} \end{aligned}$$

$$= \left\{ \frac{s(y_t)}{\log(y_t + \delta_t s(y_t)) [y_t + \delta_t s(y_t)]} - \frac{s(y_t)}{y_t + \delta_t s(y_t)} \left[ \frac{\alpha_t^2 \log(\mu_t)}{\mu_t^{1/\phi_t} \phi_t^2} \right] \right\} \frac{1}{h'(\phi_t)} z_{tj}.$$

Avaliando as expressões acima em  $\delta_0 = (0, 0, \dots, 0)^\top$  e  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$ , podemos escrever  $\Delta_\beta = X^\top \hat{T}^* \hat{M}^* S_y^*$  e  $\Delta_\gamma = Z^\top \hat{H}^* \hat{B}^* S_y^*$ , onde

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{pmatrix} X^\top \hat{T}^* \hat{M}^* S_y^* \\ Z^\top \hat{H}^* \hat{B}^* S_y^* \end{pmatrix},$$

com  $M^* = \text{diag}(m_1^*, \dots, m_n^*)$ ,  $S_y^* = \text{diag}(s^*(y_1), \dots, s^*(y_n))$  e  $\hat{B}^* = (\hat{b}_1^*, \dots, \hat{b}_n^*)$  em que  $m_t^*$  e  $b_t^*$  estão definidos, respectivamente, em (4.20) e (4.21).

## Perturbação individual de covariáveis

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

#### Perturbação na matriz $X$

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

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k (x_{tk} + \delta_t s_{xk}) + \dots + \beta_p x_{tp},$$

e  $\mu_t(\delta)$  é tal que  $g(\mu_t(\delta)) = \eta_{1t}(\delta)$ , para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$ .

O logaritmo da função de verossimilhança do modelo perturbado é

$$\ell_\delta(\mu, \phi) = \sum_{t=1}^n \phi_t \log(\alpha_t(\delta)) - \log(\Gamma(\phi_t)) + (\alpha_t(\delta) - 1) \log(y_t) + (\phi_t - 1) \log(-\log(y_t)).$$

Assim, temos que

$$\frac{\partial \ell(\beta, \gamma)}{\partial \delta_t} = \frac{\partial \ell(\beta, \gamma)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \delta_t} = (\mu_t^* + y_t^*) \frac{1}{g'(\mu_t)} \beta_k s_{xk}, \quad (\text{B.3})$$

em que  $s_{xk}$  é o desvio padrão de  $x_k$ .

De (B.3) segue que, para  $i = 1, 2, \dots, p$  com  $i \neq k$

$$\begin{aligned} \frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} &= \frac{\partial}{\partial \mu_t} (\mu_t^* + y_t^*) \left[ \frac{1}{g'(\mu_t)} \right]^2 \beta_k s_{xk} x_{ti} \\ &\quad + (\mu_t^* + y_t^*) \left\{ \frac{-g''(\mu_t)}{[g'(\mu_t)]^2} \right\} \frac{1}{g'(\mu_t)} \beta_k s_{xk} x_{ti} \\ &= \left\{ v_i^* - (\mu_t^* + y_t^*) \left[ \frac{g''(\mu_t)}{g'(\mu_t)} \right] \right\} \left[ \frac{1}{g'(\mu_t)} \right]^2 \beta_k s_{xk} x_{ti} \end{aligned}$$

ou, ainda,

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = -s_{x_k} \beta_k x_{tk} q_t^*. \quad (\text{B.4})$$

Quando  $i = k$ , temos que

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = s_{x_k} \left\{ v_t^* - (\mu_t^* + y_t^*) \left[ \frac{g''(\mu_t)}{g'(\mu_t)} \right] \right\} \left[ \frac{1}{g'(\mu_t)} \right]^2 \beta_k (x_{tk} + \delta_t s_{x_k}) + (\mu_t^* + y_t^*) \left[ \frac{1}{g'(\mu_t)} \right] s_{x_k},$$

Simplificando a expressão acima, temos

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = -s_{x_k} \left[ \beta_k (x_{tk} + \delta_t s_{x_k}) q_t^* - (\mu_t^* + y_t^*) \frac{1}{g'(\mu_t)} \right]. \quad (\text{B.5})$$

Ainda, de (B.3), segue que, para  $j = 1, \dots, q$ ,

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} = \frac{\partial}{\partial \phi_t} \left[ \frac{\partial l_\delta(\beta, \delta)}{\partial \delta_t} \right] \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_j} = c_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \beta_k s_{x_k} z_{tj}.$$

Adicionalmente,

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = -\beta_k s_{x_k} f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} z_{tj}. \quad (\text{B.6})$$

Avaliando as expressões (B.4), (B.5) e (B.6) em  $\delta_0 = (0, 0, \dots, 0)^\top$  e  $\hat{\theta}$  obtemos  $\Delta_\beta = -s_{x_k} \left( \hat{\beta}_k X^\top \hat{Q}^* - P \hat{T}^* \hat{\mathcal{E}}^* \right)$  e  $\Delta_\gamma = -\hat{\beta}_k s_{x_k} Z^\top \hat{F}^* \hat{T}^* \hat{H}^*$ , em que  $Q^*$ ,  $T^*$ ,  $\mathcal{E}^*$ ,  $F^*$  e  $H^*$ , estão definidos em (4.12), (4.14), (4.17), (4.13) e (4.15), respectivamente. Assim, a matriz  $\Delta$  fica expressa na forma

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{bmatrix} -s_{x_k} \left( \hat{\beta}_k X^\top \hat{Q}^* - P \hat{T}^* \hat{\mathcal{E}}^* \right) \\ -\hat{\beta}_k s_{x_k} Z^\top \hat{F}^* \hat{T}^* \hat{H}^* \end{bmatrix}. \quad (\text{B.7})$$

## Perturbação na matriz $Z$

Suponha que  $k \neq 2$  ou  $k \neq q$ ,

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 z_{t2} + \dots + \gamma_k (z_{tk} + \delta_t s_{zk}) + \dots + \gamma_p z_{tp},$$

e  $\phi_t(\delta)$  é tal que  $h(\phi_t(\delta)) = \eta_{2t}(\delta)$ , para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$ .

O logaritmo da função de verossimilhança do modelo perturbado é

$$\ell_\delta(\mu, \phi) = \sum_{t=1}^n \phi_t(\delta) \log(\alpha_t(\delta)) - \log(\Gamma(\phi_t(\delta))) + [\alpha_t(\delta) - 1] \log(y_t) + [\phi_t(\delta) - 1] \log(-\log(y_t)).$$

Assim, temos que

$$\frac{\partial \ell(\beta, \gamma)}{\partial \delta_t} = \frac{\partial \ell(\beta, \gamma)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \delta_t} = e_t \frac{1}{h'(\phi_t)} \gamma_k s_{z_k}, \quad (\text{B.8})$$

em que  $s_{z_k}$  é o desvio padrão de  $z_k$ .

De (B.8) segue que, para  $i = 1, 2, \dots, p$ , temos

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = \frac{\partial}{\partial \beta_i} \left[ \frac{\partial \ell(\beta, \gamma)}{\partial \delta_t} \right] = \frac{\partial}{\partial \mu_i} \left[ \frac{\partial \ell(\beta, \gamma)}{\partial \delta_t} \right] \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i} = c_t^* \frac{1}{h'(\phi)} \frac{1}{g'(\mu)} x_{ti} \gamma_k s_{z_k},$$

ou seja,

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = -\gamma_k s_{z_k} f_t^* \frac{1}{h'(\phi)} \frac{1}{g'(\mu)} x_{ti}. \quad (\text{B.9})$$

Ainda, de (B.8), segue que, para  $j = 1, \dots, q$  com  $j \neq k$ , temos

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_i} = \left\{ e_t^* - e_t \left[ \frac{h''(\phi)}{h'(\phi)} \right] \right\} \left[ \frac{1}{h'(\phi_t)} \right]^2 \gamma_k s_{z_k} z_{tj}.$$

Portanto,

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = -s_{z_k} \gamma_k z_{tj} v_t^*. \quad (\text{B.10})$$

Quando  $j = k$ , temos que

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} = s_{z_k} \left\{ e_t^* - e_t \left[ \frac{h''(\phi_t)}{h'(\phi_t)} \right] \right\} \left[ \frac{1}{h'(\phi_t)} \right]^2 \gamma_k (z_{tk} + \delta_t s_{z_k}) + e_t \left[ \frac{1}{h'(\phi_t)} \right] s_{z_k}.$$

Simplificando a expressão acima, temos

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = -s_{z_k} \left[ \gamma_k (z_{tk} + \delta_t s_{z_k}) v_t^* - e_t \frac{1}{h'(\phi_t)} \right]. \quad (\text{B.11})$$

Avaliando as expressões (B.9), (B.10) e (B.11) em  $\delta_0 = (0, 0, \dots, 0)^\top$  e  $\hat{\theta}$  obtemos

$\Delta_\beta = -\hat{\gamma}_k s_{x_k} X^\top \hat{F}^* \hat{T}^* \hat{H}^*$  e  $\Delta_\gamma = -s_{z_k} \left( \hat{\gamma}_k Z^\top \hat{V}^* - P \hat{H}^* \hat{E}^* \right)$  em que  $F^*$ ,  $T^*$ ,  $H^*$ ,  $V^*$  e  $E^*$  estão definidos em (4.13), (4.14), (4.15), (4.16) e (4.18). Assim, a matriz  $\Delta$  fica expressa na forma

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{bmatrix} -\hat{\gamma}_k s_{x_k} X^\top \hat{F}^* \hat{T}^* \hat{H}^* \\ -s_{z_k} \left( \hat{\gamma}_k Z^\top \hat{V}^* - P \hat{H}^* \hat{E}^* \right) \end{bmatrix}. \quad (\text{B.12})$$

### Matriz $Z$ igual à matriz $X$

Neste caso,

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k (x_{tk} + \delta_t s_{x_k}) + \cdots + \beta_p x_{tp},$$

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 x_{t2} + \cdots + \gamma_k (x_{tk} + \delta_t s_{x_k}) + \cdots + \gamma_p x_{tp},$$

onde  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_{1t}(\delta)$  e  $\phi_t(\delta)$  é tal que  $h[\phi_t(\delta)] = \eta_{2t}(\delta)$ , para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$ .

O logaritmo da função de verossimilhança do modelo perturbado é

$$\ell_\delta(\mu, \phi) = \sum_{t=1}^n \phi_t(\delta) \log(\alpha_t(\delta)) - \log(\Gamma(\phi_t(\delta))) + [\alpha_t(\delta) - 1] \log(y_t) + [\phi_t(\delta) - 1] \log(-\log(y_t)).$$

Assim, temos que

$$\frac{\partial \ell(\beta, \gamma)}{\partial \delta_t} = \frac{\partial l_\delta(\mu_t, \phi_t)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \delta_t} + \frac{\partial l_\delta(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \delta_t} = (\mu_t^* + y_t^*) \frac{1}{g'(\mu_t)} \beta_k s_{x_k} + e_t \frac{1}{h'(\phi_t)} \gamma_k s_{z_k}, \quad (\text{B.13})$$

em que  $s_{x_k}$  e  $s_{z_k}$  são os desvio padrão de  $x_k$  e  $z_k$ , respectivamente.

De (B.13), segue que para  $i = 1, 2, \dots, p$  com  $i \neq k$

$$\begin{aligned} \frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} &= v_t^* \left[ \frac{1}{g'(\mu_t)} \right]^2 \beta_k s_{x_k} x_{ti} + (\mu_t^* + y_t^*) \left\{ -\frac{g''(\mu_t)}{[g'(\mu_t)]^2} \right\} \beta_k s_{x_k} \frac{1}{g'(\mu_t)} x_{ti} \\ &\quad + c_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \gamma_k s_{z_k} x_{ti}. \end{aligned}$$

Dessa forma, obtemos para  $i \neq k$ ,

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_i} = \frac{\partial}{\partial \beta_i} \left[ \frac{\partial \ell_\delta(\beta, \gamma)}{\partial \delta_t} \right] = -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} x_{ti} \gamma_k s_{z_k} - q_t^* x_{ti} \beta_k s_{x_k}. \quad (\text{B.14})$$

Para  $i = 1, 2, \dots, p$  com  $i = k$ , temos

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \beta_k} = -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \gamma_k s_{x_k}(x_{tk} + \delta_t s_{x_k}) - q_t^* \beta_k s_{x_k}(x_{tk} + \delta_t s_{x_k}) + (\mu_t^* + y_t^*) \frac{1}{g'(\mu_t)} s_{x_k}. \quad (\text{B.15})$$

Observe agora que para calcular as segundas derivadas com relação ao vetor  $\gamma$ , também teremos que considerar  $j = k$  e  $j \neq k$ . Assim, com base em (B.13) para  $j = 1, \dots, q$ , com  $j \neq k$ , temos que

$$\begin{aligned} \frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} &= c_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \beta_k s_{x_k} x_{tj} + e_t^* \left[ \frac{1}{h'(\phi_t)} \right]^2 \gamma_k s_{x_k} x_{tj} + e_t \left[ -\frac{h''(\phi_t)}{h'(\phi_t)^2} \right] \frac{1}{h'(\phi_t)} \gamma_k s_{x_k} x_{tj} \\ &= -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \beta_k s_{x_k} x_{tj} - \left[ -e_t^* + e_t \frac{h''(\phi_t)}{h'(\phi_t)} \right] \left[ \frac{1}{h'(\phi_t)} \right]^2 \gamma_k s_{x_k} x_{tj}. \end{aligned}$$

Dessa forma, temos

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} = -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \beta_k s_{x_k} x_{tj} - v_t^* \gamma_k s_{x_k} x_{tj}. \quad (\text{B.16})$$

Finalmente, quando  $j = k$ , temos que

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} = -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \beta_k s_{x_k}(x_{tk} + \delta_t s_{x_k}) - v_t^* \gamma_k s_{x_k}(x_{tk} + \delta_t s_{x_k}) + e_t \frac{1}{h'(\phi_t)} s_{x_k}. \quad (\text{B.17})$$

Avaliando as expressões (B.14)-(B.17) em  $\delta_0 = (0, 0, \dots, 0)^\top$  e  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$ , obtemos  $\Delta_\beta = -s_{x_k} \left[ X^\top (\hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\gamma}_k) - P \hat{T}^* \hat{\mathcal{E}}^* \right]$  e  $\Delta_\gamma = -s_{x_k} \left[ X^\top (\hat{\gamma}_k \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k) - P \hat{H}^* \hat{E}^* \right]$ .

Logo,

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{pmatrix} -s_{x_k} \left[ X^\top (\hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\gamma}_k) - P \hat{T}^* \hat{\mathcal{E}}^* \right] \\ -s_{x_k} \left[ X^\top (\hat{\gamma}_k \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k) - P \hat{H}^* \hat{E}^* \right] \end{pmatrix}, \quad (\text{B.18})$$

em que  $P$  é uma matriz  $p \times n$  de zeros exceto a  $k$ -ésima linha que é composta por uns e as expressões  $Q^*$ ,  $F^*$ ,  $T^*$ ,  $H^*$ ,  $\mathcal{E}^*$ ,  $V^*$  e  $E^*$  estão definidas em (4.12), (4.13), (4.14), (4.15), (4.17) e (4.18), respectivamente.

**A  $k'$ -ésima coluna da matriz  $Z$  é igual à  $k$ -ésima coluna da matriz  $X$**

Aqui, estamos considerando a situação em que algumas covariáveis (não todas) que determinam a média também estão envolvidas na modelagem da precisão e pretendemos perturbar tais covariáveis. Ou seja, consideramos que  $z_{tk'} = x_{tk}$ , para algum par  $(k, k')$ ,

com  $k = 2, \dots, p$  e  $k' = 2, \dots, q$ . Neste caso,

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k(x_{tk} + \delta_t s_{xk}) + \dots + \beta_p x_{tp},$$

Suponha que se  $k' \neq 2$  ou  $k' \neq q$

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 z_{t2} + \dots + \gamma_{k'}(x_{tk} + \delta_t s_{xk}) + \dots + \gamma_q z_{tq},$$

onde  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_{1t}(\delta)$  e  $\phi_t(\delta)$  é tal que  $h[\phi_t(\delta)] = \eta_{2t}(\delta)$ , para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$ .

O logaritmo da função de verossimilhança do modelo perturbado é

$$\ell_\delta(\mu, \phi) = \sum_{t=1}^n \phi_t(\delta) \log(\alpha_t(\delta)) - \log(\Gamma(\phi_t(\delta))) + [\alpha_t(\delta) - 1] \log(y_t) + [\phi_t(\delta) - 1] \log(-\log(y_t)).$$

Assim, temos que

$$\frac{\partial \ell(\beta, \gamma)}{\partial \delta_t} = \frac{\partial \ell_\delta(\mu_t, \phi_t)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \delta_t} + \frac{\partial \ell_\delta(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \delta_t} = (\mu_t^* + y_t^*) \frac{1}{g'(\mu_t)} \beta_k s_{xk} + e_t \frac{1}{h'(\phi_t)} \gamma_{k'} s_{xk}, \quad (\text{B.19})$$

em que  $s_{xk}$  é o desvio padrão de  $x_k$ .

Note que a única diferença entre as expressões (B.13) e (B.19) é que na primeira aparece  $\gamma_k$ , e na segunda aparece  $\gamma_{k'}$ . Ocorre o mesmo com a expressão de  $\Delta_\beta$  em (B.26) e a referência ao esquema de perturbação atual, em que  $\Delta_\beta = -s_{xk} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\gamma}_{k'} \right) - P \hat{T}^* \hat{E}^* \right]$ .

Quanto as derivadas com relação ao vetor  $\gamma$ , temos que para  $j = 1, \dots, q$  com  $j \neq k$ , temos que

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} = -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} z_{tj} \beta_k s_{xk} - v_t^* z_{tj} \gamma_{k'} s_{xk}. \quad (\text{B.20})$$

e quando  $j = k$ , temos que

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} = -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \beta_k s_{xk} (x_{tk} + \delta_t s_{xk}) - v_t^* (x_{tk} + \delta_t s_{xk}) \gamma_{k'} s_{xk} + e_t \frac{1}{h'(\phi_t)} s_{xk}.$$

No entanto, considerando que  $x_{tk} = z_{tk'}$ , a expressão anterior pode ser escrita como

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_k} = -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \beta_k s_{xk} (z_{tk'} + \delta_t s_{xk}) - v_t^* (z_{tk'} + \delta_t s_{xk}) \gamma_{k'} s_{xk} + e_t \frac{1}{h'(\phi_t)} s_{xk}. \quad (\text{B.21})$$

Avaliando as expressões (B.20) e (B.21) em  $\delta_0 = (0, 0, \dots, 0)^\top$  e  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$ , obtemos  $\Delta_\gamma = -s_{xk} \left[ Z^\top \left( \hat{\gamma}_{k'} \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P \hat{H}^* \hat{E}^* \right]$ .

Logo,

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \left\{ \begin{array}{l} -s_{x_k} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\gamma}_{k'} \right) - P \hat{T}^* \hat{\mathcal{E}}^* \right] \\ -s_{x_k} \left[ Z^\top \left( \hat{\gamma}_{k'} \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P \hat{H}^* \hat{E}^* \right] \end{array} \right\}, \quad (\text{B.22})$$

em que  $P$  é uma matriz  $q \times n$  de zeros exceto a  $k$ -ésima linha que é composta por uns e  $V^*, Q^*, F^*, T^*, H^*, \mathcal{E}^*, E^*$  estão definidos em (4.16), (4.12), (4.13), (4.14), (4.15), (4.17) e (4.18) respectivamente.

### A $k'$ -ésima coluna da matriz $Z$ é uma função da $k$ -ésima coluna da matriz $X$

Aqui estamos considerando a situação em que  $z_{tk'} = \mathcal{G}(x_{tk})$ , com  $k = 2, \dots, p$  e  $k' = 2, \dots, q$ , em que  $\mathcal{G}$  é uma função contínua e diferenciável; se  $\mathcal{G}$  é ua função identidade, temos a situação tratada na seção anterior. Aqui,

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k (x_{tk} + \delta_t s_{xk}) + \dots + \beta_p x_{tp}.$$

Seja  $k' \neq 2$  e  $k' \neq q$ . Temos que

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 z_{t2} + \dots + \gamma_{k'} \mathcal{G}(x_{tk} + \delta_t s_{xk}) + \dots + \gamma_q z_{tq},$$

onde  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_{1t}(\delta)$  e  $\phi_t(\delta)$  é tal que  $h[\phi_t(\delta)] = \eta_{2t}(\delta)$ , para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$ .

O logaritmo da função de verossimilhança do modelo perturbado é

$$\ell_\delta(\mu, \phi) = \sum_{t=1}^n \phi_t(\delta) \log(\alpha_t(\delta)) - \log(\Gamma(\phi_t(\delta))) + [\alpha_t(\delta) - 1] \log(y_t) + [\phi_t(\delta) - 1] \log(-\log(y_t)).$$

Assim, temos que

$$\begin{aligned} \frac{\partial \ell(\beta, \gamma)}{\partial \delta_t} &= \frac{\partial \ell_\delta(\mu_t, \phi_t)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \delta_t} + \frac{\partial \ell_\delta(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \delta_t} \\ &= (\mu_t^* + y_t^*) \frac{1}{g'(\mu_t)} \beta_k s_{xk} + e_t \frac{1}{h'(\phi_t)} \gamma_{k'} \mathcal{G}'(x_{tk} + \delta_t s_{xk}) s_{xk}, \end{aligned}$$

em que  $s_{x_k}$  é o desvio padrão de  $x_k$ . Com base na expressão acima, chegamos a  $\Delta_\beta = -s_{x_k} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{G} \hat{\gamma}_{k'} \right) - P \hat{T}^* \hat{\mathcal{E}}^* \right]$ , em que  $\hat{G} = \text{diag} [\mathcal{G}'(x_{1k}), \dots, \mathcal{G}'(x_{nk})]$  com  $\mathcal{G}'(x_{tk}) = d\mathcal{G}(x_{tk})/dx_{tk}$ .

Quanto as derivadas com relação ao vetor  $\gamma$ , temos que para  $j = 1, \dots, q$  com  $j \neq k$ , temos que

$$\frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} = -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} z_{tj} \beta_k s_{xk} - v_t^* z_{tj} \gamma_{k'} \mathcal{G}'(x_{tk} + \delta_t s_{xk}) s_{xk} \quad (\text{B.23})$$

e, quando  $j = k$ , temos que

$$\begin{aligned} \frac{\partial^2 \ell_\delta(\beta, \gamma)}{\partial \delta_t \partial \gamma_j} &= -f_t^* \frac{1}{g'(\mu_t)} \frac{1}{h'(\phi_t)} \beta_k s_{x_k} \mathcal{G}(x_{tk} + \delta_t s_{x_k}) - v_t^* \mathcal{G}(x_{tk} + \delta_t s_{x_k}) \gamma_{k'} \mathcal{G}'(x_{tk} + \delta_t s_{x_k}) s_{x_k} \\ &\quad + e_t \frac{1}{h'(\phi_t)} \mathcal{G}'(x_{tk} + \delta_t s_{x_k}) s_{x_k}. \end{aligned} \quad (\text{B.24})$$

Avaliando as expressões (B.23) e (B.24) em  $\delta_0 = (0, 0, \dots, 0)^\top$  e  $\hat{\theta} = (\hat{\beta}^\top, \hat{\gamma}^\top)^\top$ . Considerando que  $\mathcal{G}(x_{tk} + \delta_t s_{x_k}) = \mathcal{G}(x_{tk}) = z_{tk'}$ , obtemos  $\Delta_\gamma = -s_{x_k} \left[ Z^\top \left( \hat{\gamma}_{k'} \dot{G} \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P \hat{H}^* \hat{E}^* \dot{G} \right]$ .

Logo,

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \begin{Bmatrix} -s_{x_k} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \dot{G} \hat{\gamma}_{k'} \right) - P \hat{T}^* \hat{E}^* \right] \\ -s_{x_k} \left[ Z^\top \left( \hat{\gamma}_{k'} \dot{G} \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P \hat{H}^* \hat{E}^* \dot{G} \right] \end{Bmatrix},$$

em que  $P$  é uma matriz  $q \times n$  de zeros exceto a  $k$ -ésima linha que é composta por uns e  $V^*, Q^*, F^*, T^*, H^*, E^*$  estão definidos em (4.16), (4.12), (4.13), (4.14), (4.15), (4.17) e (4.18), respectivamente.

## Perturbação conjunta de covariáveis

Neste caso,

$$\eta_{1t}(\delta) = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k (x_{tk} + \delta_t s_{x_k}) + \dots + \beta_p x_{tp},$$

$$\eta_{2t}(\delta) = \gamma_1 + \gamma_2 z_{t2} + \dots + \gamma_k (z_{tk} + \delta_t s_{z_k}) + \dots + \gamma_q z_{tq},$$

onde  $\mu_t(\delta)$  é tal que  $g[\mu_t(\delta)] = \eta_{1t}(\delta)$  e  $\phi_t(\delta)$  é tal que  $h[\phi_t(\delta)] = \eta_{2t}(\delta)$ , para esse tipo de perturbação  $\delta_0 = (0, 0, \dots, 0)^\top$ .

O logaritmo da função de verossimilhança do modelo perturbado é

$$\ell_\delta(\mu, \phi) = \sum_{t=1}^n \phi_t(\delta) \log(\alpha_t(\delta)) - \log(\Gamma(\phi_t(\delta))) + [\alpha_t(\delta) - 1] \log(y_t) + [\phi_t(\delta) - 1] \log(-\log(y_t)).$$

Assim, temos que

$$\frac{\partial \ell(\beta, \gamma)}{\partial \delta_t} = \frac{\partial \ell_\delta(\mu_t, \phi_t)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \delta_t} + \frac{\partial \ell_\delta(\mu_t, \phi_t)}{\partial \phi_t} \frac{d\phi_t}{d\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \delta_t} = (\mu_t^* + y_t^*) \frac{1}{g'(\mu_t)} \beta_k s_{x_k} + e_t \frac{1}{h'(\phi_t)} \gamma_k s_{z_k}, \quad (\text{B.25})$$

em que  $s_{x_k}$  e  $s_{z_k}$  são os desvios padrões de  $x_k$  e  $z_k$ , respectivamente.

A matriz  $\Delta$  é semelhante a matriz encontrada na Seção , onde a matriz  $\Delta_\beta$  encontrada da Seção é idêntica, enquanto que na matriz  $\Delta_\gamma$  é suficiente trocarmos  $s_{x_k}$ ,  $X^\top$  e  $P$  por  $s_{z_k}$ ,  $Z^\top$  e  $P_p$ , respectivamente. Daí obtemos a seguinte expressão para a matriz  $\Delta$ :

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\gamma \end{pmatrix} = \left\{ \begin{array}{l} -s_{x_k} \left[ X^\top \left( \hat{\beta}_k \hat{Q}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\gamma}_k \right) - P_m \hat{T}^* \hat{\mathcal{E}}^* \right] \\ -s_{z_k} \left[ Z^\top \left( \hat{\gamma}_k \hat{V}^* + \hat{F}^* \hat{T}^* \hat{H}^* \hat{\beta}_k \right) - P_p \hat{H}^* \hat{E}^* \right] \end{array} \right\}, \quad (\text{B.26})$$

em que  $P_m$  é  $P_p$  são matrizes  $p \times n$  e  $q \times n$ , respectivamente, de zeros exceto a  $k$ -ésima linha que é composta por uns e as expressões  $Q^*$ ,  $F^*$ ,  $T^*$ ,  $H^*$ ,  $\mathcal{E}^*$ ,  $V^*$  e  $E^*$  estão definidas respectivamente em (4.12), (4.13), (4.14), (4.15), (4.17), (4.16) e (4.18).

## APÊNDICE C – PROPRIEDADE

Neste apêndice provaremos uma propriedade importante descoberta nesta tese para a distribuição gama unitária.

**Propriedade 1.** *Se  $X \sim ug$  então  $(1 - X) \sim ug$ .*

**Demonstração.** *Se  $X \sim ug$  então  $F_X(x) = \int_{-\infty}^x ug(s, \alpha, \phi) ds = \int_0^x ug(s, \alpha, \phi) ds$ .*

*Então, a função acumulada de  $X$  é dada por*

$$F_X(x) = \begin{cases} 0, & \text{se } x < 0 \\ \int_0^x ug(s, \alpha, \phi) ds, & \text{se } x \in [0, 1) \\ 1, & \text{se } x \geq 1 \end{cases}$$

*Logo, a função de densidade é*

$$f_X(x) = \begin{cases} 0, & \text{se caso contrário} \\ ug(x, \alpha, \phi) & \text{se } x \in (0, 1) \end{cases}$$

*Observe que*

$$\begin{aligned} F_{1-X}(z) &= \mathbb{P}(1 - X \leq z) \\ &= \mathbb{P}(X \geq 1 - z) \\ &= 1 - F_X(1 - z) \\ &= 1 - \int_{-\infty}^{1-z} ug(s, \alpha, \phi) ds \\ &= 1 - \int_0^{1-z} ug(s, \alpha, \phi) ds \end{aligned}$$

*Dai, concluímos a função acumulada de  $1 - X$  é*

$$F_{1-X}(x) = \begin{cases} 0, & \text{se } 1 - x \geq 1 \\ 1 - \int_0^{1-x} ug(s, \alpha, \phi) ds, & \text{se } 1 - x \in [0, 1) \\ 1, & \text{se } 1 - x \leq 0 \end{cases}$$

Então, a sua função de densidade é

$$f_{1-X}(x) = \begin{cases} 0, & \text{se } \text{caso contrário} \\ ug(1-x, \alpha, \phi) & \text{se } 1-x \in [0, 1] \end{cases}$$

De onde segue que  $1-X \sim ug$ .